

Probenahme
 Statistische Grundlagen der Probenahme
 bei zweifacher Aufteilung
 nach zwei einander nachgeordneten Gesichtspunkten

DIN
53 803
 Teil 4

Sampling; statistical basis; hierarchical classification
 Échantillonnage; base statistique; classification hiérarchique

Inhalt

		Seite		Seite
1	Grundlagen	2	6.3	Stichprobenaufbau bei vorgegebenen Gesamtkosten
				10
2	Begriffe	2	6.4	Stichprobenaufbau bei vorgegebenen Gesamtkosten und kleinster erreichbarer Weite des Vertrauensbereiches für den Mittelwert der Gesamtheit
				10
2.1	Begriffe für die Probenahme	2		
2.2	Statistische Begriffe	2		
2.3	Weitere Begriffe	2		
3	Anordnungsschema der Einzelwerte bei zweifacher Aufteilung	2	7	Programmablaufpläne für die Berechnung des Stichprobenaufbaues
				10
4	Statistische Auswertung bei zweifacher Aufteilung nach zwei einander nachgeordneten Gesichtspunkten	2	8	Probenahme bei reduzierbarer Aufteilung
				17
4.1	Modellvorstellung	2	Anhang A	Beispiele aus der Textiltechnik
4.2	Auswertung	3		17
4.3	Schätzwerte für die Modellparameter	3	A.1	Aufteilung nach zwei einander nachgeordneten Gesichtspunkten; Durchführung der Varianzanalyse; Bestimmung von Schätzwerten für die Varianzen; Testen von Hypothesen über die Modellparameter; Vertrauensbereich für den Mittelwert der Gesamtheit; Feinheit von Filamentgarn
4.4	Testen von Hypothesen über die Modellparameter	3		17
4.5	Vertrauensbereiche für die Modellparameter	5	A.2	Stichprobenaufbau bei vorgegebener relativer Weite für den Vertrauensbereich der Gesamtheit; varianzbestimmte Stichprobe; Feinheit von Filamentgarn
				19
5	Stichprobenaufbau	5	A.3	Stichprobenaufbau bei vorgegebener relativer Weite für den Vertrauensbereich der Gesamtheit und minimalen Gesamtkosten; varianzbestimmte kostenminimale Stichprobe; Feinheit von Filamentgarn
				19
5.1	Stichprobenaufbau bei vorgegebener Weite des Vertrauensbereiches für den Mittelwert der Gesamtheit; varianzbestimmte Stichprobe	5	A.4	Stichprobenaufbau bei vorgegebenen Gesamtkosten; kostenbestimmte Stichprobe; Rundversuch zur Messung der Dickenminderung einer Teppichware
5.2	Stichprobenaufbau bei vorgegebener Weite des Vertrauensbereiches für den Mittelwert der Gesamtheit und minimalen Gesamtkosten; varianzbestimmte kostenminimale Stichprobe	7		20
5.3	Stichprobenaufbau bei vorgegebenen Gesamtkosten; kostenbestimmte Stichprobe	8	A.5	Stichprobenaufbau bei vorgegebenen Gesamtkosten und kleinster erreichbarer Weite des Vertrauensbereiches für den Mittelwert der Gesamtheit; kostenbestimmte varianzminimale Stichprobe; Rundversuch zur Messung der Dickenminderung einer Teppichware
5.4	Stichprobenaufbau bei vorgegebenen Gesamtkosten und kleinster erreichbarer Weite des Vertrauensbereiches für den Mittelwert der Gesamtheit; kostenbestimmte varianzminimale Stichprobe	9		21
6	Rundungsempfehlungen	9	Anhang B	Formelzeichen
				22
6.1	Stichprobenaufbau bei vorgegebener Weite des Vertrauensbereiches für den Mittelwert der Gesamtheit	9		Zitierte Normen und andere Unterlagen
6.2	Stichprobenaufbau bei vorgegebener Weite des Vertrauensbereiches für den Mittelwert der Gesamtheit und minimalen Gesamtkosten	10		23
				Weitere Normen und andere Unterlagen
				23
				Erläuterungen
				23

Fortsetzung Seite 2 bis 23

Normenausschuß Materialprüfung (NMP) im DIN Deutsches Institut für Normung e.V.
 Ausschuß Qualitätssicherung und angewandte Statistik (AQS) im DIN

1 Grundlagen

In DIN 53 803 Teil 1 werden Verfahren zur Berechnung des Aufbaues und des Umfanges von Stichproben beschrieben, falls das zu prüfende Material in Gruppen zu unterteilen ist. Eine Unterteilung ist immer dann vorzunehmen, wenn Inhomogenitäten des Materials vorhanden sind oder vermutet werden.

Bei einer nach **einem** sachlichen Gesichtspunkt gewählten Einteilung in Gruppen ist eine einfache Aufteilung der Einheiten gegeben. Zur Auswertung werden Methoden der einfachen Varianzanalyse in DIN 53 803 Teil 1 herangezogen. Da in vielen Fällen eine einfache Aufteilung nicht ausreichend ist, wird man mehrfach aufteilen.

In dieser Norm und in DIN 53 803 Teil 3 wird der Aufbau und der Umfang von Stichproben behandelt, sofern man das zu untersuchende Material nach **zwei** sachlichen Gesichtspunkten unterteilt.

In dieser Norm wird die Unterteilung nach zwei **einander nachgeordneten** Gesichtspunkten behandelt (Schachtelmodell; Hierarchical classification), und zwar nur das Modell mit Zufallskomponenten. Zu Modellen mit systematischen Komponenten und gemischten Modellen wird auf die zitierten Unterlagen verwiesen. [1], [2]

Die Unterteilung in mehr als zwei Gruppen wird hier nicht behandelt. Falls eine weitergehende Unterteilung beabsichtigt ist, sollte man vorher überlegen, wie weit das sinnvoll ist und später befriedigend interpretiert werden kann. Zum Verständnis dieser Norm wird die Kenntnis von DIN 53 803 Teil 1 und Teil 2 vorausgesetzt.

2 Begriffe

2.1 Begriffe für die Probenahme

Begriffe für die Probenahme nach DIN 53 803 Teil 1, Ausgabe März 1979, Abschnitt 1.2.1 bis Abschnitt 1.2.11.

2.2 Statistische Begriffe

Statistische Begriffe nach DIN 53 803 Teil 1, Ausgabe März 1979, Abschnitt 1.3.1 bis Abschnitt 1.3.18. Gegenüber der dortigen Kennzeichnungsweise wird in dieser Norm jeder Einzelwert mit drei Indices versehen. Sie kennzeichnen die Gruppe, zu der jeder Einzelwert gehört.

2.3 Weitere Begriffe

2.3.1 Schachtelmodell

Bei der Gruppeneinteilung einer Gesamtheit nach zwei einander nachgeordneten Gesichtspunkten in Einheiten erster, zweiter und dritter Stufe ist das Schachtelmodell mit drei Stufen ein statistisches Auswerteverfahren, das auf der Aufteilung der gesamten Summe der quadrierten Abweichungen (S.d.q.A.) in drei Anteile beruht:

- die S.d.q.A. innerhalb der ersten Stufe (übergeordneter Gesichtspunkt)
- die S.d.q.A. innerhalb der zweiten Stufe (nachgeordneter Gesichtspunkt)
- die S.d.q.A. innerhalb der dritten Stufe (Wiederholung)

Daraus ergeben sich drei Varianzen, und zwar:

- die Varianz in der ersten Stufe,
- die Varianz in der zweiten Stufe,
- die Varianz in der dritten Stufe.

Beim Schachtelmodell ist die Reihenfolge der Unterteilung zu beachten; die drei Stufen dürfen nicht miteinander vertauscht werden.

3 Anordnungsschema der Einzelwerte bei zweifacher Aufteilung

Die Einzelwerte werden nach dem Schema in Tabelle 1 (linker Teil) zusammengestellt. Das Schema enthält n_1 Zeilen ($i = 1, 2, 3 \dots n_1$) entsprechend der Unterteilung nach dem übergeordneten Gesichtspunkt und n_2 Spalten ($j = 1, 2, 3 \dots n_2$) entsprechend der Unterteilung nach dem nachgeordneten Gesichtspunkt. An den Schnittpunkten von Zeilen und Spalten stehen in jeder Zelle (gekennzeichnet durch das Indexpaar ij) n_3 Einzelwerte x_{ijk} ($k = 1, 2, 3 \dots n_3$). Dabei wird in dieser Norm die Anzahl der Einzelwerte in jeder Zelle als gleich groß vorausgesetzt.

Aus Gründen der Zweckmäßigkeit trägt man in dieses oder ein separates Schema sogleich die Quadrate x_{ijk}^2 der Einzelwerte in gleicher Weise ein. Man bildet sodann innerhalb jeder Zelle die Zellensumme für die Einzelwerte

$$A_{ij} = \sum_{k=1}^{n_3} x_{ijk} \quad (1)$$

sowie die Quadrate der Einzelwerte

$$B_{ij} = \sum_{k=1}^{n_3} x_{ijk}^2 \quad (2)$$

Bis hierher ist der Gang der Rechnung unabhängig davon, ob man die Aufteilung nach zwei einander nachgeordneten oder zwei gleichberechtigten Gesichtspunkten wählt.

4 Statistische Auswertung bei zweifacher Aufteilung nach zwei einander nachgeordneten Gesichtspunkten

4.1 Modellvorstellung

Bei der Aufteilung nach zwei einander nachgeordneten Gesichtspunkten setzt sich jeder mögliche Einzelwert a_{ijk} der Gesamtheit aus mehreren Komponenten zusammen:

$$a_{ijk} = \mu + \alpha_i + \beta_{ij} + \gamma_{ijk} \quad (3)$$

Es ist

- | | |
|-------------------------------------|--|
| μ | der Mittelwert der Gesamtheit |
| $\alpha_i = \mu_i - \mu$ | die zufällige Abweichung des Mittelwertes μ_i der i -ten Einheit erster Stufe vom Mittelwert μ der Gesamtheit |
| $\beta_{ij} = \mu_{ij} - \mu_i$ | die zufällige Abweichung des Mittelwertes μ_{ij} der j -ten Einheit zweiter Stufe aus der i -ten Einheit erster Stufe von deren Mittelwert μ_i |
| $\gamma_{ijk} = a_{ijk} - \mu_{ij}$ | die zufällige Abweichung des Einzelwertes a_{ijk} aus der j -ten Einheit zweiter Stufe innerhalb der i -ten Einheit erster Stufe vom zugehörigen Mittelwert μ_{ij} |

In jedem Fall bezeichnet der Index i die Einheiten erster Stufe, der Index j die Einheiten zweiter Stufe innerhalb einer Einheit erster Stufe und der Index k die Einheiten dritter Stufe innerhalb einer Einheit zweiter Stufe.

Für die Zufallskomponenten in den einzelnen Stufen gilt folgendes:

- a) Der Mittelwert aller $\alpha_i = \mu_i - \mu$ ist Null; die Varianz der α_i wird mit σ_1^2 bezeichnet.
- b) Der Mittelwert aller $\beta_{ij} = \mu_{ij} - \mu_i$ bei festem Wert i ist Null; die Varianzen der β_{ij} werden mit σ_{2i}^2 bezeichnet. Da man voraussetzt, daß die Varianzen in allen Einheiten einer Stufe übereinstimmen, gilt weiter:

$$\sigma_{2i}^2 = \sigma_2^2 \text{ für alle } i$$
- c) Der Mittelwert aller $\gamma_{ijk} = a_{ijk} - \mu_{ij}$ bei festen Werten i und j ist Null; die Varianzen der γ_{ijk} werden mit σ_{3ij}^2 bezeichnet. Analog zur Voraussetzung unter Punkt b) gilt weiter:

$$\sigma_{3ij}^2 = \sigma_3^2 \text{ für alle } ij$$

Auf Grund dieser Modellvorstellung enthält also das Schachtelmodell mit drei Stufen vier unbekannte Parameter, nämlich einen Gesamtmittelwert μ und drei Varianzen σ_1^2 , σ_2^2 und σ_3^2 für die einzelnen Stufen.

Die Zufallsgrößen α_i , β_{ij} und γ_{ijk} sind normalverteilt.

4.2 Auswertung

Analog zur Modellgleichung wird jeder Meßwert x_{ijk} in folgende Anteile zerlegt:

$$x_{ijk} = \bar{x}_{...} + (\bar{x}_{i..} - \bar{x}_{...}) + (\bar{x}_{ij.} - \bar{x}_{i..}) + (x_{ijk} - \bar{x}_{ij.}) \quad (4)$$

Dabei ist:

ein Einzelwert der Stichprobe:

$$x_{ijk}$$

der Gesamtmittelwert der Stichprobe:

$$\bar{x}_{...} = \frac{1}{n_1 \cdot n_2 \cdot n_3} \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} \sum_{k=1}^{n_3} x_{ijk} \quad (5)$$

der Mittelwert aus den n_2 Einzelwerten der zweiten Stufe und den n_3 Einzelwerten der dritten Stufe, die zur Einheit i der ersten Stufe gehören:

$$\bar{x}_{i..} = \frac{1}{n_2 \cdot n_3} \sum_{j=1}^{n_2} \sum_{k=1}^{n_3} x_{ijk} \quad (6)$$

der Mittelwert aus den n_3 Einzelwerten der dritten Stufe, die zur Einheit j zweiter Stufe aus der Einheit i erster Stufe gehören:

$$\bar{x}_{ij.} = \frac{1}{n_3} \sum_{k=1}^{n_3} x_{ijk} \quad (7)$$

Ausgehend von dem Anordnungsschema im dick umrandeten Teil der Tabelle 1 für die Einzelwerte x_{ijk} (und die quadrierten Einzelwerte x_{ijk}^2) werden die folgenden auf dem rechten Rand eingetragenen Zwischensummen für jede der n_1 Zeilen gebildet:

Summe aller Einzelwerte innerhalb der Zeile i :

$$\sum_{j=1}^{n_2} A_{ij} = \sum_{j=1}^{n_2} \sum_{k=1}^{n_3} x_{ijk} \quad (8)$$

Summe aller quadrierten Einzelwerte innerhalb der Zeile i :

$$\sum_{j=1}^{n_2} B_{ij} = \sum_{j=1}^{n_2} \sum_{k=1}^{n_3} x_{ijk}^2 \quad (9)$$

Quadrat der Summe aller Einzelwerte innerhalb der Zeile i :

$$\left(\sum_{j=1}^{n_2} A_{ij} \right)^2 = \left(\sum_{j=1}^{n_2} \sum_{k=1}^{n_3} x_{ijk} \right)^2 \quad (10)$$

Zeilensumme der quadrierten Zellensummen:

$$\sum_{j=1}^{n_2} A_{ij}^2 = \sum_{j=1}^{n_2} \left(\sum_{k=1}^{n_3} x_{ijk} \right)^2 \quad (11)$$

Diese für jede Zeile i vorliegenden vier Zwischensummen werden jeweils über alle n_1 Zeilen zu den folgenden Hilfsgrößen summiert:

Summe aller Einzelwerte:

$$A = \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} A_{ij} = \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} \sum_{k=1}^{n_3} x_{ijk} \quad (12)$$

Summe aller quadrierten Einzelwerte:

$$B = \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} B_{ij} = \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} \sum_{k=1}^{n_3} x_{ijk}^2 \quad (13)$$

Summe der Quadrate der Summe aller Einzelwerte innerhalb der Zeilen i :

$$C = \sum_{i=1}^{n_1} \left(\sum_{j=1}^{n_2} A_{ij} \right)^2 = \sum_{i=1}^{n_1} \left(\sum_{j=1}^{n_2} \sum_{k=1}^{n_3} x_{ijk} \right)^2 \quad (14)$$

Gesamtsumme der quadrierten Zellensummen:

$$E = \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} A_{ij}^2 = \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} \left(\sum_{k=1}^{n_3} x_{ijk} \right)^2 \quad (15)$$

Die Zwischensummen und Hilfsgrößen dürfen nicht gerundet werden (siehe DIN 53803 Teil 1, Ausgabe März 1979, Abschnitt 4, und DIN 53804 Teil 1, Ausgabe September 1981, Abschnitt 4.3).

4.3 Schätzwerte für die Modellparameter

Mit Hilfe der Summe der quadrierten Abweichungen Q und der zugehörigen Zahl der Freiheitsgrade f wird die Zerlegungstafel aufgestellt (siehe Tabelle 2).

Die Summe Q_{total} läßt sich nach der Zerlegungstafel aus den Hilfsgrößen B und A berechnen; außerdem gilt jedoch die Beziehung:

$$Q_{\text{total}} = Q_1 + Q_2 + Q_3 \quad (16)$$

was zweckmäßig zur Kontrolle der Rechnung benutzt wird.

Analog gilt für die Zahl der Freiheitsgrade

$$f_{\text{total}} = n_1 \cdot n_2 \cdot n_3 - 1 = f_1 + f_2 + f_3 \quad (17)$$

wobei $n_1 \cdot n_2 \cdot n_3$ der gesamte Stichprobenumfang ist. Als Ergebnis der Auswertung erhält man:

Schätzwert für μ ist

$$\hat{\mu} = \bar{x}_{...} = \frac{A}{n_1 \cdot n_2 \cdot n_3} \quad (18)$$

Schätzwert für σ_1^2 ist

$$\hat{\sigma}_1^2 = \frac{s_1^2 - s_2^2}{n_2 \cdot n_3} \quad (19)$$

Schätzwert für σ_2^2 ist

$$\hat{\sigma}_2^2 = \frac{s_2^2 - s_3^2}{n_3} \quad (20)$$

Schätzwert für σ_3^2 ist

$$\hat{\sigma}_3^2 = s_3^2 \quad (21)$$

Falls $\hat{\sigma}_1^2$ oder $\hat{\sigma}_2^2$ negativ sind, werden σ_1^2 oder σ_2^2 durch Null geschätzt.

4.4 Testen von Hypothesen über die Modellparameter

Ausgehend von der Zerlegungstafel läßt sich mit dem F -Test prüfen, ob die Varianzen σ_1^2 und σ_2^2 als von Null verschieden anzusehen sind. Dazu stellt man die Nullhypothese H_0 für σ_2^2 auf:

$$H_0: \sigma_2^2 = 0, \text{ d. h. die Varianz in der zweiten Stufe ist nicht von Null verschieden.}$$

Dieser Hypothese stellt man die Alternativhypothese H_1 entgegen:

$$H_1: \sigma_2^2 \neq 0, \text{ d. h. die Varianz in der zweiten Stufe ist von Null verschieden.}$$

Die Nullhypothese wird mit dem Prüfwert s_2^2/s_3^2 geprüft, der mit dem Tabellenwert $F_{f_2; f_3; 1-\alpha}$ der F -Verteilung zum Signifikanzniveau α verglichen wird; die Tabellenwerte erhält man beispielsweise aus [1].

Ist $s_2^2/s_3^2 \leq F_{f_2; f_3; 1-\alpha}$, wird die Nullhypothese nicht verworfen und die Varianz σ_2^2 als nicht von Null verschieden angesehen. Ist $s_2^2/s_3^2 > F_{f_2; f_3; 1-\alpha}$, wird die Nullhypothese verworfen und die Varianz σ_2^2 als von Null verschieden betrachtet.

Tabelle 1. Anordnungsschema für die Einzelwerte und deren Quadrate; Zwischensummen und Hilfsgrößen für das Schachtelmodell

$i \backslash j$	1	2	bis	j	bis	n_2	$\sum_j \sum_k x_{ijk}$	$\sum_j \sum_k x_{ijk}^2$	$\left(\sum_j \sum_k x_{ijk}\right)^2$	$\sum_j \left(\sum_k x_{ijk}\right)^2$
1	x_{111} . . x_{11k} . . x_{11n_3} A_{11}	x_{121} . . x_{12k} . . x_{12n_3} A_{12}	bis	x_{1j1} . . x_{1jk} . . x_{1jn_3} A_{1j}	bis	x_{1n_21} . . x_{1n_2k} . . $x_{1n_2n_3}$ A_{1n_2}	$\sum_j A_{1j}$	$\sum_j B_{1j}$	$\left(\sum_j A_{1j}\right)^2$	$\sum_j A_{1j}^2$
2	x_{211} . . x_{21k} . . x_{21n_3} A_{21}	x_{221} . . x_{22k} . . x_{22n_3} A_{22}	bis	x_{2j1} . . x_{2jk} . . x_{2jn_3} A_{2j}	bis	x_{2n_21} . . x_{2n_2k} . . $x_{2n_2n_3}$ A_{2n_2}	$\sum_j A_{2j}$	$\sum_j B_{2j}$	$\left(\sum_j A_{2j}\right)^2$	$\sum_j A_{2j}^2$
bis	bis	bis		bis		bis	bis	bis	bis	bis
i	x_{i11} . . x_{i1k} . . x_{i1n_3} A_{i1}	x_{i21} . . x_{i2k} . . x_{i2n_3} A_{i2}	bis	x_{ij1} . . x_{ijk} . . x_{ijn_3} A_{ij}	bis	x_{in_21} . . x_{in_2k} . . $x_{in_2n_3}$ A_{in_2}	$\sum_j A_{ij}$	$\sum_j B_{ij}$	$\left(\sum_j A_{ij}\right)^2$	$\sum_j A_{ij}^2$
bis	bis	bis		bis		bis	bis	bis	bis	bis
n_1	x_{n_111} . . x_{n_11k} . . $x_{n_11n_3}$ A_{n_11}	x_{n_121} . . x_{n_12k} . . $x_{n_12n_3}$ A_{n_12}	bis	x_{n_1j1} . . x_{n_1jk} . . $x_{n_1jn_3}$ A_{n_1j}	bis	$x_{n_1n_21}$. . $x_{n_1n_2k}$. . $x_{n_1n_2n_3}$ $A_{n_1n_2}$	$\sum_j A_{n_1j}$	$\sum_j B_{n_1j}$	$\left(\sum_j A_{n_1j}\right)^2$	$\sum_j A_{n_1j}^2$
							$\sum_i \sum_j A_{ij}$	$\sum_i \sum_j B_{ij}$	$\sum_i \left(\sum_j A_{ij}\right)^2$	$\sum_i \sum_j A_{ij}^2$
							$\sum_i \sum_j \sum_k x_{ijk}$	$\sum_i \sum_j \sum_k x_{ijk}^2$	$\sum_i \left(\sum_j \sum_k x_{ijk}\right)^2$	$\sum_i \sum_j \left(\sum_k x_{ijk}\right)^2$
							A	B	C	E

Tabelle 2. Zerlegungstafel für das Schachtelmodell

Schwan- kungen	Summe der quadrierten Abweichungen Q	Zahl der Freiheitsgrade f	Varianz $s^2 = \frac{Q}{f}$	Die Varianz ist Schätzwert für
in der ersten Stufe	$Q_1 = n_2 \cdot n_3 \cdot \sum_i (x_{i..} - x_{...})^2 = \frac{C}{n_2 \cdot n_3} - \frac{A^2}{n_1 \cdot n_2 \cdot n_3}$	$f_1 = n_1 - 1$	s_1^2	$\sigma_3^2 + n_3 \sigma_2^2 + n_2 n_3 \sigma_1^2$
in der zweiten Stufe	$Q_2 = n_3 \cdot \sum_i \sum_j (x_{ij.} - x_{i..})^2 = \frac{E}{n_3} - \frac{C}{n_2 \cdot n_3}$	$f_2 = n_1 \cdot (n_2 - 1)$	s_2^2	$\sigma_3^2 + n_3 \sigma_2^2$
in der dritten Stufe	$Q_3 = \sum_i \sum_j \sum_k (x_{ijk} - x_{ij.})^2 = B - \frac{E}{n_3}$	$f_3 = n_1 \cdot n_2 \cdot (n_3 - 1)$	s_3^2	σ_3^2
total	$Q_{\text{total}} = \sum_i \sum_j \sum_k (x_{ijk} - x_{...})^2 = B - \frac{A^2}{n_1 \cdot n_2 \cdot n_3}$	$f_{\text{total}} = n_1 \cdot n_2 \cdot n_3 - 1$		

Weiterhin stellt man die Nullhypothese H_0 für σ_1^2 auf:

$H_0: \sigma_1^2 = 0$, d. h. die Varianz in der ersten Stufe ist nicht von Null verschieden.

Die Alternativhypothese H_1 lautet:

$H_1: \sigma_1^2 \neq 0$, d. h. die Varianz in der ersten Stufe ist von Null verschieden.

In diesem Fall ist der Prüfwert s_1^2/s_2^2 zu berechnen und mit dem Tabellenwert $F_{f_1; f_2; 1-\alpha}$ zu vergleichen. Die Nullhypothese wird verworfen, falls der Prüfwert größer als der Tabellenwert ist.

4.5 Vertrauensbereiche für die Modellparameter

Der Vertrauensbereich für den Mittelwert der Gesamtheit μ zum Vertrauensniveau $1 - \alpha$ ergibt sich zu:

$$\bar{x}_{...} - t_{f_1; 1-\alpha/2} \cdot \frac{s_1}{\sqrt{n_1 \cdot n_2 \cdot n_3}} \leq \mu \leq \bar{x}_{...} + t_{f_1; 1-\alpha/2} \cdot \frac{s_1}{\sqrt{n_1 \cdot n_2 \cdot n_3}} \quad (22)$$

wobei $t_{f_1; 1-\alpha/2}$ der Tabellenwert der t -Verteilung bei zweiseitiger Abgrenzung ist [1].

Auch für die Varianzen σ_1^2 , σ_2^2 und σ_3^2 lassen sich Vertrauensbereiche angeben [1].

5 Stichprobenaufbau

Eine vorgeschriebene halbe Weite des Vertrauensbereiches für den Mittelwert der Gesamtheit μ , d. h. ein Vertrauensbereich der Form $\bar{x}_{...} - W \leq \mu \leq \bar{x}_{...} + W$ wird bei außerdem vorgeschriebenem Vertrauensniveau $1 - \alpha$ erreicht, wenn die unten berechneten Anzahlen n_1 , n_2 und n_3 von Einheiten in den drei Stufen untersucht werden.

Aus dem vorgegebenen Wert W berechnet man zum Vertrauensniveau $1 - \alpha$ die Varianz σ_0^2 nach der Gleichung:

$$\sigma_0^2 = \left(\frac{W}{u_{1-\alpha/2}} \right)^2 \quad (23)$$

wobei $u_{1-\alpha/2}$ der Tabellenwert der standardisierten Normalverteilung bei zweiseitiger Abgrenzung ist.

5.1 Stichprobenaufbau bei vorgegebener Weite des Vertrauensbereiches für den Mittelwert der Gesamtheit; varianzbestimmte Stichprobe

Die Varianz s_G^2 für den Mittelwert der Stichprobe $\bar{x}_{...}$ ist:

$$s_G^2 = \frac{\hat{\sigma}_1^2}{n_1} + \frac{\hat{\sigma}_2^2}{n_1 \cdot n_2} + \frac{\hat{\sigma}_3^2}{n_1 \cdot n_2 \cdot n_3} \quad (24)$$

Mit der Bedingung $s_G^2 \leq \sigma_0^2$ erhält man zunächst folgende nicht-ganzzahligen Werte:

$$n'_1 = \frac{\hat{\sigma}_1^2}{\sigma_0^2} \quad n'_2 = \frac{\hat{\sigma}_2^2}{n_{1+} \cdot \sigma_0^2 - \hat{\sigma}_1^2} \quad n'_3 = \frac{\hat{\sigma}_3^2}{n_{1+} \cdot n_{2+} \cdot \sigma_0^2 - n_{2+} \cdot \hat{\sigma}_1^2 - \hat{\sigma}_2^2} \quad (25)$$

Dabei sind n_{1+} und n_{2+} die zur nächst größeren ganzen Zahl **aufgerundeten** Werte n'_1 und n'_2 , oder die um 1 größeren Werte, falls n'_1 und/oder n'_2 ganzzahlig sind. n'_3 wird zu n_{3+} aufgerundet oder bleibt unverändert, falls n'_3 ganzzahlig ist. Weitere Ausführungen zu dieser Rundungsempfehlung siehe Abschnitt 6.

Die Forderung $s_G^2 \leq \sigma_0^2$ wird erfüllt durch die Anzahlen n_{1+} , n_{2+} und n_{3+} in den drei Stufen; dabei ist n_{1+} die kleinstmögliche Anzahl der Einheiten in der ersten Stufe, mit der die Bedingung $s_G^2 \leq \sigma_0^2$ erfüllbar ist. Zu **diesem** n_{1+} gehören **mindestens** die Anzahlen n_{2+} und n_{3+} in der zweiten und dritten Stufe. Der Stichprobenumfang $n_{1+} \cdot n_{2+} \cdot n_{3+}$ ist jedoch trotz des Mindestwertes n_{1+} in der ersten Stufe nicht der kleinstmögliche. Man kann ihn verkleinern, wenn man bei unverändertem n_{1+} den kleineren der beiden Werte n_{2+} und n_{3+} etwas vergrößert und den anderen dieser Werte deutlich verkleinert.

Den gesamten Stichprobenumfang $n_1 \cdot n_2 \cdot n_3$ kann man weiterhin durch Vergrößern von n_{1+} auf verschiedene Werte n_1 ($n_1 > n_{1+}$) verkleinern, wenn man die Anzahlen n_{2+} und n_{3+} beibehält oder einzeln oder beide **verkleinert** ($n_2 \leq n_{2+}$ und $n_3 \leq n_{3+}$).

Setzt man $n_3 = 1$ für $n_1 = n_{1+}$, erhält man die größte Anzahl n_{2++} von Einheiten in der zweiten Stufe, die nicht überschritten zu werden braucht, um die Bedingung $s_G^2 \leq \sigma_0^2$ einzuhalten. Zunächst erhält man:

$$n''_2 = \frac{\hat{\sigma}_2^2 + \hat{\sigma}_3^2}{n_{1+} \cdot \sigma_0^2 - \hat{\sigma}_1^2} \quad (26)$$

Dieser Wert wird auf die nächst größere ganze Zahl n_{2++} **aufgerundet**.

Den kleinsten Stichprobenumfang erhält man, wenn man $n_2 = 1$ und $n_3 = 1$ setzt. Die dazu gehörende Anzahl von Einheiten in der ersten Stufe erhält man aus der Beziehung $s_G^2 = \sigma_0^2$ mit $n_2 = n_3 = 1$ und $n_1 = n'_1$ zu

$$n''_1 = \frac{1}{\sigma_0^2} (\hat{\sigma}_1^2 + \hat{\sigma}_2^2 + \hat{\sigma}_3^2) \quad (27)$$

aus dem man durch Aufrundung den Wert n_{1++} erhält. Diese kleinstmögliche Anzahl n_{1++} ist im allgemeinen die kleinste

die kostengünstigste. Der Stichprobenumfang $n_1 \cdot n_2 \cdot n_3$ liegt damit zwischen den Grenzen $n_{1+} \cdot n_{2+} \cdot n_{3+}$ und n_{1++} . Mit einer Anzahl von Einheiten in der ersten Stufe unterhalb n_{1+} kann man die Bedingung $s_G^2 \leq \sigma_0^2$ nicht mehr einhalten, mit einer Anzahl von Einheiten in der ersten Stufe oberhalb n_{1++} wird s_G^2 zunehmend kleiner gegenüber σ_0^2 .

Die Bereiche für die Anzahl der Einheiten in den drei Stufen sind damit:

$$1 \leq n_1 \leq n_{1+} \leq n_{1++} \quad 1 \leq n_2 \leq n_{2+} \leq n_{2++} \quad 1 \leq n_3 \leq n_{3+} \leq n_{3++} \quad (28)$$

Beginnend mit $(n_{1+}; n_{2+}; n_{3+})$ werden die Wertetripel innerhalb dieser Bereiche variiert. Dabei werden der Stichprobenumfang $n_1 \cdot n_2 \cdot n_3$ sowie alle Wertetripel notiert, für die s_G^2 nur wenig kleiner als σ_0^2 ist.

Nach fallweise verschiedenen Gesichtspunkten wählt man aus den möglichen Zahlentripeln ein Tripel aus.

Falls einer der zunächst als nicht-ganzzahlig berechneten Werte n'_1 , n'_2 und n'_3 kleiner als 1 ist, ist nach Abschnitt 8 vorzugehen.

Beispiel 1:

Bei einer Untersuchung haben sich folgende Schätzwerte für die Varianzen ergeben:

$$\hat{\sigma}_1^2 = 25 \quad \hat{\sigma}_2^2 = 9 \quad \hat{\sigma}_3^2 = 64$$

Auf dem Vertrauensniveau $1 - \alpha$ soll die Varianz $\sigma_0^2 = 3$ nicht überschritten werden. (Für dieses und die folgenden grundsätzlichen Beispiele bleibt die Einheit der Varianzen außer Betracht.) Nach den Gleichungen (25), (26) und (27) berechnet man:

$$n'_1 = 8,3 \text{ aufgerundet auf } n_{1+} = 9$$

$$n'_2 = 4,5 \text{ aufgerundet auf } n_{2+} = 5$$

$$n'_3 = 64 \text{ also } n_{3+} = 64$$

$$n''_1 = 32,7 \text{ aufgerundet auf } n_{1++} = 33$$

$$n''_2 = 36,5 \text{ aufgerundet auf } n_{2++} = 37$$

Die Bereiche, in denen n_1 , n_2 und n_3 variiert werden können, sind damit in diesem Falle:

$$9 \leq n_1 \leq 33 \quad 1 \leq n_2 \leq 37 \quad 1 \leq n_3 \leq 64$$

In der Tabelle 3 sind für dieses Beispiel **sämtliche** günstigen Kombinationen $(n_1; n_2; n_3)$ der Anzahl von Einheiten in den drei Stufen angegeben sowie die dazugehörige Varianz s_G^2 für den Mittelwert der Stichprobe.

Diese Kombinationen erhält man auf folgende Weise: Man beginnt mit dem Wertetripel $(n_{1+}; n_{2+}; n_{3+}) = (9; 5; 64)$. $n_{1+} = 9$ ist die kleinstmögliche Anzahl von Einheiten in der ersten Stufe, mit der man die Bedingung $s_G^2 \leq \sigma_0^2$ einhalten kann; zu **diesem** n_{1+} gehören $n_{2+} = 5$ und $n_{3+} = 64$ Einheiten in den anderen Stufen. Weitere Wertetripel, für die s_G^2 möglichst wenig kleiner als σ_0^2 ist, erhält man, wenn man n_2 um mindestens 1 vergrößert und **gleichzeitig** n_3 um mindestens 1 verkleinert. Für $n_1 = 9$ ergibt sich als nächstes das Wertetripel $(9; 6; 22)$; dabei wurde $n_2 = 6$ als der um 1 vergrößerte vorhergehende Wert angesetzt, und aus der Gleichung (25) wurde $n'_3 = 21,3$ berechnet, also aufgerundet $n_3 = 22$. Genau so verfährt man für $n_1 = 9$ mit den für n_2 nacheinander festgesetzten Werten 7 bis 11. Für $n_2 = 12$ ist $n'_3 = 4,27$, also aufgerundet $n_3 = 5$. Dieses Tripel $(9; 12; 5)$ ergibt gegenüber dem Tripel $(9; 11; 5)$ eine Varianz s_G^2 , die merklich kleiner als 3 ist, weil nicht gleichzeitig n_2 um mindestens 1 größer und n_3 um mindestens 1 kleiner wird. Deshalb wird das Wertetripel $(9; 12; 5)$ weggelassen. Erst das Tripel $(9; 13; 4)$ ergibt wieder einen Wert s_G^2 , der etwas kleiner als 3 ist; gegenüber diesem Tripel werden $(9; 14; 4)$ und $(9; 15; 4)$ weggelassen. Bei $(9; 16; 3)$, $(9; 21; 2)$ und $(9; 37; 1)$ erhält man die letzten drei

Tabelle 3. Wertetripel $(n_1; n_2; n_3)$ und Varianz für den Mittelwert der Stichprobe s_G^2

n_1	n_2	n_3	$n_1 \cdot n_2 \cdot n_3$	s_G^2
9	5	64	2880	3,000
	6	22	1188	2,998
	7	13	819	2,999
	8	10	720	2,992
	9	8	648	2,988
	10	6	540	2,996
	11	5	495	2,998
	13	4	468	2,991
	16	3	432	2,988
	21	2	378	2,994
	37	1	333	2,997
10	2	64	1280	3,000
	3	11	330	2,994
	4	6	240	2,992
	5	4	200	3,000
	7	3	210	2,933
	9	2	180	2,956
	15	1	150	2,987
11	2	10	220	2,973
	3	5	165	2,933
	4	3	132	2,962
	6	2	132	2,894
	10	1	110	2,936
12	1	32	384	3,000
	2	5	120	2,992
	3	3	108	2,926
	4	2	96	2,938
	7	1	84	2,952
13	1	13	169	2,994
	2	4	104	2,885
	3	2	78	2,974
	6	1	78	2,859
14	1	8	112	3,000
	2	3	84	2,869
	3	2	84	2,762
	5	1	70	2,829
15	1	6	90	2,978
	2	3	90	2,678
	3	2	90	2,578
	4	1	60	2,883
16	1	5	80	2,925
	2	2	64	2,844
	4	1	64	2,703
17	1	4	68	2,941
	2	2	68	2,677
	3	1	51	2,902
19	1	3	57	2,912
	2	2	76	2,395
	3	1	57	2,596
22	1	2	44	3,000
	2	1	44	2,796
33	1	1	33	2,970

günstigen Kombinationen für $n_1 = 9$, weil n_3 nicht kleiner als 1 werden kann.

Weiter setzt man $n_1 = 10$ und berechnet aus der Gleichung (25) $n'_2 = 1,8$, das zu $n_2 = 2$ aufgerundet wird; es folgt $n'_3 = 64$, also $n_3 = 64$. Mit diesem Wertetripel (10; 2; 64) beginnend rechnet man weiter wie oben angegeben. (Bei diesem Beispiel taucht zufällig $n_3 = 64$ für $n_1 = 9$ und für $n_1 = 10$ auf.) Beim Wertetripel (10; 15; 1) wird $n_3 = 1$, die weitere Vergrößerung von n_2 über $n_2 = 15$ hinaus ist ungünstig.

Bei $n_{1++} = 33$ erhält man das Wertetripel (33; 1; 1) und als Stichprobenumfang $n_1 \cdot n_2 \cdot n_3 = 33$, den für dieses Beispiel kleinstmöglichen Stichprobenumfang.

Nach von Fall zu Fall unterschiedlichen Gesichtspunkten wählt man eines der in der Tabelle 3 angegebenen Wertetripel oder eines der in der Nähe dieser Werte liegenden Tripel mit **größerem** Gesamtumfang $n_1 \cdot n_2 \cdot n_3$ für die zu ziehende Stichprobe aus. Wünscht man z. B. $n_1 = 10$ und $n_3 = 10$, dann muß $n_2 \geq 4$ gewählt werden; für $n_2 = 3$ wird in diesem Falle $s_G^2 > \sigma_0^2$.

Wie Beispiel 1 zeigt, kann man durch mäßiges Vergrößern von n_1 über n_{1+} hinaus die Werte für n_2 und n_3 verändern, wobei sich in jedem Falle der Stichprobenumfang $n_1 \cdot n_2 \cdot n_3$ drastisch reduziert.

5.2 Stichprobenaufbau bei vorgegebener Weite des Vertrauensbereiches für den Mittelwert der Gesamtheit und minimalen Gesamtkosten; varianzbestimmte kostenminimale Stichprobe

Es seien c_1 und c_2 die Kosten (in willkürlichen Einheiten) für die Bereitstellung einer Einheit in der ersten und in der zweiten Stufe, und c_3 die Kosten für die Entnahme und Herstellung einer Endprobe und ihre Messung. [3] Damit sind die Gesamtkosten der Untersuchung, die durch einen bestimmten Stichprobenumfang minimiert werden sollen:

$$K_G = n_1 \cdot c_1 + n_1 \cdot n_2 \cdot c_2 + n_1 \cdot n_2 \cdot n_3 \cdot c_3 \quad (29)$$

Mit Hilfe der Varianzen $\hat{\sigma}_1^2$, $\hat{\sigma}_2^2$ und $\hat{\sigma}_3^2$ sowie der in Gleichung (23) genannten Varianz σ_0^2 wird die Anzahl der auszuwählenden Einheiten in den drei Stufen für die varianzbestimmte kostenminimale Stichprobe:

$$n_1^* = \frac{1}{\sigma_0^2} \cdot \frac{\hat{\sigma}_1}{\sqrt{c_1}} (\hat{\sigma}_1 \sqrt{c_1} + \hat{\sigma}_2 \sqrt{c_2} + \hat{\sigma}_3 \sqrt{c_3}) \quad (30)$$

$$n_2^* = \frac{\hat{\sigma}_2}{\hat{\sigma}_1} \cdot \sqrt{\frac{c_1}{c_2}} \quad n_3^* = \frac{\hat{\sigma}_3}{\hat{\sigma}_2} \cdot \sqrt{\frac{c_2}{c_3}}$$

Die Werte n_1^* , n_2^* und n_3^* werden **aufgerundet** auf die nächst größeren ganzen Zahlen n_{1+} , n_{2+} und n_{3+} .

Beginnend mit $(n_{1+}; n_{2+}; n_{3+})$ wird die Anzahl der Einheiten in den drei Stufen in der Umgebung von $(n_1^*; n_2^*; n_3^*)$ variiert. Dabei werden der Stichprobenumfang $n_1 \cdot n_2 \cdot n_3$ sowie alle Wertetripel notiert, für die s_G^2 nur wenig kleiner als σ_0^2 und K_G nur wenig größer als $K_{G \min}$ ist. $K_{G \min}$ erhält man aus Gleichung (29), wenn man für die Anzahl von Einheiten n_1^* , n_2^* und n_3^* einsetzt.

Nach fallweise verschiedenen Gesichtspunkten wählt man aus den möglichen Wertetripeln ein Tripel aus.

Falls einer der zunächst als nicht-ganzzahlig berechneten Werte kleiner als 1 ist, ist nach Abschnitt 8 vorzugehen.

Beispiel 2:

Für das Beispiel 1 mit den dort genannten Varianzen mögen die Kosten (in willkürlichen Einheiten) sein:

$$c_1 = 18 \quad c_2 = 6 \quad c_3 = 1$$

Mit den Gleichungen (30) wird jetzt:

$$n_1^* = 14,36 \quad n_2^* = 1,04 \quad n_3^* = 6,53$$

was aufgerundet wird zu

$$n_{1+} = 15 \quad n_{2+} = 2 \quad n_{3+} = 7$$

Da in diesem Falle $n_2^* \approx 1$ ist, ist zunächst zu entscheiden, ob unter diesen Ausgangsbedingungen die Aufteilung nach der zweiten Stufe überhaupt sinnvoll ist. Will man die Aufteilung nach der zweiten Stufe beibehalten, muß man in dieser Stufe eine erhebliche Rundung in Kauf nehmen. In der ersten und der dritten Stufe ist die Rundung in diesem Beispiel ebenfalls relativ groß. Damit entfernt man sich von dem theoretischen Kostenminimum $K_{G \min} = 445,61$, das man aus den nicht gerundeten Werten n_1^* , n_2^* und n_3^* erhält.

In solchen Fällen kann man oft andere Wertetripel $(n_1; n_2; n_3)$ finden, die in der Nähe des Tripels $(n_1^*; n_2^*; n_3^*)$ liegen, für die s_G^2 nur wenig kleiner als σ_0^2 ist, und bei denen die Kosten näher am theoretischen Minimum liegen.

Tabelle 4. Wertetripel $(n_1; n_2; n_3)$, Varianz s_G^2 und Gesamtkosten K_G

n_1	n_2	n_3	$n_1 \cdot n_2 \cdot n_3$	s_G^2	K_G
14,36	1,04	6,53	97,5	3,000	445,6
15	2	7	210	2,271	660
11	2	10	220	2,973	550
	3	5	165	2,933	561
	4	3	132	2,962	594
	6	2	132	2,894	726
	10	1	110	2,936	968
12	1	32	384	3,000	672
	2	5	120	2,992	480
	3	3	108	2,926	540
	4	2	96	2,938	600
	7	1	84	2,952	804
13	1	13	169	2,994	481
	2	4	104	2,885	494
	3	2	78	2,974	546
	6	1	78	2,859	780
14	1	8	112	3,000	448
	2	3	84	2,869	504
	3	2	84	2,762	588
	5	1	70	2,829	742
15	1	6	90	2,978	450
	2	3	90	2,678	540
	3	2	90	2,578	630
	4	1	60	2,883	690
16	1	5	80	2,925	464
	2	2	64	2,844	544
	4	1	64	2,703	736

Geht man von den im Beispiel 1 angegebenen in der Umgebung von $(n_1^*; n_2^*; n_3^*)$ liegenden Wertetripeln aus, erhält man die in Tabelle 4 angegebenen Gesamtkosten K_G .

Mit den Wertetripeln (14; 1; 8), (15; 1; 6) und (16; 1; 5) kann man also die rundungsbedingte Erhöhung der Kosten des zunächst berechneten Tripels (15; 2; 7) weitgehend ausgleichen. Daneben gibt es etliche weitere Tripel, bei denen die Kosten kleiner sind als die des Tripels (15; 2; 7).