

Induktive Statistik – Rückschlüsse von einer Stichprobe auf die Allgemeinheit ziehen

18



Wie kann man aus
Stichproben Rückschlüsse
auf die Allgemeinheit
ziehen?

Was ist ein signifikantes
Ergebnis?

Was ist ein t-Test?

18.1	Statistische Schätzverfahren	452
18.2	Hypothesentests	459
	Aufgaben	472

Wie bereits angesprochen geht es bei der induktiven Statistik darum, aus einer Stichprobe Rückschlüsse auf die zugrunde liegende Grundgesamtheit zu ziehen bzw. Vermutungen über die Grundgesamtheit zu stützen oder zu verwerfen.

Im Wesentlichen stehen hierbei zwei Fragestellungen im Mittelpunkt. Zuerst werden wir besprechen, wie man unbekannte Parameter einer Verteilung schätzt, wenn man zwar die Art der Verteilung kennt, aber nicht die notwendigen Parameter. Man weiß z. B., dass eine Zufallsvariable normalverteilt ist, kennt aber μ nicht.

Danach testen wir Vermutungen(Hypothesen). Das ist immer dann wichtig, wenn man in Verhandlungen seinen Standpunkt vertreten muss, z. B. bei der Qualitätskontrolle bei Lieferungen. Auch hier gehen wir davon aus, dass man nicht die komplette Lieferung untersuchen kann, weil sie entweder zu groß ist oder man bei der Überprüfung die Produkte zerstört.

18.1 Statistische Schätzverfahren

Es gibt zwei Methoden, wie man unbekannte Parameter schätzen kann. Zum einen kann man durch bestimmte Berechnungsverfahren einen Wert für den unbekannt Parameter ermitteln (Punktschätzungen). Das hat den Vorteil, dass man mit der Zahl direkt weiterrechnen kann, hat aber den Nachteil, dass man nicht weiß, inwiefern man sich auf die Zahl verlassen kann.

Zum anderen kann man ein Intervall berechnen, in dem sich der unbekannte Parameter mit einer vorgegebenen Wahrscheinlichkeit befindet (Konfidenzintervalle bzw. Intervallschätzungen). Diese Methode hat den Nachteil, dass man mit Intervallen schlecht weiter rechnen kann, bietet aber direkt ein Gütemaß mit an, nämlich die Wahrscheinlichkeit, dass der unbekannte Parameter wirklich in dem Intervall liegt. (Ohne vorgeifen zu wollen, aber diese Wahrscheinlichkeit heißt Konfidenzniveau.)

Werte für unbekannte Parameter in Wahrscheinlichkeitsverteilungen erhält man mithilfe von Punktschätzungen

Die Idee der **Punktschätzung** ist die folgende: Man denkt sich eine Funktion $\hat{\vartheta}$ aus, von der man annimmt, dass sie bestimmte Eigenschaften besitzt, und die auf einer zu ziehenden Stichprobe basiert: $\hat{\vartheta} = \hat{\vartheta}(X_1, X_2, \dots, X_n)$.

In diese Schätzfunktion $\hat{\vartheta}$ setzt man die Messwerte der Stichprobe ein und erhält hierdurch einen Wert für den unbekannt Parameter ϑ , zumindest wenn man die Schätzfunktion geschickt gewählt hat.

Sinnvoll sind Schätzfunktionen, die eine oder mehrere der folgenden Eigenschaften besitzen.

Definition: Erwartungstreue, Verzerrung und Konsistenz

- Eine Schätzfunktion $\hat{\vartheta}(X_1, X_2, \dots, X_n)$ heißt **erwartungstreu** für einen unbekannt Parameter ϑ , wenn gilt:

$$E[\hat{\vartheta}(X_1, X_2, \dots, X_n)] = \vartheta$$

Das bedeutet, man fordert, dass die Werte der Schätzfunktion zumindest im Mittel tatsächlich den unbekannt Parameter als Ergebnis haben sollen.

- $b_n = E[\hat{\vartheta}] - \vartheta$ heißt **Verzerrung** von $\hat{\vartheta}$. Hiermit kann man den Unterschied zwischen dem Erwartungswert der Schätzfunktion und dem unbekannt Parameter angeben.
- Eine Schätzfunktion $\hat{\vartheta}(X_1, X_2, \dots, X_n)$ heißt **asymptotisch erwartungstreu** für einen unbekannt Parameter ϑ , wenn gilt:

$$\lim_{n \rightarrow \infty} E[\hat{\vartheta}(X_1, X_2, \dots, X_n)] = \vartheta$$

Diese Eigenschaft erfordert also, dass der Erwartungswert der Schätzfunktion zumindest für unendlich große Stichproben dem unbekannt Parameter, den es zu schätzen gilt, entspricht.

- Die Folge der Schätzfunktionen $\hat{\vartheta}_n$ heißt **konsistent** für ϑ , wenn gilt:

$$\hat{\vartheta}(X_1, \dots, X_n) \xrightarrow{\text{stoch}} \vartheta$$

Das bedeutet, dass sich die Ergebnisse der Schätzfunktion gemäß der oben erläuterten stochastischen Konvergenz immer mehr dem unbekannt Parameter annähern.

Dass eine Schätzfunktion erwartungstreu sein soll, leuchtet unmittelbar ein. Schließlich sollte im langfristigen Mittel der unbekannt Parameter bei der Rechnung herauskommen. Wenn das nicht exakt klappt, dann doch zumindest rein hypothetisch im Unendlichen, auch wenn man in der Realität nicht unendlich viele Elemente in die Stichprobe packen kann. Dann hätten wir aber zumindest die Eigenschaft der asymptotischen Erwartungstreue erfüllt. Um die Konsistenz zu verstehen, müssen wir grob eine Ahnung entwickelt haben, was die stochastische Konvergenz aus Definition 17.8 eigentlich bedeutet: Bei der stochastischen Konvergenz wird gefordert, dass eine Stichprobe mit der Wahrscheinlichkeit 1 (was nicht ganz genau dasselbe ist wie auf jeden Fall) so nahe wie möglich an ihrem Grenzwert liegt. Also besagt die Eigenschaft der Konsistenz, dass sich die Schätzfunktion mit 100 %iger Wahrscheinlichkeit dem unbekannt Parameter nähern soll.

Anhand dieser Eigenschaften können wir nun entscheiden, ob eine gegebene Schätzfunktion sinnvoll ist. Zusätzlich gibt es noch ein Qualitätskriterium, das auf der Varianz basiert. Denn

zumindest bei erwartungstreuen Schätzern ist man natürlich daran interessiert, dass die Varianz der Schätzfunktion möglichst gering ist. Das bedeutet, dass die Werte der Schätzfunktion im Durchschnitt nur wenig vom unbekanntem Parameter ϑ abweichen, und stellt ein allgemein akzeptiertes Maß für die Zuverlässigkeit der Schätzung dar.

Definition: Mittleres Fehlerquadrat

Das **mittlere Fehlerquadrat** einer Schätzfunktion $\hat{\vartheta}_n$ ist definiert als

$$\text{MFQ}(\hat{\vartheta}_n) = E \left[(\hat{\vartheta}_n(X_1, \dots, X_n) - \vartheta)^2 \right]$$

Wenn $\hat{\vartheta}_n$ erwartungstreu ist, gilt:

$$\text{MFQ}(\hat{\vartheta}_n) = \text{Var}(\hat{\vartheta}_n).$$

Für wichtige Parameter gibt es erprobte Punktschätzer

Bei einigen unbekanntem Parametern ist schon längst bekannt, wie gute bzw. die besten Schätzfunktionen aussehen. Diese muss man also lediglich kennen und kann sie dann immer verwenden.

Wichtige Punktschätzer

- Sei X eine Zufallsvariable mit unbekanntem Erwartungswert $\vartheta = \mu = E[X]$, der aus n unabhängigen Wiederholungen X_1, \dots, X_n von X geschätzt werden soll. Dann ist das arithmetische Mittel der Beobachtungswerte x_1, \dots, x_n

$$\hat{\mu}_n(x_1, \dots, x_n) = \frac{1}{n} \sum_{i=1}^n x_i = \bar{x}_n$$

eine erwartungstreu und konsistente Schätzfunktion für μ mit $\text{MFQ}(\hat{\mu}_n) = \frac{\sigma^2}{n}$. Unter allen linearen Schätzfunktionen ist diese Schätzfunktion diejenige mit der geringsten Varianz.

- Sei X eine Zufallsvariable mit unbekannter Varianz $\vartheta = \sigma^2 = \text{Var}(X)$, die aus n unabhängigen Wiederholungen X_1, \dots, X_n von X geschätzt werden soll. Dann ist die Stichprobenvarianz

$$\hat{\sigma}_n^2(x_1, \dots, x_n) = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x}_n)^2$$

eine erwartungstreu und konsistente Schätzfunktion für σ^2 .

Achtung Die in der deskriptiven Statistik gebräuchliche empirische Varianz s^2 ist nicht erwartungstreu für σ^2 , sondern lediglich asymptotisch erwartungstreu. ◀

- Sei X eine Zufallsvariable mit unbekannter Standardabweichung $\vartheta = \sigma$, die aus n unabhängigen Wiederholungen X_1, \dots, X_n von X geschätzt werden soll. Dann ist die Stichprobenstandardabweichung

$$\hat{\sigma}_n(x_1, \dots, x_n) = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x}_n)^2}$$

eine asymptotisch erwartungstreu Schätzfunktion für σ .

Zur Berechnung von $\hat{\sigma}_n^2$ verwendet man üblicherweise die Formel

$$\hat{\sigma}_n^2 = \frac{n}{n-1} (\overline{x^2} - \bar{x}^2)$$

Den Ausdruck in der Klammer kennen wir bereits, er bedeutet nichts anderes als s^2 . Also können wir auch festhalten, dass gilt:

$$\hat{\sigma}_n^2 = \frac{n}{n-1} \cdot s^2.$$

Mit der Maximum-Likelihood-Methode kann man Punktschätzer konstruieren

Nun kennen wir einige wichtige Schätzfunktionen, und wir sind in der Lage, gegebene Schätzfunktionen auf verschiedene Gütemaße hin zu untersuchen. Es bleibt noch, dass wir selbst in der Lage sind, sinnvolle Schätzfunktionen zu konstruieren. Dazu lernen wir nun die **Maximum-Likelihood-Methode (ML-Methode)** kennen. Maximum Likelihood bedeutet hierbei größte Wahrscheinlichkeit, und der Name ist, wie wir gleich feststellen werden, Programm.

Die Idee der ML-Methode besteht darin, den unbekanntem Parameter ϑ so zu wählen bzw. zu berechnen, dass die Wahrscheinlichkeit, mit diesem Wert für den unbekanntem Parameter die gegebene Stichprobe zu erhalten, am größten, also maximal wird. Diese Idee muss man sich durch den Kopf gehen lassen, dann erklärt sich die folgende Vorgehensweise eigentlich sehr gut.

Vorgehensweise der Maximum-Likelihood-Methode

Gegeben sei eine Stichprobe x_1, \dots, x_n als Realisierung der Zufallsvariablen X_1, \dots, X_n . Wir nehmen an, dass wir

die Art der Verteilung kennen, aber wie immer in diesem Kapitel ist ein Parameter, nämlich ϑ , unbekannt und soll geschätzt werden.

Zur Durchführung der ML-Methode müssen wir folgende Schritte durchführen:

1. Wir bestimmen die **Likelihood-Funktion** $L(\vartheta)$.

Im diskreten Fall lautet

$$L(\vartheta) = \prod_{i=1}^n P(X = x_i),$$

im stetigen Fall lautet sie

$$L(\vartheta) = \prod_{i=1}^n f(x_i).$$

Hierdurch berechnen wir die gemeinsame Wahrscheinlichkeitsfunktion bzw. die gemeinsame Dichte der o. g. Zufallsvariablen, also (zumindest im diskreten Fall) die Wahrscheinlichkeit dafür, dass die gegebenen Ergebnisse tatsächlich alle gemeinsam erzielt wurden. Wegen der unterstellten Unabhängigkeit der Elemente einer Stichprobe kann man diese gemeinsame Wahrscheinlichkeit bzw. Dichte als Produkt der einzelnen Wahrscheinlichkeiten bzw. Dichten berechnen. Übrigens steht da mit voller Absicht jeweils x_i . Wir müssen also bei den gegebenen Wahrscheinlichkeitsfunktionen und Dichten erst das k bzw. x durch x_i ersetzen.

2. Nun soll der Parameter ϑ so bestimmt werden, dass eben die Funktion $L(\vartheta)$ maximal wird. Das bedeutet, wir müssten eigentlich $L'(\vartheta)$ berechnen, gleich 0 setzen und prüfen, ob wir tatsächlich das globale Maximum gefunden haben. Leider ist $L(\vartheta)$ ein Produkt, und beim Ableiten von Produkten muss man die Produktregel beachten, was hier oftmals recht aufwendig wird. Daher macht man normalerweise einen Kunstgriff: Statt $L(\vartheta)$ zu betrachten, ermittelt man die **Log-Likelihood-Funktion** $l(\vartheta)$, das ist

$$l(\vartheta) = \ln L(\vartheta).$$

Im diskreten Fall ist

$$l(\vartheta) = \ln L(\vartheta) = \sum_{i=1}^n \ln(P(X = x_i)).$$

Im stetigen Fall berechnet sich $l(\vartheta)$ durch

$$l(\vartheta) = \ln L(\vartheta) = \sum_{i=1}^n \ln(f(x_i)).$$

Diese Funktion lässt sich viel einfacher ableiten, hat aber an derselben Stelle ihr globales Maximum wie die Likelihood-Funktion. Das liegt daran, dass der Logarithmus eine monoton wachsende Funktion ist. Der

Wert des Maximums ist logischerweise ein anderer, nämlich der Logarithmus des ursprünglichen Wertes, aber dieser interessiert uns sowieso nicht.

3. Nun müssen wir all unser Wissen über Logarithmusgesetze ausgraben und $l(\vartheta)$ vereinfachen. Schließlich wollen wir uns beim Ableiten nicht übernehmen.
4. Aus dem gleichen Grund wenden wir auch alles an, was wir noch zur Vereinfachung von Summen wissen.
5. Jetzt bilden wir $l'(\vartheta)$.

Achtung Immer daran denken, dass wir nach ϑ ableiten. Alle anderen Buchstaben, die bei $l(\vartheta)$ auftauchen, sind nur Konstanten. ◀

6. Jetzt suchen wir das globale Maximum von $l(\vartheta)$. Hierzu macht man am einfachsten eine Monotoniebeurteilung und prüft, wo $l(\vartheta)$ steigt und wo $l(\vartheta)$ fällt, also wo $l'(\vartheta) \geq 0$ bzw. $l'(\vartheta) \leq 0$. An den Stellen, an denen ein Wechsel von Steigen zu Fallen stattfindet, verstecken sich (zumindest bei stetigen Funktionen) lokale Maxima.
7. Nun überprüfen wir, ob es sich bei dem Kandidaten um ein globales Maximum handelt, wir suchen also die Stelle, an der $l(\vartheta)$ den absolut größten Wert annimmt. Hierzu muss man alle lokalen Maxima und die Ränder des Definitionsbereichs berücksichtigen, um sicherzustellen, dass man wirklich ein globales Maximum erhält. Die Stelle, an der $l(\vartheta)$ den größten Funktionswert annimmt, ist das globale Maximum der Funktion.
8. Das gefundene globale Maximum nennen wir nun noch $\hat{\vartheta}_n$. Wir haben die Schätzfunktion für den unbekannten Parameter ϑ gefunden.

Intervallschätzungen liefern eine Vertrauenswahrscheinlichkeit

Punktschätzer können wir nun berechnen. Allerdings haben diese, wie oben schon angesprochen, den kleinen Nachteil, dass man die Zahl, die herauskommt, glauben muss oder nicht. Es gibt keine Garantie, dass der Wert mit dem wirklichen Parameter übereinstimmt. Etwas anders sieht das aus, wenn man darauf verzichtet, eine Zahl als Ergebnis zu haben, sondern sich mit einem Intervall zufrieden gibt. Diese Intervalle heißen **Konfidenzintervalle** oder **Konfidenzbereiche**.

Der wirklich wahre Parameter soll also innerhalb gewisser Grenzen liegen. Dann kann man zusätzlich noch die Wahrscheinlichkeit berechnen, dass das Intervall tatsächlich den unbekannt Parameter überdeckt, bzw. – und das ist noch viel besser – man kann diese Wahrscheinlichkeit vorgeben, sozusagen als Anspruchsniveau, und dann ein Intervall berechnen, das dieses Anspruchsniveau erfüllt.

Anwendung: Maximum-Likelihood-Methode

In der Qualitätskontrolle eines Betriebs, der Glühbirnen produziert, sollen diese auf Funktionsfähigkeit (intakt/defekt) überprüft werden. Dazu wird der Produktion eine Stichprobe vom Umfang n entnommen, die durch die stochastisch unabhängigen Zufallsgrößen X_1, X_2, \dots, X_n beschrieben wird, wobei $X_i = 1$, falls Birne i defekt ist, und $X_i = 0$, falls Birne i intakt ist. Dabei ist die Ausschusswahrscheinlichkeit $p = P(X_i = 1)$ natürlich von großer Wichtigkeit und soll daher geschätzt werden. Es gilt also $X_i \sim \text{Bin}(1, p)$, $i = 1, \dots, n$; $0 < p < 1$, wobei p unbekannt ist.

Um p mit der ML-Methode zu schätzen, brauchen wir noch einige Informationen zur Verteilung der X_i . Die Wahrscheinlichkeitsfunktion der $\text{Bin}(1, p)$ -Verteilung lautet $P(X = k) = p^k \cdot (1-p)^{1-k}$, $k \in \{0, 1\}$ mit $E[X] = p$ und $\text{Var}(X) = p(1-p)$. Die Likelihood-Funktion lautet also

$$L(p) = \prod_{i=1}^n P(X = x_i) = \prod_{i=1}^n p^{x_i} \cdot (1-p)^{1-x_i}.$$

Wir bilden die Log-Likelihood-Funktion:

$$\begin{aligned} l(p) &= \ln L(p) = \ln \prod_{i=1}^n P(X = x_i) = \sum_{i=1}^n \ln P(X = x_i) \\ &= \sum_{i=1}^n \ln p^{x_i} \cdot (1-p)^{1-x_i}. \end{aligned}$$

Das sieht zwar auch schwierig aus, aber wenn man die Logarithmusgesetze und Summenregeln beherrscht, kann man einiges vereinfachen.

Wir vereinfachen mit den Logarithmusgesetzen:

$$\begin{aligned} l(p) &= \ln L(p) = \ln \prod_{i=1}^n P(X = x_i) = \sum_{i=1}^n \ln P(X = x_i) \\ &= \sum_{i=1}^n \ln p^{x_i} \cdot (1-p)^{1-x_i} = \sum_{i=1}^n \ln p^{x_i} + \ln(1-p)^{1-x_i} \\ &= \sum_{i=1}^n x_i \cdot \ln p + (1-x_i) \cdot \ln(1-p). \end{aligned}$$

Und nun vereinfachen wir noch mithilfe der Summenregeln:

$$\begin{aligned} l(p) &= \ln L(p) = \ln \prod_{i=1}^n P(X = x_i) = \sum_{i=1}^n \ln P(X = x_i) \\ &= \sum_{i=1}^n x_i \cdot \ln p + (1-x_i) \cdot \ln(1-p) \end{aligned}$$

$$\begin{aligned} &= \sum_{i=1}^n x_i \cdot \ln p + \sum_{i=1}^n (1-x_i) \cdot \ln(1-p) \\ &= \ln p \sum_{i=1}^n x_i + \ln(1-p) \sum_{i=1}^n (1-x_i) \end{aligned}$$

Das kann man ableiten:

$$l'(p) = 1/p \sum_{i=1}^n x_i - 1/(1-p) \sum_{i=1}^n (1-x_i)$$

Damit führen wir eine Monotoniebetrachtung durch. Wir interessieren uns dafür, wo $l(p)$ steigt, wo also $l'(p) \geq 0$ ist:

$$\begin{aligned} l'(p) &= \frac{1}{p} \sum_{i=1}^n x_i - \frac{1}{1-p} \sum_{i=1}^n (1-x_i) \geq 0 \\ \Leftrightarrow (1-p) \sum_{i=1}^n x_i - p \sum_{i=1}^n (1-x_i) &\geq 0 \\ \Leftrightarrow \sum_{i=1}^n x_i - p \sum_{i=1}^n x_i - p \sum_{i=1}^n 1 + p \sum_{i=1}^n x_i &\geq 0 \\ \Leftrightarrow \sum_{i=1}^n x_i - p \sum_{i=1}^n 1 &\geq 0 \\ \Leftrightarrow \sum_{i=1}^n x_i - n \cdot p &\geq 0 \\ \Leftrightarrow \sum_{i=1}^n x_i &\geq n \cdot p \\ \Leftrightarrow p &\leq \frac{1}{n} \sum_{i=1}^n x_i \end{aligned}$$

Bis zum arithmetischen Mittel steigt die Funktion, ab dort fällt sie. Damit liegt bei $p = \bar{x}$ ein globales Maximum, und $\hat{p} = \bar{x}$ ist der gesuchte ML-Schätzer.

Das ist auch nicht verwunderlich, denn wir haben oben bereits festgehalten, dass bei der $\text{Bin}(1, p)$ -Verteilung $E[X] = p$ gilt. Und gerade eben haben wir gesehen, dass das arithmetische Mittel der beste Schätzer für den Erwartungswert, in diesem Fall für p , ist.

Definition: Konfidenzniveau, $(1 - \alpha)$ -Konfidenzintervall

Die Wahrscheinlichkeit, dass ein Konfidenzintervall einen unbekanntem Parameter enthält, heißt **Vertrauenswahrscheinlichkeit** bzw. **Konfidenzniveau** und wird mit $1 - \alpha$ abgekürzt. α heißt auch **Misstrauenswahrscheinlichkeit**. Ein Konfidenzintervall, das einen unbekanntem Parameter mit Wahrscheinlichkeit $1 - \alpha$ enthält, heißt $(1 - \alpha)$ -Konfidenzintervall.

Wir wollen uns an dieser Stelle nicht mit der Theorie beschäftigen, die hinter Konfidenzintervallen steckt. Wichtiger ist, dass wir in der Lage sind, zu gegebenen Problemstellungen und unbekanntem Parametern das geeignete Konfidenzintervall zu berechnen. Dazu ist es ausgesprochen wichtig, dass man die Aufgabenstellung genau liest bzw. in der Realität die Problemstellung genau analysiert. Denn je nachdem, welcher Parameter geschätzt werden soll und welche Verteilung unterstellt wird, muss man völlig unterschiedliche Berechnungen anstellen.

Wenn wir ein Konfidenzintervall für den unbekanntem Erwartungswert einer Verteilung suchen, müssen wir berücksichtigen, mit welcher Verteilung wir es zu tun haben. Beginnen wir mit der Situation, dass wir guten Gewissens unterstellen können, dass wir es mit einer Normalverteilung zu tun haben. In der Realität müssen wir das zuerst testen. Wie das geht, lernen wir in Abschn. 18.2.

Konfidenzintervalle für den Erwartungswert einer Normalverteilung bei bekannter Varianz

Wenn die Stichprobe normalverteilt und die Standardabweichung σ bekannt ist, dann ist ein zweiseitiges $1 - \alpha$ -Konfidenzintervall für μ das Intervall:

$$\left[\bar{x} - x_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}; \bar{x} + x_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}} \right]$$

Das einseitige untere $(1 - \alpha)$ -Konfidenzintervall für μ lautet:

$$\left(-\infty; \bar{x} + x_{1-\alpha} \cdot \frac{\sigma}{\sqrt{n}} \right]$$

Das einseitige obere $(1 - \alpha)$ -Konfidenzintervall für μ lautet:

$$\left[\bar{x} - x_{1-\alpha} \cdot \frac{\sigma}{\sqrt{n}}; \infty \right)$$

In der Realität kann man diese Konfidenzintervalle nicht sehr oft brauchen, denn man ist selten in der Situation, dass man den Erwartungswert nicht kennt, aber die Varianz, für die man den Erwartungswert braucht, schon.

Aber für akademische Zwecke sind diese Konfidenzintervalle gut geeignet. Man sollte, wenn man bis hierhin alles verstanden hat, selbstständig in der Lage sein, die obige Formel anzuwenden. Da steht nichts drin, was wir noch nicht kennen: Mittelwert,

Quantil der Standardnormalverteilung (x_p). Der Rest muss angegeben sein, sonst hat man das falsche Intervall gewählt.

Sehen wir uns ein Beispiel an:

Beispiel

Auf zwölf Versuchsfeldern wurden neue Rapsorten zur Produktion von Biokraftstoff angebaut. Die Flächen erbrachten die Hektarerträge (in dz):

35.6, 33.7, 37.8, 31.2, 37.2, 34.1,
35.6, 36.6, 37.1, 34.9, 35.6, 34.0

Aus Erfahrung ist bekannt, dass die Hektarerträge als Realisierung unabhängiger $N(\mu, \sigma)$ -verteilter Zufallsgrößen angesehen werden können. Gehen wir davon aus, dass man (woher auch immer) weiß, dass $\sigma^2 = 3.24$ gilt. Dann können wir ein zweiseitiges 0.95-Konfidenzintervall für den durchschnittlichen Hektarertrag (μ) wie folgt ausrechnen:

$$\begin{aligned} \bar{x} &= 35.2833, \quad n = 12, \\ \alpha = 0.05 &\Rightarrow x_{1-\frac{\alpha}{2}} = x_{0.975} = 1.96, \\ \sigma &= \sqrt{3.24} = 1.8 \end{aligned}$$

Also gilt:

$$\begin{aligned} &\left[\bar{x} - x_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}; \bar{x} + x_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}} \right] \\ &= \left[35.2833 - 1.96 \cdot \frac{1.8}{\sqrt{12}}; 35.2833 + 1.96 \cdot \frac{1.8}{\sqrt{12}} \right] \\ &= [34.265; 36.302] \end{aligned}$$

Der durchschnittliche Ertrag der Rapsfelder liegt also mit einer Wahrscheinlichkeit von 95 % zwischen 34.265 und 36.302 Doppelzentner pro Hektar. ◀

Prinzipiell kann man die Länge eines beliebigen zweiseitigen Konfidenzintervalls ganz leicht durch „obere Grenze – untere Grenze“ ausrechnen. In dem soeben besprochenen Fall kann man diese Formel aber nutzen, um den benötigten Stichprobenumfang zu schätzen.

Die Länge eines zweiseitigen Konfidenzintervalls für μ bei Normalverteilung und bekannter Varianz beträgt

$$l = 2 \cdot x_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$$

Wenn ein solches Konfidenzintervall höchstens die Länge l besitzen soll, berechnet sich der Mindeststichprobenumfang wie folgt:

$$n \geq 4 \cdot \sigma^2 \cdot \left(\frac{x_{1-\frac{\alpha}{2}}}{l} \right)^2$$

Beispiel

Kommen wir noch einmal auf die zwölf Rapsfelder aus dem obigen Beispiel zurück. Das Konfidenzintervall, das wir berechnet haben, hat eine Länge von

$$l = 2 \cdot x_{1-\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} = 2 \cdot 1.96 \cdot \frac{\sqrt{3.24}}{\sqrt{12}} = 2.0369.$$

Gehen wir nun davon aus, dass das Konfidenzintervall zu lang ist und dass wir ein Konfidenzintervall haben wollen, das die Länge 0.5 besitzt. An der Verlässlichkeit wollen wir keine Abstriche machen, daher müssen wir sicherlich eine größere Stichprobe ziehen.

Genauer gesagt brauchen wir eine Stichprobe vom Umfang

$$n \geq 4 \cdot \sigma^2 \cdot x_{1-\frac{\alpha}{2}}^2 l^2 = 4 \cdot 3.24 \cdot \left(\frac{1.96}{0.5}\right)^2 = 199.148.$$

Also beträgt der Mindeststichprobenumfang $n = 200$, mit dem dies gewährleistet werden kann. ◀

Konfidenzintervalle für den Erwartungswert einer Normalverteilung bei unbekannter Varianz

Wenn die Stichprobe normalverteilt und die Standardabweichung σ unbekannt ist, muss man zuerst die Standardabweichung bzw. die Varianz schätzen, also $\hat{\sigma}_n^2 = \frac{n}{n-1} \cdot (\overline{x^2} - (\bar{x})^2)$ berechnen.

Dann ist das Intervall

$$\left[\bar{x} - t_{n-1} \left(1 - \frac{\alpha}{2}\right) \cdot \frac{\hat{\sigma}_n}{\sqrt{n}}; \bar{x} + t_{n-1} \left(1 - \frac{\alpha}{2}\right) \cdot \frac{\hat{\sigma}_n}{\sqrt{n}} \right]$$

ein zweiseitiges $(1 - \alpha)$ -Konfidenzintervall für μ .

Das einseitige untere $(1 - \alpha)$ -Konfidenzintervall für μ lautet:

$$\left(-\infty; \bar{x} + t_{n-1} (1 - \alpha) \cdot \frac{\hat{\sigma}_n}{\sqrt{n}} \right]$$

Das einseitige obere $(1 - \alpha)$ -Konfidenzintervall für μ lautet:

$$\left[\bar{x} - t_{n-1} (1 - \alpha) \cdot \frac{\hat{\sigma}_n}{\sqrt{n}}; \infty \right)$$

Hierbei bezeichnet $t_{n-1} \left(1 - \frac{\alpha}{2}\right)$ das $\left(1 - \frac{\alpha}{2}\right)$ -Quantil der **t-Verteilung**, die wie die Standardnormalverteilung tabelliert vorliegt. Im Anhang in Tab. A.12 sind allgemein Werte von $t_f(p)$ tabelliert. Das bedeutet, wir müssen in der ersten Spalte das p suchen, in der ersten Zeile das f , und im Schnittpunkt befindet sich der gesuchte Wert.

Beispiel

Gehen wir von fünf Messwerten aus und einem Konfidenzniveau $1 - \alpha = 0.95$. Damit ist $\alpha = 0.05$. Dann benötigen wir für das zweiseitige Konfidenzintervall für μ bei unbekannter Varianz

$$t_{5-1} \left(1 - \frac{0.05}{2}\right) = t_4(0.975) = 2.776450856. \quad \blacktriangleleft$$

Die Student t-Verteilung geht auf den englischen Chemiker *William Sealy Gosset* (1876–1937) zurück, der bei der Guinness-Brauerei beschäftigt war und sich mit der Statistik kleiner Stichproben beschäftigte. Er veröffentlichte sein Verteilungsmodell 1908 unter dem Pseudonym *Student*.

Wenn X_1, X_2, \dots, X_n eine Folge von Stichprobenvariablen ist, die jeweils normalverteilt mit den Parametern μ und σ sind, wobei σ unbekannt ist, dann ist der Stichprobenmittelwert \bar{X} t-verteilt mit $n - 1$ Freiheitsgraden. Freiheitsgrade werden üblicherweise in der Statistik mit *df* (degrees of freedom) abgekürzt.

Beispiel

Gehen wir davon aus, dass bei den zwölf Rapsfeldern die Varianz nicht bekannt ist, denn wie bereits erwähnt ist diese Annahme eher unrealistisch. Wenn wir aber begründet unterstellen können, dass eine Normalverteilung vorliegt, dann berechnen wir das zweiseitige Konfidenzintervall für den Erwartungswert mit der folgenden Formel, wobei gilt, dass

$$t_{n-1} \left(1 - \frac{\alpha}{2}\right) = t_{11}(0.975) = 2.201$$

und

$$\hat{\sigma}_n = \sqrt{\frac{12}{11} \cdot (1248.06 - 1244.9136)} = 1.8517.$$

Damit ergibt sich

$$\begin{aligned} & \left[\bar{x} - t_{n-1} \left(1 - \frac{\alpha}{2}\right) \cdot \frac{\hat{\sigma}_n}{\sqrt{n}}; \bar{x} + t_{n-1} \left(1 - \frac{\alpha}{2}\right) \cdot \frac{\hat{\sigma}_n}{\sqrt{n}} \right] \\ &= \left[35.2833 - 2.201 \cdot \frac{1.8517}{\sqrt{12}}; \right. \\ & \quad \left. 35.2833 + 2.201 \cdot \frac{1.8517}{\sqrt{12}} \right] \\ &= [34.107; 36.46]. \quad \blacktriangleleft \end{aligned}$$

Aber auch wenn man die Verteilung nicht kennt, kann man ein Konfidenzintervall für μ angeben. Falls σ bekannt ist, nimmt man dasselbe Konfidenzintervall wie bei der Normalverteilung, falls nicht benutzt man die folgende Formel.

Konfidenzintervalle für den Erwartungswert einer beliebigen Verteilung

Wenn die Stichprobe beliebig verteilt ist und die Standardabweichung σ unbekannt ist, muss man zuerst die Standardabweichung bzw. die Varianz schätzen, also $\hat{\sigma}_n^2 = \frac{n}{n-1} \cdot (\overline{x^2} - \bar{x}^2)$ berechnen.

Dann ist das Intervall:

$$\left[\bar{x} - x_{1-\frac{\alpha}{2}} \cdot \frac{\hat{\sigma}_n}{\sqrt{n}}; \bar{x} + x_{1-\frac{\alpha}{2}} \cdot \frac{\hat{\sigma}_n}{\sqrt{n}} \right]$$

ein zweiseitiges $(1 - \alpha)$ -Konfidenzintervall für μ .

Das einseitige untere $(1 - \alpha)$ -Konfidenzintervall für μ lautet

$$\left(-\infty; \bar{x} + x_{1-\alpha} \cdot \frac{\hat{\sigma}_n}{\sqrt{n}} \right].$$

Das einseitige obere $(1 - \alpha)$ -Konfidenzintervall für μ lautet

$$\left[\bar{x} - x_{1-\alpha} \cdot \frac{\hat{\sigma}_n}{\sqrt{n}}; \infty \right).$$

Beispiel

Kommen wir noch einmal auf die zwölf Rapsfelder aus dem obigen Beispiel zurück. Gehen wir nun davon aus, dass wir weder die Verteilung noch die Varianz kennen. Dann können wir ein zweiseitiges 95%-Konfidenzintervall für den unbekanntem Erwartungswert mit der folgenden Formel berechnen:

$$\begin{aligned} & \left[\bar{x} - x_{1-\frac{\alpha}{2}} \cdot \frac{\hat{\sigma}_n}{\sqrt{n}}; \bar{x} + x_{1-\frac{\alpha}{2}} \cdot \frac{\hat{\sigma}_n}{\sqrt{n}} \right] \\ &= \left[35.2833 - 1.96 \cdot \frac{1.8517}{\sqrt{12}}; 35.2833 + 1.96 \cdot \frac{1.8517}{\sqrt{12}} \right] \\ &= [34.236; 36.331]. \end{aligned}$$

Konfidenzintervalle für die Varianz einer Normalverteilung

Wenn die Stichprobe normalverteilt ist, muss man wieder zuerst die Standardabweichung bzw. die Varianz schätzen, also $\hat{\sigma}_n^2 = \frac{n}{n-1} \cdot (\overline{x^2} - (\bar{x})^2)$ berechnen.

Dann ist das Intervall

$$\left[\frac{(n-1) \cdot \hat{\sigma}_n^2}{\chi_{n-1}^2(1-\frac{\alpha}{2})}; \frac{(n-1) \cdot \hat{\sigma}_n^2}{\chi_{n-1}^2(\frac{\alpha}{2})} \right]$$

ein zweiseitiges $(1 - \alpha)$ -Konfidenzintervall für σ^2 .

Das einseitige untere $(1 - \alpha)$ -Konfidenzintervall für σ^2 lautet

$$\left[0; \frac{(n-1) \cdot \hat{\sigma}_n^2}{\chi_{n-1}^2(\alpha)} \right].$$

Das einseitige obere $(1 - \alpha)$ -Konfidenzintervall für σ^2 lautet

$$\left[\frac{(n-1) \cdot \hat{\sigma}_n^2}{\chi_{n-1}^2(1-\alpha)}; \infty \right).$$

Hierbei bezeichnet $\chi_{n-1}^2(1 - \frac{\alpha}{2})$ das $(1 - \frac{\alpha}{2})$ -Quantil der χ^2 -Verteilung, die ebenfalls tabelliert vorliegt. In Tab. A.13 sind allgemein Werte von $\chi_f^2(p)$ tabelliert. Das bedeutet, wir müssen in der ersten Spalte das p suchen, in der ersten Zeile das f , und im Schnittpunkt befindet sich der gesuchte Wert.

Beispiel

Gehen wir von fünf Messwerten und einem Konfidenzniveau $1 - \alpha = 0.95$ aus. Damit ist $\alpha = 0.05$. Dann benötigt man für das zweiseitige Konfidenzintervall für σ^2 $\chi_{5-1}^2(1 - \frac{0.05}{2}) = \chi_4^2(0.975) = 11.143262$.

Die χ^2 -Verteilung geht auf den deutschen Mathematiker *Friedrich Robert Helmert* (1843–1917) und den englischen Mathematiker *Karl Pearson* (1857–1936) zurück.

Wenn Z_1, Z_2, \dots, Z_n eine Folge von Stichprobenvariablen ist, die stochastisch unabhängig und jeweils standardnormalverteilt sind, dann ist die Zufallsvariable $X^2 = Z_1^2 + Z_2^2 + \dots + Z_n^2$ χ^2 -verteilt mit n Freiheitsgraden.

Beispiel

Bislang sind wir bei den Berechnungen für unsere Rapsfelder immer von der Punktschätzung für die unbekannte Varianz $\hat{\sigma}_n^2$ ausgegangen.

Nun können wir ein Konfidenzintervall für die Varianz angeben, z. B. ein zweiseitiges zum Konfidenzniveau 95 %. Dazu benötigen wir die Formel

$$\begin{aligned} & \left[\frac{(n-1) \cdot \hat{\sigma}_n^2}{\chi_{n-1}^2(1-\frac{\alpha}{2})}; \frac{(n-1) \cdot \hat{\sigma}_n^2}{\chi_{n-1}^2(\frac{\alpha}{2})} \right] \\ &= \left[\frac{11 \cdot 3.432436}{21.92}; \frac{11 \cdot 3.432436}{3.816} \right] \\ &= [1.722; 9.894] \end{aligned}$$

Mit einer Wahrscheinlichkeit von 95 % liegt die unbekannte Varianz also zwischen 1.722 und 9.894.

Konfidenzintervalle für die Standardabweichung einer Normalverteilung

Wenn die Stichprobe normalverteilt ist, darf man aus den Grenzen des Intervalls für σ^2 die Wurzel ziehen. Also ist ein zweiseitiges $(1 - \alpha)$ -Konfidenzintervall für σ das Intervall

$$\left[\sqrt{\frac{(n-1) \cdot \hat{\sigma}_n^2}{\chi_{n-1}^2 \left(1 - \frac{\alpha}{2}\right)}}, \sqrt{\frac{(n-1) \cdot \hat{\sigma}_n^2}{\chi_{n-1}^2 \left(\frac{\alpha}{2}\right)}} \right]$$

Das einseitige untere $(1 - \alpha)$ -Konfidenzintervall für σ lautet:

$$\left[0; \sqrt{\frac{(n-1) \cdot \hat{\sigma}_n^2}{\chi_{n-1}^2(\alpha)}} \right].$$

Das einseitige obere $(1 - \alpha)$ -Konfidenzintervall für σ lautet

$$\left[\sqrt{\frac{(n-1) \cdot \hat{\sigma}_n^2}{\chi_{n-1}^2(1 - \alpha)}}, \infty \right).$$

Beispiel

Aufbauend auf dem letzten Beispiel ist es nicht besonders schwierig, ein 95 %-Konfidenzintervall für die unbekannte Standardabweichung zu berechnen: Wir ziehen aus den oben ermittelten Intervallgrenzen einfach die Wurzel und erhalten

$$\begin{aligned} & \left[\sqrt{\frac{(n-1)\hat{\sigma}_n^2}{\chi_{n-1}^2 \left(1 - \frac{\alpha}{2}\right)}}, \sqrt{\frac{(n-1)\hat{\sigma}_n^2}{\chi_{n-1}^2 \left(\frac{\alpha}{2}\right)}} \right] \\ &= \left[\sqrt{1.722}; \sqrt{9.894} \right] \\ &= [1.312; 3.145]. \end{aligned}$$

18.2 Hypothesentests

Bei den Hypothesentests ist die Fragestellung etwas anders als bei den Schätzungen, denn wir betreiben zwar weiterhin induktive Statistik, schließen also von einer Stichprobe auf die Allgemeinheit zurück, aber statt einen unbekannt Parameter zu schätzen, haben wir schon zumindest eine Vermutung (also eine Hypothese) über einen der unbekannt Parameter. In diesem Fall spricht man auch von Parametertests.

Wenn die Vermutung nicht einen Parameter der Verteilung betrifft, sondern z. B. die gesamte Verteilung, oder wenn er Unabhängigkeit prüft, spricht man von nicht parametrischen bzw. parameterfreien Tests. Diese behandeln wir am Schluss.

Hypothesentests besitzen einen festgelegten Aufbau

Tests sind immer so aufgebaut, dass man eine Aussage als Hypothese H_0 (Nullhypothese) aufstellt; das Gegenteil dieser Aussage wird ebenfalls (als Alternativhypothese A bzw. H_1) notiert. Hierbei wird darauf geachtet, dass die Meinung desjenigen, der den Test durchführt, als Alternativhypothese formuliert wird. Das klingt merkwürdig, hat aber sehr gute Gründe. Zuerst einmal existiert nach wie vor das Problem, dass aus einer Stichprobe Rückschlüsse auf die Allgemeinheit gezogen werden. Das bedeutet, dass Tests streng wissenschaftlich gesehen nichts beweisen können, sondern dass man nur folgendermaßen argumentieren kann: Wenn eine Nullhypothese (aus irgendwelchen Gründen) verworfen wird, so wird die Alternativhypothese statistisch untermauert. Daher arbeiten alle Tests so, dass man versucht, die Nullhypothese zu verwerfen und so die eigene Vermutung, die ja in der Alternativhypothese steht, zu stützen.

Hinzu kommt noch ein wissenschaftstheoretischer Grund. Ein Prinzip der Forschung lautet: Wenn es mehrere Möglichkeiten der Wahrheit gibt, soll man alle unmöglichen ausschließen, und die einzig mögliche Wahrheit bleibt übrig. Auch aus diesem Grund schreibt man die eigene Meinung in die Alternative und versucht, das Gegenteil, also die Nullhypothese, zu verwerfen, um so die eigene Meinung zu unterstützen.

Parametertests gibt es aus diesem Grund übrigens immer in drei Variationen, sogenannten Fragestellungen. Auch hierbei bezeichnet ϑ wieder den unbekannt Parameter. ϑ_0 steht für einen Zahlenwert, also den vermuteten Wert. Man sagt auch, dass man ϑ gegen den Wert ϑ_0 testet.

Fragestellungen bei Parametertests

- **Erste einseitige (rechtsseitige) Fragestellung:**
 $H_0 : \vartheta \leq \vartheta_0 \quad A : \vartheta > \vartheta_0$
- **Zweite einseitige (linksseitige) Fragestellung:**
 $H_0 : \vartheta \geq \vartheta_0 \quad A : \vartheta < \vartheta_0$
- **Zweiseitige Fragestellung:**
 $H_0 : \vartheta = \vartheta_0 \quad A : \vartheta \neq \vartheta_0$

Anschließend werden bei allen Tests bestimmte Testgrößen (Teststatistiken) berechnet, und je nach Ergebnis wird die Hypothese verworfen oder nicht. Dieses Vorgehen nennt man Testvorschrift.

Aber kommen wir noch einmal auf die Problematik des induktiven Vorgehens zurück. Auch wenn wir die Testvorschrift 100 %ig richtig befolgt haben, kann es passieren, dass der Test ein falsches Ergebnis liefert. Genauer gesagt können uns zwei Fehler unterlaufen.

Fehler bei Hypothesentests

Einerseits kann der Test als Ergebnis liefern, dass die Nullhypothese H_0 abgelehnt werden soll, obwohl die Hy-

pothese richtig ist. Dies nennen wir **Fehler erster Art**. Oder die Testvorschrift sagt, dass die Nullhypothese H_0 nicht abgelehnt werden kann, obwohl sie in Wirklichkeit falsch ist. Dann unterläuft uns gerade der **Fehler zweiter Art**. Der Grund liegt darin, dass man einfach Pech mit der Stichprobe haben kann.

Betrachten wir zwei Beispiele, um zu verstehen, was gemeint ist:

Beispiel

- Ein Qualitätsprüfer kann zwei Arten von Fehlern machen: Entweder er sortiert funktionierende Teile fälschlicherweise aus, oder er lässt mangelhafte Teile fälschlicherweise durch die Kontrolle.
- Stellen wir uns vor, wir arbeiten für einen Automobilzulieferer für Bremsanlagen. Eines Morgens finden wir auf unserem Schreibtisch einen Erpresserbrief, in dem jemand behauptet, er hätte Bremsleitungen durchtrennt. Dann können uns ebenfalls zwei Fehler unterlaufen: Entweder wir tauschen alle Leitungen aus, obwohl sie gar nicht defekt sind, oder wir verkaufen die Bremsanlagen weiter, und sie sind defekt. ◀

Diese Beispiele sollen verdeutlichen, dass normalerweise einer der beiden (tatsächlich nicht ganz auszuschließenden) Fehler schwerwiegender ist als der andere.

Unmittelbar einleuchtend ist, dass man daran interessiert ist, die Wahrscheinlichkeit für das Eintreten des schwerwiegenden Fehlers möglichst gering zu halten. Genau so klar ist, dass wenn wir eine Fehlerwahrscheinlichkeit kleinhalten, die andere Fehlerwahrscheinlichkeit automatisch größer wird. (Allerdings addieren sie sich nicht zu 1, das wäre dramatisch!)

Definition: Fehlerwahrscheinlichkeiten und Signifikanzniveau

Die Wahrscheinlichkeit des Eintretens des Fehlers 1. Art, also des Fehlers, H_0 abzulehnen, obwohl H_0 richtig ist, heißt **Fehlerwahrscheinlichkeit erster Art** bzw. α . Sie wird auch als **Signifikanzniveau** eines Tests bezeichnet.

Die Wahrscheinlichkeit für das Eintreten des Fehlers zweiter Art, also des Fehlers, H_0 nicht abzulehnen, obwohl H_0 falsch ist, heißt **Fehlerwahrscheinlichkeit zweiter Art** bzw. β .

Mathematisch sind Tests so aufgebaut, dass man α beschränken kann bzw. einfach vorgibt, z. B. $\alpha = 1\%$ etc. Auch hierin liegt wieder eine Begründung, warum die eigene Vermutung in die Alternative gehört: Wenn α beschränkt werden kann, bedeutet das, dass die Wahrscheinlichkeit, dass man die Nullhypothese H_0 ablehnt, obwohl sie richtig ist, beschränkt werden kann.

Wenn dort nun das drin steht, was man eigentlich nicht vermutet, arbeitet man besonders wissenschaftlich, da man die Fehlerwahrscheinlichkeit, dass man sich „in die Tasche lügt“, gering hält.

Einstichprobentests überprüfen Vermutungen, die einen Parameter einer Verteilung betreffen

In diesem Abschnitt geht es um Hypothesentests, die sich mit einem Parameter einer Verteilung beschäftigen, also z. B. um μ oder um σ^2 . Man unterscheidet hierbei Tests, die mit einer Stichprobe arbeiten, und Tests, die mehrere Stichproben miteinander vergleichen. (Da dies eine Grundlagengeschichte ist, betrachten wir nur zwei Stichproben. Etwas umfangreicher wird es aber natürlich, wenn man drei, vier oder mehr Stichproben miteinander vergleichen will.)

Beginnen wir mit den Einstichprobentests. Die folgenden Parametertests arbeiten alle mit einer Stichprobe.

Gauß-Test

Der **Gauß-Test** testet Hypothesen bzgl. des Erwartungswertes bei Normalverteilung. Die Varianz σ^2 muss dabei bekannt sein.

Hypothesen und Alternativen lauten dann:

- Erste einseitige (rechtsseitige) Fragestellung:
 $H_0 : \mu \leq \mu_0 \quad A : \mu > \mu_0$
- Zweite einseitige (linksseitige) Fragestellung:
 $H_0 : \mu \geq \mu_0 \quad A : \mu < \mu_0$
- Zweiseitige Fragestellung:
 $H_0 : \mu = \mu_0 \quad A : \mu \neq \mu_0$

Die Testvorschrift lautet in den jeweiligen Fragestellungen:

Lehne H_0 ab, falls

- $\bar{x} > \mu_0 + x_{1-\alpha} \cdot \frac{\sigma}{\sqrt{n}}$
- $\bar{x} < \mu_0 - x_{1-\alpha} \cdot \frac{\sigma}{\sqrt{n}}$
- $\bar{x} > \mu_0 + x_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$
oder
 $\bar{x} < \mu_0 - x_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$.

Im zweiseitigen Fall muss also zuerst die eine Ungleichung überprüft werden. Wenn diese nicht erfüllt ist, muss noch die zweite Ungleichung überprüft werden. Nur wenn beide Ungleichungen nicht erfüllt sind, kann die Hypothese nicht abgelehnt werden.

Beim Gauß-Test ist es relativ einfach, die Fehlerwahrscheinlichkeit zweiter Art zu berechnen. Dazu benutzt man den zentralen Grenzwertsatz und erhält Folgendes:

Anwendung: Gauß-Test

Zwischen dem Zulieferer für Schaumstoff und einem LKW-Hersteller gibt es einen Liefervertrag, in dem geregelt ist, dass der LKW-Hersteller einen Preisnachlass fordern darf, wenn nachgewiesen werden kann, dass bei einer Lieferung die Schaumstoffpakete weniger als die vereinbarten (und auf die Pakete aufgedruckten) 100 kg wiegen. Da eine Vollerhebung zu teuer und zu aufwendig ist, haben die Parteien vereinbart, aus jeder Lieferung eine Stichprobe vom Umfang 10 zu ziehen, und sie konnten sich auf ein Signifikanzniveau von 5 % einigen.

Beide Parteien gehen davon aus, dass das Gewicht der Schaumstoffpakete normalverteilt ist, wobei generell eine Varianz von $\sigma^2 = 4$ vorliegt. Bei der letzten Lieferung ergab die Stichprobe folgende Werte: 101 kg, 98 kg, 96 kg, 102 kg, 99 kg, 101 kg, 102 kg, 103 kg, 96 kg, 97 kg.

Der Geschäftsführer des LKW-Herstellers will nun wissen, ob er einen Preisnachlass fordern darf. Hierzu verwendet er die zweite einseitige Fragestellung mit $H_0 : \mu \geq 100$ gegen $A : \mu < 100$, denn er möchte ja gerne belegen, dass er weniger als 100 kg erhält, und das muss wie bereits besprochen in die Alternative.

Seine zu befolgende Testvorschrift lautet also:

$$\text{Lehne } H_0 \text{ ab, falls } \bar{x} < \mu_0 - x_{1-\alpha} \cdot \frac{\sigma}{\sqrt{n}}.$$

Suchen wir zuerst alle Werte, die wir benötigen:

$\mu_0 = 100$, denn das ist immer die Zahl, die in Nullhypothese und Alternative auftaucht. $\sigma = 2$, denn die Varianz war ja mit 4 angegeben. $n = 10$, das ist der Stichprobenumfang. $\bar{x} = 99.5$. Bleibt noch das Quantil der Standardnormalverteilung: $x_{1-\alpha} = x_{0.95} = 1.645$.

Insgesamt gilt es also zu prüfen, ob

$$99.5 < 100 - 1.645 \cdot \frac{2}{\sqrt{10}} = 98.9596.$$

Da das nicht stimmt, kann der Geschäftsführer die Nullhypothese auf dem 5 %-Signifikanzniveau nicht ablehnen. Er kann also seine eigene Meinung nicht untermauern, denn diese steht in der Alternative, und somit kann er keinen Preisnachlass fordern.

Nun nehmen wir die Perspektive des Zulieferers ein. Dieser möchte nachweisen, dass er nicht nur 100 kg Schaumstoff verpackt hat, sondern sogar mehr als das. Also verwendet er die erste einseitige Fragestellung mit $H_0 : \mu \leq 100$ gegen $A : \mu > 100$ und der Testvorschrift:

$$\text{Lehne } H_0 \text{ ab, falls } \bar{x} > \mu_0 + x_{1-\alpha} \cdot \frac{\sigma}{\sqrt{n}}.$$

Wir müssen also prüfen, ob

$$99.5 > 100 + 1.645 \cdot \frac{2}{\sqrt{10}} = 101.0404.$$

Da dies ebenfalls nicht stimmt, kann der Zulieferer seine Behauptung auch nicht belegen, jedenfalls nicht bei einer Fehlerwahrscheinlichkeit von 5 %. Ehrlich gesagt wird er seine Behauptung auf keinem Signifikanzniveau belegen können, denn egal wie klein der zweite Term wird, er bleibt auf jeden Fall positiv, und daher wird die rechte Seite immer > 100 sein.

Als letzter Akteur mischt sich nun der Hersteller der Anlage ein, auf der die Schaumstoffpakete gepackt werden. Er will überprüfen, ob seine Anlage noch korrekt arbeitet. Daher verwendet er die zweiseitige Fragestellung: $H_0 : \mu = 100$ gegen $A : \mu \neq 100$ mit der Testvorschrift:

$$\text{Lehne } H_0 \text{ ab, falls } \bar{x} > \mu_0 + x_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}} \text{ oder } \bar{x} < \mu_0 - x_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}.$$

Hier benötigen wir ein anderes Quantil der Standardnormalverteilung, nämlich $x_{1-\frac{\alpha}{2}} = x_{0.975} = 1.960$. Beginnen wir mit der ersten Bedingung und prüfen, ob

$$99.5 > 100 + 1.96 + \frac{2}{\sqrt{10}} = 101.2396.$$

Das stimmt natürlich nicht, also müssen wir die zweite Bedingung prüfen:

$$99.5 < 100 - 1.96 - \frac{2}{\sqrt{10}} = 98.76.$$

Auch das ist nicht erfüllt. Also kann auch der Anlagenbauer die Nullhypothese nicht ablehnen.

Aber im Gegensatz zu den beiden vorherigen Fragestellungen steht in der Nullhypothese das, was der Anlagenbauer zeigen möchte, nämlich dass die Anlage Pakete mit einem Gewicht von 100 kg packt. Er kann also davon ausgehen, dass seine Anlage exakt arbeitet. Die zu dieser Vermutung gehörende Fehlerwahrscheinlichkeit ist aber β , nämlich die Wahrscheinlichkeit, H_0 nicht abzulehnen, obwohl die Nullhypothese falsch ist.

Beim Gauß-Test beträgt die Fehlerwahrscheinlichkeit zweiter Art bei der ersten einseitigen Fragestellung

$$\beta = \Phi \left((\mu_0 - \mu) \cdot \frac{\sqrt{n}}{\sigma} + x_{1-\alpha} \right).$$

Bei der zweiten einseitigen Fragestellung beträgt sie

$$\beta = 1 - \Phi \left((\mu_0 - \mu) \cdot \frac{\sqrt{n}}{\sigma} - x_{1-\alpha} \right).$$

Auf die zweiseitige Fragestellung verzichten wir an dieser Stelle. Es sollte uns klar sein, dass wir zur Berechnung natürlich den wirklichen und wahren Erwartungswert μ und die Fehlerwahrscheinlichkeit erster Art α kennen muss. Sonst geht es nicht!

Beispiel

Der Geschäftsführer des LKW-Herstellers möchte nun gerne wissen, mit welcher Wahrscheinlichkeit er irrtümlich auf einen Preisnachlass verzichtet, falls das wirklich wahre Gewicht der Schaumstoffpakete μ genauso groß ist wie der Mittelwert der Stichprobe, also 99.5 kg. Er interessiert sich also für die Fehlerwahrscheinlichkeit zweiter Art, d. h. β .

Da er die zweite einseitige Fragestellung verwendet hat, gilt

$$\begin{aligned} \beta &= 1 - \Phi \left((\mu_0 - \mu) \cdot \frac{\sqrt{n}}{\sigma} - x_{1-\alpha} \right) \\ &= 1 - \Phi \left((100 - 99.5) \cdot \frac{\sqrt{10}}{2} - 1.645 \right) \\ &= 1 - \Phi(-0.85) = \Phi(0.85) = 0.8023. \end{aligned}$$

Die Wahrscheinlichkeit beträgt also 80.23 %.

Der Schaumstoffzulieferer wiederum möchte nun wissen, mit welcher Wahrscheinlichkeit der Test als Ergebnis liefert, dass die Schaumstoffpakete höchstens 100 kg schwer sind, wenn das zu erwartende Gewicht der Pakete μ in Wirklichkeit 101 kg beträgt. Auch er muss also die Fehlerwahrscheinlichkeit zweiter Art, also β , berechnen, aber für die erste Fragestellung. Daher gilt

$$\begin{aligned} \beta &= \Phi \left((\mu_0 - \mu) \cdot \frac{\sqrt{n}}{\sigma} + x_{1-\alpha} \right) \\ &= \Phi \left((100 - 101) \cdot \frac{\sqrt{10}}{2} + 1.645 \right) \\ &= \Phi(0.06) = 0.5239. \end{aligned}$$

Der Test ist also in diesem Sinne recht ungenau, da er in mehr als der Hälfte der Fälle falsch entscheidet. ◀

Genauso kann man auch einen Mindeststichprobenumfang angeben, der bei gegebenem α und β verwendet werden muss. Dazu muss man die obigen Formeln nach n umstellen:

Beim Gauß-Test beträgt der Mindeststichprobenumfang bei gegebenem α und β bei der ersten einseitigen Fragestellung

$$n \geq \left(\frac{x_{1-\alpha} + x_{1-\beta}}{\mu_0 - \mu} \cdot \sigma \right)^2.$$

Bei der zweiten einseitigen Fragestellung beträgt er

$$n \geq \left(\frac{x_{\beta} - x_{1-\alpha}}{\mu_0 - \mu} \cdot \sigma \right)^2.$$

Das sehen wir uns an einem Beispiel an:

Beispiel

Der Geschäftsführer des LKW-Herstellers möchte nun beide Fehlerwahrscheinlichkeiten begrenzen: α soll nach wie vor höchstens 5 % betragen, β aber höchstens 10 %. Daher berechnet er einen Stichprobenumfang von

$$\begin{aligned} n &\geq \left(\frac{x_{1-\alpha} + x_{1-\beta}}{\mu_0 - \mu} \cdot \sigma \right)^2 \\ &= \left(\frac{x_{0.95} + x_{0.9}}{100 - 99.5} \cdot 2 \right)^2 \\ &= \left(\frac{1.645 + 1.282}{0.5} \cdot 2 \right)^2 = 137.077. \end{aligned}$$

Es müssen also mindestens 138 Schaumstoffpakete gewogen werden.

Der Schaumstoffzulieferer muss für seine Frage

$$\begin{aligned} n &\geq \left(\frac{x_{\beta} - x_{1-\alpha}}{\mu_0 - \mu} \cdot \sigma \right)^2 \\ &= \left(\frac{x_{0.1} - x_{0.95}}{100 - 101} \cdot 2 \right)^2 \\ &= \left(\frac{-1.282 - 1.645}{100 - 101} \cdot 2 \right)^2 = 34.269, \end{aligned}$$

also 35 Pakete wiegen. ◀

Der bekannteste Test für Erwartungswerte ist sicherlich der t-Test.

t-Test

Der **t-Test** testet Hypothesen bzgl. des Erwartungswertes bei Normalverteilung. Die Varianz σ^2 ist dabei unbekannt. Daher muss zuerst die Varianz geschätzt werden, indem $\hat{\sigma}_n^2 = \frac{n}{n-1} \cdot (\overline{x^2} - \bar{x}^2)$ berechnet wird.

Hypothesen und Alternativen lauten dann:

- $H_0 : \mu \leq \mu_0 \quad A : \mu > \mu_0$
- $H_0 : \mu \geq \mu_0 \quad A : \mu < \mu_0$
- $H_0 : \mu = \mu_0 \quad A : \mu \neq \mu_0$

Die Testvorschrift lautet in den jeweiligen Fällen:

Lehne H_0 ab, falls

- $\bar{x} > \mu_0 + t_{n-1}(1-\alpha) \cdot \frac{\hat{\sigma}_n}{\sqrt{n}}$
- $\bar{x} < \mu_0 - t_{n-1}(1-\alpha) \cdot \frac{\hat{\sigma}_n}{\sqrt{n}}$
- $\bar{x} > \mu_0 + t_{n-1}\left(1-\frac{\alpha}{2}\right) \cdot \frac{\hat{\sigma}_n}{\sqrt{n}}$
oder
 $\bar{x} < \mu_0 - t_{n-1}\left(1-\frac{\alpha}{2}\right) \cdot \frac{\hat{\sigma}_n}{\sqrt{n}}$.

Im zweiseitigen Fall muss also zuerst die eine Ungleichung überprüft werden. Wenn diese nicht erfüllt ist, muss noch die zweite Ungleichung überprüft werden. Nur wenn beide Ungleichungen nicht erfüllt sind, kann die Hypothese nicht abgelehnt werden.

Approximativer Gauß-Test

Der **approximative Gauß-Test** testet Hypothesen bzgl. des Erwartungswerts bei beliebiger Verteilung. Der Stichprobenumfang muss mindestens $n \geq 30$ betragen. Die Varianz σ^2 kann dabei bekannt sein, muss aber nicht.

Hypothesen und Alternativen lauten dann:

- $H_0 : \mu \leq \mu_0 \quad A : \mu > \mu_0$
- $H_0 : \mu \geq \mu_0 \quad A : \mu < \mu_0$
- $H_0 : \mu = \mu_0 \quad A : \mu \neq \mu_0$

Die Testvorschrift lautet in den jeweiligen Fällen, wenn σ^2 bekannt ist:

Lehne H_0 ab, falls

- $\bar{x} > \mu_0 + x_{1-\alpha} \cdot \frac{\sigma}{\sqrt{n}}$
- $\bar{x} < \mu_0 - x_{1-\alpha} \cdot \frac{\sigma}{\sqrt{n}}$
- $\bar{x} > \mu_0 + x_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$ oder $\bar{x} < \mu_0 - x_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$.

Im dritten Fall muss also zuerst die eine Ungleichung überprüft werden. Wenn diese nicht erfüllt ist, muss noch die zweite Ungleichung überprüft werden. Nur wenn beide Ungleichungen nicht erfüllt sind, kann die Hypothese nicht abgelehnt werden.

Wenn σ^2 nicht bekannt ist, muss in der Teststatistik σ durch $\hat{\sigma}_n$ ersetzt werden (wobei $\hat{\sigma}_n^2 = \frac{n}{n-1} \cdot (\overline{x^2} - \bar{x}^2)$ ist).

Hypothesen und Alternativen lauten dann:

- $H_0 : \mu \leq \mu_0 \quad A : \mu > \mu_0$
- $H_0 : \mu \geq \mu_0 \quad A : \mu < \mu_0$
- $H_0 : \mu = \mu_0 \quad A : \mu \neq \mu_0$

Die Testvorschrift lautet in den jeweiligen Fällen, wenn σ^2 unbekannt ist:

Lehne H_0 ab, falls

- $\bar{x} > \mu_0 + x_{1-\alpha} \cdot \frac{\hat{\sigma}_n}{\sqrt{n}}$
- $\bar{x} < \mu_0 - x_{1-\alpha} \cdot \frac{\hat{\sigma}_n}{\sqrt{n}}$
- $\bar{x} > \mu_0 + x_{1-\frac{\alpha}{2}} \cdot \frac{\hat{\sigma}_n}{\sqrt{n}}$ oder $\bar{x} < \mu_0 - x_{1-\frac{\alpha}{2}} \cdot \frac{\hat{\sigma}_n}{\sqrt{n}}$.

Im dritten Fall muss also zuerst die eine Ungleichung überprüft werden. Wenn diese nicht erfüllt ist, muss noch die zweite Ungleichung überprüft werden. Nur wenn beide Ungleichungen nicht erfüllt sind, kann die Hypothese nicht abgelehnt werden.

Damit wird klar, dass bei bekannter Varianz der Gauß-Test und der approximative Gauß-Test identisch sind. Wir müssen also nur den Fall betrachten, dass die Varianz unbekannt ist und dass sich nun doch herausgestellt hat, dass die Gewichte der Schaumstoffpakete nicht normalverteilt sind.

 χ^2 -Test für die Varianz

Der **χ^2 -Test für die Varianz** testet Hypothesen bzgl. der Varianz bei Normalverteilung.

Hypothesen und Alternativen lauten dann:

- $H_0 : \sigma^2 \leq \sigma_0^2 \quad A : \sigma^2 > \sigma_0^2$
- $H_0 : \sigma^2 \geq \sigma_0^2 \quad A : \sigma^2 < \sigma_0^2$
- $H_0 : \sigma^2 = \sigma_0^2 \quad A : \sigma^2 \neq \sigma_0^2$

Die Testvorschrift lautet in den jeweiligen Fällen:

Lehne H_0 ab, falls

- $\frac{(n-1)\hat{\sigma}_n^2}{\sigma_0^2} > \chi_{n-1}^2(1-\alpha)$
- $\frac{(n-1)\hat{\sigma}_n^2}{\sigma_0^2} < \chi_{n-1}^2(\alpha)$
- $\frac{(n-1)\hat{\sigma}_n^2}{\sigma_0^2} < \chi_{n-1}^2\left(\frac{\alpha}{2}\right)$ oder $\frac{(n-1)\hat{\sigma}_n^2}{\sigma_0^2} > \chi_{n-1}^2\left(1-\frac{\alpha}{2}\right)$.

Im dritten Fall muss also zuerst die eine Ungleichung überprüft werden. Wenn diese nicht erfüllt ist, muss noch die zweite Ungleichung überprüft werden. Nur wenn beide Ungleichungen nicht erfüllt sind, kann die Hypothese nicht abgelehnt werden.

Beispiel

Nun beschäftigen wir uns noch mit der Frage, ob die Varianz, die bei den beiden Geschäftspartnern im ersten Beispiel (also beim Gauß-Test) vom Himmel gefallen

Anwendung: t-Test

Wir betrachten noch einmal unsere beiden Geschäftspartner aus den vorherigen Beispielen. Aber diesmal gehen wir davon aus, dass (realistischerweise) die Varianz der Gewichte der Schaumstoffpakete unbekannt ist. Ohne es überprüft zu haben, unterstellen wir weiterhin, dass das Gewicht der Pakete normalverteilt ist.

Wir müssen also zuerst die unbekannte Varianz schätzen: $\hat{\sigma}_n^2 = \frac{n}{n-1} (\overline{x^2} - \bar{x}^2) = \frac{10}{9} \cdot (9906.5 - 99.5^2) = 6.944$, also ist $\hat{\sigma}_n = 2.635$.

Der Geschäftsführer des LKW-Herstellers will nun wissen, ob er einen Preisnachlass fordern darf. Hierzu verwendet er die zweite einseitige Fragestellung mit $H_0 : \mu \geq 100$ gegen $A : \mu < 100$, denn er möchte ja gerne belegen, dass er im Durchschnitt weniger als 100 kg erhält, und das muss wie bereits besprochen in die Alternative.

Seine zu befolgende Testvorschrift lautet also:

Lehne H_0 ab, falls $\bar{x} < \mu_0 - t_{n-1}(1 - \alpha) \cdot \frac{\hat{\sigma}_n}{\sqrt{n}}$.

Suchen wir zuerst alle Werte, die wir benötigen: $\mu_0 = 100$, denn das ist immer noch die Zahl, die in Nullhypothese und Alternative auftaucht. $\hat{\sigma}_n = 2.635$. $n = 10$, das ist der Stichprobenumfang. $\bar{x} = 99.5$. Bleibt noch das Quantil der t-Verteilung: $t_{n-1}(1 - \alpha) = t_9(0.95) = 1.833$.

Insgesamt gilt es also zu prüfen, ob

$$99.5 < 100 - 1.833 \cdot \frac{2.635}{\sqrt{10}} = 98.4725.$$

Da das nicht stimmt, kann der Geschäftsführer die Nullhypothese auf dem 5 %-Signifikanzniveau nicht ablehnen. Er kann also seine eigene Meinung nicht untermauern, denn diese steht in der Alternative, und somit kann er keinen Preisnachlass fordern.

Nun nehmen wir die Perspektive des Zulieferers ein. Dieser möchte nachweisen, dass er nicht nur 100 kg Schaumstoff in

die Pakete gepackt hat, sondern sogar mehr als das. Also verwendet er die erste einseitige Fragestellung mit $H_0 : \mu \leq 100$ gegen $A : \mu > 100$ und der Testvorschrift:

Lehne H_0 ab, falls $\bar{x} > \mu_0 + t_{n-1}(1 - \alpha) \cdot \frac{\hat{\sigma}_n}{\sqrt{n}}$.

Wir müssen also prüfen, ob $99.5 > 100 + 1.833 \cdot \frac{2.635}{\sqrt{10}} = 101.5275$. Da dies ebenfalls nicht stimmt, kann der Zulieferer seine Behauptung auch nicht belegen, jedenfalls nicht bei einer Fehlerwahrscheinlichkeit von 5 %.

Als letzter Akteur mischt sich nun der Hersteller der Verpackungsstraße ein, auf der die Schaumstoffpakete gepackt wurden. Er will überprüfen, ob seine Maschine noch korrekt arbeitet. Daher verwendet er die zweiseitige Fragestellung:

$H_0 : \mu = 100$ gegen $A : \mu \neq 100$ mit der Testvorschrift:

Lehne H_0 ab, falls $\bar{x} > \mu_0 + t_{n-1}(1 - \frac{\alpha}{2}) \cdot \frac{\hat{\sigma}_n}{\sqrt{n}}$ oder

$\bar{x} < \mu_0 - t_{n-1}(1 - \frac{\alpha}{2}) \cdot \frac{\hat{\sigma}_n}{\sqrt{n}}$.

Hier benötigen wir ein anderes Quantil der t-Verteilung, nämlich $t_{n-1}(1 - \frac{\alpha}{2}) = t_9(0.975) = 2.262$.

Beginnen wir mit der ersten Bedingung und prüfen, ob $99.5 > 100 + 2.262 \cdot \frac{2.635}{\sqrt{10}} = 101.8848$. Das stimmt natürlich nicht, also müssen wir die zweite Bedingung prüfen:

$99.5 < 100 - 2.262 \cdot \frac{2.635}{\sqrt{10}} = 98.12$. Auch das ist nicht erfüllt.

Also kann auch der Anlagenbauer die Nullhypothese nicht ablehnen. Aber im Gegensatz zu den beiden vorherigen Fragestellungen steht in der Nullhypothese das, was der Anlagenbauer zeigen möchte, nämlich dass die Anlage Pakete mit einem Gewicht von 100 kg packt. Er kann also belegen, dass seine Fertigungsstraße exakt arbeitet. Die zugehörige Fehlerwahrscheinlichkeit ist aber β , nämlich die Wahrscheinlichkeit, H_0 nicht abzulehnen, obwohl die Nullhypothese falsch ist.

war, wirklich $\sigma^2 = 4$ ist, oder ob sie nicht eher größer oder kleiner als 4 ist. Wir testen also die Varianz gegen die Zahl $\sigma_0^2 = 4$.

Wir müssen wieder zuerst die unbekannte Varianz schätzen:

$$\hat{\sigma}_n^2 = \frac{n}{n-1} (\overline{x^2} - \bar{x}^2) = \frac{10}{9} (9906.5 - 99.5^2) = 6.944.$$

Da man bei diesem Test bei jeder Fragestellung auf der linken Seite der Ungleichungen immer denselben Ausdruck berechnen muss, nämlich

$$\frac{(n-1)\hat{\sigma}_n^2}{\sigma_0^2} = \frac{9 \cdot 6.944}{4} = 15.625,$$

betrachten wir diesmal nur die zweiseitige Fragestellung

$$H_0 : \sigma^2 = 4 \quad A : \sigma^2 \neq 4.$$

Wir brauchen zum Vergleich mit unserer Testgröße zwei Quantile der χ^2 -Verteilung:

$$\chi_{n-1}^2 \left(\frac{\alpha}{2} \right) = \chi_9^2(0.025) = 2.7$$

und

$$\chi_{n-1}^2 \left(1 - \frac{\alpha}{2} \right) = \chi_9^2(0.975) = 19.023.$$

Da weder $15.625 < 2.7$ noch $15.625 > 19.023$ stimmt, kann H_0 zum Niveau 5 % nicht abgelehnt werden. Die Schätzung, dass $\sigma^2 = 4$ ist, ist also in Ordnung. ◀

Anwendung: Approximativer Gauß-Test

Wir müssen wieder zuerst die unbekannte Varianz schätzen: $\hat{\sigma}_n^2 = \frac{n}{n-1}(\bar{x}^2 - \bar{x}^2) = 10/9(9906.5 - 99.5^2) = 6.944$, also ist $\hat{\sigma}_n = 2.635$.

Der Geschäftsführer des LKW-Herstellers will immer noch wissen, ob er einen Preisnachlass fordern darf. Hierzu verwendet er die zweite einseitige Fragestellung mit $H_0 : \mu \geq 100$ gegen $A : \mu < 100$, denn er möchte ja gerne belegen, dass er weniger als 100 kg erhält, und das muss wie bereits besprochen in die Alternative. Seine zu befolgende Testvorschrift lautet also:

Lehne H_0 ab, falls $\bar{x} < \mu_0 - x_{1-\alpha} \cdot \frac{\hat{\sigma}_n}{\sqrt{n}}$. Suchen wir zuerst alle Werte, die wir benötigen.

$\mu_0 = 100$, denn das ist immer die Zahl, die in Nullhypothese und Alternative auftaucht. $\sigma_n = 2.635$, $n = 10$. $\bar{x} = 99.5$. Bleibt noch das Quantil der Standardnormalverteilung: $x_{1-\alpha} = x_{0.95} = 1.645$.

Insgesamt gilt es also zu prüfen, ob

$$99.5 < 100 - 1.645 \cdot \frac{2.635}{\sqrt{10}} = 98.6293.$$

Da das nicht stimmt, kann der Geschäftsführer die Nullhypothese auf dem 5 %-Signifikanzniveau wieder nicht ablehnen. Er kann also seine eigene Meinung immer noch nicht untermauern, denn diese steht in der Alternative, und somit kann er keinen Preisnachlass fordern.

Nun nehmen wir die Perspektive des Herstellers ein. Dieser möchte immer noch nachweisen, dass er nicht nur 100 kg in die Tafeln packt, sondern sogar mehr als das. Also verwendet er wieder die erste einseitige Fragestellung mit $H_0 : \mu \leq 100$ gegen $A : \mu > 100$ und der Testvorschrift:

Lehne H_0 ab, falls $\bar{x} > \mu_0 + x_{1-\alpha} \cdot \frac{\hat{\sigma}_n}{\sqrt{n}}$.

Wir müssen also prüfen, ob

$$99.5100 + 1.645 \cdot \frac{2.635}{\sqrt{10}} = 101.3707.$$

Da dies ebenfalls nicht stimmt, kann der Schaumstofffabrikant seine Behauptung auch noch immer nicht belegen, jedenfalls nicht bei einer Fehlerwahrscheinlichkeit von 5 %.

Als letzter Akteur mischt sich nun wieder der Hersteller der Maschine ein, auf der die Schaumstoffpakete gepackt werden. Er will überprüfen, ob seine Maschine noch korrekt befüllt. Daher verwendet er die zweiseitige Fragestellung mit $H_0 : \mu = 100$ gegen $A : \mu \neq 100$ und der Testvorschrift:

Lehne H_0 ab, falls $\bar{x} > \mu_0 + x_{1-\frac{\alpha}{2}} \cdot \frac{\hat{\sigma}_n}{\sqrt{n}}$

oder

$$\bar{x} < \mu_0 - x_{1-\frac{\alpha}{2}} \cdot \frac{\hat{\sigma}_n}{\sqrt{n}}.$$

Hier benötigen wir ein anderes Quantil der Standardnormalverteilung, nämlich $x_{1-\frac{\alpha}{2}} = x_{0.975} = 1.960$.

Beginnen wir mit der ersten Bedingung und prüfen, ob

$$99.5 > 100 + 1.96 \cdot \frac{2.635}{\sqrt{10}} = 101.633.$$

Das stimmt natürlich nicht, also müssen wir die zweite Bedingung prüfen:

$$99.5 < 100 - 1.96 \cdot \frac{2.635}{\sqrt{10}} = 98.37.$$

Auch das ist nicht erfüllt. Also kann auch der Maschinenhersteller die Nullhypothese nicht ablehnen. Aber im Gegensatz zu den beiden vorherigen Fragestellungen steht in der Nullhypothese das, was der Maschinenhersteller zeigen möchte, nämlich dass die Maschine Pakete mit einem Gewicht von 100 kg packt. Er kann also nachweisen, dass seine Maschine exakt arbeitet. Einschränkend müssen wir aber festhalten, dass unser Zahlenbeispiel die Voraussetzungen zur Anwendung des approximativen Gauß-Tests nicht vollständig erfüllt – der Stichprobenumfang ist zu gering.

Zweistichprobentests vergleichen die Parameter zweier Stichproben miteinander

In der Realität vergleicht man zwar auch Parameter einer Stichprobe gegen einen hypothetischen Wert, aber genauso wichtig sind Verfahren, die die unbekannt Parameter zweier Stichproben miteinander vergleichen. Beispielsweise braucht man diese Verfahren, wenn man zwei Produktionsanlagen miteinander vergleichen will oder die Folgen eines Prozesseingriffs abschätzen

möchte und eine Stichprobe von vor dem Eingriff mit einer Stichprobe vergleicht, die nach dem Eingriff gezogen wurde.

Am häufigsten vergleicht man in der Realität die Erwartungswerte und die Varianzen bzw. Standardabweichungen zweier Stichproben. Daher betrachten wir diese beiden Verfahren hier näher.

Im Falle des Vergleichs zweier Erwartungswerte muss man unterscheiden, ob man es mit **verbundenen** (d. h. **abhängigen**) oder **unverbundenen** (d. h. **unabhängigen**) **Stichproben** zu tun hat. Damit ist gemeint, ob man für jedes Merkmal

unterschiedliche Stichprobenelemente hat oder ob bei jedem Stichprobenelement beide Merkmale erhoben wurden.

Es gilt die Faustregel: Wenn man die Merkmalsträger der ersten und der zweiten Stichprobe physisch trennen kann, sind es unverbundene Stichproben. Wenn das nicht möglich ist, arbeitet man gerade mit verbundenen Stichproben.

Beispiel

Wenn man einen Vergleich des Konsumverhaltens von Frauen und Männern durchführt, arbeitet man mit unverbundenen Stichproben. Denn man wird das Kaufverhalten von einigen Frauen beobachten und davon getrennt das Kaufverhalten von ausgewählten Männern. Man könnte also alle Frauen in ein Geschäft bringen und die Männer in ein anderes und die beiden Stichproben so physisch voneinander trennen.

Wenn man hingegen die Wirksamkeit einer Nachbearbeitung testet, wird man sinnvollerweise die ausgewählten Werkstücke vor der Bearbeitung messen und dieselben Werkstücke nach der Bearbeitung noch einmal messen. In diesem Fall ist es nicht möglich, die Vorher- und die Nachherstichprobe physisch voneinander zu trennen, also arbeitet man mit verbundenen Stichproben. ◀

Im Falle von verbundenen Stichproben erhebt man für jedes Stichprobenelement die Differenz der beiden Merkmale und führt für die Differenzwerte je nach erfüllten Voraussetzungen einen Gauß-Test, t-Test oder approximativen Gauß-Test durch. Das bedeutet, dass Nullhypothese und Alternative umgeformt werden müssen.

Vergleich von Erwartungswerten bei verbundenen Stichproben

Aus

- $H_0 : \mu_1 \leq \mu_2 \quad A : \mu_1 > \mu_2$
- $H_0 : \mu_1 \geq \mu_2 \quad A : \mu_1 < \mu_2$
- $H_0 : \mu_1 = \mu_2 \quad A : \mu_1 \neq \mu_2$

wird

- $H_0 : \mu_1 - \mu_2 \leq 0 \quad A : \mu_1 > \mu_2$
- $H_0 : \mu_1 - \mu_2 \geq 0 \quad A : \mu_1 < \mu_2$
- $H_0 : \mu_1 - \mu_2 = 0 \quad A : \mu_1 \neq \mu_2$

Dementsprechend müssen alle Messwerte umgeformt werden, indem man die Differenzen berechnet.

Beispiel

Betrachten wir ein Werkstück. Vor Beginn der Nachbearbeitung wurde der Durchmesser bei zehn Werkstücken gemessen, nach der Nachbearbeitung noch einmal. Die

Ergebnisse können der folgenden Tabelle entnommen werden (inkl. Veränderungen):

Werkstück	Durchmesser vorher (x_1)	Durchmesser nachher (x_2)	Veränderung ($d = x_1 - x_2$)
1	156	147	9
2	167	159	8
3	134	134	0
4	147	149	-2
5	120	120	0
6	154	140	14
7	155	151	4
8	140	136	4
9	139	130	9
10	128	125	3

Zur Vereinfachung gehen wir davon aus, dass der Durchmesser normalverteilt, die Varianz aber genau wie der Erwartungswert unbekannt ist. Die Frage ist, ob sich der Durchmesser durch die Nachbearbeitung signifikant verkleinert hat ($\alpha = 5\%$). Dementsprechend führen wir einen t-Test für die Veränderung durch.

Daher wenden wir die erste einseitige Fragestellung an und testen $H_0 : \mu_1 \leq \mu_2$ gegen $A : \mu_1 > \mu_2$ bzw., umgeformt, $H_0 : \mu_1 - \mu_2 \leq 0$ gegen $A : \mu_1 - \mu_2 > 0$.

Es ergibt sich $\bar{x} = 4.9$, $\hat{\sigma}_n = 5,021$, $t_9(0.95) = 1.833$.

Da $4.9 > 0 + 1.833 \cdot \frac{5,021}{\sqrt{10}} = 2.91$, kann H_0 zum Niveau 5% abgelehnt werden. Die Nachbearbeitung scheint also zu signifikanten Verkleinerungen der Durchmesser zu führen. ◀

Im Falle von unverbundenen Stichproben (x_1, \dots, x_{n_1} und y_1, \dots, y_{n_2}) unterscheidet man wieder einige Fälle.

Zweistichproben-Gauß-Test

Beide Stichproben sind normalverteilt, beide Varianzen (σ_1^2 und σ_2^2) sind bekannt.

Dann lauten Nullhypothesen und Alternativen wie folgt:

- $H_0 : \mu_1 \leq \mu_2 \quad A : \mu_1 > \mu_2$
- $H_0 : \mu_1 \geq \mu_2 \quad A : \mu_1 < \mu_2$
- $H_0 : \mu_1 = \mu_2 \quad A : \mu_1 \neq \mu_2$

Die Testvorschrift lautet in den jeweiligen Fällen:

Lehne H_0 ab, falls

- $\frac{\bar{x} - \bar{y}}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} > x_{1-\alpha}$
- $\frac{\bar{x} - \bar{y}}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} < -x_{1-\alpha}$

$$\blacksquare \frac{\bar{x} - \bar{y}}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} > x_{1-\frac{\alpha}{2}} \quad \text{oder} \quad \frac{\bar{x} - \bar{y}}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} < -x_{1-\frac{\alpha}{2}}.$$

Beispiel

Nun können wir das Kaufverhalten zwischen den Geschlechtern näher untersuchen. Gehen wir davon aus, dass eine Forschergruppe 40 Frauen und 50 Männer in einen Baumarkt zum Einkaufen geschickt hat. Im Durchschnitt haben die Frauen $\bar{x} = 95.6$ Euro und die Männer $\bar{y} = 93.3$ Euro ausgegeben.

Heißt das, dass Frauen im Baumarkt mehr Geld ausgeben als Männer? Nehmen wir ein Signifikanzniveau von 5 % an.

Gehen wir davon aus, dass die Stichproben normalverteilt sind und dass die wirklichen Varianzen bekannt sind, nämlich $\sigma_1 = 10.5$ und $\sigma_2 = 10.9$. Wir müssen die erste einseitige Fragestellung benutzen mit $H_0: \mu_1 \leq \mu_2$ gegen $A: \mu_1 > \mu_2$.

Wir dürfen H_0 ablehnen, falls $\frac{\bar{x} - \bar{y}}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \geq x_{1-\alpha}$.

In unserem Fall gilt

$$\begin{aligned} \frac{\bar{x} - \bar{y}}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} &= \frac{95.6 - 93.3}{\sqrt{\frac{10.5^2}{40} + \frac{10.9^2}{50}}} \\ &= 1.015 \cdot x_{1-\alpha} = x_{0.95} = 1.645. \end{aligned}$$

H_0 kann also zum Niveau 5 % nicht abgelehnt werden. Frauen scheinen somit nicht mehr Geld als Männer im Baumarkt auszugeben. \blacktriangleleft

Zweistichproben-t-Test

Beide Stichproben sind normalverteilt, die Varianzen sind unbekannt, aber man weiß, dass $\sigma_1^2 = \sigma_2^2$.

Dann lauten Nullhypothesen und Alternativen wie folgt:

- $H_0: \mu_1 \leq \mu_2 \quad A: \mu_1 > \mu_2$
- $H_0: \mu_1 \geq \mu_2 \quad A: \mu_1 < \mu_2$
- $H_0: \mu_1 = \mu_2 \quad A: \mu_1 \neq \mu_2$

Die Testvorschrift lautet in den jeweiligen Fällen:

Lehne H_0 ab, falls

$$\begin{aligned} \blacksquare \frac{\bar{x} - \bar{y}}{\sqrt{\frac{(n_1-1)\hat{\sigma}_1^2 + (n_2-1)\hat{\sigma}_2^2}{n_1+n_2-2} \cdot \frac{n_1+n_2}{n_1 \cdot n_2}}} &> t_{n_1+n_2-2}(1-\alpha) \\ \blacksquare \frac{\bar{x} - \bar{y}}{\sqrt{\frac{(n_1-1)\hat{\sigma}_1^2 + (n_2-1)\hat{\sigma}_2^2}{n_1+n_2-2} \cdot \frac{n_1+n_2}{n_1 \cdot n_2}}} &< -t_{n_1+n_2-2}(1-\alpha) \end{aligned}$$

$$\begin{aligned} \blacksquare \frac{\bar{x} - \bar{y}}{\sqrt{\frac{(n_1-1)\hat{\sigma}_1^2 + (n_2-1)\hat{\sigma}_2^2}{n_1+n_2-2} \cdot \frac{n_1+n_2}{n_1 \cdot n_2}}} &> t_{n_1+n_2-2}(1-\frac{\alpha}{2}) \\ \text{oder} \\ \frac{\bar{x} - \bar{y}}{\sqrt{\frac{(n_1-1)\hat{\sigma}_1^2 + (n_2-1)\hat{\sigma}_2^2}{n_1+n_2-2} \cdot \frac{n_1+n_2}{n_1 \cdot n_2}}} &< -t_{n_1+n_2-2}(1-\frac{\alpha}{2}). \end{aligned}$$

Beispiel

Wir betrachten wieder die 40 Frauen und 50 Männer die in einen Baumarkt zum Einkaufen geschickt wurden. Wie gesagt haben die Frauen im Durchschnitt $\bar{x} = 95.6$ Euro ausgegeben, die Männer $\bar{y} = 93.3$ Euro. Dieses Mal sind aber die wahren Varianzen unbekannt, daher wurden sie aus den Stichproben geschätzt. Es ergaben sich folgende Werte: $\hat{\sigma}_1 = 10.5$ und $\hat{\sigma}_2 = 10.9$.

Können wir nun belegen, dass Frauen im Baumarkt mehr Geld ausgeben als Männer? Wir arbeiten wiederum mit einem Signifikanzniveau von 5 %.

Gehen wir nun davon aus, dass die Stichproben (die wir hier nicht kennen) normalverteilt waren und dass wir guten Gewissens (dafür sprechen auch die ähnlichen Varianzschätzungen) davon ausgehen dürfen, dass die Varianzen zwar unbekannt, aber gleich sind. Diese Unterstellung muss man eigentlich noch mit einem Zweistichproben-F-Test überprüfen, der gleich vorgestellt wird.

Uns interessiert immer noch die erste einseitige Fragestellung mit $H_0: \mu_1 \leq \mu_2$ gegen $A: \mu_1 > \mu_2$.

Wir dürfen H_0 ablehnen, falls

$$\frac{\bar{x} - \bar{y}}{\sqrt{\frac{(n_1-1)\hat{\sigma}_1^2 + (n_2-1)\hat{\sigma}_2^2}{n_1+n_2-2} \cdot \frac{n_1+n_2}{n_1 \cdot n_2}}} > t_{n_1+n_2-2}(1-\alpha).$$

Hier ergibt sich

$$\begin{aligned} \frac{\bar{x} - \bar{y}}{\sqrt{\frac{(n_1-1)\hat{\sigma}_1^2 + (n_2-1)\hat{\sigma}_2^2}{n_1+n_2-2} \cdot \frac{n_1+n_2}{n_1 \cdot n_2}}} &= \frac{95.6 - 93.3}{\sqrt{\frac{39 \cdot 10.5^2 + 49 \cdot 10.9^2}{88} \cdot \frac{90}{2000}}} \\ &= 1.011, \\ t_{n_1+n_2-2}(1-\alpha) &= t_{88}(0.95) = 1.662. \end{aligned}$$

Da $1.011 < 1.662$, kann H_0 zum Niveau 5 % nicht abgelehnt werden.

Auch in diesem Fall scheinen Frauen nicht mehr Geld im Baumarkt auszugeben als Männer. \blacktriangleleft

Approximativer Zweistichproben-Gauß-Test

Beide Stichproben sind beliebig verteilt, beide Varianzen sind unbekannt, $n_1, n_2 > 30$.

Dann lauten Nullhypothesen und Alternativen wie folgt:

- $H_0 : \mu_1 \leq \mu_2 \quad A : \mu_1 > \mu_2$
- $H_0 : \mu_1 \geq \mu_2 \quad A : \mu_1 < \mu_2$
- $H_0 : \mu_1 = \mu_2 \quad A : \mu_1 \neq \mu_2$

Die Testvorschrift lautet in den jeweiligen Fällen:

Lehne H_0 ab, falls

- $\frac{\bar{x} - \bar{y}}{\sqrt{\frac{\hat{\sigma}_1^2}{n_1} + \frac{\hat{\sigma}_2^2}{n_2}}} > x_{1-\alpha}$
- $\frac{\bar{x} - \bar{y}}{\sqrt{\frac{\hat{\sigma}_1^2}{n_1} + \frac{\hat{\sigma}_2^2}{n_2}}} < -x_{1-\alpha}$
- $\frac{\bar{x} - \bar{y}}{\sqrt{\frac{\hat{\sigma}_1^2}{n_1} + \frac{\hat{\sigma}_2^2}{n_2}}} > x_{1-\frac{\alpha}{2}}$ oder $\frac{\bar{x} - \bar{y}}{\sqrt{\frac{\hat{\sigma}_1^2}{n_1} + \frac{\hat{\sigma}_2^2}{n_2}}} < -x_{1-\frac{\alpha}{2}}$.

Beispiel

Wir betrachten wieder die 40 Frauen und 50 Männer die in einen Baumarkt zu Einkaufen geschickt wurden. Wie gesagt haben die Frauen im Durchschnitt $\bar{x} = 95.6$ Euro ausgegeben, die Männer $\bar{y} = 93.3$ Euro. Dieses Mal sind aber die wahren Varianzen unbekannt, daher wurden sie aus den Stichproben geschätzt. Es ergaben sich folgende Werte: $\hat{\sigma}_1 = 10.5$ und $\hat{\sigma}_2 = 10.9$.

Können wir nun belegen, dass Frauen im Baumarkt mehr Geld ausgeben als Männer? Wir arbeiten wiederum mit einem Signifikanzniveau von 5 %.

Diesmal unterstellen wir keine Normalverteilung. Das ist auch nicht nötig, da die Varianzen unbekannt, die Stichprobenumfänge aber groß genug sind, um den approximativen Gauß-Test durchzuführen.

Wir wollen überprüfen, ob $\mu_1 > \mu_2$, also müssen wir die erste einseitige Fragestellung anwenden mit $H_0 : \mu_1 \leq \mu_2$ gegen $A : \mu_1 > \mu_2$.

Wir können H_0 ablehnen, falls $\frac{\bar{x} - \bar{y}}{\sqrt{\frac{\hat{\sigma}_1^2}{n_1} + \frac{\hat{\sigma}_2^2}{n_2}}} > x_{1-\alpha}$.

Es gilt

$$\frac{\bar{x} - \bar{y}}{\sqrt{\frac{\hat{\sigma}_1^2}{n_1} + \frac{\hat{\sigma}_2^2}{n_2}}} = \frac{95.6 - 93.3}{\sqrt{\frac{10.5^2}{40} + \frac{10.9^2}{50}}} = 1.015,$$

$$x_{1-\alpha} = x_{0.95} = 1.645.$$

H_0 kann also zum Niveau 5 % nicht abgelehnt werden. Auch in diesem Fall scheinen Frauen nicht mehr Geld im Baumarkt auszugeben als Männer. ◀

Zweistichproben-F-Test

Möchte man die Varianzen zweier unabhängiger Stichproben miteinander vergleichen, so führt man einen Zwei-

stichproben-F-Test durch. Die Hypothesen und Alternativen lauten dann wie folgt:

- $H_0 : \sigma_1^2 \leq \sigma_2^2 \quad A : \sigma_1^2 > \sigma_2^2$
- $H_0 : \sigma_1^2 \geq \sigma_2^2 \quad A : \sigma_1^2 < \sigma_2^2$
- $H_0 : \sigma_1^2 = \sigma_2^2 \quad A : \sigma_1^2 \neq \sigma_2^2$.

Falls beide Stichproben normalverteilt sind, lautet die Testvorschrift in den jeweiligen Fällen: Lehne H_0 ab, falls

- $\frac{\hat{\sigma}_1^2}{\hat{\sigma}_2^2} > F_{n_1-1, n_2-1}(1 - \alpha)$
- $\frac{\hat{\sigma}_1^2}{\hat{\sigma}_2^2} < \frac{1}{F_{n_2-1, n_1-1}(1 - \alpha)}$
- $\frac{\hat{\sigma}_1^2}{\hat{\sigma}_2^2} > F_{n_1-1, n_2-1}(1 - \frac{\alpha}{2})$ oder $\frac{\hat{\sigma}_1^2}{\hat{\sigma}_2^2} < \frac{1}{F_{n_2-1, n_1-1}(1 - \frac{\alpha}{2})}$.

$F_{m,n}(p)$ bezeichnet hierbei die Quantile der F-Verteilung mit den Freiheitsgraden m und n . Die entsprechenden Quantile sind in Tab. A.14 aufgeführt. Wichtig ist hierbei die Reihenfolge der Indizes. Denn $F_{n,m}(p)$ ist etwas anderes als $F_{m,n}(p)$. In Tab. A.14 steht der erste Index in der ersten Spalte und der zweite Index in der ersten Zeile.

Beispiel

Gehen wir davon aus, dass in der ersten Stichprobe sieben Objekte und in der zweiten Stichprobe acht Objekte waren. Falls wir auf dem Signifikanzniveau 10 % einen Test durchführen wollen, ob die Varianzen der beiden Stichproben gleich sind, brauchen wir also einmal

$$F_{6,7}(0.95) = 3.86597776$$

und einmal

$$F_{7,6}(0.95) = 4.206668791. \quad \blacktriangleleft$$

In diesem Buch findet man wie in allen anderen Statistiklehrbüchern nur Tabellen für $p = 0.95$ und $p = 0.99$. Ein häufiger Anfängerfehler besteht darin, die beiden Tabellen zu verwechseln.

Für einige weitere p -Werte kann man die Quantile der F-Verteilung mit folgender Formel berechnen:

Es gilt der folgende Zusammenhang:

$$F_{m,n}(p) = \frac{1}{F_{n,m}(1-p)}$$

Den F-Test braucht man übrigens in der Praxis sehr oft als „Vor-test“. Denn man wundert sich beim Zweistichproben-t-Test ja

schon. Die Varianzen müssen unbekannt, aber gleich sein?! Wie soll das gehen?

Die Antwort ist, man hat zuerst einen F-Test in der zweiseitigen Fragestellung durchgeführt und konnte die Nullhypothese nicht verwerfen. Dann kennt man zwar die Varianzen immer noch nicht, aber man darf annehmen, dass sie gleich sind.

Beispiel

Wir kommen noch einmal auf unsere Studie zum Kaufverhalten von Frauen und Männern zurück. Wir überprüfen, ob die Varianzen gleich sind, denn wir wollen Klarheit haben, ob das Beispiel, das gewählt wurde, auch wirklich für den Zweistichproben-t-Test geeignet ist. Ein Signifikanzniveau von 2% wird für ausreichend exakt gehalten. Es werden 41 Frauen und 51 Männer in einen Baumarkt zum Einkaufen geschickt. (Den leicht veränderten Stichprobenumfang brauchen wir, damit wir in den Tabellen ablesen können.)

Die geschätzten Varianzen betragen $\hat{\sigma}_1^2 = 10.5^2$ und $\hat{\sigma}_2^2 = 10.9^2$. Wir überprüfen die zweiseitige Fragestellung, also $H_0 : \sigma_1^2 = \sigma_2^2$ gegen $A : \sigma_1^2 \neq \sigma_2^2$,

und wir dürfen H_0 ablehnen, falls

$$\frac{\hat{\sigma}_1^2}{\hat{\sigma}_2^2} \geq F_{n_1-1; n_2-1} \left(1 - \frac{\alpha}{2}\right)$$

oder

$$\frac{\hat{\sigma}_1^2}{\hat{\sigma}_2^2} \leq \frac{1}{F_{n_2-1; n_1-1} \left(1 - \frac{\alpha}{2}\right)}.$$

Es gilt $\frac{\hat{\sigma}_1^2}{\hat{\sigma}_2^2} = \frac{10.5^2}{10.9^2} = 0.928$.

$F_{n_1-1; n_2-1} \left(1 - \frac{\alpha}{2}\right) = F_{40; 50}(0.99) = 2.01$. Die erste Bedingung ist also nicht erfüllt.

Vielleicht die zweite:

$$\frac{1}{F_{n_2-1; n_1-1} \left(1 - \frac{\alpha}{2}\right)} = \frac{1}{F_{50; 40}(0.99)} = \frac{1}{2.06} = 0.485,$$

also ist auch die zweite Bedingung nicht erfüllt. Wir dürfen die Nullhypothese nicht verwerfen. Die Varianzen scheinen also gleich zu sein (aber immer noch unbekannt). ◀

Parameterfreie Tests überprüfen Hypothesen, die nicht die Parameter einer Verteilung betreffen

Bislang haben wir Tests besprochen, die Hypothesen bezüglich eines Parameters einer Verteilung geprüft haben. Das ist natürlich nicht alles, was uns interessiert. Weitere wichtige Fragestellungen sind z. B. „Welche Verteilung haben meine Daten?“

und „Sind meine Daten unabhängig?“. Solche Fragestellungen können mit **parameterfreien Tests** (oder auch synonym **nicht parametrischen Tests**) bearbeitet werden.

In vielen Fällen ist man daran interessiert zu wissen, ob die unbekannte Verteilung einer Grundgesamtheit gleich einer vermuteten Verteilung ist. Beispielsweise sollte man vor der Durchführung eines t-Tests in der Praxis erst einmal überprüfen, ob die Daten eigentlich normalverteilt sind.

Ein möglicher Test, um eine vorgegebene Verteilung zu testen, ist der **χ^2 -Anpassungstest**. Das ist nicht unbedingt der beste Test, da er eine relativ große Stichprobe benötigt, um hinreichend exakt zu sein. Statistikprogramme wie SPSS bieten auch den **Kolmogorov-Smirnov-Test** an, der bei kleineren Stichproben sinnvoller ist. Aber den χ^2 -Anpassungstest kann man mit überschaubarem Aufwand mit dem Taschenrechner durchführen.

Den χ^2 -Anpassungstest gibt es in drei Varianten, je nachdem, welche Verteilung man vermutet und welche Informationen man über die Verteilung besitzt.

χ^2 -Anpassungstest für diskrete Verteilungen

Hier betrachten wir den Fall, dass eine diskrete Verteilung mit den Ausprägungen x_1, x_2, \dots, x_m und der Wahrscheinlichkeitsfunktion P vorliegt, wobei $p_i = P(x_j)$ unbekannt ist.

Etwaige Parameter der zu überprüfenden Verteilung werden aber als bekannt vorausgesetzt.

Dann lauten die Hypothese und die Alternative wie folgt:

$$H_0 : p_j = \bar{p}_j \text{ für alle } j = 1, \dots, m$$

gegen

$$A : p_j \neq \bar{p}_j \text{ für mindestens ein } j$$

Die Testvorschrift lautet dann:

Lehne H_0 ab, falls $V > \chi_{m-1}^2(1 - \alpha)$, wobei

$$V = \sum_{j=1}^m \frac{(n_j - n \cdot \bar{p}_j)^2}{n \cdot \bar{p}_j} = \frac{1}{n} \sum_{j=1}^m \frac{n_j^2}{\bar{p}_j} - n.$$

Dabei muss gelten: $n \cdot \bar{p}_j \geq 5$ für alle $1 \leq j \leq m$.

Als Hilfe legt man sich normalerweise eine Tabelle an.

Beispiel

Nehmen wir an, wir finden uns nachts in einer zwielichtigen Kneipe wieder und werden zum Würfelspiel (natürlich gegen Geldeinsatz) eingeladen. Da sollte es von Interesse sein herauszufinden, ob der Würfel gefälscht ist.

Wir leihen uns also (kurz) den Würfel und würfeln 100-mal. Die folgende Tabelle zeigt das Ergebnis:

Ergebnis j	1	2	3	4	5	6
Häufigkeit n_j	23	26	17	14	11	9

Unser Hypothesensystem lautet dann natürlich:

$$H_0 : p_j = \frac{1}{6} \text{ für } j = 1, \dots, 6 \text{ (dann ist der Würfel fair)}$$

gegen

$$A : p_j \neq \frac{1}{6} \text{ für mindestens ein } j \text{ (dann ist der Würfel gefälscht).}$$

Es gilt, dass wir H_0 ablehnen (und wegrennen) müssen, falls

$$V > \chi_{m-1}^2(1 - \alpha),$$

$$\text{wobei } V = \sum_{j=1}^m \frac{(n_j - n \cdot \bar{p}_j)^2}{n \cdot \bar{p}_j} = \frac{1}{n} \sum_{j=1}^m \frac{n_j^2}{\bar{p}_j} - n.$$

Ein Signifikanzniveau von 5% bietet ausreichende Sicherheit.

Wir machen uns zur Berechnung von V eine Tabelle:

\bar{p}_j	$n \cdot \bar{p}_j$	n_j
1/6	$100 \cdot \frac{1}{6} = 16.67$	23
1/6	$100 \cdot \frac{1}{6} = 16.67$	26
1/6	$100 \cdot \frac{1}{6} = 16.67$	17
1/6	$100 \cdot \frac{1}{6} = 16.67$	14
1/6	$100 \cdot \frac{1}{6} = 16.67$	11
1/6	$100 \cdot \frac{1}{6} = 16.67$	9

Die Bedingung $n \cdot \bar{p}_j \geq 5$ ist für alle Ergebnisse erfüllt, wir können also weiterrechnen.

Wir nehmen die zweite Variante, um V auszurechnen:

$$\begin{aligned} V &= \frac{1}{n} \sum_{j=1}^m \frac{n_j^2}{\bar{p}_j} - n \\ &= \frac{1}{100} \left(\frac{23^2}{\frac{1}{6}} + \frac{26^2}{\frac{1}{6}} + \frac{17^2}{\frac{1}{6}} + \frac{14^2}{\frac{1}{6}} + \frac{11^2}{\frac{1}{6}} + \frac{9^2}{\frac{1}{6}} \right) \\ &\quad - 100 \\ &= 13.52 \end{aligned}$$

Das gesuchte Quantil ist

$$\chi_{m-1}^2(1 - \alpha) = \chi_5^2(0.95) = 11.07.$$

Da $V = 13.52 > \chi_{m-1}^2(1 - \alpha) = \chi_5^2(0.95) = 11.07$, sollten wir unsere Nullhypothese ablehnen und schleunigst den Laden verlassen. ▶

χ^2 -Anpassungstest für stetige Verteilungen

Hier betrachten wir den Fall, dass eine stetige Verteilung mit der unbekanntem Verteilungsfunktion F^X vorliegt. Etwaige Parameter der Verteilung werden wiederum als bekannt vorausgesetzt.

Dann lauten die Hypothese und die Alternative wie folgt:

$$H_0 : F^X = F_0^X \quad A : F^X \neq F_0^X$$

Die Testvorschrift lautet dann ebenfalls: Lehne H_0 ab, falls $V > \chi_{m-1}^2(1 - \alpha)$.

Hier müssen aber zuerst einige Berechnungen durchgeführt werden: Zuerst wird die x -Achse in m Intervalle zerlegt, nämlich

$$A_1 = (-\infty; x_1], A_2 = (x_1; x_2], \dots, A_m = (x_{m-1}; \infty[.$$

Dann ist

$$\begin{aligned} \bar{p}_1 &= F_0^X(x_1), \\ \bar{p}_2 &= F_0^X(x_2) - F_0^X(x_1), \dots, \\ \bar{p}_m &= 1 - F_0^X(x_{m-1}). \end{aligned}$$

Dann berechnet sich V wie oben, also

$$V = \sum_{j=1}^m \frac{(n_j - n \cdot \bar{p}_j)^2}{n \cdot \bar{p}_j} = \frac{1}{n} \sum_{j=1}^m \frac{n_j^2}{\bar{p}_j} - n.$$

Dabei muss auch hier gelten: $n \cdot \bar{p}_j \geq 5$ für alle $1 \leq j \leq m$.

Wenn die Bedingung nicht erfüllt ist, fasst man benachbarte Intervalle so lange zusammen, bis die Bedingung für alle Intervalle erfüllt ist.

Als Hilfe legt man sich normalerweise dieselbe Tabelle wie oben an.

χ^2 -Anpassungstest bei unbekanntem Parametern

Hier betrachten wir den Fall, dass eine Verteilung vermutet wird (egal ob diskret oder stetig) und dass zusätzlich k Parameter der Verteilung unbekannt sind.

In diesem Fall müssen die k unbekanntem Parameter zuerst mithilfe der Maximum-Likelihood-Methode geschätzt werden.

Anschließend werden die entsprechenden Hypothesen und Alternativen entweder des diskreten oder des stetigen Tests benutzt, also im diskreten Fall:

$$H_0 : p_j = \bar{p}_j \text{ für alle } j = 1, \dots, m$$

Anwendung: χ^2 -Anpassungstest für stetige Verteilungen

In einer Sägemühle werden Balken der Länge 2 m gesägt. Die Qualitätssicherung erlaubt eine Standardabweichung von $\sigma = 0.05$ m.

Um zu überprüfen, ob die Länge der Balken normalverteilt, genauer gesagt $\mathcal{N}(2; 0.05)$ -verteilt, ist, wurde eine Stichprobe vom Umfang 1000 gezogen. Das Ergebnis kann der folgenden Tabelle entnommen werden:

Klasse A_j	Häufigkeit n_j
$(-\infty; 1.90]$	0
$(1.90; 2.00]$	530
$(2.00; 2.10]$	470
$(2.10; \infty)$	0

Zur Überprüfung legen wir eine Tabelle wie beim diskreten χ^2 -Anpassungstest an, aber diesmal müssen wir die \bar{p}_j zuerst gesondert berechnen:

$$\begin{aligned}\bar{p}_1 &= P(X \leq 1.9) = \Phi\left(\frac{1.9 - 2}{0.05}\right) = \Phi(-2) \\ &= 1 - \Phi(2) = 1 - 0.977 = 0.023, \\ \bar{p}_2 &= P(X \leq 2) - P(X \leq 1.9) \\ &= \Phi\left(\frac{2 - 2}{0.05}\right) - \Phi\left(\frac{1.9 - 2}{0.05}\right) = \Phi(0) - \Phi(-2) \\ &= 0.5 - (1 - \Phi(2)) = -0.5 + 0.977 = 0.477, \\ \bar{p}_3 &= P(X \leq 2.1) - P(X \leq 2) \\ &= \Phi\left(\frac{2.1 - 2}{0.05}\right) - \Phi\left(\frac{2 - 2}{0.05}\right) \\ &= \Phi(2) - \Phi(0) = 0.977 - 0.5 = 0.477,\end{aligned}$$

$$\begin{aligned}\bar{p}_4 &= 1 - P(X \leq 2.1) = 1 - \Phi\left(\frac{2.1 - 2}{0.05}\right) \\ &= 1 - \Phi(2) = 1 - 0.977 = 0.023.\end{aligned}$$

Jetzt können wir folgende Tabelle aufstellen:

A_j	\bar{p}_j	$n \cdot \bar{p}_j$	n_j
$(-\infty; 1.90]$	0.023	$1000 \cdot 0.023 = 23$	0
$(1.90; 2.00]$	0.477	$1000 \cdot 0.477 = 477$	530
$(2.00; 2.10]$	0.477	$1000 \cdot 0.477 = 477$	470
$(2.10; \infty)$	0.023	$1000 \cdot 0.023 = 23$	0

Die Bedingung $n \cdot \bar{p}_j \geq 5$ ist für alle Klassen erfüllt, wir müssen also keine Klassen zusammenfassen.

Dann ist auch die Berechnung von V ganz einfach:

$$\begin{aligned}V &= \frac{1}{n} \sum_{j=1}^m \frac{n_j^2}{\bar{p}_j} - n \\ &= \frac{1}{1000} \cdot \left(\frac{0^2}{0.023} + \frac{530^2}{0.477} + \frac{470^2}{0.477} + \frac{0^2}{0.023} \right) - 1000 \\ &= 51.99.\end{aligned}$$

Diesen Wert müssen wir mit

$$\chi_{m-1}^2(1 - \alpha) = \chi_3^2(0.95) = 7.81$$

vergleichen und dürfen die Nullhypothese verwerfen. Die Länge der Balken scheint nicht $(2; 0.05)$ -verteilt zu sein.

gegen

$$A : p_j \neq \bar{p}_j \text{ für mindestens ein } j$$

und im stetigen Fall:

$$H_0 : F^X = F_0^X \quad A : F^X \neq F_0^X.$$

Die Testvorschrift lautet dann:

Lehne H_0 ab, falls $V > \chi_{m-1-k}^2(1 - \alpha)$.

Auch hierbei berechnet sich V wie oben, also

$$V = \sum_{j=1}^m \frac{(n_j - n \cdot \bar{p}_j)^2}{n \cdot \bar{p}_j} = \frac{1}{n} \sum_{j=1}^m \frac{n_j^2}{\bar{p}_j} - n.$$

Dabei muss auch hier gelten: $n \cdot \bar{p}_j \geq 5$ für alle $1 \leq j \leq m$.

Wenn die Bedingung nicht erfüllt ist, fasst man benachbarte Intervalle so lange zusammen, bis die Bedingung für alle Intervalle erfüllt ist.

Achtung Auch wenn es nur ganz klein ist, aber in dem Index steht ein $-k$. ▶

Ansonsten geht man genauso vor, wie wenn man keine Parameter geschätzt hat. Mit der Hilfstabelle geht dies relativ schnell.

Anwendung: χ^2 -Anpassungstest bei unbekanntem Parametern

Kommen wir noch einmal auf die Balken aus dem obigen Beispiel zurück. Vielleicht sind sie ja normalverteilt, aber die Parameter, die vorgegeben waren, stimmten gar nicht?!

Nehmen wir an, dass die Stichprobe die folgenden Schätzungen ergeben hat:

$\bar{x} = 2.03$ und $\hat{\sigma}_n = 0.09$.

Diesmal berechnen sich die \bar{p}_j wie folgt:

$$\begin{aligned} \bar{p}_1 &= P(X \leq 1.9) = \Phi\left(\frac{1.9 - 2.03}{0.09}\right) \\ &= \Phi(-1.44) = 1 - \Phi(1.44) = 1 - 0.925 = 0.075, \\ \bar{p}_2 &= P(X \leq 2) - P(X \leq 1.9) \\ &= \Phi\left(\frac{2 - 2.03}{0.09}\right) - \Phi\left(\frac{1.9 - 2.03}{0.09}\right) \\ &= \Phi(-0.33) - \Phi(-1.44) = \Phi(1.44) - \Phi(0.33) \\ &= 0.925 - 0.629 = 0.296, \\ \bar{p}_3 &= P(X \leq 2.1) - P(X \leq 2) \\ &= \Phi\left(\frac{2.1 - 2.03}{0.09}\right) - \Phi\left(\frac{2 - 2.03}{0.09}\right) \\ &= \Phi(0.78) - \Phi(-0.33) = 0.782 - (1 - 0.629) = 0.357, \\ \bar{p}_4 &= 1 - P(X \leq 2.1) = 1 - \Phi\left(\frac{2.1 - 2.03}{0.09}\right) \\ &= 1 - \Phi(0.78) = 1 - 0.782 = 0.218. \end{aligned}$$

Jetzt können wir folgende Tabelle aufstellen:

A_j	\bar{p}_j	$n \cdot \bar{p}_j$	n_j
$(-\infty; 1.90]$	0.075	$1000 \cdot 0.075 = 75$	0
$(1.90; 2.00]$	0.296	$1000 \cdot 0.296 = 296$	530
$(2.00; 2.10]$	0.357	$1000 \cdot 0.357 = 357$	470
$(2.10; \infty)$	0.218	$1000 \cdot 0.218 = 218$	0

Die Bedingung $n \cdot \bar{p}_j \geq 5$ ist für alle Klassen erfüllt, wir müssen also keine Klassen zusammenfassen.

Dann ist auch die Berechnung von V ganz einfach:

$$\begin{aligned} V &= \frac{1}{n} \sum_{j=1}^m \frac{n_j^2}{\bar{p}_j} - n \\ &= \frac{1}{1000} \cdot \left(\frac{0^2}{0.075} + \frac{530^2}{0.296} + \frac{470^2}{0.357} + \frac{0^2}{0.218} \right) - 1000 \\ &= 567.75. \end{aligned}$$

Diesen Wert müssen wir mit

$$\chi_{m-k-1}^2(1 - \alpha) = \chi_1^2(0.95) = 3.84$$

vergleichen. Wir haben ja zwei Parameter geschätzt, also ist $k = 2$.

Wiederum dürfen wir die Nullhypothese verwerfen. Die Länge der Balken scheint also gar nicht normalverteilt zu sein.

Aufgaben

18.1 Die Zufallsvariablen X_1, \dots, X_5 seien unabhängig. Weiter soll angenommen werden, dass sie dieselbe Verteilung haben und dass $E[X_i] = \mu$ und $\text{Var}(X_i) = \sigma^2$ für $i = 1, \dots, 5$ gilt. Welche der folgenden Schätzfunktionen

- a) sind erwartungstreu für μ ?
- b) ist unter den erwartungstreuen diejenige mit der kleinsten Varianz?
- $A = X_1 + \frac{3}{5}X_2 - \frac{1}{5}X_3 - \frac{1}{5}X_4 - \frac{1}{5}X_5$
- $B = \frac{1}{5}X_1 + \frac{1}{5}X_2 + \frac{2}{5}X_3 + \frac{3}{7}X_4 + \frac{1}{8}X_5$
- $C = \frac{1}{2}(X_1 + X_2)$
- $D = 1 + \frac{1}{5}(X_1 + X_2 + X_3 + X_4 + X_5)$
- $E = \frac{1}{5}(X_1 + X_2 + X_3 + X_4 + X_5)$

18.2 Ein Autohersteller produziert Luxusautos. Der Verkaufsleiter hat für die letzten 24 Monate notiert, wie viele Autos in jedem Monat verkauft wurden. Diese Zahlen sind in der nach-

folgenden Tabelle angegeben:

Anzahl k der in einem Monat verkauften Autos	0	1	2	3	4	5	6	≥ 7
Anzahl der Monate mit k verkauften Autos	2	5	6	5	3	2	1	0

- a) Die Ereignisse, 6, 7 oder mehr Luxusautos in einem Monat zu verkaufen, sind offensichtlich sehr selten. Welche der Ihnen bekannten Verteilungen können Sie dem Verkaufsleiter zur Modellierung für das Auftreten seltener Ereignisse vorschlagen?
- b) Die in a) bestimmte Verteilung hängt nur von einem Parameter ab. Schätzen Sie diesen Parameter mithilfe der Maximum-Likelihood-Methode für das hier angegebene Beispiel.
- c) Bestimmen Sie die relativen Häufigkeiten für den Verkauf von 0, 1, 2, ..., 6 Autos in einem Monat.
- d) Berechnen Sie für die in a) bestimmte Verteilung unter Verwendung des in b) ermittelten Parameters die Wahrscheinlichkeiten für die Ereignisse 0, 1, 2, ..., 6 Autos in einem Monat zu verkaufen.
- e) Vergleichen Sie die relativen Häufigkeiten aus c) mit den in d) bestimmten Wahrscheinlichkeiten.

18.3 Die Zufallsvariable X sei $\mathcal{N}(\mu; 2)$ -verteilt, wobei μ unbekannt ist. Eine Stichprobe vom Umfang 100 ergab einen Mittelwert von 4.5. Bestimmen Sie ein (zweiseitiges) Intervall, welches den gesuchten Erwartungswert μ mit einer Wahrscheinlichkeit von mindestens 0.9 enthält.

18.4 Für das Produkt A stehe eine neue Produktionsmethode zur Verfügung. Es kann angenommen werden, dass die Produktionsdauer $X \mathcal{N}(\mu; \sigma^2)$ -verteilt ist, wobei μ und σ unbekannt sind. Eine Stichprobe vom Umfang 10 brachte das folgende Ergebnis, wobei die jeweilige Produktionsdauer in Minuten notiert ist:

i	1	2	3	4	5	6	7	8	9	10
x_i	55	45	50	54	56	58	52	49	61	60

- a) Bestimmen Sie zum Konfidenzniveau 0.95 ein zweiseitiges Konfidenzintervall für μ .
- b) Bestimmen Sie zum Konfidenzniveau 0.95 ein zweiseitiges Konfidenzintervall für σ .

18.5 Eine Krankenversicherung hat zur Berechnung eines Tarifs 1000 zufällig ausgewählte Versicherte befragt, wie viele Tage X im Jahr sie Ski fahren. Die Auswertung der Befragung ergab

$$\bar{x}_n = 4.5 \text{ und } s_n = 3.2.$$

Bestimmen Sie ein zweiseitiges Konfidenzintervall für μ zum Konfidenzniveau 0.95.

18.6 Ein Vertrag über die Lieferung von Ravioli in 750-ml-Dosen sieht vor, dass der Abnehmer einen Preisnachlass fordern darf, wenn er (bei einer Fehlerwahrscheinlichkeit $\alpha = 5\%$) nachweist, dass der zu erwartende Doseninhalt μ kleiner als 750 ml ist. Er entnimmt dazu der Lieferung unabhängig und zufällig $n = 25$ Dosen und bestimmt den jeweiligen Inhalt [in ml]:

748	753	741	761	759
756	749	737	749	755
744	742	755	751	743
753	758	748	757	735
743	742	759	741	742

Der Doseninhalt sei eine $\mathcal{N}(\mu; 4)$ -verteilte Zufallsgröße.

- a) Wie ist die Fragestellung als statistisches Testproblem zu formulieren?
- b) Welches Ergebnis resultiert aus dem zugehörigen statistischen Test? Darf der Abnehmer einen Preisnachlass fordern?
- c) Mit welcher Wahrscheinlichkeit verzichtet der Abnehmer aufgrund der in b) benutzten Testvorschrift irrtümlich (beim Stichprobenumfang $n = 25$) auf den vertragsmäßig vereinbarten Rabatt, falls der zu erwartende Doseninhalt gleich 748 ml ist? Welche Wahrscheinlichkeit ist damit im Sinne der Testtheorie gemeint?
- d) Wie groß muss n gewählt werden, damit die in c) genannte Wahrscheinlichkeit höchstens 10 % beträgt?

e) Welche Testvorschrift müsste der Ravioliproduzent benutzen, wenn er (bei einer Fehlerwahrscheinlichkeit von 5 %) nachweisen will, dass der zu erwartende Doseninhalt μ tatsächlich größer als 750 ml ist?

18.7 Ein Unternehmen hatte in einer Mitteilung vor dem Börsengang bekannt gegeben, dass die durchschnittliche Produktionsdauer für das Gut A weniger als 12 Stunden betrage. Von der Produktionsdauer kann angenommen werden, dass sie normalverteilt ist. Eine Überprüfung vom Umfang 25 ergab nun ein Stichprobenmittel von 11.5 Stunden und eine Stichprobenstandardabweichung von 3 Stunden.

Muss das Unternehmen die Hypothese, dass die durchschnittliche Produktionsdauer weniger als 12 h beträgt, aufgrund der Stichprobe bei einem Signifikanzniveau von 0.01 revidieren?

18.8 Als Portfoliomanager investieren Sie in die Aktien A und B. Für diese beiden Papiere werden folgende Stichproben von Schlusskursen (in Euro) notiert:

Zeitpkt.	1	2	3	4	5	6
Aktie A	12.10	12.50	10.50	10.30	11.60	k. A.
Aktie B	14.30	16.00	16.50	16.40	15.30	14.50

Es wird (hier) unterstellt, dass die Schlusskurse für den beobachteten Zeitraum normalverteilt und stochastisch unabhängig sind.

Für das Risikocontrolling Ihres Arbeitgebers ist es wichtig, das „Risiko“ von Aktien zu messen. Als Maß für das Risiko einer Aktie werde vereinfachend die Standardabweichung der Schlusskurse herangezogen.

- a) Testen Sie die Nullhypothese „Das Risikomaß der Aktie A ist gleich 0.60 Euro“ (Irrtumswahrscheinlichkeit 5 %).
- b) Können Sie die Vermutung, dass die Aktie A risikoreicher ist als die Aktie B, durch einen Test zum Niveau 1 % statistisch untermauern?

18.9 Ein Hersteller von Kühlschränken bietet fünf verschiedene Modelle an. Die Firma geht davon aus, dass 30 % der verkauften Kühlschränke vom Typ 1, 35 % vom Typ 2, 15 % vom Typ 3, 10 % vom Typ 4 und 10 % vom Typ 5 sind. Von 500 zufällig ausgewählten Käufern eines Kühlschranks dieses Herstellers haben 120 den Kühlschrank vom Typ 1, 135 den vom Typ 2, 120 den vom Typ 3, 65 den vom Typ 4 und 60 den vom Typ 5 gekauft.

Formulieren Sie Hypothese und Alternative. Kann man die Hypothese H_0 zum Signifikanzniveau 0.01 verwerfen?

18.10 Ein Unternehmen möchte überprüfen, ob die Lebensdauer eines von ihm produzierten Elektronikbauteils exponential mit Parameter $\lambda = 0.25$ verteilt ist. Eine Stichprobe vom Um-

fang 30 ergab folgende Werte:

i	1	2	3	4	5	6	7	8	9
x_i	3	4	2.5	5	8	3	5	4.5	5.8
i	10	11	12	13	14	15	16	17	18
x_i	4	2.5	7	1.5	2.5	6	3.5	4	5.5
i	19	20	21	22	23	24	25	26	27
x_i	5	10	3	5	4.5	8	4	1.5	5
i	28	29	30						
x_i	4	2	6						

Überprüfen Sie die Hypothese zum Signifikanzniveau 0.05. Verwenden Sie folgende Intervalleinteilung: $A_1 = [0; 1], A_2 = (1; 3], A_3 = (3; 5]$ und $A_4 = (5; \infty)$.

18.11

X sei die Zeitdauer für die Abfertigung eines Kunden an der Kasse eines Supermarktes. Es wurde eine Stichprobe vom Umfang 50 entnommen, um zu überprüfen, ob X eine Exponentialverteilung besitzt, dabei lagen die Abfertigungszeiten (in Sekunden) mit folgenden Häufigkeiten n_i im Intervall A_i :

i	I_i	n_i
1	$[0; 15]$	2
2	$(15; 30]$	7
3	$(30; 45]$	7
4	$(45; 60]$	10
5	$(60; 90]$	12
6	$(90; 120]$	6
7	$(120; 180]$	4
8	$(180; \infty)$	2

Das Stichprobenmittel betrug 54. Testen Sie die Behauptung zum Niveau $\alpha = 0.05$.