



## **Leitlinien zu Stichprobenverfahren für Prüfbehörden**

(gemäß Artikel 62 der Verordnung (EG) Nr. 1083/2006 des Rates, Artikel 16, einschließlich Anhang IV, der Verordnung (EG) Nr. 1828/2006 der Kommission und Artikel 61 der Verordnung (EG) Nr. 1198/2006 des Rates, Artikel 42, einschließlich Anhang IV, der Verordnung (EG) Nr. 498/2007 der Kommission)

### **HAFTUNGSAUSSCHLUSS:**

*„Diese Arbeitsunterlage wurde von den Dienststellen der Kommission erstellt. Ausgehend vom geltenden Gemeinschaftsrecht bietet sie öffentlichen Verwaltungen, praktischen Anwendern, Begünstigten und möglichen Begünstigten sowie sonstigen mit der Überwachung, Kontrolle oder Durchführung der Kohäsionspolitik befassten Stellen technische Unterstützung bei der richtigen Auslegung und Anwendung der gemeinschaftlichen Rechtsvorschriften in diesem Bereich. Die Kommissionsdienststellen erläutern und interpretieren hierin die genannten Vorschriften, um die Durchführung operativer Programme zu erleichtern und bewährte Verfahrensweisen zu fördern. Dieser Leitfaden greift jedoch einer Auslegung durch den Europäischen Gerichtshof und das Gericht sowie Entscheidungen der Kommission nicht vor.“*

## INHALTSVERZEICHNIS

<b>HAFTUNGSAUSSCHLUSS:</b> .....	<b>1</b>
<b>1 EINLEITUNG</b> .....	<b>7</b>
<b>2 BEZUG AUF DIE RECHTSGRUNDLAGEN – RECHTSRAHMEN</b> .....	<b>9</b>
<b>3 PRÜFUNGSRISIKOMODELL UND PRÜFVERFAHREN</b> .....	<b>11</b>
3.1 RISIKOMODELL.....	11
3.2 SICHERHEITS-/KONFIDENZNIVEAU FÜR DIE PRÜFUNG VON VORHABEN.....	16
<b>4 BEHANDLUNG VON FEHLERN</b> .....	<b>18</b>
4.1 SYSTEMBEDINGTE UND BEKANNTE FEHLER.....	19
4.2 ZUFALLSFEHLER.....	20
4.3 ANOMALE FEHLER .....	20
4.4 PROGNOSTIZIERTE GESAMTFEHLERQUOTE (TPER) .....	21
<b>5 STATISTISCHE BEGRIFFE IM ZUSAMMENHANG MIT DER PRÜFUNG VON VORHABEN</b> .....	<b>23</b>
5.1 STICHPROBENVERFAHREN .....	23
5.2 AUSWAHLVERFAHREN .....	24
5.3 HOCHRECHNUNG (SCHÄTZUNG).....	25
5.4 GENAUIGKEIT (STICHPROBENFEHLER).....	26
5.5 GRUNDGESAMTHEIT .....	26
5.6 SCHICHTUNG .....	28
5.7 STICHPROBENEINHEIT.....	28
5.8 ERHEBLICHKEIT .....	28
5.9 ZULÄSSIGER FEHLER UND GEPLANTE GENAUIGKEIT.....	29
5.10 STREUUNG .....	30
5.11 KONFIDENZINTERVALL UND OBERE FEHLERGRENZE .....	31
5.12 KONFIDENZNIVEAU .....	33
5.13 FEHLERQUOTE .....	33
<b>6 STICHPROBENVERFAHREN FÜR DIE PRÜFUNG VON VORHABEN</b> .....	<b>35</b>
6.1 ÜBERSICHT .....	35
6.2 BEDINGUNGEN FÜR DIE ANWENDUNG VON STICHPROBENDESIGNS .....	37
6.3 PROGRAMME IM RAHMEN DER EUROPÄISCHEN TERRITORIALEN ZUSAMMENARBEIT (ETZ).....	<b>3940</b>
6.4 BEZEICHNUNGEN.....	40
<b>7 STICHPROBENVERFAHREN</b> .....	<b>42</b>
7.1 EINFACHE ZUFALLSSTICHPROBENVERFAHREN .....	42
7.1.1 <i>Standardansatz</i> .....	42
7.1.1.1 Einleitung .....	42
7.1.1.2 Stichprobenumfang.....	42
7.1.1.3 Prognostizierter Fehler.....	43
7.1.1.4 Genauigkeit .....	44
7.1.1.5 Bewertung.....	45
7.1.1.6 Beispiel .....	<b>4647</b>
7.1.2 <i>Das geschichtete einfache Stichprobenverfahren</i> .....	51
7.1.2.1 Einleitung .....	51
7.1.2.2 Stichprobenumfang.....	52
7.1.2.3 Prognostizierter Fehler.....	53

7.1.2.4	Genauigkeit .....	54
7.1.2.5	Bewertung.....	56
7.1.2.6	Beispiel .....	56
7.1.3	<i>Einfaches Zufallsstichprobenverfahren – zwei Zeiträume</i> .....	64
7.1.3.1	Einleitung .....	64
7.1.3.2	Stichprobenumfang.....	64
7.1.3.3	Prognostizierter Fehler.....	67
7.1.3.4	Genauigkeit .....	67
7.1.3.5	Bewertung.....	68
7.1.3.6	Beispiel .....	69
7.2	DIFFERENZENSCHÄTZUNG .....	75
7.2.1	<i>Standardansatz</i> .....	75
7.2.1.1	Einleitung .....	75
7.2.1.2	Stichprobenumfang.....	76
7.2.1.3	Extrapolation.....	76
7.2.1.4	Genauigkeit .....	77
7.2.1.5	Bewertung.....	77
7.2.1.6	Beispiel .....	78
7.2.2	<i>Geschichtete Differenzschätzung</i> .....	81
7.2.2.1	Einleitung .....	81
7.2.2.2	Stichprobenumfang.....	82
7.2.2.3	Extrapolation.....	83
7.2.2.4	Genauigkeit .....	83
7.2.2.5	Bewertung.....	84
7.2.2.6	Beispiel .....	84
7.2.3	<i>Differenzschätzung – zwei Zeiträume</i> .....	89
7.2.3.1	Einleitung .....	89
7.2.3.2	Stichprobenumfang.....	90
7.2.3.3	Extrapolation.....	90
7.2.3.4	Genauigkeit .....	90
7.2.3.5	Bewertung.....	91
7.2.3.6	Beispiel .....	91
7.3	WERTBEZOGENES STICHPROBENVERFAHREN .....	96
7.3.1	<i>Standardansatz</i> .....	96
7.3.1.1	Einleitung .....	96
7.3.1.2	Stichprobenumfang.....	97
7.3.1.3	Stichprobenauswahl.....	99
7.3.1.4	Prognostizierter Fehler.....	99
7.3.1.5	Genauigkeit .....	100
7.3.1.6	Bewertung.....	101
7.3.1.7	Beispiel .....	102
7.3.2	<i>Geschichtetes wertbezogenes Stichprobenverfahren</i> .....	107
7.3.2.1	Einleitung .....	107
7.3.2.2	Stichprobenumfang.....	108
7.3.2.3	Stichprobenauswahl.....	109
7.3.2.4	Prognostizierter Fehler.....	110
7.3.2.5	Genauigkeit .....	111
7.3.2.6	Bewertung.....	112
7.3.2.7	Beispiel .....	112
7.3.3	<i>Wertbezogenes Stichprobenverfahren – zwei Zeiträume</i> .....	118
7.3.3.1	Einleitung .....	118
7.3.3.2	Stichprobenumfang.....	118
7.3.3.3	Stichprobenauswahl.....	121
7.3.3.4	Prognostizierter Fehler.....	121

7.3.3.5	Genauigkeit .....	123
7.3.3.6	Bewertung.....	123
7.3.3.7	Beispiel .....	124
7.3.4	<i>Vorsichtig angelegter Ansatz</i> .....	132
7.3.4.1	Einleitung .....	132
7.3.4.2	Stichprobenumfang.....	132
7.3.4.3	Stichprobenauswahl.....	134
7.3.4.4	Prognostizierter Fehler.....	134
7.3.4.5	Genauigkeit .....	135
7.3.4.6	Bewertung.....	137
7.3.4.7	Beispiel .....	138
7.4	NICHTSTATISTISCHE STICHPROBENVERFAHREN .....	143
7.4.1	<i>Überblick</i> .....	143
7.4.2	<i>Beispiel</i> .....	146
<b>8</b>	<b>AUSGEWÄHLTE THEMEN .....</b>	<b>148</b>
8.1	BESTIMMUNG DES VORAUSSICHTLICHEN FEHLERS .....	148
8.2	ZUSÄTZLICHE STICHPROBENZIEHUNG.....	150
8.2.1	<i>Ergänzende Stichprobe (wegen unzureichender Erfassung hoch riskanter Bereiche)</i> .....	150
8.2.2	<i>Zusätzliche Stichprobe (wegen uneindeutiger Prüfergebnisse)</i> .....	151
8.3	STICHPROBENNAHME IM LAUFE DES JAHRES .....	152
8.4	ÄNDERUNG DES STICHPROBENVERFAHRENS WÄHREND DES PROGRAMMZEITRAUMS.....	153
8.5	FEHLERQUOTEN .....	153
8.6	ZWEISTUFIGES STICHPROBENVERFAHREN .....	153
8.7	NEUBERECHNUNG DES KONFIDENZNIVEAUS.....	155
8.8	STICHPROBENVERFAHREN FÜR SYSTEMPRÜFUNGEN .....	156
8.8.1	<i>Einleitung</i> .....	156
8.8.2	<i>Stichprobenumfang</i> .....	158
8.8.3	<i>Extrapolation</i> .....	159
8.8.4	<i>Genauigkeit</i> .....	159
8.8.5	<i>Bewertung</i> .....	159
8.8.6	<i>Spezielle Methoden des Attributstichprobenverfahrens</i> .....	160
<b>ANHANG 1 – HOCHRECHNUNG VON ZUFALLSFEHLERN BEI FESTSTELLUNG SYSTEMBEDINGTER FEHLER</b>		<b>161</b>
.....		
1.	EINLEITUNG .....	161
2.	EINFACHE ZUFALLSSTICHPROBENVERFAHREN .....	162
2.2	<i>Schätzung des Mittelwerts pro Einheit</i> .....	162
2.3	<i>Verhältnisschätzung</i> .....	162
3.	DIFFERENZENSCHÄTZUNG .....	163
4.	WERTBEZOGENES STICHPROBENVERFAHREN – STANDARDANSATZ .....	164
4.1	<i>MUS-Standardansatz</i> .....	165
4.2	<i>MUS-Verhältnisschätzung</i> .....	166
5.	NICHTSTATISTISCHE STICHPROBENVERFAHREN .....	167
<b>ANHANG 2 – ZUVERLÄSSIGKEITSFAKTOREN FÜR DAS MUS .....</b>		<b>170</b>
<b>ANHANG 3 – WERTE DER STANDARDISIERTEN NORMALVERTEILUNG (Z).....</b>		<b>171</b>
<b>ANHANG 4 – FUNKTIONEN IN MS EXCEL ZUR UNTERSTÜTZUNG VON STICHPROBENVERFAHREN .....</b>		<b>172</b>
<b>ANHANG 5 – GLOSSAR .....</b>		<b>173</b>

AA	Audit Authority	PB	Prüfbehörde
ACR	Annual Control Report	JKB	Jährlicher Kontrollbericht
AE	Anticipated Error	AE	Voraussichtlicher Fehler
AR	Audit Risk	AR	Prüfungsrisiko
BP	Basic Precision	BP	Grundgenauigkeit
BV	Book Value (expenditure certified to the Commission in reference year)	BV	Buchwert (die im Bezugsjahr die gegenüber der Kommission bescheinigten Ausgaben)
CBV	Corrected Book Value	CBV	Berichtigter Buchwert
COCOF	Committee of the Coordination of Funds	COCOF	Koordinierungsausschuss für die Fonds
CR	Control Risk	CR	Kontrollrisiko
DR	Detection Risk	DR	Aufdeckungsrisiko
$E_t$	Individual errors in the sample	$E_t$	Einzelfehler in der Stichprobe
$\bar{E}$	Mean error of the sample	$\bar{E}$	Mittlerer Fehler der Stichprobe
EC	European Community	EG	Europäische Gemeinschaft
EE	Projected Error	EE	Prognostizierter Fehler
EDR	Extrapolated Deviation Rate	DER	Hochgerechnete Abweichungsquote
EF	Expansion Factor	EF	Expansionsfaktor
ETC	European Territorial Cooperation	ETZ	Europäische territoriale Zusammenarbeit
IA	Incremental Allowance	IA	Inkrementelle Toleranz
IR	Inherent Risk	IR	Inhärentes Risiko
IT	Information Technologies	IT	Informationstechnologien
LL	Lower Limit	LL	Untergrenze
RF	Reliability Factor	RF	Zuverlässigkeitsfaktor
MCS	Management and Control System	MCS	Verwaltungs- und Kontrollsystem

MUS	Monetary Unit Sampling	MUS	Wertbezogenes Stichprobenverfahren
PPS	Probability Proportional to Size	PPS	Wahrscheinlichkeit proportional zur Größe
RF	Reliability Factor	RF	Zuverlässigkeitsfaktor
SE	(Effective, i.e., after performing audit work) Sampling Error (precision)	SE	(Effektiver, d. h. nach Durchführung der Prüfung) Stichprobenfehler (Genauigkeit)
SI	Sampling Interval	SI	Stichprobenintervall
TE	Maximum Tolerable Error	TE	Maximal zulässiger Fehler
TPER	Total Projected Error Rate	TPER	Prognostizierte Gesamtfehlerquote
TPE	Total Projected Error	TPE	Prognostizierter Gesamtfehler
ULD	Upper Limit of Deviation	ULD	Obere Abweichungsgrenze
ULE	Upper Limit of Error	ULE	Obere Fehlergrenze

# 1 Einleitung

Die vorliegenden Leitlinien für statistische Stichproben zu Prüfzwecken wurden erstellt, um den Prüfbehörden in den Mitgliedstaaten einen aktualisierten Überblick über die geeignetsten und am häufigsten verwendeten Stichprobenverfahren zu geben und sie auf diese Weise bei der Umsetzung des Rechtsrahmens im derzeitigen und gegebenenfalls im nächsten Programmplanungszeitraum zu unterstützen.

Um die Anforderungen von Artikel 62 der Verordnung (EG) Nr. 1083/2006 des Rates, Artikel 16 und Anhang IV der Verordnung (EG) Nr. 1828/2006 der Kommission, Artikel 61 der Verordnung (EG) Nr. 1198/2006 des Rates, Artikel 42 und Anhang IV der Verordnung (EG) Nr. 498/2007 der Kommission zu erfüllen, wählt die Prüfbehörde auf der Grundlage ihres fachlichen Urteilsvermögens das geeignetste Stichprobenverfahren aus. Zweck dieses Leitfadens ist es also, den Prüfbehörden bei der Durchführung ihrer Prüfungsarbeit zu helfen.

Das ausgewählte Verfahren ist in der Prüfstrategie im Sinne von Artikel 62 Absatz 1 Buchstabe c der Verordnung (EG) Nr. 1083/2006 und Artikel 61 der Verordnung (EG) Nr. 1198/2006, die gemäß dem Muster in Anhang V der Verordnungen (EG) Nr. 1828/2006 und Nr. 498/2007 der Kommission zu erstellen ist, darzulegen. Alle Änderungen am ausgewählten Verfahren sind in nachfolgenden Versionen der Prüfstrategie anzugeben und der Kommission im nächsten jährlichen Kontrollbericht zu übermitteln.

Für die Planung der Prüfverfahren zur Erhebung von Prüfungsstichproben bieten die internationalen Rechnungslegungsstandards und neueste stichprobentheoretische Erkenntnisse Hinweise zur Verwendung von repräsentativen Stichproben und anderen Auswahlverfahren zu Prüfzwecken.

Die vorliegenden Leitlinien ersetzen die bisherigen Leitlinien zum selben Thema (COCOF 08/0021/02-DE vom 15.9.2008). Allerdings gilt das vorliegende Dokument unbeschadet der anderen ergänzenden Leitfäden der Kommission:

- Leitfaden zu den jährlichen Kontrollberichten und Stellungnahmen vom 18.2.2009, COCOF 09/0004/01-EN und EFFC/0037/2009-EN vom 23.2.2009;
- Leitfaden zur Behandlung von in den jährlichen Kontrollberichten gemeldeten Fehlern vom 7.12.2011, COCOF 11/0041/01-DE und EFFC/87/2012 vom 9.11.2012;
- Leitlinien zu einer einheitlichen Methode für die Bewertung von Verwaltungs- und Kontrollsystemen in den Mitgliedstaaten, COCOF 08/0019/01-DE und EFFC/27/2008 vom 12.9.2008.

Es wird empfohlen, weitere Lektüre zu diesen zusätzlichen Dokumenten hinzu zu ziehen, um einen vollständigen Überblick über die Leitlinien für die Anfertigung jährlicher Kontrollberichte zu erhalten.

## 2 Bezug auf die Rechtsgrundlagen – Rechtsrahmen

In Artikel 62 Absatz 1 Buchstaben a und b der Verordnung (EG) Nr. 1083/2006 des Rates vom 11. Juli 2006 mit allgemeinen Bestimmungen über den Europäischen Fonds für regionale Entwicklung, den Europäischen Sozialfonds und den Kohäsionsfonds sowie Artikel 61 der Verordnung (EG) Nr. 1198/2006 vom 27. Juli 2006 mit allgemeinen Bestimmungen über den Europäischen Fischereifonds wird darauf hingewiesen, dass die Prüfbehörde die Aufgabe hat, zu gewährleisten, dass das effektive Funktionieren des Verwaltungs- und Kontrollsystems geprüft wird, und sicherzustellen, dass Vorhaben anhand geeigneter Stichproben geprüft werden.

Die Verordnungen (EG) Nr. 1828/2006 vom 8. Dezember 2006 und Nr. 498/2007 vom 26. März 2007 (nachstehend „Verordnungen“) zur Festlegung von Durchführungsvorschriften zu den Verordnungen (EG) Nrn. 1083/2006 und 1198/2006 enthalten in den Artikeln 16<sup>1</sup> und 17<sup>2</sup> sowie in Anhang IV bzw. in den Artikeln 42 und 43 sowie in Anhang IV detaillierte Bestimmungen in Bezug auf die Stichprobenverfahren zur Prüfung von Vorhaben.

In den Verordnungen sind die Anforderungen an die Systemprüfungen und an die Prüfungen von Vorhaben festgelegt, die im Rahmen der Strukturfonds durchzuführen sind, sowie die Bedingungen für die Entnahme von Stichproben für die Prüfung von Vorhaben, die die Prüfbehörden bei der Aufstellung und Genehmigung des Stichprobenverfahrens beachten müssen. Sie enthalten bestimmte technische Kriterien für eine Zufallsstichprobe sowie Faktoren, die bei einer ergänzenden Zufallsstichprobe zu berücksichtigen sind.

Hauptzweck der Systemprüfungen und der Prüfungen von Vorhaben ist die Prüfung der effizienten Funktionsweise des Verwaltungs- und Kontrollsystems des operationellen Programms sowie der geltend gemachten Ausgaben<sup>3</sup>.

In den Verordnungen sind auch der Zeitplan für die Prüfungen und die Berichterstattung durch die Prüfbehörde festgelegt.

---

<sup>1</sup> Artikel 16 Absatz 1 lautet: „Die Prüfungen gemäß Artikel 62 Absatz 1 Buchstabe b der Verordnung (EG) Nr. 1083/2006 erfolgen ab 1. Juli 2008 jeweils für einen Zwölfmonatszeitraum anhand einer Stichprobe von Vorhaben, die nach einer von der Prüfbehörde aufgestellten oder genehmigten Methode gemäß Artikel 17 der vorliegenden Verordnung ausgewählt werden.“

<sup>2</sup> Artikel 17 Absatz 2 lautet: „Das zur Auswahl der Stichprobe und für Schlussfolgerungen aus den Ergebnissen herangezogene Verfahren muss international anerkannte Prüfstandards berücksichtigen und dokumentiert sein. Unter Berücksichtigung der Ausgabenbeträge, der Zahl und Art der Vorhaben und anderer relevanter Faktoren bestimmt die Prüfbehörde das anzuwendende angemessene statistische Stichprobenverfahren. Die technischen Parameter der Stichprobe werden gemäß Anhang IV festgelegt.“

<sup>3</sup> Artikel 62 Absatz 1 Buchstabe b der Verordnung (EG) Nr. 1083/2006 des Rates (ABl. L 210 vom 31.7.2006, S. 25) und Artikel 61 Absatz 1 Buchstabe c der Verordnung (EG) Nr. 1198/2006 (ABl. L 223 vom 15.8.2006, S. 1).

Bei den Vorhaben werden die Ausgaben geprüft, die im Bezugsjahr (Bezugszeitraum für die Zufallsstichprobe) bei der Kommission geltend gemacht werden. Um eine jährliche Stellungnahme abgeben zu können, sollte die Prüfbehörde die Prüfungen einschließlich der Systemprüfungen und der Prüfungen der Vorhaben gründlich vorbereiten.

### 3 Prüfungsrisikomodell und Prüfverfahren

#### 3.1 Risikomodell

Das Prüfungsrisiko bezeichnet das Risiko, dass der Prüfer trotz wesentlicher Fehlangaben in der Ausgabenerklärung ein uneingeschränktes Prüfungsurteil ausspricht.

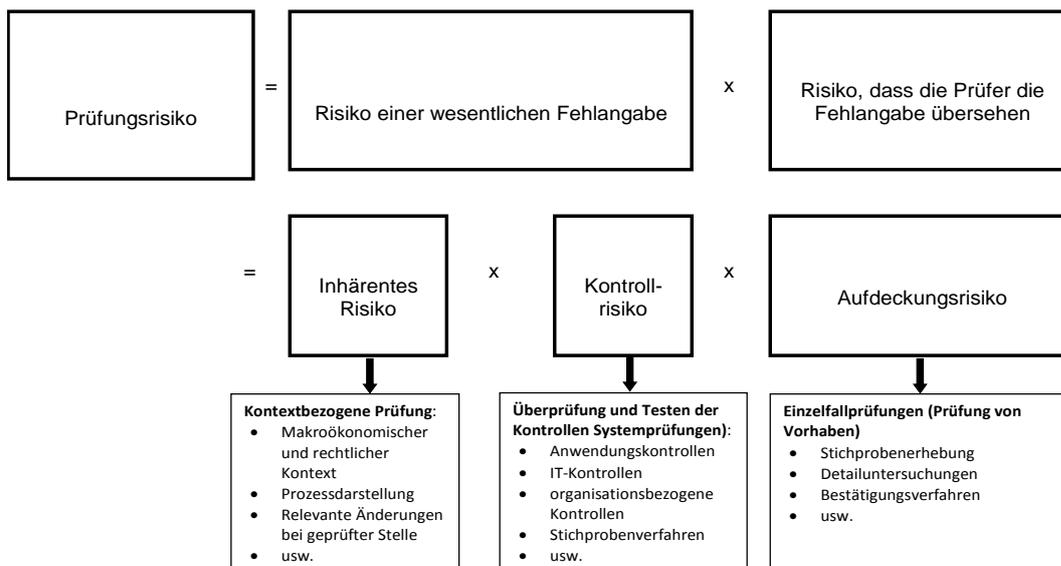


Abb. 1. Prüfungsrisikomodell

Die drei Bestandteile, aus denen sich das Prüfungsrisiko (AR – audit risk) zusammensetzt, sind das inhärente Risiko (IR – inherent risk), das Kontrollrisiko (CR – control risk) und das Aufdeckungsrisiko (DR – detection risk). Somit ist das Prüfungsrisikomodell

$$AR = IR \times CR \times DR.$$

Dabei gilt Folgendes:

- **IR**, das inhärente Risiko, ist das Risiko, dass in den der Kommission bescheinigten Ausgabenaufstellungen oder in den zugrunde liegenden Aggregationsebenen bei nicht vorhandenen internen Kontrollverfahren ein wesentlicher Fehler vorkommen kann. Das inhärente Risiko ist direkt mit der Art der Tätigkeit der geprüften Stelle verknüpft und hängt von externen Faktoren (kultureller, politischer, wirtschaftlicher, geschäftstätigkeits-, kunden- und lieferantenbezogener und ähnlicher Art) sowie von internen Faktoren (Art der Organisation, Verfahren, Kompetenz der Mitarbeiter, Veränderungen in den

Abläufen oder Neubesetzung von Führungsstellen usw.) ab. Das inhärente Risiko muss vor Inangriffnahme detaillierter Prüfungsverfahren bewertet werden (Befragungen des Managements und des Schlüsselpersonals, Überprüfung kontextbezogener Informationen, z. B. Organigramme, Handbücher und interne/externe Unterlagen). Für die Strukturfonds und den Fischereifonds wird das inhärente Risiko üblicherweise mit einem höheren Prozentsatz angesetzt.

- **CR**, das Kontrollrisiko, ist das Risiko, dass in den der Kommission bescheinigten Ausgabenaufstellungen oder in den zugrunde liegenden Aggregationsebenen mit den internen Kontrollverfahren der Verwaltung ein wesentlicher Fehler nicht verhindert, entdeckt oder behoben wird. Die Kontrollrisiken sind mit dem Grad verknüpft, zu dem die inhärenten Risiken beherrscht (kontrolliert) werden, und dementsprechend mit dem internen Kontrollsystem, das u. a. Anwendungskontrollen, IT-Kontrollen und organisatorische Kontrollen umfasst. Die Kontrollrisiken können anhand der **Systemprüfung** bewertet werden – einer detaillierten Prüfung der Kontrollen und Berichte, die Aufschluss darüber geben soll, wie wirksam Konzeption und Funktionsweise eines Kontrollsystems sind, um wesentliche Fehler zu verhindern oder zu entdecken, sowie darüber, wie die Organisation in der Lage ist, Daten zu erfassen, zu verarbeiten, zusammenzufassen und zu melden.

Das Produkt aus dem inhärenten und dem Kontrollrisiko (d. h. **IR × CR**) wird als das Risiko eines wesentlichen Fehlers bezeichnet. Das **Risiko eines wesentlichen Fehlers** hängt mit dem Ergebnis der **Systemprüfungen** zusammen.

- **DR**, das Aufdeckungsrisiko, ist das Risiko, dass der Prüfer in den der Kommission bescheinigten Ausgabenaufstellungen oder in den zugrunde liegenden Aggregationsebenen einen wesentlichen Fehler nicht aufdeckt. Aufdeckungsrisiken sind mit dem Grad verknüpft, zu dem die Prüfungen ordnungsgemäß durchgeführt werden, also mit dem angewandten Stichprobenverfahren, der Kompetenz der Mitarbeiter, den Prüfmethoden, den Prüfwerkzeugen usw. Die Aufdeckungsrisiken sind mit der Durchführung von Prüfungen der Vorhaben verbunden. Dazu gehören vertiefte Prüfungen von Details oder Transaktionen im Zusammenhang mit den Vorhaben eines Programms, in der Regel auf der Grundlage einer Stichprobenprüfung der Vorhaben.

**FIGURE 5.2**  
An Illustration of  
Audit Risk

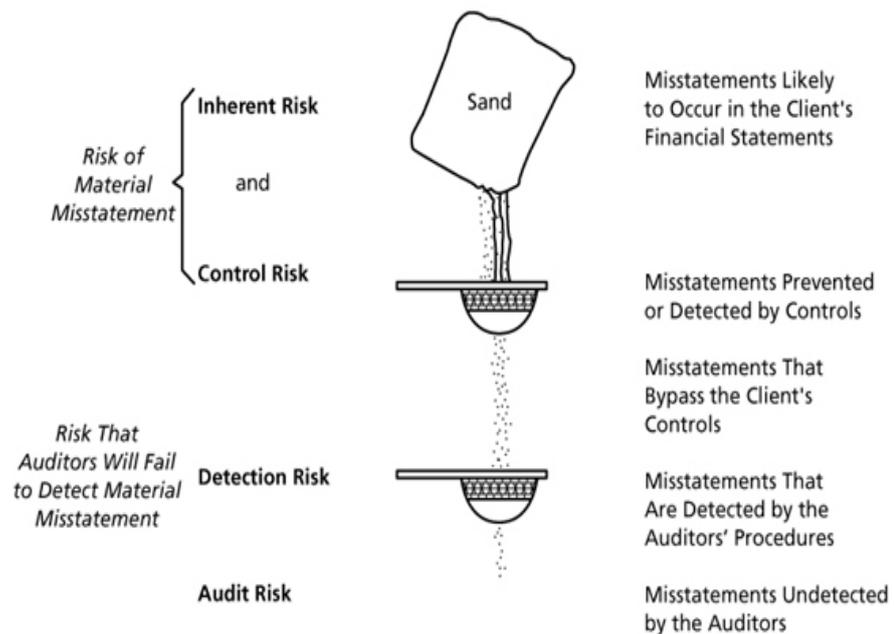


Abb. 2 Veranschaulichung des Prüfungsrisikos (nach einer unbekanntem Quelle)

Das Sicherheitsmodell ist das Gegenteil des Risikomodells. Wird das Prüfungsrisiko auf 5 % eingeschätzt, so wird von einer Prüfungssicherheit von 95 % ausgegangen.

Die Verwendung des Prüfungsrisiko- bzw. des Prüfungssicherheitsmodells bezieht sich auf die Planung und die zugrunde liegende Zuteilung der Ressourcen für ein bestimmtes operationelles Programm oder für mehrere operationelle Programme, wobei ein zweifacher Zweck verfolgt wird:

- Gewährleistung eines hohen Sicherheitsniveaus: Die Sicherheit wird auf einem bestimmten Niveau gewährleistet, beispielsweise liegt bei einer Sicherheit von 95 % das Prüfungsrisiko bei 5 %.
- Durchführung effizienter Prüfungen: Bei einem gegebenen Sicherheitsniveau von beispielsweise 95 % sollte der Prüfer seine Prüfungsverfahren unter Berücksichtigung des inhärenten Risikos und des Kontrollrisikos entwickeln. Das Prüfteam hat somit die Möglichkeit, in einigen Bereichen weniger eingehende Prüfungen vorzunehmen, um sich auf die Bereiche mit höherem Risiko zu konzentrieren.

Zu beachten ist, dass die Festlegung eines Wertes für das Aufdeckungsrisiko, der wiederum den Stichprobenumfang für die Stichprobenprüfung der Vorhaben beeinflusst, ein unmittelbares Ergebnis darstellt, sofern das *IR* und das *CR* zuvor bewertet wurden. So gilt

$$AR = IR \times CR \times DR \Rightarrow DR = \frac{AR}{IR \times CR}$$

wobei das **AR** in der Regel mit 5 % angesetzt wird, **IR** und **CR** werden vom Prüfer bewertet.

### **Erläuterung**

**Geringe Kontrollsicherheit:** Im Falle eines beabsichtigten und akzeptierten Prüfungsrisikos von 5 % sollte der Prüfer immer dann ein sehr niedriges Aufdeckungsrisiko von etwa 10 % anstreben, wenn es sich um eine Stelle mit hohem Risiko handelt, der es an geeigneten internen Kontrollverfahren zur Risikobeherrschung mangelt, d. h. wenn das inhärente Risiko (= 100 %) und das Kontrollrisiko (= 50 %) hoch sind. Um das Aufdeckungsrisiko niedrig zu halten, müssen der Umfang der vertieften Prüfungen und dementsprechend der Stichprobenumfang groß sein.

$$AR = IR \times CR \times DR = 1 \times 0.5 \times 0.1 = 0.05$$

**Hohe Kontrollsicherheit:** In einem andersgearteten Umfeld mit hohem inhärenten Risiko (100 %), in dem jedoch geeignete Kontrollen eingerichtet sind, kann das Kontrollrisiko auf 12,5 % geschätzt werden. Um auf ein Prüfungsrisiko von 5 % zu kommen, könnte das Aufdeckungsrisikoniveau bei 40 % liegen; der Prüfer kann somit durch die Reduzierung des Stichprobenumfangs ein höheres Risiko eingehen. Dies würde schließlich eine weniger detaillierte und finanziell weniger aufwendige Prüfung bedeuten.

$$AR = IR \times CR \times DR = 1 \times 0.125 \times 0.4 = 0.05$$

Trotz unterschiedlicher Umfeldler ist das Prüfungsrisiko in beiden Beispielen gleich, nämlich 5 %.

Zur Planung der Prüfungstätigkeit sollten anhand einer Ablaufplanung die verschiedenen Risikoniveaus abgeschätzt werden, wobei in erster Linie das inhärente Risiko abzuschätzen ist, da in Abhängigkeit davon das Kontrollrisiko überprüft werden muss. Anhand dieser beiden Faktoren setzt das Prüfteam das Aufdeckungsrisiko fest und wählt geeignete Verfahren für die Prüfung der Vorhaben aus.

So sehr das Prüfungsrisikomodell als Rahmen für die Konzipierung des Prüfplans und die Ressourcenzuteilung dienen kann, so schwierig wird es sich mitunter in der Praxis erweisen, eine genaue Quantifizierung des inhärenten und des Kontrollrisikos vorzunehmen.

Das Sicherheitsniveau/Konfidenzniveau für die Prüfung von Vorhaben hängt in erster Linie von der Qualität des internen Kontrollsystems ab. Zur Beurteilung der Risikokomponenten verwenden die Prüfer aufgrund ihres Wissens und ihrer Erfahrung anstelle genauer Wahrscheinlichkeitsangaben eher Begriffe wie NIEDRIG, MODERAT/DURCHSCHNITTLLICH oder HOCH. Werden während der

Systemprüfung bedeutende Schwachstellen festgestellt, so liegt ein hohes Kontrollrisiko vor, was zur Folge hätte, dass die sich aus dem System ergebende Sicherheit niedrig wäre. Bestehen keine bedeutenden Schwachstellen, so liegt ein niedriges Kontrollrisiko vor, und sollte auch das inhärente Risiko niedrig sein, so wäre die aus dem System resultierende Sicherheit hoch.

Im Zusammenhang mit den Strukturfonds und dem Fischereifonds heißt es in Anhang IV der Verordnungen: „Um eine hohe Sicherheit zu erlangen, also ein geringes Prüfungsrisiko, muss die Prüfbehörde die Ergebnisse der Systemprüfungen (*dies entspricht der inhärenten Sicherheit und der Kontrollsicherheit*) und der Prüfungen von Vorhaben (*Aufdeckungssicherheit*) kombinieren. Das kombinierte Sicherheitsniveau, das sich aus den Systemprüfungen und den Prüfungen von Vorhaben ergibt, muss hoch sein. Die Prüfbehörde muss im jährlichen Kontrollbericht beschreiben, auf welche Weise die Sicherheit erlangt wurde.“ Es wird davon ausgegangen, dass die Prüfbehörde eine Sicherheit von 95 % erreichen muss, um in ihrem Prüfungsurteil von „hinreichender Gewähr“ sprechen zu können. Das Prüfungsrisiko liegt dann bei 5 %. Den Verordnungen liegt die Annahme zugrunde, dass selbst bei einem mangelhaft funktionierenden System eine Mindestsicherheit erreicht wird (d. h. das Risiko der wesentlichen Fehlangebe höchstens 50 % beträgt) und dass die restliche Sicherheit (90 %) durch die Prüfung von Vorhaben zustande kommen muss.

Falls die Prüfbehörde im Ausnahmefall zu dem Schluss kommt, dass das System keinerlei Sicherheit bietet, muss durch die Prüfung von Vorhaben ein Sicherheitsniveau/Konfidenzniveau von 95 % erzielt werden.

Werden im Zuge der Systemprüfungen bedeutende Schwachstellen vorgefunden, so ist, wie bereits festgestellt, das Risiko des Auftretens eines wesentlichen Fehlers (Kontrollrisiken in Verbindung mit inhärenten Risiken) hoch und das Sicherheitsniveau niedrig. Laut Anhang IV der Verordnungen muss bei niedrigem Sicherheitsniveau, das sich aus dem System ergibt, das Konfidenzniveau für die Stichprobenprüfung von Vorhaben mindestens 90 % betragen.

Sind die Systeme frei von bedeutenden Schwachstellen, so ist die Gefahr des Auftretens wesentlicher Fehler gering, und die Sicherheit, die sich aus dem System ergibt, wäre hoch, d. h. dass das Konfidenzniveau für die Stichprobenprüfung von Vorhaben mindestens 60 % erreichen muss.

Abschnitt 3.2 enthält einen ausführlichen Rahmen für die Auswahl des Sicherheitsniveaus-/Konfidenzniveaus für die Prüfung von Vorhaben.

### 3.2 Sicherheits-/Konfidenzniveau für die Prüfung von Vorhaben

Gemäß Anhang IV der Verordnungen sollen vertiefte Prüfungen an Stichproben durchgeführt werden, deren Umfang von dem Konfidenzniveau abhängt, das entsprechend dem aus den Systemprüfungen resultierenden Sicherheitsniveau bestimmt wird, d. h.

- mindestens 60 % bei hoher Sicherheit;
- durchschnittliche Sicherheit (für dieses Sicherheitsniveau ist in der Verordnung der Kommission kein Konfidenzniveau angegeben, obwohl ein Niveau von 70 % bis 80 % angeraten ist);
- mindestens 90 % bei niedriger Sicherheit.

Darüber hinaus heißt es in Anhang IV, dass die Prüfbehörde Kriterien für Systemprüfungen erstellt, um die Zuverlässigkeit der Verwaltungs- und Kontrollsysteme zu bestimmen. Diese Kriterien beinhalten eine quantifizierte Bewertung aller wichtigen Elemente des Systems (Kernanforderungen) und beziehen die an Verwaltung und Kontrolle des operationellen Programms beteiligten wichtigsten Behörden und zwischengeschalteten Stellen ein.

Die Kommission hat gemeinsam mit dem Europäischen Rechnungshof einen Leitfaden für die Methodik zur Bewertung der Verwaltungs- und Kontrollsysteme erarbeitet. Dieser lässt sich sowohl auf die Mainstream- als auch auf die ETZ-Programme anwenden. Der Prüfbehörde wird empfohlen, dieser Methodik Rechnung zu tragen.

Laut dieser Methodik sind vier Zuverlässigkeitsniveaus vorgesehen:

- sehr zuverlässig, nur geringfügige Verbesserungen notwendig;
- zuverlässig, jedoch einige Verbesserungen notwendig;
- teilweise zuverlässig, erhebliche Verbesserungen notwendig;
- im Wesentlichen unzuverlässig.

Gemäß der Verordnung wird das Konfidenzniveau für Stichprobenprüfungen anhand des in den Systemprüfungen festgestellten Zuverlässigkeitsniveaus ermittelt.

Wie bereits erwähnt, sind in der Verordnung für Systeme nur drei Sicherheitsniveaus vorgesehen: hoch, durchschnittlich und niedrig. Das durchschnittliche Niveau entspricht im Prinzip den Kategorien 2 und 3 der Methodik zur Bewertung der Verwaltungs- und Kontrollsysteme, die eine genauere Differenzierung zwischen den beiden Extremen hoch/„sehr zuverlässig“ und niedrig/„unzuverlässig“ ermöglichen.

Die empfohlenen Beziehungen sind in der nachstehenden Tabelle zusammengefasst:

<b>Sicherheitsniveau aus den Systemprüfungen</b>	<b>Entsprechende Zuverlässigkeit/Sicherheit aus den Systemprüfungen laut Verordnung</b>	<b>Konfidenzniveau</b>	<b>Aufdeckungsrisiko</b>
sehr zuverlässig, nur geringfügige Verbesserungen notwendig	hoch	mindestens 60 %	weniger als oder gleich 40 %
zuverlässig, jedoch einige Verbesserungen notwendig	durchschnittlich	70 %	30 %
teilweise zuverlässig, erhebliche Verbesserungen notwendig	durchschnittlich	80 %	20 %
im Wesentlichen unzuverlässig	niedrig	mindestens 90 %	maximal 10 %

Tabelle 1. Konfidenzniveau für die Prüfung von Vorhaben entsprechend der sich aus den Systemprüfungen ergebenden Sicherheit

Zu Beginn des Programmzeitraums ist von einem niedrigen Sicherheitsniveau auszugehen, da bis dahin noch gar keine oder nur wenige Systemprüfungen stattgefunden haben. Daher darf das Konfidenzniveau nicht unter 90 % liegen. Falls die Systeme gegenüber dem vorherigen Programmzeitraum allerdings unverändert sind und ihr Sicherheitsniveau durch Prüfungen zuverlässig nachgewiesen ist, ist es den Mitgliedstaaten freigestellt, ein anderes Konfidenzniveau (zwischen 60 % und 90 %) zu verwenden. Das Niveau kann auch während eines Programmzeitraums abgesenkt werden, wenn keine wesentlichen Fehler festgestellt werden oder die Systeme im Laufe der Zeit nachweislich verbessert wurden. Die Methodik zur Bestimmung des Konfidenzniveaus ist in der Prüfstrategie unter Angabe der dazu herangezogenen Prüfnachweise zu erläutern.

Die Festlegung eines angemessenen Konfidenzniveaus ist für die Prüfung von Vorhaben von größter Bedeutung, da der Stichprobenumfang maßgeblich von diesem Niveau abhängt (mit dem Ansteigen des Konfidenzniveaus erhöht sich auch der Stichprobenumfang). Die Verordnungen bieten daher die Möglichkeit, für Systeme mit einer niedrigen Fehlerquote (und somit hoher Sicherheit) das Konfidenzniveau abzusenken und somit das Arbeitspensum zu reduzieren, während für Systeme mit einer potenziell hohen Fehlerquote (und somit niedriger Sicherheit) auch weiterhin ein hohes Konfidenzniveau (und folglich ein größerer Stichprobenumfang) gefordert wird.

## **Bestimmung des Sicherheitsniveaus bei der Zusammenfassung von Programmen**

Die Prüfbehörde sollte, wenn sie Programme zusammenfasst, nur **ein** Sicherheitsniveau verwenden.

Ergeben sich bei den Systemprüfungen der zusammengefassten Programme unterschiedliche Schlussfolgerungen hinsichtlich der Funktionsweise der einzelnen Programme, so bestehen folgende Möglichkeiten:

- Aufteilung in zwei (oder mehr) Gruppen, beispielsweise die erste für Programme mit niedriger Sicherheit (Konfidenzniveau 90 %) und die zweite für Programme mit hoher Sicherheit (Konfidenzniveau 60 %) usw. Die beiden Gruppen werden als zwei unterschiedliche Grundgesamtheiten behandelt. Es müssten entsprechend mehr Kontrollen durchgeführt werden, weil von jeder Gruppe eine Stichprobe genommen werden muss.
- Anwendung des niedrigsten Sicherheitsniveaus eines Einzelprogramms auf die gesamte Programmgruppe. Die Programmgruppe wird als eine Grundgesamtheit behandelt. Prüfungsschlussfolgerungen werden in diesem Fall in Bezug auf die gesamte Programmgruppe gezogen, so dass Schlussfolgerungen zu den Einzelprogrammen normalerweise nicht möglich sind.

Im letztgenannten Fall ist die Verwendung eines nach Programmen geschichteten Stichprobenplans möglich, der normalerweise einen geringeren Stichprobenumfang zulässt. Aber auch hier ist nur ein einziges Sicherheitsniveau zu verwenden, und Schlussfolgerungen sind auch weiterhin nur für die gesamte Programmgruppe möglich.

## **4 Behandlung von Fehlern**

Wie bereits an anderer Stelle erklärt, sollte dieses Dokument im Zusammenhang mit dem Leitfaden zur Behandlung von in den jährlichen Kontrollberichten gemeldeten Fehlern betrachtet werden, in dem sich eine detaillierte Darstellung verschiedener Fehlerarten findet. Die genaue Analyse und die Definitionen der verschiedenen Fehlerarten sollen daher in diesem Abschnitt nicht noch einmal umfassend wiedergegeben werden, vielmehr wird nur eine kurze Zusammenfassung von Themen vorgenommen, die sich direkt auf das Stichprobenverfahren und die Hochrechnung des Gesamtfehlers auswirken.

## **4.1 Systembedingte und bekannte Fehler**

### **Systembedingte Fehler**

Systembedingte Fehler sind in der geprüften Stichprobe festgestellte Fehler, die sich auf die nicht geprüfte Grundgesamtheit auswirken und unter klar definierten und ähnlichen Umständen auftreten. Diese Fehler weisen im Allgemeinen ein gemeinsames Merkmal auf (z. B. Art des Vorhabens, zwischengeschaltete Stelle, Ort oder Zeitraum). Sie gehen im Allgemeinen auf unwirksame Kontrollverfahren in (Teilen von) Verwaltungs- und Kontrollsystemen zurück. Wird ein potenzieller systembedingter Fehler festgestellt, so sind weitere Arbeiten erforderlich, um die gesamte Reichweite des Fehlers zu ermitteln und eine Quantifizierung vorzunehmen. Das bedeutet, dass alle Situationen, in denen ein Fehler derselben Art wie in der Zufallsstichprobe auftreten könnte, ermittelt werden müssen, damit die Gesamtwirkung dieses Fehlers auf die Grundgesamtheit eingegrenzt werden kann.

Wenn die Prüfbehörde hinreichende Gewähr hat, dass die von systembedingten Fehlern betroffene Teilgesamtheit genau eingegrenzt ist und keine weiteren Einheiten in der Grundgesamtheit für ähnliche Fehler anfällig sind, sollten zur Bestimmung des Gesamtfehlers der Betrag der systembedingten Fehler, der prognostizierte Zufallsfehler und der unkorrigierte anomale Fehler addiert werden<sup>4</sup>. Bei der Extrapolation der in der Stichprobe gefundenen Zufallsfehler auf die Grundgesamtheit sollte die Prüfbehörde berücksichtigen, dass die Zufallsfehler nur auf die restlichen Ausgaben extrapoliert werden (Gesamtausgaben minus Betrag der systembedingten Fehler). Zu diesem Zweck erfolgt erforderlichenfalls eine Umwandlung der für die Hochrechnung der Fehler und die Berechnung der Genauigkeit verwendeten Formeln, wie sie in Anhang 1 enthalten sind. Der Betrag der in der Stichprobe festgestellten systembedingten Fehler gilt nicht als Zufallsfehler und wird im prognostizierten Zufallsfehler nicht berücksichtigt. Gleichwohl sollten alle Zufallsfehler, die in den von systembedingten Fehlern betroffenen Vorhaben gefunden werden, extrapoliert und im prognostizierten Zufallsfehler mit erfasst werden.

### **Bekannte Fehler**

Es ist auch möglich, dass der Prüfer aufgrund eines in der Stichprobe festgestellten Fehlers einen oder mehrere Fehler außerhalb der Stichprobe feststellt. Diese werden als „bekannte Fehler“ eingestuft. Wenn beispielsweise festgestellt wurde, dass ein Vertrag nicht den Vorschriften für die Vergabe öffentlicher Aufträge entspricht, ist es wahrscheinlich, dass ein Teil der damit zusammenhängenden unregelmäßigen Ausgaben in einem Zahlungsantrag oder einer Rechnung geltend gemacht wurde, die Teil der Stichprobe sind, und dass die restlichen Ausgaben in Zahlungsanträgen oder Rechnungen geltend gemacht wurden, die nicht in der Stichprobe enthalten waren. Der

---

<sup>4</sup> Und nicht der Ausgabenbetrag der Teilgesamtheit, auf die sich die systembedingten Fehler beziehen.

Hauptunterschied zu systembedingten Fehlern besteht darin, dass die vom bekannten Fehler betroffenen Ausgaben normalerweise auf ein Vorhaben begrenzt sind.

Wenn die Prüfbehörde die bekannten Fehler zur prognostizierten Gesamtfehlerquote (TPER – siehe Abschnitt 4.4)<sup>5</sup> hinzuaddiert, sollte sie berücksichtigen, dass die Zufallsfehler in der Stichprobe (einschließlich des Fehlers, der zur Feststellung des bekannten Fehlers führte)<sup>6</sup> nur auf die restlichen Ausgaben extrapoliert werden (Gesamtausgaben abzüglich des Betrags der bekannten Fehler). Zu diesem Zweck erfolgt erforderlichenfalls eine Umwandlung der für die Hochrechnung der Fehler und die Berechnung der Genauigkeit verwendeten Formeln, wie sie in Anhang 1 enthalten sind.

Eine einfachere Vorgehensweise ist die Extrapolation der Zufallsfehler in der Stichprobe (einschließlich des Fehlers, der zur Feststellung des bekannten Fehlers führte) auf die Gesamtausgaben (ohne Abzug des Betrags der bekannten Fehler). In diesem Fall wird der bekannte Fehler nicht zur TPER hinzugerechnet.

## **4.2 Zufallsfehler**

Fehler, die nicht als systembedingt oder bekannt gelten, werden als Zufallsfehler eingestuft. Dabei wird davon ausgegangen, dass Zufallsfehler, die in der geprüften Stichprobe gefunden wurden, wahrscheinlich auch in der nicht geprüften Grundgesamtheit vorhanden sind, da es sich um eine repräsentative Stichprobe handelt. Diese Fehler sind daher bei der Berechnung der Fehler zu berücksichtigen.

## **4.3 Anomale Fehler**

Ein Fehler, der für die Grundgesamtheit nachweislich nicht repräsentativ ist, wird als anomaler Fehler bezeichnet. Da eine statistische Stichprobe für die Grundgesamtheit repräsentativ ist, sollten „anomale Fehler“ nur in sehr außergewöhnlichen, gut begründeten Fällen akzeptiert werden.

Der häufige und unbegründete Rückgriff auf dieses Konzept würde die Zuverlässigkeit der Prüfstellnahme beeinträchtigen.

Die Prüfbehörde hat im jährlichen Kontrollbericht ein hohes Maß an Sicherheit dafür zu bieten, dass ein solcher anomaler Fehler für die Grundgesamtheit nicht repräsentativ ist (nicht anderswo in der Grundgesamtheit auftritt), und zu erläutern, welche zusätzlichen

---

<sup>5</sup> Entsprechend den Vorgaben im Leitfaden zur Behandlung von in den jährlichen Kontrollberichten gemeldeten Fehlern (COCOF 11/0041/01-DE vom 7.12.2011 und EFFC/87/2012 vom 9.11.2012).

<sup>6</sup> Anders als bei den systembedingten Fehlern lässt sich bei den bekannten Fehlern durch eine Eingrenzung nicht sicherstellen, dass nicht auch andere Vorhaben von dieser Fehlerart oder sonstigen Unregelmäßigkeiten betroffen sind.

Prüfverfahren sie durchgeführt hat, um auf die Existenz von anomalen Fehlern gemäß ISA Nr. 530 zu schließen. In ISA Nr. 530 heißt es dazu:

*„A.19. Wenn eine falsche Darstellung als Anomalie ermittelt wurde, kann diese bei der Hochrechnung falscher Darstellungen auf die Grundgesamtheit ausgeschlossen werden. Die Auswirkungen solcher falschen Darstellungen, sofern sie nicht korrigiert wurden, müssen jedoch weiterhin zusätzlich zu der Hochrechnung der nicht anomalen falschen Darstellungen berücksichtigt werden.“*

*„A.22. Bei Einzelfallprüfungen stellt die hochgerechnete falsche Darstellung zuzüglich etwaiger anomaler falscher Darstellungen die beste Schätzung des Abschlussprüfers zu falschen Darstellungen in der Grundgesamtheit dar. Wenn die hochgerechnete falsche Darstellung zuzüglich etwaiger anomaler falscher Darstellungen die tolerierbare falsche Darstellung überschreitet, bildet die Stichprobe keine hinreichende Grundlage für Schlussfolgerungen über die geprüfte Grundgesamtheit. (...)“*

Beschließt die Prüfbehörde also, einen anomalen Fehler von der Berechnung des prognostizierten Fehlers auszuschließen, so sollte der Betrag des anomalen Fehlers in die Berechnung der prognostizierten Gesamtfehlerquote einfließen, wenn er nicht korrigiert wurde. Eine Möglichkeit für diese Berechnung besteht darin, das betreffende Vorhaben sowohl aus der Grundgesamtheit als auch aus der Stichprobe herauszunehmen, bevor der prognostizierte Fehler anhand der Stichprobe berechnet wird. Nach der Bestimmung des prognostizierten Fehlers ist der Betrag des einzelnen unkorrigierten anomalen Fehlers hinzuaddieren, um zum prognostizierten Gesamtfehler zu gelangen. Wurde der anomale Fehler korrigiert, so bleibt er bei der prognostizierten Gesamtfehlerquote unberücksichtigt.

#### **4.4 Prognostizierte Gesamtfehlerquote (TPER)**

Die Prüfbehörde sollte im jährlichen Kontrollbericht die prognostizierte Gesamtfehlerquote angeben und diese mit der Erheblichkeitsschwelle vergleichen, um Schlussfolgerungen für die Grundgesamtheit zu ziehen (siehe Artikel 17 Absatz 4 zweiter Unterabsatz der Verordnung (EG) Nr. 1828/2006 der Kommission und Artikel 43 Absatz 4 der Verordnung (EG) Nr. 498/2007 der Kommission). Im zweiten Unterabsatz der erstgenannten Verordnung heißt es: „Bei operationellen Programmen, deren prognostizierte Fehlerquote über der Erheblichkeitsschwelle liegt, analysiert die Prüfbehörde die Signifikanz und ergreift die erforderlichen Maßnahmen, darunter geeignete Empfehlungen, die im jährlichen Kontrollbericht mitgeteilt werden.“

Die prognostizierte Gesamtfehlerquote entspricht der geschätzten Wirkung der Fehler in den Verwaltungs- und Kontrollsystemen als prozentualer Anteil an der Grundgesamtheit. Diese Quote muss die von der Prüfbehörde durchgeführte Analyse in Bezug auf die Fehler widerspiegeln, die im Zusammenhang mit den Prüfungen von Vorhaben gemäß Artikel 62 Absatz 1 Buchstabe b der Verordnung (EG) Nr. 1083/2006

und Artikel 61 Absatz 1 Buchstabe b der Verordnung (EG) Nr. 1198/2006 festgestellt wurden.

**Der prognostizierte Gesamtfehler (TE) entspricht der Summe folgender Fehler: prognostizierte Zufallsfehler, systembedingte Fehler und nicht berichtigte anomale Fehler. Bekannte Fehler können dem TE ebenfalls hinzuaddiert werden, ein einfacheres Vorgehen ist jedoch in Abschnitt 4.1 dargelegt.**

Werden in der geprüften Stichprobe systembedingte Fehler festgestellt und wird deren Umfang in der nicht geprüften Grundgesamtheit genau eingegrenzt, so werden die mit der Grundgesamtheit zusammenhängenden systembedingten Fehler zum prognostizierten Gesamtfehler addiert. Kann eine solche Eingrenzung nicht vor der Übermittlung des jährlichen Kontrollberichts erfolgen, so sind die systembedingten Fehler – für die Berechnung des prognostizierten Zufallsfehlers – als Zufallsfehler zu behandeln.

Bei den Zufallsfehlern fällt die Hochrechnung der Fehler je nach gewähltem und in der Prüfstrategie beschriebenem Stichprobenverfahren unterschiedlich aus. Die Prüfbehörde sollte alle Fehler quantifizieren und in die prognostizierte Gesamtfehlerquote aufnehmen, mit Ausnahme der korrigierten anomalen Fehler. Ohne diese Quantifizierung kann die Fehlerquote nicht als zuverlässig angesehen werden, da sie wahrscheinlich zu niedrig ist. Im Allgemeinen müssen alle festgestellten Fehler bei der Berechnung der prognostizierten Gesamtfehlerquote berücksichtigt werden.

Wenn der systembedingte Fehler **Ausgaben** betrifft, die **in früheren Jahren gemeldet** wurden, und diese Fehler bei früher gemeldeten prognostizierten Gesamtfehlerquoten unberücksichtigt geblieben waren, sollte die Prüfbehörde diese Fehlerquoten entsprechend revidieren, wobei die Durchführung der Korrekturmaßnahmen Auswirkungen auf die Revision haben kann. Werden bei den im Jahr N gemeldeten Ausgaben und auch bei den im nachfolgenden Jahr/in den nachfolgenden Jahren gemeldeten Ausgaben systembedingte oder bekannte Fehler festgestellt, so sollten diese bei der Berechnung der prognostizierten Gesamtfehlerquote im nachfolgenden Jahr/in den nachfolgenden Jahren berücksichtigt werden.

Wie aus dem Leitfaden zu den jährlichen Kontrollberichten (COCOF 09/0004/01-EN vom 18.2.2009 und EFF/0037/2009-EN) hervorgeht, werden bei Systemprüfungen (Kontrolltests) festgestellte Fehler nicht zum prognostizierten Gesamtfehler addiert, sondern sollten berichtet und in Abschnitt 4 des jährlichen Kontrollberichts mitgeteilt werden. Natürlich sollten die Schlussfolgerungen aus den Systemprüfungen ebenso wie die Ergebnisse der Prüfung von Vorhaben in das im jährlichen Kontrollbericht enthaltene Prüfungsurteil einfließen.

## **5 Statistische Begriffe im Zusammenhang mit der Prüfung von Vorhaben**

### **5.1 Stichprobenverfahren**

Das Stichprobenverfahren umfasst zwei Elemente, und zwar den Stichprobenplan (z. B. gleiche Wahrscheinlichkeit, Wahrscheinlichkeit proportional zur Größe) und das Verfahren der Hochrechnung (Schätzung). Zusammen bilden diese beiden Elemente den Rahmen für die Berechnung des Stichprobenumfangs.

Die bekanntesten Stichprobenverfahren, die für die Prüfung von Vorhaben geeignet sind, werden in Abschnitt 6.1 dargelegt. Zu beachten ist, dass dabei zunächst zwischen statistischen und nichtstatistischen Stichprobenverfahren unterschieden wird.

Ein statistisches Stichprobenverfahren weist folgende Merkmale auf:

- Jedes Element in der Grundgesamtheit hat eine bekannte und positive Auswahlwahrscheinlichkeit.
- Die Zufälligkeit muss mithilfe eines geeigneten (spezialisierten oder nicht spezialisierten) Zufallsgenerators (z. B. kann man mit MS Excel Zufallszahlen generieren) sichergestellt werden.

Statistische Stichprobenverfahren ermöglichen die Auswahl einer Stichprobe, die die Grundgesamtheit „repräsentiert“ (weshalb die statistische Auswahl so wichtig ist). Ziel ist es, den in einer Stichprobe beobachteten Wert eines Parameters („Variable“) auf die Grundgesamtheit hochzurechnen (zu extrapolieren oder zu schätzen), woraus geschlossen werden kann, ob in Bezug auf eine Grundgesamtheit wesentliche Fehlangaben gemacht wurden oder nicht und, wenn ja, in welchem Ausmaß (fehlerhafter Betrag).

Bei nichtstatistischen Stichprobenverfahren ist die Berechnung der Genauigkeit nicht möglich, folglich gibt es keine Kontrolle des Prüfungsrisikos und es kann nicht gewährleistet werden, dass die Stichprobe die Grundgesamtheit repräsentiert. Der Fehler muss daher empirisch bewertet werden.

Statistische Stichproben sind gemäß Verordnung (EG) Nr. 1083/2006 und Verordnung (EG) Nr. 1198/2006 des Rates sowie Verordnung (EG) Nr. 1828/2006 und Verordnung (EG) Nr. 498/2007 der Kommission für die Durchführung von vertieften Prüfungen (Prüfung von Vorhaben) vorgeschrieben. Eine nichtstatistische Auswahl sollte nur in extremen Fällen angewandt werden, in denen die statistische Auswahl unmöglich ist, z. B. in Verbindung mit sehr kleinen Grundgesamtheiten oder Stichprobenumfängen (siehe Abschnitt 6.2).

## 5.2 Auswahlverfahren

Das Auswahlverfahren kann zu einer der folgenden zwei großen Kategorien gehören:

- statistische Auswahl oder
- nichtstatistische Auswahl.

Die statistische Auswahl umfasst zwei mögliche Methoden:

- Zufallsauswahl
- systematische Auswahl.

Bei der Zufallsauswahl werden für die einzelnen Einheiten der Grundgesamtheit Zufallszahlen generiert, die als Grundlage dienen für die Auswahl für die zur Stichprobe bildenden Einheiten.

Beim systematischen Stichprobenverfahren wird willkürlich ein Anfangspunkt ausgewählt und anschließend eine systematische Regel angewandt, um die zusätzlichen Elemente auszuwählen (z. B. jedes 20. Element nach dem willkürlich gewählten Anfangspunkt).

Die von gleicher Wahrscheinlichkeit ausgehenden Methoden beruhen normalerweise auf der Zufallsauswahl, doch findet beim wertbezogenen Stichprobenverfahren (MUS) eine systematische Auswahl Anwendung.

Für eine nichtstatistische Auswahl stehen (u. a.) folgende Möglichkeiten zur Verfügung:

- ungeordnete Auswahl
- Blockauswahl
- Ermessensauswahl
- risikoorientiertes Stichprobenverfahren unter Einbeziehung von Elementen aus allen drei vorgenannten Auswahlverfahren.

Die ungeordnete Auswahl ist eine „falsche zufällige“ Auswahl in dem Sinne, dass eine Person die zu prüfenden Elemente „zufällig“ auswählt und diese Auswahl ohne messbare Voreingenommenheit vornimmt (z. B. einfach zu analysierende Elemente, leicht zugängliche Elemente, Elemente aus einer gerade „zufällig“ am Bildschirm angezeigten Liste usw.).

Die Blockauswahl ähnelt einem Cluster-Stichprobenverfahren, bei dem das Cluster (eine Gruppe von Elementen der Grundgesamtheit) nicht nach dem Zufallsprinzip ausgewählt wird.

Die Ermessensauswahl beruht ausschließlich auf dem Ermessen des Prüfers, unabhängig vom angewendeten Grundprinzip (z. B. Elemente mit ähnlichen Namen, alle Vorhaben im Zusammenhang mit einem bestimmten Forschungsbereich usw.).

Das risikoorientierte Stichprobenverfahren ist eine nichtstatistische Auswahl von Elementen, die auf verschiedenen willkürlich festgelegten Bestandteilen beruht, die vielfach allen drei nichtstatistischen Auswahlverfahren entnommen werden.

### 5.3 Hochrechnung (Schätzung)

Wie bereits festgestellt, wird mit der Anwendung eines Stichprobenverfahrens das Ziel verfolgt, das in einer Stichprobe beobachtete Fehlerniveau (falsche Darstellung) auf die Grundgesamtheit hochzurechnen (zu extrapolieren oder zu schätzen), woraus geschlossen werden kann, ob in Bezug auf eine Grundgesamtheit wesentliche Fehlangaben gemacht wurden und, wenn ja, in welchem Ausmaß (fehlerhafter Betrag). Das in der Stichprobe festgestellte Fehlerniveau selbst ist daher nicht von Interesse, sondern dient lediglich als Mittel zur Hochrechnung des Fehlers auf die Grundgesamtheit.

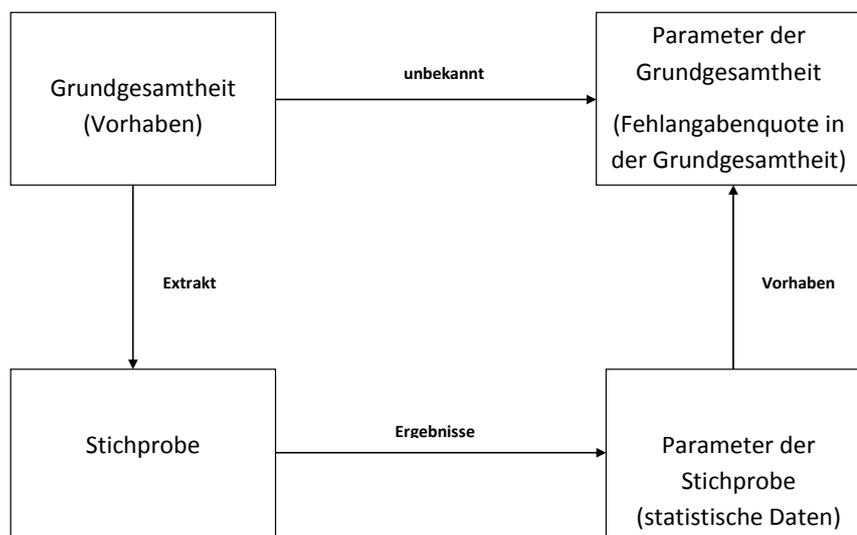


Abb. 3 Stichprobenauswahl und Hochrechnung

Statistische Daten von Stichproben, die zur Hochrechnung des Fehlers auf die Grundgesamtheit genutzt werden, werden als Schätzfunktionen bezeichnet. Der Vorgang der Hochrechnung wird als Schätzung bezeichnet, und der aus der Stichprobe errechnete Wert (prognostizierter Wert) ist der Schätzwert. Zweifelsohne ist dieser nur auf einem Teil der Grundgesamtheit basierende Schätzwert mit einem Fehler behaftet, dem sogenannten Stichprobenfehler.

## 5.4 Genauigkeit (Stichprobenfehler)

Dieser Fehler entsteht dadurch, dass nicht die ganze Grundgesamtheit betrachtet wird. Bei einem Stichprobenverfahren kommt stets ein Schätzungsfehler (Extrapolationsfehler) vor, weil aus Stichprobendaten Rückschlüsse auf die Grundgesamtheit gezogen werden. Der Stichprobenfehler entspricht der Differenz zwischen dem anhand der Stichprobe hochgerechneten Parameter (Schätzwert) und dem tatsächlichen (unbekannten) Parameter der Grundgesamtheit (Wert des Fehlers). Er ist Ausdruck der Unsicherheit, die die Hochrechnung der Ergebnisse auf die Grundgesamtheit mit sich bringt. Ein Maß für diesen Fehler ist normalerweise die **Genauigkeit** oder Präzision der Schätzung. Sie hängt in erster Linie vom **Stichprobenumfang** und der **Streuung der Grundgesamtheit** ab und in geringerem Maße von der **Größe der Grundgesamtheit**.

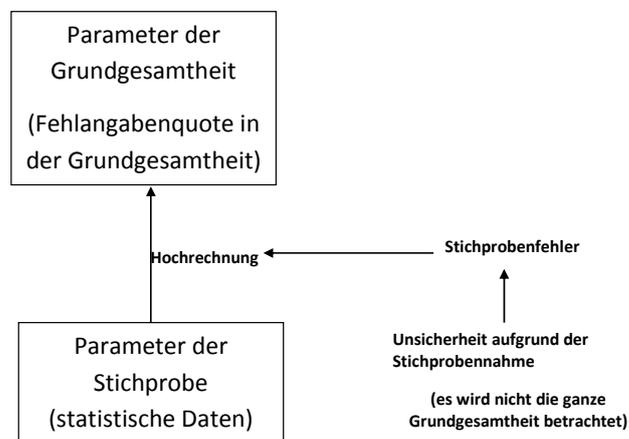


Abb. 4 Stichprobenfehler

Es sollte zwischen geplanter Genauigkeit und effektiver Genauigkeit (SE in den Formeln in Abschnitt 7) unterschieden werden. Während die geplante Genauigkeit der größtmögliche geplante Stichprobenfehler für die Ermittlung des Stichprobenumfangs ist (in der Regel die möglichst auf einen Wert unterhalb der Erheblichkeitsschwelle zu setzende Differenz zwischen dem maximal zulässigen Fehler und dem voraussichtlichen Fehler), entspricht der effektive Stichprobenfehler der Differenz zwischen dem anhand der Stichprobe hochgerechneten Parameter (Schätzwert) und dem tatsächlichen (unbekannten) Parameter der Grundgesamtheit (Wert des Fehlers) und ist Ausdruck der Unsicherheit, die die Hochrechnung der Ergebnisse auf die Grundgesamtheit mit sich bringt.

## 5.5 Grundgesamtheit

Für Stichprobenzwecke umfasst die Grundgesamtheit die gegenüber der Kommission für Vorhaben im Rahmen eines Programms bzw. einer Gruppe von Programmen im Bezugsjahr bescheinigten Ausgaben. Es sollten all jene Vorhaben in die zu prüfende Grundgesamtheit aufgenommen werden, für die geltend gemachte Ausgaben in den

bescheinigten Ausgabenerklärungen enthalten sind, die der Kommission im zu prüfenden Jahr übermittelt wurden, es sei denn, die Grundgesamtheit der Vorhaben ist für die statistische Stichprobenprüfung zu klein (siehe Abschnitt 5.7).

Generell gilt, dass für alle mit der Stichprobe ausgewählten Vorhaben auch alle der Kommission gemeldeten Ausgaben geprüft werden müssen. Wenn jedoch für die ausgewählten Vorhaben sehr viele **Zahlungsanträge** oder Rechnungen vorliegen, können sie anhand von Stichproben geprüft werden, wobei die Zahlungsanträge/Rechnungen nach denselben Prinzipien auszuwählen sind wie die Vorhaben. In diesem Fall müssen innerhalb der einzelnen Vorhaben angemessene Stichprobenumfänge<sup>7</sup> berechnet werden. Bei diesem Ansatz ist das angewandte Stichprobenverfahren im Prüfbericht oder in den Arbeitsunterlagen festzuhalten.

Bei der Aufschlüsselung der Ausgaben, die durch die der Bescheinigungsbehörde vorgelegten Zahlungsanträge bestätigt wurden, können entsprechend den von der Verwaltungsbehörde vorgenommenen Korrekturen Forderungen mit **negativen Beträgen** auftreten. In diesem Fall oder ähnlichen Fällen müssen die negativen Beträge eine separate Grundgesamtheit bilden und separat<sup>8</sup> geprüft werden, da hier zu verifizieren ist, ob der korrigierte Betrag dem entspricht, was Mitgliedstaat oder Kommission beschlossen haben. Gelangt die Prüfbehörde zu dem Schluss, dass der korrigierte Betrag unter dem beschlossenen Betrag liegt, ist dies im jährlichen Kontrollbericht unter Abschnitt „8 – Sonstige Informationen“ zu melden, speziell wenn diese Nichtübereinstimmung auf Mängel in der Fähigkeit des Mitgliedstaats hindeutet, Korrekturen gemäß Artikel 61 Buchstabe f der Verordnung (EG) Nr. 1083/2006 vorzunehmen. Die Berechnung des prognostizierten Fehlers betrifft in diesem Zusammenhang nur die in der Grundgesamtheit der positiven Beträge festgestellten Fehler, wobei es sich um den bei der prognostizierten Gesamtfehlerquote zu berücksichtigenden Buchwert handelt.

Vor der Berechnung der prognostizierten Fehlerquote sollte die Prüfbehörde untersuchen, ob die festgestellten Fehler im Bezugsjahr bereits korrigiert wurden (d. h. wie oben dargestellt, in die Grundgesamtheit der negativen Beträge integriert wurden). Ist dies der Fall, sind sie bei der prognostizierten Fehlerquote nicht zu berücksichtigen.

Die Prüfbehörde kann zur Steigerung der Effizienz der Prüfungen diese auf sonstige verwandte, für die ausgewählten Vorhaben gemeldete Ausgaben außerhalb des Bezugszeitraums ausdehnen. Die Ergebnisse aus der Prüfung zusätzlicher Ausgaben außerhalb des Bezugszeitraums sind bei der Berechnung der prognostizierten Gesamtfehlerquote nicht zu berücksichtigen.

---

<sup>7</sup> Dieser Ansatz entspricht einem zweistufigen Stichprobenplan. Die genaue Bestimmung des Stichprobenumfangs für ein zweistufiges Stichprobenverfahren lässt sich nicht im Rahmen dieser Erläuterungen darlegen. Abgesehen von den Methoden zur Bestimmung von Stichprobenumfängen bei statistischen Stichprobenverfahren gilt die Faustregel, dass niemals Stichproben mit weniger als 30 Beobachtungen verwendet werden sollten.

<sup>8</sup> Natürlich kann die Prüfbehörde auch aus der separaten Grundgesamtheit eine Stichprobe ziehen.

## **5.6 Schichtung**

Eine Schichtung liegt immer dann vor, wenn die Grundgesamtheit in Teilgesamtheiten („Schichten“) unterteilt wird und aus jeder Schicht unabhängige Stichproben gezogen werden.

Mit der Schichtung werden zwei Hauptziele verfolgt: Zum einen ermöglicht sie in der Regel eine höhere Genauigkeit (bei gleichem Stichprobenumfang) oder eine Verringerung des Stichprobenumfangs (bei gleichem Genauigkeitsgrad); zum anderen gewährleistet sie, dass in der Stichprobe alle Teilgesamtheiten vertreten sind, die den einzelnen Schichten entsprechen.

Wird mit unterschiedlichen Fehlerniveaus (Fehlangaben) für verschiedene Gruppen in der Grundgesamtheit gerechnet (z. B. nach Programm, Region, zwischengeschalteter Stelle, Risiko des Vorhabens), sind diese Einteilungskategorien gut für eine Schichtung geeignet. Eine Schichtung nach Ausgabenniveau je Vorhaben ist ebenfalls möglich und gewollt, hauptsächlich in Verbindung mit der Stichprobenziehung mit gleicher Wahrscheinlichkeit.

Für die einzelnen Schichten können unterschiedliche Stichprobenverfahren angewandt werden. Beispielsweise ist es üblich, bei den hochwertigen Elementen eine 100%ige Prüfung vorzunehmen und ein statistisches Stichprobenverfahren für die Prüfung einer Stichprobe aus den verbleibenden Elementen mit niedrigerem Wert einzusetzen, die eine zusätzliche Schicht bzw. zusätzliche Schichten bilden. Das ist dann hilfreich, wenn die Grundgesamtheit einige herausragende Elemente enthält, weil auf diese Weise die Streuung in jeder Schicht verringert wird und damit eine höhere Genauigkeit (oder eine Verringerung des Stichprobenumfangs) möglich ist.

## **5.7 Stichprobeneinheit**

Die zur Prüfung auszuwählende Einheit ist grundsätzlich das Vorhaben. Besteht ein Vorhaben aus mehreren separaten Projekten, können diese für Stichprobenezwecke getrennt ausgewiesen werden. Ist die Grundgesamtheit der Vorhaben zu klein für ein statistisches Stichprobenverfahren (d. h. 50 bis 150 Einheiten), kann der Zahlungsantrag des Begünstigten als Prüfeinheit gewählt werden.

## **5.8 Erheblichkeit**

Gemäß Anhang IV der Verordnungen gilt für Ausgaben, die im Bezugsjahr bei der Kommission geltend gemacht werden, eine Erheblichkeitsschwelle von maximal 2 %. Die Prüfbehörde kann zu Planungszwecken eine Absenkung der Erheblichkeitsschwelle erwägen (zulässiger Fehler). Die Erheblichkeit wird verwendet:

- als Schwellenwert zum Vergleich des prognostizierten Fehlers bei den Ausgaben
- zur Festlegung des zulässigen/vertretbaren Fehlers, der für die Ermittlung des Stichprobenumfangs angewandt wird.

## 5.9 Zulässiger Fehler und geplante Genauigkeit

Der zulässige Fehler ist die höchstzulässige Fehlerquote in der Grundgesamtheit in einem bestimmten Jahr. Bei einer Erheblichkeitsschwelle von 2 % beträgt daher der maximal zulässige Fehler 2 % der Ausgaben, die der Kommission für das Bezugsjahr bescheinigt wurden.

Die geplante Genauigkeit ist der größtmögliche Stichprobenfehler, der zur Hochrechnung der Fehler in einem bestimmten Jahr akzeptiert wird, d. h. die maximale Abweichung zwischen dem wahren Fehler der Grundgesamtheit und der anhand von Stichprobendaten vorgenommenen Hochrechnung. Sie sollte vom Prüfer auf einen Wert unterhalb des zulässigen Fehlers gesetzt werden, da ansonsten die Gefahr hoch ist, dass die Ergebnisse der Stichprobenprüfungen der Vorhaben nicht schlüssig sind und eine ergänzende Stichprobe erforderlich wird.

Beispielsweise beträgt für eine Grundgesamtheit mit einem Gesamtbuchwert von 10 000 000 EUR der entsprechende zulässige Fehler 200 000 EUR (2 % des Gesamtbuchwerts). Beläuft sich der prognostizierte Fehler auf 5000 EUR und setzt der Prüfer die Genauigkeit auf exakt 200 000 EUR (Ursache für diesen Fehler ist die ausschließliche Betrachtung eines kleinen Teils der Grundgesamtheit, d. h. der Stichprobe), liegt die obere Fehlergrenze (Obergrenze des Konfidenzintervalls) bei rund 205 000 EUR. Das ist kein schlüssiges Ergebnis, da einem sehr geringen prognostizierten Fehler eine Obergrenze gegenübersteht, die über der Erheblichkeitsschwelle liegt.

Die Bestimmung der geplanten Genauigkeit erfolgt am besten als Differenz zwischen dem zulässigen Fehler und dem voraussichtlichen Fehler (dem prognostizierten Fehler, den der Prüfer am Ende der Prüfung vorzufinden erwartet). Maßgeblich für den voraussichtlichen Fehler ist natürlich das fachliche Urteil des Prüfers, unterstützt durch die Nachweise aus der Prüftätigkeit in vorangegangenen Jahren in Bezug auf die gleiche oder eine ähnliche Grundgesamtheit oder aus einer Vorab-/Pilotstichprobe.

Zu beachten ist, dass die Wahl eines realistischen voraussichtlichen Fehlers wichtig ist, da der Stichprobenumfang sehr stark von diesem Wert abhängig ist. Siehe auch Abschnitt 8.1.

Abschnitt 7 enthält genaue Formeln zur Verwendung bei der Bestimmung des Stichprobenumfangs.

## 5.10 Streuung

Die Streuung der Grundgesamtheit ist ein Parameter, der den Stichprobenumfang sehr stark beeinflusst. Normalerweise wird die Streuung anhand eines als Standardabweichung bezeichneten Parameters<sup>9</sup> gemessen und mit  $\sigma$  angegeben. Bei einer Grundgesamtheit von 100 Vorhaben beispielsweise, die alle das gleiche Fehlerniveau von 1 000 000 EUR (durchschnittlicher Fehler von  $\mu = 1\,000\,000$  EUR) haben, besteht keine Streuung (die Standardabweichung der Fehler ist Null). Jedoch ist bei einer Grundgesamtheit von 100 Vorhaben, von denen 50 allesamt einen Fehler von 0 EUR und die übrigen 50 jeweils einen Fehler von 2 000 000 EUR aufweisen (der gleiche durchschnittliche Fehler von  $\mu = 1\,000\,000$  EUR) die Standardabweichung der Fehler hoch (1 000 000 EUR).

Für die Prüfung einer Grundgesamtheit mit niedriger Streuung wird ein kleinerer Stichprobenumfang benötigt als bei einer Grundgesamtheit mit hoher Streuung. Im Extremfall des ersten Beispiels (mit einer Varianz von 0) wäre eine Stichprobe bestehend aus einem Vorhaben ausreichend, um den Fehler der Grundgesamtheit exakt hochzurechnen.

Die Standardabweichung ist das gebräuchlichste Maß für die Streuung, da sie leichter verständlich ist als die Varianz. Sie wird in den Einheiten der Variablen angegeben, deren Streuung mit ihr gemessen werden soll. Die Varianz hingegen ist das Quadrat der Einheiten der Variablen, deren Streuung sie misst, und sie ist der einfache Durchschnitt der Quadrate der Abweichung der Variablenwerte vom Mittelwert:

$$s^2 = \frac{1}{\text{Anzahl Einheiten}} \sum_{i=1}^{\text{Anzahl Einheiten}} (V_i - \bar{V})^2$$

wobei  $V_i$  die einzelnen Werte der Variablen  $V$  und  $\bar{V} = \frac{\sum_{i=1}^{\text{Anzahl Einheiten}} V_i}{\text{Anzahl Einheiten}}$  den mittleren Fehler bezeichnen. Die Standardabweichung ist einfach die Quadratwurzel der Varianz:

$$s = \sqrt{s^2}$$

Die Standardabweichung der Fehler bei den am Anfang dieses Abschnitts genannten Beispielen kann wie folgt berechnet werden:

---

<sup>9</sup> Die Standardabweichung ist ein Maß für die Streuung der Grundgesamtheit um ihren Mittelwert. Sie kann anhand von Fehlern oder von Buchwerten berechnet werden. Bei der Berechnung anhand der Grundgesamtheit wird sie normalerweise mit  $\sigma$  angegeben, bei der Berechnung anhand der Stichprobe mit  $s$ . Je größer die Standardabweichung ist, desto heterogener ist die Grundgesamtheit (oder die Stichprobe). Die Varianz ist das Quadrat der Standardabweichung.

a) Fall 1

a.  $N = 100$

b. Alle Vorhaben weisen das gleiche Fehlerniveau von 1 000 000 EUR auf.

c. Mittlerer Fehler

$$\frac{\sum_{i=1}^{100} 1\,000\,000}{1000} = \frac{100 \times 1\,000\,000}{100} = 1\,000\,000$$

d. Standardabweichung der Fehler

$$s = \sqrt{\frac{1}{100} \sum_{i=1}^{100} (1\,000\,000 - 1\,000\,000)^2} = 0$$

b) Fall 2

a.  $N = 100$

b. 50 Vorhaben weisen einen Fehler von 0 EUR und 50 Vorhaben einen Fehler von 2 000 000 EUR auf.

c. Mittlerer Fehler

$$\frac{\sum_{i=1}^{50} 0 + \sum_{i=1}^{50} 2\,000\,000}{1000} = \frac{50 \times 2\,000\,000}{100} = 1\,000\,000$$

d. Standardabweichung der Fehler

$$\begin{aligned} s &= \sqrt{\frac{1}{100} \left( \sum_{i=1}^{50} (0 - 1\,000\,000)^2 + \sum_{i=1}^{50} (2\,000\,000 - 1\,000\,000)^2 \right)} \\ &= \sqrt{\frac{50 \times 1\,000\,000^2 + 50 \times 1\,000\,000^2}{100}} \\ &= \sqrt{1\,000\,000^2} = 1\,000\,000 \end{aligned}$$

## 5.11 Konfidenzintervall und obere Fehlergrenze

Es handelt sich um das Intervall, das mit einer gewissen Wahrscheinlichkeit (als Konfidenzniveau bezeichnet) den wahren (unbekannten) Wert der Grundgesamtheit (Fehler) enthält und allgemein folgendermaßen ausgedrückt wird:

$$[EE - SE; EE + SE],$$

wobei

- EE den prognostizierten oder extrapolierten Fehler darstellt, der dem Begriff der wahrscheinlichsten Fehlerquote (MLE) beim MUS entspricht;
- SE die Genauigkeit (Stichprobenfehler) darstellt.

Der prognostizierte/extrapolierte Fehler (EE) und die obere Fehlergrenze (EE + SE) sind die beiden wichtigsten Faktoren, anhand derer geschlussfolgert werden kann, ob in

Bezug auf eine Grundgesamtheit von Vorhaben wesentliche Fehlangaben gemacht wurden oder nicht. Natürlich kann die ULE nur bei Verwendung eines statistischen Stichprobenverfahrens berechnet werden, bei nichtstatistischen Verfahren ist der EE folglich immer die beste Schätzung des Fehlers in der Grundgesamtheit.

Bei der Verwendung von statistischen Stichprobenverfahren können die folgenden Situationen entstehen:

- Liegt der EE über der Erheblichkeitsschwelle (nachfolgend der Einfachheit halber „2 %“), schlussfolgert die Prüfbehörde, dass ein wesentlicher Fehler vorliegt.
- Sind sowohl der EE als auch die ULE niedriger als 2 %, schlussfolgert die Prüfbehörde, dass die Grundgesamtheit ausgehend von der festgelegten Höhe des Stichprobenrisikos nicht mehr als 2 % Fehlangaben enthält.
- Liegt der EE unter 2 %, die ULE jedoch über 2 %, schlussfolgert die Prüfbehörde, dass zusätzliche Arbeiten erforderlich sind. Dabei kann es sich gemäß der INTOSAI-Leitlinie Nr. 23<sup>10</sup> um Folgendes handeln:
  - *„Aufforderung der geprüften Stelle zur Untersuchung der aufgedeckten Fehler/Ausnahmen und zur Ermittlung des weiteren Fehler-/Ausnahmepotentials. Dies führt möglicherweise dazu, dass die Stelle bereit ist, Berichtigungen an den Rechnungsabschlüssen vorzunehmen.*
  - *Durchführung weiterer Prüfungen, um den Stichprobenfehler und die bei der Bewertung der Ergebnisse aufgrund des wahrscheinlichen Stichprobenfehlers vorhandene Unsicherheit zu verringern.*
  - *Anwendung alternativer Prüfungsverfahren, um zusätzliche Bestätigung zu erhalten.“*

Die Prüfbehörde sollte sich bei der Auswahl einer der oben genannten Optionen auf ihr fachliches Urteil stützen und diese im jährlichen Kontrollbericht vermelden.

Es wird darauf hingewiesen, dass sich eine ULE von deutlich mehr als 2 % in den meisten Fällen vermeiden oder verringern lässt, wenn die Prüfbehörde bei der Berechnung des ursprünglichen Stichprobenumfangs einen realistischen voraussichtlichen Fehler in Betracht zieht (genauere Ausführungen dazu siehe Abschnitt 8.2.2).

Bei der dritten Variante (prognostizierter Fehler unter 2 %, ULE jedoch über 2 %) könnte die Prüfbehörde in einigen Fällen feststellen, dass die Ergebnisse auch bei einem niedrigeren Konfidenzniveau als das ursprünglich geplante immer noch schlüssig sind. Ist dieses Neuberechnete Konfidenzniveau mit einer Bewertung der Qualität der Verwaltungs- und Kontrollsysteme nach wie vor vereinbar, kann sogar ohne weitere

---

<sup>10</sup> Siehe <http://eca.europa.eu/portal/pls/portal/docs/1/133817.PDF>.

Prüfarbeiten mit Sicherheit geschlussfolgert werden, dass die Grundgesamtheit keine wesentlichen Fehlangaben enthält. Daher ist es angeraten, die Neuberechnung des Konfidenzniveaus vorzunehmen und nur in Situationen, in denen die Neuberechnete Konfidenz nicht akzeptabel ist (d. h. mit der Bewertung der Systeme nicht in Einklang steht) die oben vorgeschlagenen zusätzlichen Arbeiten durchzuführen. Erläuterungen zur Neuberechnung der Konfidenzniveaus finden sich in Abschnitt 8.7.

## 5.12 Konfidenzniveau

In der Verordnung sind mit Blick auf die Bestimmung des Stichprobenumfangs für vertiefte Prüfungen Festlegungen zum Konfidenzniveau getroffen worden.

Da der Umfang der Stichprobe vom Konfidenzniveau direkt beeinflusst wird, zielt die Verordnung eindeutig darauf ab, dass für Systeme mit einer feststehenden niedrigen Fehlerquote (und somit hoher Sicherheit) die Möglichkeit gegeben sein soll, das Arbeitspensum zu reduzieren, es für Systeme mit einer potenziell hohen Fehlerquote (und somit niedriger Sicherheit) jedoch möglich sein muss, eine höhere Zahl von Elementen zu prüfen.

Am einfachsten ausgedrückt, bezeichnet das Konfidenzniveau die Wahrscheinlichkeit, dass ein anhand der Stichprobendaten erstelltes Konfidenzintervall den wahren (unbekannten) Fehler der Grundgesamtheit enthält. Beträgt beispielsweise der prognostizierte Fehler in der Grundgesamtheit 6 000 000 EUR und das Intervall zum Konfidenzniveau 90 % ist

[~~5 000 000 €~~; 7 000 000 €]

so heißt das, dass der wahre (jedoch unbekannt) Fehler der Grundgesamtheit mit einer Wahrscheinlichkeit von 90 % zwischen diesen beiden Grenzwerten liegt. Die Auswirkungen dieser strategischen Entscheidungen auf die Planung der Prüfungen und die Stichprobenprüfung der Vorhaben werden in den folgenden Kapiteln erläutert.

## 5.13 Fehlerquote

Die **Stichprobenfehlerquote** wird berechnet als Quotient aus dem Gesamtfehler in der Stichprobe und dem Gesamtbuchwert der Elemente der Stichprobe, die **prognostizierte Fehlerquote** ist der Quotient aus dem **prognostizierten Fehler der Grundgesamtheit** und dem Gesamtbuchwert. Erneut ist zu beachten, dass der Stichprobenfehler als

solcher nicht von Interesse ist, da er lediglich als Instrument zur Berechnung des prognostizierten Fehlers gesehen werden sollte.<sup>11</sup>

---

<sup>11</sup> Bei einigen Stichprobenverfahren, nämlich denen auf der Grundlage der Auswahl mit gleicher Wahrscheinlichkeit, kann die Stichprobenfehlerquote zur Hochrechnung der Fehlerquote der Grundgesamtheit verwendet werden.

## 6 Stichprobenverfahren für die Vorhabensprüfung

### 6.1 Übersicht

Im Rahmen der Vorhabensprüfung dient das Stichprobenverfahren dazu, die zu prüfenden Vorhaben für Einzelfallprüfungen auszuwählen, wobei die Grundgesamtheit die gegenüber der Kommission für Vorhaben im Rahmen eines Programms bzw. einer Gruppe von Programmen bescheinigten Ausgaben im Bezugsjahr umfasst.

Abbildung 5 enthält eine Zusammenfassung der bei Prüfungen am häufigsten verwendeten Stichprobenverfahren.

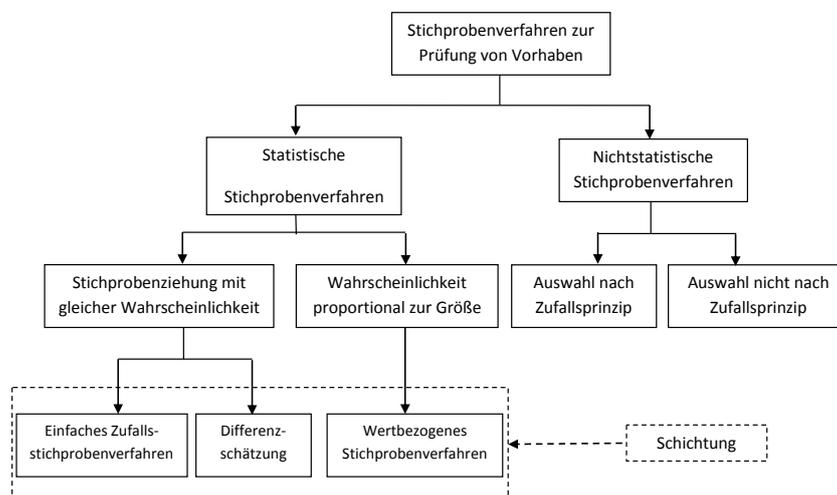


Abb. 5 Stichprobenverfahren zur Vorhabensprüfung

Wie bereits vermerkt, wird bei den Stichprobenverfahren zunächst zwischen statistischen und nichtstatistischen Verfahren unterschieden.

Statistische Stichproben sind gemäß Verordnung (EG) Nr. 1083/2006 und Verordnung (EG) Nr. 1198/2006 des Rates sowie Verordnung (EG) Nr. 1828/2006 und Verordnung (EG) Nr. 498/2007 der Kommission für die Durchführung von vertieften Prüfungen (Vorhabensprüfung) vorgeschrieben. Eine nichtstatistische Auswahl sollte nur in extremen Fällen angewandt werden, in denen die statistische Auswahl unmöglich ist, z. B. in Verbindung mit sehr kleinen Grundgesamtheiten oder Stichprobenumfängen. In Abschnitt 6.2 werden die Bedingungen für die Anwendung der verschiedenen Stichprobenpläne diskutiert und auf die spezifischen Extremsituationen verwiesen, in denen nichtstatistische Stichprobenverfahren zulässig sind.

Bei den statistischen Verfahren unterscheiden sich die Methoden hauptsächlich nach der Auswahlwahrscheinlichkeit: Es gibt Methoden mit gleicher Auswahlwahrscheinlichkeit (einschließlich einfacher Zufallsstichprobenverfahren und Differenzschätzung) und Methoden mit einer Wahrscheinlichkeit proportional zur Größe, von denen insbesondere das bekannte wertbezogene Stichprobenverfahren (MUS) zu nennen ist.

Beim wertbezogenen Stichprobenverfahren (MUS) ist die Wahrscheinlichkeit de facto proportional zur Größe (PPS). Die Vorhaben werden demnach mit Wahrscheinlichkeiten ausgewählt, die zu ihrem Geldwert proportional sind. Je höher der Geldwert, umso höher die Wahrscheinlichkeit der Auswahl. Günstige Bedingungen für die Anwendung der einzelnen spezifischen Methoden werden, wie bereits festgestellt, im folgenden Abschnitt diskutiert.

Ungeachtet des gewählten spezifischen Stichprobenverfahrens sollte die Prüfung der Vorhaben immer nach folgendem grundlegenden Ablauf durchgeführt werden:

1. **Festlegung der Ziele der vertieften Prüfungen:** im Regelfall die Ermittlung der Fehlerquote, die auf einer Hochrechnung anhand einer Stichprobe bei der Kommission vorgelegten bescheinigten Ausgabenaufstellungen für ein bestimmtes Jahr und für ein Programm (oder eine Gruppe von Programmen) basiert
2. **Festlegung der Grundgesamtheit:** die der Kommission vorgelegten bescheinigten Ausgabenaufstellungen für ein bestimmtes Jahr für ein Programm oder für eine Gruppe von Programmen, und die **Stichprobeneinheit**, d. h. das für die Stichprobe zu ziehende Element (normalerweise das Vorhaben, jedoch sind auch andere Möglichkeiten wie beispielsweise Zahlungsanträge denkbar).
3. **Festlegung der Parameter der Grundgesamtheit:** beinhaltet die Festlegung des zulässigen Fehlers (2 % der der Kommission bescheinigten Ausgaben), des voraussichtlichen Fehlers (wie er vom Prüfer erwartet wird), des Konfidenzniveaus (unter Berücksichtigung des Prüfungsrisikomodells) und (normalerweise) des Streuungsmaßes der Grundgesamtheit.
4. **Festlegung des Stichprobenumfangs** entsprechend dem verwendeten Stichprobenverfahren. **Es ist unbedingt zu beachten, dass der endgültige Stichprobenumfang immer auf die nächstliegende ganze Zahl aufgerundet wird.**
5. **Auswahl der Stichprobe und Durchführung der Prüfung.**
6. **Hochrechnung der Ergebnisse, Berechnung der Genauigkeit und Ziehen von Schlussfolgerungen:** beinhaltet die Berechnung der Genauigkeit und des prognostizierten Fehlers und den Vergleich dieser Ergebnisse mit der Erheblichkeitsschwelle.

Diese Grundstruktur wird durch die Wahl eines bestimmten Stichprobenverfahrens, das eine Formel zur Ermittlung des Stichprobenumfangs und einen Rahmen für die Hochrechnung der Ergebnisse vorsieht, verfeinert.

Zu beachten ist auch, dass die spezifischen Formeln zur Ermittlung des Stichprobenumfangs in Abhängigkeit vom gewählten Stichprobenverfahren unterschiedlich sind. Drei Parameter jedoch sind in jedem Falle für den Stichprobenumfang maßgeblich:

- Das Konfidenzniveau (je höher das Konfidenzniveau umso größer der Stichprobenumfang).
- Die Streuung der Grundgesamtheit (d. h. wie variabel sind die Werte der Grundgesamtheit; wenn alle Vorhaben in der Grundgesamtheit ähnliche Fehlerwerte aufweisen, wird von einer geringeren Streuung ausgegangen als bei einer Grundgesamtheit, bei der die Vorhaben höchst unterschiedliche Fehlerwerte haben). Je höher die Streuung der Grundgesamtheit umso größer der Stichprobenumfang.
- Die vom Prüfer festgelegte geplante Genauigkeit; die geplante Genauigkeit ist normalerweise die Differenz zwischen dem zulässigen Fehler von 2 % der Ausgaben und dem voraussichtlichen Fehler. Angenommen, der voraussichtliche Fehler liegt unter 2 %, dann ist der Stichprobenumfang umso größer, je größer der voraussichtliche Fehler (oder je kleiner die geplante Genauigkeit) ist.

Der Stichprobenumfang hängt außerdem von der Größe der Grundgesamtheit ab, jedoch in deutlich geringerem Maße. Bei verhältnismäßig großen Grundgesamtheiten ist deren Größe für den Stichprobenumfang nahezu unerheblich. Das heißt, dass für die Darstellung einer Grundgesamtheit von 10 000 Vorhaben unter normalen Umständen fast der gleiche Stichprobenumfang benötigt wird wie für die Bewertung einer Grundgesamtheit von 20 000 Vorhaben. Zu beachten ist, dass in extremen Fällen, in denen sehr kleine Grundgesamtheiten vorliegen, deren Größe zuweilen nicht ignoriert werden kann.

Spezifische Formeln zur Ermittlung des Stichprobenumfangs werden in Abschnitt 7 angeboten. Als Faustregel jedoch gilt, dass niemals Stichproben mit weniger als 30 Einheiten verwendet werden sollten (damit die für die Konfidenzintervallbestimmung herangezogenen Annahmen zur Verteilung aufrechterhalten werden können).

## **6.2 Bedingungen für die Anwendung von Stichprobendesigns**

Vorab ist in Bezug auf die Entscheidung für eine Methode zur Auswahl der zu prüfenden Vorhaben zu erwähnen, dass – auch wenn diese Entscheidung von zahlreichen Kriterien beeinflusst wird – aus statistischer Sicht die Erwartung hinsichtlich der Fehlerstreuung und ihres Zusammenhangs mit den Ausgaben am wichtigsten ist.

In der nachstehenden Tabelle sind in Abhängigkeit von den Kriterien die geeignetsten Methoden aufgeführt.

<b>Stichprobenverfahren</b>	<b>Günstige Bedingungen</b>
MUS-Standardansatz	Fehler haben eine hohe Streuung <sup>12</sup> und verhalten sich nahezu proportional zur Ausgabenhöhe (d. h. die Streuung der Fehlerquoten ist gering). Die Werte der Ausgaben pro Vorhaben zeigen eine hohe Streuung.
Vorsichtiger MUS-Ansatz	Fehler haben eine hohe Streuung und verhalten sich nahezu proportional zur Ausgabenhöhe. Die Werte der Ausgaben pro Vorhaben zeigen eine hohe Streuung. Es wird von einem niedrigen Fehleranteil ausgegangen <sup>13</sup> .
Differenzschätzung	Fehler sind relativ konstant oder haben eine geringe Streuung. Es muss eine Schätzung der korrigierten Gesamtausgaben in der Grundgesamtheit vorgenommen werden.
Einfaches Zufallsstichprobenverfahren	Allgemeines vorgeschlagenes Verfahren für die Fälle, in denen die zuvor genannten Bedingungen nicht zutreffen. Kann zusammen mit der Schätzung des Mittelwerts pro Einheit oder der Verhältnisschätzung angewendet werden (Abschnitt 7.1.1.3 enthält Anleitungen für die Entscheidung zwischen diesen beiden Schätzmethoden).
Nichtstatistische Verfahren	Wenn die Anwendung des statistischen Verfahrens nicht möglich ist (siehe Diskussion im Folgenden).
Schichtung	Kann in Kombination mit jeder der obengenannten Methoden angewendet werden. Sie ist immer dann besonders hilfreich, wenn bei der Fehlerquote mit sehr großen Unterschieden zwischen Gruppen der Grundgesamtheit (Teilgesamtheiten) gerechnet wird.

Tabelle 3. Günstige Bedingungen für die einzelnen Stichprobenverfahren

Zwar sollten die bislang erteilten Ratschläge befolgt werden, jedoch kann keines der genannten Verfahren als das einzig geeignete oder gar als das „beste Verfahren“

<sup>12</sup> Eine hohe Streuung bedeutet, dass die Fehler bei den Vorhaben nicht vergleichbar sind, d. h. es kommen kleine und große Fehler vor. Das Gegenstück dazu sind Fälle, bei denen alle Fehler mehr oder weniger den gleichen Wert haben (vgl. Abschnitt 5.10).

<sup>13</sup> Da der vorsichtige MUS-Ansatz auf einer Verteilung für seltene Ereignisse basiert, ist er besonders dann geeignet, wenn der Quotient aus der Anzahl der Fehler und der Gesamtanzahl der Vorhaben in der Grundgesamtheit (Fehleranteil) voraussichtlich niedrig ist.

eingestuft werden. Generell sind alle anwendbar. Fällt die Entscheidung dabei auf ein Verfahren, das für die betreffende Sachlage nicht das am besten geeignete ist, so muss der Stichprobenumfang größer sein als bei Verwendung eines besser geeigneten Verfahrens. Es wird jedoch immer möglich sein, durch eines der Verfahren eine repräsentative Stichprobe zu ziehen, sofern von einem angemessenen Stichprobenumfang ausgegangen wird.

Zu beachten ist in diesem Zusammenhang, dass bei allen Stichprobenverfahren die Möglichkeit der Schichtung besteht. Mithilfe der Schichtung lässt sich eine Grundgesamtheit in Gruppen (Schichten) aufteilen, die homogener sind (eine geringere Streuung aufweisen) als die Grundgesamtheit selbst. Statt einer Grundgesamtheit mit hoher Streuung stehen dann zwei oder mehr Teilgesamtheiten mit niedrigerer Streuung zur Verfügung. Auf die Schichtung sollte zurückgegriffen werden, **um die Streuung zu verringern oder Fehler generierende Teilmengen der Grundgesamtheit zu isolieren**. In beiden Fällen verringert sich der Umfang der benötigten Stichprobe.

Wie bereits festgestellt, sollten mithilfe statistischer Stichprobenverfahren Schlussfolgerungen zum Fehlerbetrag in einer Grundgesamtheit gezogen werden. Es gibt jedoch auch spezielle Fälle, in denen ein statistisches Verfahren nicht anwendbar ist.

Nichtstatistische Stichprobenverfahren sollten nur angewendet werden, wenn

- eine extrem kleine Grundgesamtheit vorliegt, deren Umfang die Ziehung einer hinreichend großen Stichprobe nicht zulassen würde (die Grundgesamtheit ist kleiner oder nahezu genauso groß wie der empfohlene Stichprobenumfang)<sup>14</sup>,
- der für ein statistisches Verfahren angeratene Stichprobenumfang aufgrund von unkontrollierbaren Beschränkungen nicht eingehalten werden kann.

Die Prüfbehörde muss nach Möglichkeit eine ausreichend große Grundgesamtheit schaffen, etwa indem Programme, die zu einem gemeinsamen System gehören, zusammengefasst und/oder regelmäßige Auszahlungsanträge der Begünstigten als Einheit gewählt werden. Die Prüfbehörde sollte zudem beachten, dass selbst in einer extremen Situation, wenn zu Beginn des Programmzeitraums der statistische Ansatz nicht möglich ist, dieser angewendet werden sollte, sobald die Voraussetzungen gegeben sind.

### **6.3 Programme im Rahmen der europäischen territorialen Zusammenarbeit (ETZ)**

ETZ-Programme weisen eine Reihe von Besonderheiten auf: Sie können normalerweise nicht zu Gruppen zusammengefasst werden, da jedes Programmsystem unterschiedlich

---

<sup>14</sup> Vgl. Abschnitt 7.4.1.

ist. Die Zahl der Vorhaben ist häufig gering. Normalerweise gibt es für die einzelnen Vorhaben einen federführenden Partner und mehrere weitere Projektpartner.

Bei Programmen mit wenigen Vorhaben sollten die vorstehenden Leitlinien befolgt und die nachstehenden zusätzlichen Verfahren berücksichtigt werden.

Um eine hinreichend große Grundgesamtheit zu erhalten, auf die ein statistisches Stichprobenverfahren angewendet werden kann, ist es erstens gegebenenfalls möglich, die zu Grunde liegenden genehmigten Auszahlungsanträge der einzelnen begünstigten Partner eines Vorhabens als Stichprobeneinheit zu verwenden. In diesem Fall erfolgt die Prüfung bei jedem ausgewählten Begünstigten, nicht aber notwendigerweise bei dem federführenden Partner des Vorhabens.

Kann keine ausreichend große Grundgesamtheit für das statistische Stichprobenverfahren erreicht werden, so sind die in Abschnitt 7.4.1 genannten Optionen 1 oder 2 anzuwenden.

Im Zusammenhang mit den ausgewählten Verfahren sollten bei der Prüfung der federführenden Partner stets deren Ausgaben und das Verfahren zur Aggregation der Auszahlungsanträge der Projektpartner geprüft werden. Ist es aufgrund der Anzahl der Projektpartner nicht möglich, sie alle zu prüfen, so kann eine Zufallsstichprobe ausgewählt werden. Die kombinierte Stichprobe aus federführenden Partnern und Projektpartnern muss hinreichend groß sein, damit die Prüfbehörde hinreichend abgesicherte Schlussfolgerungen ziehen kann.

## 6.4 Bezeichnungen

Bevor die wichtigsten Stichprobenverfahren für die Prüfung von Vorhaben dargelegt werden, sollen zunächst einige damit zusammenhängende Begriffe definiert werden, die bei allen Verfahren anzutreffen sind. Es gilt Folgendes:

- $z$  ist ein Parameter aus der Normalverteilung, der mit dem Konfidenzniveau im Zusammenhang steht, das mithilfe von Systemprüfungen ermittelt wird. Die möglichen Werte von  $z$  werden in der folgenden Tabelle angegeben. Eine vollständige Tabelle mit Werten der Normalverteilung ist im Anhang enthalten.

Konfidenzniveau	60 %	70 %	80 %	90 %	95 %
Systemsicherheitsniveau	Hoch	Moderat	Moderat	Niedrig	Keine Sicherheit
$z$	0,842	1,036	1,282	1,645	1,960

Tabelle 4. Werte von  $z$  nach Konfidenzniveau

- $N$  ist die Größe der Grundgesamtheit (z. B. Anzahl der Vorhaben in einem Programm oder der Zahlungsanträge). Im Falle der Schichtung der Grundgesamtheit wird ein Index  $h$  verwendet, um die jeweilige Schicht anzugeben,  $N_{h,t} = 1, 2, \dots, H$ , wobei  $H$  die Anzahl der Schichten ist.
- $n$  ist der Stichprobenumfang. Im Falle der Schichtung der Grundgesamtheit wird ein Index  $h$  verwendet, um die jeweilige Schicht anzugeben,  $n_{h,t} = 1, 2, \dots, H$ , wobei  $H$  die Anzahl der Schichten ist.
- $TE$  ist der nach der Verordnung maximal zulässige Fehler, d. h. 2 % der gegenüber der Kommission bescheinigten Gesamtausgaben (Buchwert,  $BV$ ).
- $BV_{i,t} = 1, 2, \dots, N$  ist der Buchwert (der gegenüber der Kommission bescheinigten Ausgaben) für ein Element (Vorhaben/Zahlungsantrag).
- $CBV_{i,t} = 1, 2, \dots, N$  ist der korrigierte Buchwert, d. h. die nach Durchführung der Prüfverfahren für ein Element (Vorhaben/Zahlungsantrag) bestimmten Ausgaben.
- $E_i = BV_i - CBV_{i,t} = 1, 2, \dots, N$  ist der Fehlerbetrag eines Elements und definiert als die Differenz zwischen dem Buchwert des  $i$ -ten Elements in der Stichprobe und dem entsprechenden korrigierten Buchwert. Im Falle der Schichtung der Grundgesamtheit wird ein Index  $h$  verwendet, um die jeweilige Schicht anzugeben,  $E_{ht} = BV_{ht} - CBV_{ht,t} = 1, 2, \dots, N_{h,t}, h = 1, 2, \dots, H$ , wobei  $H$  die Anzahl der Schichten ist..
- $AE$  ist der voraussichtliche Fehler, der auf der Grundlage des erwarteten Fehlerniveaus auf der Ebene der Vorhaben vom Prüfer festgelegt wird (z. B. eine voraussichtliche Fehlerquote multipliziert mit den Gesamtausgaben auf der Ebene der Grundgesamtheit).  $AE$  kann anhand von historischen Daten (prognostizierter Fehler im vergangenen Zeitraum) oder einer Vorabstichprobe/Pilotstichprobe mit geringer Stichprobengröße (die auch zur Bestimmung der Standardabweichung verwendet wird) ermittelt werden.

## 7 Stichprobenverfahren

### 7.1 Einfache Zufallsstichprobenverfahren

#### 7.1.1 Standardansatz

##### 7.1.1.1 Einleitung

Beim einfachen Zufallsstichprobenverfahren handelt es sich um ein statistisches Stichprobenverfahren. Unter den Auswahlverfahren mit gleicher Wahrscheinlichkeit ist es das bekannteste, es dient der Hochrechnung der in der Stichprobe ermittelten Fehlerquote auf die gesamte Grundgesamtheit.

Die zu untersuchende statistische Einheit ist das Vorhaben (oder der Auszahlungsantrag). Die in der Stichprobe enthaltenen Einheiten werden nach dem Zufallsprinzip mit gleichen Wahrscheinlichkeiten ausgewählt. Das einfache Zufallsstichprobenverfahren ist ein generisches Verfahren, das zu jeder Art von Grundgesamtheit passt, obwohl dafür – da keine Hilfsinformationen verwendet werden – in der Regel ein größerer Stichprobenumfang benötigt wird als beim MUS (wenn das Ausgabenniveau zwischen den Vorhaben sehr unterschiedlich ist und ein positiver Zusammenhang zwischen Ausgaben und Fehlern vorliegt). Die Hochrechnung der Fehler kann sich auf zwei Teilverfahren gründen: Schätzung des Mittelwerts pro Einheit oder Verhältnisschätzung (siehe Abschnitt 7.1.1.3).

Wie alle anderen Verfahren lässt sich dieses Verfahren mit einer Schichtung kombinieren (günstige Bedingungen für die Schichtung werden in Abschnitt 6.2 erörtert und konkrete Formeln in Abschnitt 7.1.2 vorgestellt).

##### 7.1.1.2 Stichprobenumfang

Die Berechnung des Stichprobenumfangs  $n$  innerhalb der Rahmenbedingungen des einfachen Zufallsstichprobenverfahrens beruht auf den folgenden Informationen:

- dem Umfang der Grundgesamtheit  $N$
- dem Konfidenzniveau, das mithilfe von Systemprüfungen und dem damit verwandten Koeffizienten  $z$  aus einer Normalverteilung ermittelt wird (siehe Abschnitt 6.4)
- dem maximal zulässigen Fehler  $TE$  (in der Regel 2 % der gesamten Ausgaben)
- dem voraussichtlichen Fehler  $AE$ , der vom Prüfer anhand seines fachlichen Urteils und vorhandener Informationen ausgewählt wird
- der Standardabweichung  $\sigma_e$  der Fehler.

Der Stichprobenumfang wird wie folgt berechnet<sup>15</sup>:

$$n = \left( \frac{N \times z \times \sigma_e}{TE - AE} \right)^2$$

wobei  $\sigma_e$  die Standardabweichung der Fehler in der Grundgesamtheit ist. Zu beachten ist, dass die Standardabweichung der Fehler für die gesamte Grundgesamtheit in der vorstehenden Berechnung als bekannt angenommen wird. In der Praxis wird dies fast nie der Fall sein, und die Mitgliedstaaten müssen sich entweder auf historische Daten (Standardabweichung der Fehler für die Grundgesamtheit im vergangenen Zeitraum) bzw. auf eine Vorabstichprobe/Pilotstichprobe mit geringem Stichprobenumfang (für den nicht weniger als 20 bis 30 Einheiten empfohlen werden) stützen. Im letztgenannten Fall wird eine Stichprobe des Umfangs  $n^p$  gezogen, und ein Vorabschätzwert der Fehlervarianz (Quadrat der Standardabweichung) wird wie folgt errechnet:

$$\sigma_e^2 = \frac{1}{n^p - 1} \sum_{i=1}^{n^p} (E_i - \bar{E})^2.$$

Dabei sind  $E_i$  die einzelnen Fehler der Einheiten in der Stichprobe und  $\bar{E} = \frac{\sum_{i=1}^{n^p} E_i}{n^p}$  der mittlere Fehler der Stichprobe.

Es sei darauf hingewiesen, dass die Pilotstichprobe im weiteren Verlauf als Teil der zur Prüfung ausgewählten Stichprobe verwendet werden kann.

### 7.1.1.3 Prognostizierter Fehler

Für die Hochrechnung des Stichprobenfehlers auf die Grundgesamtheit gibt es zwei mögliche Ansätze. Der erste beruht auf der Schätzung des Mittelwerts pro Einheit (absolute Fehler) und der zweite auf der Verhältnisschätzung (Fehlerquoten).

---

<sup>15</sup> Bei einer kleinen Größe der Grundgesamtheit, d. h. wenn der endgültige Stichprobenumfang einen großen Teil der Grundgesamtheit repräsentiert (als Faustregel mehr als 10 % der Grundgesamtheit), kann eine genauere Formel verwendet werden:  $n = \left( \frac{N \times z \times \sigma_e}{TE - AE} \right)^2 / \left( 1 + \left( \frac{N \times z \times \sigma_e}{TE - AE} \right)^2 \right)$ . Diese Korrektur gilt für das einfache Zufallsstichprobenverfahren und für die Differenzschätzung. Sie lässt sich auch in zwei Schritten durchführen, indem der Stichprobenumfang  $n$  mit der üblichen Formel berechnet und unter Verwendung von  $n' = \frac{n \times N}{n + N - 1}$  sequentiell berichtigt wird.

### Schätzung des Mittelwerts pro Einheit (absolute Fehler)

Der aus der Stichprobe aufgedeckte durchschnittliche Fehler pro Vorhaben wird mit der Zahl der Vorhaben in der Grundgesamtheit multipliziert, so dass sich der prognostizierte Fehler ergibt:

$$EE_1 = N \times \frac{\sum_{i=1}^n E_i}{n}$$

### Verhältnisschätzung (Fehlerquoten)

Die aus der Stichprobe aufgedeckte durchschnittliche Fehlerquote wird mit dem Buchwert auf der Ebene der Grundgesamtheit multipliziert:

$$EE_2 = BV \times \frac{\sum_{i=1}^n E_i}{\sum_{i=1}^n BV_i}$$

Die Stichprobenfehlerquote in der obigen Formel errechnet sich durch einfache Division des Gesamtfehlerbetrags in der Stichprobe durch den Gesamtbetrag der Ausgaben der Einheiten in der Stichprobe (geprüfte Ausgaben).

Es lässt sich nicht von vornherein abschätzen, welches Extrapolationsverfahren das Beste ist, da die jeweiligen Vorteile von der Höhe des Zusammenhangs zwischen Fehlern und Ausgaben abhängen. Als Faustregel sollte gelten, dass das zweite Verfahren zur Anwendung kommt, wenn ein enger Zusammenhang zwischen Fehlern und Ausgaben (höhere Werteelemente weisen eher höhere Fehler auf) erwartet wird, und das erste Verfahren (Mittelwert pro Einheit), wenn die Erwartung besteht, dass die Fehler relativ unabhängig von der Ausgabenhöhe sind (höhere Fehler finden sich entweder in Einheiten eines höheren oder niedrigen Ausgabenniveaus)<sup>16</sup>. Diese Bewertung lässt sich anhand von Stichprobendaten vornehmen, da die Entscheidung über das Extrapolationsverfahren nach der Auswahl und Prüfung der Stichprobe getroffen werden kann.

#### 7.1.1.4 Genauigkeit

Genauigkeit (Stichprobenfehler) ist das Maß der mit der Hochrechnung (Extrapolation) verbundenen Unsicherheit. Sie errechnet sich je nach dem gewählten Extrapolationsverfahren auf unterschiedliche Weise.

---

<sup>16</sup> Exakte Formeln zur Bestimmung des besten Verfahrens lassen sich im Rahmen dieser Erläuterungen nicht darlegen. Formell ist das zweite Verfahren besser geeignet, wenn  $\frac{COV_{E,BV}}{VAR_{BV}} > R/2$  ist, wobei  $COV_{E,BV}$  die Kovarianz zwischen Fehlern und Ausgaben,  $VAR_{BV}$  die Varianz der Ausgaben und R die Fehlerquote in der Grundgesamtheit ist.

### Schätzung des Mittelwerts pro Einheit (absolute Fehler)

Für die Genauigkeit gilt die Formel

$$SE_1 = N \times z \times \frac{s_e}{\sqrt{n}}$$

wobei  $s_e$  die Standardabweichung der Fehler in der Stichprobe ist (nunmehr aus der Stichprobe berechnet, die auch zur Hochrechnung der Fehler auf die Grundgesamtheit verwendet wurde)

$$s_e^2 = \frac{1}{n-1} \sum_{i=1}^n (E_i - E)^2$$

### Verhältnisschätzung (Fehlerquoten)

Für die Genauigkeit gilt die Formel

$$SE_2 = N \times z \times \frac{s_q}{\sqrt{n}}$$

wobei  $s_q$  die Stichprobenabweichung der Variablen  $q$  ist:

$$q_i = E_i - \frac{\sum_{i=1}^n E_i}{\sum_{i=1}^n BV_i} \times BV_i$$

Diese Variable berechnet sich für jede Einheit in der Stichprobe als Differenz zwischen ihrem Fehler und dem Produkt aus ihrem Buchwert und der Fehlerquote in der Stichprobe.

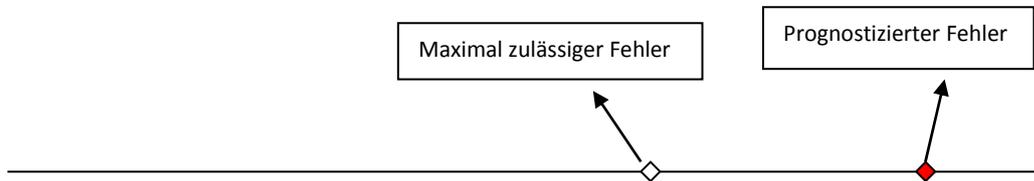
#### 7.1.1.5 Bewertung

Um auf die Erheblichkeit der Fehler zu schließen, ist zunächst die obere Fehlergrenze (ULE) zu berechnen. Diese Obergrenze entspricht der Summe aus dem prognostizierten Fehler  $EE$  selbst und der Genauigkeit der Extrapolation

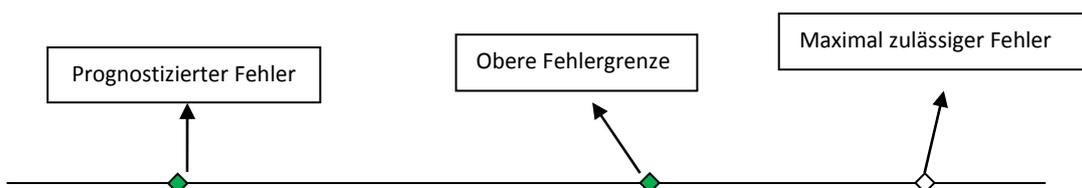
$$ULE = EE + SE$$

Anschließend sollten der prognostizierte Fehler und die Obergrenze jeweils mit dem maximal zulässigen Fehler verglichen werden, um Prüfungsschlussfolgerungen zu ziehen:

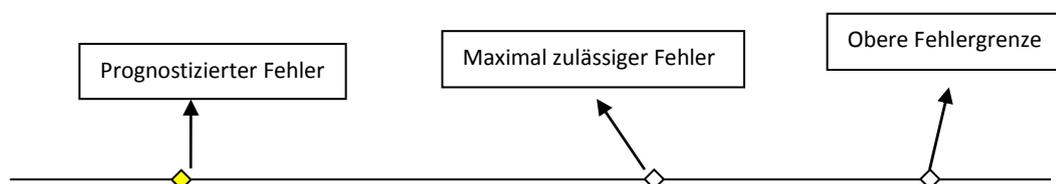
- Liegt der prognostizierte Fehler über dem maximal zulässigen Fehler, würde der Prüfer schlussfolgern, dass hinreichende Beweise dafür vorliegen, dass die Fehler in der Grundgesamtheit über der Erheblichkeitsschwelle liegen:



- Liegt die obere Fehlergrenze unter dem maximal zulässigen Fehler, sollte der Prüfer schlussfolgern, dass die Fehler in der Grundgesamtheit unter der Erheblichkeitsschwelle liegen.



- Liegt der prognostizierte Fehler unter dem maximal zulässigen Fehler, die obere Fehlergrenze jedoch darüber, sind weitere Schritte erforderlich, da nicht hinreichend belegt ist, dass die Grundgesamtheit keine wesentlichen Fehlangaben enthält. Die erforderlichen Schritte werden in Abschnitt 5.11 erörtert.



#### 7.1.1.6 Beispiel

Angenommen wird eine Grundgesamtheit von Ausgaben, die gegenüber der Kommission in einem bestimmten Jahr für Vorhaben in einem Programm oder einer Gruppe von Programmen bescheinigt wurden. Die von der Prüfbehörde durchgeführten

Systemprüfungen ergaben ein moderates Sicherheitsniveau. Daher ist für die Prüfung der Vorhaben ein Konfidenzniveau von 80 % angemessen.

Größe der Grundgesamtheit (Anzahl der Vorhaben)	3852
Buchwert (Summe der Ausgaben im Bezugsjahr)	3 199 654 543 EUR

Eine Vorabstichprobe von 20 Vorhaben ergab einen Vorabschätzwert für die Standardabweichung der Fehler von 62 194 EUR (in MS Excel berechnet als „=STDEV(D2:D21)“):

	A	B	C	D
1	<b>Operation</b>	<b>Book Value (BV)</b>	<b>Audited Value (AV)</b>	<b>Error</b>
2	2,299	1,368,071 €	1,368,071 €	- €
3	2,579	2,324,672 €	2,267,903 €	56,768 €
4	3,736	4,521,748 €	4,245,900 €	275,848 €
5	4,068	11,905,284 €	11,905,284 €	- €
6	2,020	1,237,076 €	1,237,076 €	- €
7	13	850,647 €	850,647 €	- €
8	2,761	987,452 €	987,452 €	- €
9	3,384	6,863,730 €	6,863,730 €	- €
10	4,266	737,302 €	737,302 €	- €
11	1,273	1,748,231 €	1,748,231 €	- €
12	1,506	1,426,326 €	1,426,326 €	- €
13	1,882	885,025 €	885,025 €	- €
14	143	721,564 €	705,720 €	15,845 €
15	2,246	3,398,994 €	3,398,994 €	- €
16	2,841	1,913,322 €	1,913,322 €	- €
17	1,757	845,404 €	845,404 €	- €
18	4,998	3,701,924 €	3,701,924 €	- €
19	4,828	1,633,711 €	1,633,711 €	- €
20	646	1,981,079 €	1,981,079 €	- €
21	2,635 €	975,033 €	975,033 €	- €
22	<b>Total</b>	50,026,595 €	49,678,134 €	348,461 €
23	<b>Sample error rate:=D22/B22</b>			0.7%
24	<b>Sample standard deviation of errors:=STDEV(D2:D21)-----&gt;</b>			62,194 €

Der erste Schritt besteht in der Berechnung des erforderlichen Stichprobenumfangs unter Verwendung der folgenden Formel:

$$n = \left( \frac{N \times z \times \sigma_e}{TE - AE} \right)^2$$

wobei sich  $z$  auf 1,282 beläuft (Koeffizient, der einem Konfidenzniveau von 80 % entspricht),  $\sigma_e$  einen Wert von 62 194 EUR hat und  $TE$ , der zulässige Fehler, 2 % (in der Verordnung festgelegter Höchstwert für die Erheblichkeitsschwelle) des Buchwerts, d. h. 2 % x 3 199 654 543 EUR = 63 993 091 EUR, beträgt. Diese Vorabstichprobe ergibt eine Stichprobenfehlerquote von 0,7 %. Ausgehend von den Erfahrungen des Vorjahres und der Schlussfolgerung im Bericht zu den Verwaltungs- und Kontrollsystemen rechnet die Prüfbehörde außerdem mit einer Fehlerquote von höchstens 0,7 %. Damit beträgt  $AE$ , der voraussichtliche Fehler, 0,6 % der Gesamtausgaben, d. h. 0,7 % x 3 199 654 543 EUR = 22 397 582 EUR.

$$n = \left( \frac{3,852 \times 1,282 \times 62,194}{63,993,091 - 22,397,582} \right)^2 \approx 55$$

Der Mindestumfang der Stichprobe liegt daher bei 55 Vorhaben.

Die vorherige Vorabstichprobe von 20 Vorhaben wird im Rahmen der Hauptstichprobe verwendet. Daher muss der Prüfer nur 35 weitere Vorhaben nach dem Zufallsprinzip auswählen. Die folgende Tabelle weist die Ergebnisse der gesamten Stichprobe von 55 Vorhaben aus:

	A	B	C	D	E	F
1	Operation ID (1)	Book Value (BV) (2)	Audited Value (AV) (3)	Error (4)	Error rate (4)/(2)	q (4)-SUM(4)/SUM(2)*2
2	2,299	1,368,071 €	1,368,071 €	- €	0.0%	7,016 €
3	2,579	2,324,672 €	2,267,903 €	56,768 €	2.4%	44,847 €
4	3,736	4,521,748 €	4,245,900 €	275,848 €	0.7%	252,660 €
5	4,068	11,905,284 €	11,905,284 €	- €	0.0%	61,053 €
6	2,020	1,237,076 €	1,237,076 €	- €	0.0%	6,344 €
7	13	850,647 €	850,647 €	- €	0.0%	4,362 €
8	2,761	987,452 €	987,452 €	- €	0.0%	5,064 €
9	3,384	6,863,730 €	6,863,730 €	- €	0.0%	35,199 €
10	4,266	737,302 €	737,302 €	- €	0.0%	3,781 €
11	1,273	1,748,231 €	1,748,231 €	- €	0.0%	8,965 €
12	1,506	1,426,326 €	1,426,326 €	- €	0.0%	7,315 €
13	1,882	885,025 €	885,025 €	- €	0.0%	4,539 €
14	143	721,564 €	705,720 €	15,845 €	2.2%	12,144 €
15	2,246	3,398,994 €	3,398,994 €	- €	0.0%	17,431 €
16	2,841	1,913,322 €	1,913,322 €	- €	0.0%	9,812 €
17	1,757	845,404 €	845,404 €	- €	0.0%	4,335 €
18	4,998	3,701,924 €	3,701,924 €	- €	2.5%	18,984 €
19	4,828	1,633,711 €	1,633,711 €	- €	0.0%	8,378 €
20	646	1,981,079 €	1,981,079 €	- €	0.0%	10,160 €
21	2,635	975,033 €	975,033 €	- €		
22	(...)	(...)	(...)	(...)	(...)	(...)
56	3,873	470,763 €	470,763 €	- €	0.0%	2,354 €
57	<b>Total</b>	<b>52,056,328 €</b>	<b>51,509,453 €</b>	<b>546,875 €</b>		
58	<b>Sample standard deviation:=STDEV(D2:D56)--&gt;</b>			56,539 €		58,412 €

Der Gesamtbuchwert der 55 Vorhaben in der Stichprobe beträgt 52 056 328 EUR (in MS Excel berechnet als „=SUM(B2:B56)“). Der Gesamtfehlerbetrag in der Stichprobe beläuft sich auf 546 875 EUR (in MS Excel berechnet als „=SUM(D2:D56)“). Dieser Betrag – dividiert durch den Umfang der Stichprobe – entspricht dem durchschnittlichen Fehler eines Vorhabens innerhalb der Stichprobe.

Bei Verwendung einer Schätzung des Mittelwerts pro Einheit erfolgt die Hochrechnung des Fehlers auf die Grundgesamtheit, indem dieser durchschnittliche Fehler mit der Größe der Grundgesamtheit (3852 in diesem Beispiel) multipliziert wird. Diese Zahl ist der prognostizierte Fehler auf der Ebene des Programms:

$$EE_1 = N \times \frac{\sum_{i=1}^{55} E_i}{n} = 3,852 \times \frac{546,875}{55} = 38,301,136.$$

Bei Verwendung einer Verhältnisschätzung kann die Hochrechnung der Fehler auf die Grundgesamtheit durch Multiplikation der in der Stichprobe festgestellten durchschnittlichen Fehlerquote mit dem Buchwert auf der Ebene der Grundgesamtheit erfolgen:

$$EE_2 = BV \times \frac{\sum_{i=1}^{55} E_i}{\sum_{i=1}^{55} BV_i} = 3,199,654,543 \times \frac{546,875}{52,056,328} = 33,613,802$$

Die Stichprobenfehlerquote in dieser Formel errechnet sich durch einfache Division des Gesamtfehlerbetrags in der Stichprobe durch den Gesamtbetrag der Ausgaben der Vorhaben in der Stichprobe.

Die prognostizierte Fehlerquote ist der Quotient aus dem prognostizierten Fehler und dem Buchwert der Grundgesamtheit (Gesamtausgaben). Wird eine Schätzung des Mittelwerts pro Einheit vorgenommen, so beträgt der prognostizierte Fehler:

$$r_1 = \frac{38,301,136}{3,199,654,543} = 1.20 \%$$

und bei der Verhältnisschätzung ist:

$$r_2 = \frac{33,613,802}{3,199,654,543} = 1.05 \%$$

In beiden Fällen liegt der prognostizierte Fehler unterhalb der Erheblichkeitsschwelle. Endgültige Schlussfolgerungen können allerdings erst nach Berücksichtigung des Stichprobenfehlers (Genauigkeit) gezogen werden.

Der erste Schritt zur Ermittlung der Genauigkeit ist die Berechnung der Standardabweichung der Fehler in der Stichprobe (in MS Excel berechnet als „=STDEV(D2:D56)“):

$$s_e = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (E_i - E)^2} = \sqrt{\frac{1}{54} \sum_{i=1}^{55} (E_i - E)^2} = 56,539.$$

Die Genauigkeit der Schätzung des Mittelwerts pro Einheit ergibt sich wie folgt:

$$SE_1 = N \times z \times \frac{s_e}{\sqrt{n}} = 3,852 \times 1,282 \times \frac{56,539}{\sqrt{55}} = 37,647,928.$$

Für die Verhältnisschätzung muss folgende Variable erstellt werden:

$$q_i = E_i - \frac{\sum_{t=1}^{55} E_t}{\sum_{t=1}^{55} BV_t} \times BV_i.$$

Diese Variable befindet sich in der letzten Spalte der Tabelle (Spalte F). Beispielsweise ergibt sich der Wert in Zelle F2 aus dem Wert des Fehlers des ersten Vorhabens (0 EUR) minus der Summe der Stichprobenfehler in Spalte D, 546 875 EUR („=SUM(D2:D56)“) dividiert durch die Summe der Buchwerte in Spalte B, 52 056 328 EUR („=SUM(B2:B56)“) und multipliziert mit dem Buchwert des Vorhabens (768 071 EUR):

$$q_1 = 0 - \frac{546,875}{52,056,328} \times 1,368,071 = -7,016.$$

Ausgehend von der Standardabweichung dieser Variablen,  $s_q = 58,412$  (in MS Excel berechnet als „=STDEV(F2:F56)“) ergibt sich die Genauigkeit für die Verhältnisschätzung anhand der folgenden Formel:

$$SE_2 = N \times z \times \frac{s_q}{\sqrt{n}} = 3,852 \times 1,282 \times \frac{58,412}{\sqrt{55}} = 38,895,113$$

Um auf die Erheblichkeit der Fehler zu schließen, ist zunächst die obere Fehlergrenze (ULE) zu berechnen. Diese Obergrenze entspricht der Summe aus dem prognostizierten Fehler  $EE$  selbst und der Genauigkeit der Hochrechnung

$$ULE = EE + SE$$

Anschließend sollten der prognostizierte Fehler und die Obergrenze jeweils mit dem maximal zulässigen Fehler verglichen werden, um Prüfungsschlussfolgerungen zu ziehen:

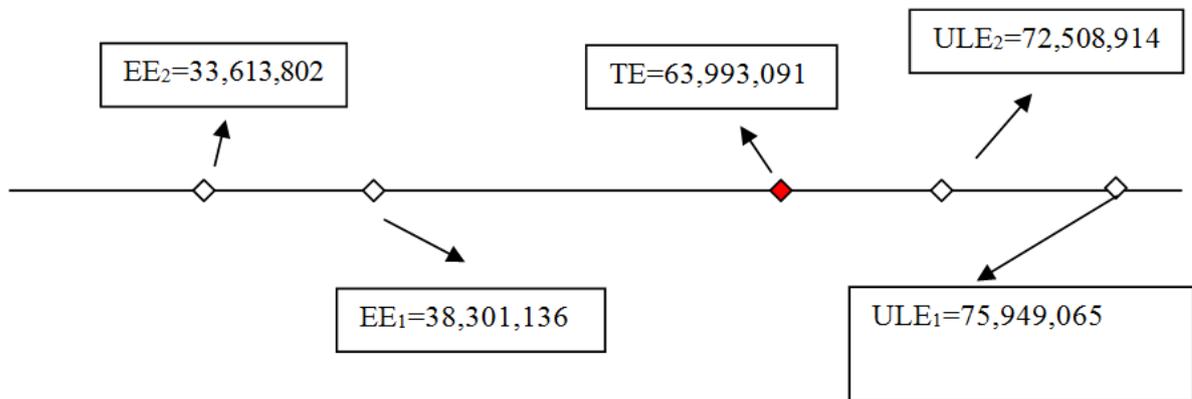
$$ULE_1 = EE_1 + SE_1 = 38,301,136 + 37,647,928 = 75,949,065$$

oder

$$ULE_2 = EE_2 + SE_2 = 33,613,802 + 38,895,113 = 72,508,914$$

Beim Vergleich mit der Erheblichkeitsschwelle von 2 % des gesamten Buchwerts des Programms (2 % x 3 199 654 543 EUR = 63 993 091 EUR) mit dem prognostizierten

Fehler und der oberen Fehlergrenze lautet die Schlussfolgerung bei Anwendung beider Ansätze (Schätzung des Mittelwerts pro Einheit und Verhältnisschätzung), dass der prognostizierte Fehler unter dem maximal zulässigen Fehler, die obere Fehlergrenze aber darüber liegt. Der Prüfer kann daraus schließen, dass weitere Schritte erforderlich sind, da nicht hinreichend belegt ist, dass die Grundgesamtheit keine wesentlichen Fehlangaben enthält. Die erforderlichen Schritte werden in Abschnitt 5.11 erörtert.



## 7.1.2 Das geschichtete einfache Stichprobenverfahren

### 7.1.2.1 Einleitung

Beim geschichteten einfachen Stichprobenverfahren wird die Grundgesamtheit in Teilgesamtheiten („Schichten“) unterteilt, und es werden unter Anwendung des üblichen Zufallsstichprobenverfahrens unabhängige Stichproben gezogen.

Bei der Auswahl der Schichtungskriterien ist zu berücksichtigen, dass das Ziel in der Bildung von Gruppen (Schichten) besteht, deren Streuung geringer ist als die der Grundgesamtheit. Beim einfachen Zufallsstichprobenverfahren stellt die Schichtung nach Ausgabenniveau je Vorhaben in der Regel immer dann einen guten Ansatz dar, wenn ein Zusammenhang zwischen Fehlerquote und Ausgabenniveau zu erwarten ist. Andere Variablen, von denen Aufklärung über das Fehlerniveau in den Vorhaben erwartet werden kann, sind somit ebenfalls gut für die Schichtung geeignet. In Frage kommen beispielsweise Programme, Regionen, zwischengeschaltete Stellen und Klassen, die auf dem Risiko des Vorhabens beruhen.

Wird eine Schichtung nach Ausgabenniveau vorgenommen, so sollte eine hochwertige Schicht<sup>17</sup> ermittelt, eine 100%ige Prüfung dieser Elemente vorgenommen und das

<sup>17</sup> Eine generelle Regel für die Bestimmung des Schwellenwertes, ab dem es sich um eine hochwertige Schicht handelt, besteht nicht. Als Faustregel könnte gelten, dass alle Vorhaben, deren Ausgaben bei der Multiplikation mit den Gesamtausgaben der Grundgesamtheit oberhalb der Erheblichkeitsschwelle liegen

einfache Zufallsstichprobenverfahren zur Prüfung der Stichproben aus den verbleibenden Elementen mit niedrigerem Wert eingesetzt werden, die eine zusätzliche Schicht bzw. zusätzliche Schichten bilden. Das ist dann hilfreich, wenn die Grundgesamtheit herausragende Elemente enthält. Die zu der 100 %-Schicht gehörenden Elemente sind in diesem Fall aus der Grundgesamtheit herauszunehmen, und alle in den übrigen Abschnitten betrachteten Schritte beziehen sich nur auf die Grundgesamtheit der niedrigwertigen Elemente. Es sei darauf hingewiesen, dass eine 100%ige Prüfung der Einheiten der hochwertigen Schicht nicht obligatorisch ist. Die Prüfbehörde kann eine Strategie auf der Grundlage mehrerer Schichten, die unterschiedlichen Ausgabenhöhen entsprechen, entwickeln und alle Schichten durch Stichprobennahme prüfen lassen.

### 7.1.2.2 Stichprobenumfang

Der Stichprobenumfang wird wie folgt berechnet:

$$n = \left( \frac{N \times E \times \sigma_w}{TS - AE} \right)^2$$

Darin bezeichnet  $\sigma_w^2$  den gewichteten Mittelwert der Fehlervarianzen für die gesamte Schichtenmenge:

$$\sigma_w^2 = \sum_{h=1}^H \frac{N_h}{N} \sigma_{eh}^2, \quad h = 1, 2, \dots, H,$$

$\sigma_{eh}^2$  ist die Fehlervarianz in jeder Schicht. Die Varianz der Fehler wird für jede Schicht als unabhängige Grundgesamtheit berechnet als

$$\sigma_{eh}^2 = \frac{1}{n_h^p - 1} \sum_{i=1}^{n_h^p} (E_{hi} - \bar{E}_h)^2, \quad h = 1, 2, \dots, H.$$

Dabei sind  $E_{hi}$  die einzelnen Fehler der Einheiten in der Stichprobe von Schicht  $h$  und  $\bar{E}_h$  der mittlere Fehler der Stichprobe in der Schicht  $h$ .

Wie bereits im Zusammenhang mit dem Standardansatz für das einfache Zufallsstichprobenverfahren dargelegt wurde, können diese Werte auf vorherigen Erkenntnissen oder auf einer Vorabstichprobe/Pilotstichprobe mit geringem Stichprobenumfang beruhen. Im letztgenannten Fall kann die Pilotstichprobe in der Folge wie üblich als Teil der zur Prüfung ausgewählten Stichprobe verwendet werden.

---

(2 %) liegen, zu dieser Kategorie gehören. Bei vorsichtiger angelegten Ansätzen wird durch Division der Erheblichkeit durch 2 oder 3 ein kleinerer Schwellenwert verwendet, doch hängt der Schwellenwert von den Merkmalen der Grundgesamtheit ab und sollten dem fachlichen Urteil unterliegen.

Stehen am Anfang eines Programmplanungszeitraums keine historischen Informationen zur Verfügung und ist der Zugriff auf eine Pilotstichprobe nicht möglich, so kann der Stichprobenumfang (für das erste Jahr des Zeitraums) mit dem Standardansatz berechnet werden. Die in der Prüfungsstichprobe dieses ersten Jahres erfassten Daten können zur Präzisierung der Stichprobenumfangberechnung in den Folgejahren herangezogen werden. Der Preis für das Fehlen dieser Informationen besteht darin, dass der Stichprobenumfang im ersten Jahr wahrscheinlich größer sein wird, als er es sein müsste, wenn Hilfsinformationen über Schichten zur Verfügung stünden.

Nach Berechnung des Gesamtstichprobenumfangs  $n$  ist die nach Schicht vorgenommene Aufteilung der Stichprobe wie folgt:

$$n_h = \frac{N_h}{N} \times n.$$

Dabei handelt es sich um eine allgemeine Aufteilungsmethode, die auch als proportionale Aufteilung bezeichnet wird. Mit einer spezifischeren Aufteilung lassen sich in einigen Fällen ein zusätzlicher Genauigkeitsgewinn bzw. eine Verringerung des Stichprobenumfangs erreichen. Entscheidungen hinsichtlich der Eignung anderer Aufteilungsmethoden für einzelne spezifische Grundgesamtheiten erfordern ein gewisses Maß an stichprobentheoretischen Kenntnissen.

### 7.1.2.3 Prognostizierter Fehler

Anhand von  $H$  nach dem Zufallsprinzip ausgewählten Stichproben von Vorhaben, deren Umfang jeweils nach der oben genannten Formel berechnet wurde, lässt sich der prognostizierte Fehler auf der Ebene der Grundgesamtheit anhand der beiden üblichen Methoden berechnen: Schätzung des Mittelwerts pro Einheit und Verhältnisschätzung.

#### Schätzung des Mittelwerts pro Einheit

In jeder Gruppe der Grundgesamtheit (Schicht) wird der in der Stichprobe festgestellte durchschnittliche Fehler je Vorhaben mit der Anzahl der Vorhaben in der Schicht ( $N_h$ ) multipliziert, und alle für die einzelnen Schichten erzielten Ergebnisse werden addiert, woraus sich der prognostizierte Fehler ergibt:

$$EE_1 = \sum_{h=1}^H N_h \times \frac{\sum_{t=1}^{m_h} E_t}{n_h}.$$

#### Verhältnisschätzung

In jeder Gruppe der Grundgesamtheit (Schicht) wird die in der Stichprobe festgestellte durchschnittliche Fehlerquote mit dem Buchwert der Grundgesamtheit auf der Schichtebene ( $BV_h$ ) multipliziert:

$$EE_2 = \sum_{h=1}^H BV_h \times \frac{\sum_{t=1}^{n_h} E_t}{\sum_{t=1}^{n_h} BV_t}$$

Die Stichprobenfehlerquote in den einzelnen Schichten errechnet sich durch einfache Division des Gesamtfehlerbetrags in der Stichprobe der Schicht durch den Gesamtbetrag der Ausgaben in der gleichen Stichprobe.

Bei der Entscheidung für eine der beiden Methoden sollten die für den Standardansatz des einfachen Zufallsstichprobenverfahrens angestellten Überlegungen als Grundlage dienen.

Wurde eine 100-%-Schicht betrachtet und zuvor der Grundgesamtheit entnommen, ist der in dieser umfassenden Schicht festgestellte Gesamtfehlerbetrag dem oben genannten Schätzwert ( $EE_1$  oder  $EE_2$ ) hinzuzuaddieren, um die endgültige Hochrechnung für den Fehlerbetrag in der gesamten Grundgesamtheit zu erhalten.

#### 7.1.2.4 Genauigkeit

Wie beim Standardverfahren ist die Genauigkeit (Stichprobenfehler) das Maß der mit der Hochrechnung (Extrapolation) verbundenen Unsicherheit. Sie errechnet sich je nach dem gewählten Extrapolationsverfahren auf unterschiedliche Weise.

#### Schätzung des Mittelwerts pro Einheit (absolute Fehler)

Für die Genauigkeit gilt die Formel

$$SE_1 = N \times z \times \frac{s_w^2}{\sqrt{n}}$$

wobei  $s_w^2$  der gewichtete Mittelwert der Varianz der Fehler für die gesamte Schichtenmenge ist (nunmehr aus der Stichprobe berechnet, die auch zur Hochrechnung der Fehler auf die Grundgesamtheit verwendet wurde):

$$s_w^2 = \sum_{h=1}^H \frac{N_h}{N} s_{e_h}^2, h = 1, 2, \dots, H;$$

und  $s_{e_h}^2$  ist die geschätzte Varianz der Fehler für die Stichprobe von Schicht h

$$s_{eh}^2 = \frac{1}{n_h - 1} \sum_{i=1}^{n_h} (E_{ht} - \bar{E}_h)^2, h = 1, 2, \dots, H$$

## Verhältnisschätzung (Fehlerquoten)

Für die Genauigkeit gilt die Formel

$$SE_x = N \times x \times \frac{S_{qW}}{\sqrt{n}}$$

wobei

$$S_{qW}^2 = \sum_{h=1}^H \frac{N_h}{N} S_{qh}^2$$

ein gewichtetes Mittel der Stichprobenvarianzen der Variablen  $q_h$  ist und

$$q_{ih} = E_{ih} - \frac{\sum_{i=1}^{N_h} E_{ih}}{\sum_{i=1}^{N_h} BV_{ih}} \times BV_{ih}$$

Diese Variable berechnet sich für jede Einheit in der Stichprobe als Differenz zwischen ihrem Fehler und dem Produkt aus ihrem Buchwert und der Fehlerquote in der Stichprobe.

### 7.1.2.5 Bewertung

Um auf die Erheblichkeit der Fehler zu schließen, ist zunächst die obere Fehlergrenze (ULE) zu berechnen. Diese Obergrenze entspricht der Summe aus dem prognostizierten Fehler **EE** selbst und der Genauigkeit der Extrapolation

$$ULE = EE + SE$$

Anschließend sollten der prognostizierte Fehler und die Obergrenze jeweils mit dem maximal zulässigen Fehler verglichen werden, um Prüfungsschlussfolgerungen unter Verwendung des Ansatzes zu ziehen, der in Abschnitt 7.1.1.5 dargelegt wurde.

### 7.1.2.6 Beispiel

Es wird ausgegangen von einer Grundgesamtheit von Ausgaben, die gegenüber der Kommission in einem bestimmten Jahr für Vorhaben in einer Gruppe von Programmen bescheinigt wurden. Das Verwaltungs- und Kontrollsystem wird für die gesamte Gruppe von Programmen verwendet, und die von der Prüfbehörde durchgeführten Systemprüfungen ergaben ein moderates Sicherheitsniveau. Daher beschloss die

Prüfbehörde, Vorhabensprüfungen mit einem Konfidenzniveau von 80 % durchzuführen.

Die Prüfbehörde hat Grund zu der Annahme, dass bei hochwertigen Vorhaben unabhängig vom jeweiligen Programm erhebliche Fehlerrisiken bestehen. Auch gibt es Anhaltspunkte dafür, dass in den Programmen mit unterschiedlichen Fehlerquoten zu rechnen ist. Unter Berücksichtigung all dieser Aspekte entscheidet sich die Prüfbehörde für eine Schichtung der Grundgesamtheit nach Programm und nach Ausgaben (wobei in einer Stichprobenschicht mit 100 %-Prüfung alle Vorhaben mit einem über der Erheblichkeitsschwelle liegenden Buchwert isoliert werden).

In der nachstehenden Tabelle sind die vorhandenen Informationen zusammengefasst:

Größe der Grundgesamtheit (Anzahl der Vorhaben)	3852
Größe der Grundgesamtheit – Schicht 1 (Anzahl der Vorhaben in Programm 1)	2520
Größe der Grundgesamtheit – Schicht 2 (Anzahl der Vorhaben in Programm 2)	1327
Größe der Grundgesamtheit – Schicht 3 (Anzahl der Vorhaben mit BV > Erheblichkeitsschwelle)	5
Buchwert (Summe der Ausgaben im Bezugsjahr)	4 199 882 024 EUR
Buchwert – Schicht 1 (Gesamtausgaben in Programm 1)	2 168 367 291 EUR
Buchwert – Schicht 2 (Gesamtausgaben in Programm 2)	1 447 155 510 EUR
Buchwert – Schicht 3 (Gesamtausgaben der Vorhaben mit BV > Erheblichkeitsschwelle)	584 359 223 EUR

Die Stichprobenschicht mit 100%iger Prüfung, die die fünf hochwertigen Vorhaben enthält, sollte – wie in Abschnitt 7.1.2.1 dargelegt – separat behandelt werden. Daher entspricht nachfolgend der Wert von  $N$  der Gesamtzahl der Vorhaben in der Grundgesamtheit abzüglich der Zahl der in der Stichprobenschicht mit 100 %-Prüfung enthaltenen Vorhaben, d. h. 3847 (= 3852 – 5) Vorhaben.

Der erste Schritt besteht in der Berechnung des erforderlichen Stichprobenumfangs unter Verwendung der folgenden Formel:

$$n = \left( \frac{N \times z \times \sigma_w}{TE \cdot AE} \right)^2$$

wobei  $z$  den Wert 1,282 hat (Koeffizient, der einem Konfidenzniveau von 80 % entspricht)) und  $TE$ , der zulässige Fehler, 2 % (in der Verordnung festgelegter Höchstwert für die Erheblichkeitsschwelle) des Buchwerts beträgt, d. h. 2 % x 4 199 882 024 EUR = 83 997 640 EUR. Ausgehend von den Erfahrungen des Vorjahres und der Schlussfolgerung im Bericht zu den Verwaltungs- und Kontrollsystemen

rechnet die Prüfbehörde außerdem mit einer Fehlerquote von höchstens 0,4 %. Damit beträgt  $AE$ , der voraussichtliche Fehler, 0,4 % der Gesamtausgaben, d. h. 0,4 % x 4 199 882 024 EUR = 16 799 528 EUR.

Da es sich bei der dritten Schicht um eine komplett zu prüfende Stichprobenschicht handelt, hat sie einen fixen Stichprobenumfang, der der Größe der Grundgesamtheit (mit den fünf hochwertigen Vorhaben) entspricht. Die Berechnung des Stichprobenumfangs für die übrigen zwei Schichten erfolgt anhand der nachstehenden Formel, wobei  $\sigma_w^2$  den gewichteten Durchschnitt der Varianzen der Fehler für die beiden übrigen Schichten bezeichnet:

$$\sigma_w^2 = \sum_{i=1}^2 \frac{N_i}{N} \sigma_{eh}^2, h = 1, 2;$$

$\sigma_{eh}^2$  ist die Fehlervarianz in jeder Schicht. Die Fehlervarianz wird für jede Schicht als unabhängige Grundgesamtheit berechnet als

$$\sigma_{eh}^2 = \frac{1}{n_h^p - 1} \sum_{i=1}^{n_h^p} (E_{hi} - E_h)^2, h = 1, 2, \dots, H.$$

Dabei sind  $E_{hi}$  die einzelnen Fehler der Einheiten in der Stichprobe von Schicht  $h$  und  $E_h$  der mittlere Fehler der Stichprobe in Schicht  $h$ .

Eine Vorabstichprobe von 20 Vorhaben der Schicht 1 ergab einen Schätzwert für die Standardabweichung der Fehler von 5370 EUR:

	A	B	C	D
1	<b>Operation ID</b>	<b>Book Value (BV)</b>	<b>Audited Value (AV)</b>	<b>Error</b>
2	143	721,564 €	705,822 €	15,743 €
3	2,246	398,994 €	398,994 €	- €
4	2,841	113,322 €	113,322 €	- €
5	1,757	445,404 €	445,404 €	- €
6	4,998	701,924 €	684,715 €	17,209 €
7	4,828	233,711 €	233,711 €	- €
8	646	81,079 €	81,079 €	- €
9	2,635	575,033 €	575,033 €	10,000 €
10	3,873	470,763 €	470,763 €	- €
11	834	381,364 €	381,364 €	- €
12	4,738	208,882 €	208,882 €	- €
13	1,695	590,045 €	590,045 €	- €
14	2,589	422,135 €	422,135 €	- €
15	2,917	171,645 €	171,645 €	- €
16	1,434	474,949 €	474,949 €	- €
17	4,077	86,929 €	86,303 €	1,200 €
18	953	293,675 €	293,675 €	- €
19	232	116,948 €	116,948 €	- €
20	1,097	275,418 €	275,418 €	- €
21	166	313,678 €	313,678 €	- €
22	<b>Total</b>	<b>7,077,466 €</b>	<b>7,043,888 €</b>	<b>44,151 €</b>
23	<b>Sample error rate:=D22/B22</b>			<b>0.6%</b>
24	<b>Sample standard deviation of errors:=STDEV(D2:D21)</b>			<b>5,370 €</b>

Das gleiche Verfahren wurde für die Grundgesamtheit von Schicht 2 angewandt.

Eine Vorabstichprobe von 20 Vorhaben der Schicht 2 ergab einen Schätzwert für die Standardabweichung der Fehler von 177 582 EUR:

Schicht 1 – Vorabschätzwert für die Standardabweichung der Fehler	5370 EUR
Schicht 2 – Vorabschätzwert für die Standardabweichung der Fehler	177 582 EUR

Daher beträgt der gewichtete Durchschnitt der Varianzen der Fehler dieser beiden Schichten

$$\sigma_w^2 = \frac{2520}{3,847} 5,370^2 + \frac{1,327}{3,847} 177,582^2 = 10,896,828,862$$

Der Stichprobenumfang errechnet sich wie folgt:

$$n = \left( \frac{3,847 \times 1,282 \times \sqrt{10,896,828,862}}{83,997,640 - 16,799,528} \right)^2 \approx 59$$

Der Gesamtstichprobenumfang ergibt sich aus diesen 59 Vorhaben plus den 5 Vorhaben der Stichprobenschicht mit 100 %-Prüfung, also 64 Vorhaben.

Die Aufteilung der Stichprobe nach Schicht ist wie folgt:

$$n_1 = \frac{N_1}{N_1 + N_2} \times n = \frac{2,520}{3,847} \times 59 \approx 39,$$

$$n_2 = n - n_1 = 20$$

und

$$n_3 = N_3 = 5$$

Die Prüfung von 39 Vorhaben in Schicht 1, von 20 Vorhaben in Schicht 2 und von fünf Vorhaben in Schicht 3 liefert dem Prüfer einen Gesamtfehler für die Vorhaben der Stichprobe. Die vorherigen Vorabstichproben von 20 Vorhaben in Schicht 1 und 2 werden als Teil der Hauptstichprobe verwendet. Somit braucht der Prüfer in Schicht 1 nur 19 weitere Vorhaben nach dem Zufallsprinzip auszuwählen. Die folgende Tabelle weist die Stichprobenergebnisse für die 64 geprüften Vorhaben aus:

<b>Stichprobenergebnisse – Schicht 1</b>		
A	Stichprobenbuchwert	11 966 658 EUR
B	Stichprobengesamtfehler	190 866 EUR
C	Durchschnittlicher Stichprobenfehler (C = B ÷ 39)	4 894 EUR
D	Standardabweichung der Fehler in der Stichprobe	4 329 EUR
<b>Stichprobenergebnisse – Schicht 2</b>		
E	Stichprobenbuchwert	572 607 646 EUR
F	Stichprobengesamtfehler	400 825 EUR
G	Durchschnittlicher Stichprobenfehler (G = F ÷ 20)	20 041 EUR
H	Standardabweichung der Fehler in der Stichprobe	177 582 EUR
<b>Stichprobenergebnisse – Schicht 3</b>		
I	Stichprobenbuchwert	584 359 223 EUR
J	Stichprobengesamtfehler	7 240 855 EUR
K	Durchschnittlicher Stichprobenfehler (K = J ÷ 5)	1 448 171 EUR

Die folgende Abbildung veranschaulicht die Ergebnisse für Schicht 1:

	A	B	C	D	E	F
	Operation ID (1)	Book Value (BV) (2)	Audited Value (AV) (3)	Error (4)	Error rate (4)/(2)	q (4)-SUM(4)/SUM(2)*2
1						
2	143	721,564 €	705,822 €	15,743 €	2.2%	4,224 €
3	2,246	398,994 €	398,994 €	- €	0.0%	6,369 €
4	2,841	113,322 €	113,322 €	- €	0.0%	1,809 €
5	1,757	445,404 €	445,404 €	- €	0.0%	7,110 €
6	4,998	701,924 €	684,715 €	17,209 €	2.5%	6,004 €
7	4,828	233,711 €	233,711 €	- €	0.0%	3,731 €
8	646	81,079 €	81,079 €	- €	0.0%	1,294 €
9	2,635	575,033 €	575,033 €	10,000 €	1.7%	821 €
10	3,873	470,763 €	470,763 €	- €	0.0%	7,515 €
11	834	381,364 €	381,364 €	- €	0.0%	6,088 €
12	4,738	208,882 €	208,882 €	- €	0.0%	3,334 €
13	1,695	590,045 €	590,045 €	- €	0.0%	9,419 €
14	2,589	422,135 €	422,135 €	- €	0.0%	6,739 €
15	2,917	171,645 €	171,645 €	- €	0.0%	2,740 €
16	1,434	474,949 €	474,949 €	- €	0.0%	7,582 €
17	4,077	86,929 €	86,303 €	1,200 €	1.4%	188 €
18	953	293,675 €	293,675 €	- €	0.0%	4,688 €
19	232	116,948 €	116,948 €	- €	0.0%	1,867 €
20	1,097	275,418 €	275,418 €	- €	0.0%	4,397 €
21	166	313,678 €	313,678 €	- €	0.0%	5,007 €
22	(...)	(...)	(...)	(...)	(...)	(...)
39	1699	393,458 €	392,487 €	972 €	0.2%	5,309 €
40	2392	294,702 €	290,764 €	3,938 €	1.3%	767 €
41	<b>Total</b>	11,956,658 €	11,765,792 €	190,866 €		
42	<b>Sample standard deviation:=STDEV()</b>			4,329 €		76,286 €

Bei der Schätzung des Mittelwerts pro Einheit erfolgt die Extrapolation des Fehlers für die beiden Stichprobenschichten durch Multiplikation des durchschnittlichen Stichprobenfehlers mit der Größe der Grundgesamtheit. Die Summe dieser beiden Zahlen ist zum Fehler, der in der Stichprobenschicht mit 100 %-Prüfung festgestellt wurde, hinzuaddieren, um den Fehler auf die Grundgesamtheit hochzurechnen:

$$EE_1 = \sum_{h=1}^3 N_h \times \frac{\sum_{i=1}^{n_h} E_i}{n_h} = 2,520 \times 4,894 + 1,327 \times 20,041 + 7,240,855 = 46,168,474$$

Ein alternatives Schätzergebnis unter Anwendung der Verhältnisschätzung ergibt sich, wenn die in der Stichprobe der Schicht festgestellte durchschnittliche Fehlerquote mit dem Buchwert auf der Schichtebene (für die beiden Stichprobenschichten) multipliziert wird. Anschließend ist die Summe dieser beiden Zahlen zum Fehler, der in der Stichprobenschicht mit 100 %-Prüfung festgestellt wurde, hinzuaddieren, um den Fehler auf die Grundgesamtheit hochzurechnen:

$$\begin{aligned}
EE_2 &= \sum_{h=1}^3 BV_h \times \frac{\sum_{t=1}^{n_h} E_t}{\sum_{t=1}^{n_h} BV_t} \\
&= 2,168,267,291 \times \frac{190,866}{11,966,958} + 1,447,155,510 \times \frac{400,825}{572,607,646} \\
&\quad + 7,240,055 = 42,030,924.
\end{aligned}$$

Die prognostizierte Fehlerquote ist der Quotient aus dem prognostizierten Fehler und dem Buchwert der Grundgesamtheit (Gesamtausgaben). Wird eine Schätzung des Mittelwerts pro Einheit vorgenommen, so beträgt der prognostizierte Fehler:

$$r_1 = \frac{46,168,142}{4,199,882,024} = 1.10 \%$$

und bei der Verhältnisschätzung ist:

$$r_2 = \frac{42,838,056}{4,199,002,024} = 1.02 \%$$

In beiden Fällen liegt der prognostizierte Fehler unterhalb der Erheblichkeitsschwelle. Endgültige Schlussfolgerungen können allerdings erst nach Berücksichtigung des Stichprobenfehlers (Genauigkeit) gezogen werden. Es ist zu beachten, dass die Schichten 1 und 2 die einzigen Stichprobenfehlerquellen sind, da die hochwertige Schicht einer 100%igen Stichprobenprüfung unterzogen wird. Im Folgenden werden nur die beiden Stichprobenschichten betrachtet.

Ausgehend von den Standardabweichungen der Fehler in den Stichproben der beiden Schichten (Tabelle mit Stichprobenergebnissen) beträgt der gewichtete Durchschnitt der Varianz der Fehler für alle Stichprobenschichten:

$$s_w^2 = \sum_{h=1}^2 \frac{N_h}{N} s_{eh}^2 = \frac{2520}{3,847} \times 4,329^2 + \frac{1,327}{3,847} \times 177,582^2 = 10,890,214,986.$$

Daher errechnet sich die Genauigkeit des absoluten Fehlers mit der folgenden Formel:

$$SE_1 = N \times z \times \frac{s_w}{\sqrt{n}} = 3,847 \times 1,282 \times \frac{\sqrt{10,890,214,986}}{\sqrt{59}} = 67,004,263.$$

Für die Verhältnisschätzung muss folgende Variable erstellt werden:

$$q_{th} = E_{th} - \frac{\sum_{t=1}^{n_h} E_{th}}{\sum_{t=1}^{n_h} BV_{th}} \times BV_{th}.$$

Schicht 1 ist in der letzten Spalte der vorigen Tabelle (Spalte F) dargestellt. Beispielsweise ergibt sich der Wert in Zelle F2 aus dem Wert des Fehlers des ersten Vorhabens (15 743 EUR) minus der Summe der Stichprobenfehler in Spalte E, 190 866 EUR („=SUM(D2:D40)“), dividiert durch die Summe der Buchwerte in Spalte B, 11 956 658 EUR („=SUM(B2:B40)“), multipliziert mit dem Buchwert des Vorhabens (721 564 EUR):

$$q_{11} = 0 - \frac{190,866}{11,956,658} \times 721,564 = 4,224.$$

Die Standardabweichung dieser Variablen für Schicht 1 ist  $s_{q1} = 76,286$  (berechnet in MS Excel als „=STDEV(F2:F40)“). Unter Anwendung der gerade beschriebenen Methodik ist die Standardabweichung für Schicht 2  $s_{q2} = 92,907$ . Daher gilt für die gewichtete Summe der Varianzen von  $q_{ih}$ :

$$s_{q_{gw}}^2 = \sum_{h=1}^3 \frac{N_h}{N} s_{q_h}^2 = \frac{2,520}{3,847} \times 76,286^2 + \frac{1,327}{3,847} \times 92,907^2 = 6,789,595,201.$$

Die Genauigkeit für die Verhältnisschätzung ergibt sich aus

$$SE_2 = N \times z \times \frac{s_{q_{gw}}}{\sqrt{n}} = 3,847 \times 1,282 \times \frac{\sqrt{6,789,595,201}}{\sqrt{59}} = 52,906,131.$$

Um auf die Erheblichkeit der Fehler zu schließen, ist zunächst die obere Fehlergrenze (ULE) zu berechnen. Diese Obergrenze entspricht der Summe aus dem prognostizierten Fehler  $EE$  selbst und der Genauigkeit der Extrapolation

$$ULE = EE + SE$$

Anschließend sollten der prognostizierte Fehler und die Obergrenze jeweils mit dem maximal zulässigen Fehler verglichen werden, um Prüfungsschlussfolgerungen zu ziehen:

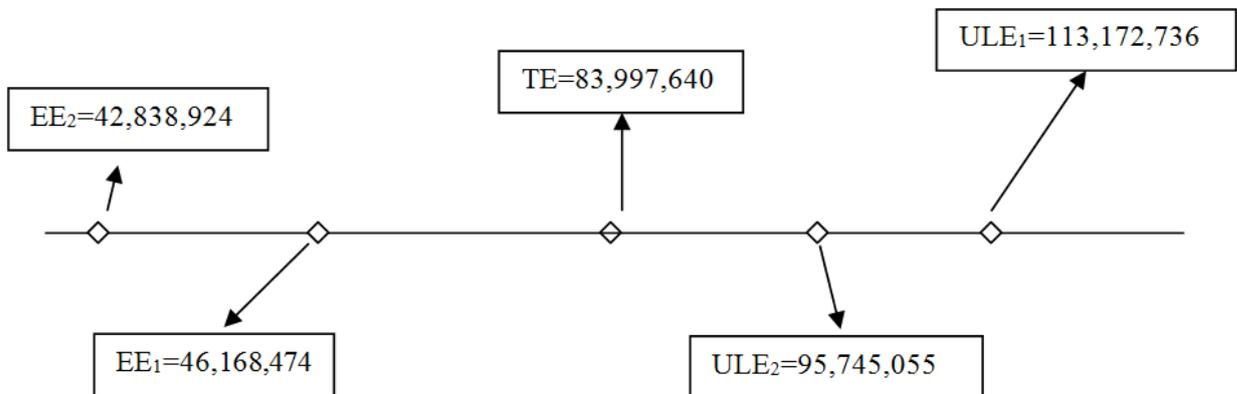
$$ULE_1 = EE_1 + SE_1 = 46,168,474 + 67,004,263 = 113,172,736$$

oder

$$ULE_2 = EE_2 + SE_2 = 42,838,924 + 52,906,131 = 95,745,055$$

Beim Vergleich der prognostizierten Ergebnisse mit der Erheblichkeitsschwelle von 2 % des gesamten Buchwerts der Grundgesamtheit (2 % x 4 199 882 023 EUR = 83 997 640 EUR) wird festgestellt, dass der maximal zulässige Fehler größer ist als die prognostizierten Fehler (bei Verwendung beider Methoden), jedoch unterhalb der Obergrenze liegt. Daher sind zusätzliche Arbeitsschritte (wie in Abschnitt 5.11

beschrieben) erforderlich, denn es ist nicht hinreichend belegt, dass die Grundgesamtheit keine wesentlichen Fehlangaben enthält.



### 7.1.3 Einfaches Zufallsstichprobenverfahren – zwei Zeiträume

#### 7.1.3.1 Einleitung

Die Prüfbehörde kann das Stichprobenverfahren in mehreren Zeiträumen während des Jahres (typischerweise zwei Halbjahre) durchführen. Der wichtigste Vorteil dieses Ansatzes steht nicht im Zusammenhang mit einer Reduzierung des Stichprobenumfangs, sondern mit der Möglichkeit, den Prüfungsaufwand über das Jahr zu verteilen und somit den Aufwand zu verringern, der bei nur einer Beobachtung am Ende des Jahres entstünde.

Mit diesem Ansatz wird die Jahresgrundgesamtheit in zwei Teilgesamtheiten aufgeteilt, die jeweils den Vorhaben und Ausgaben eines Halbjahres entsprechen. Für jedes Halbjahr werden unabhängige Stichproben unter Anwendung des üblichen Zufallsstichprobenverfahrens gezogen.

#### 7.1.3.2 Stichprobenumfang

##### Erstes Halbjahr

Im ersten Prüfzeitraum (z. B. Halbjahr) wird der Gesamtstichprobenumfang (für beide Halbjahre) wie folgt errechnet:

$$n = \left( \frac{N \times z \times \sigma_w}{TE - AE} \right)^2$$

wobei  $\sigma_w^2$  der gewichtete Mittelwert der Fehlervarianzen in jedem Halbjahr ist:

$$\sigma_w^2 = \frac{N_1}{N} \sigma_{e1}^2 + \frac{N_2}{N} \sigma_{e2}^2$$

und  $\sigma_{e_t}^2$  die Fehlervarianz in jedem Zeitraum t (Halbjahr) bezeichnet. Die Fehlervarianz der beiden Halbjahre wird jeweils als eigenständige Grundgesamtheit berechnet

$$\sigma_{e_t}^2 = \frac{1}{n_t^p - 1} \sum_{i=1}^{n_t^p} (E_{it} - \bar{E}_t)^2, t = 1, 2.$$

Dabei sind  $E_{it}$  die einzelnen Fehler der Einheiten in der Stichprobe des Halbjahres t und  $\bar{E}_t$  der mittlere Fehler der Stichprobe im Halbjahr t.

Hierbei ist zu beachten, dass die Festlegung der Werte für die erwarteten Varianzen bei beiden Halbjahreswerten nach fachlichem Urteil und unter Verwendung von historischem Wissen erfolgen muss. Die Option einer Vorabstichprobe/Pilotstichprobe mit geringem Stichprobenumfang, wie sie bereits im Zusammenhang mit dem Standardansatz für das einfache Zufallsstichprobenverfahren dargelegt wurde, ist nach wie vor gegeben, lässt sich jedoch nur für das erste Halbjahr realisieren. Die Ausgaben für das zweite Halbjahr sind zum Zeitpunkt der ersten Beobachtung noch nicht getätigt, und es sind keine objektiven Daten (außer den vorherigen) verfügbar. Falls Pilotstichproben gezogen werden, können sie danach wie üblich als Teil der zur Prüfung ausgewählten Stichprobe verwendet werden.

Stehen zur Einschätzung der Datenstreuung im zweiten Halbjahr keine vorherigen Daten oder Kenntnisse zur Verfügung, so kann ein vereinfachter Ansatz genutzt werden, bei dem der Gesamtstichprobenumfang wie folgt errechnet wird:

$$n = \left( \frac{N \times z \times \sigma_{e1}}{TE - AE} \right)^2$$

Bei diesem vereinfachten Ansatz werden übrigens nur Informationen zur Fehlerstreuung im ersten Beobachtungszeitraum benötigt. Es wird davon ausgegangen, dass die Fehlerstreuung in beiden Halbjahren annähernd gleich groß ist.

Außerdem ist zu beachten, dass die Formeln zur Berechnung des Stichprobenumfangs Werte für  $N_1$  und  $N_2$  erfordern, d. h. für die Anzahl der Vorhaben in der Grundgesamtheit des ersten und des zweiten Halbjahres. Bei der Berechnung des Stichprobenumfangs wird der Wert für  $N_1$  bekannt sein, der Wert für  $N_2$  jedoch nicht. Er muss entsprechend den Erwartungen des Prüfers (und auch auf der Grundlage von

vorherigen Informationen) bestimmt werden. Das stellt normalerweise kein Problem dar, da alle im zweiten Halbjahr aktiven Vorhaben bereits im ersten Halbjahr vorhanden sind und daher  $N_1 = N_2$  gilt.

Nach Berechnung des Gesamtstichprobenumfangs  $n$  ist die Aufteilung der Stichprobe nach Halbjahr wie folgt:

$$n_1 = \frac{N_1}{N} n$$

und

$$n_2 = \frac{N_2}{N} n$$

### Zweites Halbjahr

Im ersten Beobachtungszeitraum wurden Annahmen getroffen, die sich auf die nachfolgenden Beobachtungszeiträume (normalerweise das nächste Halbjahr) beziehen. Weichen die Merkmale der Grundgesamtheit in den nachfolgenden Zeiträumen erheblich von den Annahmen ab, so muss der Stichprobenumfang für den nachfolgenden Zeitraum möglicherweise berichtigt werden.

Im zweiten Prüfzeitraum (z. B. Halbjahr) stehen mehr Informationen zur Verfügung:

- Die Anzahl der im Halbjahr  $N_2$  aktiven Vorhaben ist genau bekannt.
- Die aus der Stichprobe des ersten Halbjahres errechnete Stichprobenvarianz der Fehler  $s_{e1}^2$  ist bereits vorhanden.
- Die Standardabweichung der Fehler für das zweite Halbjahr  $\sigma_{e2}$  lässt sich anhand realer Daten nunmehr genauer bewerten.

Unterscheiden sich diese Parameter nicht wesentlich von denen, die im ersten Halbjahr entsprechend den Erwartungen des Prüfers geschätzt wurden, so braucht der ursprünglich für das zweite Halbjahr geplante Stichprobenumfang ( $n_2$ ) nicht korrigiert zu werden. Sieht der Prüfer jedoch große Unterschiede zwischen den ursprünglichen Annahmen und den tatsächlichen Merkmalen der Grundgesamtheit, so können Korrekturen des Stichprobenumfangs erforderlich sein, um den ungenauen Schätzungen Rechnung zu tragen. In diesem Fall ist der Stichprobenumfang des zweiten Halbjahres wie folgt neu zu berechnen:

$$n_2 = \frac{(z \cdot N_2 - \sigma_{e2})^2}{(TE - AE)^2 - z^2 \cdot \frac{N_1^2}{n_1} \cdot s_{e1}^2}$$

wobei  $s_{e1}$  die aus der Stichprobe des ersten Halbjahres errechnete Standardabweichung der Fehler ist und  $\sigma_{e2}$  eine Schätzung der Standardabweichung der Fehler im zweiten Halbjahr basierend auf vorherigen Erkenntnissen (möglicherweise angepasst an

Informationen aus dem ersten Halbjahr) oder einer Vorabstichprobe/Pilotstichprobe des zweiten Halbjahres ist.

### 7.1.3.3 Prognostizierter Fehler

Auf der Grundlage der beiden Teilstichproben jedes Halbjahres lässt sich der prognostizierte Fehler auf der Ebene der Grundgesamtheit anhand der beiden üblichen Methoden berechnen: Schätzung des Mittelwerts pro Einheit und Verhältnisschätzung.

#### Schätzung des Mittelwerts pro Einheit

In jedem Halbjahr wird der in der Stichprobe festgestellte durchschnittliche Fehler je Vorhaben mit der Anzahl der Vorhaben in der Grundgesamtheit ( $N_t$ ) multipliziert, und die Ergebnisse für beide Halbjahre werden addiert, woraus sich der prognostizierte Fehler ergibt:

$$EE_1 = \frac{N_1}{n_1} \sum_{i=1}^{n_1} E_{1if} + \frac{N_2}{n_2} \sum_{i=1}^{n_2} E_{2if}$$

#### Verhältnisschätzung

In jedem Halbjahr wird die in der Stichprobe festgestellte durchschnittliche Fehlerquote mit dem Buchwert der Grundgesamtheit des entsprechenden Halbjahrs ( $BV_t$ ) multipliziert:

$$EE_2 = BV_1 \times \frac{\sum_{i=1}^{n_1} E_{1if}}{\sum_{i=1}^{n_1} BV_{1if}} + BV_2 \times \frac{\sum_{i=1}^{n_2} E_{2if}}{\sum_{i=1}^{n_2} BV_{2if}}$$

Die Stichprobenfehlerquote in jedem Halbjahr errechnet sich durch einfache Division des Gesamtfehlerbetrags in der Stichprobe des Halbjahres durch den Gesamtbetrag der Ausgaben in der gleichen Stichprobe.

Bei der Entscheidung für eine der beiden Methoden sollten die für den Standardansatz des einfachen Zufallsstichprobenverfahrens angestellten Überlegungen als Grundlage dienen.

### 7.1.3.4 Genauigkeit

Wie beim Standardverfahren ist die Genauigkeit (Stichprobenfehler) das Maß der mit der Hochrechnung (Extrapolation) verbundenen Unsicherheit. Sie errechnet sich je nach dem gewählten Extrapolationsverfahren auf unterschiedliche Weise.

### Schätzung des Mittelwerts pro Einheit (absolute Fehler)

Für die Genauigkeit gilt die Formel

$$SE = z \times \sqrt{\left( N_1^2 \times \frac{s_{e1}^2}{n_1} + N_2^2 \times \frac{s_{e2}^2}{n_2} \right)}$$

wobei  $s_{et}$  die Standardabweichung der Fehler in der Stichprobe des Halbjahres  $t$  bezeichnet (nunmehr aus den Stichproben berechnet, die auch zur Hochrechnung der Fehler auf die Grundgesamtheit verwendet wurden)

$$s_{et}^2 = \frac{1}{n_t - 1} \sum_{i=1}^{n_t} (E_{it} - \bar{E}_t)^2$$

### Verhältnisschätzung (Fehlerquoten)

Für die Genauigkeit gilt die Formel

$$SE = z \times \sqrt{\left( N_1^2 \times \frac{s_{q1}^2}{n_1} + N_2^2 \times \frac{s_{q2}^2}{n_2} \right)}$$

Darin ist  $s_{qt}$  die Standardabweichung der Variablen  $q$  in der Stichprobe des Halbjahres  $t$ , wobei

$$q_{it} = E_{it} - \frac{\sum_{i=1}^{n_t} E_{it}}{\sum_{i=1}^{n_t} BV_{it}} \times BV_{it}$$

#### 7.1.3.5 Bewertung

Um auf die Erheblichkeit der Fehler zu schließen, ist zunächst die obere Fehlergrenze (ULE) zu berechnen. Diese Obergrenze entspricht der Summe aus dem prognostizierten Fehler  $EE$  selbst und der Genauigkeit der Extrapolation

$$ULE = EE + SE$$

Anschließend sollten der prognostizierte Fehler und die Obergrenze jeweils mit dem maximal zulässigen Fehler verglichen werden, um Prüfungsschlussfolgerungen unter Verwendung des Ansatzes zu ziehen, der in Abschnitt 7.1.1.5 dargelegt wurde.

### 7.1.3.6 Beispiel

Eine Prüfbehörde hat sich entschieden, den Prüfaufwand auf zwei Zeiträume aufzuteilen. Am Ende des ersten Halbjahres betrachtet sie die entsprechend den beiden Halbjahren in zwei Gruppen unterteilte Grundgesamtheit, die sich zu diesem Zeitpunkt durch folgende Merkmale auszeichnet:

Geltend gemachte Ausgaben am Ende des ersten Halbjahres	1 237 952 015 EUR
Größe der Grundgesamtheit (Vorhaben – erstes Halbjahr)	3852

Ausgehend von den vorliegenden Erfahrungen ist der Prüfbehörde bekannt, dass in der Regel alle in den Programmen am Ende des Bezugszeitraums enthaltenen Vorhaben bereits in der Grundgesamtheit des ersten Halbjahres aktiv sind. Ferner wird erwartet, dass die am Ende des ersten Halbjahres geltend gemachten Ausgaben etwa 30 % der zum Ende des Bezugsjahres insgesamt geltend gemachten Ausgaben ausmachen. Auf der Grundlage dieser Annahmen gibt die folgende Tabelle eine Übersicht über diese Grundgesamtheit:

Geltend gemachte Ausgaben am Ende des ersten Halbjahres	1 237 952 015 EUR
Geltend gemachte Ausgaben am Ende des zweiten Halbjahres (Vorhersage)	2 888 554 703 EUR
Größe der Grundgesamtheit (Vorhaben – Zeitraum 1)	3852
Größe der Grundgesamtheit (Vorhaben – Zeitraum 2, Vorhersage)	3852

Da die von der Prüfbehörde durchgeführten Systemprüfungen ein moderates Sicherheitsniveau ergaben, kann das Stichprobeverfahren für dieses Programm mit einem Konfidenzniveau von 80 % durchgeführt werden.

Im ersten Zeitraum wird der Gesamtstichprobenumfang (für beide Halbjahre) wie folgt errechnet:

$$n = \left( \frac{N \times z \times \sigma_w}{TE - AE} \right)^2$$

wobei  $\sigma_w^2$  der gewichtete Mittelwert der Varianzen der Fehler in jedem Halbjahr ist:

$$\sigma_w^2 = \frac{N_1}{N} \sigma_{e1}^2 + \frac{N_2}{N} \sigma_{e2}^2$$

und  $\sigma_{et}^2$  die Varianz der Fehler in jedem Zeitraum (Halbjahr) bezeichnet. Die Fehlervarianz der beiden Halbjahre wird jeweils als eigenständige Grundgesamtheit berechnet:

$$\sigma_{et}^2 = \frac{1}{n_t^e - 1} \sum_{i=1}^{n_t^e} (E_{it} - \bar{E}_t)^2, t = 1, 2$$

Dabei sind  $E_{it}$  die einzelnen Fehler der Einheiten in der Stichprobe des Halbjahres  $t$  und  $\bar{E}_t$  der mittlere Fehler der Stichprobe im Halbjahr  $t$ .

Da der Wert von  $\sigma_{et}^2$  unbekannt ist, beschloss die Prüfbehörde, am Ende des ersten laufenden Halbjahres eine Vorabstichprobe von 20 Vorhaben zu ziehen. Die Standardabweichung der Fehler in dieser Vorabstichprobe im ersten Halbjahr beträgt 72 091 EUR. Auf der Grundlage einer fachlichen Beurteilung und in dem Wissen, dass die Ausgaben im zweiten Halbjahr gewöhnlich höher sind als im ersten, hat die Prüfbehörde in einer vorläufigen Vorhersage die Standardabweichung der Fehler für das zweite Halbjahr um 40 % höher angesetzt als im ersten Halbjahr, d. h. 100 475 EUR. Daher beträgt der gewichtete Durchschnitt der Fehlervarianzen:

$$\begin{aligned} \sigma_w^2 &= \frac{N_1}{N_1 + N_2} \sigma_{e1}^2 + \frac{N_2}{N_1 + N_2} \sigma_{e2}^2 \\ &= \frac{3852}{3852 + 3852} \times 72,091^2 + \frac{3852}{3852 + 3852} \times 100,475^2 \\ &= 7,646,166,953. \end{aligned}$$

Hinweis: Die Größe der Grundgesamtheit in einem gegebenen Halbjahr entspricht der Anzahl der aktiven Vorhaben (mit Ausgaben) in dem entsprechenden Halbjahr.

Im ersten Halbjahr beträgt der für das gesamte Jahr geplante Stichprobenumfang:

$$n = \left( \frac{(N_1 + N_2) \times z \times \sigma_w}{TE - AE} \right)^2$$

wobei  $z$  den Wert 0,842 hat (Koeffizient, der einem Konfidenzniveau von 60 % entspricht) und  $TE$ , der zulässige Fehler, 2 % des Buchwerts (in der Verordnung festgelegter Höchstwert für die Erheblichkeitsschwelle) beträgt. Der Gesamtbuchwert umfasst den tatsächlichen Buchwert am Ende des ersten Halbjahres zuzüglich des

vorhergesagten Buchwerts für das zweite Halbjahr (1 237 952 015 EUR + 2 888 554 703 EUR = 4 126 506 718 EUR), d. h. der zulässige Fehler beträgt 2 % x 4 126 506 718 EUR = 82 530 134 EUR. Die Vorabstichprobe auf die Grundgesamtheit des ersten Halbjahres ergibt eine Stichprobenfehlerquote von 0,6 %. Die Prüfbehörde erwartet, dass diese Fehlerquote über das gesamte Jahr konstant bleibt. Es ergibt sich ein voraussichtlicher Fehler **AE** von 0,6 % x 4 126 506 718 EUR = 24 759 040 EUR. Der geplante Stichprobenumfang für das gesamte Jahr ist:

$$n = \left( \frac{(3852 + 3852) \times 0.842 \times \sqrt{7,646,168,953}}{82,530,134 - 24,759,040} \right)^2 \approx 97$$

Die Aufteilung der Stichprobe nach Halbjahr ist wie folgt:

$$n_1 = 0.5 n \approx 49$$

und

$$n_2 = n - n_1 = 49$$

Die Stichprobe des ersten Halbjahres führte zu den folgenden Ergebnissen:

Stichprobenbuchwert – erstes Halbjahr	13 039 581 EUR
Stichprobengesamtfehler – erstes Halbjahr	199 185 EUR
Stichprobenstandardabweichung der Fehler – erstes Halbjahr	69 815 EUR

Am Ende des zweiten Halbjahres stehen mehr Informationen zur Verfügung. Insbesondere die Anzahl der im zweiten Halbjahr aktiven Vorhaben ist genau bekannt, die aus der Stichprobe des ersten Halbjahres errechnete Stichprobenvarianz der Fehler  $s_{e1}$  ist bereits vorhanden, und die Standardabweichung der Fehler für das zweite Halbjahr  $\sigma_{e2}$  lässt sich anhand einer Vorabstichprobe realer Daten nunmehr genauer bewerten.

Die Prüfbehörde stellt fest, dass die am Ende des ersten Halbjahres getroffene Annahme zur Gesamtzahl der Vorhaben richtig war. Dennoch gibt es zwei Parameter, für die aktualisierte Zahlen verwendet werden sollten.

Erstens ergab die Schätzung der Standardabweichung der Fehler auf der Grundlage der Stichprobe des ersten Halbjahres von 49 Vorhaben einen Wert von 69 815 EUR. Dieser neue Wert sollte nun zur Neubewertung des geplanten Stichprobenumfangs herangezogen werden. Zweitens schätzt die Prüfbehörde auf der Grundlage einer neuen Vorabstichprobe von 20 Vorhaben der Grundgesamtheit des zweiten Halbjahres die Standardabweichung der Fehler für das zweite Halbjahr auf 108 369 EUR (dieser Wert liegt dicht bei dem am Ende des ersten Zeitraums vorhergesagten Wert, ist jedoch

genauer). Die Schlussfolgerung lautet dementsprechend, dass die zur Planung des Stichprobenumfangs verwendeten Standardabweichungen der Fehler in beiden Halbjahren den am Ende des ersten Halbjahres erzielten Werten nahekommen. Dennoch hat sich die Prüfbehörde für eine Neuberechnung des Stichprobenumfangs unter Verwendung der verfügbaren aktualisierten Daten entschieden. Die Stichprobe für das zweite Halbjahr wird also korrigiert.

Darüber hinaus sollte der vorhergesagte Gesamtbuchwert der Grundgesamtheit des zweiten Halbjahres – 2 888 554 703 EUR – durch den realen Wert 2 961 930 008 EUR ersetzt werden.

Parameter	Ende des ersten Halbjahres	Ende des zweiten Halbjahres
Standardabweichung der Fehler im ersten Halbjahr	72 091 EUR	69 815 EUR
Standardabweichung der Fehler im zweiten Halbjahr	100 475 EUR	108 369 EUR
Gesamtausgaben im zweiten Halbjahr	2 888 554 703 EUR	2 961 930 008 EUR

Unter Berücksichtigung dieser drei Anpassungen beträgt der neu berechnete Stichprobenumfang des zweiten Halbjahres

$$\begin{aligned}
 n_2 &= \frac{(z \times N_2 \times \sigma_{e2})^2}{(TE - AE)^2 - z^2 \times \frac{N_1^2}{n_1} \times s_{e1}^2} \\
 &= \frac{(0,842 \times 3,852 \times 108,369)^2}{(83,997,640 - 25,199,292)^2 - 0,842^2 \times \frac{3,852^2}{49} \times 69,815^2} = 52
 \end{aligned}$$

Mit der Prüfung der 49 Vorhaben im ersten Halbjahr zuzüglich dieser 52 Vorhaben im zweiten Halbjahr erhält der Prüfer Informationen über den Gesamtfehler für die Vorhaben der Stichprobe. Die vorherige Vorabstichprobe von 20 Vorhaben wird als Teil der Hauptstichprobe verwendet. Daher braucht der Prüfer im zweiten Halbjahr nur 32 weitere Vorhaben auszuwählen.

Die Stichprobe des zweiten Halbjahres führte zu den folgenden Ergebnissen:

Stichprobenbuchwert – zweites Halbjahr	34 323 574 EUR
Stichprobengesamtfehler – zweites Halbjahr	374 790 EUR
Stichprobenstandardabweichung der Fehler – zweites Halbjahr	59 489 EUR

Auf der Grundlage der beiden Teilstichproben lässt sich der prognostizierte Fehler auf der Ebene der Grundgesamtheit anhand der beiden üblichen Methoden berechnen: Schätzung des Mittelwerts pro Einheit und Verhältnisschätzung. Bei der ersten Methode wird der in der Stichprobe festgestellte durchschnittliche Fehler pro Vorgang mit der Zahl der Vorhaben in der Grundgesamtheit ( $N_t$ ) multipliziert, und die Ergebnisse für beide Halbjahre werden addiert, woraus sich der prognostizierte Fehler ergibt:

$$EE_1 = \frac{N_1}{n_1} \sum_{i=1}^{49} E_{1i} + \frac{N_2}{n_2} \sum_{i=1}^{52} E_{2i} = \frac{3,852}{49} \times 199,185 + \frac{3,852}{52} \times 374,790$$

$$= 43,318,770$$

Bei der zweiten Methode wird die in der Stichprobe festgestellte durchschnittliche Fehlerquote mit dem Buchwert der Grundgesamtheit des entsprechenden Halbjahrs ( $BV_t$ ) multipliziert:

$$EE_2 = BV_1 \times \frac{\sum_{i=1}^{n_1} E_{1i}}{\sum_{i=1}^{n_1} BV_{1i}} + BV_2 \times \frac{\sum_{i=1}^{n_2} E_{2i}}{\sum_{i=1}^{n_2} BV_{2i}}$$

$$= 1,237,952,015 \times \frac{199,185}{13,039,581} + 2,961,930,008 \times \frac{374,790}{34,323,574}$$

$$= 51,252,484$$

Wird eine Schätzung des Mittelwerts pro Einheit vorgenommen, so beträgt der prognostizierte Fehler:

$$r_1 = \frac{43,421,670}{1,237,952,015 + 2,961,930,008} = 1.03 \%$$

und bei der Verhältnisschätzung ist:

$$r_2 = \frac{51,252,451}{1,237,952,015 + 2,961,930,008} = 1.22 \%$$

Die Genauigkeit wird entsprechend der für die Hochrechnung verwendeten Methode auf unterschiedliche Weise berechnet. Bei einer Schätzung des Mittelwerts pro Einheit gilt für die Genauigkeit die folgende Formel:

$$SE_1 = z \times \sqrt{\left( N_1^2 \times \frac{s_{e1}^2}{n_1} + N_2^2 \times \frac{s_{e2}^2}{n_2} \right)}$$

$$= 0.842 \times \sqrt{3,852^2 \times \frac{69,815^2}{49} + 3,852^2 \times \frac{59,489^2}{52}} = 41,980,051$$

Bei der Verhältnisschätzung ist die Standardabweichung der Variablen  $q$  zu berechnen (Abschnitt 7.1.2.6):

$$q_{tt} = E_{tt} - \frac{\sum_{t=1}^{n_t} E_{tt}}{\sum_{t=1}^{n_t} BV_{tt}} \times BV_{tt}$$

Diese Standardabweichung beträgt für die Halbjahre 54 897 EUR bzw. 57 659 EUR. Die Genauigkeit errechnet sich daher folgendermaßen:

$$SE_2 = z \times \sqrt{\left( N_1^2 \times \frac{s_{q1}^2}{n_1} + N_2^2 \times \frac{s_{q2}^2}{n_2} \right)}$$

$$= 0.842 \times \sqrt{3,852^2 \times \frac{54,897^2}{49} + 3,852^2 \times \frac{57,659^2}{52}} = 36,325,544$$

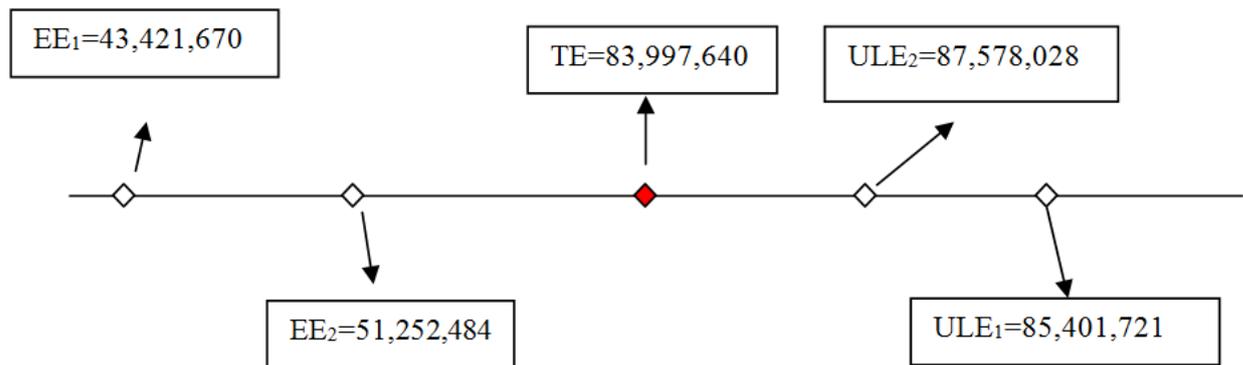
Anschließend sollten der prognostizierte Fehler und die Obergrenze jeweils mit dem maximal zulässigen Fehler verglichen werden, um Prüfungsschlussfolgerungen zu ziehen:

$$ULE_1 = EE_1 + SE_1 = 43,421,670 + 41,980,051 = 85,401,721$$

oder

$$ULE_2 = EE_2 + SE_2 = 51,252,484 + 36,325,544 = 87,578,028$$

Beim Vergleich der prognostizierten Ergebnisse mit der Erheblichkeitsschwelle von 2 % des gesamten Buchwerts der Grundgesamtheit (2 % x 4 199 882 023 EUR = 83 997 640 EUR) wird festgestellt, dass der maximal zulässige Fehler größer ist als die prognostizierten Fehler (bei Verwendung beider Methoden). Daher sind zusätzliche Arbeitsschritte (wie in Abschnitt 5.11 beschrieben) erforderlich, denn es ist nicht hinreichend belegt, dass die Grundgesamtheit keine wesentlichen Fehlangaben enthält.



## 7.2 Differenzenschätzung

### 7.2.1 Standardansatz

#### 7.2.1.1 Einleitung

Die Differenzenschätzung ist ebenfalls ein statistisches Stichprobenverfahren, das auf einer Auswahl mit gleichen Wahrscheinlichkeiten beruht. Das Verfahren stützt sich auf die Extrapolation des Fehlers in der Stichprobe und die Subtraktion des prognostizierten Fehlers von den gesamten geltend gemachten Ausgaben in der Grundgesamtheit, um die korrekten Ausgaben in der Grundgesamtheit abzuschätzen (d. h. die Ausgaben, die sich bei Prüfung sämtlicher Vorhaben in der Grundgesamtheit ergeben würden).

Dieses Verfahren kommt dem einfachen Zufallsstichprobenverfahren sehr nahe. Der Hauptunterschied besteht in der Verwendung eines komplexeren Extrapolationsansatzes.

Besonders nützlich ist dieses Verfahren, wenn die korrekten Ausgaben in der Grundgesamtheit prognostiziert werden sollen, wenn das Fehlerniveau in der Grundgesamtheit relativ konstant ist und wenn der Buchwert verschiedener Vorhaben ähnliche Größenordnungen aufweist (niedrige Streuung). Es ist in der Regel besser als das MUS-Verfahren, wenn Fehler eine geringe Streuung aufweisen oder ein schwacher oder negativer Zusammenhang mit den Buchwerten besteht. Bei einer hohen Streuung der Fehler und des Bestehens eines positiven Zusammenhangs mit den Buchwerten hingegen ist es tendenziell schlechter als das MUS-Verfahren.

Wie alle anderen Verfahren lässt sich dieses Verfahren mit einer Schichtung kombinieren (günstige Bedingungen für die Schichtung werden in Abschnitt 6.2 erörtert und konkrete Formeln in Abschnitt 7.2.2 vorgestellt).

### 7.2.1.2 Stichprobenumfang

Die Berechnung des Stichprobenumfangs  $n$  im Rahmen der Differenzschätzung beruht auf genau den gleichen Informationen und Formeln, die beim einfachen Zufallsstichprobenverfahren verwendet werden:

- Größe der Grundgesamtheit  $N$
- Konfidenzniveau, das mithilfe von Systemprüfungen und dem damit verwandten Koeffizienten  $z$  aus einer Normalverteilung ermittelt wird (siehe Abschnitt 6.4)
- maximal zulässiger Fehler  $TE$  (in der Regel 2 % der gesamten Ausgaben)
- voraussichtlicher Fehler  $AE$ , der vom Prüfer anhand seines fachlichen Urteils und vorhandener Informationen ausgewählt wird
- Standardabweichung  $\sigma_e$  der Fehler.

Der Stichprobenumfang wird wie folgt berechnet:

$$n = \left( \frac{N \times z \times \sigma_e}{TE - AE} \right)^2$$

Darin ist  $\sigma_e$  die Standardabweichung der Fehler in der Grundgesamtheit. Zu beachten ist, dass – wie bereits im Zusammenhang mit dem einfachen Zufallsstichprobenverfahren diskutiert – diese Standardabweichung fast nie im Voraus bekannt ist und sich die Mitgliedstaaten entweder auf historische Daten (Standardabweichung der Fehler für die Grundgesamtheit im vorangegangenen Zeitraum) oder auf eine Vorabstichprobe/Pilotstichprobe mit geringem Stichprobenumfang (für den nicht weniger als 20 bis 30 Einheiten empfohlen werden) stützen müssen. Außerdem kann die Pilotstichprobe in der Folge als Teil der zur Prüfung ausgewählten Stichprobe verwendet werden. Weitere Informationen zur Berechnung dieser Standardabweichung sind dem Abschnitt 7.1.1.2 zu entnehmen.

### 7.2.1.3 Extrapolation

Anhand einer nach dem Zufallsprinzip ausgewählten Stichprobe von Vorhaben, deren Umfang nach der oben genannten Formel berechnet wurde, lässt sich der prognostizierte Fehler auf der Ebene der Grundgesamtheit berechnen, indem der in der Stichprobe beobachtete durchschnittliche Fehler pro Vorhaben mit der Zahl der Vorhaben in der Grundgesamtheit multipliziert wird, woraus sich der prognostizierte Fehler ergibt

$$EE = N \times \frac{\sum_{i=1}^n E_i}{n}$$

Darin sind  $E_i$  die einzelnen Fehler der Einheiten in der Stichprobe und  $\bar{E}$  der mittlere Fehler der Stichprobe.

In einem zweiten Schritt lässt sich der korrekte Buchwert (die korrekten Ausgaben, die bei einer Prüfung aller Vorhaben in der Grundgesamtheit gefunden werden würden) prognostizieren, indem der prognostizierte Fehler (EE) vom Buchwert (BV) der Grundgesamtheit (geltend gemachte Ausgaben) abgezogen wird. Die Hochrechnung auf den korrekten Buchwert (CBV) errechnet sich aus folgender Formel:

$$CBV = BV - EE$$

#### 7.2.1.4 Genauigkeit

Für die Genauigkeit der Hochrechnung (Maß der mit der Hochrechnung verbundenen Unsicherheit) gilt

$$SE = N \times z \times \frac{s_e}{\sqrt{n}}$$

wobei  $s_e$  die Standardabweichung der Fehler in der Stichprobe ist (nunmehr aus der Stichprobe berechnet, die auch zur Hochrechnung der Fehler auf die Grundgesamtheit verwendet wurde)

$$s_e^2 = \frac{1}{n-1} \sum_{i=1}^n (E_i - \bar{E})^2$$

#### 7.2.1.5 Bewertung

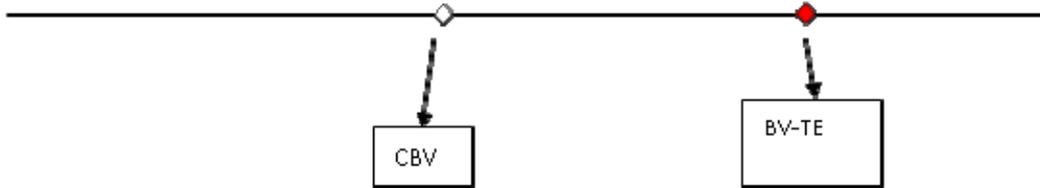
Um Rückschlüsse auf die Erheblichkeit der Fehler zu ermöglichen, wird zunächst die Untergrenze für den korrigierten Buchwert berechnet. Diese Untergrenze entspricht

$$LL = CBV - SE$$

Die Hochrechnungen auf den korrekten Buchwert und die Untergrenze sind jeweils mit der Differenz zwischen dem Buchwert (geltend gemachte Ausgaben) und dem maximal zulässigen Fehler (TE) zu vergleichen, die dem Produkt aus Erheblichkeitsschwelle und Buchwert entspricht:

$$BV - TE = BV - 2\% \times BV = 98\% \times BV$$

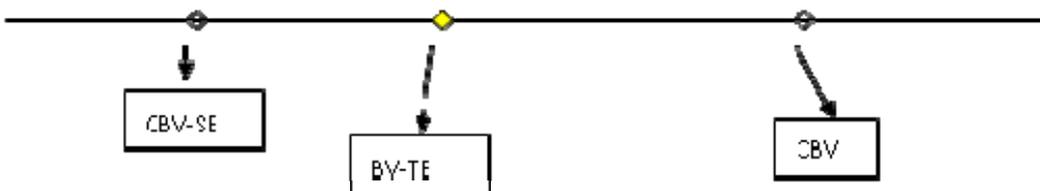
- Ist  $BV - TE$  größer als der  $CBV$ , sollte der Prüfer schlussfolgern, dass hinreichende Beweise dafür vorliegen, dass die Fehler im Programm oberhalb der Erheblichkeitsschwelle liegen:



- Liegt der Wert von  $BV - TE$  unter dem der Untergrenze  $CBV - SE$ , so bestehen hinreichende Beweise dafür, dass die Fehler im Programm unterhalb der Erheblichkeitsschwelle liegen.



- Liegt  $BV - TE$  zwischen der Untergrenze  $CBV - SE$  und dem  $CBV$ , sind weitere Schritte erforderlich, da nicht hinreichend belegt ist, dass die Grundgesamtheit keine wesentlichen Fehlangaben enthält. Die erforderlichen Schritte werden in Abschnitt 5.11 erörtert:



#### 7.2.1.6 Beispiel

Angenommen wird eine Grundgesamtheit von Ausgaben, die gegenüber der Kommission in einem bestimmten Jahr für Vorhaben in einem Programm bescheinigt wurden. Die von der Prüfbehörde durchgeführten Systemprüfungen ergaben ein hohes

Sicherheitsniveau. Das Stichprobeverfahren kann für dieses Programm daher mit einem Konfidenzniveau von 60 % durchgeführt werden.

In der nachstehenden Tabelle werden die Angaben zur Grundgesamtheit zusammengefasst:

Größe der Grundgesamtheit (Anzahl der Vorhaben)	3852
Buchwert (Summe der Ausgaben im Bezugsjahr)	4 199 882 024 EUR

Ausgehend von der Vorjahresprüfung rechnet die Prüfbehörde mit einer Fehlerquote von 0,7 % (Fehlerquote des letzten Jahres) und veranschlagt die Standardabweichung der Fehler auf 168 397 EUR.

Der erste Schritt besteht in der Berechnung des erforderlichen Stichprobenumfangs unter Verwendung der folgenden Formel:

$$n = \left( \frac{N \times z \times \sigma_e}{TE - AE} \right)^2$$

wobei  $z$  den Wert 0,842 hat (Koeffizient, der einem Konfidenzniveau von 60 % entspricht),  $\sigma_e$  168 397 EUR beträgt,  $TE$ , der zulässige Fehler, 2 % des Buchwerts (in der Verordnung festgelegter Höchstwert für die Erheblichkeitsschwelle) ausmacht, d. h. 2 % x 4 199 882 024 EUR = 83 997 640 EUR, und  $AE$ , der voraussichtliche Fehler, bei 0,7 % liegt, d. h. 0,7 % x 4 199 882 024 EUR = 29 399 174 EUR:

$$n = \left( \frac{3,852 \times 0,842 \times 168,397}{83,997,640 - 29,399,174} \right)^2 \approx 101$$

Der Mindestumfang der Stichprobe beläuft sich daher auf 101 Vorhaben.

Die Prüfung dieser 101 Vorhaben liefert dem Prüfer den Gesamtfehler für die Vorhaben der Stichprobe.

Die Ergebnisse der Stichprobe sind in der nachstehenden Tabelle zusammengefasst:

Stichprobenbuchwert („=SUM(B2:B102)“)	124 944 535 EUR
Stichprobengesamtfehler („=SUM(D2:D102)“)	1 339 765 EUR
Standardabweichung der Fehler in der Stichprobe („=STDEV(D2:D102)“)	162 976 EUR

Die folgende Abbildung veranschaulicht die Ergebnisse des Stichprobenverfahrens:

	A	B	C	D	
	Operation ID (1)	Book Value (BV) (2)	Audited Value (AV) (3)	Error (4)	
1					
2	4371	265,032 €	259,894 €	5,138 €	
3	2238	566,474 €	566,474 €	- €	
4	4972	388,768 €	388,768 €	- €	
5	3050	231,114 €	231,114 €	- €	
6	459	869,080 €	869,080 €	- €	
7	4232	11,856,372 €	11,856,372 €	- €	
8	1371	640,087 €	640,087 €	- €	
9	1278	372,259 €	372,259 €	- €	
10	4217	626,949 €	626,949 €	- €	
11	2381	221,181 €	221,181 €	- €	
12	4214	661,927 €	661,927 €	- €	
13	2868	611,219 €	611,219 €	- €	
14	2350	326,179 €	326,179 €	- €	
15	3121	835,153 €	835,153 €	- €	
16	2936	114,856 €	114,856 €	- €	
17	203	16,098,649 €	16,098,649 €	- €	
18	2073	859,992 €	859,992 €	- €	
19	1057	272,282 €	272,282 €	- €	
20	3270	759,543 €	698,060 €	61,483 €	
21	3915	767,864 €	767,864 €	- €	
22	(...)	(...)	(...)	(...)	
101	1482	601,270 €	593,783 €	7,486 €	
102	1109	644,971 €	644,971 €	- €	
103	Total:=SUM([Row 2]:[Row 102])		124,944,535 €	123,604,770 €	1,339,765
104	Sample standard deviation:=STDEV(D2:D102)---->				162,976
105					

Der prognostizierte Fehler auf der Ebene der Grundgesamtheit ist:

$$EE = N \times \frac{\sum_{i=1}^{101} E_i}{n} = 3,852 \times \frac{1,339,765}{101} = 51,096,780,$$

Dies entspricht einer prognostizierten Fehlerquote von:

$$r = \frac{51,096,780}{4,199,882,024} = 1,22 \%$$

Zur Hochrechnung auf den korrekten Buchwert (korrekter Ausgabenbetrag, der festgestellt werden könnte, wenn alle Vorhaben in der Grundgesamtheit geprüft werden würden) wird der prognostizierte Fehler (**EE**) vom Buchwert (**BV**) der Grundgesamtheit (geltend gemachte Ausgaben) abgezogen. Die Hochrechnung auf den korrekten Buchwert (**CBV**) erfolgt daher nach folgender Formel:

$$CBV = 4,199,882,024 - 51,096,780 = 4,148,785,244$$

Für die Genauigkeit der Hochrechnung gilt

$$SE = N \times z \times \frac{s_e}{\sqrt{n}} = 3,852 \times 0,842 \times \frac{162,976}{\sqrt{101}} = 52,597,044.$$

Durch Zusammenfassung des prognostizierten Fehlers und der Genauigkeit lässt sich eine Obergrenze für die Fehlerquote errechnen. Diese Obergrenze ist der Quotient aus der oberen Fehlergrenze und dem Buchwert der Grundgesamtheit. Die Obergrenze für die Fehlerquote ist daher:

$$f_{UL} = \frac{EF + SE}{BV} = \frac{51,096,780 + 52,597,044}{4,199,882,024} = 1,47 \%$$

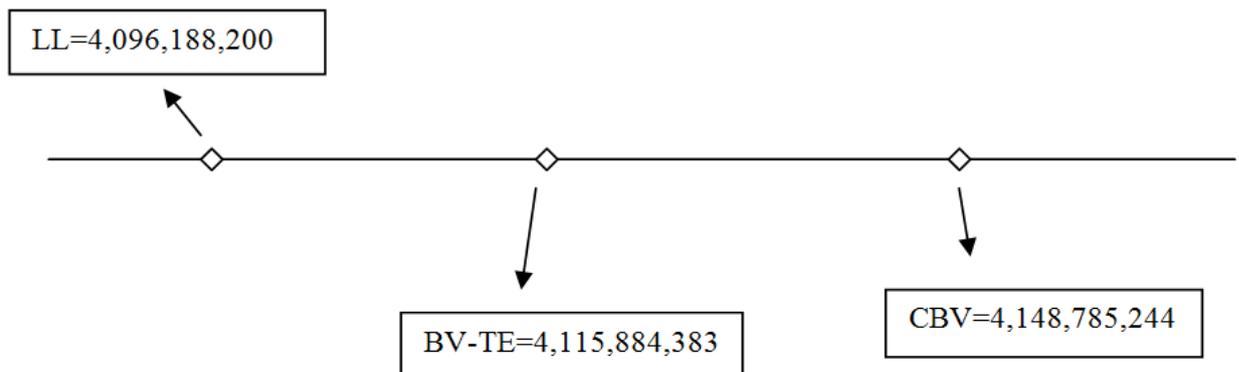
Um Rückschlüsse auf die Erheblichkeit der Fehler zu ermöglichen, wird zunächst die Untergrenze für den korrekten Buchwert berechnet. Diese Untergrenze entspricht

$$LL = CBV - SE = 4,148,785,244 - 52,597,044 = 4,096,188,200$$

Die Hochrechnungen auf den korrekten Buchwert und die Untergrenze sollten jeweils mit der Differenz zwischen dem Buchwert (geltend gemachte Ausgaben) und dem maximal zulässigen Fehler ( $TE$ ) verglichen werden:

$$BV - TE = 4,199,882,024 - 83,997,640 = 4,115,884,383$$

Liegt  $BV - TE$  zwischen der Untergrenze  $LL = CBV - SE$  und dem  $CBV$ , sind weitere Schritte erforderlich, um nachzuweisen, dass die Grundgesamtheit keine wesentlichen Fehlangaben enthält. Die erforderlichen Schritte werden in Abschnitt 5.11 erörtert:



## 7.2.2 Geschichtete Differenzschätzung

### 7.2.2.1 Einleitung

Bei der geschichteten Differenzschätzung wird die Grundgesamtheit in Teilgesamtheiten („Schichten“) unterteilt. Aus jeder Schicht werden unter Anwendung des Verfahrens der Differenzschätzung unabhängige Stichproben gezogen.

Die Gründe für die Anwendung der Schichtung und die Auswahlkriterien für die Schichtung sind die gleichen wie beim einfachen Zufallsstichprobenverfahren (siehe Abschnitt 7.1.2.1). Wie auch schon bei diesem Verfahren, ist die Schichtung nach Ausgabenniveau pro Vorhaben in der Regel immer dann ein guter Ansatz, wenn ein Zusammenhang zwischen Fehlerquote und Ausgabenniveau zu erwarten ist.

Erfolgt die Schichtung nach Ausgabenniveau und findet sich eine bestimmte Anzahl extrem hochwertiger Vorhaben, so wird empfohlen, diese in eine hochwertige Schicht einzufügen, die zu 100 % geprüft wird. Die zu dieser Schicht gehörenden Elemente sind in diesem Fall getrennt zu behandeln, so dass sich die einzelnen Schritte des Stichprobenverfahrens nur auf die Grundgesamtheit der niedrigwertigen Elemente beziehen. Dabei ist beachten, dass die geplante Genauigkeit für die Ermittlung des Stichprobenumfangs dennoch auf dem Gesamtbuchwert der Grundgesamtheit beruhen sollte. Da die Schicht der niedrigwertigen Elemente die Fehlerquelle darstellt, die geplante Genauigkeit jedoch auf der Ebene der Grundgesamtheit gelten soll, sind der zulässige Fehler und der voraussichtliche Fehler ebenfalls auf der Ebene der Grundgesamtheit zu berechnen.

#### 7.2.2.2 Stichprobenumfang

Der Stichprobenumfang wird nach dem gleichen Ansatz wie beim einfachen Zufallsstichprobenverfahren berechnet

$$n = \left( \frac{N \times z \times \sigma_w}{TE - AE} \right)^2$$

Darin bezeichnet  $\sigma_w^2$  den gewichteten Mittelwert der Fehlervarianzen für die gesamte Schichtenmenge (Einzelheiten siehe Abschnitt 7.1.2.2).

Wie üblich können die Varianzen auf historischem Wissen oder auf einer Vorabstichprobe/Pilotstichprobe mit geringem Stichprobenumfang beruhen. Im letztgenannten Fall kann die Pilotstichprobe in der Folge wie üblich als Teil der zur Prüfung ausgewählten Hauptstichprobe verwendet werden.

Nach Berechnung des Gesamtstichprobenumfangs  $n$  ist die nach Schicht vorgenommene Aufteilung der Stichprobe wie folgt:

$$n_h = \frac{N_h}{N} \times n.$$

Diese allgemeine Aufteilungsmethode wird in gleicher Form auch beim einfachen Zufallsstichprobenverfahren angewendet und als proportionale Aufteilung bezeichnet. Allerdings stehen auch hier andere Aufteilungsmethoden zur Verfügung.

### 7.2.2.3 Extrapolation

Anhand von  $H$  nach dem Zufallsprinzip ausgewählten Stichproben von Vorhaben, deren Umfang jeweils nach der oben genannten Formel berechnet wurde, lässt sich der prognostizierte Fehler auf der Ebene der Grundgesamtheit errechnen als:

$$EE = \sum_{h=1}^H N_h \frac{\sum_{t=1}^{n_h} E_t}{n_h}$$

In der Praxis wird in jeder Gruppe der Grundgesamtheit (Schicht) der Durchschnitt der festgestellten Fehler in der Stichprobe mit der Anzahl der Vorhaben in der Schicht ( $N_h$ ) multipliziert, und alle für die einzelnen Schichten erzielten Ergebnisse werden addiert.

In einem zweiten Schritt lässt sich der korrekte Buchwert (die korrekten Ausgaben, die bei Prüfung aller Vorhaben in der Grundgesamtheit gefunden werden würden) mit der folgenden Formel hochrechnen:

$$CBV = BV - \sum_{h=1}^H N_h \frac{\sum_{t=1}^{n_h} E_t}{n_h}$$

Mit dieser Formel wird 1) in jeder Schicht der Durchschnitt der festgestellten Fehler in der Stichprobe berechnet, 2) in jeder Schicht der durchschnittliche Stichprobenfehler mit der Schichtgröße ( $N_h$ ) multipliziert, 3) die Summe der Ergebnisse für alle Schichten gebildet, 4) dieser Wert vom Gesamtbuchwert der Grundgesamtheit (BV) subtrahiert. Das Ergebnis der Summe ist eine Hochrechnung des korrekten Buchwerts (CBV) in der Grundgesamtheit.

### 7.2.2.4 Genauigkeit

Genauigkeit (Stichprobenfehler) ist das Maß der mit der Hochrechnung (Extrapolation) verbundenen Unsicherheit. Bei der geschichteten Differenzschätzung ergibt sie sich aus der folgenden Formel:

$$SE = N \times z \times \frac{s_w}{\sqrt{n}}$$

wobei  $s_w^2$  der gewichtete Mittelwert der Varianz der Fehler für die gesamte Schichtenmenge ist, berechnet aus der Stichprobe, die auch zur Hochrechnung der Fehler auf die Grundgesamtheit verwendet wurde:

$$s_w^2 = \sum_{h=1}^H \frac{N_h}{N} s_{\hat{e}_h}^2, h = 1, 2, \dots, H;$$

und  $s_{\hat{e}_h}^2$  ist die geschätzte Varianz der Fehler für die Stichprobe von Schicht  $h$

$$s_{\hat{e}_h}^2 = \frac{1}{n_h - 1} \sum_{i=1}^{n_h} (E_{hit} - \bar{E}_h)^2, h = 1, 2, \dots, H$$

### 7.2.2.5 Bewertung

Um Rückschlüsse auf die Erheblichkeit der Fehler zu ermöglichen, wird zunächst die Untergrenze für den korrigierten Buchwert berechnet. Diese Untergrenze entspricht

$$LL = CBV - SE$$

Die Hochrechnungen auf den korrekten Buchwert und die Untergrenze sollten jeweils mit der Differenz zwischen dem Buchwert (geltend gemachte Ausgaben) und dem maximal zulässigen Fehler ( $TE$ ) verglichen werden:

$$BV - TE = BV - 2\% \times BV = 98\% \times BV$$

Außerdem sollten Prüfungsschlussfolgerungen unter Verwendung des Ansatzes gezogen werden, der in Abschnitt 7.2.1.5 für die Standarddifferenzschätzung dargelegt wurde.

### 7.2.2.6 Beispiel

Ausgegangen wird von einer Grundgesamtheit von Ausgaben, die gegenüber der Kommission in einem bestimmten Jahr für Vorhaben in einer Gruppe von Programmen bescheinigt wurden. Das Verwaltungs- und Kontrollsystem wird für die gesamte Gruppe von Programmen verwendet, und die von der Prüfbehörde durchgeführten Systemprüfungen ergaben ein hohes Sicherheitsniveau. Das Stichprobeverfahren kann für dieses Programm daher mit einem Konfidenzniveau von 60 % durchgeführt werden.

Die Prüfbehörde hat Grund zu der Annahme, dass bei hochwertigen Vorhaben unabhängig vom jeweiligen Programm erhebliche Fehlerrisiken bestehen. Auch gibt es Anhaltspunkte dafür, dass in den Programmen mit unterschiedlichen Fehlerquoten zu rechnen ist. Unter Berücksichtigung all dieser Aspekte entscheidet sich die Prüfbehörde für eine Schichtung der Grundgesamtheit nach Programm und nach Ausgaben (wobei in einer Stichprobenschicht mit 100 %-Prüfung alle Vorhaben mit einem über der Erheblichkeitsschwelle liegenden Buchwert isoliert werden).

In der nachstehenden Tabelle sind die vorhandenen Informationen zusammengefasst:

Größe der Grundgesamtheit (Anzahl der Vorhaben)	4872
Größe der Grundgesamtheit – Schicht 1 (Anzahl der Vorhaben in Programm 1)	1520
Größe der Grundgesamtheit – Schicht 2 (Anzahl der Vorhaben in Programm 2)	3347
Größe der Grundgesamtheit – Schicht 3 (Anzahl der Vorhaben mit BV > Erheblichkeitsschwelle)	5
Buchwert (Summe der Ausgaben im Bezugsjahr)	6 440 727 190 EUR
Buchwert – Schicht 1 (Gesamtausgaben in Programm 1)	3 023 598 442 EUR
Buchwert – Schicht 2 (Gesamtausgaben in Programm 2)	2 832 769 525 EUR
Buchwert – Schicht 3 (Gesamtausgaben der Vorhaben mit BV > Erheblichkeitsschwelle)	584 359 223 EUR

Die Stichprobenschicht mit 100 %-Prüfung, die die fünf hochwertigen Vorhaben enthält, sollte – wie in Abschnitt 7.2.2.1 dargelegt – separat behandelt werden. Daher entspricht nachfolgend der Wert von  $N$  der Gesamtzahl der Vorhaben in der Grundgesamtheit abzüglich der Zahl der in der Stichprobenschicht mit 100 %-Prüfung enthaltenen Vorhaben, d. h. 4867 (= 4872 – 5) Vorhaben.

Der erste Schritt besteht in der Berechnung des erforderlichen Stichprobenumfangs unter Verwendung der folgenden Formel:

$$n = \left( \frac{N \times z \times \sigma_w}{TE \quad AE} \right)^2$$

wobei  $z$  den Wert 0,842 hat (Koeffizient, der einem Konfidenzniveau von 60 % entspricht) und  $TE$ , der zulässige Fehler, 2 % (in der Verordnung festgelegter Höchstwert für die Erheblichkeitsschwelle) des Buchwerts beträgt, d. h. 2 % x 6 440 727 190 EUR = 128 814 544 EUR. Ausgehend von den Erfahrungen des Vorjahres und der Schlussfolgerung im Bericht zu den Verwaltungs- und Kontrollsystemen rechnet die Prüfbehörde mit einer Fehlerquote von höchstens 0,4 %. Damit beträgt  $AE$ , der voraussichtliche Fehler, 0,4 %, d. h. 0,4 % x 6 440 727 190 EUR = 25 762 909 EUR.

Da es sich bei der dritten Schicht um eine komplett zu prüfende Stichprobenschicht handelt, hat sie einen fixen Stichprobenumfang, der der Größe der Grundgesamtheit (mit den fünf hochwertigen Vorhaben) entspricht. Die Berechnung des Stichprobenumfangs für die übrigen zwei Schichten erfolgt anhand der nachstehenden Formel, wobei  $\sigma_w^2$  den gewichteten Durchschnitt der Varianzen der Fehler für die beiden übrigen Schichten bezeichnet:

$$\sigma_w^2 = \sum_{i=1}^2 \frac{N_i}{N} \sigma_{eh,i}^2, h = 1, 2;$$

$\sigma_{eh}^2$  ist die Fehlervarianz in jeder Schicht. Die Varianz der Fehler wird für jede Schicht als unabhängige Grundgesamtheit berechnet als

$$\sigma_{eh}^2 = \frac{1}{n_h^p - 1} \sum_{i=1}^{n_h^p} (E_{hit} - \bar{E}_h)^2, h = 1, 2, \dots, H$$

Dabei sind  $E_{hit}$  die einzelnen Fehler der Einheiten in der Stichprobe von Schicht  $h$  und  $\bar{E}_h$  der mittlere Fehler der Stichprobe in der Schicht  $h$ . Eine Vorabstichprobe von 20 Vorhaben der Schicht 1 ergab einen Schätzwert für die Standardabweichung der Fehler von 21 312 EUR:

	A	B	C	D
1	<b>Operation ID</b>	<b>Book Value (BV)</b>	<b>Audited Value (AV)</b>	<b>Error</b>
2	143	626,949 €	626,949 €	- €
3	2,246	265,032 €	265,032 €	- €
4	2,841	372,259 €	372,259 €	- €
5	1,757	611,219 €	609,878 €	1,342 €
6	4,998	566,474 €	566,474 €	- €
7	4,828	231,114 €	231,114 €	- €
8	646	326,179 €	326,179 €	- €
9	2,635	11,856,372 €	11,856,372 €	- €
10	3,873	221,181 €	221,181 €	- €
11	834	388,768 €	388,768 €	- €
12	4,738	114,856 €	114,856 €	- €
13	1,695	759,543 €	759,543 €	- €
14	2,589	272,282 €	272,282 €	- €
15	2,917	16,098,649 €	16,098,649 €	- €
16	1,434	661,927 €	661,927 €	- €
17	4,077	767,864 €	672,495 €	95,370 €
18	953	869,080 €	869,080 €	- €
19	232	835,153 €	835,153 €	- €
20	1,097	859,992 €	859,992 €	- €
21	166	640,087 €	640,087 €	- €
22	<b>Total</b>	37,344,981 €	37,248,270 €	96,712 €
23	<b>Sample error rate:=D22/B22</b>			0.3%
24	<b>Sample standard deviation of errors:=STDEV(D2:D21)----&gt;</b>			21,312 €
25				

Das gleiche Verfahren wurde für die Grundgesamtheit von Schicht 2 angewandt. Eine Vorabstichprobe von 20 Vorhaben der Schicht 2 ergab einen Schätzwert für die Standardabweichung der Fehler von 215 546 EUR:

Schicht 1 – Vorabschätzwert für die Standardabweichung der Fehler	21 312 EUR
Schicht 2 – Vorabschätzwert für die Standardabweichung der Fehler	215 546 EUR

Daher beträgt der gewichtete Mittelwert der Fehlervarianzen dieser beiden Schichten

$$\sigma_w^2 = \frac{1,520}{4,867} \times 21,312^2 + \frac{3,347}{4,867} \times 215,546^2 = 32,059,168,205$$

Der Mindestumfang der Stichproben errechnet sich wie folgt:

$$n = \left( \frac{4,867 \times 0,845 \times \sqrt{32,092,103,451}}{128,814,544 - 25,762,909} \right)^2 \approx 51$$

Die Aufteilung dieser 51 Vorhaben nach Schicht ist wie folgt:

$$n_1 = \frac{1,520}{4,867} \times 51 \approx 16,$$

$$n_2 = n - n_1 = 35$$

und

$$n_3 = N_3 = 5$$

Der Gesamtumfang der Stichprobe beträgt daher 56 Vorhaben.

- 16 Vorhaben der Vorabstichprobe von Schicht 1 plus
- 35 Vorhaben von Schicht 2 (die 20 Vorhaben der Vorabstichprobe plus eine zusätzliche Stichprobe von 15 Vorhaben) plus
- 5 hochwertige Vorhaben.

Die folgende Tabelle weist die Ergebnisse der gesamten Stichprobe von 60 Vorhaben aus:

<b>Stichprobenergebnisse – Schicht 1</b>		
A	Stichprobenbuchwert	37 344 981 EUR
B	Stichprobengesamtfehler	77 376 EUR
C	Durchschnittlicher Stichprobenfehler (C = B ÷ 16)	4 836 EUR
D	Standardabweichung der Fehler in der Stichprobe	16 783 EUR
<b>Stichprobenergebnisse – Schicht 2</b>		
E	Stichprobenbuchwert	722 269 643 EUR
F	Stichprobengesamtfehler	264 740 EUR
G	Durchschnittlicher Stichprobenfehler (G = F ÷ 35)	7 564 EUR
H	Standardabweichung der Fehler in der Stichprobe	117 335 EUR
<b>Stichprobenergebnisse – Schicht mit 100 %-Prüfung</b>		
I	Stichprobenbuchwert	584 359 223 EUR
J	Stichprobengesamtfehler	7 240 855 EUR
K	Durchschnittlicher Stichprobenfehler (I = J ÷ 5)	1 448 171 EUR

Die Berechnung des prognostizierten Fehlers für die beiden Stichprobenschichten erfolgt durch Multiplikation des durchschnittlichen Stichprobenfehlers mit der Größe der Grundgesamtheit. Die Summe dieser beiden Zahlen, addiert zu dem in der Stichprobenschicht mit 100 %-Prüfung festgestellten Fehler, ist der erwartete Fehler auf der Ebene der Grundgesamtheit:

$$EE = \sum_{h=1}^3 1520 \times 4,836 + 3,347 \times 7,564 + 7,240,855 = 39,908,283$$

Die prognostizierte Fehlerquote ist der Quotient aus dem hochgerechneten Fehler und dem Buchwert der Grundgesamtheit (Gesamtausgaben):

$$r_1 = \frac{39,908,283}{6,440,727,190} = 0.62 \%$$

Der korrekte Buchwert (die korrekten Ausgaben, die bei Prüfung aller Vorhaben in der Grundgesamtheit gefunden würden) lässt sich mit der folgenden Formel prognostizieren:

$$CBV = BV - EE = 6,440,727,190 - 39,908,283 = 6,400,818,907$$

Ausgehend von den Standardabweichungen der Fehler in den Stichproben der beiden Schichten (Tabelle mit Stichprobenergebnissen) beträgt der gewichtete Mittelwert der Varianz der Fehler für alle Stichprobenschichten:

$$s_{wr}^2 = \sum_{h=1}^2 \frac{N_h}{N} s_{wh}^2 = \frac{1,520}{4,867} \times 16,783^2 + \frac{3,347}{4,867} \times 117,335^2 = 9,555,777,062$$

Für die Genauigkeit der Hochrechnung gilt

$$SE = N \times z \times \frac{s_{wr}}{\sqrt{n}} = 4,867 \times 0,842 \times \frac{\sqrt{9,555,777,062}}{\sqrt{51}} = 56,094,639$$

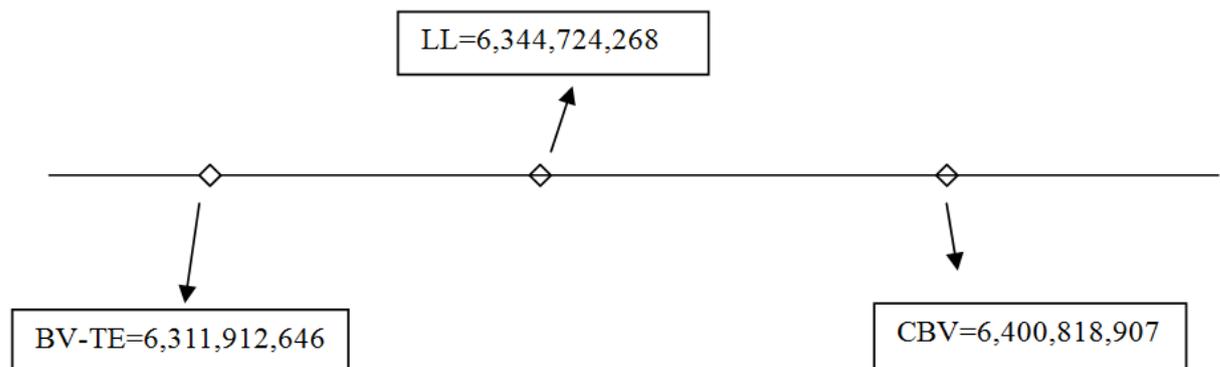
Um Rückschlüsse auf die Erheblichkeit der Fehler zu ermöglichen, wird zunächst die Untergrenze für den korrigierten Buchwert berechnet. Diese Untergrenze entspricht

$$LL = CBV - SE = 6,400,818,907 - 56,094,639 = 6,344,724,268$$

Die Hochrechnungen auf den korrekten Buchwert und die Untergrenze sollten jeweils mit der Differenz zwischen dem Buchwert (geltend gemachte Ausgaben) und dem maximal zulässigen Fehler ( $TE$ ) verglichen werden:

$$BV - TE = 6,440,727,190 - 128,814,544 = 6,311,912,646$$

Da  $BV - TE$  einen geringeren Wert aufweist als die Untergrenze  $CBV - SE$ , bestehen hinreichende Beweise dafür, dass die Fehler im Programm unterhalb der Erheblichkeitsschwelle liegen.



### 7.2.3 Differenzenschätzung – zwei Zeiträume

#### 7.2.3.1 Einleitung

Die Prüfbehörde kann das Stichprobenverfahren in mehreren Zeiträumen während des Jahres (typischerweise zwei Halbjahre) durchführen. Der wichtigste Vorteil dieses Ansatzes steht nicht im Zusammenhang mit einer Reduzierung des

Stichprobenumfangs, sondern mit der Möglichkeit, den Prüfungsaufwand über das Jahr zu verteilen und somit den Aufwand zu verringern, der bei nur einer Beobachtung am Ende des Jahres entstünde.

Mit diesem Ansatz wird die Jahresgrundgesamtheit in zwei Teilgesamtheiten aufgeteilt, die jeweils den Vorhaben und Ausgaben eines Halbjahres entsprechen. Für jedes Halbjahr werden unabhängige Stichproben unter Anwendung des üblichen Zufallsstichprobenverfahrens gezogen.

### 7.2.3.2 Stichprobenumfang

Der Stichprobenumfang wird nach dem gleichen Ansatz wie beim einfachen Zufallsstichprobenverfahren in zwei Halbjahren berechnet. Einzelheiten dazu siehe Abschnitt 7.1.3.2.

### 7.2.3.3 Extrapolation

Auf der Basis der beiden Teilstichproben jedes Halbjahres lässt sich der prognostizierte Fehler auf der Ebene der Grundgesamtheit wie folgt berechnen:

$$EE = N_1 \cdot \frac{\sum_{i=1}^{n_1} E_{1i}}{n_1} + N_2 \cdot \frac{\sum_{i=1}^{n_2} E_{2i}}{n_2}$$

In der Praxis wird in jedem Halbjahr der Durchschnitt der festgestellten Fehler in der Stichprobe mit der Anzahl der Vorhaben in der Grundgesamtheit ( $N_i$ ) multipliziert, und die Ergebnisse für die beiden Halbjahre werden so dann addiert.

In einem zweiten Schritt lässt sich der korrekte Buchwert (die korrekten Ausgaben, die bei Prüfung aller Vorhaben in der Grundgesamtheit gefunden werden würden) mit der folgenden Formel hochrechnen:

$$CBV = BV - EE$$

wobei  $BV$  der jährliche Buchwert (beide Halbjahre) ist und  $EE$  den genannten prognostizierten Fehler bezeichnet.

### 7.2.3.4 Genauigkeit

Genauigkeit (Stichprobenfehler) ist das Maß der mit der Hochrechnung (Extrapolation) verbundenen Unsicherheit. Sie ergibt sich aus der folgenden Formel

$$SE = z \times \sqrt{\left( N_1^2 \times \frac{s_{e1}^2}{n_1} + N_2^2 \times \frac{s_{e2}^2}{n_2} \right)}$$

wobei  $s_{et}$  die Standardabweichung der Fehler in der Stichprobe des Halbjahres  $t$  bezeichnet (nunmehr aus den Stichproben berechnet, die auch zur Hochrechnung der Fehler auf die Grundgesamtheit verwendet wurden)

$$s_{et}^2 = \frac{1}{n_t - 1} \sum_{i=1}^{n_t} (E_{it} - E_t)^2$$

### 7.2.3.5 Bewertung

Um Rückschlüsse auf die Erheblichkeit der Fehler zu ermöglichen, wird zunächst die Untergrenze für den korrigierten Buchwert berechnet. Diese Untergrenze entspricht

$$LL = CBV - SE$$

Die Hochrechnungen auf den korrekten Buchwert und die Untergrenze sollten jeweils mit der Differenz zwischen dem Buchwert (geltend gemachte Ausgaben) und dem maximal zulässigen Fehler ( $TE$ ) verglichen werden:

$$BV - TE = BV - 2\% \times BV = 98\% \times BV$$

Außerdem sollten Prüfungsschlussfolgerungen unter Verwendung des Ansatzes gezogen werden, der in Abschnitt 7.2.1.5 für die Standarddifferenzschätzung dargelegt wurde.

### 7.2.3.6 Beispiel

Eine Prüfbehörde hat sich entschieden, den Prüfaufwand zwischen den beiden Halbjahren aufzuteilen. Am Ende des ersten Halbjahres stellen sich die Merkmale der Grundgesamtheit wie folgt dar:

Geltend gemachte Ausgaben (DE) am Ende des ersten Halbjahrs	1 237 952 015 EUR
Größe der Grundgesamtheit (Vorhaben – erstes Halbjahr)	3852

Ausgehend von den vorliegenden Erfahrungen ist der Prüfbehörde bekannt, dass in der Regel alle in den Programmen am Ende des Bezugszeitraums enthaltenen Vorhaben

bereits in der Grundgesamtheit des ersten Halbjahres aktiv sind. Ferner wird erwartet, dass die am Ende des ersten Halbjahres geltend gemachten Ausgaben etwa 30 % der zum Ende des Bezugsjahres insgesamt geltend gemachten Ausgaben ausmachen. Auf der Grundlage dieser Annahmen gibt die folgende Tabelle eine Übersicht über diese Grundgesamtheit:

Geltend gemachte Ausgaben (DE) am Ende des ersten Halbjahrs	1 237 952 015 EUR
Geltend gemachte Ausgaben (DE) am Ende des zweiten Halbjahres (Vorhersage)	2 888 554 703 EUR
Größe der Grundgesamtheit (Vorhaben – Zeitraum 1)	3852
Größe der Grundgesamtheit (Vorhaben – Zeitraum 2, Vorhersage)	3852

Die von der Prüfbehörde durchgeführten Systemprüfungen ergaben ein niedriges Sicherheitsniveau. Die Stichprobenerhebung für dieses Programm sollte daher mit einem Konfidenzniveau von 90 % durchgeführt werden.

Am Ende des ersten Halbjahres wird der Gesamtstichprobenumfang (für beide Semester) wie folgt errechnet:

$$n = \left( \frac{N \times z \times \sigma_w}{TE - AE} \right)^2$$

wobei  $\sigma_w^2$  der gewichtete Mittelwert der Fehlervarianzen in jedem Halbjahr ist:

$$\sigma_w^2 = \frac{N_1}{N} \sigma_{e1}^2 + \frac{N_2}{N} \sigma_{e2}^2$$

und  $\sigma_{et}^2$  die Varianz der Fehler in jedem Zeitraum  $t$  (Halbjahr) bezeichnet. Die Fehlervarianz der beiden Halbjahre wird jeweils als eigenständige Grundgesamtheit berechnet

$$\sigma_{et}^2 = \frac{1}{n_t^p - 1} \sum_{i=1}^{n_t^p} (E_{it} - \bar{E}_t)^2, t = 1, 2$$

Dabei sind  $E_{it}$  die einzelnen Fehler der Einheiten in der Stichprobe des Halbjahres  $t$  und  $\bar{E}_t$  der mittlere Fehler der Stichprobe im Halbjahr  $t$ .

Da der Wert von  $\sigma_{et}^2$  unbekannt ist, beschloss die Prüfbehörde, am Ende des ersten laufenden Halbjahres eine Vorabstichprobe von 20 Vorhaben zu ziehen. Die Standardabweichung der Fehler in dieser Vorabstichprobe im ersten Halbjahr beträgt

69 534 EUR. Auf der Grundlage einer fachlichen Beurteilung und in dem Wissen, dass die Ausgaben im zweiten Halbjahr gewöhnlich höher sind als im ersten, hat die Prüfbehörde in einer vorläufigen Vorhersage die Standardabweichung der Fehler für das zweite Halbjahr um 20 % höher angesetzt als im ersten Halbjahr, d. h. 83 441 EUR. Daher beträgt der gewichtete Durchschnitt der Fehlervarianzen:

$$\sigma_w^2 = \frac{N_1}{N_1 + N_2} \sigma_{e1}^2 + \frac{N_2}{N_1 + N_2} \sigma_{e2}^2 = 0.5 \times 69,534^2 + 0.5 \times 83,441^2 = 5,898,672,130.$$

Hinweis: Die Größe der Grundgesamtheit in einem gegebenen Halbjahr entspricht der Anzahl der aktiven Vorhaben (mit Ausgaben) in dem entsprechenden Halbjahr.

Am Ende des ersten Halbjahres beträgt der Stichprobenumfang für das gesamte Jahr:

$$n = \left( \frac{N \times z \times \sigma_w}{TE - AE} \right)^2$$

wobei  $\sigma_w^2$  den gewichteten Durchschnitt der Varianzen der Fehler für die gesamte Schichtenmenge (Einzelheiten siehe Abschnitt 7.1.2.2) bezeichnet,  $z$  1,645 (Koeffizient, der einem Konfidenzniveau von 90 % entspricht) beträgt und  $TE$ , der zulässige Fehler, 2 % des Buchwerts (in der Verordnung festgelegter Höchstwert für die Erheblichkeitsschwelle) beträgt. Der Gesamtbuchwert umfasst den tatsächlichen Buchwert am Ende des ersten Halbjahres zuzüglich des vorhergesagten Buchwerts für das zweite Halbjahr 4 126 506 718, d. h. der zulässige Fehler beträgt 2 % x 4 126 506 718 EUR = 82 530 134 EUR. Die Vorabstichprobe auf die Grundgesamtheit des ersten Halbjahres ergibt eine Stichprobenfehlerquote von 0,6 %. Die Prüfbehörde erwartet, dass diese Fehlerquote über das gesamte Jahr konstant bleibt. Es ergibt sich ein voraussichtlicher Fehler  $AE$  von 0,6 % x 4 126 506 718 EUR = 24 759 040 EUR. Der Stichprobenumfang für das gesamte Jahr ist:

$$n = \left( \frac{3852 \times 2 \times 1.645 \times \sqrt{5,898,672,130}}{82,530,134 - 24,759,040} \right)^2 \approx 264$$

Die Aufteilung der Stichprobe nach Halbjahr ist wie folgt:

$$n_1 = 0.5 n = 142$$

und

$$n_2 = n - n_1 = 142$$

Die Stichprobe des ersten Halbjahres führte zu den folgenden Ergebnissen:

Stichprobenbuchwert – erstes Halbjahr	41 009 806 EUR
Stichprobengesamtfehler – erstes Halbjahr	577 230 EUR
Standardabweichung der Fehler – erstes Halbjahr	65 815 EUR

Am Ende des zweiten Halbjahres stehen mehr Informationen zur Verfügung. Insbesondere die Anzahl der im zweiten Halbjahr aktiven Vorhaben ist genau bekannt, die aus der Stichprobe des ersten Halbjahres errechnete Stichprobenvarianz der Fehler  $s_{e1}$  ist bereits vorhanden, und die Standardabweichung der Fehler für das zweite Halbjahr  $\sigma_{e2}$  lässt sich anhand einer Vorabstichprobe realer Daten nunmehr genauer bewerten.

Die Prüfbehörde stellt fest, dass die am Ende des ersten Halbjahres getroffene Annahme zur Gesamtzahl der Vorhaben richtig war. Dennoch gibt es zwei Parameter, für die aktualisierte Zahlen verwendet werden sollten.

Erstens ergab die Schätzung der Standardabweichung der Fehler auf der Grundlage der Stichprobe des ersten Halbjahres von 142 Vorhaben einen Wert von 65 815 EUR. Dieser neue Wert sollte nun zur Neubewertung des geplanten Stichprobenumfangs herangezogen werden. Zweitens schätzt die Prüfbehörde auf der Grundlage einer neuen Vorabstichprobe von 20 Vorhaben der Grundgesamtheit des zweiten Halbjahres die Standardabweichung der Fehler für das zweite Halbjahr auf 107 369 EUR (weit entfernt vom am Ende des ersten Zeitraums vorhergesagten Wert). Die Schlussfolgerung lautet dementsprechend, dass die zur Planung des Stichprobenumfangs verwendete Standardabweichung der Fehler im ersten Halbjahr dem am Ende des ersten Halbjahres erzielten Wert nahekommt. Dennoch ist die zur Planung des Stichprobenumfangs verwendete Standardabweichung der Fehler im zweiten Halbjahr weit von der durch die neue Vorabstichprobe gelieferten Zahl entfernt. Die Stichprobe für das zweite Halbjahr muss also korrigiert werden.

Darüber hinaus sollte der vorhergesagte Gesamtbuchwert der Grundgesamtheit des zweiten Halbjahres - 2 888 554 703 EUR - durch den realen Wert 5 202 775 175 EUR ersetzt werden.

Parameter	Ende des ersten Halbjahres	Ende des zweiten Halbjahres
Standardabweichung der Fehler im ersten Halbjahr	69 534 EUR	65 815 EUR
Standardabweichung der Fehler im zweiten Halbjahr	83 441 EUR	107 369 EUR
Gesamtausgaben im zweiten Halbjahr	2 888 554 703 EUR	5 202 775 175 EUR

Unter Berücksichtigung dieser beiden Anpassungen beträgt der neu berechnete Stichprobenumfang des zweiten Halbjahres

$$n_2 = \frac{(z \times N_2 \times \sigma_{z2})^2}{(TE - AE)^2 - z^2 \times \frac{N_1^2}{n_1} \times s_{z1}^2}$$

$$= \frac{(1.645 \times 3.852 \times 107.369)^2}{(128.814.544 - 38.644.363) - 1.645^2 \times \frac{3.852^2}{142} \times 65.815^2} \approx 68$$

Mit der Prüfung der 142 Vorhaben im ersten Halbjahr zuzüglich dieser 68 Vorhaben im zweiten Halbjahr erhält der Prüfer Informationen über den Gesamtfehler für die Vorhaben der Stichprobe. Die vorherige Vorabstichprobe von 20 Vorhaben wird als Teil der Hauptstichprobe verwendet. Daher braucht der Prüfer im zweiten Halbjahr nur 48 weitere Vorhaben auszuwählen.

Die Stichprobe des zweiten Halbjahres führte zu den folgenden Ergebnissen:

Stichprobenbuchwert – zweites Halbjahr	59 312 212 EUR
Stichprobengesamtfehler – zweites Halbjahr	588 336 EUR
Stichprobenstandardabweichung – erstes Halbjahr	53 489 EUR

Auf der Grundlage beider Stichproben lässt sich der prognostizierte Fehler auf der Ebene der Grundgesamtheit berechnen als:

$$EE = N_1 \times \frac{\sum_{t=1}^{n_1} F_{1t}}{n_1} + N_2 \times \frac{\sum_{t=1}^{n_2} F_{2t}}{n_2} = 3.852 \times \frac{577.230}{142} + 3.852 \times \frac{588.336}{68}$$

$$= 48.985.884$$

In einem zweiten Schritt lässt sich der korrekte Buchwert (die korrekten Ausgaben, die bei Prüfung aller Vorhaben in der Grundgesamtheit gefunden würden) mit der folgenden Formel hochrechnen:

$$CBV = BV - EE = 6.440.727.190 - 48.985.884 = 6.391.741.306$$

wobei **BV** der jährliche Buchwert (beide Halbjahre) ist und **EE** den genannten prognostizierten Fehler bezeichnet.

Die Genauigkeit (Stichprobenfehler) ist ein Maß der mit der Hochrechnung (Extrapolation) verbundenen Unsicherheit und wird anhand folgender Formel berechnet:

$$SE = z \times \sqrt{\left( N_1^2 \times \frac{s_{E1}^2}{n_1} + N_2^2 \times \frac{s_{E2}^2}{n_2} \right)}$$

$$= 1.645 \times \sqrt{\left( 3852^2 \times \frac{65,815^2}{142} + 3852^2 \times \frac{53,489^2}{68} \right)} = 53,983,055$$

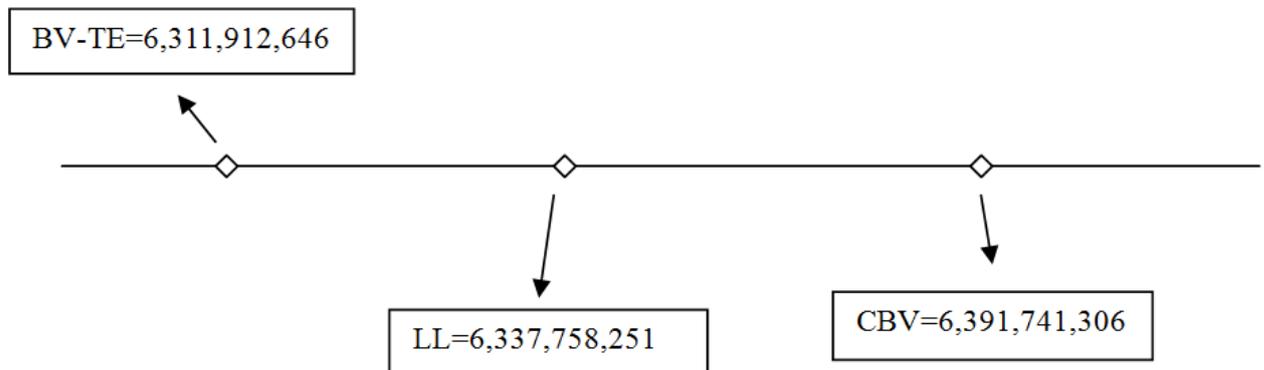
Um Rückschlüsse auf die Erheblichkeit der Fehler zu ermöglichen, wird zunächst die Untergrenze für den korrigierten Buchwert berechnet. Diese Untergrenze entspricht

$$LL = CBV - SE = 6,391,741,306 - 53,983,055 = 6,337,758,251$$

Die Hochrechnungen auf den korrekten Buchwert und die Untergrenze sollten jeweils mit der Differenz zwischen dem Buchwert (geltend gemachte Ausgaben) und dem maximal zulässigen Fehler (**TE**) verglichen werden:

$$BV - TE = 6,440,727,190 - 120,014,544 = 6,311,912,646$$

Da der Wert **BV-TE** unter der Untergrenze **LL = CBV - SE** liegt, kann geschlussfolgert werden, dass hinreichende Beweise dafür vorliegen, dass die Fehler im Programm unter der Erheblichkeitsschwelle liegen.



### 7.3 Wertbezogenes Stichprobenverfahren

#### 7.3.1 Standardansatz

##### 7.3.1.1 Einleitung

Das wertbezogene Stichprobenverfahren (MUS) ist ein statistisches Stichprobenverfahren, bei dem die Geldeinheit als Hilfsvariable für die Probenahme verwendet wird. Dieser Ansatz beruht in der Regel auf einer systematischen Stichprobenziehung, bei der die Wahrscheinlichkeit in einem proportionalen Verhältnis zur Größe steht (PPS = Probability Proportional to Size), d. h. in einem proportionalen

Verhältnis zum Geldwert der Stichprobeneinheit (Elemente mit höheren Beträgen werden mit größerer Wahrscheinlichkeit ausgewählt).

Dabei handelt es sich um das wohl beliebteste Stichprobenverfahren für die Buchprüfung. Besonders nützlich ist es, wenn die Buchwerte eine hohe Streuung aufweisen und eine positive Korrelation (Assoziation) zwischen Fehlern und Buchwerten besteht, also immer wenn zu erwarten ist, dass Elemente mit höheren Werten eher höhere Fehler aufweisen. Eine derartige Situation kommt im Prüfungszusammenhang häufig vor.

Bei Vorliegen der genannten Bedingungen – also Buchwerte mit hoher Variabilität und Fehler positiv korreliert (assoziiert) mit Buchwerten – ergibt das MUS bei gleichem Genauigkeitsgrad eher kleinere Stichprobenumfänge als Stichprobenverfahren mit gleicher Wahrscheinlichkeit.

Außerdem sei darauf hingewiesen, dass bei diesem Stichprobenverfahren normalerweise hochwertige Elemente überrepräsentiert und niedrigwertige Elemente unterrepräsentiert sind. Zwar stellt dies an sich kein Problem dar, da im Extrapolationsprozess ein Ausgleich geschaffen wird, doch werden Stichprobenergebnisse (z. B. die Stichprobenfehlerquote) damit nicht interpretierbar (nur extrapolierte Ergebnisse lassen sich interpretieren).

Wie die Verfahren mit gleicher Wahrscheinlichkeit lässt sich dieses Verfahren mit einer Schichtung kombinieren (günstige Bedingungen für die Schichtung werden in Abschnitt 6.2 erörtert und konkrete Formeln in Abschnitt 7.3.2 vorgestellt).

### 7.3.1.2 Stichprobenumfang

Die Berechnung des Stichprobenumfangs  $n$  innerhalb der Rahmenbedingungen des wertbezogenen Stichprobenverfahrens beruht auf den folgenden Informationen:

- dem Buchwert der Grundgesamtheit (insgesamt geltend gemachte Ausgaben)  $BV$
- dem Konfidenzniveau, das mithilfe von Systemprüfungen und dem damit verwandten Koeffizienten  $z$  aus einer Normalverteilung ermittelt wird (siehe Abschnitt 6.3)
- dem maximal zulässigen Fehler  $TE$  (in der Regel 2 % der gesamten Ausgaben)
- dem voraussichtlichen Fehler  $AE$ , der vom Prüfer anhand seines fachlichen Urteils und vorhandener Informationen ausgewählt wird
- der Standardabweichung  $\sigma_f$  der Fehlerquoten (aus einer MUS-Stichprobe).

Der Stichprobenumfang wird wie folgt berechnet:

$$n = \left( \frac{z \times BV \times \sigma_r}{TE - AE} \right)^2$$

wobei  $\sigma_r$  die Standardabweichung der Fehlerquoten aus einer MUS-Stichprobe bezeichnet. Für eine Näherung an diese Standardabweichung vor Durchführung der Prüfung müssen sich die Mitgliedstaaten entweder auf ihre vorherigen Erkenntnisse (Varianz der Fehlerquoten in einer Stichprobe des vergangenen Zeitraums) oder auf eine Vorabstichprobe/Pilotstichprobe mit geringem Stichprobenumfang  $n^p$  stützen (für die Vorabstichprobe wird ein Stichprobenumfang von nicht weniger als 20 bis 30 Einheiten empfohlen). In jedem Fall wird die Varianz der Fehlerquoten (Quadrat der Standardabweichung) wie folgt berechnet:

$$\sigma_r^2 = \frac{1}{n^p - 1} \sum_{i=1}^{n^p} (r_i - \bar{r})^2;$$

wobei  $r_i = \frac{E_i}{BV_i}$  die Fehlerquote eines Vorhabens ist und als Quotient aus  $E_i$  und dem Buchwert (die gegenüber der Kommission bescheinigten Ausgaben,  $BV_i$ ) des  $i$ -ten Vorhabens in der Stichprobe definiert wird und  $\bar{r}$  die mittlere Fehlerquote in der Stichprobe darstellt, das heißt<sup>18</sup>:

$$\bar{r} = \frac{1}{n^p} \sum_{i=1}^{n^p} \frac{E_i}{BV_i}$$

Beruhet die Standardabweichung wie üblich auf einer Vorabstichprobe, so kann die Stichprobe auch in diesem Fall anschließend als Teil der zur Prüfung ausgewählten Stichprobe verwendet werden. Dennoch ist die Auswahl und Beobachtung einer Vorabstichprobe im Rahmen des wertbezogenen Stichprobenverfahrens weitaus komplizierter als beim einfachen Zufallsstichprobenverfahren oder bei der Differenzschätzung, da für die Stichprobe hochwertige Elemente mit größerer Häufigkeit ausgewählt werden. Daher stellt die Betrachtung einer Stichprobe von 20 bis 30 Vorhaben oft eine schwierige Aufgabe dar. Aus diesem Grunde empfiehlt es sich im Rahmen des MUS nachdrücklich, dass die Schätzung der Standardabweichung  $\sigma_r$  auf vorherigen Daten beruht, weil dann die Notwendigkeit einer Vorabstichprobe entfällt.

---

<sup>18</sup> Ist der Buchwert von Einheit  $i$  ( $BV_i$ ) größer als der Schwellenwert  $BV/n^p$ , dann wird der Quotient  $\frac{E_i}{BV_i}$  durch den Quotienten  $\frac{E_i}{BV/n^p}$  ersetzt.

### 7.3.1.3 Stichprobenauswahl

Nach der Festlegung des Stichprobenumfangs sind die hochwertigen Einheiten der Grundgesamtheit (sofern vorhanden) zu ermitteln, die einer zu 100 % zu prüfenden hochwertigen Schicht angehören werden. Der Schwellenwert für die Bestimmung dieser oberen Schicht entspricht dem Quotienten aus dem Buchwert ( $BV$ ) und dem geplanten Stichprobenumfang ( $n$ ). Alle Elemente, deren Buchwert größer ist als dieser Schwellenwert (wenn  $BV_i > BV/n$ ), werden der Schicht für die 100 %-Prüfung zugeordnet.

Der der nicht umfassenden Schicht zuzuordnende Stichprobenumfang  $n_s$  wird als Differenz zwischen  $n$  und der Anzahl der Stichprobeneinheiten (z. B. Vorhaben) in der umfassenden Schicht ( $n_e$ ) berechnet.

Die Auswahl der Stichprobe in der nicht umfassenden Schicht erfolgt mit einer Wahrscheinlichkeit proportional zur Größe (PPS), d. h. proportional zu den Buchwerten der Elemente  $BV_i$ . Eine gängige Möglichkeit zur Durchführung der Auswahl ist eine systematische Auswahl unter Verwendung eines Stichprobenintervalls, das dem Quotienten aus Gesamtausgaben in der nicht umfassenden Schicht ( $BV_s$ ) und der Stichprobengröße ( $n_s$ ) entspricht, d. h.

$$SI = \frac{BV_s}{n_s}$$

In der Praxis wird die Stichprobe aus einer in zufälliger Reihenfolge angeordneten Liste von Elementen (in der Regel Vorhaben) ausgewählt, wobei jedes Element, welches die  $x$ -te Geldeinheit enthält, ausgewählt wird, und  $x$  dem Stichprobenintervall entspricht und einen willkürlich ausgewählten Anfangspunkt zwischen 1 und  $SI$  aufweist. Wenn beispielsweise eine Grundgesamtheit einen Buchwert von 10 000 000 EUR besitzt und eine Stichprobe von 40 Vorhaben gebildet wird, so wird jedes Vorhaben ausgewählt, das den 250 000-sten Euro enthält.

### 7.3.1.4 Prognostizierter Fehler

Die Hochrechnung der Fehler auf die Grundgesamtheit erfolgt für die Einheiten in der umfassenden Schicht anders als für die Einheiten in der nicht umfassenden Schicht.

Bei der umfassenden Schicht, d. h. der Schicht mit den Stichprobenelementen, deren Buchwert größer ist als der Schwellenbetrag ( $BV_{it} > \frac{BV_t}{n_t}$ ), entspricht der prognostizierte Fehler einfach der Summe der Fehler, die bei den Elementen in dieser Schicht festgestellt werden:

$$EE_e = \sum_{i=1}^{n_s} E_i$$

Bei der nicht umfassenden Schicht, d. h. der Schicht mit den Stichprobenelementen, deren Buchwert kleiner oder gleich dem Schwellenbetrag ist ( $BV_i \leq \frac{BF}{n}$ ), ergibt sich der prognostizierte Fehler wie folgt:

$$EE_s = \frac{BV_s}{n_s} \sum_{i=1}^{n_s} \frac{E_i}{BV_i}$$

Zur Berechnung dieses prognostizierten Fehlers wird folgendermaßen vorgegangen:

- 1) für jede Einheit in der Stichprobe wird die Fehlerquote berechnet, d. h. der Quotient aus dem Fehler und dem jeweiligen Ausgabenbetrag  $\frac{E_i}{BV_i}$  wird gebildet;
- 2) die ermittelten Fehlerquoten werden für alle Einheiten in der Stichprobe aufsummiert;
- 3) das Ergebnis des vorigen Schritts wird mit den Gesamtausgaben in der Grundgesamtheit der nicht umfassenden Schicht ( $BV_s$ ) multipliziert; diese Ausgaben entsprechen auch den Gesamtausgaben in der Grundgesamtheit abzüglich der Ausgaben der zur umfassenden Schicht gehörenden Elemente;
- 4) das Ergebnis des vorigen Schritts wird durch die Stichprobengröße in der nicht umfassenden Schicht ( $n_s$ ) dividiert.

Der prognostizierte Fehler auf der Ebene der Grundgesamtheit ist genau die Summe dieser beiden Komponenten:

$$EE = EE_e + EE_s$$

### 7.3.1.5 Genauigkeit

Die Genauigkeit ist das Maß der mit der Extrapolation verbundenen Unsicherheit. Sie steht für den Stichprobenfehler und ist zu berechnen, um anschließend ein Konfidenzintervall zu erhalten.

Für die Genauigkeit gilt die Formel:

$$SE = z \times \frac{BV_s}{\sqrt{n_s}} \times s_r$$

wobei  $s_r$  die Standardabweichung der Fehlerquoten in der Stichprobe der nicht umfassenden Schicht ist (aus der gleichen Stichprobe berechnet, die auch zur Extrapolation der Fehler auf die Grundgesamtheit verwendet wurde)

$$s_r^2 = \frac{1}{n_g - 1} \sum_{i=1}^{n_g} (r_i - \bar{r}_g)^2$$

Dabei entspricht  $\bar{r}_g$  dem einfachen Durchschnitt der Fehlerquoten in der Stichprobe der Schicht:

$$\bar{r}_g = \frac{\sum_{i=1}^{n_g} \frac{E_i}{BV_i}}{n_g}$$

Zu beachten ist, dass der Stichprobenfehler nur für die nicht umfassende Schicht berechnet wird, da in der umfassenden Schicht kein Stichprobenfehler nachgewiesen werden muss.

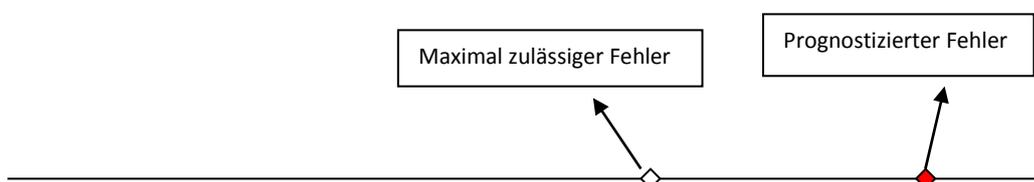
### 7.3.1.6 Bewertung

Um auf die Erheblichkeit der Fehler zu schließen, ist zunächst die obere Fehlergrenze (ULE) zu berechnen. Diese Obergrenze entspricht der Summe aus dem prognostizierten Fehler **EE** selbst und der Genauigkeit der Extrapolation

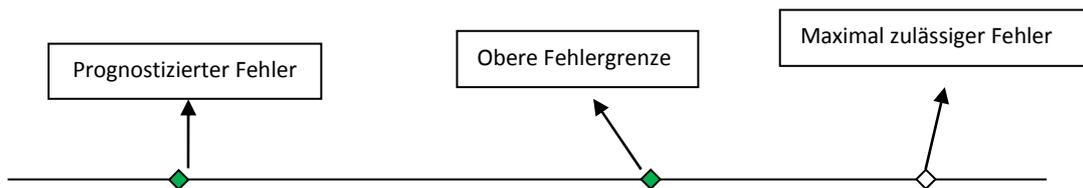
$$ULE = EE + SE$$

Anschließend sollten der prognostizierte Fehler und die Obergrenze jeweils mit dem maximal zulässigen Fehler verglichen werden, um Prüfungsschlussfolgerungen zu ziehen:

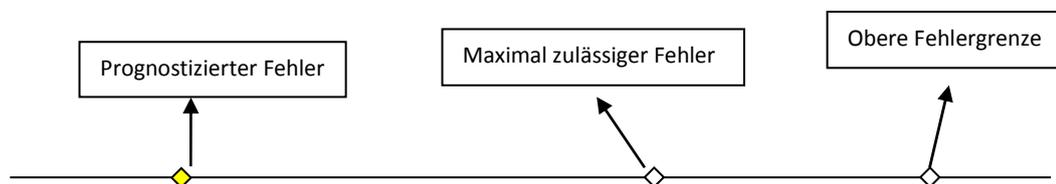
- Liegt der prognostizierte Fehler über dem maximal zulässigen Fehler, würde der Prüfer schlussfolgern, dass hinreichende Beweise dafür vorliegen, dass die Fehler in der Grundgesamtheit über der Erheblichkeitsschwelle liegen.



- Liegt die obere Fehlergrenze unter dem maximal zulässigen Fehler, sollte der Prüfer schlussfolgern, dass die Fehler in der Grundgesamtheit unter der Erheblichkeitsschwelle liegen.



- Liegt der prognostizierte Fehler unter dem maximal zulässigen Fehler, die obere Fehlergrenze jedoch darüber, sind weitere Schritte erforderlich, da nicht hinreichend belegt ist, dass die Grundgesamtheit keine wesentlichen Fehlangaben enthält. Die erforderlichen Schritte werden in Abschnitt 5.11 erörtert.



### 7.3.1.7 Beispiel

Angenommen wird eine Grundgesamtheit von Ausgaben, die gegenüber der Kommission in einem bestimmten Jahr für Vorhaben in einem Programm bescheinigt wurden. Die von der Prüfbehörde durchgeführten Systemprüfungen ergaben ein geringes Sicherheitsniveau. Die Stichprobenerhebung für dieses Programm sollte daher mit einem Konfidenzniveau von 90 % durchgeführt werden.

Die Grundgesamtheit ist in der nachstehenden Tabelle zusammengefasst:

Größe der Grundgesamtheit (Anzahl der Vorhaben)	3852
Buchwert (Summe der Ausgaben im Bezugsjahr)	4 199 882 024 EUR

Der Stichprobenumfang wird wie folgt berechnet:

$$n = \left( \frac{z \times BV \times \sigma_p}{TE - AE} \right)^2$$

wobei  $\sigma_p$  die Standardabweichung der Fehlerquoten aus einer MUS-Stichprobe bezeichnet. Um eine Näherung an diese Standardabweichung zu ermitteln, beschloss die

Prüfungsbehörde, die Standardabweichung des Vorjahres heranzuziehen. Die Stichprobe des Vorjahres bestand aus 50 Vorhaben, von denen fünf einen Buchwert hatten, der über dem Stichprobenintervall lag.

Die folgende Tabelle zeigt für diese fünf Vorhaben die Ergebnisse der Vorjahresprüfung:

Vorhaben-kenn-nummer	Buchwert (BV)	Korrekturer Buchwert (CBV)	Fehler	Fehler-quote
1850	115 382 867 EUR	115 382 867 EUR	- EUR	-
4327	129 228 811 EUR	129 228 811 EUR	- EUR	-
4390	142 151 692 EUR	138 029 293 EUR	4 122 399 EUR	0,0491
1065	93 647 323 EUR	93 647 323 EUR	- EUR	-
1817	103 948 529 EUR	100 830 073 EUR	3 118 456 EUR	0,0371

Hierbei ist besonders zu beachten, dass die Fehlerquote (letzte Spalte) berechnet wird als  $r_i = \frac{E_i}{BV/n}$ , dem Quotienten aus dem Vorhabenfehler und dem BV dividiert durch den anfänglichen Stichprobenumfang, also 50, da diese Vorhaben einen Buchwert aufweisen, der größer ist als das Stichprobenintervall (Einzelheiten dazu siehe Abschnitt 7.3.1.2).

In der folgenden Tabelle sind die Ergebnisse der Vorjahresprüfung für die Stichprobe von 45 Vorhaben mit Buchwert unter dem Schwellenwert zusammengefasst.

	A	B	C	D	E
1	<b>Operation ID</b>	<b>Book Value (BV)</b>	<b>Audit Value (AV)</b>	<b>Error</b>	<b>Error rate</b>
2	239	10,173,875 €	9,962,918 €	210,956 €	0.0207
3	424	23,014,045 €	23,014,045 €	- €	
4	2327	32,886,198 €	32,886,198 €	- €	
5	5009	34,595,201 €	34,595,201 €	- €	
6	1491	78,695,230 €	78,695,230 €	- €	
7	(...)	(...)	(...)	(...)	(...)
39	2596	8,912,999 €	8,909,491 €	3,508 €	0.00039
40	779	26,009,790 €	26,009,790 €	- €	-
41	1250	264,950 €	264,950 €	- €	-
42	3895	30,949,004 €	30,949,004 €	- €	-
43	2011	617,668 €	617,668 €	- €	-
44	4796	335,916 €	335,916 €	- €	-
45	3632	7,971,113 €	7,971,113 €	- €	-
46	2451	17,470,048 €	17,470,048 €	- €	-
47	<b>Sample standard deviation:=STDEV(E2:E46;0;0.0491;0;0.0371)-----&gt;</b>				0.085

Auf der Grundlage dieser Vorabstichprobe beträgt die Standardabweichung der Fehlerquoten  $\sigma_p$  0,085 (in MS Excel berechnet als „=STDEV(E2:E46;0;0;0.0491;0;0.0371)“).

Ausgehend von diesen Schätzwerten für die Standardabweichung der Fehlerquoten, den maximal zulässigen Fehler und den voraussichtlichen Fehler lässt sich der Stichprobenumfang errechnen. Bei einem zulässigen Fehler von 2 % des Gesamtbuchwertes,  $2\% \times 4\,199\,882\,024 = 83\,997\,640$  (in der Verordnung festgelegter Erheblichkeitswert), und einer voraussichtlichen Fehlerquote von 0,4 %,  $0,4\% \times 4\,199\,882\,024 = 16\,799\,528$  (dies entspricht der festen Überzeugung der Prüfbehörde auf der Grundlage der Vorjahresinformationen und der Ergebnisse des Berichts über die Bewertung von Verwaltungs- und Kontrollsystemen), gilt

$$n = \left( \frac{1.645 \times 4,199,882,024 \times 0.085}{83,997,640 - 16,799,528} \right)^2 \approx 77$$

Zunächst einmal sind die hochwertigen Einheiten der Grundgesamtheit (sofern vorhanden) zu ermitteln, die zur hochwertigen Schicht für eine Prüfung zu 100 % gehören werden. Der Schwellenwert für die Bestimmung dieser oberen Schicht entspricht dem Quotienten aus dem Buchwert ( $BV$ ) und der geplanten Stichprobengröße ( $n$ ). Alle Elemente, deren Buchwert größer ist als dieser Schwellenwert (wenn  $BV_i > BV/n$ ), werden der Schicht für die 100 %-Prüfung zugeordnet. In diesem Fall beträgt der Schwellenwert  $4\,199\,882\,024/77=54\,593\,922$  EUR.

In eine isolierte Schicht ordnete die Prüfbehörde alle Vorhaben mit einem Buchwert von über 54 593 922 EUR ein, was acht Vorhaben und 786 837 081 EUR entspricht.

Das Stichprobenintervall für die übrige Grundgesamtheit ist gleich dem Buchwert in der nicht umfassenden Schicht ( $BV_s$ ) (Differenz zwischen dem Gesamtbuchwert und dem Buchwert der acht Vorhaben der oberen Schicht) dividiert durch die Anzahl der auszuwählenden Vorhaben (77 minus die 8 Vorhaben in der oberen Schicht).

$$\text{Sampling interval} = \frac{BV_s}{n_s} = \frac{4,199,882,024 - 786,837,081}{69} = 49,464,419$$

Aus einer in zufälliger Reihenfolge angeordneten Liste von Vorhaben wird die Stichprobe durch Auswahl jedes Elements gezogen, dass die 49 464 419-te Geldeinheit enthält.

Eine Datei mit den übrigen 3844 Vorhaben (3852 minus 8 hochwertige Vorhaben) der Grundgesamtheit wird nach dem Zufallsprinzip sortiert und eine sequenzielle Variable für den kumulierten Buchwert erstellt. Es wird eine Stichprobe von 69 Vorhaben (77 minus 8 hochwertige Vorhaben) nach genau folgendem Verfahren gezogen.

Es wurde ein Zufallswert zwischen 1 und dem Stichprobenintervall 49 464 419 generiert (22 006 651). Die erste Auswahl entspricht dem ersten Vorhaben in der Datei, dessen kumulierter Buchwert größer oder gleich 22 006 651 ist.

Die zweite Auswahl entspricht dem ersten Vorhaben, das die 71 471 070-ste Geldeinheit (~~22,006,651~~ + 49,464,419 = 71,471,070, Ausgangspunkt plus Stichprobenintervall) enthält. Das dritte auszuwählende Vorhaben entspricht dem ersten Vorhaben mit der 120 935 489-ten Geldeinheit (71,471,070 + 49,464,419 = 120,935,489, vorheriger Geldeinheitpunkt plus Stichprobenintervall) usw.

Vorhaben-kenn-nummer	Buchwert (BV)	AcumBV	Stich-probe
239	10 173 875 EUR	10 173 875 EUR	Nein
424	23 014 045 EUR	33 187 920 EUR	Ja
2327	32 886 198 EUR	66 074 118 EUR	Nein
5009	34 595 201 EUR	100 669 319 EUR	Ja
1491	78 695 230 EUR	179 364 549 EUR	Ja
(...)	(...)	(...)	...
2596	8 912 999 EUR	307 654 321 EUR	Nein
779	26 009 790 EUR	333 664 111 EUR	Ja
1250	264 950 EUR	333 929 061 EUR	Nein
3895	30 949 004 EUR	364 878 065 EUR	Nein
2011	617 668 EUR	365 495 733 EUR	Nein
4796	335 916 EUR	365 831 649 EUR	Nein
3632	7 971 113 EUR	373 802 762 EUR	Nein
2451	17 470 048 EUR	391 272 810 EUR	Nein
(...)	(...)	(...)	...

Nach der Prüfung der 77 Vorhaben kann die Prüfbehörde den Fehler hochrechnen.

Von den acht hochwertigen Vorhaben (Gesamtbuchwert 786 837 081 EUR) enthalten drei Vorhaben Fehler, was einem Fehlerbetrag von 7 616 805 EUR entspricht.

Bei der restlichen Stichprobe gestaltet sich die Fehlerbehandlung anders. Für diese Vorhaben wird folgendes Verfahren verwendet:

1) für jede Einheit in der Stichprobe wird die Fehlerquote berechnet, d. h. der Quotient aus dem Fehler und dem jeweiligen Ausgabenbetrag  $\frac{F_i}{EV_i}$  gebildet;

2) diese Fehlerquoten werden für alle Einheiten in der Stichprobe aufsummiert (in MS Excel berechnet als „=SUM(E2:E70)“)

3) das Ergebnis des vorigen Schritts wird mit dem Stichprobenintervall (SI) multipliziert:

$$EE_s = SI \sum_{i=1}^{n_s} \frac{E_i}{BV_i}$$

	A	B	C	D	E
1	<b>Operation ID</b>	<b>Book Value (BV)</b>	<b>Audited Value (AV)</b>	<b>Error</b>	<b>Error rate</b>
2	5002	48,725,645 €	48,725,645 €	- €	-
3	779	26,009,790 €	333,664,111 €	- €	-
4	2073	859,992 €	859,992 €	- €	-
5	239	10,173,875 €	9,962,918 €	210,956 €	0.02
6	989	394,316 €	394,316 €	- €	-
7	65	25,234,699 €	25,125,915 €	108,784 €	0
8	5010	34,595,201 €	34,595,201 €	- €	-
9	(...)	(...)	(...)	(...)	(...)
64	1841	768,278 €	768,278 €	- €	-
65	3672	624,882 €	624,882 €	- €	-
66	2355	343,462 €	301,886 €	41,576 €	0.12
67	959	204,847 €	204,847 €	- €	-
68	608	15,293,716 €	15,293,716 €	- €	-
69	4124	6,773,014 €	6,773,014 €	- €	-
70	262	662 €	662 €	- €	-
71	<b>Total:=SUM(E2:E70)</b> ----->				1.096
72	<b>Sample standard deviation:=STDEV(E2:E70)</b> ----->				0.09

$$EE_s = 49,464,419 \times 1.096 = 54,213,004$$

Der prognostizierte Fehler auf der Ebene der Grundgesamtheit ist genau die Summe dieser beiden Komponenten:

$$EE = 7,616,805 + 54,213,004 = 61,829,809$$

Die prognostizierte Fehlerquote ist der Quotient aus dem prognostizierten Fehler und den Gesamtausgaben:

$$r = \frac{61,829,809}{4,199,882,024} = 1.47 \%$$

Die Standardabweichung der Fehlerquoten in der Stichprobenschicht beträgt 0,09 (in MS Excel berechnet als „=STDEV(E2:E70)“).

Die Genauigkeit errechnet sich folgendermaßen:

$$SE = z \times \frac{BV_s}{\sqrt{n_s}} \times s_r = 1,645 \times \frac{4,199,882,024 - 786,837,081}{\sqrt{69}} \times 0,09 = 60,831,129$$

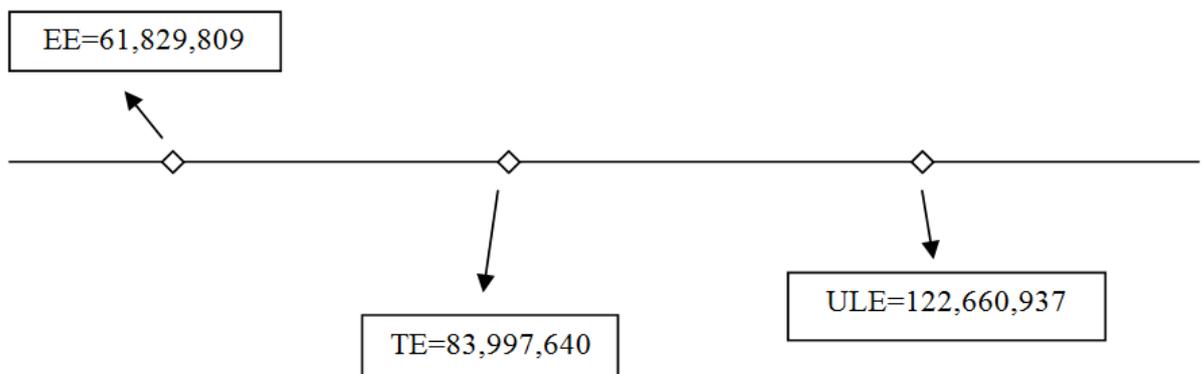
Zu beachten ist, dass der Stichprobenfehler nur für die nicht umfassende Schicht berechnet wird, da in der umfassenden Schicht kein Stichprobenfehler nachgewiesen werden muss.

Um auf die Erheblichkeit der Fehler zu schließen, ist zunächst die obere Fehlergrenze (ULE) zu berechnen. Diese Obergrenze entspricht der Summe aus dem prognostizierten Fehler **EE** selbst und der Genauigkeit der Extrapolation

$$ULE = 61,829,809 + 60,831,129 = 122,660,937$$

Anschließend sollten der prognostizierte Fehler und die Obergrenze jeweils mit dem maximal zulässigen Fehler 83 997 640 EUR verglichen werden, um Prüfungsschlussfolgerungen zu ziehen.

Da der maximal zulässige Fehler über dem prognostizierten Fehler, aber unter der oberen Fehlergrenze liegt, sind weitere Schritte erforderlich, um zu belegen, dass keine wesentlichen Fehlangaben in der Grundgesamtheit enthalten sind. Die erforderlichen Schritte werden in Abschnitt 5.11 erörtert.



### 7.3.2 Geschichtetes wertbezogenes Stichprobenverfahren

#### 7.3.2.1 Einleitung

Beim geschichteten MUS wird die Grundgesamtheit in Teilgesamtheiten („Schichten“) unterteilt. Aus jeder Schicht werden unter Anwendung des MUS-Standardansatzes unabhängige Stichproben gezogen.

Wie üblich ist bei der Auswahl der Schichtungskriterien zu berücksichtigen, dass das Ziel in der Bildung von Gruppen (Schichten) besteht, deren Streuung geringer ist als die der Grundgesamtheit. Alle Variablen, von denen Aufklärung über das Fehlerniveau in den Vorhaben erwartet werden kann, sind somit ebenfalls gut für die Schichtung geeignet. In Frage kommen beispielsweise Programme, Regionen, verantwortliche Stellen oder Klassen, die auf dem Risiko des Vorhabens beruhen. Anders als bei der Stichprobenziehung mit gleicher Wahrscheinlichkeit ist übrigens eine Schichtung nach Ausgabenniveau beim geschichteten MUS nicht von Interesse, da die Ausgabenhöhe hier bereits bei der Auswahl der Stichprobeneinheit berücksichtigt wird.

### 7.3.2.2 Stichprobenumfang

Der Stichprobenumfang wird wie folgt berechnet:

$$n = \left( \frac{z \times BV \times \sigma_{r_{rw}}}{TE - AE} \right)^2$$

Dabei ist  $\sigma_{r_{rw}}^2$  ein gewichtetes Mittel der Varianzen der Fehlerquoten für die Gesamtheit der Schichten; das Gewicht jeder Schicht ist gleich dem Quotienten aus dem Buchwert der Schicht ( $BV_h$ ) und dem Buchwert der Grundgesamtheit ( $BV$ )

$$\sigma_{r_{rw}}^2 = \sum_{h=1}^H \frac{BV_h}{BV} \sigma_{r_{rh}}^2, h = 1, 2, \dots, H;$$

und  $\sigma_{r_{rh}}^2$  ist die Varianz der Fehlerquoten in jeder Schicht. Die Varianz der Fehlerquoten wird für jede Schicht als unabhängige Grundgesamtheit berechnet

$$\sigma_{r_{rh}}^2 = \frac{1}{n_h^p - 1} \sum_{i=1}^{n_h^p} (r_{hi} - \bar{r}_h)^2, h = 1, 2, \dots, H$$

wobei  $r_{hi} = \frac{E_i}{BV_i}$  die einzelnen Fehlerquoten für die Einheiten in der Stichprobe von Schicht  $h$  darstellt und  $\bar{r}_h$  die mittlere Fehlerquote der Stichprobe von Schicht  $h$  angibt.<sup>19</sup>

---

<sup>19</sup> Ist der Buchwert von Einheit  $i$  ( $BV_i$ ) größer als der Schwellenwert  $BV_h/n_h$ , so wird der Quotient  $\frac{E_i}{BV_i}$  durch den Quotienten  $\frac{E_i}{BV_h/n_h}$  ersetzt.

Im Zusammenhang mit dem MUS-Standardansatz wurde bereits erklärt, dass diese Werte auf historischem Wissen oder auf einer Vorabstichprobe/Pilotstichprobe mit geringem Umfang beruhen können. Im letztgenannten Fall kann die Pilotstichprobe in der Folge wie üblich als Teil der zur Prüfung ausgewählten Stichprobe verwendet werden. Auch hier gilt die Empfehlung, diese Parameter anhand vorheriger Daten zu berechnen, weil dann die Notwendigkeit einer Vorabstichprobe entfällt. Wenn die geschichtete MUS erstmals angewandt wird, sind jedoch möglicherweise keine vorherigen geschichteten Daten verfügbar. In diesem Fall kann der Stichprobenumfang unter Zuhilfenahme der Formeln für den MUS-Standardansatz (siehe Abschnitt 7.3.1.2) bestimmt werden. Die fehlenden vorherigen Erkenntnisse schlagen natürlich nachteilig zu Buche: Im ersten Prüfzeitraum ist die Stichprobe zwangsläufig größer, als es sonst notwendig wäre. Andererseits können die Informationen, die im ersten Zeitraum der Anwendung der geschichteten MUS gewonnen werden, in künftigen Zeiträumen zur Ermittlung des Stichprobenumfangs herangezogen werden.

Nach Berechnung des Gesamtstichprobenumfangs  $n$  ist die nach Schicht vorgenommene Aufteilung der Stichprobe wie folgt:

$$n_h = \frac{BV_h}{BV} n.$$

Hierbei handelt es sich um eine allgemeine Aufteilungsmethode, bei der die Aufteilung der Stichprobe auf die Schichten proportional zu den Ausgaben (dem Buchwert) der Schichten erfolgt. Es stehen auch andere Methoden zur Verfügung. Mit einer spezifischeren Aufteilung lassen sich in einigen Fällen ein zusätzlicher Genauigkeitsgewinn bzw. eine Verringerung des Stichprobenumfangs erreichen. Entscheidungen hinsichtlich der Eignung anderer Aufteilungsmethoden für einzelne spezifische Grundgesamtheiten erfordern ein gewisses Maß an stichprobentheoretischen Kenntnissen.

### 7.3.2.3 Stichprobenauswahl

In jeder Schicht  $h$  wird es zwei Komponenten geben: die umfassende Gruppe innerhalb der Schicht  $h$  (d. h. die Gruppe mit den Stichprobeneinheiten, deren Buchwert größer ist als der Schwellenwert,  $BV_{ht} > \frac{BV_h}{n_h}$ ) und die Stichprobengruppe innerhalb der Schicht  $h$  (d. h. die Gruppe mit den Stichprobeneinheiten, deren Buchwert kleiner oder gleich dem Schwellenwert ( $BV_{ht} \leq \frac{BV_h}{n_h}$ ) ist).

Nach der Festlegung des Stichprobenumfangs sind in jeder der ursprünglichen Schichten ( $h$ ) die hochwertigen Einheiten der Grundgesamtheit (sofern vorhanden) zu ermitteln, die einer zu 100 % zu prüfenden hochwertigen Schicht angehören werden.

Der Schwellenwert für die Bestimmung dieser oberen Gruppe entspricht dem Quotienten aus dem Buchwert der Schicht ( $BV_h$ ) und dem geplanten Stichprobenumfang ( $n_h$ ). Alle Elemente, deren Buchwert größer ist als dieser Schwellenwert (wenn  $BV_{ht} > \frac{BV_h}{n_h}$ ), werden der Stichprobenschicht der 100%-Prüfung zugeordnet.

Der der nicht umfassenden Schicht zuzuordnende Stichprobenumfang  $n_{hs}$  wird als Differenz zwischen  $n_h$  und der Anzahl der Stichprobeneinheiten (z. B. Vorhaben) in der umfassenden Gruppe der Schicht ( $n_{hs}$ ) berechnet.

Die Auswahl der Stichproben schließlich erfolgt in der nicht umfassenden Gruppe einer jeden Schicht mit einer Wahrscheinlichkeit proportional zur Größe, d. h. proportional zu den Buchwerten der Elemente  $BV_i$ . Eine gängige Möglichkeit zur Durchführung der Auswahl ist eine systematische Auswahl unter Verwendung eines Auswahlintervalls, das dem Quotienten aus den Gesamtausgaben in der nicht umfassenden Gruppe der ( $BV_{hs}$ ) und dem Stichprobenumfang ( $n_{hs}$ ) entspricht, d. h.

$$SI_h = \frac{BV_{hs}}{n_{hs}}$$

Es ist zu beachten, dass mehrere unabhängige Stichproben ausgewählt werden, nämlich eine für jede ursprüngliche Schicht.

#### 7.3.2.4 Prognostizierter Fehler

Die Hochrechnung der Fehler auf die Grundgesamtheit erfolgt für die Einheiten in den umfassenden Gruppen anders als für die Elemente in den nicht umfassenden Gruppen.

Bei den umfassenden Gruppen, d. h. den Gruppen mit den Stichprobeneinheiten, deren Buchwert größer ist als der Schwellenwert ( $BV_{ht} > \frac{BV_h}{n_h}$ ), entspricht der prognostizierte Fehler der Summe der Fehler, die bei den Elementen in dieser Gruppe festgestellt werden:

$$EE_g = \sum_{h=1}^H \sum_{i=1}^{n_h} E_{ht}$$

In der Praxis bedeutet dies:

- 1) Für jede Schicht  $h$  werden die zur umfassenden Gruppe gehörenden Einheiten ermittelt und ihre Fehler summiert,
- 2) die Ergebnisse des vorigen Schritts werden für die Gesamtheit der  $H$  Schichten addiert.

Bei den nicht umfassenden Gruppen, d. h. den Gruppen mit den Stichprobeneinheiten, deren Buchwert kleiner oder gleich dem Schwellenwert ( $BV_{ht} \leq \frac{BV_h}{n_h}$ ) ist, errechnet sich der prognostizierte Fehler anhand folgender Formel

$$EE_g = \sum_{h=1}^H \frac{BV_{sh}}{n_{sh}} \sum_{i=1}^{n_{sh}} \frac{E_{ht}}{BV_{ht}}$$

Zur Berechnung dieses prognostizierten Fehlers wird folgendermaßen vorgegangen:

- 1) in jeder Schicht  $h$  wird für jede Einheit in der Stichprobe die Fehlerquote berechnet, d. h. der Quotient aus dem Fehler und den jeweiligen Ausgaben  $\frac{E_{ht}}{BV_{ht}}$ ,
- 2) in jeder Schicht  $h$  werden die ermittelten Fehlerquoten für alle Einheiten in der Stichprobe aufsummiert,
- 3) in jeder Schicht  $h$  wird das Ergebnis des vorigen Schritts mit den Gesamtausgaben in der Grundgesamtheit der nicht umfassenden Gruppe ( $BV_{sh}$ ) multipliziert; diese Ausgaben entsprechen auch den Gesamtausgaben in der Schicht abzüglich der Ausgaben der zur umfassenden Gruppe gehörenden Elemente,
- 4) in jeder Schicht  $h$  wird das Ergebnis des vorigen Schritts durch den Stichprobenumfang in der nicht umfassenden Gruppe ( $n_{sh}$ ) dividiert,
- 5) die vorherigen Ergebnisse werden für die Gesamtheit der  $H$  Schichten addiert.

Der prognostizierte Fehler auf der Ebene der Grundgesamtheit ist genau die Summe dieser beiden Komponenten:

$$EE = EE_g + EE_s$$

### 7.3.2.5 Genauigkeit

Wie beim MUS-Standardansatz, ist die Genauigkeit das Maß der mit der Extrapolation verbundenen Unsicherheit. Sie steht für den Stichprobenfehler und ist zu berechnen, um anschließend ein Konfidenzintervall zu erhalten.

Für die Genauigkeit gilt die Formel:

$$SE = z \times \sqrt{\sum_{h=1}^H \frac{BV_{sh}^2}{n_{sh}} \cdot s_{rsh}^2}$$

wobei  $s_{rsh}$  die Standardabweichung der Fehlerquoten in der Stichprobe der nicht umfassenden Gruppe von Schicht  $h$  ist (berechnet aus der gleichen Stichprobe, die auch zur Extrapolation der Fehler auf die Grundgesamtheit verwendet wurde)

$$s_{r_{sh}}^2 = \frac{1}{n_{sh} - 1} \sum_{i=1}^{n_{sh}} (r_{hi} - \bar{r}_{sh})^2, h = 1, 2, \dots, H$$

Dabei entspricht  $\bar{r}_{sh}$  dem einfachen Durchschnitt der Fehlerquoten in der Stichprobe der nichtumfassenden Gruppe von Schicht  $h$ .

Der Stichprobenfehler wird nur für die nicht umfassenden Gruppen berechnet, da sich aus den umfassenden Gruppen kein Stichprobenfehler ergibt.

### 7.3.2.6 Bewertung

Um auf die Erheblichkeit der Fehler zu schließen, ist zunächst die obere Fehlergrenze (ULE) zu berechnen. Diese Obergrenze entspricht der Summe aus dem prognostizierten Fehler **EE** selbst und der Genauigkeit der Extrapolation

$$ULE = EE + SE$$

Anschließend sollten der prognostizierte Fehler und die Obergrenze jeweils mit dem maximal zulässigen Fehler verglichen werden, um Prüfungsschlussfolgerungen unter Verwendung des Ansatzes zu ziehen, der in Abschnitt 7.3.1.6 dargelegt wurde.

### 7.3.2.7 Beispiel

Es wird ausgegangen von einer Grundgesamtheit von Ausgaben, die gegenüber der Kommission in einem bestimmten Jahr für Vorhaben in einer Gruppe von zwei Programmen bescheinigt wurden. Die von der Prüfbehörde durchgeführten Systemprüfungen ergaben ein geringes Sicherheitsniveau. Die Stichprobenerhebung für dieses Programm sollte daher mit einem Konfidenzniveau von 90 % durchgeführt werden.

Die Prüfbehörde hat Grund zu der Annahme, dass in den Programmen mit unterschiedlichen Fehlerquoten zu rechnen ist. Unter Berücksichtigung all dieser Aspekte entschied sich die Prüfbehörde für eine Schichtung der Grundgesamtheit nach Programm.

In der nachstehenden Tabelle sind die vorhandenen Informationen zusammengefasst:

Größe der Grundgesamtheit (Anzahl der Vorhaben)	6252
Größe der Grundgesamtheit – Schicht 1	4520
Größe der Grundgesamtheit – Schicht 2	1732
Buchwert (Summe der Ausgaben im Bezugsjahr)	4 199 882 024 EUR
Buchwert – Schicht 1	2 506 626 292 EUR
Buchwert – Schicht 2	1 693 255 732 EUR

Der erste Schritt besteht in der Berechnung des erforderlichen Stichprobenumfangs unter Verwendung der folgenden Formel:

$$n = \left( \frac{z \times BV \times \sigma_{r_{rw}}}{TE - AE} \right)^2$$

Dabei ist  $\sigma_{r_{rw}}^2$  ein gewichtetes Mittel der Varianzen der Fehlerquoten für die Gesamtheit der Schichten; das Gewicht jeder Schicht ist gleich dem Quotienten aus dem Buchwert der Schicht ( $BV_h$ ) und dem Buchwert der Grundgesamtheit ( $BV$ ):

$$\sigma_{r_{rw}}^2 = \sum_{h=1}^H \frac{BV_h}{BV} \sigma_{r_{rh}}^2, h = 1, 2, \dots, H;$$

wobei  $\sigma_{r_{rh}}$  die Standardabweichung der Fehlerquoten aus einer MUS-Stichprobe bezeichnet. Um eine Näherung an diese Standardabweichung zu ermitteln, beschloss die Prüfungsbehörde, die Standardabweichung des Vorjahres heranzuziehen. Die Stichprobe des Vorjahres bestand aus 110 Vorhaben, 70 Vorhaben aus dem ersten Programm (Schicht) und 40 aus dem zweiten Programm.

Auf der Grundlage dieser Vorjahresstichprobe erfolgt die Berechnung der Varianz der Fehlerquoten wie folgt (Einzelheiten siehe Abschnitt 7.3.1.7):

$$\sigma_{r_{r1}}^2 = \frac{1}{70 - 1} \sum_{i=1}^{70} (r_{1i} - \bar{r}_{r1})^2 = 0.000045$$

und

$$\sigma_{r_{r2}}^2 = \frac{1}{40 - 1} \sum_{i=1}^{40} (r_{2i} - \bar{r}_{r2})^2 = 0.010909$$

Das führt zu folgendem Ergebnis:

$$\sigma_{rvr}^2 = \frac{2,506,626,292}{4,199,882,024} \times 0,000045 + \frac{1,693,255,732}{4,199,882,024} \times 0,010909 = 0,004425$$

Ausgehend von diesem Schätzwert für die Varianz der Fehlerquoten lässt sich der Stichprobenumfang errechnen. Wie bereits festgestellt, rechnet die Prüfbehörde mit erheblichen Unterschieden zwischen beiden Schichten. Ausgehend vom Bericht zur Funktionsweise des Verwaltungs- und Kontrollsystems erwartet sie eine Fehlerquote von rund 1,1 %. Bei einem angenommenen zulässigen Fehler von 2 % des Gesamtbuchwertes (in der Verordnung festgelegter Höchstwert für die Erheblichkeitsschwelle), d. h.  $TE=2\% \times 4\,199\,882\,024=83\,997\,640$ , und dem voraussichtlichen Fehler, d. h.  $AE=1,1\% \times 4\,199\,882\,024=46\,198\,702$ , beträgt der Stichprobenumfang

$$n = \left( \frac{1,645 \times 4,199,882,024 \times \sqrt{0,004425}}{83,997,640 - 46,198,702} \right)^2 \approx 148$$

Die Aufteilung der Stichprobe nach Schicht ist wie folgt:

$$n_1 = \frac{BV_1}{BV} \times n = \frac{2,506,626,292}{4,199,882,024} \times 148 = 89$$

$$n_2 = n - n_1 = 148 - 89 = 59.$$

Diese beiden Stichprobenumfänge haben folgende Schwellenwerte für hochwertige Schichten zur Folge:

$$\text{Cut-off}_1 = \frac{BV_1}{n_1} = \frac{2,506,626,292}{89} = 28,164,340$$

und

$$\text{Cut-off}_2 = \frac{BV_2}{n_2} = \frac{1,693,255,731}{59} = 28,699,250$$

Bei Verwendung dieser beiden Schwellenwerte befinden sich in der Schicht 1 16 und in der Schicht 2 12 hochwertige Vorhaben.

Der Stichprobenumfang für den in die Stichprobe einzubeziehenden Teil von Schicht 1 ergibt sich aus dem Gesamtstichprobenumfang (89), von dem die 16 hochwertigen Vorhaben abgezogen werden. Sie umfasst also 73 Vorhaben. Bei gleichem Vorgehen ergibt sich für Schicht 2 ein Stichprobenumfang für den einzubeziehenden Teil von  $59 - 12 = 47$  Vorhaben.

Der nächste Schritt ist die Berechnung der Stichprobenintervalle für die Stichprobenschichten. Sie betragen

$$SI_1 = \frac{BV_{s1}}{n_{s1}} = \frac{1,643,963,924}{73} = 22,520,054$$

bzw.

$$SI_2 = \frac{BV_{s2}}{n_{s2}} = \frac{1,059,467,667}{47} = 22,541,865$$

In der nachstehenden Tabelle werden die bisherigen Ergebnisse zusammengefasst:

Größe der Grundgesamtheit (Anzahl der Vorhaben)	6252
Größe der Grundgesamtheit – Schicht 1	4520
Größe der Grundgesamtheit – Schicht 2	1732
Buchwert (Summe der Ausgaben im Bezugsjahr)	4 199 882 024 EUR
Buchwert – Schicht 1	2 506 626 292 EUR
Buchwert – Schicht 2	1 693 255 731 EUR
<b>Stichprobenergebnisse – Schicht 1</b>	
Schwellenwert	28 164 340 EUR
Anzahl der Vorhaben über dem Schwellenwert	16
Buchwert der Vorhaben über dem Schwellenwert	862 662 369 EUR
Buchwert der Vorhaben (nicht umfassende Grundgesamtheit)	1 643 963 923 EUR
Stichprobenintervall (nicht umfassende Grundgesamtheit)	22 520 054 EUR
Anzahl der Vorhaben (nicht umfassende Grundgesamtheit)	4504
<b>Stichprobenergebnisse – Schicht 2</b>	
Schwellenwert	28 699 250 EUR
Anzahl der Vorhaben über dem Schwellenwert	12
Buchwert der Vorhaben über dem Schwellenwert	633 788 064 EUR
Buchwert der Vorhaben (nicht umfassende Grundgesamtheit)	1 059 467 668 EUR
Stichprobenintervall (nicht umfassende Grundgesamtheit)	22 541 865 EUR
Anzahl der Vorhaben (nicht umfassende Grundgesamtheit)	1720

Für Schicht 1 wird eine Datei mit den restlichen 4504 Vorhaben (4520 minus 16 hochwertige Vorhaben) der Grundgesamtheit nach dem Zufallsprinzip sortiert und eine sequenzielle Variable für den kumulierten Buchwert erstellt. Es wird eine Stichprobe

von 73 Vorhaben (89 minus 16 hochwertige Vorhaben) nach genau dem Verfahren gezogen, das in Abschnitt 7.3.1.7 beschrieben wird.

Für Schicht 2 wird eine Datei mit den restlichen 1720 Vorhaben (1732 minus 12 hochwertige Vorhaben) der Grundgesamtheit nach dem Zufallsprinzip sortiert und eine sequenzielle Variable für den kumulierten Buchwert erstellt. Es wird eine Stichprobe von 47 Vorhaben (59 minus 12 hochwertige Vorhaben) nach dem im vorangegangenen Abschnitt beschriebenen Verfahren gezogen.

In Schicht 1 wurden bei den 16 hochwertigen Vorhaben keine Fehler entdeckt.

In Schicht 2 wurden bei 6 der 12 hochwertigen Vorhaben Fehler in Höhe von 15 460 340 EUR festgestellt.

Bei den restlichen Stichproben gestaltet sich die Fehlerbehandlung anders. Für diese Vorhaben wird folgendes Verfahren verwendet:

- 1) für jede Einheit in der Stichprobe wird die Fehlerquote berechnet, d. h. der Quotient aus dem Fehler und dem jeweiligen Ausgabenbetrag  $\frac{E_i}{BV_i}$  gebildet;
- 2) die ermittelten Fehlerquoten werden für alle Einheiten in der Stichprobe aufsummiert;
- 3) das Ergebnis des vorigen Schritts wird mit dem Stichprobenintervall (SI) multipliziert

$$EE_{ns} = SI_{ns} \sum_{i=1}^{n_{st}} \frac{E_{st}}{BV_{st}}$$

Die Summe der Fehlerquoten für die nicht umfassende Grundgesamtheit in Schicht 1 beträgt 1,0234,

$$EE_{1s} = 22,520,054 \times 1,0234 = 23,047,023$$

und in Schicht 2 beträgt sie 1,176,

$$EE_{2s} = 22,541,865 \times 1,176 = 26,509,234.$$

Der prognostizierte Fehler auf der Ebene der Grundgesamtheit ist einfach die Summe aller Komponenten, d. h. des im umfassenden Teil beider Schichten gefundenen Fehlbetrags (15 460 340 EUR) und des prognostizierten Fehlers für beide Schichten:

$$EE = 15,460,340 + 23,047,023 + 26,509,234 = 65,016,597.$$

Dies entspricht einer prognostizierten Fehlerquote von 1,55 %.

Zur Berechnung der Genauigkeit müssen unter Verwendung des in Abschnitt 7.3.1.7 beschriebenen Verfahrens die Varianzen der Fehlerquoten für beide Stichprobenschichten ermittelt werden:

$$s_{1r}^2 = \frac{1}{72-1} \sum_{i=1}^{72} (r_{1i} - \bar{r}_{1r})^2 = 0.000036$$

und

$$s_{2r}^2 = \frac{1}{48-1} \sum_{i=1}^{48} (r_{2i} - \bar{r}_{2r})^2 = 0.0081$$

Die Genauigkeit errechnet sich folgendermaßen:

$$SE = z \times \sqrt{\sum_{h=1}^H \frac{BV_{sh}^2}{n_{sh}} \times s_{rsh}^2}$$

$$SE = 1.645 \times \sqrt{\frac{1,643,963,923^2}{73} \times 0.000036 + \frac{1,059,467,667^2}{47} \times 0.0081} = 22,958,216$$

Zu beachten ist, dass der Stichprobenfehler nur für die nicht umfassenden Teile der Grundgesamtheit berechnet wird, da in der umfassenden Schicht kein Stichprobenfehler nachgewiesen werden muss.

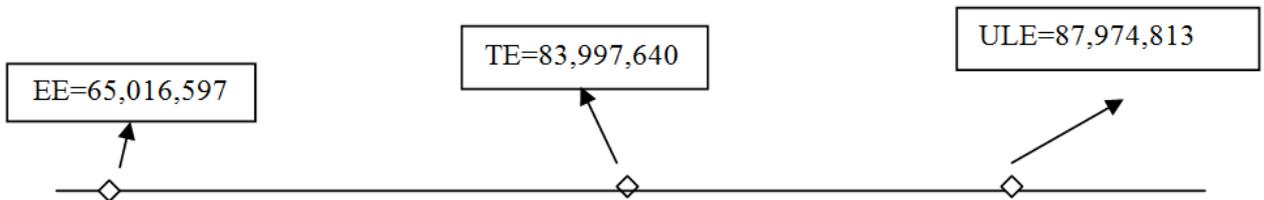
Um auf die Erheblichkeit der Fehler zu schließen, ist zunächst die obere Fehlergrenze (ULE) zu berechnen. Diese Obergrenze entspricht der Summe aus dem prognostizierten Fehler **EE** selbst und der Genauigkeit der Extrapolation

$$ULE = 65,016,597 + 22,958,216 = 87,974,813$$

Anschließend sollten der prognostizierte Fehler und die Obergrenze jeweils mit dem maximal zulässigen Fehler verglichen werden, um Prüfungsschlussfolgerungen zu ziehen:

Beim Vergleich der prognostizierten Ergebnisse mit der Erheblichkeitsschwelle von 2 % des gesamten Buchwerts der Grundgesamtheit (2 % x 4 199 882 024 EUR = 83 997 640 EUR) wird festgestellt, dass der maximal zulässige Fehler größer ist als der prognostizierte Fehler, jedoch unterhalb der Obergrenze liegt. Daher sind zusätzliche

Arbeitsschritte (wie in Abschnitt 5.11 beschrieben) erforderlich, denn es ist nicht hinreichend belegt, dass die Grundgesamtheit keine wesentlichen Fehlangaben enthält.



### 7.3.3 Wertbezogenes Stichprobenverfahren – zwei Zeiträume

#### 7.3.3.1 Einleitung

Die Prüfbehörde kann das Stichprobenverfahren in mehreren Zeiträumen während des Jahres (üblicherweise zwei Halbjahre) durchführen. Wie bei allen anderen Stichprobenverfahren steht der wichtigste Vorteil dieses Ansatzes nicht im Zusammenhang mit einer Reduzierung des Stichprobenumfangs, sondern mit der Möglichkeit, den Prüfungsaufwand über das Jahr zu verteilen und somit den Aufwand zu verringern, der bei nur einer Beobachtung am Ende des Jahres entstünde.

Mit diesem Ansatz wird die Jahresgrundgesamtheit in zwei Teilgesamtheiten aufgeteilt, die jeweils den Vorhaben und Ausgaben eines Halbjahres entsprechen. Für jedes Halbjahr werden unabhängige Stichproben unter Anwendung des MUS-Standardansatzes gezogen.

#### 7.3.3.2 Stichprobenumfang

##### Erstes Halbjahr

Im ersten Prüfzeitraum (z. B. Halbjahr) wird der Gesamtstichprobenumfang (für beide Semester) wie folgt errechnet:

$$n = \left( \frac{z \times BV \times \sigma_{rw}}{TE - AE} \right)^2$$

Dabei ist  $\sigma_{rw}^2$  ein gewichtetes Mittel der Varianzen der Fehlerquoten im jeweiligen Halbjahr; das Gewicht jedes Halbjahres entspricht dem Quotienten aus dem Buchwert des Halbjahres ( $BV_1$ ) und dem Buchwert der Grundgesamtheit ( $BV$ )

$$\sigma_w^2 = \frac{BV_1}{BV} \sigma_{r1}^2 + \frac{BV_2}{BV} \sigma_{r2}^2$$

und  $\sigma_{r_{it}}^2$  ist die Varianz der Fehlerquoten in jedem Halbjahr. Die Varianz der Fehlerquoten wird für jedes Halbjahr berechnet als

$$\sigma_{r_{it}}^2 = \frac{1}{n_t^p - 1} \sum_{i=1}^{n_t^p} (r_{it} - \bar{r}_t)^2, t = 1, 2$$

Dabei sind  $r_{it} = \frac{E_{it}}{BV_{1t}}$  die einzelnen Fehlerquoten der Einheiten in der Stichprobe des Halbjahres t und  $\bar{r}_t$  der mittlere Fehler der Stichprobe im Halbjahr t<sup>20</sup>.

Die Festlegung der Werte für die erwarteten Standardabweichungen der Fehlerquoten in beiden Halbjahren muss nach fachlichem Urteil und unter Verwendung von bereits vorhandenem Wissen erfolgen. Die Option einer Vorabstichprobe/Pilotstichprobe mit geringem Stichprobenumfang, wie sie bereits im Zusammenhang mit dem Standard-MUS dargelegt wurde, ist nach wie vor gegeben, lässt sich jedoch nur für das erste Halbjahr realisieren. Die Ausgaben für das zweite Halbjahr sind zum Zeitpunkt der ersten Beobachtung noch nicht getätigt, und es sind keine objektiven Daten (außer den bereits vorhandenen) verfügbar. Falls Pilotstichproben gezogen werden, können sie danach wie üblich als Teil der zur Prüfung ausgewählten Stichprobe verwendet werden.

Stehen zur Einschätzung der Datenstreuung im zweiten Halbjahr keine vorherigen Daten oder Kenntnisse zur Verfügung, so kann ein vereinfachter Ansatz genutzt werden, bei dem der Gesamtstichprobenumfang wie folgt errechnet wird:

$$n = \left( \frac{z \times BV \times \sigma_{r1}}{TE - AE} \right)^2$$

Bei diesem vereinfachten Ansatz werden übrigens nur Informationen zur Streuung der Fehlerquoten im ersten Beobachtungszeitraum benötigt. Es wird davon ausgegangen, dass die Streuung der Fehlerquoten in beiden Halbjahren annähernd gleich groß ist.

Probleme wegen fehlender zusätzlicher vorhandener Informationen treten in der Regel nur im ersten Jahr des Programmzeitraums auf. Die im ersten Prüfljahr gewonnenen Informationen können dann künftig für die Ermittlung des Stichprobenumfangs verwendet werden.

Außerdem ist zu beachten, dass die Formeln zur Berechnung des Stichprobenumfangs Werte für  $BV_1$  und  $BV_2$  erfordern, d. h. für den Gesamtbuchwert (geltend gemachte Ausgaben) des ersten und des zweiten Halbjahres. Bei der Berechnung des Stichprobenumfangs wird der Wert für  $BV_1$  bekannt sein, der Wert für  $BV_2$  jedoch nicht.

<sup>20</sup> Ist der Buchwert von Einheit i ( $BV_i$ ) größer als  $BV_t/n_t$ , so wird der Quotient  $\frac{E_{it}}{BV_t}$  durch den Quotienten  $\frac{E_{it}}{BV_i/n_t}$  ersetzt.

Er muss entsprechend den Erwartungen des Prüfers (und auch auf der Grundlage von historischen Informationen) bestimmt werden.

Nach Berechnung des Gesamtstichprobenumfangs  $n$  ist die Aufteilung der Stichprobe nach Halbjahr wie folgt:

$$n_1 = \frac{BV_1}{BV} n$$

und

$$n_2 = \frac{BV_2}{BV} n$$

### Zweites Halbjahr

Im ersten Beobachtungszeitraum wurden Annahmen getroffen, die sich auf die nachfolgenden Beobachtungszeiträume (normalerweise das nächste Halbjahr) beziehen. Weichen die Merkmale der Grundgesamtheit in den nachfolgenden Zeiträumen erheblich von den Annahmen ab, so muss der Stichprobenumfang für den nachfolgenden Zeitraum möglicherweise berichtigt werden.

Im zweiten Prüfzeitraum (z. B. Halbjahr) stehen mehr Informationen zur Verfügung:

- Der Gesamtbuchwert im zweiten Halbjahr  $BV_2$  ist genau bekannt.
- Die aus der Stichprobe des ersten Halbjahres errechnete Stichprobenstandardabweichung der Fehlerquoten  $s_{r1}$  ist bereits vorhanden.
- Die Standardabweichung der Fehlerquoten für das zweite Halbjahr  $\sigma_{r2}$  lässt sich anhand realer Daten nunmehr genauer bewerten.

Unterscheiden sich diese Parameter nicht wesentlich von denen, die im ersten Halbjahr entsprechend den Erwartungen des Prüfers geschätzt wurden, so braucht der ursprünglich für das zweite Halbjahr geplante Stichprobenumfang ( $n_2$ ) nicht korrigiert zu werden. Sieht der Prüfer jedoch große Unterschiede zwischen den ursprünglichen Annahmen und den tatsächlichen Merkmalen der Grundgesamtheit, so können Korrekturen des Stichprobenumfangs erforderlich sein, um den ungenauen Schätzungen Rechnung zu tragen. In diesem Fall ist der Stichprobenumfang des zweiten Halbjahres wie folgt neu zu berechnen:

$$n_2 = \frac{(z \times BV_2 \times \sigma_{r2})^2}{(TE - AE)^2 - z^2 \times \frac{BV_1^2}{n_1} \times s_{r1}^2}$$

wobei  $s_{t-1}$  die aus der Stichprobe des ersten Halbjahres errechnete Standardabweichung der Fehlerquoten ist und  $\sigma_{t2}$  eine Schätzung der Standardabweichung der Fehlerquoten im zweiten Halbjahr auf der Grundlage vorheriger Erkenntnisse (möglicherweise angepasst an Informationen aus dem ersten Halbjahr) oder einer Vorabstichprobe/Pilotstichprobe des zweiten Halbjahres darstellt.

### 7.3.3.3 Stichprobenauswahl

Die Auswahl der Stichprobe erfolgt in jedem Halbjahr genau nach dem Verfahren, das für den MUS-Standardansatz beschrieben wurde. Es wird an dieser Stelle zur besseren Verdeutlichung noch einmal dargelegt.

Für jedes Halbjahr sind nach der Festlegung des Stichprobenumfangs die hochwertigen Einheiten der Grundgesamtheit (sofern vorhanden) zu ermitteln, die einer zu 100 % zu prüfenden hochwertigen Gruppe angehören werden. Der Schwellenwert zur Bestimmung dieser oberen Gruppe entspricht dem Quotienten aus dem Buchwert des Halbjahres ( $BV_t$ ) und dem geplanten Stichprobenumfang ( $n_t$ ). Alle Elemente, deren Buchwert höher ist als dieser Schwellenwert (wenn  $BV_{it} > \frac{BV_t}{n_t}$ ), werden der Gruppe der 100 %-Prüfung zugeordnet.

Der der nicht umfassenden Gruppe zuzuordnende Stichprobenumfang  $n_{ts}$  wird als Differenz zwischen  $n_t$  und der Anzahl der Stichprobeneinheiten (z. B. Vorhaben) in der umfassenden Gruppe ( $n_{ts}$ ) berechnet.

Die Auswahl der Stichproben schließlich erfolgt in jedem Halbjahr in der nicht umfassenden Gruppe mit einer Wahrscheinlichkeit proportional zur Größe, d. h. proportional zu den Buchwerten der Elemente  $BV_{it}$ . Eine gängige Möglichkeit zur Durchführung der Auswahl ist eine systematische Auswahl unter Verwendung eines Auswahlintervalls, das dem Quotienten aus Gesamtausgaben in der nicht umfassenden Gruppe ( $BV_{ts}$ ) und dem Stichprobenumfang ( $n_{ts}$ ) entspricht, d. h.

$$SI_n = \frac{BV_{ts}}{n_{ts}}$$

### 7.3.3.4 Prognostizierter Fehler

Die Hochrechnung der Fehler auf die Grundgesamtheit erfolgt für die Einheiten in den umfassenden Gruppen anders als für die Elemente in den nicht umfassenden Gruppen.

Bei den umfassenden Gruppen, d. h. den Gruppen mit den Stichprobeneinheiten, deren Buchwert größer ist als der Schwellenbetrag ( $BV_{ct} > \frac{BV_t}{n_t}$ ), ist der prognostizierte Fehler die Summe der Fehler, die bei den Elementen in dieser Gruppe festgestellt werden:

$$EE_e = \sum_{t=1}^{n_1} E_{1t} + \sum_{t=1}^{n_2} E_{2t}$$

In der Praxis bedeutet dies:

- 1) Für jedes Halbjahr  $t$  werden die zur umfassenden Gruppe gehörenden Einheiten ermittelt und ihre Fehler aufsummiert.
- 2) Die Ergebnisse des vorherigen Schritts werden für beide Halbjahre aufsummiert.

Bei den nicht umfassenden Gruppen, d. h. den Gruppen mit den Stichprobeneinheiten, deren Buchwert kleiner oder gleich dem Schwellenwert ( $BV_{ct} \leq \frac{BV_t}{n_t}$ ) ist, errechnet sich der prognostizierte Fehler anhand folgender Formel

$$EE_s = \frac{BV_{s1}}{n_{s1}} \times \sum_{t=1}^{n_{s1}} E_{1t} + \frac{BV_{s2}}{n_{s2}} \times \sum_{t=1}^{n_{s2}} E_{2t}$$

Zur Berechnung dieses prognostizierten Fehlers wird folgendermaßen vorgegangen:

- 1) In jedem Halbjahr  $t$  wird für jede Einheit in der Stichprobe die Fehlerquote berechnet, d. h. der Quotient aus dem Fehler und dem jeweiligen Ausgabenbetrag  $\frac{E_{ct}}{BV_{ct}}$ .
- 2) In jedem Halbjahr  $t$  werden diese Fehlerquoten für alle Einheiten in der Stichprobe aufsummiert.
- 3) In jedem Halbjahr  $t$  wird das Ergebnis des vorigen Schritts mit den Gesamtausgaben in der Grundgesamtheit der nicht umfassenden Gruppe ( $BV_{st}$ ) multipliziert; diese Ausgaben entsprechen auch den Gesamtausgaben des Halbjahres abzüglich der Ausgaben der zur umfassenden Gruppe gehörenden Elemente.
- 4) In jedem Halbjahr  $t$  wird das Ergebnis des vorigen Schritts durch den Stichprobenumfang in der nicht umfassenden Gruppe ( $n_{st}$ ) dividiert.
- 5) Die Ergebnisse des vorherigen Schritts werden für beide Halbjahre aufsummiert.

Der prognostizierte Fehler auf der Ebene der Grundgesamtheit ist genau die Summe dieser beiden Komponenten:

$$EE = EE_e + EE_s$$

### 7.3.3.5 Genauigkeit

Wie beim MUS-Standardansatz ist die Genauigkeit das Maß der mit der Extrapolation verbundenen Unsicherheit. Sie steht für den Stichprobenfehler und ist zu berechnen, um anschließend ein Konfidenzintervall zu erhalten.

Für die Genauigkeit gilt die Formel:

$$SE = z \times \sqrt{\frac{BV_{s1}^2}{n_{s1}} \times s_{rs1}^2 + \frac{BV_{s2}^2}{n_{s2}} \times s_{rs2}^2}$$

wobei  $s_{rs2}$  die Standardabweichung der Fehlerquoten in der Stichprobe der nicht umfassenden Gruppe von Halbjahr  $t$  ist (berechnet aus der Stichprobe, die auch zur Hochrechnung der Fehler auf die Grundgesamtheit verwendet wurde)

$$s_{rs2}^2 = \frac{1}{n_{st} - 1} \sum_{i=1}^{n_{st}} (r_{it} - \bar{r}_{st})^2, t = 1,2$$

Dabei entspricht  $\bar{r}_{st}$  dem einfachen Durchschnitt der Fehlerquoten in der Stichprobe der nicht umfassenden Gruppe von Halbjahr  $t$ .

Der Stichprobenfehler wird nur für die nicht umfassenden Gruppen berechnet, da sich aus den umfassenden Gruppen kein Stichprobenfehler ergibt.

### 7.3.3.6 Bewertung

Um auf die Erheblichkeit der Fehler zu schließen, ist zunächst die obere Fehlergrenze (ULE) zu berechnen. Diese Obergrenze entspricht der Summe aus dem prognostizierten Fehler  $EE$  selbst und der Genauigkeit der Extrapolation

$$ULE = EE + SE$$

Anschließend sollten der prognostizierte Fehler und die Obergrenze jeweils mit dem maximal zulässigen Fehler verglichen werden, um Prüfungsschlussfolgerungen unter Verwendung des Ansatzes zu ziehen, der in Abschnitt 7.3.1.6 dargelegt wurde.

### 7.3.3.7 Beispiel

Um die sich normalerweise am Ende eines Prüffjahres häufenden Prüfarbeiten besser zu verteilen, hat sich die Prüfbehörde entschieden, den Prüfaufwand auf zwei Zeiträume aufzuteilen. Am Ende des ersten Halbjahres betrachtete sie die entsprechend den beiden Halbjahren in zwei Gruppen unterteilte Grundgesamtheit, die sich zu diesem Zeitpunkt durch folgende Merkmale auszeichnet:

Geltend gemachte Ausgaben am Ende des ersten Halbjahres	1 827 930 259 EUR
Größe der Grundgesamtheit (Vorhaben – erstes Halbjahr)	2344

Ausgehend von den vorliegenden Erfahrungen, ist der Prüfbehörde bekannt, dass in der Regel alle in den Programmen am Ende des Bezugszeitraums enthaltenen Vorhaben bereits in der Grundgesamtheit des ersten Halbjahres aktiv sind. Ferner wird erwartet, dass die am Ende des ersten Halbjahres geltend gemachten Ausgaben etwa 35 % der zum Ende des Bezugsjahres insgesamt geltend gemachten Ausgaben ausmachen. Auf der Grundlage dieser Annahmen gibt die folgende Tabelle eine Übersicht über diese Grundgesamtheit:

Geltend gemachte Ausgaben (DE) am Ende des ersten Halbjahrs	1 827 930 259 EUR
Geltend gemachte Ausgaben (DE) am Ende des zweiten Halbjahres (Vorhersage) 1 827 930 259 EUR ÷ 35 % – 1 827 930 259 EUR) = 3 394 727 624 EUR)	3 394 727 624 EUR
Prognostizierte Gesamtausgaben im ganzen Jahr	5 222 657 883 EUR
Größe der Grundgesamtheit (Vorhaben – erstes Halbjahr)	2344
Größe der Grundgesamtheit (Vorhaben – zweites Halbjahr, Vorhersage)	2344

Im ersten Zeitraum wird der Gesamtstichprobenumfang (für beide Halbjahre) wie folgt errechnet:

$$n = \left( \frac{z \times BV \times \sigma_{rw}}{TE - AE} \right)^2$$

Dabei ist  $\sigma_{rw}^2$  der gewichtete Durchschnitt der Varianzen der Fehlerquoten im jeweiligen Halbjahr; das Gewicht jedes Halbjahres entspricht dem Quotienten aus dem Buchwert des Halbjahres ( $BV_1$ ) und Buchwert der Grundgesamtheit ( $BV$ )

$$\sigma_{rw}^2 = \frac{BV_1}{BV} \sigma_{r1}^2 + \frac{BV_2}{BV} \sigma_{r2}^2$$

und  $\sigma_{r_t}^2$  ist die Varianz der Fehlerquoten in jedem Halbjahr. Die Varianz der Fehlerquoten wird für jedes Halbjahr berechnet als

$$\sigma_{r_t}^2 = \frac{1}{n_t^p - 1} \sum_{i=1}^{n_t^p} (r_{it} - \bar{r}_t)^2, t = 1, 2, \dots, T$$

Da diese Varianzen nicht bekannt sind, beschloss die Prüfbehörde, am Ende des ersten laufenden Halbjahres eine Vorabstichprobe von 20 Vorhaben zu ziehen. Die Standardabweichung der Fehlerquoten bei dieser Vorabstichprobe im ersten Halbjahr beträgt 0,12. Auf der Basis einer fachlichen Beurteilung und in dem Wissen, dass die Ausgaben im zweiten Halbjahr gewöhnlich höher sind als im ersten, hat die Prüfbehörde vorläufig vorausgesagt, dass die Standardabweichung der Fehlerquoten für das zweite Halbjahr 110 % höher ist als im ersten Halbjahr, d. h. 0,25 beträgt. Daraus ergibt sich folgender gewichteter Durchschnitt der Varianzen der Fehlerquoten:

$$\begin{aligned} \sigma_{rw}^2 &= \frac{1,827,930,259}{1,827,930,259 + 3,394,727,624} \times 0,12^2 \\ &+ \frac{3,394,727,624}{1,827,930,259 + 3,394,727,624} \times 0,25^2 = 0,0457 \end{aligned}$$

Im ersten Halbjahr erachtet die Prüfbehörde angesichts der Funktionsweise des Verwaltungs- und Kontrollsystems ein Konfidenzniveau von 60 % für angemessen. Der Gesamtstichprobenumfang für das gesamte Jahr ist:

$$n = \left( \frac{0,842 \times (1,827,930,259 + 3,394,727,624) \times \sqrt{0,0457}}{104,453,158 - 20,890,632} \right)^2 \approx 127$$

wobei  $z$  den Wert 0,842 hat (Koeffizient, der einem Konfidenzniveau von 60 % entspricht) und  $TE$ , der zulässige Fehler, 2 % des Buchwerts (in der Verordnung festgelegter Höchstwert für die Erheblichkeitsschwelle) beträgt. Der Gesamtbuchwert umfasst den tatsächlichen Buchwert am Ende des ersten Halbjahres zuzüglich des vorhergesagten Buchwerts für das zweite Halbjahr 3 394 727 624 EUR, d. h. der zulässige Fehler beträgt 2 % x 5 222 657 883 EUR = **104,453,158** EUR. Bei der Vorjahresprüfung wurde eine Fehlerquote von 0,4 % prognostiziert. Es ergibt sich ein voraussichtlicher Fehler **AE** von 0,4 % x 5 222 657 883 EUR = 20 890 632 EUR.

Die Aufteilung der Stichprobe nach Halbjahr ist wie folgt:

$$n_1 = \frac{BV_1}{BV_1 + BV_2} = \frac{1,827,930,259}{1,827,930,259 + 3,394,727,624} \times 127 \approx 45$$

und

$$n_2 = n - n_1 = 83$$

Für das erste Halbjahr sind die hochwertigen Einheiten der Grundgesamtheit (sofern vorhanden) zu ermitteln, die zu einer hochwertigen Schicht für eine Prüfung zu 100 % gehören. Der Schwellenwert zur Bestimmung dieser oberen Schicht entspricht dem Quotienten aus dem Buchwert ( $BV_1$ ) und dem geplanten Stichprobenumfang ( $n_1$ ). Alle Elemente, deren Buchwert größer ist als dieser Schwellenwert (wenn  $BV_{i1} > BV_1/n_1$ ), werden der Schicht für die 100-%-Prüfung zugeordnet. In diesem Fall beträgt der Schwellenwert 40 620 672 EUR. Es gibt 11 Vorhaben, deren Buchwert über diesem Schwellenwert liegt. Der Gesamtbuchwert dieser Vorhaben beläuft sich auf 891 767 519 EUR.

Der der nicht umfassenden Schicht zuzuordnende Stichprobenumfang ( $n_{s1}$ ) ist die Differenz zwischen  $n_1$  und der Zahl der Stichprobeneinheiten in der umfassenden Schicht ( $n_2$ ), d. h. 34 Vorhaben.

Die Auswahl der Stichprobe in der nicht umfassenden Schicht erfolgt schließlich mit einer Wahrscheinlichkeit proportional zur Größe (PPS), d. h. proportional zu den Buchwerten der Elemente  $BV_{is1}$ , durch systematische Auswahl unter Verwendung eines Stichprobenintervalls, das dem Quotienten aus den Gesamtausgaben in der nicht umfassenden Schicht ( $BV_{s1}$ ) und dem Stichprobenumfang ( $n_{s1}$ ) entspricht, d. h.

$$SI_{s1} = \frac{BV_{s1}}{n_{s1}} = \frac{1,827,930,259 - 891,767,519}{34} = 27,534,198$$

Der Buchwert in der nicht umfassenden Schicht ( $BV_{s1}$ ) ist einfach die Differenz zwischen dem Gesamtbuchwert und dem Buchwert der fünf Vorhaben in der obersten Schicht.

Die Ergebnisse sind in der folgenden Tabelle zusammengefasst:

Schwellenwert – erstes Halbjahr	40 620 672 EUR
Zahl der Vorhaben, deren Buchwert über dem Schwellenwert liegt – erstes Halbjahr	11
Buchwert der Vorhaben, deren Buchwert über dem Schwellenwert liegt – erstes Halbjahr	891 767 519 EUR
$BV_{s1}$ – erstes Halbjahr	936 162 740 EUR

$n_{s1}$ – erstes Halbjahr	34
$SF_{s1}$ – erstes Halbjahr	27 534 198 EUR

Sechs der elf Vorhaben, deren Buchwert über dem Stichprobenintervall liegt, weisen Fehler auf. Der Gesamtfehler in dieser Schicht beläuft sich auf 19 240 855 EUR.

Eine Datei mit den restlichen 2333 Vorhaben der Grundgesamtheit wird nach dem Zufallsprinzip sortiert und eine sequenzielle Variable für den kumulierten Buchwert erstellt. Unter Anwendung des systematischen PPS-Verfahrens wird eine Stichprobe von 34 Vorhaben gezogen.

Der Wert der 34 Vorhaben wird geprüft. Die Summe der Fehlerquoten für das erste Halbjahr beträgt:

$$\sum_{i=1}^{34} \frac{E_{fs1}}{BV_{fs1}} = 1,4256$$

Die Standardabweichung der Fehlerquoten in der Stichprobe der nicht umfassenden Grundgesamtheit des ersten Halbjahres ergibt sich folgendermaßen (Näheres dazu siehe Abschnitt 7.3.1.7):

$$s_{rs1} = \sqrt{\frac{1}{34-1} \sum_{i=1}^{34} (r_{fs1} - \bar{r}_{s1})^2} = 0,085$$

Dabei entspricht  $\bar{r}_{s2}$  dem einfachen Durchschnitt der Fehlerquoten in der Stichprobe der nicht umfassenden Gruppe des ersten Halbjahres.

Am Ende des zweiten Halbjahres stehen mehr Informationen zur Verfügung. Insbesondere der Gesamtbetrag der im zweiten Halbjahr aktiven Vorhaben ist genau bekannt, die aus der Stichprobe des ersten Halbjahres errechnete Stichprobenvarianz der Fehlerquoten  $s_{r1}$  ist bereits vorhanden, und die Standardabweichung der Fehlerquoten für das zweite Halbjahr  $\sigma_{r2}$  lässt sich anhand einer Vorabstichprobe realer Daten nunmehr genauer bewerten.

Die Prüfbehörde stellt fest, dass in der am Ende des ersten Halbjahres getroffenen Annahme zu den Gesamtausgaben (3 394 727 624 EUR) der wirkliche Wert (2 961 930 008 EUR) überschätzt wurde. Es gibt noch zwei Parameter, für die aktualisierte Zahlen verwendet werden sollten.

Erstens ergab die Schätzung der Standardabweichung der Fehlerquoten auf der Grundlage der Stichprobe des ersten Halbjahres von 34 Vorhaben einen Wert von 0,085. Dieser neue Wert sollte nun zur Neubewertung des geplanten Stichprobenumfangs herangezogen werden. Zweitens sind die Ausgaben im zweiten

Halbjahr höher als zunächst geschätzt, so dass es die Prüfbehörde für sinnvoller hält, die Standardabweichung der Fehlerquoten für das zweite Halbjahr nicht wie anfänglich mit 0,25, sondern mit 0,30 anzusetzen. Die aktualisierten Zahlen für die Standardabweichung der Fehlerquoten in den beiden Halbjahren sind weit von den ursprünglichen Schätzwerten entfernt. Die Stichprobe für das zweite Halbjahr muss also korrigiert werden.

Parameter	Vorhersage im ersten Halbjahr	Ende des zweiten Halbjahres
Standardabweichung der Fehlerquoten erstes Halbjahr	0,12	0,085
Standardabweichung der Fehlerquoten zweites Halbjahr	0,25	0,30
Gesamtausgaben im zweiten Halbjahr	3 394 727 624 EUR	2 961 930 008 EUR

Unter Berücksichtigung dieser drei Anpassungen beträgt der neu berechnete Stichprobenumfang des zweiten Halbjahres

$$n_2 = \frac{(z \times BV_2 \times \sigma_{r2})^2}{(TE - AE)^2 - z^2 \times \frac{BV_1^2}{n_1} \times s_{r1}^2}$$

Dabei ist  $s_{r1}$  die Standardabweichung der Fehlerquoten, die (ebenso wie der prognostizierte Fehler) anhand der Stichprobe des ersten Halbjahres berechnet wurde, und  $\sigma_{r2}$  eine Schätzung der Standardabweichung der Fehlerquoten im zweiten Halbjahr:

$$n_2 = \frac{(0.842 \times 2,961,930,008 \times 0.30)^2}{(95,797,205 - 19,159,441)^2 - 0.842^2 \times \frac{1,827,930,259^2}{45} \times 0.085^2} \approx 102$$

Darin sind:

- $TE = (1\,827\,930\,259\text{€} + 2\,961\,930\,008\text{€}) \times 2\% = 95,797,205\text{€}$
- $AE = (1\,827\,930\,259\text{€} + 2\,961\,930\,008\text{€}) \times 0,4\% = 19,159,441\text{€}$

Es sind die hochwertigen Einheiten der Grundgesamtheit (sofern vorhanden) zu ermitteln, die zur hochwertigen Schicht für eine Prüfung zu 100 % gehören werden. Der Schwellenwert für die Bestimmung dieser oberen Schicht entspricht dem Quotienten aus dem Buchwert ( $BV_2$ ) und dem geplanten Stichprobenumfang ( $n_2$ ). Alle Elemente, deren Buchwert größer ist als dieser Schwellenwert (wenn  $BV_{i2} > BV_2/n_2$ ), werden der Schicht für die 100 %-Prüfung zugeordnet. In diesem Fall beträgt der Schwellenwert 29 038 529 EUR. Es gibt sechs Vorhaben, deren Buchwert über diesem Schwellenwert liegt. Der Gesamtbuchwert dieser Vorhaben beläuft sich auf 415 238 983 EUR.

Der der nicht umfassenden Schicht zuzuordnende Stichprobenumfang  $n_{s2}$  wird als Differenz zwischen  $n_2$  und der Anzahl der Stichprobeneinheiten (z. B. Vorhaben) in der umfassenden Schicht ( $n_{e2}$ ) berechnet; es handelt sich also um 96 Vorhaben (Stichprobenumfang 102 minus 6 hochwertige Vorhaben). Folglich muss der Prüfer bei der Auswahl folgendes Stichprobenintervall verwenden:

$$SI_{s2} = \frac{BV_{s2}}{n_{s2}} = \frac{2,961,930,008 - 415,238,983}{96} = 26,528,032$$

Der Buchwert in der nicht umfassenden Schicht ( $BV_{s2}$ ) ist einfach nur die Differenz zwischen dem Gesamtbuchwert und dem Buchwert der sechs Vorhaben in der oberen Schicht.

Die Ergebnisse sind in der folgenden Tabelle zusammengefasst:

Schwellenwert – zweites Halbjahr	29 038 529 EUR
Anzahl der Vorhaben, deren Buchwert größer ist als der Schwellenwert – erstes Halbjahr	6
Anzahl der Vorhaben, deren Buchwert größer ist als der Schwellenwert – zweites Halbjahr	415 238 983 EUR
$BV_{s2}$ – zweites Halbjahr	2 546 691 025 EUR
$n_{s2}$ – zweites Halbjahr	96
$SI_{s2}$ – zweites Halbjahr	26 528 032 EUR

Von den sechs Vorhaben, deren Buchwert über dem Schwellenwert liegt, weisen vier Fehler auf. Der Gesamtfehler in dieser Schicht beläuft sich auf 9 340 755 EUR.

Eine Datei mit den übrigen 2338 Vorhaben der Grundgesamtheit des zweiten Halbjahres wird nach dem Zufallsprinzip sortiert und eine sequenzielle Variable für den kumulierten Buchwert erstellt. Unter Anwendung des systematischen PPS-Verfahrens wird eine Stichprobe von 96 Vorhaben gezogen.

Der Wert dieser 96 Vorhaben wird geprüft. Die Summe der Fehlerquoten für das zweite Halbjahr beläuft sich auf:

$$\sum_{i=1}^{96} \frac{E_{2i}}{BV_{2i}} = 1,1875$$

Die Standardabweichung der Fehlerquoten in der Stichprobe der nicht umfassenden Grundgesamtheit des zweiten Halbjahres ist:

$$s_{r_{s2}} = \sqrt{\frac{1}{96-1} \sum_{i=1}^{96} (r_{is2} - \bar{r}_{s2})^2} = 0.29$$

Dabei entspricht  $\bar{r}_{s2}$  dem einfachen Durchschnitt der Fehlerquoten in der Stichprobe der nicht umfassenden Schicht des zweiten Halbjahres.

Die Hochrechnung der Fehler auf die Grundgesamtheit erfolgt bei den Einheiten in der umfassenden Schicht anders als bei den Elementen in der nicht umfassenden Schicht.

Bei den umfassenden Schichten, d. h. den Schichten mit den Stichprobenelementen, deren Buchwert größer ist als der Schwellenbetrag ( $BV_{ct} > \frac{BV_t}{n_t}$ ), entspricht der prognostizierte Fehler der Summe der Fehler, die bei den Elementen in diesen Schichten festgestellt werden:

$$EE_s = \sum_{t=1}^{n_1} E_{1t} + \sum_{t=1}^{n_2} E_{2t} = 19,240,855 + 9,340,755 = 28,581,610$$

In der Praxis bedeutet dies:

- 1) Für jedes Halbjahr  $t$  werden die zur umfassenden Gruppe gehörenden Einheiten ermittelt und ihre Fehler aufsummiert.
- 2) Die Ergebnisse des vorherigen Schritts werden für beide Halbjahre aufsummiert.

Bei der nicht umfassenden Schicht, d. h. der Schicht mit den Stichprobenelementen, deren Buchwert kleiner oder gleich dem Schwellenbetrag ist ( $BV_{ct} \leq \frac{BV_t}{n_t}$ ), ist der prognostizierte Fehler

$$EE_s = \frac{BV_{s1}}{n_{s1}} \times \sum_{i=1}^{n_{s1}} \frac{E_{1i}}{BV_{1i}} + \frac{BV_{s2}}{n_{s2}} \times \sum_{i=1}^{n_{s2}} \frac{E_{2i}}{BV_{2i}}$$

$$= \frac{936,162,740}{34} \times 1.4256 + \frac{2,546,691,025}{96} \times 1.1875 = 70,754,790$$

Zur Berechnung dieses prognostizierten Fehlers wird folgendermaßen vorgegangen:

- 1) In jedem Halbjahr  $t$  wird für jede Einheit in der Stichprobe die Fehlerquote berechnet, d. h. der Quotient aus dem Fehler und dem jeweiligen Ausgabenbetrag  $\frac{E_{it}}{BV_{it}}$ .
- 2) In jedem Halbjahr  $t$  werden diese Fehlerquoten für alle Einheiten in der Stichprobe aufsummiert.
- 3) In jedem Halbjahr  $t$  wird das Ergebnis des vorigen Schritts mit den Gesamtausgaben in der Grundgesamtheit der nicht umfassenden Gruppe ( $BV_{st}$ ) multipliziert; diese

Ausgaben entsprechen auch den Gesamtausgaben des Halbjahres abzüglich der Ausgaben der zur umfassenden Gruppe gehörenden Elemente.

4) In jedem Halbjahr  $t$  wird das Ergebnis des vorigen Schritts durch den Stichprobenumfang in der nicht umfassenden Gruppe ( $n_{st}$ ) dividiert.

5) Die Ergebnisse des vorherigen Schritts werden für beide Halbjahre aufsummiert. Der prognostizierte Fehler auf der Ebene der Grundgesamtheit ist genau die Summe dieser beiden Komponenten:

$$EE = EE_{\text{I}} + EE_{\text{II}} = 28,581,610 + 70,754,790 = 99,336,400$$

Dies entspricht einer prognostizierten Fehlerquote von 2,07 %.

Die Genauigkeit ist das Maß der mit der Hochrechnung verbundenen Unsicherheit. Für die Genauigkeit gilt die Formel:

$$\begin{aligned} SE &= z \times \sqrt{\frac{BV_{st1}^2}{n_{st1}} \times s_{\text{Fst1}}^2 + \frac{BV_{st2}^2}{n_{st2}} \times s_{\text{Fst2}}^2} \\ &= 0.842 \times \sqrt{\frac{936,162,740^2}{34} \times 0.085^2 + \frac{2,546,691,025^2}{96} \times 0.29^2} \\ &= 64,499,188 \end{aligned}$$

wobei  $s_{\text{Fst}}$  die Standardabweichung der bereits berechneten Fehlerquoten ist.

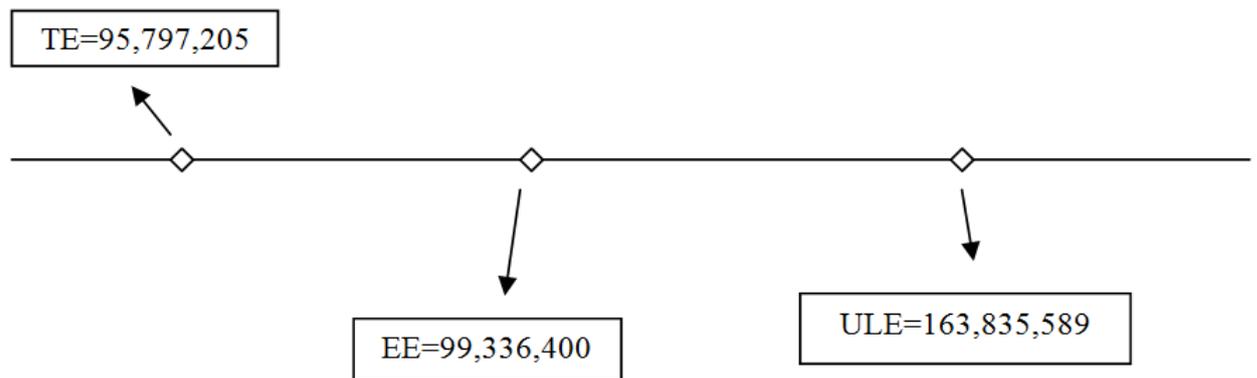
Der Stichprobenfehler wird nur für die nicht umfassenden Schichten berechnet, da sich aus den umfassenden Gruppen kein Stichprobenfehler ergibt.

Um auf die Erheblichkeit der Fehler zu schließen, ist zunächst die obere Fehlergrenze (ULE) zu berechnen. Diese Obergrenze entspricht der Summe aus dem prognostizierten Fehler  $EE$  selbst und der Genauigkeit der Hochrechnung

$$ULE = EE + SE = 99,336,400 + 64,499,188 = 163,835,589$$

Anschließend sollten der prognostizierte Fehler und die Obergrenze jeweils mit dem maximal zulässigen Fehler verglichen werden, um Prüfungsschlussfolgerungen zu ziehen:

In diesem konkreten Fall liegt der prognostizierte Fehler über dem maximal zulässigen Fehler. Somit würde der Prüfer schlussfolgern, dass hinreichende Beweise dafür vorliegen, dass die Fehler in der Grundgesamtheit die Erheblichkeitsschwelle überschreiten:



### 7.3.4 *Vorsichtig angelegter Ansatz*

#### 7.3.4.1 *Einleitung*

Beim wertbezogenen Stichprobenverfahren verwenden die Prüfer meist einen vorsichtig angelegten Ansatz. Der Vorteil liegt darin, dass weniger Kenntnisse über die Grundgesamtheit erforderlich sind (z. B. werden für die Berechnung des Stichprobenumfangs keine Angaben zur Streuung der Grundgesamtheit benötigt). Außerdem wird dieser Ansatz von mehreren Prüfsoftwarepaketen automatisch implementiert, was seine Anwendung zusätzlich erleichtert. Wenn eine ausreichende Unterstützung durch derartige Softwarepakete gegeben ist, setzt die vorsichtig angelegte Methode weit weniger technische und statistische Kenntnisse voraus als der sogenannte Standardansatz. Genau diese unproblematische Anwendung des konservativen Ansatzes hat aber auch einen großen Nachteil: Da weniger detaillierte Informationen für die Berechnung des Stichprobenumfangs und die Bestimmung der Genauigkeit verwendet werden, erfordert diese Methode in der Regel größere Stichproben und führt zu größeren geschätzten Stichprobenfehlern als die genaueren Formeln des Standardansatzes. Wenn jedoch die Stichprobe bereits einen überschaubaren Umfang aufweist und dem Prüfer keine größeren Probleme bereitet, kann dieser Ansatz wegen seiner Einfachheit eine gute Wahl sein.

Es ist nicht möglich, diese Methode mit einer Schichtung zu kombinieren oder die Prüfungen innerhalb des Bezugsjahres auf zwei oder mehr Zeiträume aufzuteilen, da die Formeln zur Bestimmung der Genauigkeit zu unhandlich wären. Daher wird den Prüfbehörden nahegelegt, für diese Zwecke den Standardansatz zu verwenden.

#### 7.3.4.2 *Stichprobenumfang*

Die Berechnung des Stichprobenumfangs  $n$  innerhalb der Rahmenbedingungen des wertbezogenen Stichprobenverfahrens beruht auf den folgenden Informationen:

- dem Buchwert der Grundgesamtheit (insgesamt geltend gemachte Ausgaben)  $BV$

- einer als Zuverlässigkeitsfaktor (**RF**) bezeichneten Konstante, die sich durch das Konfidenzniveau bestimmt
- dem maximal zulässigen Fehler *TE* (in der Regel 2 % der gesamten Ausgaben)
- dem voraussichtlichen Fehler *AE*, der vom Prüfer anhand seines fachlichen Urteils und vorhandener Informationen ausgewählt wird
- dem Expansionsfaktor **EF**, einer Konstante, die auch mit dem Konfidenzniveau zusammenhängt und verwendet wird, wenn Fehler erwartet werden.

Der Stichprobenumfang wird wie folgt berechnet:

$$n = \frac{BV \times RF}{TE - (AE \times EF)}$$

Der Zuverlässigkeitsfaktor **RF** ist eine Konstante aus der Poisson-Verteilung für einen erwarteten Null-Fehler. Er ist vom Konfidenzniveau abhängig. In der folgenden Tabelle sind die jeweils anzuwendenden Werte angegeben.

Konfidenzniveau	99 %	95 %	90 %	85 %	80 %	75 %	70 %	60 %	50 %
Zuverlässigkeitsfaktor (RF)	4,61	3,00	2,31	1,90	1,61	1,39	1,21	0,92	0,70

Tabelle 5. Zuverlässigkeitsfaktoren nach Konfidenzniveau

Der Expansionsfaktor **EF** ist ein Faktor, der bei der Berechnung der MUS-Stichprobe verwendet wird, wenn Fehler erwartet werden, und beruht auf dem Risiko der fälschlichen Akzeptanz. Er reduziert den Stichprobenfehler. Werden keine Fehler erwartet, dann ist der voraussichtliche Fehler (AE) gleich Null, und es wird kein Expansionsfaktor verwendet. Werte für den Expansionsfaktor sind der folgenden Tabelle zu entnehmen.

Konfidenzniveau	99 %	95 %	90 %	85 %	80 %	75 %	70 %	60 %	50 %
Expansionsfaktor (EF)	1,9	1,6	1,5	1,4	1,3	1,25	1,2	1,1	1,0

Tabelle 6. Expansionsfaktoren nach Konfidenzniveau

Aus der Formel für die Bestimmung des Stichprobenumfangs geht hervor, warum dieser Ansatz als vorsichtig bezeichnet wird. Der Stichprobenumfang hängt weder von der Größe der Grundgesamtheit noch von deren Streuung ab. Die Formel soll auf alle Arten von Grundgesamtheiten ungeachtet ihrer spezifischen Merkmale anwendbar sein und führt daher in der Regel zu Stichprobenumfängen, die die praktischen Erfordernisse übersteigen.

### 7.3.4.3 Stichprobenauswahl

Nach der Bestimmung des Stichprobenumfangs erfolgt die Auswahl der Stichprobe mit einer Wahrscheinlichkeit proportional zur Größe (PPS), d. h. proportional zu den Buchwerten der Elemente  $BV_i$ . Eine gängige Möglichkeit zur Durchführung der Auswahl ist eine systematische Auswahl unter Verwendung eines Stichprobenintervalls, das dem Quotienten aus den Gesamtausgaben ( $BV$ ) und dem Stichprobenumfang ( $n$ ) entspricht, d. h.

$$SI = \frac{BV}{n}$$

Typischerweise wird die Stichprobe aus einer in zufälliger Reihenfolge angeordneten Liste aller Elemente ausgewählt, wobei jedes Element, welches die  $x$ -te Geldeinheit enthält, ausgewählt wird, und  $x$  **die Schrittweite ist, die dem Buchwert – dividiert durch den Stichprobenumfang – entspricht**, also dem Stichprobenintervall.

Einige Elemente können mehrfach ausgewählt werden (wenn ihr Buchwert über dem Stichprobenintervall liegt). In diesem Fall sollte der Prüfer eine umfassende Schicht bilden, die alle Elemente enthält, deren Buchwert größer ist als das Stichprobenintervall. Bei dieser Schicht gestaltet sich die Behandlung der Fehlerprognose anders als sonst üblich.

### 7.3.4.4 Prognostizierter Fehler

Die Hochrechnung der Fehler auf die Grundgesamtheit erfolgt nach dem Verfahren, das im Zusammenhang mit dem MUS-Standardansatz vorgestellt wurde. Auch hier erfolgt die Hochrechnung für die Einheiten in der umfassenden Schicht anders als für die Einheiten in der nicht umfassenden Schicht.

Bei der umfassenden Schicht, d. h. der Schicht mit den Stichprobenelementen, deren Buchwert größer ist als das Stichprobenintervall ( $BV_i > \frac{BV}{n}$ ), entspricht der prognostizierte Fehler einfach der Summe der Fehler, die bei den Elementen in dieser Schicht festgestellt werden:

$$EE_g = \sum_{i=1}^{n_g} E_i$$

Bei der nicht umfassenden Schicht, d. h. der Schicht mit den Stichprobenelementen, deren Buchwert kleiner oder gleich dem Stichprobenintervall ( $BV_i \leq \frac{BP}{n}$ ) ist, ist der prognostizierte Fehler:

$$EE_s = SI \sum_{i=1}^{n_s} \frac{E_i}{BV_i}$$

Zur Berechnung dieses prognostizierten Fehlers wird folgendermaßen vorgegangen:

- 1) für jede Einheit in der Stichprobe wird die Fehlerquote berechnet, d. h. der Quotient aus dem Fehler und dem jeweiligen Ausgabenbetrag  $\frac{E_i}{BV_i}$  gebildet;
- 2) die ermittelten Fehlerquoten werden für alle Einheiten in der Stichprobe aufsummiert;
- 3) das Ergebnis des vorigen Schritts wird mit dem Stichprobenintervall (SI) multipliziert.

Der prognostizierte Fehler auf der Ebene der Grundgesamtheit ist genau die Summe dieser beiden Komponenten:

$$EE = EE_s + EE_g$$

#### 7.3.4.5 Genauigkeit

Die Genauigkeit, also die Messung des Stichprobenfehlers, hat zwei Komponenten: die Grundgenauigkeit  $BP$  und die inkrementelle Toleranz  $IA$ .

Die Grundgenauigkeit ist das Produkt aus Stichprobenintervall und Zuverlässigkeitsfaktor (bereits zur Berechnung des Stichprobenumfangs verwendet):

$$BP = SI \times RF.$$

Die inkrementelle Toleranz wird für jede mit einem Fehler behaftete Stichprobeneinheit in der nicht umfassenden Schicht berechnet.

Zuerst werden die fehlerbehafteten Elemente nach absteigendem Wert des Fehlers geordnet.

Danach wird die inkrementelle Toleranz für jedes dieser (fehlerbehafteten) Elemente nach folgender Formel berechnet:

$$IA_i = (RF(n) - RF(n-1) - 1) \times SI \times \frac{E_i}{BV_i}$$

Dabei ist  $RF(n)$  der Zuverlässigkeitsfaktor für den Fehler, der an n-ter Ordnung bei einem vorgegebenen Konfidenzniveau (in der Regel dasselbe wie bei der Berechnung der Stichprobengröße) auftritt, und  $RF(n-1)$  der Zuverlässigkeitsfaktor für den Fehler an (n-1)-ter Ordnung bei einem vorgegebenen Konfidenzniveau. Bei einer Konfidenz von 90 % ergibt sich folgende Tabelle von Zuverlässigkeitsfaktoren:

Ordnung des Fehlers	Zuverlässigkeitsfaktor (RF)	$RF(n) - RF(n-1) - 1$
Nullte Ordnung	2,31	
1.	3,89	0,58
2.	5,33	0,44
3.	6,69	0,36
4.	8,00	0,31
...		

Tabelle 7. Zuverlässigkeitsfaktoren nach Fehlerordnung

Beläuft sich beispielsweise der größere Fehler in der Stichprobe auf 10 000 EUR (25 % der Ausgaben von 40 000 EUR) und das Stichprobenintervall auf 200 000 EUR, so ist die individuelle inkrementelle Toleranz für diesen Fehler gleich  $0,58 \times 0,25 \times 200\,000 = 29\,000$  EUR.

Eine Tabelle mit Zuverlässigkeitsfaktoren für unterschiedliche Konfidenzniveaus und unterschiedliche Fehlerzahlen ist im Anhang enthalten.

Die inkrementelle Toleranz ist die Summe der inkrementellen Toleranzen aller Elemente:

$$IA = \sum_{i=1}^{n_g} IA_i$$

Die Gesamtgenauigkeit ( $SE$ ) entspricht der Summe der beiden Komponenten: Grundgenauigkeit ( $BP$ ) und inkrementelle Toleranz ( $IA$ ):

$$SE = BP + IA$$

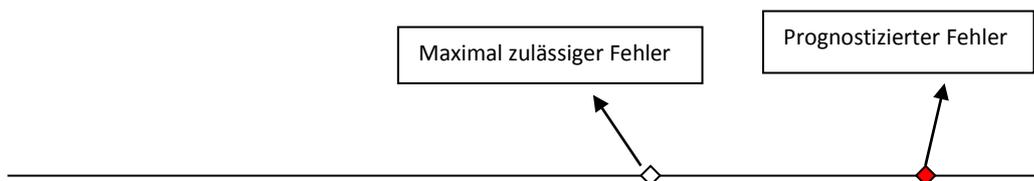
### 7.3.4.6 Bewertung

Um auf die Erheblichkeit der Fehler zu schließen, ist zunächst die obere Fehlergrenze (ULE) zu berechnen. Diese Obergrenze entspricht der Summe aus dem prognostizierten Fehler  $EE$  selbst und der Genauigkeit der Hochrechnung:

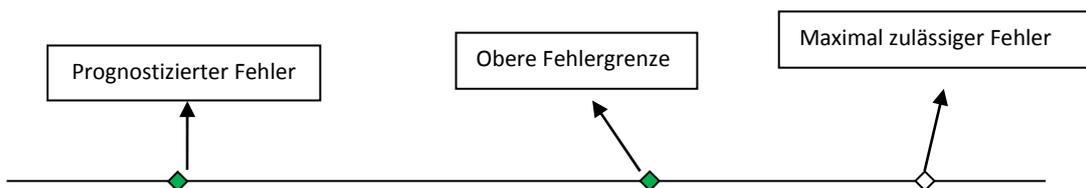
$$ULE = EE + SE$$

Anschließend sollten der prognostizierte Fehler und die Obergrenze jeweils mit dem maximal zulässigen Fehler verglichen werden, um Prüfungsschlussfolgerungen zu ziehen:

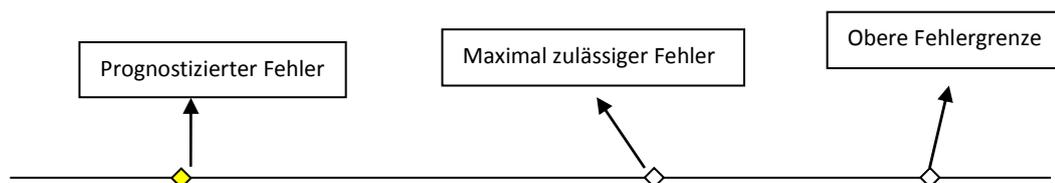
- Liegt der prognostizierte Fehler über dem maximal zulässigen Fehler, würde der Prüfer schlussfolgern, dass hinreichende Beweise dafür vorliegen, dass die Fehler in der Grundgesamtheit über der Erheblichkeitsschwelle liegen:



- Liegt die obere Fehlergrenze unter dem maximal zulässigen Fehler, sollte der Prüfer schlussfolgern, dass die Fehler in der Grundgesamtheit unter der Erheblichkeitsschwelle liegen.



- Liegt der prognostizierte Fehler unter dem maximal zulässigen Fehler, die obere Fehlergrenze jedoch darüber, sind weitere Schritte erforderlich, da nicht hinreichend belegt ist, dass die Grundgesamtheit keine wesentlichen Fehlangaben enthält. Die erforderlichen Schritte werden in Abschnitt 5.11 erörtert.



### 7.3.4.7 Beispiel

Es wird ausgegangen von einer Grundgesamtheit von Ausgaben, die gegenüber der Kommission in einem bestimmten Jahr für Vorhaben in einem Programm bescheinigt wurden. Die von der Prüfbehörde durchgeführten Systemprüfungen ergaben ein geringes Sicherheitsniveau. Die Stichprobenerhebung für dieses Programm sollte daher mit einem Konfidenzniveau von 90 % durchgeführt werden.

Die Grundgesamtheit ist in der nachstehenden Tabelle zusammengefasst:

Größe der Grundgesamtheit (Anzahl der Vorhaben)	3852
Buchwert (Summe der Ausgaben im Bezugsjahr)	4 199 882 024 EUR

Der Stichprobenumfang wird wie folgt berechnet:

$$n = \frac{BV \times RF}{TE - (AE \times EF)}$$

Dabei ist **BV** der Gesamtbuchwert der Grundgesamtheit, d. h. der Gesamtbetrag der Ausgaben, die im Bezugsjahr gegenüber der Kommission bescheinigt wurden; **RF** ist der dem Konfidenzniveau von 90 % entsprechende Zuverlässigkeitsfaktor (2,31); und **EF** ist der Hochrechnungsfaktor, der dem Konfidenzniveau entspricht, wenn Fehler erwartet werden (1,5). Bezüglich dieser spezifischen Grundgesamtheit hat die Prüfbehörde ausgehend von den Erfahrungen der letzten Jahre und in Anbetracht der erfolgten Verbesserungen am Verwaltungs- und Kontrollsystem entschieden, dass eine erwartete Fehlerquote von 0,2 % zuverlässig ist:

$$n = \frac{4,199,882,024 \times 2,31}{0,02 \times 4,199,882,024 - (0,002 \times 4,199,882,024 \times 1,5)} \approx 136$$

Die Auswahl der Stichprobe erfolgt mit einer Wahrscheinlichkeit proportional zur Größe (PPS), d. h. proportional zu den Buchwerten der Elemente (**BV<sub>i</sub>**). Es handelt sich um eine systematische Auswahl unter Verwendung eines Stichprobenintervalls, das dem Quotienten aus Gesamtausgaben (**BV**) und Stichprobenumfang (**n**) entspricht:

$$SI = \frac{BV}{n} = \frac{4,199,882,024}{136} = 30,881,485$$

Eine Datei mit den 3852 Vorhaben der Grundgesamtheit wird nach dem Zufallsprinzip sortiert und eine sequenzielle Variable für den kumulierten Buchwert erstellt.

Aus dieser randomisierten Liste aller Vorhaben wird die Stichprobe durch Auswahl jedes Elements gebildet, das die 30 881 485-ste Geldeinheit enthält.

Vorhaben	Buchwert (BV)	AcumBV
239	10 173 875 EUR	10 173 875 EUR
424	23 014 045 EUR	33 187 920 EUR
2327	32 886 198 EUR	66 074 118 EUR
5009	34 595 201 EUR	100 669 319 EUR
1491	78 695 230 EUR	179 364 549 EUR
(...)	(...)	(...)

Es wird ein Zufallswert zwischen 0 und dem Stichprobenintervall (30 881 485) generiert (16 385 476). Als erstes wird das Element ausgewählt, das die 16 385 476-ste Geldeinheit enthält. Als zweites folgt das erste Vorhaben in der Datei, dessen kumulierter Buchwert größer oder gleich 16 385 476 + 30 881 485 ist, usw. usf....

Vorhaben	Buchwert (BV)	AcumBV	Stichprobe
239	10 173 875 EUR	10 173 875 EUR	Nein
424	23 014 045 EUR	33 187 920 EUR	Ja
2327	32 886 198 EUR	66 074 118 EUR	Ja
5009	34 595 201 EUR	100 669 319 EUR	Ja
1491	78 695 230 EUR	179 364 549 EUR	Ja
(...)	(...)	(...)	(...)
2596	8 912 999 EUR	307 654 321 EUR	Ja
779	26 009 790 EUR	333 664 111 EUR	Nein
1250	264 950 EUR	333 929 061 EUR	Nein
3895	30 949 004 EUR	364 878 065 EUR	Ja
2011	617 668 EUR	365 495 733 EUR	Nein
4796	335 916 EUR	365 831 649 EUR	Nein
3632	7 971 113 EUR	373 802 762 EUR	Nein
2451	17 470 048 EUR	391 272 810 EUR	Ja
(...)	(...)	(...)	(...)

Es gibt 24 Vorhaben, deren Buchwert größer ist als das Stichprobenintervall, so dass jedes von ihnen mindestens einmal ausgewählt wird (z. B. wird Vorhaben 1491 dreimal

ausgewählt, vgl. vorstehende Tabelle). Der Buchwert dieser 24 Vorhaben beläuft sich auf 1 375 130 377 EUR. Bei vier dieser 24 Vorhaben sind Fehler aufgetreten, die einem Fehlerbetrag von 7 843 574 EUR entsprechen.

Bei der restlichen Stichprobe gestaltet sich die Fehlerbehandlung anders. Für diese Vorhaben wird folgendes Verfahren verwendet:

- 1) für jede Einheit in der Stichprobe wird die Fehlerquote berechnet, d. h. der Quotient aus dem Fehler und dem jeweiligen Ausgabenbetrag  $\frac{E_i}{BV_i}$  gebildet;
- 2) die ermittelten Fehlerquoten werden für alle Einheiten in der Stichprobe aufsummiert;
- 3) das Ergebnis des vorigen Schritts wird mit dem Stichprobenintervall (SI) multipliziert

$$EE_s = SI \sum_{i=1}^{n_s} \frac{E_i}{BV_i}$$

Vorhaben	Buchwert (BV)	Korrigierter Buchwert (CBV)	Fehler	Fehlerquote
2596	8 912 999 EUR	8 912 999 EUR	- EUR	-
459	869 080 EUR	869 080 EUR	- EUR	-
2073	859 992 EUR	859 992 EUR	- EUR	-
239	10 173 875 EUR	9 962 918 EUR	210 956 EUR	0,02
989	394 316 EUR	394 316 EUR	- EUR	-
65	25 234 699 EUR	25 125 915 EUR	108 784 EUR	0,00
5010	34 595 201 EUR	34 595 201 EUR	- EUR	-
...	...	...	...	...
3632	7 971 113 EUR	7 971 113 EUR	- EUR	-
3672	624 882 EUR	624 882 EUR	- EUR	-
2355	343 462 EUR	301 886 EUR	41 576 EUR	0,12
959	204 847 EUR	204 847 EUR	- EUR	-
608	15 293 716 EUR	15 293 716 EUR	- EUR	-
4124	6 773 014 EUR	6 773 014 EUR	- EUR	-
262	662 EUR	662 EUR	- EUR	-
<b>Gesamt</b>				<b>1,077</b>

$$EE_s = 30,881,485 \times 1,077 = 33,259,360$$

Der prognostizierte Fehler auf der Ebene der Grundgesamtheit ist genau die Summe dieser beiden Komponenten:

$$EE = 7,843,574 + 33,259,360 = 41,102,934$$

Dies entspricht einer prognostizierten Fehlerquote von 0,98 %.

Zur Bestimmung der oberen Fehlergrenze müssen die beiden Komponenten der Genauigkeit – die Grundgenauigkeit  $BP$  und die inkrementelle Toleranz  $IA$  – berechnet werden.

Die Grundgenauigkeit ist das Produkt aus Stichprobenintervall und Zuverlässigkeitsfaktor (das bereits zur Berechnung des Stichprobenumfangs verwendet wurde):

$$BP = 30,881,485 \times 2,31 = 71,336,231$$

Die inkrementelle Toleranz wird für jede mit einem Fehler behaftete Stichprobeneinheit in der nicht umfassenden Schicht berechnet.

Zuerst werden die fehlerbehafteten Elemente nach absteigendem Wert des Fehlers geordnet. Danach wird die inkrementelle Toleranz für jedes dieser (fehlerbehafteten) Elemente nach folgender Formel berechnet:

$$IA_i = (RF(n) - RF(n-1) - 1) \times SI \times \frac{E_i}{BV_i}$$

Darin ist  $RF(n)$  der Zuverlässigkeitsfaktor für den Fehler, der an n-ter Ordnung bei einem vorgegebenen Konfidenzniveau (in der Regel dasselbe wie bei der Berechnung des Stichprobenumfangs) auftritt, und  $RF(n-1)$  der Zuverlässigkeitsfaktor für den Fehler an (n-1)-ter Ordnung bei einem vorgegebenen Konfidenzniveau (siehe Tabelle im Anhang).

Die inkrementelle Toleranz ist die Summe der inkrementellen Toleranzen aller Elemente:

$$IA = \sum_{i=1}^{n_g} IA_i$$

In der nachstehenden Tabelle werden diese Ergebnisse für die 16 Vorhaben zusammengefasst, bei denen Fehler aufgetreten sind:

Ordnung	Fehler (A)	Fehlerquote (B): =(A)/BV	Prognostizierter Fehler: =(B)*SI	RF(n)	(RF(n)-RF(n-1))-1	IA <sub>i</sub>
0				2,31		
1	4 705 321 EUR	0,212	6 546 875 EUR	3,89	0,58	3 797 187 EUR
...	...	...	...	...	...	...
12	26 952 EUR	0,001	29 488 EUR	18,21	0,29	8 552 EUR
13	12 332 EUR	0,024	741 156 EUR	19,50	0,30	218 916 EUR
14	7 706 EUR	0,012	370 578 EUR	20,80	0,30	109 458 EUR
15	6 822 EUR	0,020	617 630 EUR	22,09	0,30	182 430 EUR
16	4 787 EUR	0,008	247 052 EUR	23,39	0,30	72 972 EUR
Gesamt		1,077	38 264 277 EUR			14 430 761 EUR

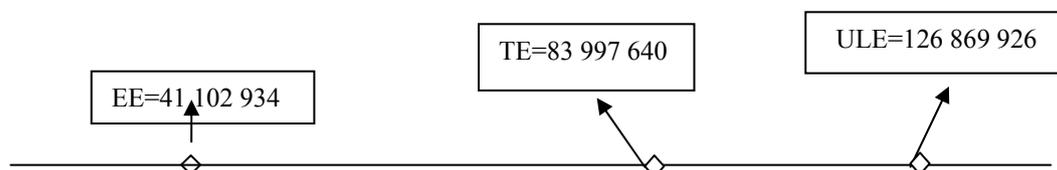
Die Gesamtgenauigkeit (**SE**) entspricht der Summe der beiden Komponenten Grundgenauigkeit (**BP**) und inkrementelle Toleranz (**IA**):

$$SE = 71,336,231 + 14,430,761 = 85,766,992$$

Um auf die Signifikanz der Fehler zu schließen, ist zunächst die obere Fehlergrenze zu berechnen. Diese Fehlergrenze entspricht der Summe aus dem prognostizierten Fehler **SE** und der Gesamtgenauigkeit der Hochrechnung:

$$ULE = 41,102,933 + 85,766,992 = 126,869,926$$

Jetzt wird der maximal zulässige Fehler ( $TE = 2 \% \times 4\,199\,882\,024 = 83\,997\,640$  EUR) sowohl mit dem prognostizierten Fehler als auch mit der oberen Fehlergrenze verglichen. Der maximal zulässige Fehler ist größer als der prognostizierte Fehler, aber kleiner als die obere Fehlergrenze. Im Fazit ist also nicht hinreichend belegt, dass die Grundgesamtheit keine wesentlichen Fehlangaben enthält. Daher sind weitere Schritte erforderlich, auf die in Abschnitt 5.11 eingegangen wurde.



## 7.4 Nichtstatistische Stichprobenverfahren

### 7.4.1 Überblick

Wie bereits ausgeführt (vgl. Abschnitt 6.2), sollte das nichtstatistische Stichprobenverfahren nur verwendet werden, wenn

- eine extrem kleine Grundgesamtheit vorliegt, deren Umfang die Ziehung einer hinreichend großen Stichprobe nicht zulassen würde (die Grundgesamtheit ist kleiner oder nahezu genauso groß wie der empfohlene Stichprobenumfang),
- der für ein statistisches Verfahren angeratene Stichprobenumfang aufgrund von unkontrollierbaren Beschränkungen nicht eingehalten werden kann.

Eine nichtstatistische Stichprobenziehung sollte mit allen Mitteln vermieden und nur für die Fälle vorbehalten werden, in denen keine ausreichend große Grundgesamtheit für ein statistisches Stichprobenverfahren gegeben ist. Der Prüfer greift also nur dann auf ein nichtstatistisches Verfahren zurück, wenn nicht der Stichprobenumfang erreicht wird, der für die Anwendung der relevanten Formeln empfohlen wird. Allerdings kann nicht exakt festgelegt werden, unterhalb welcher Grundgesamtheitsgröße ein nichtstatistisches Verfahren erforderlich ist, denn dabei spielen verschiedene Merkmale der Grundgesamtheit eine Rolle. Mit Sicherheit lässt sich lediglich sagen, dass diese Schwelle irgendwo im Bereich zwischen 50 und 150 Einheiten der Grundgesamtheit liegt.<sup>21</sup>

Bei Anwendung des nichtstatistischen Verfahrens können sich zwei Optionen ergeben:

- Option 1: Enthält die Grundgesamtheit Vorhaben mit hohem Wert, ist eine Schichtung nach Ausgaben ratsam. Dazu wird wie folgt vorgegangen:
  - Zunächst wird der Schwellenwert der Ausgaben für die Elemente festgelegt, die in die Schicht mit den hohen Werten aufgenommen werden sollen. In der Regel entspricht dieser Schwellenwert dem maximal zulässigen Fehler (2 % der Gesamtausgaben) der Grundgesamtheit. Dieser Schwellenwert kann und sollte in Abhängigkeit von den Merkmalen der Grundgesamtheit verändert werden. So wird er bei sehr kleinen Grundgesamtheiten (unter 50 Einheiten) in der Regel höher angesetzt. Hierbei sollte sich der Prüfer vor allem auf sein fachliches Urteilsvermögen stützen. Kann er Elemente ausmachen, die sich von den anderen durch signifikant höhere Ausgaben abheben, sollte er die Bildung einer Schicht aus diesen Elementen ins Auge fassen.
  - Die Elemente mit hohem Wert sollten zu 100 % geprüft werden.
  - Der erforderliche Stichprobenumfang für die verbleibende Grundgesamtheit wird nach fachlichem Urteil unter Berücksichtigung des bei den Systemprüfungen erreichten Sicherheitsniveaus ermittelt. Einer Faustregel zufolge sollten mindestens 10 % der verbleibenden

---

<sup>21</sup> Außerdem besagt eine Faustregel, dass der Stichprobenumfang bei statistischen Stichprobenverfahren mindestens 30 betragen sollte.

Grundgesamtheit der Vorhaben in die Stichprobe aufgenommen werden, doch kann sich dieser Wert je nach dem fachlichen Urteil des Prüfers ändern.

- Option 2: Wenn die Grundgesamtheit keine Vorhaben mit hohem Wert (also mit Ausgaben oberhalb der empfohlenen Schwelle) enthält, kann auch keine hochwertige Schicht gebildet werden. In diesem Fall wird wie folgt vorgegangen:
  - Der erforderliche Stichprobenumfang wird nach fachlichem Urteil unter Berücksichtigung des bei den Systemprüfungen erreichten Sicherheitsniveaus ermittelt. Auch hier gilt als Faustregel, dass mindestens 10 % der verbleibenden Grundgesamtheit der Vorhaben in die Stichprobe aufgenommen werden, wobei dieser Wert je nach dem fachlichen Urteil des Prüfers geändert werden kann.

Auf jeden Fall wird empfohlen, die Stichprobe nach einem Zufallsverfahren zu ziehen. In Frage kommen die Auswahl mit gleicher Wahrscheinlichkeit, wie sie beim einfachen Zufallsstichprobenverfahren erfolgt (vgl. Abschnitt 7.1.1.6 bzw. 7.1.2.6), die Differenzschätzung (vgl. Abschnitt 7.2.1.6 bzw. 7.2.2.6) oder die Auswahl mit einer Wahrscheinlichkeit proportional zur Größe (Ausgabenhöhe) (siehe Abschnitt 7.4.2). Bei der Entscheidung sind die Streuung der Ausgaben in der hochwertigen Schicht und der erwartete Zusammenhang zwischen Fehlern und Ausgaben zu berücksichtigen. Sollte bei den Elementen in der niedrigwertigen Schicht noch immer eine starke Streuung der Ausgaben bestehen und ein starker positiver Zusammenhang zwischen Fehlern und Ausgaben erwartet werden, dann empfiehlt sich das PPS-Verfahren (Wahrscheinlichkeit proportional zur Größe). Ansonsten wird die Stichprobe durch Auswahl mit gleicher Wahrscheinlichkeit gebildet werden. Nicht randomisierte Auswahlverfahren sind ebenfalls möglich, was vor allem im Falle sehr kleiner Grundgesamtheiten und Stichprobenumfänge gilt. Soll bei einer Grundgesamtheit von 15 Einheiten eine Stichprobe im Umfang von zwei Einheiten gezogen werden, so kann die Stichprobenauswahl auch auf der Grundlage eines fachlichen Urteils erfolgen, wenn der Prüfer über die Merkmale und die Risiken bei den 15 Einheiten in der Grundgesamtheit im Bilde ist.

Bei der umfassenden Schicht, d. h. der Schicht mit den Stichprobeneinheiten, deren Buchwert größer ist als der Schwellenbetrag, wird der prognostizierte Fehler wie folgt berechnet:

$$EE_g = \sum_{i=1}^{n_g} E_i$$

Dabei ist  $E_i = BV_i - CBV_i, i = 1, 2, \dots, n_g$ , also wie üblich der Fehlerbetrag eines Vorhabens, d. h. die Differenz zwischen dem i-ten Vorhaben in der Stichprobe und dem entsprechenden korrigierten Buchwert.

Anders ausgedrückt ist der prognostizierte Fehler die Summe der Einzelfehler bei allen Elementen in der hochwertigen Schicht.

Bei der Schicht mit den Stichprobeneinheiten, deren Buchwert unter dem Schwellenbetrag liegt, richtet sich die Berechnung des prognostizierten Fehlers danach, ob die Auswahl mit gleicher Wahrscheinlichkeit oder mit einer Wahrscheinlichkeit proportional zur Größe getroffen wurde.

Wurden die Einheiten mit gleicher Wahrscheinlichkeit ausgewählt, dann errechnet sich der prognostizierte Zufallsfehler für die niedrigwertige Schicht wie folgt:

$$EE_s = N_s \frac{\sum_{i=1}^{n_s} E_i}{n_s},$$

wobei  $N_s$  die Größe der Grundgesamtheit und  $n_s$  der Stichprobenumfang in der niedrigwertigen Schicht sind.

Dieser prognostizierte Fehler entspricht dem Durchschnitt der Fehler in der Stichprobe multipliziert mit der Grundgesamtheitsgröße in der niedrigwertigen Schicht.

Wurden die Einheiten mit Wahrscheinlichkeiten ausgewählt, die dem Ausgabenwert proportional sind, dann errechnet sich der prognostizierte Zufallsfehler für die niedrigwertige Schicht wie folgt:

$$EE_s = \frac{BV_s}{n_s} \sum_{i=1}^{n_s} \frac{E_i}{BV_i}$$

Darin sind  $BV_s$  der Gesamtbuchwert und  $n_s$  der Stichprobenumfang in der niedrigwertigen Schicht.

Dieser prognostizierte Fehler entspricht dem Durchschnitt der Fehlerquoten in der Stichprobe multipliziert mit dem Gesamtbuchwert der niedrigwertigen Schicht.

Der prognostizierte Gesamtfehler auf der Ebene der Grundgesamtheit ist einfach die Summe dieser beiden Komponenten:

$$EE = EE_h + EE_s$$

Abschließend wird der prognostizierte Fehler mit dem maximal zulässigen Fehler (Erheblichkeit multipliziert mit den Ausgaben der Grundgesamtheit) verglichen:

- Liegt er unterhalb des zulässigen Fehlers, dann ist daraus zu schließen, dass die Grundgesamtheit keinen wesentlichen Fehler enthält.
- Liegt er oberhalb des zulässigen Fehlers, dann ist daraus zu schließen, dass die Grundgesamtheit einen wesentlichen Fehler enthält.

Trotz der vorhandenen Beschränkungen (keine Möglichkeit der Berechnung der oberen Fehlergrenze und daher keine Kontrolle des Prüfrisikos) ist die prognostizierte Fehlerquote die bestmögliche Schätzung des Fehlers in der Grundgesamtheit und kann daher mit der Erheblichkeitsschwelle verglichen werden, um Rückschlüsse auf mögliche wesentliche Fehlangaben in der Grundgesamtheit zu ziehen.

#### **7.4.2 Beispiel**

Angenommen wird, dass für eine Grundgesamtheit von 36 Vorhaben Ausgaben in Höhe von 22 031 228 EUR geltend gemacht wurden.

Diese Grundgesamtheit ist für eine Prüfung mit statistischer Stichprobenziehung kaum ausreichend. Außerdem ist keine Stichprobenziehung von Zahlungsanträgen möglich. Daher entscheidet sich die Prüfbehörde für einen nichtstatistischen Ansatz mit Schichtung der hochwertigen Vorhaben, da die Ausgaben bei einigen Vorhaben extrem hoch sind.

Nachstehend eine Zusammenfassung der Merkmale der Grundgesamtheit:

Geltend gemachte Ausgaben (DE) im Bezugszeitraum	22 031 228 EUR
Größe der Grundgesamtheit (Vorhaben)	36
Erheblichkeitsschwelle (maximal 2 %)	2 %
Zulässige Fehlangebe (TE)	440 625 EUR

Als Erstes ermitteln die Prüfer diejenigen Vorhaben, die einzeln betrachtet einen erheblichen Wert repräsentieren oder von Natur aus wesentlich sind. Die wesentlichen Einzelbeträge werden der Erheblichkeitsschwelle gleichgesetzt (2 % von 22 031 228 EUR).

In der nachstehenden Tabelle sind die Ergebnisse zusammengefasst:

Zahl der Einheiten über dem Schwellenwert	4
Buchwert der Grundgesamtheit über dem Schwellenwert	8 411 965 EUR

Wert der restlichen Grundgesamtheit	13 619 623 EUR
-------------------------------------	----------------

Diese Projekte werden von der Stichprobenziehung ausgenommen und separat behandelt. Ihr Gesamtwert beläuft sich auf 8 411 965 EUR. Der festgestellte Fehlerbetrag bei diesen vier Vorhaben beträgt

$$EE_s = 80,028.$$

Die Prüfbehörde schätzt die Qualität des Verwaltungs- und Kontrollsystems als durchschnittlich ein und entscheidet sich daher für einen Stichprobenumfang von 20 % der restlichen Grundgesamtheit von Vorhaben. Dies sind 20 % x 32 = 6,4, exzessiv aufgerundet auf 7.

Aufgrund der großen Streuung der Ausgaben bei dieser Grundgesamtheit beschließt der Prüfer, die Stichprobe in der restlichen Grundgesamtheit nach dem MUS-Verfahren (Wahrscheinlichkeit proportional zur Größe) zu bilden. Das Stichprobenintervall ist gleich den Gesamtausgaben in der nicht umfassenden Schicht ( $BV_s$ ) dividiert durch den Stichprobenumfang ( $n_s$ ), d. h.

$$SI = \frac{BV_s}{n_s} = \frac{13,619,623}{7} = 1,945,609$$

Eine Datei mit den restlichen 32 Vorhaben der Grundgesamtheit wird nach dem Zufallsprinzip sortiert und eine sequenzielle Variable für den kumulierten Buchwert erstellt. Die Stichprobe wird durch Auswahl jedes Elements gebildet, dass die 1 945 609-te Geldeinheit enthält.

Der Wert des hochgerechneten Fehlers für die niedrigwertige Schicht errechnet sich wie folgt:

$$FF_s = \frac{BV_s}{n_s} \sum_{i=1}^{n_s} \frac{E_{st}}{BV_{st}}$$

Dabei sind  $BV_s$  der Gesamtbuchwert der restlichen Grundgesamtheit und  $n_s$  der entsprechende Stichprobenumfang. Übrigens ist dieser prognostizierte Fehler gleich der Summe der Fehlerquoten multipliziert mit dem Stichprobenintervall. Die Summe der Fehlerquoten ist gleich 0,0272:

$$EE_s = \frac{13,619,623}{7} \times 0,0272 = 52,921.$$

Der prognostizierte Gesamtfehler auf der Ebene der Grundgesamtheit ist einfach die Summe dieser beiden Komponenten:

$$EE = EE_e + EE_s = 80,028 + 52,921 = 132,949$$

Abschließend wird der prognostizierte Fehler mit dem maximal zulässigen Fehler (2 % von 22 031 228 EUR = 440 625 EUR) verglichen. Der prognostizierte Fehler liegt unterhalb der Erheblichkeitsschwelle.

Somit kann der Prüfer darauf schließen, dass die Grundgesamtheit keinen wesentlichen Fehler enthält. Allerdings kann die erreichte Genauigkeit nicht bestimmt werden, und die Konfidenz der Schlussfolgerung ist unbekannt. Somit entscheidet das fachliche Urteil des Prüfers darüber, ob weitere Prüfverfahren oder alternative Strategien angewandt werden sollten, um die geltend gemachten Ausgaben zu beurteilen.

## 8 Ausgewählte Themen

### 8.1 Bestimmung des voraussichtlichen Fehlers

Der voraussichtliche Fehler ist der Fehlbetrag, den der Prüfer in der Grundgesamtheit vorzufinden erwartet. Die Fehlererwartung des Prüfers wird beeinflusst durch die Ergebnisse der Überprüfung der Kontrollen, der im vorangegangenen Zeitraum durchgeführten Prüfverfahren und anderer vertiefter Prüfungen. Mit zunehmender Abweichung des voraussichtlichen Fehlers vom wahren Fehler nimmt das Risiko zu, dass die Prüfung nicht zu schlüssigen Ergebnissen führt ( $MLE < 2\%$  und  $ULE > 2\%$ ).

Bei der Bestimmung des Wertes des voraussichtlichen Fehlers sollte der Prüfer Folgendes berücksichtigen:

1. Liegen dem Prüfer Informationen zu den Fehlerquoten aus Vorjahren vor, so sollte der voraussichtliche Fehler grundsätzlich auf dem prognostizierten Fehler des Vorjahres beruhen. Hat allerdings der Prüfer Kenntnis von Qualitätsveränderungen bei den Kontrollsystemen erlangt, kann er den voraussichtlichen Fehler ausgehend von diesen Informationen entweder niedriger oder höher ansetzen. Wenn beispielsweise die prognostizierte Fehlerquote im letzten Jahr 0,7 % betrug und keine weiteren Informationen vorliegen, kann eine voraussichtliche Fehlerquote von 0,7 % unterstellt werden. Hat der Prüfer jedoch Belege für eine Verbesserung der Systeme erhalten, die ihn überzeugt haben, dass die Fehlerquote im laufenden Jahr niedriger ausfallen wird, kann er den voraussichtlichen Fehler unter Berücksichtigung dieser Information niedriger ansetzen, also beispielsweise mit 0,4 %.
2. Liegen keine historischen Daten zu Fehlerquoten vor, kann der Prüfer anhand einer Vorabstichprobe/Pilotstichprobe eine vorläufige Schätzung der

Fehlerquote in der Grundgesamtheit vornehmen. Als voraussichtliche Fehlerquote gilt der Fehler, der anhand dieser Vorabstichprobe prognostiziert wurde. Wird bereits eine Vorabstichprobe ausgewählt, um die erforderlichen Standardabweichungen zur Berechnung der Formeln für den Stichprobenumfang zu bestimmen, dann kann dieselbe Vorabstichprobe für eine erste Hochrechnung der Fehlerquote und somit des voraussichtlichen Fehlers verwendet werden.

3. Liegen keine historischen Informationen vor, anhand derer ein voraussichtlicher Fehler bestimmt werden kann, und kann aufgrund nicht beeinflussbarer Beschränkungen keine Vorabstichprobe verwendet werden, dann sollte der Prüfer den Wert des voraussichtlichen Fehlers ausgehend von seiner beruflichen Erfahrung und seinem fachlichen Urteil festlegen. In diesem Fall wird für gewöhnlich ein voraussichtlicher Fehler im Bereich von 10 % bis 30 % der Erheblichkeit gewählt. Allerdings sollte der Wert nicht aufgrund dieser Leitlinien eingeschränkt werden, sondern in erster Linie die Erwartungen des Prüfers hinsichtlich des wahren Fehlerniveaus in der Grundgesamtheit reflektieren.

Kurz gesagt sollte der Prüfer also vorherige Daten, Hilfsdaten, sein berufliches Urteilsvermögen oder eine Kombination aus all diesen Faktoren nutzen, um einen möglichst realistischen Wert für den voraussichtlichen Fehler festzulegen.

Ein auf objektiven quantitativen Daten beruhender voraussichtlicher Fehler ist in der Regel genauer und verhindert, dass Mehrarbeiten wegen uneindeutiger Prüfergebnisse anfallen. Legt der Prüfer beispielsweise den voraussichtlichen Fehler mit 10 % der Erheblichkeit – d. h. 0,2 % der Ausgaben – fest und erhält am Ende der Prüfung einen prognostizierten Fehler von 1,5 %, so werden die Ergebnisse höchstwahrscheinlich nicht eindeutig sein, da die obere Fehlergrenze höher ist als die Erheblichkeitsschwelle. In diesem Fall sind weitere Arbeitsschritte erforderlich. Um dies zu vermeiden, sollte der Prüfer den voraussichtlichen Fehler bei künftigen Stichprobenziehungen so ansetzen, dass er den wahren Fehler in der Grundgesamtheit möglichst realistisch wiedergibt.

Eine besondere Situation kann auftreten, wenn die voraussichtliche Fehlerquote knapp über oder unter 2 % liegt. So kann sich bei einem voraussichtlichen Fehler von 1,9 % und einem hohen Konfidenzniveau (z. B. 90 %) ein extrem großer und kaum erreichbarer Stichprobenumfang ergeben. Dieses Phänomen tritt bei allen Stichprobenverfahren immer dann auf, wenn die geplante Genauigkeit sehr gering ist (0,1 % in diesem Beispiel). In diesem Fall besteht eine empfehlenswerte Möglichkeit darin, die Grundgesamtheit in zwei Teilgesamtheiten aufzuspalten, in denen der Prüfer unterschiedliche Fehlerniveaus erwartet. Wenn eine Teilgesamtheit mit einem erwarteten Fehler unter 2 % und eine weitere Teilgesamtheit mit einem erwarteten Fehler über 2 % ermittelt werden kann, ermöglicht dies dem Prüfer eine sichere Planung zweier unterschiedlicher Stichproben aus diesen Teilgesamtheiten, ohne dass ein übermäßig großer Stichprobenumfang zu befürchten ist.

Nicht zuletzt sollte die Prüfbehörde ihre Prüftätigkeit so planen, dass auch dann eine ausreichende Genauigkeit des MLE gegeben ist, wenn der voraussichtliche Fehler deutlich über der Erheblichkeitsschwelle liegt (d. h. größer oder gleich 3,5 % ist). In diesem Fall sollten die Formeln für den Stichprobenumfang mit einem voraussichtlichen Fehler berechnet werden, der zu einer geplanten Genauigkeit von maximal 1,5 % führt. Wenn dagegen vorherige Daten aus Vorhabenprüfungen sowie eventuell Ergebnisse von Systemprüfungen zu einer voraussichtlich sehr niedrigen Fehlerquote führen (d. h. nicht höher als 0,5 %, woraus sich eine geplante Genauigkeit von mehr als 1,5 % ergibt), kann der Prüfer nach eigenem Ermessen entweder diese vorherigen Daten verwenden oder den voraussichtlichen Fehler auf 0,5 % festlegen. Die Prüfbehörden sollten das Ziel dieses Vorgehens beachten: Durch entsprechende Maßnahmen soll sichergestellt werden, dass die erzielte Genauigkeit nach erfolgter Prüfung nicht über 2 % beträgt, da es andernfalls sein kann, dass die Genauigkeit des berechneten MLE nicht ausreicht.

## **8.2 Zusätzliche Stichprobenziehung**

### ***8.2.1 Ergänzende Stichprobe (wegen unzureichender Erfassung hoch riskanter Bereiche)***

Artikel 17 Absatz 5 der Verordnung (EG) Nr. 1828/2006 der Kommission und Artikel 43 Absatz 5 der Verordnung (EG) Nr. 498/2007 der Kommission verweisen auf eine ergänzende Stichprobe.

Die Ergebnisse der Zufallsstichproben müssen in Bezug auf die Ergebnisse der Risikoanalyse jedes Programms und den Erfassungsbereich der Prioritäten, die Art der Vorhaben, die Begünstigten usw. im Programm bewertet werden. Ergibt sich aus diesem Vergleich, dass die Zufallsstichprobe nicht auf die hoch riskanten Bereiche und/oder Erfassungsbereiche abgestellt ist, so sollte sie durch eine weitere Auswahl von Vorhaben, d. h. eine ergänzende Stichprobe, ergänzt werden.

Die Prüfbehörde sollte während des Durchführungszeitraums diese Bewertung regelmäßig durchführen.

Dabei werden die Ergebnisse der Prüfungen der ergänzenden Stichprobe getrennt von den Ergebnissen der Prüfungen der Zufallsstichprobe analysiert. Insbesondere werden die in der ergänzenden Stichprobe entdeckten Fehler bei der Berechnung der Fehlerquote im Ergebnis der Prüfung der Zufallsstichprobe nicht berücksichtigt. Es ist jedoch auch eine detaillierte Analyse der in der ergänzenden Stichprobe entdeckten Fehler durchzuführen, um die Fehlerart festzustellen, und es sind Empfehlungen für deren Korrektur abzugeben.

Die Ergebnisse der Prüfung der ergänzenden Stichprobe sind der Kommission im jährlichen Kontrollbericht unmittelbar nach der Prüfung der ergänzenden Stichprobe vorzulegen.

### **8.2.2 Zusätzliche Stichprobe (wegen uneindeutiger Prüfergebnisse)**

Wenn die Ergebnisse der Prüfung keine eindeutigen Schlüsse zulassen und – unter Berücksichtigung von Abschnitt 8.7 – weitere Schritte erforderlich sind (in der Regel dann, wenn der prognostizierte Fehler unter der Erheblichkeitsschwelle liegt, die obere Grenze jedoch darüber), kann beispielsweise eine zusätzliche Stichprobe ausgewählt werden. In diesem Fall wird in den Formeln zur Bestimmung des Stichprobenumfangs nicht der voraussichtliche Fehler verwendet, sondern der aus der ursprünglichen Stichprobe abgeleitete prognostizierte Fehler (übrigens ist der prognostizierte Fehler in dieser Situation die bestmögliche Schätzung des Fehlers in der Grundgesamtheit). So kann anhand der neuen, aus der ursprünglichen Stichprobe hergeleiteten Information ein neuer Stichprobenumfang berechnet werden. Der Umfang der erforderlichen zusätzlichen Stichprobe kann durch Subtraktion des ursprünglichen Stichprobenumfangs vom neuen Stichprobenumfang ermittelt werden. Abschließend kann eine neue Stichprobe (nach demselben Verfahren wie die ursprüngliche) ausgewählt werden. Die beiden Stichproben werden zusammengefasst und neue Ergebnisse (prognostizierter Fehler und Genauigkeit) anhand der Daten aus der endgültigen zusammengefassten Stichprobe berechnet.

Angenommen, bei der ursprünglichen Stichprobe (mit einem Stichprobenumfang von 60 Vorhaben) ergab sich eine prognostizierte Fehlerquote von 1,5 % bei einer Genauigkeit von 0,9 %, dann beträgt die Obergrenze für die Fehlerquote  $1,5 + 0,9 = 2,4$  %. In diesem Fall liegt die prognostizierte Fehlerquote unterhalb der Erheblichkeitsschwelle von 2 %, die Obergrenze jedoch darüber. Will der Prüfer zu einem eindeutigen Schluss kommen, sind also weitere Schritte erforderlich (vgl. Abschnitt 5.11). Eine Möglichkeit besteht darin, weitere Prüfungen anhand zusätzlicher Stichproben vorzunehmen. Entscheidet sich der Prüfer für dieses Vorgehen, wird in der Formel für die Bestimmung des Stichprobenumfangs nicht der voraussichtliche Fehler verwendet, sondern eine prognostizierte Fehlerquote von 1,5 % unterstellt. Dies führt zur Neuberechnung des Stichprobenumfangs, der in unserem Beispiel nunmehr  $n = 78$  betragen würde. Da die ursprüngliche Stichprobe 60 Vorhaben umfasste, wird dieser Wert vom neuen Stichprobenumfang abgezogen, so dass sich  $78 - 60 = 18$  neue Beobachtungen ergeben. Folglich ist aus der Grundgesamtheit eine zusätzliche Stichprobe von 18 Vorhaben auszuwählen, wobei dasselbe Verfahren anzuwenden ist wie bei der ursprünglichen Stichprobe (z. B. MUS). Nach erfolgter Auswahl werden die beiden Stichproben zusammengefasst und bilden nun eine neue Gesamtstichprobe von  $60 + 18 = 78$  Vorhaben. Abschließend werden ausgehend von dieser Gesamtstichprobe unter Verwendung der üblichen Formeln der prognostizierte Fehler und die Genauigkeit der Hochrechnung neu berechnet.

### 8.3 Stichprobennahme im Laufe des Jahres

Die Prüfbehörde kann das Stichprobenverfahren in mehreren Zeiträumen während des Jahres (normalerweise zwei Halbjahre) durchführen. Die erste Prüfung bezieht sich auf die Vorhaben und Ausgaben im Zeitraum 1.1.xx-30.6.xx, die zweite auf die Vorhaben und Ausgaben im Folgehalbjahr, d. h. vom 1.7.xx-31.12.xx. Diese Methode sollte nicht zur Verringerung des Gesamtstichprobenumfangs verwendet werden. Die Summe der Stichprobenumfänge für die verschiedenen Zeiträume ist nämlich meist höher als der Stichprobenumfang bei einer einzigen Stichprobe am Jahresende. Andererseits fällt die Summe der Stichprobenumfänge nicht dramatisch höher aus, wenn die Berechnungen auf realistischen Annahmen beruhen. Der wichtigste Vorteil dieses Ansatzes steht nicht im Zusammenhang mit einer Reduzierung des Stichprobenumfangs, sondern mit der Möglichkeit, den Prüfungsaufwand über das Jahr zu verteilen und somit den Aufwand zu verringern, der bei nur einer Beobachtung am Ende des Jahres entstünde.

Diese Methode setzt voraus, dass im ersten Beobachtungszeitraum Annahmen für die nachfolgenden Beobachtungszeiträume getroffen werden (normalerweise für das nächste Halbjahr). Eventuell muss der Prüfer die erwarteten Gesamtausgaben für die Grundgesamtheit im nächsten Halbjahr abschätzen. Angesichts der möglichen Ungenauigkeiten bei den Annahmen für die nachfolgenden Zeiträume ist dieser Ansatz nicht ohne Risiko. Wenn die Merkmale der Grundgesamtheit in den Folgezeiträumen wesentlich von den Annahmen abweichen, muss der Stichprobenumfang für den Folgezeitraum möglicherweise erhöht werden, so dass der Gesamtstichprobenumfang (für alle Zeiträume) über das erwartete und geplante Maß hinausgeht.

Abschnitt 7 dieses Leitfadens enthält spezifische Formeln und ausführliche Hinweise für die Stichprobenahme in zwei Beobachtungszeiträumen innerhalb eines Jahres. Dieser Ansatz kann im Übrigen für jedes beliebige Stichprobenverfahren (einschließlich eventueller Schichtung) verwendet werden, für das sich der Prüfer entscheidet. Es ist auch zulässig, die unterschiedlichen Zeiträume des Jahres als unterschiedliche Grundgesamtheiten für die Stichprobenplanung und -ziehung zu behandeln.<sup>22</sup> Darauf wird in den methodischen Darlegungen in Abschnitt 7 nicht eingegangen, da die Anwendung eindeutig ist und die Standardformeln für die unterschiedlichen Stichprobenverfahren zum Einsatz kommen. Der einzig notwendige Zusatzschritt bei diesem Ansatz ist die Addition der prognostizierten Teilfehler am Ende des Jahres.

Die Prüfbehörde sollte nach Möglichkeit in einem gegebenen Bezugsjahr stets ein und dasselbe Stichprobenverfahren verwenden. Die Verwendung unterschiedlicher Stichprobenverfahren innerhalb eines Bezugsjahres erfordert komplexere Formeln zur Berechnung des Fehlers für dieses Jahr und ist daher nicht ratsam. Es könnten

---

<sup>22</sup> Dadurch sind die Stichprobenumfänge natürlich größer als bei dem in Abschnitt 7 beschriebenen Ansatz.

Gesamtwerte für die Genauigkeit ermittelt werden, sofern in dem gegebenen Bezugsjahr statistische Stichprobenverfahren zur Anwendung kamen. Diese komplexeren Formeln sind aber im vorliegenden Dokument nicht enthalten. Nutzt die Prüfbehörde dennoch innerhalb eines Jahres unterschiedliche Stichprobenverfahren, so sollte sie einschlägige Experten hinzuziehen, um eine korrekte Berechnung der prognostizierten Fehlerquote zu gewährleisten.

#### **8.4 Änderung des Stichprobenverfahrens während des Programmzeitraums**

Hält die Prüfbehörde das zunächst gewählte Verfahren nicht mehr für das angemessenste, kann sie eine Änderung des Verfahrens beschließen. Dies ist der Kommission im jährlichen Kontrollbericht oder in einer überarbeiteten Prüfstrategie mitzuteilen.

#### **8.5 Fehlerquoten**

Bei den in Abschnitt 7 vorgestellten Formeln und Methoden zur Ermittlung des prognostizierten Fehlers und der entsprechenden Genauigkeit werden die Fehler in Form von Geldeinheiten ausgedrückt, d. h. als Differenz zwischen dem Buchwert der Grundgesamtheit (geltend gemachte Ausgaben) und dem korrekten/geprüften Buchwert. Allgemein üblich ist es jedoch, die Ergebnisse in Form von Fehlerquoten darzustellen, da diese intuitiv interpretierbar und damit eingängiger sind. Die Umwandlung von Fehlern in Fehlerquoten ist unkompliziert und bei allen Stichprobenverfahren gleich.

Die prognostizierte Fehlerquote ist gleich dem prognostizierten Fehler dividiert durch den Buchwert der Grundgesamtheit:

$$EER = \frac{EE}{BV}$$

Die Genauigkeit der Schätzung der Fehlerquote ist gleich der Genauigkeit des prognostizierten Fehlers dividiert durch den Buchwert:

$$SER = \frac{SE}{BV}$$

#### **8.6 Zweistufiges Stichprobenverfahren**

Generell gilt, dass für alle mit der Stichprobe ausgewählten Vorhaben auch alle der Kommission gemeldeten Ausgaben geprüft werden müssen. Wenn jedoch für die ausgewählten Vorhaben sehr viele Zahlungsanträge oder Rechnungen vorliegen, können

sie anhand von Stichproben geprüft werden, wobei die Zahlungsanträge/Rechnungen nach denselben Prinzipien auszuwählen sind wie die Vorhaben. Dies entspricht dem sogenannten zweistufigen Stichprobenverfahren. Bei diesem Ansatz ist das angewandte Stichprobenverfahren im Prüfbericht oder in den Arbeitsunterlagen festzuhalten.

In diesem Fall müssen innerhalb der einzelnen Vorhaben angemessene Stichprobenumfänge berechnet werden. Eine genaue Bestimmung des Stichprobenumfangs für zweistufige Stichprobenverfahren würde hier zu weit gehen, ähnelt jedoch in allen Punkten der Strategie, die für die Auswahl der Vorhaben bei der ersten Stufe der Stichprobenziehung verwendet wurde. Mit anderen Worten: Der Stichprobenumfang der Rechnungen bzw. Zahlungsanträge kann mit denselben Formeln berechnet werden, die erforderlich sind, um die Ausgaben für das Vorhaben mit einer bestimmten geplanten Genauigkeit zu prognostizieren (z. B. 2 % der betreffenden Ausgaben). Ungeachtet des Stichprobenverfahrens, anhand dessen die Stichprobenumfänge bestimmt werden, gilt als grobe Faustregel, dass niemals ein Stichprobenumfang von weniger als 30 Beobachtungen (d. h. Rechnungen oder Zahlungsanträge von Begünstigten) verwendet werden sollte.

Am leichtesten lässt sich die Stichprobe auf der zweiten Stufe (Zahlungsanträge oder Rechnungen) durch ein einfaches Zufallsstichprobenverfahren<sup>23</sup> bilden (Formel zur Bestimmung des Stichprobenumfangs siehe Abschnitt 7.1.1). Wird die Stichprobe mit gleichen Wahrscheinlichkeiten ausgewählt, ist die Hochrechnung völlig unkompliziert, da die Fehlerquote der Stichprobenfehler als Schätzwert für die Fehlerquote des Vorhabens verwendet werden kann. Sobald die Fehlerquoten für die einzelnen Vorhaben in der Stichprobe ermittelt worden sind, erfolgt die Hochrechnung auf die Grundgesamtheit wie üblich (als wären die Gesamtausgaben des Vorhabens beobachtet worden).

Beispiel: Bei einem Vorhaben in der Stichprobe belaufen sich die Ausgaben auf 2 500 000 EUR, und es liegen 400 Rechnungen vor. Soll eine Stichprobe von 40 Rechnungen ausgewählt werden, so werden zunächst der Gesamtfehler und die Gesamtausgaben der Stichprobe berechnet. Geht man z. B. von Gesamtausgaben in Höhe von 290 000 EUR und einem Gesamtfehler von 9280 EUR aus, beträgt die geschätzte Fehlerquote für das Vorhaben  $3,2 \% = 9280 \text{ EUR} \div 290\,000 \text{ EUR}$ , und die prognostizierte Fehlerquote für das Vorhaben liegt bei  $80\,000 \text{ EUR} = 3,2 \% \times 2\,500\,000 \text{ EUR}$ .

---

<sup>23</sup> Die Prüfbehörde kann nach eigenem Ermessen komplexere Methoden für die Auswahl der Zahlungsanträge/Rechnungen zu den Vorhaben anwenden. Die Schichtung der Zahlungsanträge/Rechnungen nach Ausgabenniveau oder die Wahl nach Wahrscheinlichkeit proportional zur Größe (wie beim MUS) sind mögliche Beispiele.

## 8.7 Neuberechnung des Konfidenzniveaus

Wenn der Prüfer nach erfolgter Prüfung feststellt, dass der prognostizierte Fehler unterhalb und die Obergrenze oberhalb der Erheblichkeitsschwelle liegt, kann er eine Neuberechnung des Konfidenzniveaus vornehmen, um eindeutige Ergebnisse zu erhalten (d. h. einen prognostizierten Fehler und eine Obergrenze, die beide unterhalb der Erheblichkeitsschwelle liegen).

Ist dieses neuberechnete Konfidenzniveau mit der Bewertung der Qualität der Verwaltungs- und Kontrollsysteme vereinbar (siehe Tabelle in Abschnitt 3.2), kann sogar ohne weitere Prüfschritte mit Sicherheit geschlussfolgert werden, dass die Grundgesamtheit keine wesentlichen Fehlangaben enthält. Nur wenn die neuberechnete Konfidenz nicht akzeptabel ist (d. h. nicht mit der Bewertung der Systeme vereinbar ist), müssen die in Abschnitt 5.11 vorgeschlagenen zusätzlichen Arbeiten durchgeführt werden.

Die Neuberechnung des Konfidenzintervalls wird wie folgt vorgenommen:

- Die Erheblichkeitsschwelle wird wertmäßig berechnet, d. h. die Erheblichkeitsschwelle (2 %) wird mit dem Gesamtbuchwert der Grundgesamtheit multipliziert.
- Der prognostizierte Fehler (EE) wird von der Erheblichkeitsschwelle subtrahiert.
- Das Ergebnis wird durch die Genauigkeit der Hochrechnung (SE) dividiert. Die Genauigkeit ist vom Stichprobenverfahren abhängig und wird in den Abschnitten zu den einzelnen Verfahren behandelt.
- Das vorstehende Ergebnis wird mit dem Parameter  $z$  multipliziert, der sowohl für die Berechnung des Stichprobenumfangs als auch die Berechnung der Genauigkeit verwendet wird. Daraus ergibt sich der neue Wert  $z^*$ :

$$z^* = z \times \frac{(0.02 \times BV) - EE}{SE}$$

- Das mit diesem neuen Parameter ( $z^*$ ) verbundene Konfidenzniveau wird einer Tabelle der Normalverteilung (im Anhang) entnommen. Alternativ kann die folgende Excel-Formel verwendet werden: „=1-(1-NORMSDIST( $z^*$ ))\*2“.

Beispiel: Nach Prüfung einer Grundgesamtheit mit einem Buchwert von 1 858 233 036 EUR und einem Konfidenzniveau von 90 % (entsprechend  $z = 1.645$ , vgl. Abschnitt 6.4) liegen folgende Ergebnisse vor:

Merkm <del>al</del>	Wert
BV	1 858 233 036 EUR
Erheblichkeitsschwelle (2 % des BV)	37 164 661 EUR
Prognostizierter Fehler (EE)	14 568 765 EUR (0,8 %)
Genauigkeit (SE)	26 195 819 EUR (1,4 %)
Obere Fehlergrenze (ULE)	40 764 584 EUR (2,2 %)

Der neue Parameter  $z^*$  errechnet sich wie folgt:

$$z^* = 1.645 \times \frac{37,164,661\text{€} - 14,568,765\text{€}}{26,195,819\text{€}} = 1.419$$

Mithilfe der MS-Excel-Funktion „=1-(1-NORMSDIST(1.419))\*2“ wird das neue Konfidenzniveau abgeleitet, nämlich 84,4 %.

Wenn das Neuberechnete Konfidenzniveau mit der Einschätzung der Qualität der Verwaltungs- und Kontrollsysteme vereinbar ist, kann geschlussfolgert werden, dass die Grundgesamtheit keine wesentlichen Fehlangaben enthält.

## 8.8 Stichprobenverfahren für Systemprüfungen

### 8.8.1 Einleitung

In Artikel 62 der Verordnung (EG) Nr. 1083/2006 des Rates heißt es: „Die für ein operationelles Programm zuständige Prüfbehörde hat insbesondere die Aufgabe: (a) zu gewährleisten, dass das effektive Funktionieren des Verwaltungs- und Kontrollsystems für das operationelle Programm geprüft wird.“ Diese Prüfungen sind sogenannte Systemprüfungen. Anhand der Systemprüfungen wird die Wirksamkeit der Kontrollen des Verwaltungs- und Kontrollsystems untersucht und eine Schlussfolgerung über das Niveau der zu erwartenden Zuverlässigkeit des Systems gezogen. Die Entscheidung über die Verwendung oder Nichtverwendung eines statistischen Stichprobenverfahrens für die Überprüfungen von Kontrollen unterliegt dem beruflichen Urteilsvermögen, unter den jeweiligen Umständen die wirksamste Methode zum Erzielen hinreichend geeigneter Prüfnachweise zu ermitteln.

Da die vom Prüfer durchgeführte Analyse der Art und der Ursache der Fehler ebenso wichtig ist wie das Nichtvorhandensein bzw. das Vorhandensein von Fehlern, kann auch ein nichtstatistischer Ansatz geeignet sein. In diesem Fall kann der Prüfer für jede wichtige Kontrollvariable einen festen Stichprobenumfang aus den zu prüfenden Elementen festlegen. Dennoch obliegt es dem Prüfer, anhand seines beruflichen

Urteilsvermögens aus den in Frage kommenden Faktoren jene von Relevanz<sup>24</sup> auszuwählen. Bei Verwendung eines nichtstatistischen Ansatzes können die Ergebnisse nicht extrapoliert werden.

Das System der Attributstichproben ist ein statistischer Ansatz, mit dessen Hilfe der Prüfer die Sicherheit des Systems ermitteln und die erwartete Fehlerquote einer Stichprobe abschätzen kann. Am häufigsten wird diese Methode bei Prüfungen eingesetzt, wenn die Fehlerquote einer vorgeschriebenen Kontrolle mit dem Ziel zu untersuchen ist, das vom Prüfer geschätzte Niveau des Kontrollrisikos zu untermauern. Die Ergebnisse können anschließend auf die Grundgesamtheit hochgerechnet werden.

Da es sich beim Attributstichprobenverfahren um eine allgemeine Methode mit mehreren Varianten handelt, wird sie als grundlegende statistische Methode bei Systemprüfungen angewendet. Sonstige Methoden, die auch bei Systemprüfungen angewendet werden können, beruhen auf den nachstehend erläuterten Konzepten.

Das Attributstichprobenverfahren verarbeitet binäre Aufgaben, wie z. B. ja oder nein, hoch oder niedrig, wahre oder falsche Antworten. Anhand dieser Methode werden die Daten der Stichprobe auf die Grundgesamtheit hochgerechnet, um daraus abzuleiten, ob die Grundgesamtheit zu der einen oder zu der anderen Kategorie gehört.

Die Verordnung schreibt für Stichprobenkontrollen im Rahmen einer Systemprüfung nicht die Anwendung eines statistischen Ansatzes vor. Daher dienen diese Kapitel und die entsprechenden Anhänge allgemeinen Informationszwecken und werden nicht weiter ausgeführt.

Weitere Informationen und Beispiele im Zusammenhang mit Stichprobenverfahren für Systemprüfungen finden Sie in der Fachliteratur über Stichprobenverfahren zu Prüfzwecken.

Beim Attributstichprobenverfahren im Rahmen einer Systemprüfung ist der folgende allgemeine sechsstufige Ablauf zu beachten:

1. Festlegung der Prüfziele: z. B. Bestimmung, ob die Fehlerquote in einer Grundgesamtheit den Kriterien für eine hohe Sicherheit entspricht;
2. Festlegung der Grundgesamtheit und der Stichprobeneinheit: z. B. die einem Programm zugeordneten Rechnungen;
3. Festlegung des Abweichungszustands: Hierbei handelt es sich um das zu beurteilende Attribut, etwa das Vorhandensein einer Unterschrift auf den einem Vorhaben innerhalb eines Programms zugeordneten Rechnungen;
4. Festlegung des Stichprobenumfangs gemäß der nachstehend aufgeführten Formel;

---

<sup>24</sup> Weiterführende Erklärungen oder Beispiele können dem „Audit Guide on Sampling“ des American Institute of Certified Public Accountants vom 1.4.2001 entnommen werden.

5. Auswahl der Stichprobe und Durchführung der Prüfung (Auswahl nach dem Zufallsprinzip);
6. Bewertung und Dokumentierung der Ergebnisse.

### 8.8.2 Stichprobenumfang

Die Berechnung des Stichprobenumfangs  $n$  im Rahmen des Attributstichprobenverfahrens stützt sich auf folgende Angaben:

- das Konfidenzniveau und den damit verwandten Koeffizienten  $z$  aus einer Normalverteilung (siehe Abschnitt 6.4);
- die vom Prüfer bestimmte maximal zulässige Abweichungsquote  $T$ , wobei die zulässigen Schwellenwerte von der Prüfbehörde des Mitgliedstaats festgelegt werden (z. B. Zahl der fehlenden Unterschriften auf Rechnungen – unterhalb dieses Schwellenwertes geht der Prüfer nicht von einem Fehler aus);
- die voraussichtliche Abweichungsquote innerhalb der Grundgesamtheit,  $p$ , die geschätzt oder aus einer früheren Stichprobe übernommen wird. Die zulässige Abweichungsquote sollte über der erwarteten Abweichungsquote innerhalb der Grundgesamtheit liegen, da die Prüfung andernfalls zwecklos wäre (wird eine Fehlerquote von 10 % erwartet, so ist es unsinnig, die zulässige Fehlerquote auf 5 % festzulegen, da in der Grundgesamtheit mehr Fehler erwartet werden, als man bereit ist zu tolerieren).

Der Stichprobenumfang wird wie folgt berechnet:

$$n = \frac{z^2 \times p \times (1 - p)}{T^2}$$

Beispiel: Bei einem angenommenen Konfidenzniveau von 95 % ( $z = 1.96$ ), einer zulässigen Abweichungsquote ( $T$ ) von 12 % und einer erwarteten Abweichungsquote ( $p$ ) der Grundgesamtheit von 6 % läge die Mindestgröße der Stichprobe bei

$$n = \frac{1.96^2 \times 0.06 \times (1 - 0.06)}{0.12^2} \approx 16.$$

Hinweis: Die Größe der Grundgesamtheit hat keinen Einfluss auf den Stichprobenumfang; in der vorstehenden Formel wird der erforderliche Stichprobenumfang für kleine Grundgesamtheiten etwas zu hoch veranschlagt, was jedoch akzeptabel ist. Um den erforderlichen Stichprobenumfang zu verringern, bietet sich die Herabsetzung des Konfidenzniveaus (d. h. Erhöhung des Risikos, das Kontrollrisiko zu gering einzuschätzen) bzw. die Erhöhung der zulässigen Abweichungsquote an.

### 8.8.3 Extrapolation

Die Zahl der in der Stichprobe festgestellten Abweichungen dividiert durch die Zahl der Elemente in der Stichprobe (also den Stichprobenumfang) ergibt die Abweichungsquote der Stichprobe:

$$EDR = \frac{\text{Anzahl der Abweichungen in der Stichprobe}}{n}$$

Dies ist zugleich der bestmögliche Schätzwert für die hochgerechnete Abweichungsquote (**EDR**), der aus der Stichprobe abgeleitet werden kann.

### 8.8.4 Genauigkeit

Genauigkeit (Stichprobenfehler) ist das Maß der mit der Hochrechnung (Extrapolation) verbundenen Unsicherheit. Für die Genauigkeit gilt die Formel

$$SE = z \times \frac{p_s \times (1 - p_s)}{\sqrt{n}}$$

Dabei ist  $p_s$  die Zahl der in der Stichprobe festgestellten Abweichungen dividiert durch den Stichprobenumfang, also die Abweichungsquote der Stichprobe.

### 8.8.5 Bewertung

Die ermittelte obere Abweichungsschwelle ist eine theoretische Zahl, die vom Stichprobenumfang und der Zahl der entdeckten Fehler abhängt:

$$ULD = EDR + SE.$$

Sie stellt die maximale Fehlerquote der Grundgesamtheit bei dem festgelegten Konfidenzniveau dar und ergibt sich aus binomialen Tabellen (so beträgt etwa bei einer Stichprobenumfang von 150 und 3 festgestellten Abweichungen (Abweichungsquote der Stichprobe von 2 %) die maximale Abweichungsquote (oder ermittelte obere Abweichungsschwelle) bei einem Konfidenzniveau von 95 %:

$$ULD = \frac{3}{150} + 1.96 \times \frac{\frac{3}{150} \times (1 - \frac{3}{150})}{\sqrt{150}} = 0.023.$$

Liegt dieser Wert über der zulässigen Abweichungsquote, dann ist die Stichprobe mit der erwarteten Fehlerquote der Grundgesamtheit bei dem gewählten Konfidenzniveau

unvereinbar. Die logische Schlussfolgerung ist daher, dass die Grundgesamtheit nicht das festgelegte Kriterium für hohe Sicherheit erfüllt und daher als durchschnittlich oder wenig sicher einzustufen ist. Hinweis: Die Schwellenwerte für niedrige, durchschnittliche und hohe Sicherheit werden von der Prüfbehörde festgelegt.

### **8.8.6 Spezielle Methoden des Attributstichprobenverfahrens**

Das Attributstichprobenverfahren ist ein allgemeines Verfahren. Deshalb wurden für bestimmte Zwecke einige Varianten entwickelt. Hier tragen die Entdeckungstichprobe und das schrittweise Stichprobenverfahren speziellen Erfordernissen Rechnung.

Die Entdeckungstichprobe kommt dann zum Einsatz, wenn ein einziger Fehler kritisch wäre; sie ist daher zum Teil auf die Aufdeckung von Betrugsfällen oder Fällen, in denen Kontrollen umgangen wurden, gerichtet. Bei der auf dem Attributstichprobenverfahren beruhenden Methode wird von einer Nullfehlerquote (oder zumindest einer sehr geringen Fehlerquote) ausgegangen. Sie ist daher für die Hochrechnung etwaiger Fehler in der Stichprobe auf die Grundgesamtheit kaum geeignet. Der Prüfer kann beim Entdeckungstichprobenverfahren anhand der Stichprobe darauf schließen, ob die angenommene geringe Fehlerquote oder Nullfehlerquote für die Grundgesamtheit zutrifft. Das Verfahren eignet sich nicht zur Beurteilung der Sicherheit interner Kontrollen und ist daher nicht auf Systemprüfungen anwendbar.

Das schrittweise Stichprobenverfahren wurde entwickelt, weil es häufig notwendig ist, den Stichprobenumfang so weit wie möglich zu reduzieren. Ziel ist es, durch Prüfung möglichst weniger Elemente zu dem Schluss zu gelangen, dass die Fehlerquote der Grundgesamtheit bei einem gegebenen Konfidenzniveau unter einem festgelegten Schwellenwert liegt. Sobald das erwartete Ergebnis erreicht wurde, wird das Stichprobenverfahren abgebrochen. Auch dieses Verfahren ist für die Hochrechnung der Ergebnisse auf die Grundgesamtheit kaum geeignet, wenngleich es für die Beurteilung der Schlussfolgerungen einer Systemprüfung von Nutzen sein kann. So kann es etwa angewandt werden, wenn das Ergebnis einer Systemprüfung bezweifelt wird und überprüft werden muss, ob das Kriterium für das gegebene Sicherheitsniveau tatsächlich erfüllt wird.

## **Anhang 1 – Hochrechnung von Zufallsfehlern bei Feststellung systembedingter Fehler**

### **1. Einleitung**

In diesem Anhang wird die Berechnung prognostizierter Zufallsfehler im Falle der Feststellung systembedingter Fehler erläutert. Im Leitfaden (vgl. Abschnitt 4.1) wurde dazu bereits Folgendes ausgeführt: Wird ein potenzieller systembedingter Fehler festgestellt, so sind weitere Arbeiten erforderlich, um die gesamte Reichweite des Fehlers zu ermitteln und eine Quantifizierung vorzunehmen. Das bedeutet, dass alle Situationen, in denen ein Fehler derselben Art wie in der Zufallsstichprobe auftreten könnte, ermittelt werden müssen, damit die Gesamtwirkung dieses Fehlers auf die Grundgesamtheit eingegrenzt werden kann. Kann eine solche Eingrenzung nicht vor der Übermittlung des jährlichen Kontrollberichts erfolgen, so sind die systembedingten Fehler – für die Berechnung des prognostizierten Zufallsfehlers – als Zufallsfehler zu behandeln.

Der prognostizierte Gesamtfehler (TPE) entspricht der Summe folgender Fehler: prognostizierte Zufallsfehler, systembedingte Fehler und nicht berichtigte anomale Fehler.

Bei der Hochrechnung der in der Stichprobe gefundenen Zufallsfehler auf die Grundgesamtheit muss die Prüfbehörde also die Menge des systembedingten Fehlers vom Buchwert (im Bezugsjahr bescheinigte Gesamtausgaben) abziehen, wenn dieser Wert Bestandteil der Hochrechnungsformel ist, wie nachstehend erläutert wird.

Im Falle der Schätzung des Mittelwerts pro Einheit<sup>25</sup> und der Differenzschätzung bleiben die im Leitfaden angegebenen Formeln für die Hochrechnung der Zufallsfehler unverändert. Für das wertbezogene Stichprobenverfahren (MUS) werden in diesem Anhang zwei mögliche Ansätze beschrieben (einer davon mit unveränderter Formel, der andere mit komplexeren Formeln und höherer Genauigkeit). Bei der Verhältnisschätzung wird zur Hochrechnung der Zufallsfehler und zur Berechnung der Genauigkeit (SE) der Gesamtbuchwert nach Abzug der systembedingten Fehler benötigt.

Für alle statistischen Stichprobenverfahren gilt Folgendes: Sind systembedingte Fehler oder anomale nicht berichtigte Fehler vorhanden, dann entspricht die obere Fehlergrenze (ULE) der Summe aus TPE und Genauigkeit (SE). Sind lediglich Zufallsfehler vorhanden, entspricht die ULE der Summe aus prognostizierten Zufallsfehlern und Genauigkeit.

---

<sup>25</sup> Vgl. Abschnitt „Einfaches Zufallsstichprobenverfahren“ in diesen Leitlinien.

In den folgenden Abschnitten werden für die wichtigsten Stichprobenverfahren ausführlichere Erläuterungen zur Hochrechnung von Zufallsfehlern bei Vorhandensein systembedingter Fehler gegeben.

## 2. Einfache Zufallsstichprobenverfahren

### 2.2 Schätzung des Mittelwerts pro Einheit

Die Hochrechnung der Zufallsfehler und die Berechnung der Genauigkeit erfolgen wie üblich:

$$EE_1 = N \times \frac{\sum_{i=1}^n E_i}{n}$$

$$SE_1 = N \times z \times \frac{s_e}{\sqrt{n}}$$

Darin sind  $E_i$  die Menge der festgestellten Zufallsfehler in den einzelnen Stichprobeneinheiten und  $s_e$  wie gewöhnlich die Standardabweichung der Zufallsfehler in der Stichprobe.

Der prognostizierte Gesamtfehler entspricht der Gesamtheit aller prognostizierten Zufallsfehler, systembedingten Fehler und nicht berichtigten anomalen Fehler.

Die obere Fehlergrenze (ULE) entspricht der Summe aus dem prognostizierten Gesamtfehler  $TPE$  und der Genauigkeit der Extrapolation:

$$ULE = TPE + SE$$

### 2.3 Verhältnisschätzung

Der Zufallsfehler wird wie folgt hochgerechnet:

$$EE_2 = BV' \times \frac{\sum_{i=1}^n E_i}{\sum_{i=1}^n BV'_i}$$

Dabei ist  $BV'$  der Gesamtbuchwert der Grundgesamtheit nach Abzug der zuvor ermittelten systembedingten Fehler:  $BV' = BV - \text{systembedingte Fehler}$ .  $BV'_i$  ist der Buchwert der Einheit  $i$  nach Abzug der Menge des systembedingten Fehlers bei dieser Einheit.

Die Stichprobenfehlerquote in der obigen Formel errechnet sich durch einfache Division des Gesamtbetrags des Zufallsfehlers in der Stichprobe durch den Gesamtbetrag der Ausgaben (nach Abzug der systembedingten Fehler) der Einheiten in der Stichprobe (geprüfte Ausgaben).

Für die Genauigkeit gilt die Formel

$$SE_2 = N \times z \times \frac{s_{q'}}{\sqrt{n}}$$

wobei  $s_{q'}$  die Stichprobenstandardabweichung der Variablen  $q'$  ist:

$$q'_i = E_i - \frac{\sum_{i=1}^n E_i}{\sum_{i=1}^n BV'_i} \times BV'_i.$$

Diese Variable berechnet sich für jede Einheit in der Stichprobe als Differenz zwischen ihrem Zufallsfehler und dem Produkt aus ihrem Buchwert (nach Abzug der systembedingten Fehler) und der Fehlerquote in der Stichprobe.

Der prognostizierte Gesamtfehler entspricht der Gesamtheit aller prognostizierten Zufallsfehler, systembedingten Fehler und nicht berichtigten anomalen Fehler.

Die obere Fehlergrenze (ULE) entspricht der Summe aus dem prognostizierten Gesamtfehler  $TPE$  und der Genauigkeit der Extrapolation:

$$ULE = TPE + SE$$

### 3. Differenzenschätzung

Der prognostizierte Zufallsfehler auf Ebene der Grundgesamtheit kann wie üblich berechnet werden, indem der in der Stichprobe beobachtete durchschnittliche Zufallsfehler pro Vorhaben mit der Zahl der Vorhaben in der Grundgesamtheit multipliziert wird, woraus sich der prognostizierte Fehler ergibt:

$$EE = N \times \frac{\sum_{i=1}^n E_i}{n}$$

In einem zweiten Schritt erfolgt die Berechnung des prognostizierten Gesamtfehlers (TPE). Dazu wird der Betrag des systembedingten Fehlers und der nicht berichtigten anomalen Fehler zum prognostizierten Zufallsfehler (EE) hinzuaddiert.

Zur Hochrechnung auf den korrekten Buchwert (korrekter Ausgabenbetrag, der festgestellt werden könnte, wenn alle Vorhaben in der Grundgesamtheit geprüft würden) wird der prognostizierte Fehler (EE) vom Buchwert (BV) der Grundgesamtheit (geltend gemachte Ausgaben ohne Abzug der systembedingten Fehler) abgezogen. Die Hochrechnung auf den korrekten Buchwert (CBV) erfolgt nach folgender Formel:

$$CBV = BV - TPE$$

Für die Genauigkeit der Hochrechnung gilt wie üblich

$$SE = N \times z \times \frac{s_e}{\sqrt{n}}$$

wobei  $s_e$  die Standardabweichung der Zufallsfehler in der Stichprobe ist.

Um Rückschlüsse auf die Erheblichkeit der Fehler zu ermöglichen, wird zunächst die Untergrenze für den korrigierten Buchwert berechnet. Diese Untergrenze entspricht

$$LL = CBV - SE$$

Die Hochrechnungen auf den korrekten Buchwert und die Obergrenze sind jeweils mit der Differenz zwischen dem Buchwert (geltend gemachte Ausgaben) und dem maximal zulässigen Fehler (TE) zu vergleichen, die dem Produkt aus Erheblichkeitsschwelle und Buchwert entspricht:

$$BV - TE = BV - 2\% \times BV = 98\% \times BV$$

Die Fehlerbewertung sollte nach Maßgabe von Abschnitt 7.2.1.5 des Leitfadens erfolgen.

#### 4. Wertbezogenes Stichprobenverfahren – Standardansatz

Beim MUS-Verfahren gibt es zwei mögliche Ansätze für die Hochrechnung von Zufallsfehlern und die Berechnung der Genauigkeit bei Vorliegen systembedingter Fehler. Sie werden hier als *MUS-Standardansatz* bzw. *MUS-Verhältnisschätzung* bezeichnet. Die letztgenannte Methode erfordert komplexere Berechnungen. Zwar kommen bei jedem Szenario beide Methoden in Frage, doch liefert die zweite Methode präzisere Ergebnisse, wenn die Zufallsfehler stärker mit den um den systembedingten

Fehler berichtigten Buchwerten korrelieren als mit den ursprünglichen Buchwerten. Bei einem geringen Anteil systembedingter Fehler in der Grundgesamtheit dagegen fällt der mit der zweiten Methode erzielte Genauigkeitsgewinn meist sehr bescheiden aus, so dass in diesem Fall die erstgenannte Methode wegen der einfacheren Anwendung zu bevorzugen ist.

#### 4.1 MUS-Standardansatz

Die Hochrechnung der Zufallsfehler und die Berechnung der Genauigkeit werden wie gewöhnlich vorgenommen.

Die Hochrechnung der Zufallsfehler auf die Grundgesamtheit erfolgt für die Einheiten in der umfassenden Schicht anders als für die Einheiten in der nicht umfassenden Schicht.

Bei der umfassenden Schicht, d. h. der Schicht mit den Stichprobenelementen, deren Buchwert größer ist als der Schwellenbetrag ( $BV_i > \frac{BV}{n}$ ), entspricht der prognostizierte Fehler einfach der Summe der Fehler, die bei den Elementen in dieser Schicht festgestellt werden:

$$EE_s = \sum_{i=1}^{n_s} E_i$$

Bei der nicht umfassenden Schicht, d. h. der Schicht mit den Stichprobenelementen, deren Buchwert kleiner oder gleich dem Schwellenbetrag ist ( $BV_i \leq \frac{BV}{n}$ ), errechnet sich der prognostizierte Zufallsfehler wie folgt:

$$EE_s = \frac{BV_s}{n_s} \sum_{i=1}^{n_s} \frac{E_i}{BV_i}$$

Zu beachten ist, dass es sich bei den Buchwerten in der obigen Formel um die Ausgaben **ohne** Abzug des Betrags des systembedingten Fehlers handelt. Die Fehlerquoten  $\frac{E_i}{BV_i}$  werden also ausgehend von den Gesamtausgaben der Stichprobenelemente berechnet, unabhängig davon, ob in den einzelnen Einheiten systembedingte Fehler gefunden wurden oder nicht.

Für die Genauigkeit gilt ebenfalls die übliche Formel, nämlich

$$SE = z \times \frac{BV_s}{\sqrt{n_s}} \times z_r$$

wobei  $s_r$  die Standardabweichung der Zufallsfehlerquoten in der Stichprobe der nicht umfassenden Schicht ist. Auch diese Fehlerquoten werden ausgehend von den ursprünglichen Buchwerten  $BV_i$  ohne Abzug des Betrags des systembedingten Fehlers berechnet.

Der prognostizierte Gesamtfehler entspricht der Gesamtheit aller prognostizierten Zufallsfehler, systembedingten Fehler und nicht berichtigten anomalen Fehler.

Die obere Fehlergrenze (ULE) entspricht der Summe aus dem prognostizierten Gesamtfehler  $TPE$  und der Präzision der Hochrechnung:

$$ULE = TPE + SE$$

#### 4.2 MUS-Verhältnisschätzung

Auch hier erfolgt die Hochrechnung der Zufallsfehler auf die Grundgesamtheit bei den Elementen in der umfassenden Schicht anders als bei den Elementen in der nicht umfassenden Schicht.

Bei der umfassenden Schicht, d. h. der Schicht mit den Stichprobenelementen, deren Buchwert größer ist als der Schwellenbetrag ( $BV_i > \frac{BV}{n}$ ), entspricht der prognostizierte Fehler einfach der Summe der Zufallsfehler, die bei den Elementen in dieser Schicht festgestellt werden:

$$EE_s = \sum_{i=1}^{n_s} E_i$$

Bei der nicht umfassenden Schicht, d. h. der Schicht mit den Stichprobenelementen, deren Buchwert kleiner oder gleich dem Schwellenbetrag ist ( $BV_i \leq \frac{BV}{n}$ ), errechnet sich der prognostizierte Zufallsfehler wie folgt:

$$EE_s = BV'_s \times \frac{\sum_{i=1}^{n_s} E_i}{\sum_{i=1}^{n_s} \frac{BV'_i}{BV_i}}$$

Dabei ist  $BV'_s$  der Gesamtbuchwert der niedrigwertigen Schicht nach Abzug der systembedingten Fehler, die zuvor in derselben Schicht abgegrenzt wurden,  $BV'_s = BV_s - \text{systembedingte Fehler in der Stichprobenschicht}$ .  $BV'_i$  ist der Buchwert der Einheit  $i$  nach Abzug des systembedingten Fehlers bei dieser Einheit.

Für die Genauigkeit gilt die Formel:

$$SE = z \times \frac{BV_s}{\sqrt{n_s}} \times s_{rq}$$

wobei  $s_{rq}$  die Standardabweichung der Fehlerquoten für den **transformierten Fehler**  $q'$  ist. Für die Berechnung dieser Formel müssen zunächst die Werte der **transformierten Fehler** für alle Einheiten in der Stichprobe berechnet werden:

$$q'_t = E_t - \frac{\sum_{i=1}^{n_s} \frac{E_i}{EV_i}}{\sum_{i=1}^{n_s} \frac{BV_i}{EV_i}} \times BV'_t$$

Zuletzt wird die Standardabweichung der Fehlerquoten in der Stichprobe der nicht umfassenden Schicht ( $s_{rq}$ ) für den transformierten Fehler  $q'$  wie folgt berechnet:

$$s_{rq} = \sqrt{\frac{1}{n_s - 1} \sum_{i=1}^{n_s} \left( \frac{q'_i}{EV_{t_i}} - \bar{rq}_s \right)^2}$$

Darin entspricht  $\bar{rq}_s$  dem einfachen Durchschnitt der transformierten Fehlerquoten in der Stichprobe der Schicht:

$$\bar{rq}_s = \frac{\sum_{i=1}^{n_s} \frac{q'_i}{EV_i}}{n_s}$$

Der prognostizierte Gesamtfehler entspricht der Gesamtheit aller prognostizierten Zufallsfehler, systembedingten Fehler und nicht berichtigten anomalen Fehler.

Die obere Fehlergrenze entspricht der Summe aus dem prognostizierten Gesamtfehler (**TPE**) und der Genauigkeit der Hochrechnung:

$$ULE = TPE + SE$$

## 5. Nichtstatistische Stichprobenverfahren

Die Hochrechnung und Berechnung der Genauigkeit werden wie üblich vorgenommen.

Ist eine umfassende Schicht vorhanden, d. h. eine Schicht mit den Stichprobeneinheiten, deren Buchwert größer ist als der Schwellenbetrag, dann entspricht der prognostizierte Fehler einfach der Summe der in dieser Gruppe festgestellten Zufallsfehler:

$$EE_s = \sum_{i=1}^{n_s} E_i$$

Wurden die Einheiten mit gleichen Wahrscheinlichkeiten ausgewählt, dann errechnet sich der prognostizierte Zufallsfehler für die Stichprobenschicht nach der üblichen Formel

$$EE_s = N_s \frac{\sum_{i=1}^{n_s} E_i}{n_s}$$

wobei  $N_s$  die Größe der Grundgesamtheit und  $n_s$  der Stichprobenumfang in der niedrigwertigen Schicht sind.

Wurden die Einheiten mit Wahrscheinlichkeiten ausgewählt, die dem Ausgabenwert proportional sind, dann errechnet sich der prognostizierte Zufallsfehler für die niedrigwertige Schicht wie folgt:

$$EE_s = \frac{BV_s}{n_s} \sum_{i=1}^{n_s} \frac{E_i}{BV_i}$$

Darin sind  $BV_s$  der Gesamtbuchwert (**ohne** Abzug des Betrags des systembedingten Fehlers),  $BV_i$  der Buchwert der Stichprobeneinheit  $i$  (**ohne** Abzug des Betrags des systembedingten Fehlers) und  $n_s$  der Stichprobenumfang in der niedrigwertigen Schicht.

Ebenso wie beim MUS-Verfahren (vgl. Abschnitt 2.4) kann alternativ die Verhältnisschätzungsformel

$$EE_s = BV'_s \times \frac{\sum_{i=1}^{n_s} \frac{E_i}{BV_i}}{\sum_{i=1}^{n_s} \frac{BV_i}{BV'_s}}$$

verwendet werden. Erneut ist  $BV'_s$  der Gesamtbuchwert der niedrigwertigen Schicht nach Abzug der systembedingten Fehler, die zuvor in derselben Schicht abgegrenzt wurden:  $BV'_s = BV_s - \text{systembedingte Fehler in der Stichprobenschicht}$ .  $BV'_i$  ist der Buchwert der Einheit  $i$  nach Abzug des systembedingten Fehlers bei dieser Einheit.

Der prognostizierte Gesamtfehler (TPE) entspricht der Gesamtheit aller prognostizierten Zufallsfehler, systembedingten Fehler und nicht berechtigten anomalen Fehler.

## Anhang 2 – Zuverlässigkeitsfaktoren für das MUS

Zahl der Fehler	Risiko einer falschen Anerkennung									
	1 %	5 %	10 %	15 %	20 %	25 %	30 %	37 %	40 %	50 %
0	4,61	3,00	2,30	1,90	1,61	1,39	1,20	0,99	0,92	0,69
1	6,64	4,74	3,89	3,37	2,99	2,69	2,44	2,14	2,02	1,68
2	8,41	6,30	5,32	4,72	4,28	3,92	3,62	3,25	3,11	2,67
3	10,05	7,75	6,68	6,01	5,52	5,11	4,76	4,34	4,18	3,67
4	11,60	9,15	7,99	7,27	6,72	6,27	5,89	5,42	5,24	4,67
5	13,11	10,51	9,27	8,49	7,91	7,42	7,01	6,49	6,29	5,67
6	14,57	11,84	10,53	9,70	9,08	8,56	8,11	7,56	7,34	6,67
7	16,00	13,15	11,77	10,90	10,23	9,68	9,21	8,62	8,39	7,67
8	17,40	14,43	12,99	12,08	11,38	10,80	10,30	9,68	9,43	8,67
9	18,78	15,71	14,21	13,25	12,52	11,91	11,39	10,73	10,48	9,67
10	20,14	16,96	15,41	14,41	13,65	13,02	12,47	11,79	11,52	10,67
11	21,49	18,21	16,60	15,57	14,78	14,12	13,55	12,84	12,55	11,67
12	22,82	19,44	17,78	16,71	15,90	15,22	14,62	13,88	13,59	12,67
13	24,14	20,67	18,96	17,86	17,01	16,31	15,70	14,93	14,62	13,67
14	25,45	21,89	20,13	19,00	18,13	17,40	16,77	15,97	15,66	14,67
15	26,74	23,10	21,29	20,13	19,23	18,49	17,83	17,02	16,69	15,67
16	28,03	24,30	22,45	21,26	20,34	19,57	18,90	18,06	17,72	16,67
17	29,31	25,50	23,61	22,38	21,44	20,65	19,96	19,10	18,75	17,67
18	30,58	26,69	24,76	23,50	22,54	21,73	21,02	20,14	19,78	18,67
19	31,85	27,88	25,90	24,62	23,63	22,81	22,08	21,17	20,81	19,67
20	33,10	29,06	27,05	25,74	24,73	23,88	23,14	22,21	21,84	20,67
21	34,35	30,24	28,18	26,85	25,82	24,96	24,20	23,25	22,87	21,67
22	35,60	31,41	29,32	27,96	26,91	26,03	25,25	24,28	23,89	22,67
23	36,84	32,59	30,45	29,07	28,00	27,10	26,31	25,32	24,92	23,67
24	38,08	33,75	31,58	30,17	29,08	28,17	27,36	26,35	25,95	24,67
25	39,31	34,92	32,71	31,28	30,17	29,23	28,41	27,38	26,97	25,67
26	40,53	36,08	33,84	32,38	31,25	30,30	29,46	28,42	28,00	26,67
27	41,76	37,23	34,96	33,48	32,33	31,36	30,52	29,45	29,02	27,67
28	42,98	38,39	36,08	34,57	33,41	32,43	31,56	30,48	30,04	28,67
29	44,19	39,54	37,20	35,67	34,49	33,49	32,61	31,51	31,07	29,67
30	45,40	40,69	38,32	36,76	35,56	34,55	33,66	32,54	32,09	30,67
31	46,61	41,84	39,43	37,86	36,64	35,61	34,71	33,57	33,11	31,67
32	47,81	42,98	40,54	38,95	37,71	36,67	35,75	34,60	34,14	32,67
33	49,01	44,13	41,65	40,04	38,79	37,73	36,80	35,63	35,16	33,67
34	50,21	45,27	42,76	41,13	39,86	38,79	37,84	36,66	36,18	34,67
35	51,41	46,40	43,87	42,22	40,93	39,85	38,89	37,68	37,20	35,67
36	52,60	47,54	44,98	43,30	42,00	40,90	39,93	38,71	38,22	36,67
37	53,79	48,68	46,08	44,39	43,07	41,96	40,98	39,74	39,24	37,67
38	54,98	49,81	47,19	45,47	44,14	43,01	42,02	40,77	40,26	38,67
39	56,16	50,94	48,29	46,55	45,20	44,07	43,06	41,79	41,28	39,67
40	57,35	52,07	49,39	47,63	46,27	45,12	44,10	42,82	42,30	40,67
41	58,53	53,20	50,49	48,72	47,33	46,17	45,14	43,84	43,32	41,67
42	59,71	54,32	51,59	49,80	48,40	47,22	46,18	44,87	44,34	42,67
43	60,88	55,45	52,69	50,87	49,46	48,27	47,22	45,90	45,36	43,67
44	62,06	56,57	53,78	51,95	50,53	49,32	48,26	46,92	46,38	44,67
45	63,23	57,69	54,88	53,03	51,59	50,38	49,30	47,95	47,40	45,67
46	64,40	58,82	55,97	54,11	52,65	51,42	50,34	48,97	48,42	46,67
47	65,57	59,94	57,07	55,18	53,71	52,47	51,38	49,99	49,44	47,67
48	66,74	61,05	58,16	56,26	54,77	53,52	52,42	51,02	50,45	48,67
49	67,90	62,17	59,25	57,33	55,83	54,57	53,45	52,04	51,47	49,67
50	69,07	63,29	60,34	58,40	56,89	55,62	54,49	53,06	52,49	50,67

### Anhang 3 – Werte der standardisierten Normalverteilung (z)

x	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.998650	0.998694	0.998736	0.998777	0.998817	0.998856	0.998893	0.998930	0.998965	0.998999
3.1	0.999032	0.999064	0.999096	0.999126	0.999155	0.999184	0.999211	0.999238	0.999264	0.999289
3.2	0.999313	0.999336	0.999359	0.999381	0.999402	0.999423	0.999443	0.999462	0.999481	0.999499
3.3	0.999517	0.999533	0.999550	0.999566	0.999581	0.999596	0.999610	0.999624	0.999638	0.999650
3.4	0.999663	0.999675	0.999687	0.999698	0.999709	0.999720	0.999730	0.999740	0.999749	0.999758
3.5	0.999767	0.999776	0.999784	0.999792	0.999800	0.999807	0.999815	0.999821	0.999828	0.999835
3.6	0.999841	0.999847	0.999853	0.999858	0.999864	0.999869	0.999874	0.999879	0.999883	0.999888
3.7	0.999892	0.999896	0.999900	0.999904	0.999908	0.999912	0.999915	0.999918	0.999922	0.999925
3.8	0.999928	0.999930	0.999933	0.999936	0.999938	0.999941	0.999943	0.999946	0.999948	0.999950
3.9	0.999952	0.999954	0.999956	0.999958	0.999959	0.999961	0.999963	0.999964	0.999966	0.999967
4.0	0.999968	0.999970	0.999971	0.999972	0.999973	0.999974	0.999975	0.999976	0.999977	0.999978

## **Anhang 4 - Funktionen in MS Excel zur Unterstützung von Stichprobenverfahren**

Die nachstehend aufgeführten Funktionen können in MS Excel zur Berechnung der verschiedenen Parameter verwendet werden, die für die in diesen Leitlinien beschriebenen Verfahren und Konzepte erforderlich sind. Weitere Informationen zu ihrer Funktionsweise wie etwa zu den zugrundeliegenden mathematischen Formeln finden Sie in der Excel-Hilfe.

In den oben aufgeführten Formeln ist (.) ein Vektor, der die Adresse der Zellen mit den Werten der Stichprobe oder Grundgesamtheit enthält.

=AVERAGE(.): Mittelwert einer Datenmenge

=VAR(.): Varianz einer Stichprobendatenmenge

=VARP(.): Varianz einer Grundgesamtheitsdatenmenge

=STDEV(.): Standardabweichung einer Stichprobendatenmenge

=STDEVP(.): Standardabweichung einer Grundgesamtheitsdatenmenge

=RAND(): Zufallszahl zwischen 0 und 1 aus einer gleichmäßigen Verteilung

=SUM(.): Summe einer Datenmenge

## Anhang 5 – Glossar

<b>Begriff</b>	<b>Definition</b>
Anomaler Fehler	Fehler/falsche Darstellung, die nachweislich nicht repräsentativ für die Grundgesamtheit ist. Da eine statistische Stichprobe für die Grundgesamtheit repräsentativ ist, sollten „anomale Fehler“ nur in sehr außergewöhnlichen, gut begründeten Fällen akzeptiert werden.
Attributstichprobenverfahren	Ein statistischer Ansatz, mit dessen Hilfe die Sicherheit des Systems ermittelt und die erwartete Fehlerquote einer Stichprobe abgeschätzt werden kann. Am häufigsten wird diese Methode bei Prüfungen eingesetzt, wenn die Fehlerquote einer vorgeschriebenen Kontrolle mit dem Ziel zu untersuchen ist, das vom Prüfer geschätzte Niveau des Kontrollrisikos zu untermauern.
Aufdeckungsrisiko	Das Risiko, dass der Prüfer in den der Kommission bescheinigten Ausgabenaufstellungen oder in den zugrunde liegenden Aggregationsebenen einen wesentlichen Fehler nicht aufdeckt. Die Aufdeckungsrisiken hängen mit der Art und Weise zusammen, wie die Prüfungen durchgeführt werden.
Bekannter Fehler	Es kann der Fall eintreten, dass der Prüfer aufgrund eines in der Stichprobe festgestellten Fehlers einen oder mehrere Fehler außerhalb der Stichprobe feststellt. Diese werden als „bekannte Fehler“ eingestuft. Der in der Stichprobe aufgedeckte Fehler wird als Zufallsfehler angesehen und in die Hochrechnung einbezogen. Somit sollte dieser Stichprobenfehler, der zur Ermittlung des bekannten Fehlers führte, ebenso wie jeder andere Zufallsfehler auf die Grundgesamtheit extrapoliert werden.
Buchwert ( <i>BV</i> )	Die gegenüber der Kommission bescheinigten Ausgaben für ein Element (Vorhaben/Zahlungsantrag), $BV_t = 1, 2, \dots, N$ . Der Gesamtbuchwert einer Grundgesamtheit ist die Summe der Buchwerte der Elemente der Grundgesamtheit.
Differenzschätzung	Ein statistisches Stichprobenverfahren, das auf einer Auswahl mit gleichen Wahrscheinlichkeiten beruht. Das Verfahren stützt sich auf die Extrapolation des Fehlers in der Stichprobe. Der prognostizierte Fehler wird von den gesamten geltend gemachten Ausgaben in der Grundgesamtheit subtrahiert, um die korrekten Ausgaben in der Grundgesamtheit abzuschätzen (d. h. die Ausgaben, die sich bei Prüfung sämtlicher Vorhaben in der Grundgesamtheit ergeben würden).

<b>Begriff</b>	<b>Definition</b>
(Effektive) Genauigkeit (SE)	Dieser Fehler entsteht dadurch, dass nicht die ganze Grundgesamtheit betrachtet wird. Bei einem Stichprobenverfahren kommt stets ein Schätzungsfehler (Extrapolationsfehler) vor, weil der Prüfer aus den Stichprobendaten Rückschlüsse auf die Grundgesamtheit zieht. Dieser effektive Stichprobenfehler entspricht der Differenz zwischen dem anhand der Stichprobe hochgerechneten Parameter (Schätzwert) und dem tatsächlichen (unbekannten) Parameter der Grundgesamtheit (Wert des Fehlers). Er ist Ausdruck der Unsicherheit, die die Hochrechnung der Ergebnisse auf die Grundgesamtheit mit sich bringt.
Einfache Zufallsstichproben- verfahren	Beim einfachen Zufallsstichprobenverfahren handelt es sich um ein statistisches Stichprobenverfahren. Die zu untersuchende statistische Einheit ist das Vorhaben (oder der Zahlungsantrag, wie vorstehend erläutert). Die in der Stichprobe enthaltenen Einheiten werden nach dem Zufallsprinzip mit gleichen Wahrscheinlichkeiten ausgewählt.
Erheblichkeit	Fehler sind dann wesentlich, wenn sie ein bestimmtes für zulässig erachtetes Fehlerniveau überschreiten. Eine Erheblichkeitsschwelle von maximal 2 % gilt für Ausgaben, die im Bezugsjahr bei der Kommission geltend gemacht werden. Die Prüfbehörde kann zu Planungszwecken eine Absenkung der Erheblichkeitsschwelle erwägen (zulässiger Fehler). Die Erheblichkeit wird als Schwellenwert zum Vergleich des prognostizierten Fehlers bei den Ausgaben verwendet.
Expansionsfaktor (EF)	Der Expansionsfaktor ist ein Faktor, der bei der Berechnung der vorsichtig angesetzten MUS-Stichprobe verwendet wird, wenn Fehler erwartet werden, und beruht auf dem Risiko der fälschlichen Akzeptanz. Er reduziert den Stichprobenfehler. Werden keine Fehler erwartet, dann ist der voraussichtliche Fehler (AE) gleich Null, und es wird kein Expansionsfaktor verwendet. Werte für den Expansionsfaktor können Abschnitt 7.3.4.2 dieses Leitfadens entnommen werden.
Fehlangaben	Gleiche Bedeutung wie Fehler.

<b>Begriff</b>	<b>Definition</b>
Fehler ( $E$ )	<p>Für die Zwecke dieses Leitfadens ist ein Fehler eine quantifizierbare Überbewertung bei den bescheinigten Ausgaben, die der Kommission gemeldet wurden.</p> <p>Er wird definiert als die Differenz zwischen dem Buchwert des <math>i</math>-ten Elements in der Stichprobe und dem entsprechenden korrekten Buchwert,  <math>E_i = BV_i - CBV_i, i = 1, 2, \dots, N</math>.</p> <p>Im Falle der Schichtung der Grundgesamtheit wird ein Index <math>h</math> verwendet, um die jeweilige Schicht anzugeben:  <math>E_{hi} = BV_{hi} - CBV_{hi}, \text{ where } i = 1, 2, \dots, N_h, h = 1, 2, \dots, H</math>.  Dabei ist <math>H</math> die Anzahl der Schichten.</p>
Geplante Genauigkeit	<p>Der größtmögliche geplante Stichprobenfehler für die Ermittlung des Stichprobenumfangs, d. h. die maximale Abweichung zwischen dem wirklichen Wert der Grundgesamtheit und dem aus den Stichprobendaten abgeleiteten Schätzwert.</p> <p>In der Regel handelt es sich um die Differenz zwischen dem maximal zulässigen Fehler und dem voraussichtlichen Fehler, die auf einen Wert unterhalb der Erheblichkeitsschwelle gesetzt werden sollte.</p>
Gesamtbuchwert	<p>Die gegenüber der Kommission für Vorhaben im Rahmen eines Programms bzw. einer Gruppe von Programmen bescheinigten Gesamtausgaben; entspricht der Grundgesamtheit, aus der die Stichprobe gezogen wird.</p>
Größe der Grundgesamtheit ( $N$ )	<p>Die Anzahl der Vorhaben oder Zahlungsanträge, die im Bezugsjahr die gegenüber der Kommission bescheinigten Ausgaben ausmachen.</p> <p>Im Falle der Schichtung der Grundgesamtheit wird ein Index <math>h</math> verwendet, um die betreffende Schicht anzugeben, <math>N_{hi}, h = 1, 2, \dots, H</math>. Darin ist <math>H</math> die Anzahl der Schichten.</p>
Grundgenauigkeit ( $BP$ )	<p>Wird bei konservativem MUS verwendet und entspricht dem Produkt aus Stichprobenintervall und Zuverlässigkeitsfaktor (RF = Reliability Factor) (bereits für die Berechnung des Stichprobenumfangs verwendet).</p>

<b>Begriff</b>	<b>Definition</b>
Grundgesamtheit	Im Sinne von Artikel 62 Absatz 1 Buchstabe b der Verordnung (EG) Nr. 1083/2006 umfasst die Grundgesamtheit die gegenüber der Kommission im Rahmen eines Programms bzw. einer Gruppe von Programmen bescheinigten Ausgaben im Bezugsjahr. Die Grundgesamtheit, aus der die Stichprobe stammt, sollte alle Vorhaben, für die geltend gemachte Ausgaben in den der Kommission übermittelten bescheinigten Ausgabenerklärungen enthalten sind, umfassen. Stichprobeneinheit sollte das Vorhaben sein; ist die Grundgesamtheit der Vorhaben zu klein für ein statistisches Stichprobenverfahren (d. h. 50 bis 150 Einheiten), kann der Zahlungsantrag des Begünstigten als Prüfeinheit gewählt werden.
Inhärentes Risiko ( <i>IR</i> )	Das Risiko, dass in den der Kommission bescheinigten Ausgabenaufstellungen oder in den zugrunde liegenden Aggregationsebenen bei nicht vorhandenen internen Kontrollverfahren ein wesentlicher Fehler vorkommen kann. Das inhärente Risiko muss vor Inangriffnahme detaillierter Prüfungsverfahren bewertet werden; dies erfolgt durch mündliche Befragungen des Managements und des Schlüsselpersonals, aber auch durch die Überprüfung kontextbezogener Informationen (wie z. B. Organigramme, Handbücher und interne/externe Unterlagen).
Inkrementelle Toleranz ( <i>IA</i> )	Misst die schrittweise Erhöhung des Genauigkeitsgrades mit jedem in der Stichprobe festgestellten Fehler. Diese Toleranz wird bei einem konservativen Ansatz für das MUS verwendet und sollte immer dann, wenn Fehler in der Stichprobe festgestellt werden, der Grundgenauigkeit hinzuaddiert werden (vgl. Abschnitt 7.3.4.5 dieses Leitfadens).
Konfidenzintervall	Das Intervall, das mit einer gewissen Wahrscheinlichkeit (Konfidenzniveau genannt) den wahren (unbekannten) Wert der Grundgesamtheit (im Allgemeinen der Fehlerbetrag oder die Fehlerquote) enthält.
Konfidenzniveau	Die Wahrscheinlichkeit, dass ein anhand der Stichprobendaten erstelltes Konfidenzintervall den wahren (unbekannten) Fehler der Grundgesamtheit enthält.
Kontrollrisiko ( <i>CR</i> )	Das Risiko, dass in den der Kommission bescheinigten Ausgabenaufstellungen oder in den zugrunde liegenden Aggregationsebenen mit den internen Kontrollverfahren der Verwaltung ein wesentlicher Fehler nicht verhindert, entdeckt oder behoben wird.
Korrektur Buchwert ( <i>CBV</i> )	Der korrekte Ausgabenbetrag, der festgestellt werden könnte, wenn alle Zahlungsanträge in der Grundgesamtheit geprüft werden würden.

<b>Begriff</b>	<b>Definition</b>
Maximal zulässiger Fehler ( <i>TE</i> )	Der maximal zulässige Fehler, der in einem bestimmten Jahr in der Grundgesamtheit auftreten darf, also die Schwelle, bei deren Überschreiten von wesentlichen Fehlangaben hinsichtlich der Grundgesamtheit auszugehen ist. Bei einer Erheblichkeitsschwelle von 2 % beträgt daher der maximal zulässige Fehler 2 % der Ausgaben, die der Kommission für das Bezugsjahr bescheinigt wurden.
Monetary Unit Sampling (MUS, wertbezogenes Stichprobenverfahren)	Ein statistisches Stichprobenverfahren, bei dem die Geldeinheit als Hilfsvariable für die Probenahme verwendet wird. Dieser Ansatz beruht in der Regel auf einer systematischen Stichprobenziehung, bei der die Wahrscheinlichkeit in einem proportionalen Verhältnis zur Größe steht (PPS = Probability Proportional to Size), d. h. in einem proportionalen Verhältnis zum Geldwert der Stichprobeneinheit (Elemente mit hohen Werten werden mit größerer Wahrscheinlichkeit ausgewählt).
Obere Fehlergrenze ( <i>ULE</i> )	Entspricht der Summe des prognostizierten Fehlers und der Genauigkeit der Extrapolation. Gleiche Bedeutung wie Obergrenze des Konfidenzintervalls, Obergrenze für die Fehlangabe der Grundgesamtheit und Obergrenze der Fehlangabe.
Prognostizierte Gesamtfehlerquote ( <i>TPER</i> )	Die prognostizierte Gesamtfehlerquote entspricht der geschätzten Wirkung der Fehler als prozentualer Anteil an der Grundgesamtheit. Die prognostizierte Gesamtfehlerquote ist das Verhältnis zwischen dem prognostizierten Gesamtfehler und dem Gesamtbuchwert der Grundgesamtheit (Ausgaben des Bezugsjahres). Die Prüfbehörde sollte die prognostizierte Gesamtfehlerquote mit der Erheblichkeitsschwelle vergleichen, um Schlussfolgerungen für die von der Stichprobe erfasste Grundgesamtheit als Ganzes ziehen zu können.
Prognostizierter Gesamtfehler ( <i>TPE</i> )	Der prognostizierte Gesamtfehler entspricht der Summe folgender Fehler: prognostizierte Zufallsfehler, systembedingte Fehler, bekannte Fehler und nicht berichtigte anomale Fehler. Die Prüfbehörde sollte alle Fehler quantifizieren und in den prognostizierten Gesamtfehler aufnehmen; dies gilt nicht für berichtigte anomale Fehler. Gleiche Bedeutung wie prognostizierte Gesamtfehlangabe.
Prognostizierter Zufallsfehler	Der prognostizierte Zufallsfehler ist das Ergebnis der Hochrechnung der in der Stichprobe (bei der Prüfung von Vorhaben) festgestellten Zufallsfehler auf die Grundgesamtheit. Das Hochrechnungs-/Prognoseverfahren hängt vom verwendeten Stichprobenverfahren ab.

<b>Begriff</b>	<b>Definition</b>
Prognostizierter/ Hochgerechneter Fehler ( <i>EE</i> )	Der prognostizierte/hochgerechnete Fehler entspricht der geschätzten Wirkung der Zufallsfehler auf der Ebene der Grundgesamtheit.
Prüfungsrisiko ( <i>AR</i> )	Das Risiko, dass der Prüfer trotz wesentlicher Fehlangaben in der Ausgabenerklärung ein uneingeschränktes Prüfungsurteil ausspricht.
Prüfungssicherheit	Das Sicherheitsmodell ist das Gegenteil des Risikomodells. Wird das Prüfungsrisiko auf 5 % eingeschätzt, so wird von einer Prüfungssicherheit von 95 % ausgegangen. Die Verwendung des Prüfungssicherheitsmodells bezieht sich auf die Planung und die zugrunde liegende Zuteilung der Ressourcen für ein bestimmtes operationelles Programm oder für mehrere operationelle Programme.
Risiko eines wesentlichen Fehlers	Das Produkt aus dem inhärenten und dem Kontrollrisiko. Das Risiko eines wesentlichen Fehlers hängt mit dem Ergebnis der Systemprüfungen zusammen.
Schichtung	Unterteilung einer Grundgesamtheit in mehrere Gruppen (Schichten) anhand des Wertes einer Hilfsvariablen (normalerweise die zu prüfende Variable, d. h. der Wert der Ausgaben pro Vorhaben in dem geprüften Programm). Beim geschichteten Stichprobenverfahren werden aus jeder Schicht unabhängige Stichproben entnommen. Mit der Schichtung werden zwei Hauptziele verfolgt: Zum einen ermöglicht sie in der Regel eine höhere Genauigkeit (bei gleichem Stichprobenumfang) oder eine Verringerung des Stichprobenumfangs (bei gleichem Genauigkeitsgrad); zum anderen gewährleistet sie, dass in der Stichprobe alle Teilgesamtheiten vertreten sind, die den einzelnen Schichten entsprechen.
Standardabweichung ( $\sigma$ oder $s$ )	Ein Maß für die Streuung der Grundgesamtheit um ihren Mittelwert. Sie kann anhand von Fehlern oder von Buchwerten berechnet werden. Bei der Berechnung anhand der Grundgesamtheit wird sie normalerweise mit $\sigma$ angegeben, bei der Berechnung anhand der Stichprobe mit $s$ . Je größer die Standardabweichung ist, desto heterogener ist die Grundgesamtheit (Stichprobe).
Stichprobeneinheit	Die zur Prüfung auszuwählende Einheit; in der Regel das Vorhaben. Besteht ein Vorhaben aus mehreren separaten Projekten, können diese für Stichprobenezwecke getrennt ausgewiesen werden. Ist die Grundgesamtheit der Vorhaben zu klein für ein statistisches Stichprobenverfahren (d. h. 50 bis 150 Einheiten), kann der Zahlungsantrag des Begünstigten als Prüfeinheit gewählt werden.

<b>Begriff</b>	<b>Definition</b>
Stichprobenfehler	Gleiche Bedeutung wie effektive Genauigkeit.
Stichprobenfehlerquote	Der Stichprobenfehler entspricht dem Quotienten aus dem Betrag der festgestellten Unregelmäßigkeiten und dem Betrag der geprüften Ausgaben.
Stichprobenintervall ( <i>SI</i> )	Bei Stichprobenverfahren, die auf einer systematischen Auswahl beruhen, ist das Stichprobenintervall der verwendete Auswahlschritt. Bei Verfahren, bei denen die Auswahl mit einer Wahrscheinlichkeit proportional zu den Ausgaben erfolgt (wie beim MUS-Verfahren), ist das Stichprobenintervall der Quotient aus dem Gesamtbuchwert der Grundgesamtheit und dem Stichprobenumfang.
Stichprobenumfang ( <i>n</i> )	Die Anzahl der Einheiten/Elemente in der Stichprobe. Im Falle der Schichtung der Grundgesamtheit wird ein Index <b>h</b> verwendet, um die betreffende Schicht anzugeben: $n_h, h = 1, 2, \dots, H$ . Darin ist <b>H</b> die Anzahl der Schichten.
Stichprobenverfahren	Das Stichprobenverfahren umfasst zwei Elemente, und zwar den Stichprobenplan (z. B. gleiche Wahrscheinlichkeit, Wahrscheinlichkeit proportional zur Größe) und das Verfahren der Hochrechnung (Schätzung). Zusammen bilden diese beiden Elemente den Rahmen für die Berechnung des Stichprobenumfangs und die Hochrechnung des Fehlers.
Systembedingter Fehler	Systembedingte Fehler sind in der geprüften Stichprobe festgestellte Fehler, die sich auf die nicht geprüfte Grundgesamtheit auswirken und unter klar definierten und ähnlichen Umständen auftreten. Diese Fehler weisen im Allgemeinen ein gemeinsames Merkmal auf, z. B. Art des Vorhabens, Ort oder Zeitraum. Sie gehen im Allgemeinen auf unwirksame Kontrollverfahren in (Teilen von) Verwaltungs- und Kontrollsystemen zurück.
Unregelmäßigkeit	Gleiche Bedeutung wie Fehler
Varianz ( $\sigma^2$ )	Quadrat der Standardabweichung
Voraussichtlicher Fehler ( <b>AE</b> )	Der Fehlerbetrag, den der Prüfer in der Grundgesamtheit vorzufinden erwartet (nach durchgeführter Prüfung). Zur Planung des Stichprobenumfangs wird die voraussichtliche Fehlerquote auf maximal 4 % des Buchwerts der Grundgesamtheit festgesetzt.
Zufallsfehler	Fehler, die nicht als systembedingt eingestuft werden, werden als Zufallsfehler bezeichnet. Dabei wird angenommen, dass Zufallsfehler, die in der geprüften Stichprobe gefunden wurden, wahrscheinlich auch in der nicht geprüften Grundgesamtheit vorhanden sind. Diese Fehler sind daher bei der Berechnung der Fehler zu berücksichtigen.
Zulässige Fehlengabe	Gleiche Bedeutung wie zulässiger Fehler.

<b>Begriff</b>	<b>Definition</b>
Zulässiger Fehler	Die maximal zulässige Fehlerquote in der Grundgesamtheit. Bei einer Erheblichkeitsschwelle von 2 % beträgt daher der maximal zulässige Fehler 2 % der Ausgaben, die der Kommission für das Bezugsjahr bescheinigt wurden.
Zuverlässigkeitsfaktor ( <i>RF</i> )	Der Zuverlässigkeitsfaktor <i>RF</i> ist eine Konstante aus der Poisson-Verteilung für einen erwarteten Null-Fehler. Er ist vom Konfidenzniveau abhängig. Die in den verschiedenen Situationen anzuwendenden Werte sind Abschnitt 7.3.4.2 dieses Leitfadens zu entnehmen.
z	Ein Parameter aus der Normalverteilung, der mit dem Konfidenzniveau im Zusammenhang steht, das mithilfe von Systemprüfungen ermittelt wird. Die möglichen Werte von z werden in Abschnitt 6.4 dieses Leitfadens angegeben.