HANSER

Regina Storm

Wahrscheinlichkeitsrechnung, mathematische Statistik und statistische Qualitätskontrolle

ISBN-10: 3-446-40906-8
ISBN-13: 978-3-446-40906-4

Leseprobe

Weitere Methoden zur Konstruktion von Schätzungen sind die **Minimum-χ^2-Methode** und die **Momentenmethode** [s. z.B. Müller (Hrsg.; 1991)].

12.3 Konfidenzschätzungen

Mit den Methoden in 12.2 können für unbekannte Parameter der Verteilung der Grundgesamtheit an Hand von Stichproben geeignete Schätzwerte ermittelt werden. Von den gebräuchlichsten dieser Schätzungen sind günstige Eigenschaften bekannt, die sich häufig auf das Verhalten für großen Stichprobenumfang beziehen (z.B. asymptotische Erwartungstreue, Konsistenz, asymptotische Normalverteilung). Hierbei fehlen aber noch Genauigkeitsangaben, insbesondere bezüglich der Abhängigkeit vom Stichprobenumfang. Dazu dienen die **Intervallschätzungen** (allgemeiner **Bereichsschätzungen**). Das sind aus der Stichprobe berechnete Intervalle (oder Bereiche), in denen der wahre, aber unbekannte Parameter der Verteilung der Grundgesamtheit mit großer Wahrscheinlichkeit (Sicherheit) zu erwarten ist. Solche Intervalle bezeichnet man als **Konfidenz-** oder **Vertrauensintervalle**.
Sie werden bei der Lösung praktischer Probleme verwendet, für die folgende Beispiele typisch sind.

Beispiel 12-6: Zur Untersuchung einer bestimmten Stahlsorte bezüglich ihrer Streckgrenze wurden 145 Messungen durchgeführt und daraus das arithmetische Mittel $\bar{x} = 314{,}0$ N/mm^2 berechnet. Es interessiert, in welchen Grenzen die mittlere Streckgrenze dieser Stahlsorte, z.B. bei der Produktion innerhalb eines Quartals, liegt. Nach 12.2.3.3 ist bekannt, daß $\bar{x}$ ein geeigneter Schätzwert für den Mittelwert μ der Grundgesamtheit ist. Kann man mit Hilfe von $\bar{x}$ Genauigkeitsschranken für μ und damit für die mittlere Streckgrenze angeben ?

Beispiel 12-7: Einem Warenposten von einer automatischen Drehmaschine gefertigter Teile mit dem unbekannten Ausschußprozentsatz p wird eine Stichprobe von n Stück entnommen und auf Ausschuß untersucht. Nach 12.2.3.1 kann p durch die relative Häufigkeit eines defekten Stückes in der Stichprobe geschätzt werden. Wie genau ist die Schätzung, und in welchen Grenzen ist der Ausschußprozentsatz des Postens zu erwarten ?

12.3.1 Konfidenzintervalle

Allgemein besteht die Aufgabe der Konfidenzschätzung darin, für den wahren, aber unbekannten Parameter Θ der Verteilungsfunktion F der Grundgesamtheit an Hand einer Stichprobe ein Intervall J anzugeben, das Θ mit einer möglichst großen Wahrscheinlichkeit überdeckt. Bezeichnet man diese Wahrscheinlichkeit mit $\varepsilon = 1 - \alpha$ und die Grenzen des Intervalls J mit G_u und G_o, d.h., gilt $J = (G_u, G_o)$ mit $G_u < G_o$, so bedeutet diese Forderung

$$P(G_u < \Theta < G_o) = \varepsilon = 1 - \alpha. \tag{12.3}$$

(12.3) sagt aus, daß der wahre, aber unbekannte Parameter Θ mit der Wahrscheinlichkeit ε vom Intervall J überdeckt wird.

Das durch (12.3) definierte Intervall $J = (G_u, G_o)$ heißt **Konfidenz-** oder **Vertrauensintervall** für Θ, seine Grenze G_u bzw G_o **untere** bzw. **obere Konfidenzgrenze** und die Wahrscheinlichkeit ε **Konfidenzkoeffizient, Konfidenzniveau** oder **Überdeckungswahrscheinlichkeit**.

Da die von der mathematischen Stichprobe $(X_1, X_2, ..., X_n)$ abhängigen Konfidenzgrenzen $G_u = G_u\,(X_1, X_2, ..., X_n)$ und $G_o = G_o\,(X_1, X_2, ..., X_n)$ in (12.3) (nach den Ausführungen in 11.2) Stichprobenfunktionen und damit Zufallsgrößen sind, ist auch das Intervall J ein **zufälliges** Intervall. Die von der konkreten Stichprobe $(x_1, x_2, ..., x_n)$ abhängigen Realisierungen von G_u und G_o bezeichnen wir mit g_u bzw. g_o.

(12.3) kann folgendermaßen interpretiert werden: Von 100 aus Stichproben derselben Grundgesamtheit mit dem Parameter Θ berechneten Konfidenzintervallen überdecken im Mittel $(1-\alpha) \cdot 100 = \varepsilon \cdot 100$ den wahren Parameter Θ. Nur im Mittel $\alpha \cdot 100$ aller Stichproben liefern Grenzen, die Θ nicht enthalten. ε heißt in diesem Zusammenhang auch **statistische Sicherheit** und α **Irrtumswahrscheinlichkeit**.
Hat man für eine konkrete Stichprobe das Intervall (g_u, g_o) berechnet, so liegt Θ entweder in diesem Intervall oder nicht. Eine Wahrscheinlichkeitsaussage ist hier nicht mehr sinnvoll.
ε bzw. α ist vor Untersuchungsbeginn vorzugeben. Man bevorzugt in der Praxis die Werte

$$\varepsilon = 0{,}95,\ 0{,}99,\ 0{,}999 \text{ bzw. } \alpha = 0{,}05,\ 0{,}01,\ 0{,}001,$$

oder, in Prozent ausgedrückt:

$$\varepsilon = 95\ \%,\ 99\ \%,\ 99{,}9\ \% \text{ bzw. } \alpha = 5\ \%,\ 1\ \%,\ 0{,}1\ \%.$$

Auf die Bedeutung dieser Vorgabe kommen wir in 12.3.2 zurück.
In den folgenden Abschnitten werden einige für die Anwendung wichtige Konfidenzintervalle hergeleitet.

12.3.2 Konfidenzintervall für den Mittelwert μ der normalverteilten Grundgesamtheit bei bekanntem σ^2

Es wird vorausgesetzt, daß das Verteilungsgesetz F der Grundgesamtheit die Normalverteilung mit den Parametern μ und σ^2 ist.
Zunächst sei die Streuung σ^2 als Erfahrungswert bekannt, der Parameter μ – und damit der Mittelwert der Grundgesamtheit – jedoch unbekannt.
Gesucht ist ein Konfidenzintervall für μ an Hand einer Stichprobe vom Umfang n aus der Grundgesamtheit, d.h. nach (12.3) ein Intervall $J = (G_u, G_o)$ mit

$$P(G_u < \mu < G_o) = \varepsilon = 1 - \alpha.$$

Zu seiner Konstruktion gehen wir von der Punktschätzung $\overline{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$ (arithmetisches Mittel) für μ aus. Nach Satz 11-2 in 11.3 besitzt die Stichprobenfunktion

$$Z = \frac{\overline{X} - \mu}{\sigma} \sqrt{n}$$

eine standardisierte Normalverteilung mit der in 5.2.2 definierten Verteilungsfunktion Φ. Damit gilt nach den Ausführungen in 5.2.2 für eine beliebige reelle Zahl κ

$$P(|Z| < \kappa) = P(-\kappa < Z < \kappa) = \Phi(\kappa) - \Phi(-\kappa) = 2\Phi(\kappa) - 1. \tag{12.4}$$

Setzt man nun die rechte Seite von (12.4) gleich $1 - \alpha$, so erhält man

$$2\Phi(\kappa) - 1 = 1 - \alpha \text{ oder}$$

$$\Phi(\kappa) = 1 - \frac{\alpha}{2}.$$

Damit ist κ gerade das Quantil der Ordnung $q = 1 - \frac{\alpha}{2}$ der standardisierten Normalverteilung (vgl.5.2.2). Wir setzen deshalb $\kappa = z_{1-\frac{\alpha}{2}}$. Die sich aus (12.4) ergebende Beziehung

$$P\left(|Z| < z_{1-\frac{\alpha}{2}}\right) = 1 - \alpha \tag{12.5}$$

entspricht (12.3) und führt durch Auflösen der Ungleichung $|Z| = \left|\frac{\overline{X} - \mu}{\sigma} \sqrt{n}\right| < z_{1-\frac{\alpha}{2}}$ nach μ auf das gewünschte **Konfidenzintervall für μ**

$$\overline{X} - z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} < \mu < \overline{X} + z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}. \tag{12.6}$$

Die Konfidenzgrenzen lauten damit

$$G_\mathrm{u} = \overline{X} - z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}, \quad G_\mathrm{o} = \overline{X} + z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$$

und ihre Realisierungen entsprechend

$$g_\mathrm{u} = \bar{x} - z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}, \quad g_\mathrm{o} = \bar{x} + z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}.$$

Die Beziehung (12.5) ist im Bild 12-1 veranschaulicht.

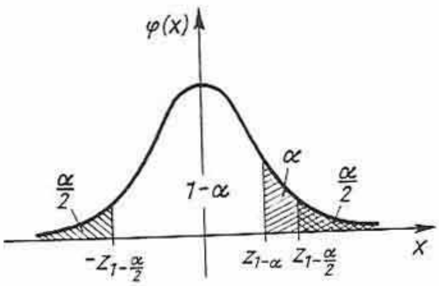

Bild 12-1. Quantile der standardisierten Normalverteilung

Für vorgegebenes ε bzw. α können die Werte $z_{1-\frac{\alpha}{2}} = z_q$ mit Hilfe von Tafel II des Anhangs ermittelt oder direkt aus Tafel III abgelesen werden.

Beispielsweise liest man für $\alpha = 0{,}01$ wegen $q = 1 - \frac{\alpha}{2} = 0{,}995$ aus Tafel III den Wert $z_q = 2{,}576$ (gerundet) ab.

Die **Länge des Konfidenzintervalls** beträgt

$$L = g_\mathrm{o} - g_\mathrm{u} = 2 z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}. \tag{12.7}$$

L ist ein **Maß für die Genauigkeit** der Schätzung: Je kleiner L, um so enger ist das Konfidenzintervall, d.h. um so genauer die Schätzung. Demgegenüber bestimmt ε die **Sicherheit** der Aussage: Größeres ε (kleineres α) bedeutet eine höhere Sicherheit. Da die Werte $z_{1-\frac{\alpha}{2}}$ mit wachsendem ε größer werden (vgl. Tafel III), wächst (bei festem n) auch L, d.h., die Schätzung wird ungenauer. Daraus ist ersichtlich, daß man bei festem n nicht gleichzeitig die Sicherheit und die Genauigkeit der Konfidenzschätzung erhöhen kann. Die Länge L des Konfidenzintervalls hängt auch vom Stichprobenumfang n ab. Sie wird um so kleiner, d.h. die Schätzung um so genauer, je größer n ist.

Beispiel 12-8: Für das oben genannte Beispiel 12-6 (145 Messungen der Streckgrenze einer Stahlsorte) ist unter der Voraussetzung, daß die Streckgrenze eine normalverteilte Zufallsgröße X ist, ein Konfidenzintervall für μ zum Konfidenzniveau $\varepsilon = 0{,}99$ zu bestimmen. Dabei wird angenommen, daß die Streuung σ^2 der Grundgesamtheit aus Erfahrungen bekannt ist: $\sigma^2 = 1000\ (\text{N/mm}^2)^2$. Mit $n = 145$, $\bar{x} = 314{,}0\ \text{N/mm}^2$, $\sigma = 31{,}6\ \text{N/mm}^2$ und $z_{1-\frac{\alpha}{2}} = z_{0{,}995} = 2{,}576$ (wegen $q = 1 - \frac{\alpha}{2} = 0{,}995$) berechnet man die Konfidenzgrenzen

$$g_u = \left(314{,}0 - 2{,}576\frac{31{,}6}{\sqrt{145}}\right)\ \text{N/mm}^2 = 307{,}2\ \text{N/mm}^2,$$

$$g_o = (314{,}0 + 6{,}8)\ \text{N/mm}^2 = 320{,}8\ \text{N/mm}^2$$

und damit das Konfidenzintervall (307,2; 320,8) für die mittlere Streckgrenze der Stahlsorte. Die Länge des Konfidenzintervalls beträgt 13,6 N/mm^2.

Um die Frage zu beantworten, wieviele Messungen mindestens durchzuführen sind, damit sich bei demselben Konfidenzniveau $\varepsilon = 0{,}99$ ein *engeres* Konfidenzintervall, z.B. das Intervall (310,0; 318,0) mit der Länge $L = 8{,}0\ \text{N/mm}^2$ ergibt, hat man (12.7) nach n aufzulösen:

$$n = \frac{4z^2_{1-\frac{\alpha}{2}}\,\sigma^2}{L^2} = \frac{z^2_{1-\frac{\alpha}{2}}\,\sigma^2}{d^2},$$

wobei $d = \frac{L}{2}$ die halbe Länge des Intervalls bezeichnet. Mit diesem **notwendigen Stichprobenumfang** n wird μ mit der Genauigkeit d geschätzt, was auch durch die Angabe $\bar{x} \pm d$ verdeutlicht werden kann.

Für Beispiel 12-8 erhält man

$$n = \frac{4 \cdot 2{,}576^2 \cdot 1000}{8{,}0^2} = 415.$$

Für eine Stichprobe vom Umfang $n = 415$ weicht die mittlere Streckgrenze der bestimmten Stahlsorte höchstens um $\pm d = \pm 4{,}0\ \text{N/mm}^2$ vom arithmetischen Mittel der Stichprobe ab.

Bei Vorgabe des Konfidenzkoeffizienten ε gibt es im Prinzip unendlich viele Möglichkeiten für die Konstruktion eines Konfidenzintervalls. Dies läuft im wesentlichen darauf hinaus, in der grafischen Darstellung für die Dichte der standardisierten Normalverteilung (vgl. Bild 12-1) den Flächenanteil α entsprechend aufzuteilen. In der Praxis sind aber sinnvollerweise nur die beiden folgenden Fälle von Interesse. Im ersten Fall wird eine symmetrische Aufteilung von α vorgenommen, d.h. auf *beiden* Seiten unter

der Kurve jeweils $\frac{\alpha}{2}$ abgeschnitten. Auf diese Weise ergibt sich das (bezüglich $\overline{X}$ symmetrische) **zweiseitige Konfidenzintervall** (12.6).
Sog. einseitige Konfidenzintervalle der Form $(-\infty, G_o')$ oder $(G_u', +\infty)$ erhält man, wenn der gesamte Flächenanteil α nur auf *einer* Seite unter der Kurve abgeschnitten wird. In diesem Fall lautet der (12.5) entsprechende Ansatz z.B. für das Konfidenzintervall $(G_u', +\infty)$ für μ

$$P(Z < z_{1-\alpha}) = 1 - \alpha,$$

wobei $z_{1-\alpha}$ das Quantil der Ordnung $q = 1 - \alpha$ der standardisierten Normalverteilung bezeichnet. Es kann wiederum für ausgewähltes ε bzw. α aus Tafel III des Anhangs abgelesen werden. Dabei gilt

$$z_{1-\alpha} < z_{1-\frac{\alpha}{2}} \qquad \text{(s. auch Bild 12-1).}$$

Durch Auflösen der Ungleichung $Z = \frac{\overline{X} - \mu}{\sigma} \sqrt{n} < z_{1-\alpha}$ nach μ erhält man das **einseitige Konfidenzintervall** $(G_u', +\infty)$ für μ mit der unteren Konfidenzgrenze

$$G_u' = \overline{X} - z_{1-\alpha} \frac{\sigma}{\sqrt{n}}.$$

Entsprechende Überlegungen führen auf das **einseitige Konfidenzintervall** $(-\infty, G_o')$ für μ mit der oberen Konfidenzgrenze

$$G_o' = \overline{X} + z_{1-\alpha} \frac{\sigma}{\sqrt{n}}.$$

Für Beispiel 12-8 berechnet man mit den dort angegebenen Werten und $\varepsilon = 0{,}99$, d.h., $z_{1-\alpha} = z_{0{,}99} = 2{,}326$, als einseitige Konfidenzintervalle für die mittlere Streckgrenze:

$$g_u' = \left(314{,}0 - 2{,}326 \frac{31{,}6}{\sqrt{145}}\right) \text{N/mm}^2 = 307{,}9 \text{ N/mm}^2, \text{ d.h. } (307{,}9; +\infty)$$

bzw.

$$g_o' = (314{,}0 + 6{,}1) \text{ N/mm}^2 = 320{,}1 \text{ N/mm}^2, \text{ d.h. } (-\infty; 320{,}1).$$

12.3.3 Konfidenzintervall für den Mittelwert μ der normalverteilten Grundgesamtheit bei unbekanntem σ^2

Bei Bestimmung eines Konfidenzintervalls für den Mittelwert μ der normalverteilten Grundgesamtheit hat man häufig bei praktischen Problemen keine Kenntnis über die Streuung σ^2 der Grundgesamtheit. Man muß deshalb für σ^2 eine geeignete Schätzung aus der Stichprobe verwenden. Nach den Ausführungen von 12.2 bietet sich als Schätzwert für σ^2 die empirische Streuung der Stichprobe $s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \overline{x})^2$ an.