

Statistik

Aufgabe 1 Es sei $\Omega = \{1, 2, 3, 4\}$ und

$$\mathcal{S} = \{\emptyset, \{3\}, \{4\}, \{1, 2\}, \{3, 4\}, \{1, 2, 3\}, \{1, 2, 4\}, \Omega\}.$$

- a) Zeigen Sie, daß $\mathcal{S}$ eine σ -Algebra ist.
- b) Geben Sie ein Erzeugendensystem von $\mathcal{S}$ an.

Aufgabe 2 Sie holen einen Freund vom Flughafen ab. Erfahrungsgemäß ist das Flugzeug in 70% aller Fälle verspätet, falls es regnet, und in 20% aller Fälle, falls es nicht regnet. Der Wetterbericht prophezeit eine Regenwahrscheinlichkeit von 40%. Wie groß ist die Wahrscheinlichkeit, daß das Flugzeug Verspätung hat? Wie groß ist die Wahrscheinlichkeit, daß es regnet, falls das Flugzeug Verspätung hat?

Aufgabe 3 Drei Kugeln $K1, K2, K3$ werden auf vier Urnen $U1, U2, U3, U4$ verteilt.

- a) Geben Sie aufgrund des Prinzips der Gleichwahrscheinlichkeit einen geeigneten Wahrscheinlichkeitsraum $(\Omega, \mathcal{S}, P)$ für dieses Zufallsexperiment an.
- b) Bestimmen Sie die Wahrscheinlichkeit der folgenden Ereignisse:

$$A = "K1 \text{ liegt in } U3", \quad B = "In \text{ jeder Urne liegt höchstens eine Kugel"}.$$

Aufgabe 4 Auf einem gegebenen Wahrscheinlichkeitsraum $(\Omega, \mathcal{S}, P)$ soll die Zufallsvariable

$$X = (X_n)_{1 \leq n \leq N} : \Omega \rightarrow \Omega' := \{0, 1, 2\}^N \quad (1)$$

das Zufallsexperiment "Tagesproduktion von N Taschenrechnern" beschreiben, X_n beschreibt dabei wiefolgt den Zustand von Taschenrechner n , $1 \leq n \leq N$:

$$X_n(\omega) = \begin{cases} 0, & \text{Taschenrechner defekt,} \\ 1, & \text{Taschenrechner ok, Gehäuse defekt,} \\ 2, & \text{Taschenrechner und Gehäuse ok.} \end{cases} \quad (2)$$

Drücken Sie mit Hilfe der Zufallsvariablen X_n folgende Ereignisse aus:

1. Der n -te Taschenrechner ist voll intakt.
2. Kein Taschenrechner ist ok.
3. Mindestens ein Rechner ist nicht ok oder am Gehäuse beschädigt.

Aufgabe 5 Die Zufallsvariable X sei Poisson-verteilt mit Parameter $\lambda = 9$. Geben Sie unter Verwendung der Ungleichung von Chebyshev-Markov mit $\kappa = 2$ eine untere Schranke für die Wahrscheinlichkeit dafür an, daß X Werte zwischen 4 und 14 annimmt.

Aufgabe 6 Die Länge der in einer (vom Konkurs bedrohten) Fabrik hergestellten Schrauben sei normalverteilt mit $\mu = 3$ [cm] und $\sigma = 0.2$ [cm]. Eine Schraube wird als unbrauchbar betrachtet, falls sie länger als 3.1 cm oder kürzer als 2.8 cm ist. Wie groß ist die Wahrscheinlichkeit, daß eine produzierte Schraube dem Ausschuß zuzurechnen ist?

Aufgabe 7 Seien $\alpha, \beta \in \mathbb{R}$ mit $\alpha \neq \beta$.

a) Man gebe Bedingungen für die Parameter α und β an, so daß

$$f(x; \alpha, \beta) := \begin{cases} \frac{\alpha\beta}{\alpha-\beta}(e^{-\beta x} - e^{-\alpha x}), & x \in (0, \infty) \\ 0, & \text{sonst} \end{cases}$$

eine Dichtefunktion ist.

b) Man berechne den Erwartungswert $\mathbb{E}(X)$ einer Zufallsvariable X mit der Wahrscheinlichkeitsdichte $f(x; 2, 1)$.

Aufgabe 8 a) (X_1, X_2) sei $\mathcal{N}(\mu_1, \Sigma_1)$ -verteilt mit

$$\Sigma_1 = \begin{pmatrix} 5 & 0 \\ 0 & 3 \end{pmatrix}.$$

Sind X_1 und X_2 unkorreliert? Sind sie unabhängig?

b) (X_3, X_4) sei $\mathcal{N}(\mu_2, \Sigma_2)$ -verteilt mit

$$\Sigma_2 = \begin{pmatrix} 5 & 1 \\ 1 & 3 \end{pmatrix}.$$

Sind X_3 und X_4 unkorreliert? Sind sie unabhängig?

Statistik

Aufgabe 9 In Baden-Württemberg wurden in den Jahren 1996-1999 218 674 Mädchen und 231 432 Jungen geboren. Entscheiden Sie mit Hilfe des zentralen Grenzwertsatzes, ob diese Zahlen die Hypothese "beide Geschlechter sind gleichwahrscheinlich" unterstützen.

Aufgabe 10 Finden Sie geeignete statistische Modelle für die folgenden Zufallsexperimente:

a) Eine zufällige diskrete Zeit $x \in \mathbb{N}$ wird beobachtet (z. B. eine Wartezeit), die als geometrisch- (π) -verteilt angenommen wird mit einem unbekannten Wert $\pi \in]0, 1[$.

b) Jetzt wird das Zufallsexperiment in a) $n \geq 2$ mal unabhängig durchgeführt, also z.B. Wartezeiten in n gleichartigen Situationen beobachtet (wobei die Unabhängigkeit unterstellt wird).

Aufgabe 11 Zwei verschiedene Fertigungsmethoden sollen hinsichtlich einer quantitativen Produkt-Charakteristik verglichen werden. Aus der Fertigung unter Methode 1 wird eine Stichprobe vom Umfang n_1 entnommen und aus der Fertigung unter Methode 2 eine Stichprobe vom Umfang n_2 . Die beobachteten Werte der Produkt-Charakteristik seien mit $x_1, \dots, x_{n_1}$ und $y_1, \dots, y_{n_2}$ bezeichnet. Die Schwankungen der Werte innerhalb der beiden Stichproben werden auf Zufallseinflüsse zurückgeführt. Geben Sie ein geeignetes Normalverteilungsmodell an.

Statistik

Aufgabe 12 Stellen Sie sich vor, ein Unternehmen muss Insolvenz anmelden und bietet allen Beschäftigten an, unter Zahlung einer Abfindung freiwillig auszuscheiden oder nicht (mit einem gewissen Risiko, dass später - dann allerdings ohne Zahlung einer Abfindung - gekündigt wird). Nehmen wir an, dass jeder Beschäftigte sich mit derselben Wahrscheinlichkeit π (jedoch unabhängig von den anderen) für das freiwillige Ausscheiden entscheidet.

Zuerst wird eine kleinere Anzahl von n Beschäftigten vor die Entscheidung gestellt, um Aussagen über den Wert des unbekannten Parameters π treffen zu können.

a) Die einzelnen Entscheidungen $x_1, \dots, x_n$ können als Werte von n unabhängigen $B(1, \pi)$ identisch verteilten Zufallsvariablen $X_1, \dots, X_n$ aufgefasst werden, wobei $\pi \in]0, 1[$ der unbekannte Parameter sei. Berechnen Sie die Likelihoodfunktion für $n = 10$ auf Grund der Beobachtungen $x_1 = 1, x_2 = 0, x_3 = 0, x_4 = 0, x_5 = 1, x_6 = 0, x_7 = 1, x_8 = 0, x_9 = 0, x_{10} = 1$.

b) Betrachten Sie die Zufallsvariable $Y = \sum_{i=1}^n X_i$. Wie wir wissen, ist Y $B(n, \pi)$ -verteilt. Also können wir $y = \sum_{i=1}^n x_i$ als beobachteten Wert von Y auffassen. $\pi \in]0, 1[$ ist der unbekannte Parameter. Berechnen Sie die Likelihoodfunktion in diesem Modell.

c) Welcher Wert von π ist auf Grund der Beobachtungen plausibler: 0.3 oder 0.5?

Aufgabe 13 Die Anzahl der Verkehrstoten pro Woche in einem Landkreis sei Poisson- (λ) -verteilt. Für die letzten 15 Wochen ergaben sich folgende Anzahlen:

0 2 0 1 1 0 0 4 0 0 1 2 1 0 0.

Schätzen Sie den unbekannten Parameter λ der Anzahl von Verkehrstoten pro Woche in dem Landkreis mittels Maximum-Likelihood-Methode.

Bitte wenden

Aufgabe 14 Die Fachschaft möchte vor den bevorstehenden Hochschulwahlen die voraussichtliche Wahlbeteiligung schätzen. Auf dem Weg zur bzw. von der Mensa befragt jeder der sechs Mitarbeiter solange andere Informatik-Studenten, bis er jemanden trifft, der angibt, nicht zur Wahl zu gehen.

X bezeichne die Anzahl der Studenten, die überprüft wurden, bis ein potenzieller Nichtwähler gefunden wurde. Eine sinnvolle Verteilung von X ist beispielsweise die geometrische Verteilung mit dem Parameter $1 - \vartheta$, d. h.

$$P_{\vartheta}(X = i) = \vartheta^{i-1} \cdot (1 - \vartheta), \quad i \in \mathbb{N}.$$

Die unbekannte Größe ϑ kann dabei als die gesuchte prozentuale Wahlbeteiligung interpretiert werden.

Man bestimme auf der Basis der ermittelten Stichprobenwerte

3 1 2 3 2 1 2 1 5 3 2 3

einen Maximum-Likelihood-Schätzwert für ϑ .

Statistik

Aufgabe 15 In der Qualitätsabteilung eines Glühlampenherstellers wurde folgender Versuch durchgeführt: 100 Glühlampen desselben Typs wurden zufällig aus der Tagesproduktion entnommen und die Lebensdauern [in h] wurden ermittelt. Das arithmetische Mittel der beobachteten Lebensdauern betrug 5488 h.

a) Wie lautet das statistische Modell, wenn Sie die Lebensdauer einer Glühlampe dieses Typs als exponential- (λ) -verteilt annehmen?

b) Berechnen Sie eine ML-Schätzung für den unbekannten Parameter λ .

Aufgabe 16 Die Zufallsvariablen $X_1, \dots, X_n$ seien u. i. v. $\mathcal{N}(\mu, 1)$ -normalverteilt mit Erwartungswert μ und bekannter Varianz $\sigma^2 = 1$. Es liegen die Beobachtungsdaten (Realisierungen) $x_1, \dots, x_n$ von $X_1, \dots, X_n$ vor. Geben Sie eine ML-Schätzung für den unbekannten Parameter μ in Abhängigkeit von $x_1, \dots, x_n$ an.

Aufgabe 17 Die Zufallsvariablen $X_1, \dots, X_n$ seien u. i. v. gleichverteilt über dem Intervall $[0, \vartheta]$. Gegeben seien die Realisierungen $x_1, \dots, x_n$ von $X_1, \dots, X_n$. Berechnen Sie eine ML-Schätzung für den unbekannten Parameter ϑ in Abhängigkeit von $x_1, \dots, x_n$.

Aufgabe 18 In der folgenden Tabelle sind für sechs verschiedene Monate jeweils der Hypothekenzinssatz (t_i) [in %] und das Volumen (y_i) [in Mrd.] eingehender Aufträge im privaten Hausbau dargestellt:

Monat i	t_i	y_i
1	6	3.0
2	5	3.2
3	7	2.5
4	7	2.3
5	8	2.0
6	9	2.0

a) Bestimmen Sie auf Grund der Daten eine Regressionsgerade für das Auftragsvolumen in Abhängigkeit vom Hypothekenzinssatz.

b) Berechnen Sie die *Residual Sum of Squares*.

c) Treffen Sie basierend auf der Regressionsgeraden eine Vorhersage über das Auftragsvolumen, wenn der Hypothekenzinssatz 7.5% beträgt.

Statistik

Aufgabe 19 10 zufällig ausgewählte Konjunkturexperten haben für das nächste Jahr unabhängig voneinander folgende Vorhersagen für das prozentuale Wachstum des Bruttosozialprodukts gemacht:

2.2, 2.8, 3.0, 2.5, 2.4, 2.6, 2.5, 2.4, 2.7, 2.6.

Wir nehmen an, dies sei eine Realisierung im normalverteilten statistischen u. i. v. Modell $((X_1, \dots, X_{10}), \vartheta = (\mu, \sigma^2) \in \mathbb{R} \times \mathbb{R}^{>0}, P_{(\mu, \sigma^2), X_i} \sim \mathcal{N}(\mu, \sigma^2))$. Geben Sie erwartungstreue Schätzungen für die folgenden reellen Parameter an:

- a) $\gamma(\vartheta) = \mu$,
- b) $\gamma(\vartheta) = \sigma^2$,
- c) $\gamma(\vartheta) = \text{Var}_{\vartheta}(\bar{X})$ (Varianz des Stichprobenmittels).

Aufgabe 20 Sei $((X_1, \dots, X_n), \vartheta \in \Theta, P_{\vartheta, X})$ ein statistisches u. i. v. Modell. Der Wert des Parameters $\mu(\vartheta) = E_{\vartheta}(X_1)$ sei bekannt.

- a) Betrachten Sie den folgenden Schätzer für $\sigma^2(\vartheta) = \text{Var}_{\vartheta}(X_1)$:

$$S^* = \frac{1}{n} \sum_{i=1}^n (X_i - \mu(\vartheta))^2.$$

Zeigen Sie die Erwartungstreue des Schätzers.

- b) Zeigen Sie, dass der Schätzer

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$$

für $\sigma^2(\vartheta) = \text{Var}_{\vartheta}(X_1)$ nicht erwartungstreu ist.

Bitte wenden

Aufgabe 21 Die Preise für einen Liter Normalbenzin an den Tankstellen einer großen Stadt werden als unabhängig und identisch normal- (μ, σ_0^2) -verteilt angenommen, wobei $\sigma_0^2 = 0.1$ bekannt sei.

Zwei unabhängige Stichproben vom Umfang m bzw. n für die Preise seien verfügbar. Die Stichprobenmittel $\bar{x}$ bzw. $\bar{y}$ sollen zu einer einzigen Schätzung für den erwarteten Preis μ pro Liter kombiniert werden.

a) Wählen Sie ein geeignetes statistisches Modell und zeigen Sie, dass die folgenden Schätzer erwartungstreu für den Parameter μ sind:

$$\hat{\mu}_1 = \hat{\mu}_1(X_1, \dots, X_m, Y_1, \dots, Y_n) = \frac{1}{2}(\bar{X} + \bar{Y}),$$

$$\hat{\mu}_2 = \hat{\mu}_2(X_1, \dots, X_m, Y_1, \dots, Y_n) = \frac{1}{m+n}(m\bar{X} + n\bar{Y}),$$

b) Berechnen Sie die Varianzen der beiden Schätzer. Welcher ist effizienter?

c) Berechnen Sie konkrete Schätzungen für μ auf Grund der Beobachtungsdaten $\bar{x} = 0.969$ ($m = 10$) und $\bar{y} = 0.975$ ($n = 6$).

Aufgabe 22 Die Zufallsvariablen $X_1, \dots, X_{2n}$ seien stochastisch unabhängig und auf $[0, \lambda]$ gleichverteilte Zufallsvariablen mit unbekanntem Parameter $\lambda > 0$. Betrachten Sie das u. i. v. Gleichverteilungsmodell $(X_1, \dots, X_{2n}, \lambda \in \mathbb{R}^{>0}, P_{\lambda, X_i})$.

(a) Man zeige, dass $S := \frac{1}{n}(X_1 + \dots + X_{2n})$ ein erwartungstreuer Schätzer für λ ist.

(b) Ein weiterer erwartungstreuer Schätzer T hat die Varianz $\text{Var}_\lambda(T) = \lambda^2$. Für welchen Schätzer entscheiden Sie sich? Begründen Sie Ihre Antwort.

Statistik

Aufgabe 23 Seien $(X_1, \dots, X_{10}, \vartheta \in \Theta, P_{\vartheta, X_i})$ ein statistisches u. i. v. Modell und

$$\gamma : \Theta \rightarrow \mathbb{R}, \quad \vartheta \mapsto \mu(\vartheta) := E_{\vartheta}(X_1).$$

Wir betrachten den Schätzer $\hat{\mu}(X_1, \dots, X_{10}) = \bar{X}$ für $\gamma(\vartheta)$. Die folgenden Werte seien Realisierungen der Stichprobenvariablen:

3, 7, 6, 2, 6, 5, 4, 9, 7, 4.

Berechnen Sie den *estimated standard error* $ese(\hat{\mu}^*)$ für die angegebene Stichprobe.

Aufgabe 24 Sie gehen an zehn aufeinander folgenden Tagen zum Bäcker, um Semmeln zu kaufen. Dabei zählen Sie jeweils die Kunden, die vor Ihnen anstehen und erhalten die folgenden Ergebnisse:

3, 7, 6, 2, 6, 5, 4, 9, 7, 4.

Wir nehmen an, dass dies eine Realisierung im statistischen u. i. v. Poisson-Modell $(X_1, \dots, X_{10}, \lambda \in \mathbb{R}^{>0}, P_{\lambda, X_i} \sim Poi(\lambda))$ ist. Für den unbekannten Parameter $\lambda = E_{\lambda}(X_1)$ betrachten wir den Schätzer $\hat{\lambda}(X) = \bar{X}$. Berechnen Sie $ese(\hat{\lambda}^*)$.

Aufgabe 25 Eine Fabrik stellt Fernseher her. Sie wollen die mittlere Lebensdauer eines Fernsehers ermitteln und untersuchen deshalb 10 Fernseher. Als mittlere Lebensdauer dieser Stichprobe erhalten Sie 5.3 [Jahre]. Wir nehmen an, dies sei das Stichprobenmittel einer Realisierung im statistischen u. i. v. Exponentialverteilungs-Modell $(X_1, \dots, X_{10}, \lambda \in \mathbb{R}^{>0}, P_{\lambda, X_i} \sim Exp(\lambda))$. Wir betrachten den Schätzer $\hat{\mu}(X) = \bar{X}$ für $E_{\lambda}(X_1)$. Berechnen Sie $ese(\hat{\mu}^*)$.

Aufgabe 26 Gegeben sei das lineare Regressionsmodell $(Y_1, \dots, Y_n, (a, c, \sigma^2) \in \mathbb{R}^2 \times \mathbb{R}^{>0}, P_{\vartheta, Y_i} \sim \mathcal{N}(y(t_i), \sigma^2))$, $Y_1, \dots, Y_n$ unabhängig und $y(t_i) = at_i + c$, $i = 1, \dots, n$.

a) Zeigen Sie, dass die ML-Schätzer $\hat{a}$ für a und $\hat{c}$ für c erwartungstreu sind.

Bitte wenden

b) Nun betrachten wir die Situation aus Aufgabe 18: In der folgenden Tabelle sind für sechs verschiedene Monate jeweils der Hypothekenzinssatz (t_i) [in %] und das Volumen (y_i) [in Mrd.] eingehender Aufträge im privaten Hausbau dargestellt:

Monat i	t_i	y_i
1	6	3.0
2	5	3.2
3	7	2.5
4	7	2.3
5	8	2.0
6	9	2.0

Berechnen Sie $ese(\hat{a}^*)$, $ese(\hat{c}^*)$ und $ese(\hat{y}(t_0)^*)$ für $t_0 = 7.5$.

Statistik

Aufgabe 27 Es sei Z eine exponentialverteilte Zufallsvariable, $P_Z \sim \text{Exp}(\lambda)$. Sei $\alpha \in]0, 1[$.

a) Geben Sie für $p \in]0, 1[$ das p -Quantil z_p an.

b) Geben Sie p_1 an, so dass gilt

$$P(z_{\frac{\alpha}{2}} < Z \leq p_1) = 1 - \alpha.$$

c) Geben Sie p_2, p_3 an, so dass gilt

$$P(Z \leq p_2) = 1 - \alpha, \quad P(p_3 < Z) = 1 - \alpha.$$

Aufgabe 28 Die Größe der an einem Rechnernetz per FTP übertragenen Dateien werde jeweils als $\mathcal{N}(\mu, \sigma^2)$ -normalverteilte Zufallsvariable angenommen mit unbekanntem Erwartungswert μ . Man gebe für die Stichprobe $x_1, \dots, x_{16}$ mit

144, 136, 156, 147, 143, 180, 176, 121, 137, 155, 191, 162, 133, 148, 140, 179

ein konkretes 95%-Konfidenzintervall für μ an.

Aufgabe 29 Ein neues Medikament gegen das HIV-Virus soll hinsichtlich seiner Wirksamkeit getestet werden. Je mehr T-Helferzellen sich nach der einwöchigen Behandlung im Blut befinden, desto besser ist das Medikament. Für 15 Patienten wurde jeweils die Anzahl der T-Helferzellen pro Mikroliter Blut vor (y_i) und nach (z_i) einwöchiger Behandlung mit dem neuen Medikament festgestellt ($i = 1, 2, \dots, 15$):

y_i	415	507	389	677	560	359	402	500	461	534	540	390	457	576	460
z_i	441	490	423	698	550	394	472	510	451	581	495	421	424	580	490

a) Berechnen Sie unter der Annahme, dass $y_1, \dots, y_{15}$ Werte von u. i. v. normal- (μ_1, σ_1^2) -verteilten Zufallsvariablen sind, 95%-Konfidenzintervalle für μ_1 und σ_1^2 .

b) Berechnen Sie unter der Annahme, dass $z_1, \dots, z_{15}$ Werte von u. i. v. normal- (μ_2, σ_2^2) -verteilten Zufallsvariablen sind, 95%-Konfidenzintervalle für μ_2 und σ_2^2 .

Bitte wenden

Aufgabe 30 Auf dem Bogen zur "Evaluation der Lehre HT03" findet sich folgender Satz, der zum Ausfüllen des Bogens auffordert: "Eine statistisch relevante Aussage [...] ergibt sich nur bei ausreichender Beteiligung".

Um die Aussage des Satzes zu präzisieren, betrachten wir speziell eine Frage, bei der Sie 5 Alternativen für Ihre Antwort haben. Diese Alternativen interpretieren wir als Noten auf einer Notenskala von 1 (sehr gut) bis 5 (mangelhaft).

Wir nehmen an, dass n Studenten unabhängig voneinander die Frage beantworten. Die Wahl einer Note sei für jeden Studenten identisch $\text{normal}(\mu, \sigma^2)$ -verteilt. Unabhängig vom Umfang n der Stichprobe seien das Stichprobenmittel $\bar{x} = 3.2$ und die Stichprobenvarianz $s^2 = 1.7$. Geben Sie für jeden der folgenden Werte von n ein 90%-Konfidenzintervall für den unbekannten Erwartungswert μ an:

$$n = 5, \quad 10, \quad 20, \quad 41.$$

Statistik

Aufgabe 31 Ein Großhändler hat eine große Lieferung eines bestimmten Produktes erhalten. 300 Stück werden zufällig aus der Lieferung ausgewählt und kontrolliert. Davon waren $k = 4$ mangelhaft.

a) Bestimmen Sie ein (approximatives) 95%-Konfidenzintervall für den wahren Anteil der mangelhaften Stücke in der Lieferung.

b) Sei n die Anzahl der kontrollierten Stücke. Wieviele Stücke müssten mindestens kontrolliert werden, wenn gesichert werden soll, dass die Länge des Konfidenzintervalls höchstens 0.01 beträgt? Wir nehmen an, dass k/n konstant ist mit $k/n = 1/75$.

c) Antworten Sie, ohne die Berechnung noch einmal durchzuführen: Wenn ein höheres Konfidenzniveau gesichert werden soll, ist dann das entsprechende Konfidenzintervall kleiner oder größer als das in a) berechnete?

Aufgabe 32 In der Qualitätsabteilung eines Glühlampenherstellers wurde folgender Versuch durchgeführt: 50 Glühlampen desselben Typs wurden zufällig aus der Tagesproduktion entnommen und die Lebensdauern [in h] wurden ermittelt. Das arithmetische Mittel der beobachteten Lebensdauern betrug 5488 h. Wir nehmen an, die Lebensdauer einer Glühlampe sei exponentialverteilt mit Parameter λ und interpretieren die Entnahme der 50 Glühlampen als Realisierung im u. i. v. Exponentialverteilungs-Modell.

a) Geben Sie ein 90%-Konfidenzintervall für den Erwartungswert $\mu(\lambda)$ der Lebensdauer einer Glühlampe an.

b) Geben Sie eine untere 90%-Konfidenzschranke und eine obere 90%-Konfidenzschranke für $\mu(\lambda)$ an.

Aufgabe 33 Die Längen der in einem industriellen Prozess produzierten Metallstäbe sind unabhängig und normal- (μ, σ^2) -verteilt mit (bekannter) Varianz $\sigma^2 = 3.24$. Von Interesse ist auf Grund einer Stichprobe vom Umfang n die Konstruktion eines 99%-Konfidenzintervalls der Intervalllänge 1[mm] für den Parameter μ .

Im u. i. v. (μ, σ^2) -Normalverteilungsmodell mit bekannter Varianz σ^2 ist ein $(1 - \alpha)$ -Intervallschätzer $\hat{I}$ für μ gegeben durch

$$\hat{I} = [\bar{X} - \frac{u_{1-\frac{\alpha}{2}}}{\sqrt{n}}\sigma, \bar{X} + \frac{u_{1-\frac{\alpha}{2}}}{\sqrt{n}}\sigma].$$

Dabei bezeichnet u_p für $p \in (0, 1)$ das p -Quantil der Standardnormalverteilung. Wie groß ist n zu wählen, um ein 99%-Konfidenzintervall zu erhalten? Ist dieses Problem lösbar, wenn σ^2 unbekannt ist?

Statistik

Aufgabe 34 Ein neues Medikament mit unbekannter Erfolgswahrscheinlichkeit π werde an 10 Patienten unabhängig voneinander getestet. Ein zugelassenes Medikament habe eine Erfolgswahrscheinlichkeit von 60%. Es soll entschieden werden, ob π größer als 0.6 ist. Es soll also zwischen den beiden Alternativen

$$H_0 : \pi \leq 0.6 \text{ und } H_1 : \pi > 0.6$$

entschieden werden.

Sei Y der Anteil der Patienten, bei denen das Medikament erfolgreich wirkt. Betrachten Sie folgende Entscheidungsregel: "Entscheidung für H_0 , falls $Y \leq 0.6$, ansonsten für H_1 ". Wie groß ist die Wahrscheinlichkeit, sich für H_0 zu entscheiden, wenn die (unbekannte) Erfolgswahrscheinlichkeit bei 40%, 60% bzw. 80% liegt?

Aufgabe 35 Ein Unternehmen möchte die Qualität einer großen Lieferung von Bauelementen untersuchen. Die Lieferung kann aus Kosten- und Imagegründen nur akzeptiert werden, wenn der Anteil π der mangelhaften Teile in der Lieferung höchstens 20% beträgt (Nullhypothese); andernfalls wird die Lieferung zurückgewiesen. Da aus Zeit- und Kostengründen natürlich nicht jedes einzelne Bauteil überprüft werden kann, wird die Entscheidung wie folgt getroffen:

Eine Stichprobe von $n = 10$ Bauteilen (zufällig mit Zurücklegen aus der Lieferung ausgewählt) wird auf Mängel untersucht und die Anzahl x der mangelhaften Teile wird festgestellt. Die Lieferung wird zurückgewiesen, falls $x > 2$; andernfalls wird die Lieferung akzeptiert.

- a) Formulieren Sie eine Nullhypothese H_0 , eine Gegenhypothese H_1 und eine Entscheidungsfunktion in der Sprache der Statistik.
- b) Beschreiben Sie verbal die Fehler 1. Art und 2. Art. Wie groß ist die Wahrscheinlichkeit, sich für H_0 zu entscheiden, wenn die Lieferung genau 10% (20%, 30%) mangelhafter Teile aufweist? Welches Signifikanzniveau hat der Test?
- c) Die Stichprobe enthalte $x = 5$ defekte Teile. Ist damit der statistische Nachweis für das Vorliegen einer mangelhaften Lieferung erbracht?
- d) Das Unternehmen möchte aus o. g. Gründen möglichst die Akzeptanz einer "schlechten" Lieferung verhindern. Deshalb wird eine Testprozedur benötigt, die sicher stellt, dass eine "schlechte" Lieferung nur mit einer Wahrscheinlichkeit von höchstens 5% angenommen wird. Leistet obige Testprozedur das? Formulieren Sie ggf. ein anderes Testproblem.

Statistik

Aufgabe 36 Ein Zigarettenhersteller wirbt damit, dass seine neue Light-Zigarette "im Durchschnitt nicht mehr als 4 mg Kondensat (Teer) enthält". Um dies zu überprüfen, wurde eine Stichprobe von $n = 25$ Zigaretten untersucht. Es ergab sich ein mittlerer Teergehalt von $\bar{x} = 4.08$ mg und $s = 0.3$ mg. Die Teergehalte werden allgemein als unabhängig und identisch normalverteilt angenommen.

a) Zeigen die Daten signifikant (zum Niveau $\alpha = 0.05$), dass die Aussage des Zigarettenherstellers falsch ist? Ein Statistikprogramm gibt für diesen Test auf Grund der vorliegenden Daten einen p -value von 0.098 aus. Was heißt das? Sind die Daten signifikant zum Niveau $\alpha = 0.1$?

b) Wie sind Null- und Alternativhypothese zu wählen, wenn herausgefunden werden soll, ob die Daten signifikant zeigen (zum Niveau $\alpha = 0.05$), dass die Aussage des Zigarettenherstellers korrekt ist?

Aufgabe 37 Die Verhältnisse $X_1, \dots, X_n$ von Monatsschlussstand zu Monatsanfangsstand des DAX für die Monate $i = 1, \dots, n$ werden als unabhängig und identisch logarithmisch-normal- (μ, σ^2) -verteilt angenommen, d. h. die transformierten Zufallsvariablen $Y_1, \dots, Y_n$ (wobei $Y_i := \ln(X_i)$ für $i = 1, \dots, n$) werden als identisch normal- (μ, σ^2) -verteilt angenommen.

Angenommen, Sie wollen in den DAX investieren. Auf Grund einiger Berechnungen haben Sie jedoch entschieden, dass dies nur in Frage kommt, falls der Wert des unbekannten Parameters μ größer ist als -0.02.

Für die 12 Monate des Jahres 2002 ergaben sich die folgenden Werte $x_1, \dots, x_{12}$ der Zufallsvariablen $X_1, \dots, X_{12}$:

1.056, 0.914, 0.939, 1.075, 0.977, 0.989, 0.956, 0.896, 0.830, 1.058, 1.094, 1.034.

Testen Sie zum Niveau $\alpha = 0.05$:

$$H_0 : \mu \leq -0.02 \text{ gegen } H_1 : \mu > -0.02.$$

Aufgabe 38 Benutzen Sie einen "guten" Test (aus der Vorlesung), um in der Situation von Aufgabe 35 auf Grund der Beobachtung $x = 1$ zum Niveau 0.05 zu testen:

$$H_0 : \pi \geq 0.2 \text{ gegen } H_1 : \pi < 0.2.$$

Bitte wenden

Aufgabe 39 Wir betrachten die Daten aus Aufgabe 29: Für 15 Patienten wurde jeweils die Anzahl der T-Helferzellen pro Mikroliter Blut vor (y_i) und nach (z_i) einwöchiger Behandlung mit einem neuen Medikament gegen HIV festgestellt ($i = 1, 2, \dots, 15$):

y_i	415	507	389	677	560	359	402	500	461	534	540	390	457	576	460
z_i	441	490	423	698	550	394	472	510	451	581	495	421	424	580	490

Die Daten werden auch dem Kassenverband zur Verfügung gestellt. Dort will man natürlich das Medikament nur zahlen, wenn es wirklich wirkt. Wir nehmen an, dass $y_1, \dots, y_{15}$ Werte von u. i. v. normal- (μ_1, σ_1^2) -verteilten Zufallsvariablen sind und $z_1, \dots, z_{15}$ Werte von u. i. v. normal- (μ_2, σ_2^2) -verteilten Zufallsvariablen. Zeigen die Daten signifikant zum Niveau $\alpha = 0.1$, dass das Medikament wirkt, dass also $\mu_1 < \mu_2$ gilt?

Statistik

Aufgabe 40 Betrachten Sie das statistische Modell mit einer hypergeometrisch- (N, s, n) -verteilten Zufallsvariablen X , wobei N und n bekannt sind ($N, n \in \mathbb{N}, n < N$) und $s \in \{0, 1, \dots, N\}$ der unbekannte Parameter ist. Betrachten Sie als Schätzer für s :

$$\hat{s}(X) = \frac{N}{n}X.$$

a) Zeigen Sie die Erwartungstreue des Schätzers und leiten Sie die folgende Formel für seine Varianz her:

$$\text{Var}_s(\hat{s}) = \frac{(N-n)s(N-s)}{(N-1)n}, \quad \text{für alle } s.$$

b) Angenommen $N = 20$, $n = 10$ und der Wert $x = 4$ wird beobachtet. Geben Sie die Schätzung an, die $\hat{s}$ liefert und bestimmen Sie $\text{ese}(\hat{s}^*)$.

Aufgabe 41 X_1 und X_2 seien unabhängige und in $(0, \lambda)$ gleichverteilte Zufallsvariablen mit unbekanntem $\lambda > 0$.

a) Man zeige, dass $S := X_1 + X_2$ und $T := 3 \cdot \min\{X_1, X_2\}$ erwartungstreue Schätzer für λ sind.

b) Wenn Sie die Wahl haben zwischen S und T , für welchen Schätzer entscheiden Sie sich? Begründen Sie Ihre Entscheidung.

Aufgabe 42 Eine statistische Datenerhebung hat $n = 30$ reelle Daten $x_1, \dots, x_{30}$ ergeben mit $\bar{x} = 9.95$ und $s = 1.64$. Die Daten $x_1, \dots, x_{30}$ werden als Werte von u.i.v. normal- (μ, σ^2) -verteilten Zufallsvariablen aufgefasst, wobei μ und σ^2 unbekannte Parameter sind.

a) Berechnen Sie eine untere 99%-Konfidenzschranke für μ .

b) Berechnen Sie eine obere 99%-Konfidenzschranke für σ^2 .

c) Liefern die Daten Signifikanz zum 1%-Niveau für die Alternativhypothese H_1 des Testproblems

$$H_0 : \mu \geq 10.5 \text{ gegen } H_1 : \mu < 10.5?$$

Bitte wenden

Aufgabe 43 Betrachten Sie das statistische u.i.v. Modell $((X_1, \dots, X_n), \vartheta \in (0, \infty), P_{\vartheta, X_i})$ wobei die Dichtefunktion von P_{ϑ, X_i} gegeben ist durch

$$f_{\vartheta} : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto \begin{cases} \vartheta x^{\vartheta-1}, & \text{falls } 0 < x < 1 \\ 0, & \text{sonst} \end{cases}.$$

Es liegen Beobachtungsdaten $x_1, \dots, x_n \in (0, 1)$ vor. Stellen Sie die Loglikelihood-Funktion auf und geben Sie die ML-Schätzung $\hat{\vartheta}^*$ an.

Aufgabe 44 Eine Datenerhebung hat die folgenden Lebensdauern von 12 gleichartigen Glühlampen ergeben (aufgerundet auf volle [h]):

2888, 849, 213, 1580, 1560, 1774, 225, 2752, 722, 517, 2025, 791.

Wir unterstellen das statistische u.i.v. Modell mit 12 geometrisch(π)-verteilten Zufallsvariablen. Berechnen Sie die ML-Schätzung für π und schätzen Sie auch die mittlere Lebensdauer $\mu(\pi) = E_{\pi}(X_i)$ sowie die Varianz $\sigma^2(\pi) = Var_{\pi}(X_i)$.

Aufgabe 45 Für die folgenden $n = 8$ Beobachtungsdaten unterstellen wir das statistische u.i.v. Normalverteilungsmodell $((X_1, \dots, X_8), (\mu, \sigma^2) \in \mathbb{R} \times \mathbb{R}^{>0}, P_{(\mu, \sigma^2), X_i} \sim \mathcal{N}(\mu, \sigma^2))$:

1.641, 4.845, 3.251, 3.497, 2.371, 3.39, 2.714, 4.368.

a) Ergeben die Beobachtungsdaten eine Signifikanz zum 5%-Niveau für die Alternativhypothese H_1 des Testproblems

$$H_0 : \mu \leq 2 \text{ gegen } H_1 : \mu > 2?$$

b) Berechnen Sie ein 95%-Konfidenzintervall für σ^2 .

Aufgabe 46 Betrachten Sie das statistische u.i.v. Normalverteilungsmodell $((X_1, \dots, X_n), (0, \sigma^2) \in \{0\} \times \mathbb{R}^{>0}, P_{(0, \sigma^2), X_i} \sim \mathcal{N}(0, \sigma^2))$, wobei $\sigma^2 > 0$ der unbekannte Parameter ist. Angenommen, es liegen Beobachtungswerte $x_1, \dots, x_n \in \mathbb{R}$ vor. Stellen Sie die Loglikelihood-Funktion auf und leiten Sie eine Formel für die ML-Schätzung $\hat{\sigma}^{2*}$ her.

Aufgabe 47 In einer oberbayerischen Flaschenabfüllfabrik wird regelmäßig überprüft, ob die mittlere Füllung $\mu_0 = 500[cm^3]$ beträgt. Zur Überprüfung werden 20 Flaschen untersucht. Wir interpretieren dies als Stichprobe im u.i.v. Normalverteilungsmodell. Es gelte $\bar{x} = 498.39$ und $s^2 = 11.7$. Entwerfen Sie einen Signifikanz-Test mit Signifikanz-Niveau $\alpha = 0.05$ bzw. $\alpha = 0.01$, um festzustellen, wann die Anlage neu eingestellt werden muss. Wird die Nullhypothese abgelehnt?

Kurztest Wahrscheinlichkeitstheorie

1. Für einen Wahrscheinlichkeitsraum $(\Omega, \mathcal{S}, P)$ gilt $P(\Omega) = \square$.
2. Für zwei disjunkte Ereignisse A, B gilt $P(A \square B) = P(A) \square P(B)$.
3. Für zwei stochastisch unabhängige Ereignisse A, B gilt $P(A \square B) = P(A) \square P(B)$.
4. Für zwei beliebige Ereignisse A, B gilt $P(A \setminus B) = P(A) - P(B)$. Richtig?
 Ja/Nein.
5. Seien $(\Omega, \mathcal{S}, P)$ ein Wahrscheinlichkeitsraum, $(M, \mathcal{S}')$ ein Meßraum und $X : \Omega \rightarrow M$ eine Zufallsvariable. Für $C \in \mathcal{S}'$ bezeichnet

$$\{X \in C\} := \{\omega \in \Omega : \square\},$$

und dies ist ein Ereignis, d. h. ein Element von $\mathcal{S}$.

6. Die Wahrscheinlichkeit für drei Erfolge ist binomial-verteilt. Sinnvoll? Ja/Nein.
7. Gegeben seien N Kugeln, darunter genau s rote. Dann ist die Anzahl der roten Kugeln in einer Stichprobe vom Umfang n beim zufälligen Ziehen
 - a) mit Zurücklegen $\square$ - ($\square$)-verteilt und
 - a) ohne Zurücklegen $\square$ - ($\square$)-verteilt.
8. Seien $(\Omega, \mathcal{S}, P)$ ein Wahrscheinlichkeitsraum, $(M, \mathcal{S}')$, $(N, \mathcal{S}'')$ Meßräume und $X : \Omega \rightarrow M$, $Y : \Omega \rightarrow N$ Zufallsvariablen.
 - a) X und Y heißen stochastisch unabhängig, wenn für alle $\square$ und $\square$ gilt

$$P(X \square, Y \square) = P(X \square) \square P(Y \square)$$

- b) Sind X und Y absolutstetige reelle Zufallsvariablen mit Dichte f_1 bzw. f_2 , so gilt für die gemeinsame Dichte f

$$f : \square \rightarrow \square, \quad \square \mapsto \square.$$

9. Der Wert der Indikatorvariablen $I_A(x)$ ist genau dann gleich 1, wenn $\square$,
 und sonst $\square$.
10. Die bedingte Wahrscheinlichkeit ist definiert als

$$P(A|B) := \frac{\square}{\square}, \text{ falls } \square.$$

11. Welche Namen haben die folgenden Wahrscheinlichkeitsverteilungen?

a) $P(X = k) = \binom{n}{k} \pi^k (1 - \pi)^{n-k}$: - () - Verteilung,

b) $P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}$: - () - Verteilung.

12. Für eine stetig verteilte reelle Zufallsvariable X mit Verteilungsfunktion F gelten

a) $F(x) = P(\text{})$,

b) $P(X = x) = \text{}$ und

c) $P(a < X \leq b) = F(\text{}) - F(\text{})$.

13. Für eine reelle Zufallsvariable X mit Erwartungswert $E(X)$ und Varianz $Var(X)$ gilt

$$Var(X) = \text{} - \text{}.$$

14. Für eine stetig verteilte nichtnegative Zufallsvariable X mit Verteilungsfunktion F gelten

a) $E(X) = \int_0^\infty \text{} dx$,

b) $E(X^2) = \int_0^\infty \text{} dx$.

15. Für eine absolut-stetige reelle Zufallsvariable X mit Dichtefunktion f gelten

a) $E(X) = \int_{-\infty}^\infty \text{} dx$,

b) $E(X^2) = \int_{-\infty}^\infty \text{} dx$.

16. Für eine absolut-stetige reelle Zufallsvariable X mit Dichtefunktion f und Verteilungsfunktion F gelten

a) $\int_{-\infty}^\infty f(y) dy = \text{}$,

b) $\int_{-\infty}^x f(y) dy = \text{}$,

c) $\int_a^b f(y) dy = \text{}$,

17. Welche Namen haben die Wahrscheinlichkeitsverteilungen mit den folgenden Dichtefunktionen?

a) $f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$: - () - Verteilung,

b) $f(x) = \lambda \cdot e^{-\lambda x} I_{(0,\infty)}(x)$: - () - Verteilung.

Statistik

Aufgabe 1 Es sei $\Omega = \{1, 2, 3, 4\}$ und

$$\mathcal{S} = \{\emptyset, \{3\}, \{4\}, \{1, 2\}, \{3, 4\}, \{1, 2, 3\}, \{1, 2, 4\}, \Omega\}.$$

- a) Zeigen Sie, daß $\mathcal{S}$ eine σ -Algebra ist.
- b) Geben Sie ein Erzeugendensystem von $\mathcal{S}$ an.

Aufgabe 2 Sie holen einen Freund vom Flughafen ab. Erfahrungsgemäß ist das Flugzeug in 70% aller Fälle verspätet, falls es regnet, und in 20% aller Fälle, falls es nicht regnet. Der Wetterbericht prophezeit eine Regenwahrscheinlichkeit von 40%. Wie groß ist die Wahrscheinlichkeit, daß das Flugzeug Verspätung hat? Wie groß ist die Wahrscheinlichkeit, daß es regnet, falls das Flugzeug Verspätung hat?

Aufgabe 3 Drei Kugeln $K1, K2, K3$ werden auf vier Urnen $U1, U2, U3, U4$ verteilt.

- a) Geben Sie aufgrund des Prinzips der Gleichwahrscheinlichkeit einen geeigneten Wahrscheinlichkeitsraum $(\Omega, \mathcal{S}, P)$ für dieses Zufallsexperiment an.
- b) Bestimmen Sie die Wahrscheinlichkeit der folgenden Ereignisse:

$$A = "K1 \text{ liegt in } U3", \quad B = "In \text{ jeder Urne liegt höchstens eine Kugel"}.$$

Aufgabe 4 Auf einem gegebenen Wahrscheinlichkeitsraum $(\Omega, \mathcal{S}, P)$ soll die Zufallsvariable

$$X = (X_n)_{1 \leq n \leq N} : \Omega \rightarrow \Omega' := \{0, 1, 2\}^N \quad (1)$$

das Zufallsexperiment "Tagesproduktion von N Taschenrechnern" beschreiben, X_n beschreibt dabei wiefolgt den Zustand von Taschenrechner $n, 1 \leq n \leq N$:

$$X_n(\omega) = \begin{cases} 0, & \text{Taschenrechner defekt,} \\ 1, & \text{Taschenrechner ok, Gehäuse defekt,} \\ 2, & \text{Taschenrechner und Gehäuse ok.} \end{cases} \quad (2)$$

Drücken Sie mit Hilfe der Zufallsvariablen X_n folgende Ereignisse aus:

1. Der n -te Taschenrechner ist voll intakt.
2. Kein Taschenrechner ist ok.
3. Mindestens ein Rechner ist nicht ok oder am Gehäuse beschädigt.

Aufgabe 5 Die Zufallsvariable X sei Poisson-verteilt mit Parameter $\lambda = 9$. Geben Sie unter Verwendung der Ungleichung von Chebyshev-Markov mit $\kappa = 2$ eine untere Schranke für die Wahrscheinlichkeit dafür an, daß X Werte zwischen 4 und 14 annimmt.

Aufgabe 6 Die Länge der in einer (vom Konkurs bedrohten) Fabrik hergestellten Schrauben sei normalverteilt mit $\mu = 3$ [cm] und $\sigma = 0.2$ [cm]. Eine Schraube wird als unbrauchbar betrachtet, falls sie länger als 3.1 cm oder kürzer als 2.8 cm ist. Wie groß ist die Wahrscheinlichkeit, daß eine produzierte Schraube dem Ausschuß zuzurechnen ist?

Aufgabe 7 Seien $\alpha, \beta \in \mathbb{R}$ mit $\alpha \neq \beta$.

a) Man gebe Bedingungen für die Parameter α und β an, so daß

$$f(x; \alpha, \beta) := \begin{cases} \frac{\alpha\beta}{\alpha-\beta}(e^{-\beta x} - e^{-\alpha x}), & x \in (0, \infty) \\ 0, & \text{sonst} \end{cases}$$

eine Dichtefunktion ist.

b) Man berechne den Erwartungswert $\mathbb{E}(X)$ einer Zufallsvariable X mit der Wahrscheinlichkeitsdichte $f(x; 2, 1)$.

Aufgabe 8 a) (X_1, X_2) sei $\mathcal{N}(\mu_1, \Sigma_1)$ -verteilt mit

$$\Sigma_1 = \begin{pmatrix} 5 & 0 \\ 0 & 3 \end{pmatrix}.$$

Sind X_1 und X_2 unkorreliert? Sind sie unabhängig?

b) (X_3, X_4) sei $\mathcal{N}(\mu_2, \Sigma_2)$ -verteilt mit

$$\Sigma_2 = \begin{pmatrix} 5 & 1 \\ 1 & 3 \end{pmatrix}.$$

Sind X_3 und X_4 unkorreliert? Sind sie unabhängig?