KARLSRUHER INSTITUT FÜR TECHNOLOGIE

Lehrstuhl für Ökonometrie und Statistik

S T A T I S T I K II

Lösungsblätter zu den Übungen

Wintersemester 2013/2014

Prof. Dr. Wolf-Dieter Heller
Frieder Conrad
Hartwig Senska

Institut für Volkswirtschaftslehre (ECON)
D-76128 Karlsruhe

Aufgabe 1

Gegeben seien eine standardnormalverteilte Zufallsvariable X sowie eine Funktion $g : \mathbb{R} \to \mathbb{R}$ mit:

$$g(x) = |x| + 4 \ , \ x \in \mathbb{R}.$$

(a) Bestimmen Sie $P(g(X) \in [a, b])$, $a, b \in \mathbb{R}$, mit $a < b$.

(b) Berechnen Sie $P(g(X) \leq 4.4)$.

(c) Berechnen Sie $E(g(X))$

Lösung: (Wahrscheinlichkeitstheorie S.84ff)

(a) $g(x)$ ist nicht invertierbar, weshalb die Transformationsformel für Dichtefunktionen nicht anwendbar ist.

Es gilt:

$$P(g(X) \in [a, b]) = F_{g(X)}(b) - F_{g(X)}(a)$$

und für $F_{g(X)}(u)$ mit $u \geq 4$ ergibt sich:

$$\begin{aligned} F_{g(X)}(u) &= \int\limits_{g(x) \leq u} f_X(x)\, dx = \int\limits_{|x|+4 \leq u} f_X(x)\, dx \\ &= \int\limits_{4-u}^{u-4} f_X(x)\, dx = F_X(u-4) - F_X(4-u) \end{aligned}$$

Für $a, b \in \mathbb{R}$ mit $4 \leq a < b$ ergibt sich somit:

$$\begin{aligned} P(g(X) \in [a, b]) &= F_{g(X)}(b) - F_{g(X)}(a) \\ &= F_X(b-4) - F_X(4-b) - F_X(a-4) + F_X(4-a) \\ &= F_X(b-4) - F_X(a-4) + (F_X(4-a) - F_X(4-b)) \\ &= P(X \in [a-4, b-4]) + P(X \in [4-b, 4-a]) \end{aligned}$$

Beispiel: $P(g(X) \in [6, 8]) = P(X \in [2, 4]) + P(X \in [-4, -2])$

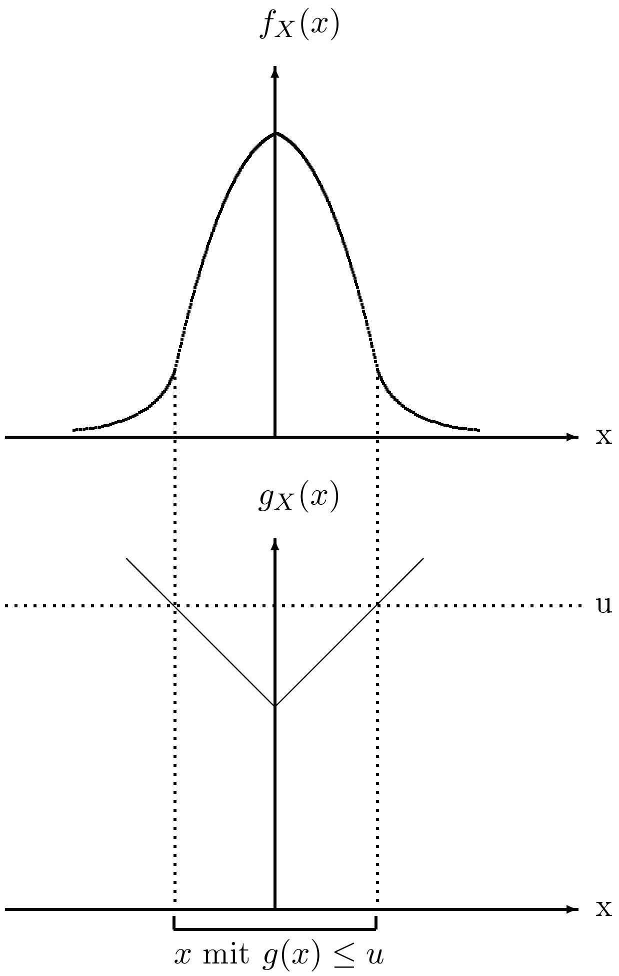


Abbildung 1: Veranschaulichung der Transformation

(b)

$$
\begin{aligned}
P(g(X) \leq 4.4) &= \int\limits_{g(x)\leq 4.4} f_X(x)\,dx = \int\limits_{-0.4}^{0.4} f_X(x)\,dx \\
&= 2\Phi(0.4) - 1 = 0.3108
\end{aligned}
$$

(c) Für den Erwartungswert von $g(X)$ ergibt sich:

$$
\begin{aligned}
E(g(X)) &= \int_{\mathbb{R}} g(x) \cdot f_X(x)\,dx = \int_{\mathbb{R}} (|x| + 4) \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\,dx \\
&= \frac{2}{\sqrt{2\pi}} \int\limits_0^{\infty} x \cdot e^{-\frac{x^2}{2}}\,dx + 4 \\
&= \sqrt{\frac{2}{\pi}} (-e^{-\frac{x^2}{2}}) \Big|_0^{\infty} + 4 = \sqrt{\frac{2}{\pi}} + 4
\end{aligned}
$$

Aufgabe 2

X sei eine reellwertige Zufallsvariable, deren Verteilung durch die folgende von einem Parameter $\gamma > 0$ abhängende Dichtefunktion $f_\gamma(x)$ bestimmt ist:

$$f_\gamma(x) = \begin{cases} \frac{1}{\gamma} \cdot (1-x)^{\left(\frac{1}{\gamma}-1\right)} & \text{für } 0 < x < 1 \\ \\ 0 & \text{sonst} \end{cases},$$

(a) Skizzieren Sie den Verlauf der Dichtefunktion $f_\gamma(x)$ für die Fälle $\gamma = 0.5$ und $\gamma = 2$.

(b) Berechnen Sie die Verteilungsfunktion $F_\gamma(t)$ der Zufallsvariablen X.

(c) Berechnen Sie für beliebiges $n \in \mathbb{N}$ folgende Momente:

$$E(1-X)^n,\ E(X),\ Var(1-X),\ Var(X)$$

(d) Berechnen Sie eine Dichte der Zufallsvariablen $Y := -\ln(1-X)$

Lösung: (Wahrscheinlichkeitstheorie, S. 64ff, 69ff, 91ff und 109 ff)

(a) Für $\gamma = 2$, 0.5 ergeben sich folgende Dichtefunktionen:

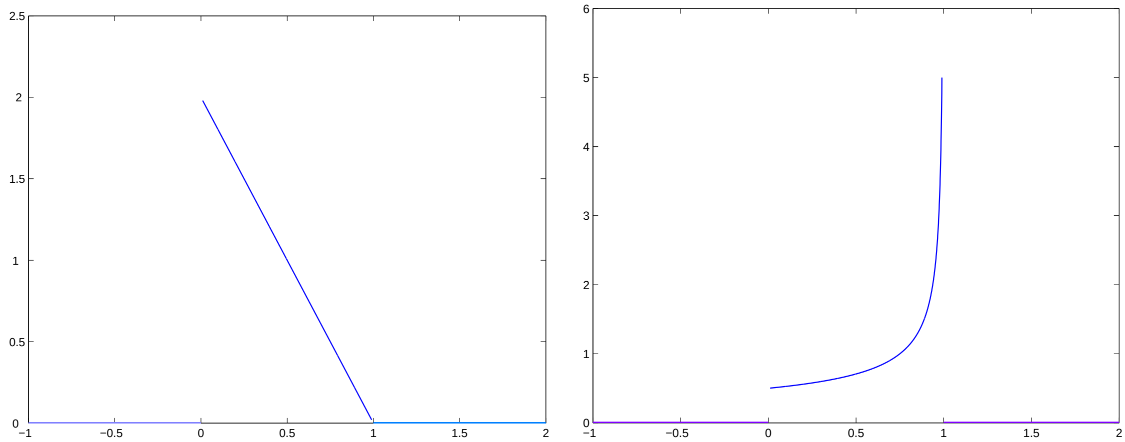
$$f_{0.5}(x) = \begin{cases} -2x+2 & \text{für } 0 < x < 1 \\ 0 & \text{sonst} \end{cases},$$

$$f_2(x) = \begin{cases} \frac{1}{2\sqrt{(1-x)}} & \text{für } 0 < x < 1 \\ 0 & \text{sonst} \end{cases}$$

(b) Für die Verteilungsfunktion $F_\gamma(t)$ gilt:

$$F_\gamma(t) = \int_{-\infty}^{t} f_\gamma(x)dx = \begin{cases} 1 & \text{für } t \geq 1 \\ \\ \int_0^t \frac{1}{\gamma} \cdot (1-x)^{\left(\frac{1}{\gamma}-1\right)} dx & \text{für } 0 < t < 1 \\ \\ 0 & \text{für } t \leq 0 \end{cases}$$

$$= \begin{cases} 1 & \text{für } t \geq 1 \\ \\ 1-(1-t)^{\frac{1}{\gamma}} & \text{für } 0 < t < 1 \\ \\ 0 & \text{für } t \leq 0 \end{cases}.$$



(a) $f_{0.5}(x) = -2x + 2 \cdot \mathbf{1}_{(0,1)(x)}$ (b) $f_2(x) = \frac{1}{2\sqrt{1-x}} \cdot \mathbf{1}_{(0,1)(x)}$

Abbildung 2: Skizzen der Dichtefuntionen

(c)

$$\begin{aligned}
E\left((1-X)^n\right) &= \int_{-\infty}^{\infty} (1-x)^n f_\gamma(x)\,dx = \int_0^1 \frac{1}{\gamma}(1-x)^{(n+\frac{1}{\gamma}-1)}\,dx \\
&= -\frac{1}{\gamma n+1}(1-x)^{(n+\frac{1}{\gamma})}\Big|_0^1 = \frac{1}{n\gamma+1} \\
E(X) &= -(E(1-X)) + 1 \overset{s.o.}{=} \frac{\gamma}{1+\gamma} \\
Var(1-X) &= E((1-X)^2) - (E(1-X))^2 \overset{s.o.}{=} \frac{1}{2\gamma+1} - \frac{1}{(\gamma+1)^2} \\
Var(X) &= 0 + (-1)^2 Var(X) = Var(1) + Var(-X) = Var(1-X)
\end{aligned}$$

(d) Für den Wertebereich von Y gilt, da $X \in (0,1)$, dass $Y \in (0,\infty)$; desweiteren ist die Transformation bijektiv. Damit erhält man bereits für die Dichte $g_Y(y) = 0$ für $y \leq 0$.

Alternative 1: Man berechnet die Verteilungsfunktion $G_Y(t)$ und erhält dann eine Dichte $g_Y(y)$ durch Differentiation:

$$\begin{aligned}
G_Y(t) &= P(Y \leq t) = P(-\ln(1-X) \leq t) = P(\ln(1-X) \geq -t) \\
&= P(1-X \geq e^{-t}) = P(X \leq 1-e^{-t}) = F_X(1-e^{-t}) \\
&= 1-(1-(1-e^{-t}))^{\frac{1}{\gamma}} = 1-e^{-\frac{t}{\gamma}} \Rightarrow \\
g_Y(y) &= \frac{1}{\gamma}e^{-\frac{y}{\gamma}} \text{ für } y > 0
\end{aligned}$$

Alternative 2: Transformationsformel mit

$$g_Y(y) = \frac{f_X(T^{-1}(y))}{|T'(T^{-1}(y))|}$$

Aufgabe 3

Eine Zufallsvariable X mit $0 < X < 1$ habe die Verteilungsfunktion

$$F_X(x) \quad := \quad \begin{cases} 0 & \text{für } x \leq 0 \\ 1 - e^{-\frac{\lambda x}{1-x}} & \text{für } 0 < x < 1 \\ 1 & \text{für } x \geq 1 \end{cases} ,$$

mit festem Parameter $\lambda > 0$.

(a) Berechnen Sie die Dichte f_X zu X.

(b) Es sei $Y := T(X)$ die Zufallsvariable, die aus X durch die Transformation

$$T : x \mapsto \frac{x}{1-x}$$

hervorgeht. Bestimmen Sie den Wertebereich von Y und geben Sie eine Dichte g_Y zu Y an. Wie heißt die Verteilung der Zufallsvariablen Y?

(c) Bestimmen Sie

$$E\left(\frac{X}{(1-X)^2}\right)$$

Hinweis: Verwenden Sie die Zerlegung $X = X(1-X) + X^2$

Lösung: (Wahrscheinlichkeitstheorie, S. 64ff, 69ff, 91ff und 109 ff)

(a) Die Dichtefunktion berechnet sich als Ableitung der Verteilungsfunktion und somit gilt für $0 < x < 1$:

$$f_X(x) = F_X^{'}(x) = e^{-\frac{\lambda x}{1-x}} \cdot \frac{\lambda(1-x) + \lambda x}{(1-x)^2}$$

und es ergibt sich:

$$f_X(x) \quad = \quad \begin{cases} \frac{\lambda}{(1-x)^2} \cdot e^{-\frac{\lambda x}{1-x}} & \text{für } 0 < x < 1 \\ 0 & \text{sonst} \end{cases}$$

(b) Wertebereich von Y:

$$Y = \frac{X}{1-X} = \frac{-(1-X)+1}{1-X} = -1 + \underbrace{\frac{1}{1-X}}_{\in(1,\infty)} \in (0, \infty)$$

Berechnung der Verteilung von Y:

$$\begin{aligned}
F_Y(t) &= P(Y \leq t) = P(\frac{X}{1-X} \leq t) \\
&= P(X \leq t(1-X)) = P((1+t)X \leq t) \\
&= P(X \leq \frac{t}{1+t}) = F_X(\frac{t}{1+t}) = \begin{cases} 0 & \text{für } \frac{t}{1+t} \leq 0 \\ 1 - e^{-\frac{\lambda \frac{t}{1+t}}{1-\frac{t}{1+t}}} & \text{für } 0 < \frac{t}{1+t} < 1 \\ 1 & \text{für } \frac{t}{1+t} \geq 1 \end{cases} \\
&= \begin{cases} 0 & \text{für } t \leq 0 \\ 1 - e^{-\lambda t} & \text{für } 0 < t \end{cases}
\end{aligned}$$

Die Dichte zu Y lautet damit

$$f_Y(t) = \begin{cases} \lambda e^{-\lambda t} & \text{für } t \geq 0 \\ 0 & \text{sonst} \end{cases}$$

Alternativ kann auch die Transformationsformel angewendet werden und führt zu den gleichen Ergebnissen.

Y ist exponentialverteilt mit Parameter λ.

(c) Mit Hilfe des Hinweises ergibt sich:

$$\begin{aligned}
E\left(\frac{X}{(1-X)^2}\right) &= E\left(\frac{X}{1-X}\right) + E\left(\left(\frac{X}{1-X}\right)^2\right) \\
&= E(Y) + E(Y^2) = \frac{1}{\lambda} + \frac{2}{\lambda^2}
\end{aligned}$$

Aufgabe 4

(a) Die Zufallsvariable X sei standardnormalverteilt. Berechnen Sie für $\mu \in \mathbb{R}$ und $\sigma \in \mathbb{R}_{>0}$ die Dichtefunktion der Zufallvariablen Y mit $Y := \exp(\sigma X + \mu)$.

(b) Berechnen Sie den Modalwert und den Median der Zufallsvariablen Y.

(c) Berechnen Sie den Erwartungswert von Y.

(d) Skizzieren Sie die Dichtefunktion von Y für eine geeignete Wahl der Parameter μ, σ und tragen Sie die berechneten Lageparameter im Schaubild ein.

Bemerkung: Die Verteilung von Y heißt Lognormalverteilung $LN(\mu, \sigma)$ und spielt in der Finanzmathematik eine bedeutende Rolle.

Lösung: (Wahrscheinlichkeitstheorie, S. 64ff, 69ff, 91ff und 109 ff)

(a) Für die Transformation T gilt:

$$\begin{aligned} T \ : \ & \mathbb{R} \to (0, \infty) \\ & x \mapsto e^{\sigma x + \mu} =: y \end{aligned}$$

und für die Ableitung bzw. Umkehrfunktion ergibt sich:

$$T'(x) = \sigma e^{\sigma x + \mu}, \quad T^{-1}(y) = \frac{1}{\sigma}(\ln y - \mu)$$

Hieraus ergibt sich mit Hilfe der Transformationsformel für Dichten:

$$\begin{aligned} g_Y(y) &= \frac{f_X(T^{-1}(y))}{|T'(T^{-1}(y))|} = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(\ln y - \mu)^2} \cdot \frac{1}{\sigma e^{\ln y}} \cdot \mathbf{1}_{(0,\infty)}(y) \\ &= \frac{1}{\sqrt{2\pi}\sigma y} e^{-\frac{1}{2\sigma^2}(\ln y - \mu)^2} \cdot \mathbf{1}_{(0,\infty)}(y) \end{aligned}$$

(b) Der Modalwert entspricht der Maximalstelle der Dichtefunktion. Da $g(y)$ für $y \to 0$ und $y \to \infty$ und $g_Y(y)$ stetig verschwindet, ist die notwendige Bedingung auch hinreichend:

$$\begin{aligned} 0 &\overset{!}{=} g_y'(y) \\ &= -\frac{1}{y^2} \cdot \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2}(\ln y - \mu)^2} + \frac{1}{\sqrt{2\pi}\sigma y} e^{-\frac{1}{2\sigma^2}(\ln y - \mu)^2} \cdot \left(-\frac{1}{\sigma^2}\right)(\ln y - \mu) \cdot \frac{1}{y} \\ \Leftrightarrow \quad & \frac{1}{\sigma^2}(\ln y - \mu) + 1 = 0 \\ \Leftrightarrow \quad & y = e^{\mu - \sigma^2}, \end{aligned}$$

d.h. der Modalwert lautet: $y_M = e^{\mu - \sigma^2}$.

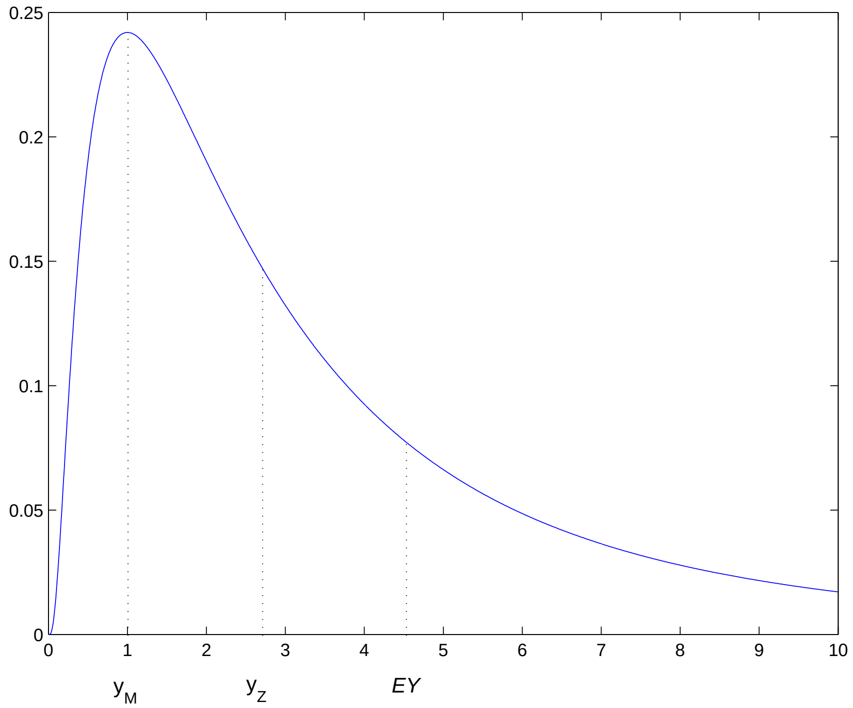


Abbildung 3: Dichtefunktion der Lognormalverteilung mit Parametern $\mu = 1$, $\sigma = 1$

Der Median von Y entspricht dem 0.5-Quantil der Verteilung von Y und lässt sich wie folgt aus der Verteilung von X berechnen:

$$\begin{aligned} P(Y \leq y) &= P(e^{\sigma X+\mu} \leq y) = P(X \leq \frac{1}{\sigma}(\log(y) - \mu)) \stackrel{!}{=} \frac{1}{2} \\ \stackrel{X\sim N(0,1)}{\Longrightarrow} & \quad \frac{1}{\sigma}(\ln y - \mu) \stackrel{!}{=} 0 \Leftrightarrow \ln y = \mu \\ \Rightarrow & \quad y = e^{\mu}, \end{aligned}$$

d.h. der Median lautet: $y_Z = e^{\mu}$.

(c) Der Erwartungswert EY lässt sich prinzipiell auf zwei Arten berechnen:

$$EY = \int_{\mathbb{R}} y \cdot g_Y(y)\, dy \quad \text{und} \quad EY = E(e^{\sigma X+\mu}) = \int_{\mathbb{R}} e^{\sigma x+\mu} \cdot f_X(x)\, dx$$

Hier wählen wir die zweite Variante:

$$\begin{aligned} E(Y) &= \int_{\mathbb{R}} e^{\sigma x+\mu} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\, dx \;=\; \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}+\sigma x+\mu}\, dx \\ &= \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}[(x-\sigma)^2-\sigma^2-2\mu]}\, dx \;=\; \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x-\sigma)^2+\mu+\frac{\sigma^2}{2}}\, dx \\ &= e^{\mu+\frac{\sigma^2}{2}} \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x-\sigma)^2}\, dx \;=\; e^{\mu+\frac{\sigma^2}{2}}, \end{aligned}$$

d.h. der Erwartungswert lautet: $EY = e^{\mu+\frac{\sigma^2}{2}}$.

Aufgabe 5

In einer automatischen Überwachungsanlage besteht $\Omega \subset \mathbb{R}^k$ aus Messwert-Kombinationen, die an k Kontrollinstrumenten prinzipiell beobachtbar sind. Nach Experteneinschätzung gilt ein Bereich $S \subset \Omega$ als kritisch ("Störfall"). Eine Modellrechnung liefert eine Wahrscheinlichkeit $P(S) = 2 \cdot 10^{-4}$. Eine weitere Zuverlässigkeitsanalyse liefert die Wahrscheinlichkeiten $P(A|S) = 0.95$ für Alarm, wenn ein Störfall vorliegt (Entdeckungswahrscheinlichkeit), und außerdem $P(A^c|S^c) = 0.99$ für Nichtalarm im unkritischen Bereich.

Für das technische System sind zu berechnen:

(a) die Wahrscheinlichkeit $P(A)$ für Alarm,

(b) die Wahrscheinlichkeit $P(S^c|A)$, dass ein Alarm ein Fehlalarm ist,

(c) die Wahrscheinlichkeit $P(S|A^c)$ für einen unentdeckten Störfall.

Lösung: (Wahrscheinlichkeitstheorie S. 91ff)

Nach Definition gilt:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}.$$

Daraus ergibt sich:

$$P(A \cap B) = P(A|B) \cdot P(B) = P(B|A) \cdot P(A)$$

Seien $A_1, \ldots, A_n$ und B Ereignisse (Hypothesen oder Zustände), mit

$$\bigcup_{i=1}^{n} A_i = \Omega, A_i \cap A_j = \emptyset \text{ für } i \neq j$$

dann gelten die folgenden beiden Aussagen:

Satz von der totalen Wahrscheinlichkeit:

$$\begin{aligned} P(B) &= \sum_{i=1}^{n} P(B|A_i) \cdot P(A_i) \\ &= \sum_{i=1}^{n} P(B \cap A_i) \end{aligned}$$

Satz von Bayes: Die Wahrscheinlichkeit für die Hypothese A_1 genügt bei Eintreten von B folgender Gleichung:

$$P(A_1|B) = \frac{P(A_1 \cap B)}{P(B)} = \frac{P(B|A_1) \cdot P(A_1)}{\sum_{i=1}^{n} P(B|A_i) \cdot P(A_i)}$$

Als Anwendung in der Aufgabe wählen wir:

- Hypothesen: Störfall(S)/kein Störfall(S^c)
- Beobachtungen: Alarm (A)/ kein Alarm (A^c)

Es ergibt sich:

$$
\begin{aligned}
\text{(a) } P(A) &= P(A|S)\cdot P(S) + P(A|S^c)\cdot P(S^c) \text{ (Satz von der totalen Wahrscheinlichkeit)}\\
&= 0.95\cdot 2\cdot 10^{-4} + 0.01\cdot(1-2\cdot 10^{-4}) = 0.01019
\end{aligned}
$$

$$
\begin{aligned}
\text{(b) } P(S^c|A) &= \frac{P(S^c\cap A)}{P(A)} = \frac{P(A|S^c)\cdot P(S^c)}{P(A)} = \frac{0.01\cdot(1-2\cdot 10^{-4})}{0.01019}\\
&= 0.9813
\end{aligned}
$$

$$
\begin{aligned}
\text{(c) } P(S|A^c) &= \frac{P(S\cap A^c)}{P(A^c)} = \frac{P(A^c|S)\cdot P(S)}{P(A^c)} = \frac{0.05\cdot 2\cdot 10^{-4}}{1-0.01019}\\
&= 0.0000101
\end{aligned}
$$

Aufgabe 6

(a) Bei einem Glückspiel werden 200 Kugeln (100 rote und 100 schwarze) gezielt auf zwei Schälchen verteilt, wobei keines der Schälchen leer sein darf. Anschließend wird mit verbundenen Augen zufällig ein Schälchen ausgewählt und hieraus rein zufällig eine Kugel gezogen.

Ist die gezogene Kugel schwarz hat man gewonnen, andernfalls verloren. Wie muss man die Kugeln aufteilen um die maximale Gewinnchance zu haben? Wie groß ist diese?

(b) Von einem Test zur Identifizierung einer ansteckenden Krankeit ist bekannt, dass der Test zu 99% bei infizierten Personen positiv (Infizierung) ausfällt, während er bei 98% der gesunden Personen negativ (keine Infizierung) ist. Ferner ist bekannt, dass 0.1% der Bevölkerung diese Krankheit hat.

Mit welcher Wahrscheinlichkeit ist man tatsächlich infiziert, wenn man ein positives Testergebnis hat?

Lösung: (Wahrscheinlichkeitstheorie S. 91ff)

(a) Es bezeichne n_A die Gesamtanzahl an Kugeln in Schälchen A und k die Anzahl an schwarzen Kugeln in Schälchen A. Die Anzahl an Kugeln in Schälchen B beträgt demnach $n_B = 200 - n_A$ und die Anzahl an schwarzen Kugeln in Schälchen B ist $100 - k$. Man stellt fest, dass für $n_A = 100$ die Gewinnwahrscheinlichkeit $P(G)$ unabhängig von k bei 50 Prozent liegt:

$$\begin{aligned} P(G) &= P(\text{schwarz} \mid A) \cdot P(A) + P(\text{schwarz} \mid B) \cdot P(B) \\ &= \frac{k}{100} \cdot 0.5 + \frac{100-k}{100} \cdot 0.5 = 0.5 \end{aligned}$$

Durch Umbenennung der Schälchen ist es im Falle $n_A \neq 100$ immer möglich, dass gilt: $n_A < 100$. In diesem Falle gilt für beliebiges aber festes n_A:

$$\begin{aligned} P(G) &= P(\text{schwarz} \mid A) \cdot P(A) + P(\text{schwarz} \mid B) \cdot P(B) \\ &= \frac{k}{n_A} \cdot 0.5 + \frac{100-k}{200-n_A} \cdot 0.5 \\ &= k \cdot \underbrace{\left(\frac{100-n_A}{n_A(200-n_A)} \right)}_{>0} + \frac{50}{200-n_A}, \quad k = 0, \ldots, n_A \end{aligned}$$

d.h die Gewinnwahrscheinlichkeit wird maximal falls gilt: $k = n_A$. Hierfür gilt wiederum:

$$\begin{aligned} P(G) &= P(\text{schwarz} \mid A) \cdot P(A) + P(\text{schwarz} \mid B) \cdot P(B) \\ &= \frac{k}{n_A} \cdot 0.5 + \frac{100-k}{200-n_A} \cdot 0.5 = \frac{1}{2} \cdot \left(1 + \frac{100-n_A}{200-n_A} \right) \\ &= \frac{1}{2} \cdot \left(2 - \frac{100}{200-n_A} \right), \end{aligned}$$

d.h die Gewinnwahrscheinlichkeit wird maximal falls gilt $n_A = 1$ und für die diese gilt dann $P(G) = \frac{149}{199} \approx 0.75$.

(b) Man berechnet analog zu Aufgabe 5 mit Hilfe des Satzes von der totalen Wahrscheinlichkeit:

$$\begin{aligned}
P(\text{positiv}) &= P(\text{positiv} \mid \text{krank}) \cdot P(\text{krank}) + P(\text{positiv} \mid \text{gesund}) \cdot P(\text{gesund}) \\
&= 0.99 \cdot 0.001 + 0.02 \cdot 0.999 = 0.02097 \\
P(\text{krank} \mid \text{positiv}) &= \frac{P(\text{krank} \cap \text{positiv})}{P(\text{positiv})} = \frac{P(\text{positiv} \mid \text{krank}) \cdot P(\text{krank})}{P(\text{positiv})} \\
&= \frac{0.99 \cdot 0.001}{0.02097} = 0.04721
\end{aligned}$$

Man muss sich somit keine allzu großen Sorgen machen, denn ein positives Testergebnis bedeutet nur in 5% aller Fälle, dass man wirklich infiziert ist.

Aufgabe 7

(a) Zwei Firmen stellen ein Produkt unter Verwendung von zwei Herstellungsverfahren her. Aus der Gesamtproduktion beider Firmen werde zufällig ein Teil entnommen. Der Zufallsvektor $X = (X_1, X_2, X_3)$ enthält folgende Angaben:

$$X_1 = \begin{cases} 1 & \text{Firma 1} \\ 2 & \text{Firma 2} \end{cases}$$

$$X_2 = \begin{cases} 1 & \text{Verfahren 1} \\ 2 & \text{Verfahren 2} \end{cases}$$

$$X_3 = \begin{cases} 0 & \text{Teil gut} \\ 1 & \text{Teil schlecht} \end{cases}$$

Die folgende Tabelle gibt die gemeinsame Wahrscheinlichkeitsverteilung an:

	Firma 1 ($X_1 = 1$)		Firma 2 ($X_1 = 2$)	
	Verfahren 1 ($X_2 = 1$)	Verfahren 2 ($X_2 = 2$)	Verfahren 1 ($X_2 = 1$)	Verfahren 2 ($X_2 = 2$)
Teil gut ($X_3 = 0$)	0.045	0.4	0.225	0.005
Teil schlecht ($X_3 = 1$)	0.005	0.05	0.225	0.045

Welches Verfahren schneidet insgesamt bzw. bei den einzelnen Firmen besser ab? Berechnen Sie zur Beantwortung dieser Frage die folgenden bedingten Wahrscheinlichkeiten:

$$\begin{aligned} P(X_3 = 1 \mid X_2 = 1) \quad &, \quad P(X_3 = 1 \mid X_2 = 2) \\ P(X_3 = 1 \mid X_1 = 1, X_2 = 1) \quad &, \quad P(X_3 = 1 \mid X_1 = 1, X_2 = 2) \\ P(X_3 = 1 \mid X_1 = 2, X_2 = 1) \quad &, \quad P(X_3 = 1 \mid X_1 = 2, X_2 = 2) \end{aligned}$$

Interpretieren Sie das Ergebnis!

(b) Die folgende Tabelle enthält das Gesamtbruttoeinkommen, sowie die daraus gezahlte Einkommenssteuer der Jahre 1974 und 1978 in den USA, aufgeschlüsselt nach verschiedenen Einkommensklassen:

Jahreseinkommen (pro Person in US$)	Einkommen (in 1000 US$)	gezahlte Steuer (in 1000 US$)
1974		
$< 5\,000$	41 651 643	2 244 467
$5\,000 - 9\,999$	146 400 740	13 646 348
$10\,000 - 14\,999$	192 688 922	21 449 597
$15\,000 - 99\,999$	470 010 790	75 038 230
$\geq 100\,000$	29 427 152	11 311 672
1978		
$< 5\,000$	19 879 622	689 318
$5\,000 - 9\,999$	122 853 315	8 819 461
$10\,000 - 14\,999$	171 858 024	17 155 758
$15\,000 - 99\,999$	865 037 814	137 860 951
$\geq 100\,000$	62 806 159	24 051 698

Modellieren Sie diesen Sachverhalt mit Hilfe eines Wahrscheinlichkeitsraums, wobei Sie als Grundraum Ω die Menge aller 1974 und 1978 verdienter Dollar wählen. Weiterhin bezeichnen $K_1, \ldots, K_5$ die Ereignisse, dass ein zufällig ausgewählter Dollar zu einer der fünf Einkommensklassen K_j gehört, B das Ereignis, dass der Dollar 1974 verdient wurde und A das Ereignis, dass der Dollar als Einkommenssteuer abgeführt wurde. Berechnen Sie die folgenden bedingten Wahrscheinlichkeiten:

$$\begin{array}{rcl} P(A \mid B) & , & P(A \mid B^c) \\ P(A \mid B, K_j) & , & j = 1 \ldots 5 \\ P(A \mid B^c, K_j) & , & j = 1 \ldots 5 \end{array}$$

Interpretieren Sie das Ergebnis!

Lösung: (Wahrscheinlichkeitstheorie S. 91ff)

(a)

$$\begin{aligned} P(X_3 = 1 \mid X_2 = 1) &= \frac{0.23}{0.5} = 0.46 \\ P(X_3 = 1 \mid X_2 = 2) &= \frac{0.095}{0.5} = 0.19 \\ P(X_3 = 1 \mid X_1 = 1, X_2 = 1) &= \frac{0.005}{0.05} = 0.1 \\ P(X_3 = 1 \mid X_1 = 1, X_2 = 2) &= \frac{0.05}{0.45} = 0.\bar{1} \\ P(X_3 = 1 \mid X_1 = 2, X_2 = 1) &= \frac{0.225}{0.45} = 0.5 \\ P(X_3 = 1 \mid X_1 = 2, X_2 = 2) &= \frac{0.045}{0.05} = 0.9 \end{aligned}$$

Interpretation: Verfahren 2 scheint insgesamt gesehen besser zu sein. Bei den beiden Firmen schneidet es aber jeweils schlechter ab. Grund ist, dass Firma 2, bei der das Verfahren 2 sehr schlecht ist, hauptsächlich Verfahren 1 einsetzt, während Firma 1, bei der die Verfahren ähnlich gut funktionieren, Verfahren 2 bevorzugt.

(b) Insgesamt wurden 1974 880 179 247 US$ verdient und davon 123 690 314 US$ als Einkommenssteuer abgeführt. 1978 wurden insgesamt 1 242 434 934 US$ verdient und hiervon wiederum 188 577 186 US$ abgeführt. Hiermit können die gesuchten Wahrscheinlichkeiten direkt aus der Tabelle abgelesen werden:

$$\begin{aligned} P(A \mid B) &= \frac{123\,690\,314}{880\,179\,247} = 0.141 \\ P(A \mid B^c) &= \frac{188\,577\,186}{1\,242\,434\,934} = 0.152 \end{aligned}$$

Für die Steuersätze in den verschiedenen Einkommensklassen 1974 ergibt sich

$$P(A \mid B, K_j) = \begin{cases} 0.054 & j = 1 \\ 0.093 & j = 2 \\ 0.111 & j = 3 \\ 0.160 & j = 4 \\ 0.384 & j = 5 \end{cases}$$

während man für 1978 erhält:

$$P(A \mid B^c, K_j) = \begin{cases} 0.035 & j = 1 \\ 0.072 & j = 2 \\ 0.100 & j = 3 \\ 0.159 & j = 4 \\ 0.383 & j = 5 \end{cases}$$

Man erkennt, dass von 1974 nach 1978 die Einkommenssteuersätze in jeder Einkommensklasse gefallen sind, dass aber trotzdem die Durchschnittssteuerbelastung von 14.1% auf 15.2% gestiegen ist.

Bemerkung: Es sei $K_1, \ldots, K_n$ eine disjunkte Zerlegung eines Grundraums Ω und A, B zwei beliebige Eregnisse. Gelten die folgenden beiden Ungleichungen, so wird dieser Sachverhalt als Simpson-Paradoxon bezeichnet:

$$\begin{aligned} P(A \mid B, K_j) &> P(A \mid B^c, K_j) \quad \text{für } j = 1, \ldots, n \\ P(A \mid B) &< P(A \mid B^c) \end{aligned}$$

Aufgabe 8

Betrachtet wird eine Maschine mit exponentialverteilter Lebensdauer ($T \sim Exp(\lambda)$).

(a) Wie groß ist die Wahrscheinlichkeit für eine Mindestlebensdauer von $a \in \mathbb{R}_+$ Zeiteinheiten? Wie ändert sich diese Wahrscheinlichkeit für wachsende Parameterwerte λ?

(b) Berechnen und interpretieren Sie die bedingte Wahrscheinlichkeit $P([0, t_0 + a] | [t_0, \infty))$ mit $a \geq 0$. Vergleichen Sie Ihr Ergebnis mit dem aus (a).

(c) Wie lautet die Dichte einer auf $[0, a]$ bzw. $[a, \infty]$ gestutzten Exponentialverteilung? Was fällt Ihnen bei der letztgenannten auf?

Lösung: (Wahrscheinlichkeitstheorie, S.64ff und S. 109ff)

Exponentialverteilung:

$$f_T(x) \;=\; \begin{cases} 0 & , \quad x < 0 \\ \lambda e^{-\lambda x} & , \quad x \geq 0 \end{cases}$$

$$F_T(x) \;=\; \begin{cases} 0 & , \quad x < 0 \\ 1 - e^{-\lambda x} & , \quad x \geq 0 \end{cases}$$

(a)

$$\begin{aligned} P(T \geq a) &= 1 - P(T \leq a) \;=\; 1 - (1 - e^{-\lambda a}) \\ &= e^{-\lambda a} \end{aligned}$$

Die Wahrscheinlichkeit nimmt mit wachsendem λ ab.

(b)

$$P_T([0, t_0 + a]) \,|\, [t_0; \infty]) = P(T \leq t_0 + a \,|\, T \geq t_0) = \frac{P(T \leq t_0 + a \text{ und } T \geq t_0)}{P(T \geq t_0)}$$

$$\begin{aligned} &= \frac{P(T \in [t_0, t_0 + a])}{P(T \geq t_0} = \frac{F(t_0 + a) - F(t_0)}{1 - F(t_0)} \\ &= \frac{1 - e^{-\lambda t_0 - \lambda a} - 1 + e^{-\lambda t_0}}{e^{-\lambda t_0}} = 1 - e^{-\lambda a} \\ &= P(T \leq a) \end{aligned}$$

Man erkennt, dass die Wahrscheinlichkeit, dass die Maschine in den nächsten a Zeiteinheiten ausfällt, wenn sie bis zu einem Zeitpunkt t_0 überlebt hat, mit der Wahrscheinlichkeit, dass die Maschine in den ersten a Zeiteinheiten nach Inbetriebnahme ausfällt, übereinstimmt. Aus diesem Grunde bezeichnet man die Exponentialverteilung auch als "Verteilung ohne Gedächtnis".

(c) Auf ein Intervall I gestutzte Verteilung: Bedingte Verteilung einer Zufallsvariable X, unter der Bedingung $X \in I$, d.h. für $I = [a, b]$ ist

$$\tilde{F}(x) = P(X \leq x \,|\, X \in I) = \frac{P(X \leq x \text{ und } x \in I)}{P(X \in I)} \quad = \quad \begin{cases} P(\emptyset) & , \quad x < a \\ \dfrac{P(a \leq X \leq x)}{P(X \in I)} & , \quad a \leq x \leq b \\ \dfrac{P(X \in I)}{P(X \in I)} = 1 & , \quad b < x \end{cases}$$

bzw.

$$\tilde{F}(x) = P(X \leq x \,|\, X \in I) \quad = \quad \begin{cases} 0 & , \quad x < a \\ \dfrac{F(x) - F(a)}{F(b) - F(a)} & , \quad x \in [a, b] \\ 1 & , \quad \text{sonst} \end{cases}$$

oder

$$\tilde{f}(x) \quad = \quad \begin{cases} \dfrac{f(x)}{F(b) - F(a)} & , \quad x \in [a, b] \\ 0 & , \quad \text{sonst} \end{cases}$$

Mit den Angaben im Aufgabentext ergibt sich nun konkret:

$$\tilde{f}_{[0,a]}(t) \quad = \quad \begin{cases} \dfrac{\lambda e^{-\lambda t}}{1 - e^{-\lambda a}} & , \quad t \in [0, a] \\ 0 & , \quad \text{sonst} \end{cases} \quad \text{und damit } \tilde{F}_{[0,a]}(t) \quad = \quad \begin{cases} 0 & , \quad t < 0 \\ \dfrac{1 - e^{-\lambda t}}{1 - e^{-\lambda a}} & , \quad t \in [0, a] \\ 1 & , \quad t > a \end{cases}$$

$$\tilde{f}_{[a,\infty)}(t) \quad = \quad \begin{cases} \lambda e^{-\lambda(t-a)} & , \quad t \geq a \\ 0 & , \quad \text{sonst} \end{cases} \quad \text{und damit } \tilde{F}_{[a,\infty]}(t) \quad = \quad \begin{cases} 0 & , \quad t < a \\ 1 - e^{-\lambda(t-a)} & , \quad t \geq a \end{cases}$$

$\tilde{f}_{[a,\infty)}(t)$ entspricht einer um a nach rechts verschobenen Dichtefunktion einer Exponentialverteilung.

Aufgabe 9

Bestimmen Sie für eine auf ein symmetrisches Intervall $[-c, c]$ gestutzte Normalverteilung $N(0; 1)$ die Dichtefunktion φ und die Verteilungsfunktion Φ, und stellen Sie diese graphisch dar.

Lösung: (Wahrscheinlichkeitstheorie S.91ff)

Es bezeichne $\varphi(x) := \frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}}$ die Dichte der Standardnormalverteilung. Dann ergibt sich analog zu Aufgabe 8 (c) die Dichtefunktion der gestutzen Normalverteilung:

$$\tilde{\varphi}_{[-c,c]}(x) = \begin{cases} \dfrac{\varphi(x)}{2\Phi(c) - 1} & , \quad x \in [-c, c] \\ 0 & , \quad \text{sonst} \end{cases}$$

bzw. die Verteilungsfunktion der gestutzten Normalverteilung:

$$\tilde{\Phi}_{[-c,c]}(x) = \int\limits_{-\infty}^{x} \tilde{\varphi}_{[-c,c]}(t)dt = \begin{cases} 0 & , \quad x < -c \\ \dfrac{\Phi(x) - \Phi(-c)}{2\Phi(c) - 1} & , \quad x \in [-c, c] \\ 1 & , \quad \text{sonst} \end{cases}$$

Aufgabe 10

Zur grafischen Darstellung des Zusammenhangs zwischen zwei Ereignissen A und B eines Wahrscheinlichkeitsraums $(\Omega, \mathcal{A}(\Omega), P)$ erstellen Sie folgende Zeichnung (siehe Abbildung 4):

Die eine Seite eines Quadrats mit der Kantenlänge 1 teilen Sie im Verhältnis $P(A) : 1-P(A)$ und teilen anschließend, entsprechend dieser Aufteilung, das Quadrat in zwei Streifen. An den beiden senkrecht dazu liegenden Seiten tragen Sie von oben die Strecken $P(B \mid A)$ und $P(B \mid \Omega \backslash A)$ ab und teilen anschließend die zu $P(A)$ bzw. $1 - P(A)$ gehörenden Streifen entsprechend.

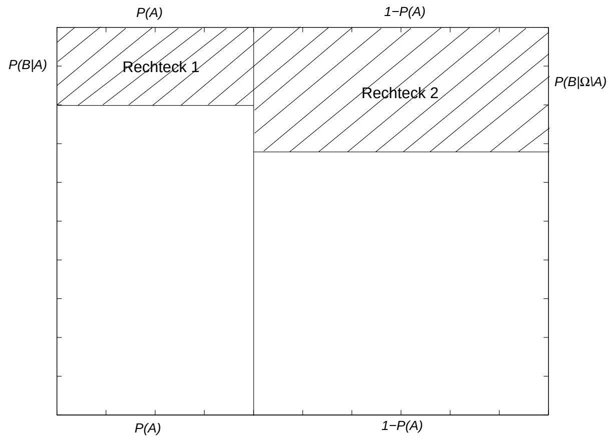


Abbildung 4: Skizze zu Aufgabe 10

(a) Zeigen Sie, dass der Flächeninhalt der schraffierten Fläche gerade $P(B)$ entspricht.

(b) Zeigen Sie, dass die Ereignisse A und B genau dann unabhängig sind, wenn die schraffierte Fläche ein Rechteck bildet.

Lösung: (Wahrscheinlichkeitstheorie S.91ff)

(a) Mit Hilfe des Satzes von der totalen Wahrscheinlichkeit ergibt sich:

$$\begin{aligned} P(B) &= \underbrace{P(B \mid A) \cdot P(A)}_{\text{=Fläche Rechteck 1}} + \underbrace{P(B \mid \Omega \backslash A) \cdot P(\Omega \backslash A)}_{\text{=Fläche Rechteck 2}} \\ &= \text{Fläche der schraffierten Fläche} \end{aligned}$$

(b)

"⇒" Seien die Ereignisse A und B unabhängig. Dann gilt für die Länge der vertikalen Kante von Rechteck 1:

$$P(B \mid A) = \frac{P(B \cap A)}{P(A)} = \frac{P(B) \cdot P(A)}{P(A)} = P(B)$$

Aus der Unabhängigkeit von A und B ergibt sich unmittelbar die Unabhängigkeit von $\Omega \backslash A$ und B (wie?) und deshalb gilt für die Länge der vertikalen Kante von Rechteck 2:

$$P(B \mid \Omega \backslash A) = \frac{P(B \cap \Omega \backslash A)}{P(\Omega \backslash A)} = \frac{P(B) \cdot P(\Omega \backslash A)}{P(\Omega \backslash A)} = P(B)$$

Die beiden Kantenlängen stimmen somit überein und die schraffierte Fläche bildet folglich ein Rechteck.

"$\Leftarrow$" Falls die schraffierte Fläche ein Rechteck darstellt so stimmt ihr Flächeninhalt mit der Länge der vertikalen Kante also mit $P(B \mid A)$ überein. Außerdem stimmt der Flächeninhalt gemäß Teilaufgabe (a) aber mit $P(B)$ überein und deshalb gilt:

$$P(A) \cdot P(B) = P(A) \cdot P(B \mid A) = P(A) \cdot \frac{P(B \cap A)}{P(A)} = P(B \cap A)$$

und somit sind die Ereignisse A und B unabhängig.

Aufgabe 11

(a) Zeigen sie, dass für die Verteilungsfunktion $F(x,y)$ einer zweidimensionalen Zufallsvariablen gilt:

(i) F ist monoton steigend

(ii) F ist in jeder der Variablen von rechts stetig.

(iii)
- $\lim\limits_{\substack{x\to+\infty \\ y\to+\infty}} F(X,Y) = 1$
- $\lim\limits_{x\to-\infty} F(x,y) = 0$ für jedes y
- $\lim\limits_{y\to-\infty} F(x,y) = 0$ für jedes x

(iv) Für alle $x_1 < x_2$ und alle $y_1 < y_2$ gilt:

$$F(x_2,y_2) - F(x_1,y_2) - F(x_2,y_1) + F(x_1,y_1) \geq 0$$

Bemerkung: Man kann zeigen, dass jede Funktion mit den Eigenschaften (i)-(iv) Verteilungsfunktion einer zweidimensionalen Zufallsvariablen ist.

(b) Überprüfen Sie unter Benutzung von Teil (a), ob die Funktion

(i)

$$F(x,y) = \begin{cases} 1 & \text{für } x \geq 0\ ,\ y \geq 0 \text{ und } x^2 + y^2 \geq 4 \\ 0 & \text{sonst} \end{cases}$$

(ii)

$$F(x,y) = \begin{cases} \min\{x,1\} \cdot \min\{y,1\} & \text{für} \quad x,y \geq 0 \\ 0 & \text{sonst} \end{cases}$$

Verteilungsfunktion einer zweidimensionalen Zufallsvariablen ist.

Lösung: (Wahrscheinlichkeitstheorie, S.129ff.)

(a) Die Verteilungsfunktion $F(x,y)$ einer zweidimensionalen Zufallsvariablen (X,Y) ist definiert als:

$$F(x,y) = P(X \leq x, Y \leq y) \text{ für } (x,y) \in \mathbb{R}^2$$

Nachweis der angegebenen Eigenschaften:

(i) Monotonie: Sei $x_1 \leq x_2$ und $y_1 \leq y_2$, dann gilt:

$$\begin{aligned} \{X \leq x_2, Y \leq y_2\} &= \{X \leq x_1, Y \leq y_1\} \cup \{X \leq x_1, y_1 < Y \leq y_2\} \cup \\ &\quad \{x_1 < X \leq x_2, Y \leq y_1\} \cup \{x_1 < X \leq x_2, y_1 < Y \leq y_2\} \end{aligned}$$

Dabei ist die Vereinigung disjunkt. Damit ergibt sich:

$$\begin{aligned} F(x_2,y_2) &= F(x_1,y_1) + P(\{X \leq x_1, y_1 < Y \leq y_2\}) + \\ &\quad + P(\{x_1 < X \leq x_2, Y \leq y_1\}) + P(\{x_1 < X \leq x_2, y_1 < Y \leq y_2\}). \end{aligned}$$

Da Wahrscheinlichkeiten nicht negativ sind, folgt somit:

$$F(x_1,y_1) = P(X \leq x_1, Y \leq y_1) \leq P(X \leq x_2, Y \leq y_2) = F(x_2,y_2)$$

(ii) Rechtsstetigkeit: Sei $(x_0, y_0) \in \mathbb{R}^2$ beliebig aber fest und $(x_n)_{n\in\mathbb{N}}$ eine Folge mit $x_n \downarrow x_0$, dann gilt analog zum eindimensionalen Fall (Wahrscheinlichkeitstheorie, S.44)

$$\{(x,y)|x > x_0 \vee y > y_0\} = \{(x,y)|x > x_1 \vee y > y_0\} \cup \bigcup_{i=1}^{\infty} \{(x,y)|x_i \geq x > x_{i+1}, y \leq y_0\}$$

und die Vereinigung ist disjunkt. Damit ist

$$\begin{aligned}
1 - F(x_0, y_0) &= P(X > x_0 \vee Y > y_0) \\
&= P(X > x_1 \vee Y > y_0) + \sum_{i=1}^{\infty} (P(x_i \geq X > x_{i+1}, Y \leq y_0) \\
&= 1 - F(x_1, y_0) + \sum_{i=1}^{\infty} (F(x_i, y_0) - F(x_{i+1} y_0)) \\
&= 1 - \lim_{i\to\infty} F(x_{i+1}, y_0)
\end{aligned}$$

woraus sich die Behauptung unmittelbar ergibt.
Analoge Beweisführung in y.

(iii) Grenzwerte: Seien $(x_n)_{n\in\mathbb{N}}, (y_n)_{n\in\mathbb{N}}$ beliebige Folgen mit $x_n \uparrow \infty$ bzw. $y_n \uparrow \infty$. Der Beweis der Behauptung

$$\lim_{n\to+\infty} F(x_n, y_n) = 1$$

erfolgt in zwei Schritten. Sei $y_0 \in \mathbb{R}$ beliebig gewählt, dann ist

$$\{(x,y)|y \leq y_0\} = \{(x,y)|x \leq x_1, y \leq y_0\} \cup \bigcup_{i=1}^{\infty} \{(x,y)|x_i < x \leq x_{i+1}, y \leq y_0\}$$

Damit gilt für die Randverteilung von Y

$$\begin{aligned}
F_Y(y_0) = P(Y \leq y_0) &= P(X \leq x_1, Y \leq y_0) + \sum_{i=1}^{\infty} (P(x_i < X \leq x_{i+1}, Y \leq y_0) \\
&= F(x_1, y_0) + \lim_{n\to\infty} \sum_{i=1}^{n} (F(x_{i+1}, y_0) - F(x_i, y_0)) \\
&= \lim_{n\to\infty} F(x_{n+1}, y_0)
\end{aligned}$$

Für die Randverteilung von Y gilt bekanntlich

$$\lim_{n\to\infty} F_Y(y_n) = 1$$

Damit ist

$$\lim_{k\to\infty} \lim_{i\to\infty} F(x_i, y_k) = 1$$

und mit

$$\lim_{n\to\infty} F(x_n, y_n) = \lim_{k\to\infty} \lim_{i\to\infty} F(x_i, y_k)$$

folgt die Behauptung.
Der Beweis für $\lim_{x\to-\infty} F(x,y)$ bzw. $\lim_{y\to-\infty} F(x,y)$ erfolgt analog zur Rechtsstetigkeit.

(iv) Ungleichung:

$$\begin{aligned}
& F(x_2,y_2) - F(x_1,y_2) - F(x_2,y_1) + F(x_1,y_1) \\
= \; & F(x_2,y_2) - F(x_1,y_2) - (F(x_2,y_1) - F(x_1,y_1)) \\
= \; & \underbrace{P(X \leq x_2, Y \leq y_2) - P(X \leq x_1, Y \leq y_2)}_{P(x_1 < X \leq x_2, Y \leq y_2)} - \underbrace{(P(X \leq x_2, Y \leq y_1) - P(X \leq x_1, Y \leq y_1))}_{P(x_1 < X \leq x_2, Y \leq y_1)} \\
= \; & P(x_1 < X \leq x_2, Y \leq y_2) - P(x_1 < X \leq x_2, Y \leq y_1) \\
= \; & P(x_1 < X \leq x_2, y_1 < Y \leq y_2) \geq 0
\end{aligned}$$

(b) (i) Die Funktion F ist keine Verteilungsfunktion, da

$$\underbrace{F(2,2)}_{=1} - \underbrace{F(0,2)}_{=1} - \underbrace{F(2,0)}_{=1} + \underbrace{F(0,0)}_{=0} = -1 < 0$$

d.h. die Bedingung (iv) aus Teilaufgabe (a) ist nicht erfüllt.

(ii) Es gilt:

$$F_{X,Y}(x,y) = \begin{cases} 0 & \text{für } x < 0 \text{ oder } y < 0 \\ x \cdot y & \text{für } 0 \leq x, y \leq 1 \\ x & \text{für } 0 \leq x \leq 1, y > 1 \\ y & \text{für } 0 \leq y \leq 1, x > 1 \\ 1 & \text{für } x, y > 1 \end{cases}$$

Die Funktion F ist stetig und monoton, die Bedingung an die Grenzwerte ist auch erfüllt. Mit vielen Fallunterscheidungen lässt sich auch zeigen, dass die Bedingung (iv) erfüllt ist, d.h. F ist eine Verteilungsfunktion. $F(x,y)$ gibt den Flächeninhalt der Fläche im Einheitsquadrat an (siehe z.B. Abb. 3(a)), der südwestlich des Punktes (x,y) liegt. Die linke Seite der Ungleichung entspricht damit dem Flächeninhalt des Teils des Rechtecks beschrieben durch die Punkte (x_1,y_1) und (x_2,y_2), der im Einheitsquadrat (siehe beispielhaft in Abb. 3(b)) liegt, und ist damit als Flächeninhalt nicht negativ.

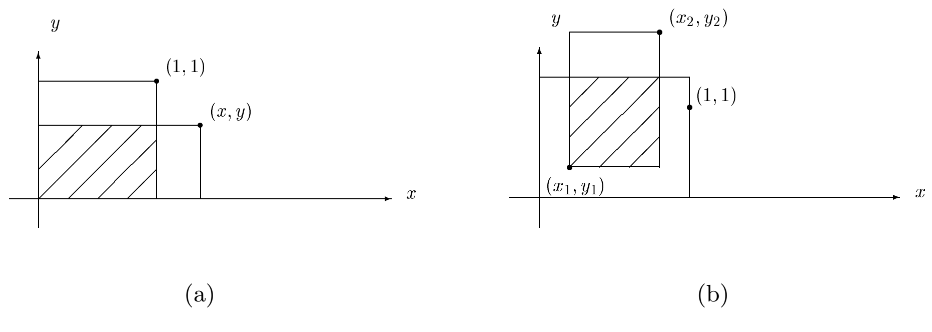


Abbildung 5: Verdeutlichung der Ungleichung

Aufgabe 12

Gegeben sei eine diskrete zweidimensionale Zufallsvariable $Y = (Y_1, Y_2)$ mit der Verteilung

(y_1, y_2)	(-1,0)	(-1,1)	(-1,c)	(2,0)	(2,1)	(2,c)	(c,0)	(c,1)	(c, c)
$P((Y_1, Y_2) = (y_1, y_2))$	1/8	1/4	1/8	1/16	c^2	1/16	1/8	1/8	1/8

(a) Bestimmen Sie c.

(b) Berechnen Sie die Randverteilungen.

(c) Wie groß ist $P(Y \leq (2, c))$?

Lösung: (Wahrscheinlichkeitstheorie, S.129ff., S.139ff.)

(a) Analog zu einer eindimensionalen Zufallsvariable muss gelten:

$$\sum_i \sum_j P(Y = (y_i, y_j)) \stackrel{!}{=} 1$$

Hieraus ergibt sich die Forderung:

$$\begin{aligned} 1 &\stackrel{!}{=} \frac{1}{8} \cdot 5 + \frac{1}{4} + \frac{1}{16} \cdot 2 + c^2 \\ &= 1 + c^2, \text{ also } c^2 = 0 \text{ bzw. } c = 0 \end{aligned}$$

d.h. die Wahrscheinlichkeitsverteilung lautet:

(y_1, y_2)	$(-1, 0)$	$(-1, 1)$	$(0, 0)$	$(0, 1)$	$(2, 0)$	$(2, 1)$
$P(Y = (y_1, y_2))$	0.25	0.25	0.25	0.125	0.125	0

(b) Die Randverteilung einer diskreten Zufallsvariable $Z = (X, Y)$ ist:

$$\begin{aligned} P(X = x_i) &= \sum_j P(X = x_i, Y = y_j) \text{ für alle } x_i \\ P(Y = y_i) &= \sum_j P(X = x_j, Y = y_i) \text{ für alle } y_i \end{aligned}$$

hier folgt mit $Y = (Y_1, Y_2)$:

y_1	-1	0	2
$P(Y_1 = y_1)$	0.5^*	0.375	0.125

bzw.

y_2	0	1
$P(Y_2 = y_2)$	0.625^{**}	0.375

wobei $^*P(Y = (-1, 0)) + P(Y = ((-1, 1))$ bzw. $^{**}P(Y = (-1, 0)) + P(Y = (2, 0)) + P(Y = (0, 0))$.

(c)

$$
\begin{aligned}
P(Y \leq (2,0)) &= P(Y_1 \leq 2 \text{ und } Y_2 \leq 0) \\
&= P(Y = (-1,0)) + P(Y = (0,0)) + P(Y = (2,0)) \\
&= \frac{1}{4} + \frac{1}{8} + \frac{1}{4} \\
&= \frac{5}{8}
\end{aligned}
$$

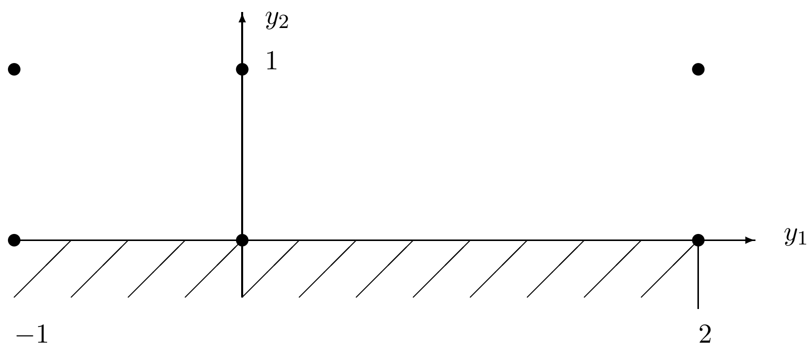


Abbildung 6: Veranschaulichung der Berechnung in (c)

Aufgabe 13

Gegeben sei die Funktion

$$\varphi(x,y) := \begin{cases} c(x^2+y^2) & -1 \le x, y \le 1 \\ \\ 0 & \text{sonst} \end{cases}$$

.

(a) Bestimmen Sie $c \in \mathbb{R}$ so, dass φ Dichte einer zweidimensionalen Zufallsvariablen (X,Y) wird.

(b) Berechnen Sie die Randdichten $\varphi_1(x)$ und $\varphi_2(y)$.

(c) Bestimmen Sie die Höhenlinien.

(d) Berechnen Sie $F_{(X,Y)}(0, 0.5)$, $E(X)$ und $E(Y \mid X = \frac{1}{2})$.

Lösung: (Wahrscheinlichkeitstheorie, S.129ff., S.139ff.,171ff.)

(a) Aus den Bedingungen an eine Dichtefunktion ergibt sich:

1. $\varphi(x,y) \ge 0$ für alle $x, y \;\Rightarrow\; c \ge 0$
2. $$\begin{aligned}
\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \varphi(x,y)dxdy &= \int_{-1}^{+1}\int_{-1}^{+1} c(x^2+y^2)dxdy \\
&= c\int_{-1}^{+1} (\frac{1}{3}x^3 + xy^2)\Big|_{-1}^{+1} dy \\
&= \frac{2}{3}c\int_{-1}^{+1}(1+3y^2)dy \\
&= \frac{2}{3}c(y+y^3)\Big|_{-1}^{+1} \\
&= \frac{2}{3}c \cdot 4 \\
&\stackrel{!}{=} 1 \;\Rightarrow\; c = \frac{3}{8}
\end{aligned}$$

(b) Für $-1 \leq x \leq 1$ gilt:

$$\begin{aligned}
\varphi_1(x) &= \int\limits_{-\infty}^{+\infty} \varphi(x,y)dy \\
&= \int\limits_{-1}^{+1} \frac{3}{8}(x^2+y^2)dy \\
&= \frac{3}{8}(yx^2 + \frac{1}{3}y^3)\Big|_{-1}^{+1} \\
&= \frac{3}{4}x^2 + \frac{1}{4}
\end{aligned}$$

Damit ist

$$\varphi_1(x) = \begin{cases} \frac{3}{4}x^2 + \frac{1}{4} & \text{für} \quad -1 \leq x \leq 1 \\ 0 & \text{sonst} \end{cases}$$

Analog ergibt sich

$$\varphi_2(y) = \begin{cases} \frac{3}{4}y^2 + \frac{1}{4} & \text{für} \quad -1 \leq y \leq 1 \\ 0 & \text{sonst} \end{cases}$$

(c) Die Definition für die Höhenlinie lautet: Menge aller Punkte$(x,y) \in \mathbb{R}^2$, für die $\varphi(x,y)$ einen vorgegebenen Wert (Höhe) h annimmt. (Schnitt durch den Graphen der Dichtefunktion auf der Höhe h)

Sei $h > 0$. Da $\varphi(x,y) \leq 2c$ ist, ergeben nur Werte von $h \leq 2c = \frac{3}{4}$ Höhenlinien:

$$\begin{aligned}
h &= \varphi(x,y)) \\
&= \tfrac{3}{8}(x^2+y^2) \\
\Leftrightarrow \quad x^2+y^2 &= \tfrac{8}{3}h
\end{aligned}$$

d.h. die Höhenlinie wird durch die Gleichung $\frac{8}{3}h = x^2+y^2$ bzw. die Menge $\{(x,y)|x,y \in [-1,1]$ und $\frac{8}{3}h = x^2+y^2\}$ beschrieben. Die Höhenlinien sind hier Kreise (bzw. Kreisausschnitte durch die Einschränkung $x,y \in [-1,1]$, s. Abb) mit Mittelpunkt (0,0) und Radius $\sqrt{\frac{8}{3}h}$, wobei $0 \leq h \leq \frac{3}{4}$ ist, d.h. mit Radien zwischen 0 und $\sqrt{2}$.

(d)

$$\begin{aligned}
F_{(X,Y)}(0,0.5) &= \int\limits_{-\infty}^{0}\int\limits_{-\infty}^{0.5} \varphi(x,y)dydx = \frac{3}{8}\int\limits_{-1}^{0}(x^2y + \frac{1}{3}y^3)\Big|_{-1}^{0.5} dx \\
&= \frac{3}{8}\int\limits_{-1}^{0}(1.5x^2 + \frac{3}{8})dx = \frac{3}{8}(\frac{1}{2}x^3 + \frac{3}{8}x)\Big|_{-1}^{0} \\
&= \frac{21}{64} \; (= 0.328125)
\end{aligned}$$

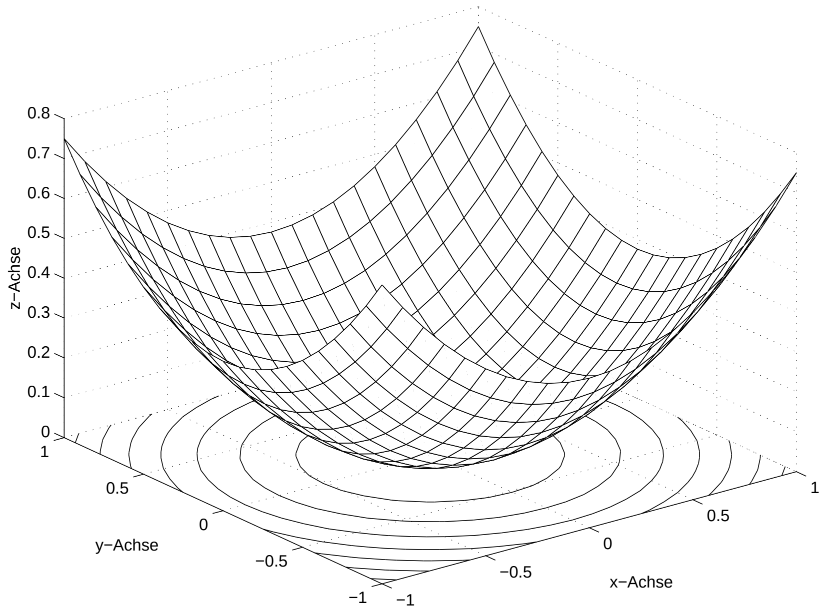


Abbildung 7: Skizze der Funktion φ mit ihren Höhenlinien zu Teilaufgabe (c)

$$\begin{aligned} E(X) &= \int_{-\infty}^{+\infty} x\varphi_1(x)dx = \int_{-1}^{1} (\frac{3}{4}x^3 + \frac{1}{4}x)dx \\ &= (\frac{3}{16}x^4 + \frac{1}{8}x^2)\Big|_{-1}^{1} = 0 \end{aligned}$$

Für die bedingte Dichte gilt:

$$\varphi_{Y\,|\,X=\frac{1}{2}}(y) = \frac{\varphi(\frac{1}{2}, y)}{\varphi_1(\frac{1}{2})} = \begin{cases} \frac{3}{14} + \frac{6}{7}y^2 & \text{für} \quad -1 \le y \le 1 \\ 0 & \text{sonst} \end{cases}$$

d.h. der bedingte Erwartungswert berechnet sich gemäß:

$$\begin{aligned} E(Y\,|\,X = \frac{1}{2}) &= \int_{-\infty}^{+\infty} y\varphi_{Y\,|\,X=\frac{1}{2}}(y)dy = \int_{-1}^{+1} (\frac{3}{14}y + \frac{6}{7}y^3)dy \\ &= (\frac{3}{28}y^2 + \frac{3}{14}y^4)\Big|_{-1}^{1} = 0. \end{aligned}$$

Aufgabe 14

Eine Straßenbahn verkehrt zwischen den Haltestellen G und H. Die Schwarzfahrer unter den Fahrgästen sind zu 60% jugendlich und zu 40% erwachsen. Um diesen das Leben zu erschweren, werden alle Fahrgäste zwischen G und H zweimal kontrolliert, und zwar zuerst von Kontrolleur Nr. 1 und dann von Kontrolleur Nr. 2. Von Kontrolleur Nr.1 entdeckte Schwarzfahrer werden natürlich nicht noch einmal kontrolliert. Kontrolleur Nr. 1 entdeckt 60% der erwachsenen Schwarzfahrer und 40% der jugendlichen Schwarzfahrer. Kontrolleur Nr.2 entdeckt 50% der erwachsenen Schwarzfahrer und 50% der jugendlichen Schwarzfahrer, die vorher noch nicht von Kontrolleur Nr. 1 entdeckt wurden.

(a) Modellieren Sie das Kontrollverfahren durch einen Wahrscheinlichkeitsraum.

(b) Wie groß ist die Wahrscheinlichkeit, dass ein Schwarzfahrer entdeckt wird ?

(c) Wie groß ist die Wahrscheinlichkeit, dass ein entdeckter Schwarzfahrer jugendlich ist ?

(d) Berechnen Sie die Wahrscheinlichkeiten weiterer Kombinationsmöglichkeiten

Lösung: (Wahrscheinlichkeitstheorie, S.129ff.)

(a) Wahrscheinlichkeitsraum:

Grundgesamtheit: Menge aller Schwarzfahrer (betrachtet nach der Kontrolle)

$$\Omega = \{\omega \,|\, \omega \text{ Schwarzfahrer}\}$$

Mengensystem der Ereignisse: Potenzmenge von Ω (da Ω endlich ist):

$$A(\Omega) = \mathcal{P}(\Omega)$$

Wahrscheinlichkeitsmaß:

$$\begin{aligned} P : \mathcal{P}(\Omega) &\rightarrow [0,1] \\ \omega &\mapsto P(\{\omega\}) = \frac{1}{N}, \ N \text{ Anzahl der Schwarzfahrer, und somit} \\ A &\mapsto P(A) = \frac{\#A}{N} \end{aligned}$$

Es handelt sich um einen "Laplaceschen Wahrscheinlichkeitsraum" mit der Menge der Schwarzfahrer als Grundgesamtheit.

Zur Differenzierung der Schwarzfahrer werden Zufallsvariablen X und Y eingeführt

$$X(\omega) = \begin{cases} 0 & \text{für } \omega \text{ erwachsen} \\ 1 & \text{für } \omega \text{ jugendlich} \end{cases}$$

$$Y(\omega) = \begin{cases} 1 & \text{für Schwarzfahrer, entdeckt von Kontrolleur Nr. 1 } (K_1) \\ 2 & \text{für Schwarzfahrer, nicht entdeckt von Kontrolleur } K_1 \text{, entdeckt von } K_2 \\ 3 & \text{für Schwarzfahrer nicht entdeckt} \end{cases}$$

Aus den Angaben ergibt sich:

$$
\begin{aligned}
&P(X=0)=0.4,\ P(X=1)=0.6\\
&P(Y=1\,|\,X=0)=0.6,\ P(Y=1\,|\,X=1)=0.4\\
&P(Y=2\,|\,X=0)=0.4\cdot 0.5,\ P(Y=2\,|\,X=1)=0.6\cdot 0.5,\\
&P(Y=3\,|\,X=0)=0.4\cdot 0.5,\ P(Y=3\,|\,X=1)=0.6\cdot 0.5
\end{aligned}
$$

(b) Gesucht: P(Schwarzfahrer entdeckt)=1 $- P(Y=3)$

$P(Y=3)$ läßt sich zerlegen in die Wahrscheinlichkeit, dass ein erwachsener Schwarzfahrer nicht entdeckt wird, und die Wahrscheinlichkeit, dass ein jugendlicher Schwarzfahrer nicht entdeckt wird:

$$
\begin{aligned}
P(Y=3) &= \underbrace{P(Y=3|X=0)}_{=0.4\cdot 0.5}\cdot\underbrace{P(X=0)}_{=0.4}+\underbrace{P(Y=3|X=1)}_{=0.6\cdot 0.5}\cdot\underbrace{P(X=1)}_{=0.6}\\
&= 0.08+0.18=0.26
\end{aligned}
$$

Folglich ergibt sich:

$$P(\text{Schwarzfahrer entdeckt})=0.74$$

Alternativ könnte die Wahrscheinlichkeit auch wie folgt berechnet werden:

$$P(\text{Schwarzfahrer entdeckt})=P(Y\in\{1,2\})=P(Y=1)+P(Y=2)$$

Für die einzelnen Summanden ergibt sich:

$$
\begin{aligned}
P(Y=1) &= P(Y=1\,|\,X=0)P(X=0)+P(Y=1\,|\,X=1)P(X=1)\\
&= 0.6\cdot 0.4+0.4\cdot 0.6=0.48\\
P(Y=2) &= P(Y=2\,|\,X=0)P(X=0)+P(Y=2\,|\text{X}=1)P(X=1)\\
&= 0.26\ (\text{analog zu } P(Y=3))
\end{aligned}
$$

Damit ergibt sich: $P(Y=1)+P(Y=2)=0.74$

(c) Gesucht wird $P(X=1\,|\,Y\in\{1,2\})$

$$
\begin{aligned}
P(X=1\,|\,Y\in\{1,2\}) &= \frac{P(X=1,Y\in\{1,2\})}{P(Y\in\{1,2\})}\\
&= \frac{P(X=1,Y=1)+P(X=1,Y=2)}{0.74}\\
&= \frac{P(Y=1\,|\,X=1)P(X=1)+P(Y=2\,|\,X=1)P(X=1)}{0.74}\\
&= \frac{0.6\cdot 0.4+0.6\cdot 0.5\cdot 0.6}{0.74}=\frac{0.42}{0.74}=0.5676
\end{aligned}
$$

(d) Die gemeinsame Wahrscheinlichkeitsverteilung von X and Y ergibt sich gemäß

$$
\begin{aligned}
P(X=0,Y=1) &= P(Y=1\,|\,X=0)\cdot P(X=0)\\
&= 0.6\cdot 0.4=0.24
\end{aligned}
$$

zu

$Y=$ / $X=$	1	2	3	$\sum$
0	0.24	0.08	0.08	0.4
1	0.24	0.18	0.18	0.6
$\sum$	0.48	0.26	0.26	1

Aufgabe 15

In einem Bahnhof steigen drei Reisende in einen Zug ein. Jeder von ihnen setzt sich zufällig in eines von drei Zugabteilen. Die Zufallsvariable N bezeichne die Anzahl der (von ihnen) besetzten Abteile, X_i ($i = 1, 2, 3$) die Anzahl der Reisenden im i-ten Abteil.

(a) Bestimmen Sie die gemeinsame Wahrscheinlichkeitsverteilung von N und X_1 sowie von X_1 und X_2.

(b) Berechnen Sie $E(N)$, $E(X_1)$, $Var(N)$ und $Var(X_1)$.

(c) Sind N und X_1 bzw. X_1 und X_2 unabhängig?

Lösung: (Wahrscheinlichkeitstheorie, S.129ff., S.155ff., S.171ff.)

(a) **1.Weg:** Tabelle mit den möglichen Konstellationen (Es wird zwischen den 3 Reisenden unterschieden, d.h. es gibt 27 Möglichkeiten, wie sich die 3 Reisenden auf die Abteile verteilen können; jeder der Reisenden hat ja 3 Möglichkeiten. Jede der 27 Möglichkeiten hat die Wahrscheinlichkeit $\frac{1}{27}$.)

X_1	X_2	X_3	N	Anzahl Möglichkeiten
3	0	0	1	1
0	3	0	1	1
0	0	3	1	1
2	1	0	2	3
2	0	1	2	3
0	2	1	2	3
1	2	0	2	3
1	0	2	2	3
0	1	2	2	3
1	1	1	3	6
				$\sum = 27$

Man erhält somit folgende Wahrscheinlichkeitsverteilung:

x_1 \ x_2	0	1	2	3	$\sum$
0	$\frac{1}{27}$	$\frac{3}{27}$	$\frac{3}{27}$	$\frac{1}{27}$	$\frac{8}{27}$
1	$\frac{3}{27}$	$\frac{6}{27}$	$\frac{3}{27}$	0	$\frac{12}{27}$
2	$\frac{3}{27}$	$\frac{3}{27}$	0	0	$\frac{6}{27}$
3	$\frac{1}{27}$	0	0	0	$\frac{1}{27}$
$\sum$	$\frac{8}{27}$	$\frac{12}{27}$	$\frac{6}{27}$	$\frac{1}{27}$	1

bzw.

x_1 \ N	1	2	3	$\sum$
0	$\frac{2}{27}$	$\frac{6}{27}$	0	$\frac{8}{27}$
1	0	$\frac{6}{27}$	$\frac{6}{27}$	$\frac{12}{27}$
2	0	$\frac{6}{27}$	0	$\frac{6}{27}$
3	$\frac{1}{27}$	0	0	$\frac{1}{27}$
$\sum$	$\frac{3}{27}$	$\frac{18}{27}$	$\frac{6}{27}$	1

2. Weg: Multinomialverteilung: $Multinomial(n, p_1, \ldots, p_r)$

Die Multinomialverteilung beschreibt folgenden Zufallsvorgang: In einer Urne befinden sich Kugeln in den Farben $1, \ldots, r$, wobei $p_i \in [0,1]$ den Anteil der Kugeln mit Farbe i sei (Es gilt also $\sum_{i=1}^{r} p_i = 1$). Es wird eine Stichprobe mit Zurücklegen vom Umfang n gezogen. X_i sei die Anzahl der Kugeln der Farbe i. Dann ist die gemeinsame Wahrscheinlichkeitsverteilung von $(X_1, \ldots, X_r)$ gegeben durch:

$$P(X_1 = x_1, \ldots, X_r = x_r) =$$

$$\begin{cases} \binom{n}{x_1,\ldots,x_r} \cdot {p_1}^{x_1} \cdots {p_r}^{x_r} = \frac{n!}{x_1! \cdots x_r!} \cdot {p_1}^{x_1} \cdots {p_r}^{x_r} & \text{für } \sum\limits_{i=1}^{r} x_i = n \\ 0 & \text{sonst} \end{cases}$$

Für $r = 2$ entspricht dies der Binomialverteilung.

Zur Aufgabe: Die Zufallsvariable (X_1, X_2, X_3) ist multinomialverteilt mit Parametern $(3; \frac{1}{3}, \frac{1}{3}, \frac{1}{3})$, da jedes Abteil hier einer Farbe im Urnenmodell entspricht.

$$P(X_1 = x_1, X_2 = x_2, X_3 = x_3) = \begin{cases} \frac{3!}{x_1!x_2!x_3!} \cdot (\frac{1}{3})^3 & \text{für } \sum\limits_{i=1}^{3} x_i = 3 \\ 0 & \text{sonst} \end{cases}$$

d.h. wir erhalten die folgenden Wahrscheinlichkeitsverteilungen:

X_2 \ X_1	0	1	2	3	
0	$(\frac{1}{3})^3$	$3 \cdot (\frac{1}{3})^3$	$3 \cdot (\frac{1}{3})^3$	$(\frac{1}{3})^3$	$(\frac{2}{3})^3$
1	$3 \cdot (\frac{1}{3})^3$	$6 \cdot (\frac{1}{3})^3$	$3 \cdot (\frac{1}{3})^3$	0	$3 \cdot (\frac{1}{3}) \cdot (\frac{2}{3})^2$
2	$3 \cdot (\frac{1}{3})^3$	$3 \cdot (\frac{1}{3})^3$	0	0	$3 \cdot (\frac{1}{3})^2 \cdot (\frac{2}{3})$
3	$(\frac{1}{3})^3$	0	0	0	$(\frac{1}{3})^3$
	$(\frac{2}{3})^3$	$3 \cdot (\frac{1}{3}) \cdot (\frac{2}{3})^2$	$3 \cdot (\frac{1}{3})^2 \cdot (\frac{2}{3})$	$(\frac{1}{3})^3$	1

die einzelnen Komponenten erhält man z.B. gemäß:

$$P(X_1 = 0, X_2 = 0) = P(X_1 = 0, X_2 = 0, X_3 = 3) = \frac{3!}{0!0!3!} (\frac{1}{3})^0 (\frac{1}{3})^0 (\frac{1}{3})^3$$

N \ X_1	0	1	2	3	
1	$2\cdot(\frac{1}{3})^3$	0	0	$(\frac{1}{3})^3$	$3\cdot(\frac{1}{3})^3$
2	$2\cdot 3\cdot(\frac{1}{3})^3$	$6\cdot(\frac{1}{3})^3$	$2\cdot 3\cdot(\frac{1}{3})^3$	0	$18\cdot(\frac{1}{3})^3$
3	0	$6\cdot(\frac{1}{3})^3$	0	0	$6\cdot(\frac{1}{3})^3$
	$(\frac{2}{3})^3$	$3\cdot(\frac{1}{3})\cdot(\frac{2}{3})^2$	$3\cdot(\frac{1}{3})^2\cdot(\frac{2}{3})$	$(\frac{1}{3})^3$	1

die einzelnen Komponenten erhält man z.B. gemäß:

$$P(X_1 = 0; N = 1) = P(X_1 = 0, X_2 = 3, X_3 = 0) + P(X_1 = 0, X_2 = 0, X_3 = 3)$$

Bei den Randverteilungen von X_1, X_2, X_3 handelt es sich um Binomialverteilungen mit $n = 3$ und $p = \frac{1}{3}$.

(b) Der Erwartungswert berechnet sich aus der Wahrscheinlichkeitsverteilung

$$E(X_1) = 0\cdot(\frac{2}{3})^3 + 1\cdot 3\cdot(\frac{1}{3})\cdot(\frac{2}{3})^2 + 2\cdot 3\cdot(\frac{1}{3})^2\cdot(\frac{2}{3}) + 3\cdot(\frac{1}{3})^3 = 1$$

oder direkt aus dem Erwartungswert der Binomialverteilung.

$$E(N) = 1\cdot 3\cdot(\frac{1}{3})^3 + 2\cdot 18\cdot(\frac{1}{3})^3 + 3\cdot 6\cdot(\frac{1}{3})^3 = \frac{57}{27} = \frac{19}{9}$$

$Var(N) = E(N^2) - [E(N)]^2$ mit

$$\begin{aligned} E(N^2) &= 1^2\cdot 3\cdot(\frac{1}{3})^3 + 2^2\cdot 18\cdot(\frac{1}{3})^3 + 3^2\cdot 6\cdot(\frac{1}{3})^3 \\ &= \frac{129}{27} = \frac{43}{9} \\ Var(N) &= \frac{43}{9} - (\frac{19}{9})^2 = \frac{26}{81} = 0.3209. \end{aligned}$$

und mit der Varianz der Binomialverteilung ergibt sich $Var(X_1) = \frac{2}{3}$. (Alternativ: $E(X_1^2) = \frac{5}{3}$)

(c) N und X_1 sind unabhängig, falls $P(N = n \text{ und } X_1 = x) = P(N = n)\cdot P(X_1 = x)$ für alle $x \in \{0, \ldots, 3\}$ und alle $n \in \{1, \ldots, 3\}$ ist.

Z.B. ist $P(N = 1, X_1 = 2) = 0 \neq \frac{3}{27}\cdot\frac{6}{27} = P(N = 1)\cdot P(X_1 = 2)$, d.h. N und X_1 sind nicht unabhängig.

Für X_1, X_2 gilt beispielsweise:

$$P(X_1 = 1 \text{ und } X_2 = 2) = \frac{3}{27} \neq \frac{12}{27}\cdot\frac{6}{27} = P(X_1 = 1)\cdot P(X_2 = 2),$$

d.h. X_1 und X_2 sind ebenfalls abhängig. (Da die Gesamtzahl der Reisenden auf 3 beschränkt ist, ändern sich durch Kenntnis der Anzahl der Personen in einem Abteil die möglichen Kombinationen für die anderen beiden Abteile.)

Aufgabe 16

Die zweidimensionale Zufallsvariable (X, Y) besitze die Dichte

$$f_{(X,Y)}(x,y) = \begin{cases} x + y & \text{für} \quad 0 \leq x, y \leq 1 \\ 0 & \text{sonst} \end{cases} .$$

(a) Bestimmen Sie die Randdichten $f_X(x)$ und $f_Y(y)$.

(b) Berechnen Sie die Verteilungsfunktion von (X, Y).

(c) Sind X und Y unabhängig ?

Lösung: (Wahrscheinlichkeitstheorie, S.139ff., S.155f.)

(a) Für $0 < x < 1$ gilt:

$$\begin{aligned} f_X(x) &= \int_{-\infty}^{+\infty} f_{(X,Y)}(x,y)\ dy \\ &= \int_0^1 (x+y)\ dy \\ &= x + \frac{1}{2} \end{aligned}$$

d.h.

$$f_X(x) = \begin{cases} x + \frac{1}{2} & \text{für} \quad x \in\ [0;1] \\ 0 & \text{sonst} \end{cases}$$

Analog ist für $0 < y < 1$:

$$\begin{aligned} f_Y(y) &= \int_{-\infty}^{+\infty} f(x,y)\ dx \\ &= \int_0^1 (x+y)\ dy \\ &= y + \frac{1}{2} \end{aligned}$$

d.h.

$$f_Y(y) = \begin{cases} y + \frac{1}{2} & \text{für} \quad y \in [0,1] \\ 0 & \text{sonst} \end{cases}$$

Bemerkung: Wegen der Symmetrie von $f_{(X,Y)}(x,y)$ kann $f_Y(y)$ auch direkt aus $f_X(x)$ abgelesen werden.

(b) Folgende Bereiche sind zu betrachten:

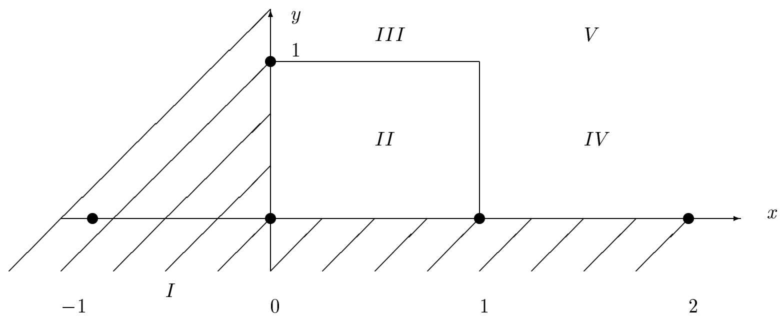


Abbildung 8: Definitionsbereiche der verschiedene Äste der Verteilungsfunktion

I: $x \leq 0$ oder $y \leq 0$:

$$F_{(X,Y)}(x,y) = 0$$

II: $x, y \in [0,1]$:

$$\begin{aligned} F_{(X,Y)}(x,y) &= \int_0^x \int_0^y (s+t)\ ds\ dt \\ &= \int_0^x st + \frac{1}{2}t^2 \Big|_0^y ds \\ &= \int_0^x sy + \frac{1}{2}y^2 = \frac{1}{2}ys^2 + \frac{1}{2}y^2 s \Big|_0^x \\ &= \frac{1}{2}yx^2 + \frac{1}{2}y^2 x = \frac{1}{2}xy(x+y) \end{aligned}$$

III: $y > 1,\ x \in [0,1]$: (vergleiche Bereich II mit $y = 1$)

$$\begin{aligned} F_{(X,Y)}(x,y) &= \int_0^1 \int_0^x s + t\ dt\ ds \\ &= \frac{1}{2}x(x+1) \end{aligned}$$

IV: $x > 1,\ y \in [0,1]$: (vergleiche Bereich II mit $x = 1$)

$$\begin{aligned} F_{(X,Y)}(x,y) &= \int_0^y \int_0^1 s + t\ dt\ ds \\ &= \frac{1}{2}y(1+y) \end{aligned}$$

V: $x, y > 1$: (vergleiche Bereich II mit $x = y = 1$)

$$F_{(X,Y)}(x,y) \quad = \quad 1$$

Damit erhält man:

$$F_{(X,Y)}(x,y) = \begin{cases} 0 & \text{für} \quad x < 0 \; y < 0 \\ \\ \frac{1}{2}xy(x+y) & \text{für} \quad x, y \in [0,1] \\ \\ \frac{1}{2}x(1+x) & \text{für} \quad x \in [0,1], \; y > 1 \\ \\ \frac{1}{2}y(1+y) & \text{für} \quad y \in [0,1], \; x > 1 \\ \\ 1 & \qquad\quad \text{sonst} \end{cases}$$

(c) Für $x, y \in [0,1]$ gilt:

$$\begin{aligned} f_X(x) \cdot f_Y(y) &= (x + \frac{1}{2}) \cdot (y + \frac{1}{2}) \\ &\neq x + y = f_{(X,Y)}(x,y) \end{aligned}$$

$\Rightarrow$ X, Y sind somit abhängig.

Aufgabe 17

Die zweidimensionale Zufallsvariable (X, Y) hat folgende Dichtefunktion:

$$f_{(X,Y)}(x,y) = \begin{cases} \frac{6}{5}x & \text{für} \quad 1 \leq x \leq 2, 0 \leq y \leq x-1 \\ 0 & \quad \text{sonst} \end{cases}$$

(a) Bestimmen Sie die Randdichten $f_X(x)$ und $f_Y(y)$.

(b) Sind X und Y unabhängig?

(c) Bestimmen Sie Erwartungswert und Varianz von X und Y.

Lösung: (Wahrscheinlichkeitstheorie, S.139ff., S.155f, S.171ff.)

(a) Für $1 \leq x \leq 2$ ergibt sich:

$$\begin{aligned} f_X(x) &= \int_{-\infty}^{+\infty} f_{(X,Y)}(x,y)dy \;=\; \int_0^{x-1} \frac{6}{5}x \; dy = \frac{6}{5}xy \Big|_0^{x-1} \\ &= \frac{6}{5}x(x-1) = \frac{6}{5}x^2 - \frac{6}{5}x. \end{aligned}$$

Für $0 \leq y \leq 1$ ergibt sich:

$$\begin{aligned} f_Y(y) &= \int_{-\infty}^{+\infty} f_{X,Y}(x,y)dx \;=\; \int_{y+1}^{2} \frac{6}{5}x \; dx = \frac{3}{5}x^2 \Big|_{y+1}^{2} \\ &= \frac{3}{5}(4-(y+1)^2) = \frac{3}{5}(-y^2-2y+3). \end{aligned}$$

Also ist :

$$f_X(x) = \begin{cases} \frac{6}{5}x(x-1) & \text{für} \quad 1 \leq x \leq 2 \\ 0 & \quad \text{sonst} \end{cases}$$

$$f_Y(y) = \begin{cases} \frac{3}{5}(3-y^2-2y) & \text{für} \quad 0 \leq y \leq 1 \\ 0 & \quad \text{sonst} \end{cases}$$

(b) Beispielsweise gilt:

$$f_{(X,Y)}(1.5, 0.8) \;=\; 0 \;\neq\; f_X(1.5) \cdot f_Y(0.8) \qquad (f_X(1.5) > 0 \text{ und } f_Y(0.8) > 0),$$

d.h. X und Y sind abhängig.

(c) Die Momente ergeben sich gemäß:

$$\begin{aligned}
E(X) &= \int_{-\infty}^{\infty} x f_X(x) dx \;=\; \frac{6}{5}\int_{1}^{2}(x^3 - x^2)dx \;=\; 1.7 \\
Var(X) &= E(X^2) - [E(X)]^2 \\
&= \int_{-\infty}^{\infty} x^2 f_X(x) dx \; - (1.7)^2 \;=\; 2.94 - 2.89 = \; 0.05 \\
E(Y) &= \int_{-\infty}^{\infty} y f_Y(y) dy \;=\; \frac{3}{5}\int_{0}^{1} y(-y^2 - 2y + 3)dy \;=\; \frac{7}{20} \\
Var(Y) &= \frac{3}{5}\int_{0}^{1}(-y^4 - 2y^3 + 3y^2)dy - (\frac{7}{20})^2 \\
&= \frac{9}{50} - (\frac{7}{20})^2 \;=\; 0.0575
\end{aligned}$$

Aufgabe 18

Eine Münze wird n-mal geworfen. Beim einmaligen Werfen gilt:

$$P(\text{Zahl})=p \qquad P(\text{Wappen})=1-p$$

mit $0 < p < 1$. Man geht davon aus, dass die Ergebnisse bei den einzelnen Würfen sich nicht gegenseitig beeinflussen. Beschreiben Sie den Zufallsvorgang "n-maliges Werfen" durch einen Zufallsvektor $X = (X_1 \ldots X_n)$ (Verwenden Sie hierbei die Codierung: Wappen=0, Zahl=1). Geben Sie die gemeinsame Wahrscheinlichkeitsverteilung von $X_1 \ldots, X_n$ an und bestimmen Sie die Wahrscheinlichkeitsverteilung von $\sum\limits_{i=1}^{n} X_i$. Wie kann $\sum\limits_{i=1}^{n} X_i$ interpretiert werden?

Lösung: (Wahrscheinlichkeitstheorie, S.161ff., S.182ff.)

Die gemeinsame Wahrscheinlichkeitsverteilung berechnet sich gemäß

$$\begin{aligned}
P(X_1 = x_1, \ldots, X_n = x_n) &= \prod_{i=1}^{n} P(X_i = x_i) \\
&= \prod_{i=1}^{n} p^{x_i} \cdot (1-p)^{1-x_i} \\
&= p^{\sum\limits_{i=1}^{n} x_i} \cdot (1-p)^{n-\sum\limits_{i=1}^{n} x_i}
\end{aligned}$$

Das Ereignis $(X_1, ..., X_n) = (x_1, ..., x_n)$ ist eine Permutation (Reihenfolge der Einsen und Nullen wird beachtet), dessen Wahrscheinlichkeit $p^{\sum_{i=1}^{n} x_i}(1-p)^{n-\sum_{i=1}^{n} x_i}$ ist. Insgesamt gibt es $\binom{n}{k}$ Permutationen $(x_1, ..., x_n)$ mit der Eigenschaft $\sum_{i=1}^{n} x_i = k$. Somit folgt

$$P(\sum_{i=1}^{n} X_i = k) = \binom{n}{k} p^k (1-p)^{n-k},$$

d.h. $\sum_{i=1}^{n} X_i$ ist binomialverteilt mit Parameter n, p.

Aufgabe 19

Überprüfen Sie die folgenden Behauptungen:

(a) Für zwei beliebige diskrete Zufallsvariablen X und Y gilt: Hängt die bedingte Wahrscheinlichkeit von Y unter der Bedingung $X = x$ nicht von x ab, ist also für alle x übereinstimmend, so stimmt die bedingte Wahrscheinlichkeitsverteilung mit der Randverteilung überein (die Umkehrung gilt offensichtlich auch).

(b) Für stetige Zufallsvariablen X und Y gilt die analoge Beziehung zwischen bedingter Dichte und Randdichte.

Lösung: (Wahrscheinlichkeitstheorie, S.139ff., S.152ff.)

(a) Formal lautet die Behauptung: Seien X,Y diskrete Zufallsvariablen mit $x_i \in \mathbb{R}$, $i = 1, \ldots, n_1$, Werte von X und y_i, $i = 1, \ldots, n_2$, Werte von Y. Dann gilt:

Ist $P(Y = y_i|X = x_j) = P(Y = y_i|X = x_k)$ für alle $i \in \{1, \ldots, n_2\}$ und alle x_j, x_k mit $j \neq k,\ j, k \in \{1, \ldots, n_1\}$, dann ist $P(Y = y_i|X = x_j) = P(Y = y_i)$, für jedes $x_j,\ j \in \{1, \ldots, n_1\}$ und alle $i \in \{1, \ldots, n_2\}$.

Der Beweis dazu: (analog zum Beweis zu Satz 8.9, Deskriptive Statistik)

$$\begin{aligned} P(Y = y_i) &= \sum_{j=1}^{n_1} P(Y = y_i, X = x_j) = \sum_{j=1}^{n_1} P(Y = y_i|X = x_j) \cdot P(X = x_j) \\ &= P(Y = y_i|X = x^*) \sum_{j=1}^{n_1} P(X = x_j) = P(Y = y_i \mid X = x^*). \end{aligned}$$

mit $x^* \in \{x_j|j = 1, \ldots, n_1\}$ beliebig (Verwendung des Bedingungsteil der Aussage).

(b) Die Behauptung lautet nun: X, Y seien stetige Zufallsvariablen mit $f_{Y|X=x}(y) = f_{Y|X=x'}(y)$ für alle $y, x, x' \in \mathbb{R}$, dann gilt $f_{Y|X=x}(y) = f_Y(y)$ für jedes x.

Der Beweis dazu:

$$\begin{aligned} f_Y(y) &= \int_{-\infty}^{\infty} f_{(X,Y)}(x, y)dx = \int_{-\infty}^{\infty} \frac{f_{(X,Y)}(x, y)}{f_x(x)} \cdot f_x(x)\, dx \\ &= \int_{-\infty}^{\infty} f_{Y|X=x}(y) \cdot f_x(x)\, dx = f_{Y|X=x^*}(y) \cdot \int_{-\infty}^{\infty} f_x(x)\, dx \\ &= f_{Y|X=x^*}(y). \end{aligned}$$

Aufgabe 20

Sei (X, Y) eine zweidimensionale Zufallsvariable mit der Dichte

$$f_{(X,Y)}(x,y) = \begin{cases} \frac{1}{4}(x+y+xy) & \text{für} \quad 0 \leq x \leq 1, 0 \leq y \leq 2 \\ 0 & \text{sonst} \end{cases}$$

(a) Berechnen Sie $E(X)$ und $E(Y)$.

(b) Bestimmen Sie die Kovarianzmatrix von (X, Y).

(c) Berechnen Sie den Korrelationskoeffizienten von X und Y.

(d) Sind X und Y unabhängig?

Lösung: (Wahrscheinlichkeitstheorie, S.171ff., S.155f.)

(a) Zur Berechnung der Erwartungswerte werden zuerst die Randdichten bestimmt. Für $0 \leq x \leq 1$ gilt:

$$\begin{aligned} f_X(x) &= \int_{-\infty}^{\infty} f_{(X,Y)}(x,y)dy = \int_0^2 \frac{1}{4}(x+y+xy)dy \\ &= \frac{1}{4}(xy + \frac{1}{2}y^2 + \frac{1}{2}xy^2)\Big|_0^2 = \frac{1}{4}(2x+2+2x) = x + \frac{1}{2}, \end{aligned}$$

und es ergibt sich

$$f_X(x) = \begin{cases} x + \frac{1}{2} & \text{für} \quad x \in [0,1] \\ 0 & \text{sonst} \end{cases}$$

Für $0 \leq y \leq 2$ ist:

$$\begin{aligned} f_Y(y) &= \frac{1}{4}\int_0^1 (x+y+xy)dx = \frac{1}{4}\,(\frac{1}{2}x^2 + xy + \frac{1}{2}x^2 y)\Big|_0^1 \\ &= \frac{1}{4}(\frac{1}{2} + y + \frac{1}{2}y) = \frac{3}{8}y + \frac{1}{8}, \end{aligned}$$

d.h. die Randdichte von y lautet:

$$f_Y(y) = \begin{cases} \frac{3}{8}y + \frac{1}{8} & \text{für} \quad y \in [0,2] \\ 0 & \text{sonst} \end{cases}$$

Für die Erwartungswerte ergibt sich somit:

$$\begin{aligned} E(X) &= \int_0^1 x(x+\frac{1}{2})dx = (\frac{1}{3}x^3 + \frac{1}{4}x^2)\Big|_0^1 = \frac{7}{12} \\ E(Y) &= \int_0^2 y(\frac{3}{8}y + \frac{1}{8}) = \frac{1}{8}(y^3 + \frac{1}{2}y^2)\Big|_0^2 = \frac{1}{8}(8+2) = \frac{5}{4} \end{aligned}$$

(b) Die Kovarianzmatrix lautet:

$$C(X,Y) = \begin{pmatrix} Var(X) & Cov(X,Y) \\ Cov(X,Y) & Var(Y) \end{pmatrix}$$

Für die einzelnen Einträge ergibt sich:

$$\begin{aligned} Var(X) &= E(X^2) - (E(X))^2 = \int_0^1 x^2(x+\frac{1}{2})dx - (\frac{7}{12})^2 \\ &= (\frac{1}{4}x^4 + \frac{1}{6}x^3)\Big|_0^1 - (\frac{7}{12})^2 = \frac{5}{12} - \frac{49}{144} = \frac{11}{144} \\ Var(Y) &= E(Y^2) - (E(Y))^2 = \int_0^2 y^2(\frac{3}{8}y + \frac{1}{8})dy - (\frac{5}{4})^2 \\ &= \frac{1}{8}(\frac{3}{4}y^4 + \frac{1}{3}y^3)\Big|_0^2 - (\frac{5}{4})^2 = \frac{1}{8}(12 + \frac{8}{3}) - \frac{25}{16} = \frac{11}{6} - \frac{25}{16} = \frac{13}{48} \end{aligned}$$

Die Formel für die Kovarianz lautet:

$$Cov(X,Y) = E(XY) - E(X)E(Y) = E[(X - E(X))(Y - E(Y))]$$

$$\begin{aligned} E(XY) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf_{(X,Y)}(x,y)dydx = \frac{1}{4}\int_0^1\int_0^2 xy(x+y+xy)dydx \\ &= \frac{1}{4}\int_0^1 (\frac{1}{2}x^2y^2 + \frac{1}{3}xy^3 + \frac{1}{3}x^2y^3)\Big|_0^2 dx = \frac{1}{4}\int_0^1 (2x^2 + \frac{8}{3}x + \frac{8}{3}x^2)dx \\ &= \frac{1}{4}(\frac{2}{3}x^3 + \frac{4}{3}x^2 + \frac{8}{9}x^3)\Big|_0^1 = \frac{1}{4}(\frac{12+24+16}{18}) = \frac{13}{18} \end{aligned}$$

somit erhält man:

$$Cov(X,Y) = \frac{13}{18} - (\frac{7}{12} \cdot \frac{5}{4}) = \frac{104}{144} - \frac{105}{144} = -\frac{1}{144}$$

Damit ergibt sich die Kovarianzmatrix:

$$\begin{pmatrix} \frac{11}{144} & -\frac{1}{144} \\ -\frac{1}{144} & \frac{13}{48} \end{pmatrix}$$

(c) Der Korrelationskoeffizient lautet:

$$\begin{aligned} \varrho_{X,Y} &= \frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}} \\ &= \frac{-\frac{1}{144}}{\sqrt{\frac{11}{144} \cdot \frac{13}{48}}} = -0.0483 \end{aligned}$$

(d) Es gilt: $\rho_{x,y} \neq 0$, d.h. X und Y sind abhängig. Analog zur deskriptiven Statistik gilt aber: X, Y unabhängig, dann $Cov(X,Y) = 0 = \varrho_{X,Y}$, aber nicht umgekehrt.

Aufgabe 21

(a) Bestimmen Sie für die Gleichverteilung über der Fläche $A = \{(x, y)|x^2 + y^2 < 1\} \subset \mathbb{R} \times \mathbb{R}$ die bedingte Verteilung $f(y|x_0)$ für ein beliebiges $x_0 \in (-1, 1)$. Wie sehen die Höhenlinien der bedingten Verteilung in Abhängigkeit von x_0 aus? Sind die beiden Komponenten eines Zufallsvektors mit Gleichverteilung über A unabhängig?

(b) Die Zufallsvariable T gebe einen entsprechend der Gleichverteilung zufällig aus dem Intervall $[0, \alpha]$, $\alpha > 0$, ausgewählten Punkt an. Von diesem Zufallsexperiment werden n unabhängige Wiederholungen gemacht (d.h. wir betrachten n unabhängige Realisationen der Zufallsvariablen T). Die Zufallsvariable X gebe an, wie oft dabei ein Wert aus einem fest vorgegebenen Teilintervall von $[0, \alpha]$ der Länge s gewählt wurde. Wie groß ist die Wahrscheinlichkeit $P(X = k)$?

Lösung: (Wahrscheinlichkeitstheorie, S.139ff., S.155f., S.161ff.)

(a) Sei (X, Y) eine zweidimensionale Zufallsvariable mit Gleichverteilung über A. Die Fläche von A ist π, also ist die Dichte von (X, Y):

$$f(x, y) = \frac{1}{\pi}\mathbf{1}_A(x, y) \quad = \quad \begin{cases} \frac{1}{\pi} & \text{für} \quad (x, y) \in A \\ 0 & \text{sonst} \end{cases}$$

Die Randdichte von X an der Stelle x_0 lautet

$$f_X(x_0) = \int_{-\infty}^{\infty} f(x_0, y)dy \cdot 1_{(-1,1)}(x_0) = \int_{-\sqrt{1-x_0^2}}^{\sqrt{1-x_0^2}} \frac{1}{\pi}dy \cdot 1_{(-1,1)}(x_0) = \frac{1}{\pi}2\sqrt{1 - x_0^2}\ \mathbf{1}_{(-1,1)}(x_0).$$

Für $x_0 \in (-1, 1)$ gilt damit:

$$\begin{aligned} f(y|x_0) &= \frac{f(x_0, y)}{f_X(x_0)} = \frac{\frac{1}{\pi}\mathbf{1}_A(x_0, y)}{\frac{1}{\pi}2\sqrt{1 - x_0^2}} = \frac{\frac{1}{\pi}\mathbf{1}_{(-\sqrt{1-x_0^2},\sqrt{1-x_0^2})}(y)}{\frac{1}{\pi}2\sqrt{1 - x_0^2}} \\ &= \begin{cases} \frac{1}{2\sqrt{1-x_0^2}} & \text{für} \quad -\sqrt{1 - x_0^2} < y < \sqrt{1 - x_0^2} \\ 0 & \text{sonst} \end{cases} \end{aligned}$$

Die bedingte Verteilung ist ebenfalls eine Gleichverteilung. Die Höhenlinie entspricht der Dichtefunktion der Gleichverteilung über dem Intervall $-\sqrt{1 - x_0^2} < y < \sqrt{1 - x_0^2}$, d.h. für x_0 ist das Niveau $\frac{1}{2\sqrt{1-x_0^2}}$. $f_Y(y_0)$ entspricht aus Symmetriegründen $f_X(x_0)$. Damit ist $f(x, y) \neq f_X(x)f_Y(y)$ für $(x, y) \in A$. Die Zufallsvariablen sind folglich abhängig.

(b) Wenn (a, b) das Teilintervall der Länge s von $[0; \alpha]$ ist, in dem k von n Werten liegen sollen, dann gilt für T (einmalige Durchführung):

$$P(T \in (a, b)) \quad = \quad F(b) - F(a) = \frac{1}{\alpha}b - \frac{1}{\alpha}a = \frac{b - a}{\alpha} = \frac{s}{\alpha},$$

d.h. die Zufallsvariable $\mathbf{1}_{\{T\in(a,b)\}}$ ist bernoulliverteilt mit $p = \frac{s}{\alpha}$. Die Zufallsvariable X, die die Anzahl der Werte im Intervall (a, b) bei $n-$facher unabhängiger Wiederholung angibt, ist binomialverteilt mit Parameter $p = \frac{s}{\alpha}$, d.h. mit $P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$ lautet die gesuchte Wahrscheinlichkeit:

$$\binom{n}{k} (\frac{s}{\alpha})^k (1 - \frac{s}{\alpha})^{n-k} = \binom{n}{k} \frac{s^k (\alpha - s)^{n-k}}{\alpha^n}$$

Aufgabe 22

Die Anzahl Y der Partikel, die eine Zählkammer in einer bestimmten Zeitspanne erreichen, ist poissonverteilt mit Parameter λ. Von den eintretenden Partikeln wird ein Spannungsstoß ausgelöst, der ein gewisses Vielfaches von Y beträgt. Der zufällig, aber unabhängig von Y erzeugte Multiplikator X hat die Dichtefunktion

$$f(x) = \frac{1}{(1+x)^2} \qquad (x \geq 0) \ .$$

Berechnen Sie die Wahrscheinlichkeit, dass die so entstehende Spannung kleiner als 1 ist.

Lösung: (Wahrscheinlichkeitstheorie, S.194ff.)

Die Poisson-Verteilung lautet: $P(Y = k) = e^{-\lambda} \cdot \frac{\lambda^k}{k!}$ $(k = 0, 1, 2, ...)$. Die Wahrscheinlichkeit, dass der Spannungsstoß kleiner 1 ist, berechnet sich gemäß:

$$\begin{aligned}
P(X \cdot Y < 1) &= P[\bigcup_{k=1}^{\infty} \{Y = k\} \cap \{X < \frac{1}{k}\} \cup \{Y = 0\}] \\
&= \sum_{k=1}^{\infty} P(Y = k \text{ und } \ X < \frac{1}{k}) + P(Y = 0) \\
&= \sum_{k=1}^{\infty} P(Y = k) \cdot P(X < \frac{1}{k}) + P(Y = 0).
\end{aligned}$$

Die erste Gleichung folgt aus der Tatsache, dass $P(X = 0) = 0$ gilt, da X stetig ist; die dritte aus der Unabhängigkeit von X und Y. Für den zweiten Faktor im letzten Ausdruck erhält man:

$$\begin{aligned}
P(X < \frac{1}{k}) &= \int_0^{\frac{1}{k}} \frac{1}{(1+x)^2} dx = (-\frac{1}{1+x})\Big|_0^{\frac{1}{k}} = 1 - \frac{1}{1+\frac{1}{k}} = 1 - \frac{k}{k+1} \\
&= \frac{k+1}{k+1} - \frac{k}{k+1} = \frac{1}{k+1}.
\end{aligned}$$

Damit ergibt sich:

$$\begin{aligned}
P(X \cdot Y < 1) &= \sum_{k=1}^{\infty} P(Y = k) \cdot \frac{1}{k+1} + P(Y = 0) = \sum_{k=1}^{\infty} e^{-\lambda} \cdot \frac{\lambda^k}{k!} \cdot \frac{1}{k+1} + e^{-\lambda} \\
&= e^{-\lambda} \cdot \sum_{k=0}^{\infty} \frac{\lambda^k}{(k+1)!} = e^{-\lambda} \frac{1}{\lambda} \sum_{k=0}^{\infty} \frac{\lambda^{k+1}}{(k+1)!} \\
&= e^{-\lambda} \cdot \frac{1}{\lambda} \cdot \sum_{k=1}^{\infty} \frac{\lambda^k}{k!} = \frac{1}{\lambda} \sum_{k=1}^{\infty} \frac{\lambda^k}{k!} e^{-\lambda} = \frac{1}{\lambda}(1 - \frac{\lambda^0}{0!} e^{-\lambda}) \\
&= \frac{1 - e^{-\lambda}}{\lambda}.
\end{aligned}$$

Aufgabe 23

Bestimmen Sie die Wahrscheinlichkeitsverteilung, den Erwartungswert und die Varianz der Summe einer $B(m,p)$-verteilten Zufallsvariablen X und einer $B(n,p)$-verteilten Zufallsvariablen Y. (X und Y seien unabhängig)

Lösung: (Wahrscheinlichkeitstheorie, S.182ff.)

Gesucht ist die Wahrscheinlichkeitsverteilung der Summe zweier unabhängiger binomialverteilter Zufallsvariablen, es handelt sich also um eine Faltung (diskreter Fall).

$$\begin{array}{ccc} X \sim Bin(n,p) & , & Y \sim Bin(m,p) \\ P(X=x) = \binom{n}{x} p^x (1-p)^{n-x} & , & P(Y=y) = \binom{m}{y} p^y (1-p)^{m-y} \end{array}$$

Für $0 \le z \le n+m$ besitzt $Z = X+Y$ die Wahrscheinlichkeitsverteilung:

$$\begin{aligned} P(Z=z) &= P(X+Y=z) \\ &= \sum_{x,y\geq 0,\ x+y=z} P(X=x)P(Y=y) \\ &= \sum_{x=0}^{z} \binom{n}{x} p^x (1-p)^{n-x} \binom{m}{z-x} p^{z-x} (1-p)^{m-z+x} \\ &= \sum_{x=0}^{z} \binom{n}{x} \binom{m}{z-x} \cdot p^z (1-p)^{m+n-z} \\ &= \binom{n+m}{z} p^z (1-p)^{m+n-z}, \text{ da } \sum_{x=0}^{z} \binom{n}{x}\binom{m}{z-x} = \binom{n+m}{z} \text{ ist.} \end{aligned}$$

Z ist damit $Bin(n+m,p)$-verteilt. Daraus folgt:

$$E(Z) = (n+m)p, \quad Var(Z) = (n+m)p(1-p).$$

Aufgabe 24

(a) Analog zu Aufgabe 18 sei der Ausgang eines Zufallsexperiments durch eine Zufallsvariable X beschrieben. Das Zufallsexperiment wird ohne gegenseitige Beeinflussung n-mal wiederholt. Beschreiben Sie diesen Vorgang durch einen Zufallsvektor $Y = (Y_1 \ldots, Y_n)$ und bestimmen Sie die Verteilungsfunktion von Y, wenn die Verteilungsfunktion von X gegeben ist.

(b) Wie lautet die Verteilungsfunktion, wenn

  (i) X poissonverteilt ist mit Parameter $\lambda > 0$

  (ii) X binomialverteilt ist mit m und p

  (iii) X exponentialverteilt ist mit Parameter $\lambda > 0$

  (iv) X normalverteilt ist mit Mittelwert μ und Varianz σ^2

(c) Wie lautet in den Fällen (b)(i)-(iv) die Verteilungsfunktion von $\sum\limits_{i=1}^{n} Y_i$.

Lösung: (Wahrscheinlichkeitstheorie, S.161ff., S.182ff.)

(a) Sei F_X die Verteilungsfunktion der Zufallsvariable X. Dann gilt aufgrund der Unabhängigkeit für die Verteilungsfunktion F_Y von $Y = (Y_1, \ldots, Y_n)$ für $y = (y_1, \ldots, y_n) \in \mathbb{R}^n$:

$$F_Y(y_1, \ldots, y_n) = F_{Y_1}(y_1) \cdot \cdots \cdot F_{Y_n}(y_n) = \prod_{i=1}^{n} F_X(y_i)$$

(b) Für die angegebenen Verteilungen gilt nun:

(i)

$$F_Y(y_1, \ldots, y_n) = \prod_{i=1}^{n} (\sum_{k=0}^{[y_i]} \frac{\lambda^k}{k!} e^{-\lambda})$$

(ii)

$$F_Y(y_1, \ldots, y_n) = \prod_{i=1}^{n} (\sum_{k=0}^{\min\{[y_i], m\}} \binom{m}{k} p^k (1-p)^{m-k})$$

(iii)

$$F_Y(y_1, \ldots, y_n) = \prod_{i=1}^{n} (1 - e^{-\lambda y_i}) \cdot 1_{[0,\infty)}(y_i) = 1_{[0,\infty)}(\min\{y_i\}) \prod_{i=1}^{n} (1 - e^{-\lambda y_i})$$

(iv)

$$F_Y(y_1, \ldots, y_n) = \prod_{i=1}^{n} \Phi(\frac{y_i - \mu}{\sigma}) = \prod_{i=1}^{n} \int_{-\infty}^{y_i} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$$

(c) (i) eine Summe von n unabhängigen identisch poissonverteilten Zufallsvariablen ist wieder poissonverteilt mit Parameter $n\cdot\lambda$, d.h. für die Verteilungsfunktion von $Z = \sum_{i=1}^{n} Y_i$ gilt:

$$F_Z(z) = \sum_{k=0}^{[z]} \frac{(n\cdot\lambda)^k}{k!} e^{-n\lambda}$$

Für den Fall $n = 2$ folgt:

$$\begin{aligned}
P(Z=z) &= P(Y_1+Y_2=z) = P(\bigcup_{k=0}^{z}\{Y_1=k\}\cap\{Y_2=z-k\}\\
&= \sum_{k=0}^{z} P(\{Y_1=k\}\cap\{Y_2=z-k\})\\
&= \sum_{k=0}^{z} P(Y_1=k)\cdot P(Y_2=z-k)\\
&= \sum_{k=0}^{z} \frac{\lambda^k}{k!}e^{-\lambda}\cdot\frac{\lambda^{z-k}}{(z-k)!}e^{-\lambda}\\
&= (\sum_{k=0}^{z} \frac{\lambda^z}{k!(z-k)!})\cdot e^{-2\lambda}\\
&= \underbrace{(\sum_{k=0}^{z} \frac{z!}{k!(z-k)!})}_{=2^z}\cdot\frac{\lambda^z}{z!}e^{-2\lambda}\\
&= 2^z\cdot\frac{\lambda^z}{z!}\cdot e^{-2\lambda} = \frac{(2\lambda)^z}{z!}\cdot e^{-2\lambda}, \text{ mit } z\in\mathbb{N}
\end{aligned}$$

$$F_Z(z) = \sum_{k=0}^{[z]} P(Z=z) = \sum_{k=0}^{[z]} \frac{(2\lambda)^k}{k!}\cdot e^{-2\lambda}\text{für } z\in\mathbb{R}$$

(ii) Eine Summe von n unabhängigen identisch $B(m,p)$-verteilten Zufallsvariablen ist wieder binomialverteilt mit Parametern $n\cdot m$ und p, d.h. für die Verteilungsfunktion von $Z = \sum_{i=1}^{n} Y_i$ gilt:

$$F_Z(z) = \sum_{k=0}^{\min\{z,n\cdot m\}} \binom{n\cdot m}{k} p^k(1-p)^{n\cdot m-k}$$

(iii) <u>Behauptung</u>:

Die Verteilungsfunktion der Zufallsvariable $Z := Y_1+\ldots+Y_n$, wobei $Y_i \overset{iid}{\sim} Exp(\lambda)$ besitzt die Form

$$F_Z(z) = \begin{cases} 0 & \text{für } z\le 0\\ 1-e^{-\lambda z}\sum_{k=0}^{n-1}\frac{(\lambda z)^k}{k!} & \text{für } z>0\end{cases}$$

<u>Beweis:</u> Die Verteilungsfunktion F_Z besitzt genau dann die angegebene Form, wenn

die Dichte f_Z die folgende Form hat (Produktregel!):

$$f_Z(z) = \begin{cases} 0 & \text{für} \quad z \leq 0 \\ \frac{\lambda^n}{(n-1)!} z^{n-1} e^{-\lambda z} & \text{für} \quad z > 0 \end{cases}$$

Wir beweisen durch vollständige Induktion, dass die Dichte f_n einer Summe von n ($n \in \mathbb{N}$) unabhängigen identisch exponentialverteilter Zufallsvariablen genau diese Gestalt besitzt.

Induktionsanfang: Für $n = 1$ gilt:

$$f_1(x) = \frac{\lambda^1}{0!} x^0 \cdot e^{-\lambda x} \cdot 1_{(0,\infty)}(x) = \begin{cases} \lambda e^{-\lambda x} & \text{für} \quad x > 0 \\ 0 & \text{für} \quad x \leq 0 \end{cases}$$

Induktionsannahme: Angenommen für ein $n \in \mathbb{N}$ gelte:

$$f_n(x) = \begin{cases} 0 & \text{für} \quad x \leq 0 \\ \frac{\lambda^n}{(n-1)!} x^{n-1} e^{-\lambda x} & \text{für} \quad x > 0 \end{cases}$$

dann gilt für f_{n+1} (für $x > 0$) : (Induktionsschritt)

$$\begin{aligned} f_{n+1}(x) &= \int\limits_{-\infty}^{\infty} f_n(t) f_1(x-t) dt = \int\limits_{0}^{x} \frac{\lambda^n}{(n-1)!} t^{n-1} e^{-\lambda t} \cdot \lambda e^{-\lambda(x-t)} dt \\ &= \lambda^{n+1} e^{-\lambda x} \int\limits_{0}^{x} \frac{t^{n-1}}{(n-1)!} dt = \frac{\lambda^{n+1}}{n!} x^n e^{-\lambda x}. \end{aligned}$$

Dies vollendet den Beweis.

(iv) Eine Summe von n unabhängigen $N(\mu, \sigma^2)$-verteilten Zufallsvariablen ist wieder normalverteilt mit Mittelwert $n \cdot \mu$ und Varianz $n \cdot \sigma^2$, d.h. die Verteilungsfunktion von $Z = \sum\limits_{i=1}^{n} Y_i$ lautet:

$$F_Z(z) = \Phi\left(\frac{z - n\mu}{\sqrt{n} \cdot \sigma}\right) = \int\limits_{-\infty}^{z} \frac{1}{\sqrt{2\pi n \sigma^2}} e^{-\frac{(x-n\mu)^2}{2n\sigma^2}} dx$$

Aufgabe 25

Veranschaulichen Sie sich den Begriff *Faltung* anhand der beiden unabhängigen diskreten Zufallsvariablen X und Y mit den Wahrscheinlichkeitsverteilungen

$X = x$	1	2	3
$P(X = x)$	0.3	0.2	0.5

$Y = y$	-1	0	1
$P(Y = y)$	0.3	0.4	0.3

Was ändert sich, wenn X und Y stetige Zufallsvariablen sind?

Lösung: (Wahrscheinlichkeitstheorie, S.182ff.)

Sei (X, Y) eine zweidimensionale Zufallsvariable (X Zufallsvariable auf $(\Omega_1, A(\Omega_1), P_1)$, Y Zufallsvariable auf $(\Omega_2, A(\Omega_2), P_2)$). Die Summe $Z = X + Y$ ist ebenfalls eine Zufallsvariable, die auf dem Wahrscheinlichkeitsraum $(\Omega_1 \times \Omega_2, A(\Omega_1 \times \Omega_2), P_1 \times P_2)$[1] definiert ist.

Anschaulich:

Die Punktmenge (X, Y diskret) bzw. Teilmenge des $\mathbb{R}^2$ (X, Y stetig) wird durch die Bildung der Summe $Z = X + Y$ so zusammengefaltet, dass alle Kombinationen (x, y) von Werten der Zufallsvariable (X, Y), deren Summe identisch ist, zusammengefasst werden (s.Abb.). Die Wahrscheinlichkeit für $Z = z$ ist gleich der Wahrscheinlichkeit dafür, eine Kombination (x, y) aus der Menge der Werte (x, y) mit $x + y = z$ zu erhalten, d.h. für diskrete X, Y:

$$P(Z = z) = \sum_{(x,y):x+y=z} P(X = x,\ Y = y)$$

und bei Unabhängigkeit

$$P(Z = z) = \sum_{(x,y):x+y=z} P(X = x)P(Y = y)$$

$Z = X + Y$ kann die Werte 0, 1, 2, 3 und 4 annehmen. Dabei ist

$$\begin{aligned}
P(Z = 0) &= P(X = 1, Y = -1) = P(X = 1)P(Y = -1) \\
&= 0.3 \cdot 0.3 = 0.09 \\
P(Z = 1) &= P(X = 1, Y = 0) + P(X = 2, Y = -1) \\
&= 0.3 \cdot 0.4 + 0.2 \cdot 0.3 = 0.12 + 0.06 = 0.18 \\
P(Z = 2) &= P(X = 1, Y = 1) + P(X = 2, Y = 0) + P(X = 3, Y = -1) \\
&= 0.3 \cdot 0.3 + 0.2 \cdot 0.4 + 0.5 \cdot 0.3 = 0.09 + 0.08 + 0.15 = 0.32 \\
P(Z = 3) &= P(X = 2, Y = 1) + P(X = 3, Y = 0) \\
&= 0.2 \cdot 0.3 + 0.5 \cdot 0.4 = 0.06 + 0.2 = 0.26 \\
P(Z = 4) &= P(X = 3, Y = 1) = 0.5 \cdot 0.3 = 0.15
\end{aligned}$$

[1] $A(\Omega_1 \times \Omega_2)$ ist die von $A(\Omega_1) \times A(\Omega_2)$ erzeugte $\sigma-$Algebra, d.h. die kleinste $\sigma-$Algebra, die $A(\Omega_1) \times A(\Omega_2)$ als Teilmenge enthält.

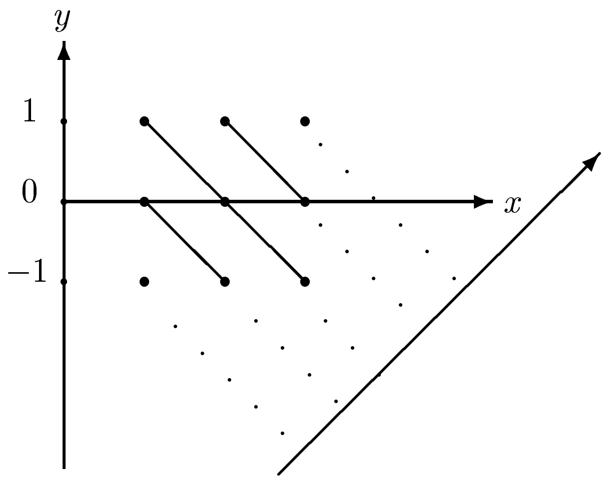


Abbildung 9: Veranschaulichung des "Faltungsprozesses"

Für stetige Zufallsvariablen X und Y gilt:

$$f_Z(z) = \int\limits_{-\infty}^{+\infty} f_{X,Y}(x, z-x)dx$$

bzw. bei Unabhängigkeit von X und Y:

$$f_Z(z) = \int\limits_{-\infty}^{+\infty} f_X(x) \cdot f_Y(z-x)\ dx.$$

Die Verteilungsfunktion erhält man durch

$$F_Z(\alpha) = \int\limits_{-\infty}^{+\alpha} f_Z(z)dz.$$

Aufgabe 26

Die Lebensdauer von einer Anlage sei exponentialverteilt mit Parameter $\lambda_1 > 0$. Bei Ausfall der Anlage wird automatisch und ohne zeitliche Verzögerung ein Notaggregat eingesetzt, dessen Lebensdauer unabhängig von der Lebensdauer der Anlage und ebenfalls exponentialverteilt ist mit einem Parameter λ_2 mit $0 < \lambda_2 \leq \lambda_1$.

(a) Bestimmen Sie die Verteilung der Gesamtlebensdauer des Systems.

(b) Wie lautet die Verteilungsfunktion im Spezialfall $\lambda_2 = \lambda_1$?

(c) Unter der Voraussetzung, dass $\lambda_1 = 2\lambda_2$ ist, ist λ_1 so zu bestimmen, dass eine Gesamtlebensdauer des Systems von mindestens 1000 Zeiteinheiten mit 90% Wahrscheinlichkeit garantiert werden kann.

Lösung: (Wahrscheinlichkeitstheorie, S.182ff.)

Sei X_1 die Zufallsvariable, die die Lebensdauer der Anlage modelliert und X_2 die Zufallsvariable, die die Lebensdauer des Notaggregats beschreibt, dann gilt:

$X_1 \sim \text{Exp}(\lambda_1)$ bzw. $X_2 \sim \text{Exp}(\lambda_2)$

(a) Gesucht ist die Verteilung der Gesamtlebensdauer, d.h. die Verteilung der Zufallsvariablen $Y = X_1 + X_2$. Mit Hilfe der Formel aus Aufgabe 24 ergibt sich für die Dichte $f_Y(y)$ der Zufallsvariablen Y für $y \geq 0$, wobei $\lambda_2 < \lambda_1$:

$$\begin{aligned}
f_Y(y) &= \int_{-\infty}^{\infty} f_{X_1}(x) f_{X_2}(y-x)dx \\
&= \int_{-\infty}^{\infty} \lambda_1 e^{-\lambda_1 x} \cdot \mathbf{1}_{[0,\infty)}(x) \cdot \lambda_2 e^{-\lambda_2(y-x)} \cdot \mathbf{1}_{[0,\infty)}(y-x)dx \\
&= \int_{-\infty}^{\infty} \lambda_1\lambda_2 e^{-\lambda_1 x - \lambda_2 y + \lambda_2 x} \cdot \mathbf{1}_{[0,\infty)}(\min\{x, y-x\})dx \\
&= \lambda_1\lambda_2 e^{-\lambda_2 y} \int_0^y e^{(-\lambda_1+\lambda_2)x} dx \quad (*) \\
&= \lambda_1\lambda_2 e^{-\lambda_2 y} \left(\frac{1}{\lambda_2 - \lambda_1} e^{(\lambda_2-\lambda_1)x} \right)\Bigg|_0^y \\
&= \lambda_1\lambda_2 e^{-\lambda_2 y} \cdot \left(\frac{1}{\lambda_2 - \lambda_1} e^{(\lambda_2-\lambda_1)y} - \frac{1}{\lambda_2 - \lambda_1} \right) \\
&= \frac{\lambda_1\lambda_2}{\lambda_2 - \lambda_1} e^{-\lambda_1 y} - \frac{\lambda_1\lambda_2}{\lambda_2 - \lambda_1} e^{-\lambda_2 y} \\
&= \frac{\lambda_1\lambda_2}{\lambda_2 - \lambda_1} (e^{-\lambda_1 y} - e^{-\lambda_2 y})
\end{aligned}$$

d.h. die Verteilungsfunktion ergibt sich gemäß:

$$\begin{aligned}
F_Y(y) &= \int_0^y \frac{\lambda_1\lambda_2}{\lambda_2-\lambda_1}(e^{-\lambda_1 t} - e^{-\lambda_2 t})dt \\
&= \frac{\lambda_1\lambda_2}{\lambda_2-\lambda_1}\left(-\frac{1}{\lambda_1}e^{-\lambda_1 t} + \frac{1}{\lambda_2}e^{-\lambda_2 t}\right)\Bigg|_0^y \\
&= \frac{\lambda_1\lambda_2}{\lambda_2-\lambda_1}\left(-\frac{1}{\lambda_1}e^{-\lambda_1 y} + \frac{1}{\lambda_2}e^{-\lambda_2 y} + \frac{1}{\lambda_1} - \frac{1}{\lambda_2}\right) \\
&= \frac{\lambda_1\lambda_2}{\lambda_2-\lambda_1}\left(\frac{1}{\lambda_2}e^{-\lambda_2 y} - \frac{1}{\lambda_1}e^{-\lambda_1 y} + \frac{\lambda_2-\lambda_1}{\lambda_1\lambda_2}\right) \\
&= 1 + \frac{\lambda_1}{\lambda_2-\lambda_1}e^{-\lambda_2 y} - \frac{\lambda_2}{\lambda_2-\lambda_1}e^{-\lambda_1 y}
\end{aligned}$$

(b) Für $\lambda = \lambda_1 = \lambda_2$ ergibt sich aus $(*)$ $(\int_0^y dx = y)$:

$$\begin{aligned}
f_Y(y) &= \lambda^2 y e^{-\lambda y} \cdot \mathbf{1}_{[0,\infty)}(y) \\
\text{bzw. } F_Y(y) &= \int_0^y \lambda^2 t e^{-\lambda t} dt = -\lambda t e^{-\lambda t}\Big|_0^y + \int_0^y \lambda e^{-\lambda t} dt \\
&= -\lambda y e^{-\lambda y} - e^{-\lambda t}\Big|_0^y = -\lambda y e^{-\lambda y} - e^{-\lambda y} + 1 \\
&= 1 - (\lambda y + 1)e^{-\lambda y} = 1 - e^{-\lambda y}(1 + \lambda y)
\end{aligned}$$

Es handelt sich hierbei um die Erlang-Verteilung mit Stufenzahl 2.

(c) Ansatz:

$$\begin{aligned}
P[Y \geq 1000] &\geq 0.9 \\
\Leftrightarrow 1 - P[Y < 1000] &\geq 0.9 \\
\Leftrightarrow P[Y < 1000] &\leq 0.1 \\
F_Z(1000) &\leq 0.1 \Rightarrow \text{ Hinreichend } F_Z(1000) = 0.1, \text{ mit } \lambda_1 = 2\lambda_2
\end{aligned}$$

d.h.:

$$\begin{aligned}
1 + \frac{2\lambda_2}{\lambda_2 - 2\lambda_2}e^{-\lambda_2 \cdot 1000} - \frac{\lambda_2}{\lambda_2 - 2\lambda_2}e^{-2\lambda_2 \cdot 1000} &= 0.1 \\
\Longleftrightarrow 0.9 - 2e^{-1000\cdot\lambda_2} + e^{-2000\cdot\lambda_2} &= 0 \\
\overset{z:=e^{-1000\cdot\lambda_2}}{\Longleftrightarrow} z^2 - 2z + 0.9 &= 0 \\
\Longleftrightarrow z_{1/2} &= 1 \pm \sqrt{0.1}
\end{aligned}$$

$z_1 = 1 + \sqrt{0.1} = e^{-1000\lambda_2}$ ergibt ein negatives λ_2, also keine Lösung, d.h

$$\lambda_2 = -\frac{\ln z}{1000} = -\frac{\ln(1-\sqrt{0.1})}{1000} = 0.00038,$$

also $\lambda_1 \approx 0.00076$.

Aufgabe 27

Die Lebensdauern T_1 und T_2 zweier elektrischer Bauteile B_1 und B_2 seien exponentialverteilt mit den Parametern $\lambda_1 = \frac{1}{500}$ und $\lambda_2 = \frac{1}{300}$ und unabhängig.

(a) Berechnen Sie die Wahrscheinlichkeit dafür, dass B_1 bzw. B_2 den Zeitpunkt $t_0 = 200$ überlebt, wenn das jeweilige Bauteil zur Zeit $t = 0$ eingesetzt wurde. Wie groß ist die Wahrscheinlichkeit dafür, dass das jeweilige Bauteil nach Erreichen von t_0 noch weitere 200 Stunden arbeitet?

(b) Bestimmen Sie die Lebensdauerverteilung eines aus B_1 und B_2 bestehenden Reihen- bzw. Parallelsystems.

(c) Bestimmen Sie die Wahrscheinlichkeiten dafür, dass das Reihen- bzw. das Parallelsystem den Zeitpunkt $t = 200$ überlebt.

Lösung: (Wahrscheinlichkeitstheorie, S.190f., S.189f.)

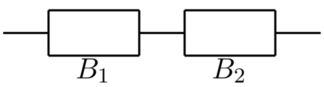
(a) Verteilungsfunktion der Exponentialverteilung ist:

$$F(t) = \begin{cases} 0 & \text{für } t < 0 \\ 1 - e^{-\lambda \cdot t} & \text{für } t \geq 0 \end{cases}$$

$$\begin{aligned} P(T_1 \geq 200) &= 1 - F(200) = e^{-\frac{200}{500}} \\ &= 0.6703 \quad (= P(T_1 \geq 400 \,|\, T_1 > 200)) \\ P(T_2 \geq 200) &= e^{-\frac{200}{300}} \\ &= 0.5134 \quad (= P(T_2 \geq 400 \,|\, T_2 > 200)) \end{aligned}$$

Die Exponentialverteilung ist eine Verteilung ohne Gedächtnis, vgl. Aufgabe 8.

(b) Das Reihensystem

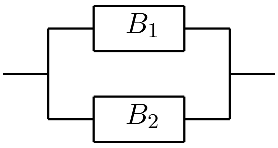


ist intakt, wenn beide Bauteile intakt sind, d.h. $T_s = \min\{T_1, T_2\}$ ist der Zeitpunkt des Systemausfalls.

$$\begin{aligned} P(T_s \leq t) &= P(\min\{T_1, T_2\} \leq t) = 1 - P(\min\{T_1, T_2\} > t) \\ &= 1 - P(T_1 > t \text{ und } T_2 > t) = 1 - P(T_1 > t) \cdot P(T_2 > t) \quad \text{(Unabhängigkeit!)} \\ &= 1 - (1 - F_{T_1}(t)) \cdot (1 - F_{T_2}(t)) = 1 - e^{-\lambda_1 \cdot t} e^{-\lambda_2 \cdot t} \\ &= 1 - e^{-(\lambda_1 + \lambda_2) \cdot t} = F_{T_s}(t) \end{aligned}$$

d.h. das Minimum und damit die Lebensdauer des Reihensystems ist exponentialverteilt mit dem Parameter $(\lambda_1 + \lambda_2)$.

Das Parallellsystem



ist intakt, solange mindestens eines der beiden Bauteile intakt ist, d.h. $T_s = \max\{T_1, T_2\}$ ist der Zeitpunkt des Systemausfalls.

$$\begin{aligned}
P(T_s \leq t) &= P(\max\{T_1, T_2\} \leq t) = P(T_1 < t \text{ und } T_2 < t) \\
&= P(T_1 \leq t) \cdot P(T_2 \leq t) = F_{T_1}(t) \cdot F_{T_2}(t) \\
&= (1 - e^{-\lambda_1 \cdot t}) \cdot (1 - e^{-\lambda_2 \cdot t}) = 1 - e^{-\lambda_1 \cdot t} - e^{-\lambda_2 \cdot t} + e^{-(\lambda_1 + \lambda_2) \cdot t} \\
&= F_{T_s}(t)
\end{aligned}$$

d.h. das Maximum und damit die Lebensdauer des Parallelsystems ist nicht exponentialverteilt.

(c) Die Überlebenswahrscheinlichkeit des Reihensystem ist:

$$\begin{aligned}
P(T_s \geq 200) &= 1 - P(T_s < 200) = e^{-(\frac{1}{500} + \frac{1}{300}) \cdot 200} = e^{-\frac{800}{150000} \cdot 200} = e^{-\frac{16}{15}} \\
&= 0.3442.
\end{aligned}$$

Die Überlebenswahrscheinlichkeit des Parallellsystem ist:

$$\begin{aligned}
P(T_s \geq 200) &= 1 - P(T_s < 200) = e^{-\frac{1}{500} \cdot 200} + e^{-\frac{1}{300} \cdot 200} - e^{-(\frac{1}{500} + \frac{1}{300}) \cdot 200} \\
&= 0.8396.
\end{aligned}$$

Aufgabe 28

Die Zufallsvariable X sei gleichverteilt auf dem Intervall $[1,3]$. $Y = (Y_1, \dots, Y_n)$ sei eine Stichprobe mit Zurücklegen zu X.

(a) Bestimmen Sie die Verteilung von Y und die Randverteilung von Y_i, $i = 1, \dots, n$.

(b) Berechnen Sie die Dichte der Spannweite R von Y.

(c) Geben Sie den Erwartungswert von R an.

Lösung: (Wahrscheinlichkeitstheorie, S.161ff., S.139ff., S.191ff., S.79ff.)

(a) Für die gemeinsame Verteilungsfunktion F_Y gilt:

$$F_Y(y_1, \dots, y_n) = \prod_{i=1}^{n} F_X(y_i)$$

Da X gleichverteilt ist, gilt:

$$f_X(x) = \begin{cases} \frac{1}{2} & \text{für} \quad x \in [1,3] \\ 0 & \text{sonst} \end{cases}$$

$$F_X(\alpha) = \begin{cases} 0 & \text{für} \quad \alpha \in (-\infty, 1] \\ \frac{1}{2}(\alpha - 1) & \text{für} \quad \alpha \in [1,3] \\ 1 & \text{sonst} \end{cases}$$

Damit ergibt sich:

$$F_Y(y_1, \dots, y_n) = \begin{cases} 0 & \text{für} \quad y_i < 1 \text{ für mindestens ein } i \in \{1, \dots, n\} \\ \frac{1}{2^n} \prod_{i=1}^{n} (y_i - 1) & \text{für} \quad y_i \in [1,3] \text{ für alle } i = 1, \dots, n \\ \frac{1}{2^{\#I}} \prod_{i \in I} (y_i - 1) & \text{für} \quad y_i \in [1,3] \text{ für } i \in I \subseteq \{1, \dots, n\} \text{ und } y_i > 3 \\ & \text{für} \quad \text{alle } i \in \{1, \dots, n\} \setminus I \\ 1 & \text{sonst} \end{cases}$$

und als Randverteilung von Y_i, $i = 1, \ldots, n$:

$$\begin{aligned}
F_{Y_i}(y_i) &= \lim_{y_j \to \infty; j \neq i} F_Y(y_1, \ldots, y_n) = F_X(y_i) \\
&= \begin{cases} 0 & \text{falls } y_i < 1 \\ \frac{1}{2}(y_i - 1) & \text{für } y_i \in [1,3] \\ 1 & y_i \in [3, \infty) \end{cases}
\end{aligned}$$

(b) $Y = (Y_1, \ldots, Y_n)$ ist nach Voraussetzung eine Stichprobe mit Zurücklegen zu X. Damit ist die gemeinsame Dichte von $\max\{Y_i\}$ und $\min\{Y_i\}$: (s. Wahrscheinlichkeitstheorie, S. 204)

$$\begin{aligned}
f(x,y) &= \begin{cases} n(n-1)(\frac{1}{2}(x-y))^{n-2}\frac{1}{2}\frac{1}{2} & \text{für} \quad 1 \leq y < x < 3 \\ 0 & \text{sonst} \end{cases} \\
&= \begin{cases} (\frac{1}{2})^n n(n-1)(x-y)^{n-2} & \text{für} \quad 1 \leq y < x < 3 \\ 0 & \text{sonst} \end{cases}
\end{aligned}$$

Die Dichte der Spannweite ergibt sich daraus zu:

$$\begin{aligned}
f_R(z) &= \int_{-\infty}^{\infty} f(x, x-z) dx \\
&= \begin{cases} 0 & \text{für} \quad z \leq 0 \text{ (da sonst } y = x - z \geq x) \\ \int_{1+z(*)}^{3} (\frac{1}{2})^n n(n-1)(x-(x-z))^{n-2}\, dx & \text{für} \quad 0 < z < 2 \\ 0 & \text{für} \quad z \geq 2 \text{ (da sonst mit } x < 3 \text{ gilt,} \\ & \qquad \text{dass } x - z < 1 \text{ ist.)} \end{cases}
\end{aligned}$$

Bemerkung zu $(*)$: Aus $1 \leq y = x - z$ folgt $x \geq z + 1$.

Das Integral im zweiten Funktionsast ergibt:

$$\begin{aligned}
\int_{z+1}^{3} (\frac{1}{2})^n n(n-1)(x-(x-z))^{n-2}\, dx &= (\frac{1}{2})^n n(n-1) \int_{z+1}^{3} z^{n-2}\, dx \\
&= (\frac{1}{2})^n n(n-1) \left(xz^{n-2}\right)\Big|_{z+1}^{3} \\
&= (\frac{1}{2})^n n(n-1)(2-z)z^{n-2}
\end{aligned}$$

(c) Der Erwartungswert berechnet sich gemäß:

$$
\begin{aligned}
E(R) &= \int\limits_{-\infty}^{+\infty} z f_R(z)\,dz = \int\limits_0^2 z(\frac{1}{2})^n n(n-1)(2-z)z^{n-2}\,dz \\
&= (\frac{1}{2})^n n(n-1) \int\limits_0^2 (2z - z^2) z^{n-2}\,dz = (\frac{1}{2})^n n(n-1) \int\limits_0^2 (2z^{n-1} - z^n)\,dz \\
&= (\frac{1}{2})^n n(n-1) \left(\frac{2}{n} z^n - \frac{1}{n+1} z^{n+1} \right) \Bigg|_0^2 = (\frac{1}{2})^n n(n-1) \left(\frac{2}{n} 2^n - \frac{1}{n+1} 2^{n+1} \right) \\
&= 2n(n-1) \left(\frac{1}{n} - \frac{1}{n+1} \right) = 2\frac{n-1}{n+1}
\end{aligned}
$$

Aufgabe 29

(a) Gegeben seien zwei unabhängige diskrete Zufallsvariablen X und Y mit den in der Aufgabe 25 definierten Wahrscheinlichkeitsverteilungen. Bestimmen Sie die Wahrscheinlichkeitsverteilung von $Z = XY$.

(b) Betrachten Sie nun die beiden unabhängigen stetigen Zufallsvariablen X und Y mit den Dichtefunktionen

$$f_X(x) = \begin{cases} \frac{2}{15}x & \text{für} \quad 1 \leq x \leq 4 \\ 0 & \quad \text{sonst} \end{cases},$$

$$f_Y(y) = \begin{cases} \frac{1}{12}y & \text{für} \quad 1 \leq y \leq 5 \\ 0 & \quad \text{sonst} \end{cases}.$$

(i) Bestimmen Sie die Dichtefunktion des Produkts $Z = XY$.

(ii) Überlegen Sie sich, wie Sie die Dichte von $Z = \frac{X}{Y}$ berechnen können.

Lösung: (Wahrscheinlichkeitstheorie, S.194ff.)

(a) Für das Produkt erhält man analog zur Summe im diskreten Fall:

$$P(\prod_{i=1}^{n} Y_i = z) = \sum_{(y_1,\ldots,y_n):\prod y_i = z} P((Y_1 \ldots Y_n) = (y_1 \ldots y_n))$$

Die Werte der Zufallsvariable XY mit X und Y aus Aufgabe 25 ergeben sich gemäß folgender Tabelle

$y \setminus x$	1	2	3
-1	-1	-2	-3
0	0	0	0
1	1	2	3

zu: -3,-2,-1,0,1,2,3.

Die Wahrscheinlichkeitsverteilung ergibt sich nach der Tabelle zu

$$P(Z = -3) = P(X \cdot Y = -3) = P((X, Y) = (3, -1)) = 0.15$$

und analog für die Werte -2,-1,1,2,3 sowie

$$P(Z = 0) = P(X, Y) \in \{(1, 0), (2, 0), (3, 0)\} = 0.4.$$

Insgesamt lautet die Wahrscheinlichkeitsverteilung damit:

z	-3	-2	-1	0	1	2	3
$P(Z=z)$	0.15	0.06	0.09	0.4	0.09	0.06	0.15

(b) (i) Dichtefunktion für $Z = X \cdot Y$ (s. Wahrscheinlichkeitstheorie, S. 207) ist:

$$f_Z(z) \;=\; \int_{-\infty}^{\infty} f_{(X,Y)}(x, \frac{z}{x})\frac{1}{|x|}\, dx$$

bzw. da X und Y unabhängig sind, ist die Dichtefunktion:

$$f_Z(z) \;=\; \int_{-\infty}^{\infty} f_X(x) f_Y(\frac{z}{x})\frac{1}{|x|}\, dx$$

Für welche Werte von x zu vorgegebenem z ist der Integrand von Null verschieden? Dies ist genau dann der Fall, wenn $f_X(x) \neq 0$ ist (also für $x \in [1,4]$) und wenn $f_Y(\frac{z}{x}) \neq 0$ ist (also für $\frac{z}{x} \in [1,5]$). Damit muß $x \geq 1$ und $\frac{z}{x} \leq 5$ d.h. $x \geq \frac{z}{5}$ sein, also $x \geq \max\{1, \frac{z}{5}\}$. Ferner folgt aus $x \leq 4$ und $\frac{z}{x} \geq 1$, d.h. $x \leq z$, $x \leq \min\{4, z\}$. Damit ergibt sich für die Dichtefunktion:

$$\begin{aligned}
f_Z(z) &= \int_{\max\{1,\frac{z}{5}\}}^{\min\{4,z\}} \frac{2}{15}x \cdot \frac{1}{12}\frac{z}{x}\frac{1}{x}\, dx \text{ für } z = x \cdot y \in [1 \cdot 1 = 1,\ 4 \cdot 5 = 20] \\
&= \int_{\max\{1;\frac{z}{5}\}}^{\min\{4,z\}} \frac{1}{90}\frac{z}{x} dx \\
&= \frac{1}{90} z(\ln x)\Big|_{\max\{1,\frac{z}{5}\}}^{\min\{4,z\}} \\
&= \frac{1}{90} z(\ln(\min\{4,z\}) - \ln(\max\{1, \frac{z}{5}\}))
\end{aligned}$$

$$\Rightarrow f_Z(z) \;=\; \begin{cases}
0 & \text{für} \quad z < 1 \quad (\max\{1,\frac{z}{5}\} = 1 > \min\{4,z\} = z) \\
\frac{1}{90} z(\ln z - 0) & \text{für} \quad z \in [1,4) \quad (\max\{1,\frac{z}{5}\} = 1, \min\{4,z\} = z) \\
\frac{1}{90} z(\ln 4 - 0) & \text{für} \quad z \in [4,5) \quad (\max\{1,\frac{z}{5}\} = 1, \min\{4,z\} = 4) \\
\frac{1}{90} z(\ln 4 - \ln\frac{z}{5}) & \text{für} \quad z \in [5,20] \quad (\max\{1,\frac{z}{5}\} = \frac{z}{5}, \min\{4,z\} = 4) \\
0 & \text{für} \quad z > 20 \quad (\max\{1,\frac{z}{5}\} = \frac{z}{5} > \min\{4,z\} = 4)
\end{cases}$$

(ii) Für den Quotienten Z zweier Zufallsvariablen X und Y gilt: $Z = X \cdot (\frac{1}{Y})$, d.h. im wesentlichen kann wie beim Produkt vorgegangen werden, wobei für die Zufallsvariable Y der Wert y=0 ausgeschlossen werden muss.

Die Dichte ergibt sich also zu (vgl. Wahrscheinlichkeitstheorie S. 203):

$$f_Z(z) = \int_{-\infty}^{\infty} f(zy, y)|y|\, dy$$

Bei Unabhängigkeit von X und Y also $f_Z(z) = \int\limits_{-\infty}^{\infty} f_X(zy) f_Y(y)|y|\,dy$

Hier sind:

$$f_X(x) \neq 0 \text{ für } x \in [1,4] \text{ und } f_Y(y) \neq 0 \text{ für } y \in [1,5].$$

Der Integrand ist genau dann positiv, falls $zy \in [1,4]$ und $y \in [1,5]$ d.h. $y \in [\frac{1}{z}, \frac{4}{z}]$ und $y \in [1,5]$ ist. Damit gilt:

$$\begin{aligned}
f_Z(z) &= \int\limits_{\max\{1;\frac{1}{z}\}}^{\min\{\frac{4}{z},5\}} f_X(zy) f_Y(y)|y|\,dy \\
&= \int\limits_{\max\{1,\frac{1}{z}\}}^{\min\{\frac{4}{z},5\}} \frac{1}{90} zy^3\,dy \\
&= \frac{1}{360} zy^4 \Big|_{\max\{1,\frac{1}{z}\}}^{\min\{\frac{4}{z},5\}}
\end{aligned}$$

und somit

$$f_Z(z) = \begin{cases}
0 & , z < \frac{1}{5} \text{ (dann ist } \frac{1}{z} = \max\{1, \frac{1}{z}\} > \min\{\frac{4}{z}, 5\} = 5) \\
\frac{1}{360} z \left(5^4 - (\frac{1}{z})^4\right) & , z \in [\frac{1}{5}, \frac{4}{5}] \ (\max\{1, \frac{1}{z}\} = \frac{1}{z}, \min\{\frac{4}{z}, 5\} = 5) \\
\frac{1}{360} z \left((\frac{4}{z})^4 - (\frac{1}{z})^4\right) & , z \in [\frac{4}{5}, 1] \ (\max\{1, \frac{1}{z}\} = \frac{1}{z}, \min\{\frac{4}{z}, 5\} = \frac{4}{z}) \\
\frac{1}{360} z \left((\frac{4}{z})^4 - (1)^4\right) & , z \in [1, 4] \ (\max\{1, \frac{1}{z}\} = 1, \min\{\frac{4}{z}, 5\} = \frac{4}{z}) \\
0 & , z > 4 \text{ (dann ist } \max\{1, \frac{1}{z}\} = 1 > \min\{\frac{4}{z}, 5\} = \frac{4}{z})
\end{cases}$$

Aufgabe 30

(a) Gegeben sei eine diskrete Zufallsvariable $X = (X_1, X_2)$, deren Komponenten unabhängig sind mit den im folgenden angegebenen Randverteilungen. Bestimmen Sie die Wahrscheinlichkeitsverteilung einer Zufallsvariablen $Y = (Y_1, Y_2)$ unter der Annahme, dass $Y_1 = X_1$ und $Y_2 = X_1 + X_2$ gilt.

x_1	-1	0	1
$P(X_1 = x_1)$	0.3	0.4	0.3

x_2	-1	0	1
$P(X_2 = x_2)$	0.3	0.4	0.3

(b) Sei X_1 eine auf dem Intervall $[1, 4]$ gleichverteilte Zufallsvariable und X_2 eine Zufallsvariable mit folgender Dichtefunktion:

$$f_{X_2}(x_2) = \begin{cases} 2 - 0.5 \cdot x_2 & \text{für} \quad 2 \leq x_2 \leq 4 \\ 0 & \text{sonst} \end{cases} .$$

(i) Bestimmen Sie die Dichtefunktion der zweidimensionalen Zufallsvariablen $Y = (Y_1, Y_2)$ unter der Annahme, dass X_1 und X_2 unabhängig sind und $Y_1 = X_1$ sowie $Y_2 = X_1 + X_2$ gilt.

(ii) Bestimmen Sie den Erwartungswert von Y.

Lösung: (Wahrscheinlichkeitstheorie, S.199ff.)

(a) $Y = (Y_1, Y_2)$ entsteht durch die Transformation $g(x_1, x_2) = (x_1, x_1 + x_2)$ aus der Zufallsvariable X =(X_1, X_2). Die Funktion g ist umkehrbar eindeutig, d.h. die Wahrscheinlichkeitsverteilung von $g(X)$ kann wie folgt berechnet werden:

$$P(g(X) = y) = P(X = g^{-1}(y))$$

Werte von Y_2 sind:

$Y_2 =$	-2	-1	0	1	2
entsteht aus $(X_1, X_2) =$	(-1,-1)	(-1,0) (0,-1)	(-1,1) (1,-1) (0;0)	(1,0) (0,1)	(1,1)

Damit ergibt sich unter der Berücksichtigung von

$$P((X_1, X_2) = (x_1, x_2)) = P(X_1 = x_1)P(X_2 = x_2)$$

folgende Wahrscheinlichkeitsverteilung für Y:

(y_1, y_2)	(-1,-2)	(-1,-1)	(-1,0)	(0,-1)	(0,0)	(0,1)	(1,0)	(1,1)	(1,2)
$g^{-1}(y_1, y_2)$	(-1,-1)	(-1,0)	(-1,1)	(0,-1)	(0,0)	(0,1)	(1,-1)	(1,0)	(1,1)
$P(X = g^{-1}(y))$	0.3^2	$0.3 \cdot 0.4$	0.3^2	$0.4 \cdot 0.3$	0.4^2	$0.4 \cdot 0.3$	0.3^2	$0.3 \cdot 0.4$	0.3^2

bzw. in anderer Darstellung:

y_2 / y_1	-2	-1	0	1	2
-1	0.09	0.12	0.09	0	0
0	0	0.12	0.16	0.12	0
1	0	0	0.09	0.12	0.09

Die einzelnen Komponenten berechnen sich z.B. wie folgt:

$$\begin{aligned}
((Y_1, Y_2) = (1, 0)) &= P(Y_1 = 1 \text{ und } Y_2 = 0) = P(X_1 = 1 \text{ und } X_1 + X_2 = 0) \\
&= P(X_1 = 1 \text{ und } X_2 = -1) = (0.3)^2 \\
P((Y_1, Y_2) = (1, -2)) &= P(Y_1 = 1 \text{ und } Y_2 = -2) = P(X_1 = 1 \text{ und } X_1 + X_2 = -2) \\
&= P(X_1 = 1 \text{ und } X_2 = -3) = 0.3 \cdot 0 = 0.
\end{aligned}$$

(b) Gegeben sind die folgenden beiden Dichtefunktionen:

$$f_{X_1}(x_1) = \begin{cases} \frac{1}{3} & \text{für} \quad x_1 \in [1, 4] \\ 0 & \text{sonst} \end{cases}$$

$$f_{X_2}(x_2) = \begin{cases} 2 - 0.5x_2 & \text{für} \quad x_2 \in [2, 4] \\ 0 & \text{sonst} \end{cases}$$

Da X_1 und X_2 unabhängig sind, gilt: $f_{(X_1, X_2)}(x_1, x_2) = f_{X_1}(x_1) f_{X_2}(x_2)$, d.h.

$$f_{(X_1, X_2)}(x_1, x_2) = \begin{cases} \frac{1}{3}(2 - 0.5x_2) & \text{für} \quad x_1 \in [1, 4] \text{ und } x_2 \in [2, 4] \\ 0 & \text{sonst} \end{cases}$$

Die Transformation lautet:

$$X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \mapsto Y = \begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} = \begin{pmatrix} X_1 \\ X_1 + X_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$$

Es handelt sich hierbei um eine lineare Transformation. Die Transformation $g(x_1, x_2) = (g_1(x_1, x_2), g_2(x_1, x_2))$ ist daher gegeben durch

$$\begin{aligned}
y_1 &= g_1(x_1, x_2) = x_1 \\
y_2 &= g_2(x_1, x_2) = x_1 + x_2
\end{aligned}$$

Die Ableitung von g lautet:

$$\begin{aligned}
g'(a) &= \begin{pmatrix} \frac{\partial g_1(a)}{\partial x_1} & \frac{\partial g_1(a)}{\partial x_2} \\ \frac{\partial g_2(a)}{\partial x_1} & \frac{\partial g_2(a)}{\partial x_2} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \text{ für alle } a \in [1, 4] \times [2, 4], \text{ d.h.} \\
\det g'(a) &= 1 \neq 0 \text{ für } a \in [1, 4] \times [2, 4]
\end{aligned}$$

Die Umkehrtransformation lautet:

$$\begin{aligned} x_1 &= g_1^{-1}(y_1, y_2) &= y_1 \\ x_2 &= g_2^{-1}(y_1, y_2) &= y_2 - y_1 \end{aligned}$$

Hieraus ergibt sich nun die transformierte Dichte gemäß dem Transformationssatz:

$$\begin{aligned} f_{(Y_1,Y_2)}(y) &= \begin{cases} f_{X_1,X_2}(g^{-1}(y_1,y_2))\dfrac{1}{|\det(g'(g^{-1}(y)))|} & \text{für } g_1^{-1}(y_1,y_2) = y_1 \in [1,4] \text{ und} \\ & \quad g_2^{-1}(y_1,y_2) = y_2 - y_1 \in [2,4] \\ 0 & \text{sonst} \end{cases} \\ &= \begin{cases} \frac{1}{3}(2 - 0.5(y_2 - y_1))\frac{1}{|1|} & \text{für } y_1 \in [1,4] \text{ und } y_2 \in [2 + y_1, 4 + y_1] \\ 0 & \text{sonst} \end{cases} \\ &= \begin{cases} \frac{2}{3} - \frac{1}{6}(y_2 - y_1) & \text{für } y_1 \in [1,4] \text{ und } y_2 \in [2 + y_1, 4 + y_1] \\ 0 & \text{sonst} \end{cases} \end{aligned}$$

Bekanntlich gilt: $E(Y) = (E(Y_1), E(Y_2))$ d.h. der Erwartungswert von Y kann komponentenweise gebildet werden. Da $Y_1 = X_1$ ist, folgt:

$$E(Y_1) = E(X_1) = 2.5$$

Der Erwartungswert von Y_2 kann gemäß

$$E(Y_2) = E(X_1 + X_2) = E(X_1) + E(X_2)$$

aus dem Erwartungswert von X_1 und X_2 berechnet werden. Für den zweiten Summand erhält man:

$$\begin{aligned} E(X_2) &= \int_2^4 x_2(2 - 0.5x_2)\,dx_2 = x_2^2 - \frac{1}{6}x_2^3\Big|_2^4 \\ &= 16 - \frac{32}{3} - 4 + \frac{4}{3} = \frac{8}{3} \end{aligned}$$

und somit gilt:

$$E(Y_2) = \frac{5}{2} + \frac{8}{3} = \frac{31}{6}$$

Aufgabe 31

Sei $X = (X_1, X_2)$ eine bivariate Zufallsvariable mit folgender Dichtefunktion:

$$f_X(x_1, x_2) = \begin{cases} 0.1 \cdot x_1 x_2 & \text{für} \quad 1 \leq x_1 \leq 3,\ 2 \leq x_2 \leq 3 \\ \\ 0 & \text{sonst} \end{cases}.$$

Bestimmen Sie die Dichtefunktion der Zufallsvariablen $Y = (Y_1, Y_2)$ mit $Y_1 = X_1$ und $Y_2 = X_1 \cdot X_2$.

Lösung: (Wahrscheinlichkeitstheorie, S.199ff.)

Die Transformation h mit $y = h(x)$ besitzt die folgende Form:

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} h_1(x_1, x_2) \\ h_2(x_1, x_2) \end{pmatrix} = \begin{pmatrix} x_1 \\ x_1 \cdot x_2 \end{pmatrix}$$

Die partiellen Ableitungen existieren und es gilt:

$$\det(h'(x)) = \det \begin{pmatrix} \frac{\partial h_1}{\partial x_1} & \frac{\partial h_1}{\partial x_2} \\ \frac{\partial h_2}{\partial x_1} & \frac{\partial h_2}{\partial x_2} \end{pmatrix} = \det \begin{pmatrix} 1 & 0 \\ x_2 & x_1 \end{pmatrix} = x_1 \neq 0 \quad \text{für } x_1 \in [1, 3]$$

Die Umkehrtransformation lautet:

$$x_1 = h_1^{-1}(y_1, y_2) = y_1; \ x_2 = h_2^{-1}(y_1, y_2) = \frac{y_2}{y_1}$$

Damit ist $\det(h'(h^{-1}(y)) = h_1^{-1}(y_1, y_2) = y_1$ und für $y_1 = h_1^{-1}(y_1, y_2) \in [1, 3]$ und $\frac{y_2}{y_1} = h_2^{-1}(y_1, y_2) \in [2, 3]$ gilt:

$$\begin{aligned} f_Y(y_1, y_2) &= f_X(h_1^{-1}(y_1, y_2), h_2^{-1}(y_1, y_2)) \cdot \frac{1}{|\det(h'(h^{-1}(y_1, y_2)))|} \\ &= 0.1 \cdot y_1 \frac{y_2}{y_1} \cdot \frac{1}{y_1} = 0.1 \cdot \frac{y_2}{y_1} \end{aligned}$$

Die Dichtefunktion von (Y_1, Y_2) ist somit:

$$f_Y(y_1, y_2) = \begin{cases} 0.1 \cdot \frac{y_2}{y_1} & \text{für } y_1 \in [1, 3],\ y_2 \in [2y_1, 3y_1] \\ 0 & \text{sonst} \end{cases}$$

Aufgabe 32

Eine Maschine produziert Stahlstifte mit einer Soll-Länge von $\mu = 110$ mm. Ungenauigkeiten bei der Produktion können leider nicht ausgeschlossen werden, so dass die Längen der produzierten Stifte Realisationen einer Zufallsvariablen X darstellen. Die Streuung der realisierten Stiftlängen ist mit $\sigma = 0.1$ mm bekannt.

Um sich ein Bild von der Güte der Produktionseinheiten zu machen, wird die Wahrscheinlichkeit dafür, dass X höchstens um den Betrag c von μ abweicht, betrachtet. c wird dabei in der Regel in Vielfachen der Standardabweichung σ des Prozesses angegeben, d.h. gesucht ist $P(X \in [\mu - k\sigma, \mu + k\sigma])$.

Bestimmen und vergleichen Sie für $k = 1, 2, 3$ die Wahrscheinlichkeiten für die Fälle, dass

(a) über den Typ der Verteilung von X keine weiteren Kenntnisse verfügbar sind bzw.

(b) X normalverteilt ist mit Mittelwert μ und Varianz σ^2.

Wie erklären Sie sich Ihre Ergebnisse?

Lösung: (Wahrscheinlichkeitstheorie, S.208ff.)

(a) Da keine genaue Kenntnisse über den Verteilungstyp vorliegt, erfolgt die Abschätzung über die Tschebyscheffsche Ungleichung:

$$\begin{aligned} P(X \in [\mu - k\sigma; \mu + k\sigma]) &= P(|X - \mu| \leq k\sigma) \\ &= 1 - P(|X - \mu| > k\sigma) \\ &\overset{Tschebyscheff}{\geq} 1 - \frac{Var(X)}{k^2\sigma^2} = 1 - \frac{1}{k^2} \end{aligned}$$

Mit $k = 1, 2, 3$ ergibt sich:

$$\begin{aligned} P(X \in [\mu - \sigma; \mu + \sigma]) &\geq 1 - \frac{1}{1} = 0 \\ P(X \in [\mu - 2\sigma; \mu + 2\sigma]) &\geq 1 - \frac{1}{4} = 0.75 \\ P(X \in [\mu - 3\sigma; \mu + 3\sigma]) &\geq 1 - \frac{1}{9} = 0.\overline{8} \ (= \frac{8}{9}) \end{aligned}$$

Für $k = 1$ ist die Abschätzung über Tschebyscheff trivial und ohne Aussagekraft.

(b) Ist die Verteilung konkret bekannt, so lassen sich die Wahrscheinlichkeiten genau berechnen. Es gilt nun: $X \sim N(110, 0.1^2)$.

$$\begin{aligned} P(X \in [\mu - k\sigma; \mu + k\sigma]) &= \Phi\left(\frac{\mu + k\sigma - \mu}{\sigma}\right) - \Phi\left(\frac{\mu - k\sigma - \mu}{\sigma}\right) \\ &= \Phi(k) - \Phi(-k) = 2\Phi(k) - 1 \end{aligned}$$

Für die konkreten Werte in der Aufgabe ergibt sich:

$$\begin{aligned} P(X \in [\mu - \sigma; \mu + \sigma]) &= 2\Phi(1) - 1 = 0.6826 \\ P(X \in [\mu - 2\sigma; \mu + 2\sigma]) &= 2\Phi(2) - 1 = 0.9544 \\ P(X \in [\mu - 3\sigma; \mu + 3\sigma]) &= 2\Phi(3) - 1 = 0.9974 \end{aligned}$$

Die Abschätzungen in (a) gelten für jede Verteilung von X. Daher ist es auch, wenn die Verteilung nicht bekannt ist, nur möglich eine untere Schranke für die Wahrscheinlichkeit anzugeben. Ist die Verteilung bekannt, so ist die Wahrscheinlichkeit (zumindest mit numerischer Genauigkeit) exakt anzugeben. Der Wert wird auf jeden Fall größer oder gleich der unteren Schranke sein. Bei der Normalverteilung erhalten wir erheblich größere Werte wegen der Symmetrie um μ und der hohen Dichtewerte im Bereich um μ.

Aufgabe 33

(a) Bei einem Großraumflugzeug ist die Auslastung pro Flug näherungsweise normalverteilt. Im Mittel fliegen 150 Passagiere mit dem Flugzeug, die Auslastung schwankt mit $\sigma = 25$ Passagiere. Mit welcher Anzahl von Fluggästen ist mit einer Wahrscheinlichkeit von 90% mindestens zu rechnen?

(b) Welche Abschätzung können Sie vornehmen, wenn der Typ der Verteilung nicht bekannt ist?

Lösung: (Wahrscheinlichkeitstheorie, S.86ff., S.102f., S.208f.)

(a) Die Anzahl X der Passagiere pro Flug ist näherungsweise $N(150; 25^2)$-verteilt. Gesucht ist: x_0 mit $P(X \geq x_0) \stackrel{!}{=} 0.9$. Sei Φ die Verteilungsfunktion der Standardnormalverteilung, dann gilt

$$\begin{aligned} P(X \geq x_0) &= 1 - P(X \leq x_0) = 1 - F_X(x_0) \\ &= 1 - \Phi\left(\frac{x_0 - \mu}{\sigma}\right) = 1 - \Phi\left(\frac{x_0 - 150}{25}\right) \stackrel{!}{=} 0.9 \end{aligned}$$

Gesucht ist also x_0 mit $\Phi(\frac{x_0-150}{25}) \stackrel{!}{=} 0.1$. Man erhält:

$$\begin{aligned} \frac{x_0 - 150}{25} &= \Phi^{-1}(0.1) = -1.28 \; (= -\Phi^{-1}(1-0.1)) \\ \Rightarrow x_0 &= -1.28 \cdot 25 + 150 = 118 \end{aligned}$$

Es ist also mit einer Wahrscheinlichkeit von 90% mit mindestens 118 Passagieren zu rechnen.

(b) Ist die Verteilung nicht bekannt, so kann eine Abschätzung mit der Tschebyscheffschen Ungleichung durchgeführt werden. Gesucht ist also c mit

$$P(|X - 150| < c) \geq 1 - \frac{Var(X)}{c^2} \stackrel{!}{=} 0.9$$

Man erhält:

$$1 - \frac{Var(X)}{c^2} = 0.9 \text{ und somit } c = \sqrt{\frac{Var(X)}{0.1}} = \frac{25}{\sqrt{0.1}} = 79.06$$

Daraus folgt: Mit mindestens 90% Wahrscheinlichkeit gilt: $|X - 150| < 79.06$. Da X ganzzahlig ist, ist dies gleichwertig mit $X \in (70, 230)$, d.h. mit mindestens 90% Wahrscheinlichkeit gilt: $70 < X < 230$.

Bemerkung: Eine genauere Abschätzung ist mit Hilfe der Tschebyscheffschen Ungleichung nicht möglich, da über die Verteilung nichts bekannt ist! Insbesondere kann die Wahrscheinlichkeit für das gesuchte einseitig offene Intervall nur sehr ungenau bestimmt werden.

Aufgabe 34

(a) Erläutern Sie anhand eines idealen Würfels das "schwache Gesetz der großen Zahlen".

(b) Verdeutlichen Sie sich die wesentlichen Aussagen des zentralen Grenzwertsatzes.

(c) Eine Vertriebsgesellschaft besitzt in einer Großstadt 200 Zigarettenautomaten. Jeder Automat hat (unabhängig von den anderen) mit der Wahrscheinlichkeit $\frac{1}{20}$ pro Woche eine Störung. Für die Entscheidung über die Größe eines ständigen Reparaturtrupps sei die Wahrscheinlichkeit dafür von Interesse, dass in einer Woche die Anzahl X der defekten Automaten zwischen 5 und 15 liegt, von Interesse. Diese Wahrscheinlichkeit (der exakte Wert beträgt übrigens 0.9292) soll

  (i) mittels der Poissonverteilung approximiert werden,

  (ii) über die Ungleichung von Tschebyscheff nach unten abgeschätzt werden und

  (iii) mit Hilfe des zentralen Grenzwertsatzes approximativ berechnet werden.

Lösung: (Wahrscheinlichkeitstheorie, S.212f., S.217ff., S.208f.)

(a) Das schwache Gesetz der großen Zahlen bezogen auf einen idealen Würfel besagt:

Bei wiederholtem unabhängigen Werfen eines Würfels konvergiert das arithmetische Mittel der gewürfelten Zahlen stochastisch gegen den Erwartungswert ($E(X) = 3.5$). D.h. für großes n ist die Wahrscheinlichkeit für eine größere Abweichung des arithmetischen Mittels vom Erwartungswert klein, oder formal:

Es bezeichne X_i, $i = 1, 2, \ldots$ eine Folge unabhängiger identisch verteilter Zufallsvariablen mit einem endlichen Erwartungswert. Dann gilt:

$$\lim_{n\to\infty} P\left(\left|\frac{1}{n}\sum_{i=1}^{n} X_i - EX\right| > \varepsilon\right) = 0 \;\forall \varepsilon > 0$$

(b) Die Aussage des Zentralen Grenzwertsatzes ist:

Seien $X_1, X_2, X_3, \ldots$ unabhängige Zufallsvariablen, deren Erwartungswerte $\mu_i = E(X_i)$ und Varianzen $\sigma_i^2 = Var(X_i)$ existieren ($i = 1, \ldots, n$). Falls $X_1, X_2, \ldots$ die Lindeberg-Bedingung erfüllen, gilt für die Verteilungsfunktion F_{Z_n} der Folge Z_n mit

$$Z_n = \frac{\sum\limits_{i=1}^{n} X_i - \sum\limits_{i=1}^{n} \mu_i}{\sqrt{\sum\limits_{i=1}^{n} \sigma_i^2}}$$

dass für alle $\alpha \in \mathbb{R}$

$$\lim_{n\to\infty} F_{Z_n}(\alpha) = \Phi(\alpha) = \int_{-\infty}^{\alpha} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\, dx$$

gilt und die Konvergenz gleichmäßig in α ist. Die Lindeberg- Bedingung lautet:

Mit $D_n^2 = \sum\limits_{i=1}^{n} \sigma_i^2$ gelte für jedes $\varepsilon > 0$:

$$\lim_{n\to\infty} \frac{1}{D_n^2} \sum_{i=1}^{n} \int\limits_{|x_i-\mu_i|>\varepsilon D_n} (x_i - \mu_i)^2 \; dF_i(x) = 0.$$

Interpretation: Jeder der unabhängigen Summanden X_i trägt zur Summe nur einen kleinen Anteil bei, d.h. kein Summand dominiert. Dies gilt z.B. insbesondere falls die X_i identisch verteilt sind.

Die Summe n unabhängiger Zufallsvariablen kann für großes n als näherungsweise normalverteilt angenommen werden, wenn einzelne Summanden keinen "prägenden" Einfluß auf die Summe haben.

Anwendung: z.B. im Bereich der Qualitätssicherung, wo unsystematische Abweichungen einer überwachten Größe von ihrem Sollwert in der Regel als Ergebnis einer Überlagerung vieler zufälliger, für sich unbedeutender Abweichungen aufgefaßt werden (andere Anwendungen z.B. in der Kapitalmarkttheorie, ...)

(c) Die Zufallsvariable X gebe die Anzahl defekter Automaten in einer Woche an, d.h. gesucht ist $P(5 \le X \le 15)$. X ist binomialverteilt mit Parametern 200 und 0.05, d.h.

$$P(5 \le X \le 15) = \sum_{k=5}^{15} \binom{200}{k} (\frac{1}{20})^k (\frac{19}{20})^{200-k} = F_X(15) - F_X(4)$$

(i) Approximation über die Poissonverteilung (gute Näherung, falls $n \ge 50, p \le \frac{1}{10}$; Parameter $\lambda := n \cdot p =$ mit $\lambda =$ Erwartungswert von $Poi(\lambda)$ und $n \cdot p =$ Erwartungswert von $Bin(n,p)$), d.h. wir betrachten X als näherungsweise Poisson-verteilt mit $\lambda = n \cdot p$.

$$P(X = x) = \begin{cases} \frac{(np)^x}{x!} e^{-np} & \text{für } x = 0, 1, 2, \ldots \\ 0 & \text{sonst} \end{cases}$$

Damit ist

$$P(5 \le X \le 15) = \sum_{k=5}^{15} \frac{10^k}{k!} e^{-10} = 0.992$$

(ii) Schranke mit Hilfe der Tschebyscheffschen Ungleichung:

$$P(|X - E(X)| < c) \ge 1 - \frac{Var(X)}{c^2}$$

$$\begin{aligned} E(X) &= n \cdot p = 10 \\ Var(X) &= n \cdot p(1-p) = 9.5 \\ \Longrightarrow P(X \in [5; 15]) &\overset{*}{=} P(|X - 10| < 6) \ge 1 - \frac{9.5}{6^2} = 0.736 \; (* \text{ da } X \text{ diskret ist}) \\ \text{bzw. } P(|X - 10| \le 5) &\ge P(|X - 10| < 5) \ge 1 - \frac{9,5}{5^2} \\ &= 0.62 \end{aligned}$$

d.h. die Schranke sollte so groß wie möglich gewählt werden, um eine möglichst gute Abschätzung zu erhalten. Auch hier zeigt sich die ziemlich grobe Abschätzung bei der Tschebyscheffschen Ungleichung.

(iii) Approximation durch die Normalverteilung: X_i sei Bernoulli-verteilt für $i = 1, \dots, n$. Dann ist $\sum_{i=1}^n X_i$ näherungsweise $N(\mu, \sigma^2)-$verteilt mit $\mu = np$ und $\sigma^2 = np(1-p)$, falls n hinreichend groß ist. (Merkregel: relativ gute Abschätzung für $np \geq 5, n(1-p) \geq 5$). Mit $\mu = 10$ und $\sigma^2 = 9.5$ gilt für $X = \sum X_i$:

$$\begin{aligned} P(X \in [5; 15]) &= P(X \leq 15) - P(X < 5) = P(X \leq 15.5) - P(X \leq 4.5) \\ &\approx^{1)} \Phi(\frac{15.5 - 10}{3.0822}) - \Phi(\frac{4.5 - 10}{3.0822}) = 0.9248. \end{aligned}$$

[1)] Stetigkeitskorrektur, d.h. beim Übergang von diskret nach stetig wird die Approximation verbessert, wenn als Schranken nicht 15 oder 5 sondern 15.5 und 4.5 gewählt werden, was bei der diskreten Zufallsvariable keine Auswirkung hat.

Aufgabe 35

Ein Unternehmen stellt Taschenlampenbatterien her. Für ein Marketingprojekt soll die Lebensdauer dieser Batterien mit einer Stichprobe untersucht werden, wobei eine Exponentialverteilung unterstellt wird. Geben Sie zu diesem Beispiel die drei Grundannahmen der schließenden Statistik an und erläutern Sie die drei Grundaufgaben. Geben Sie zu jeder der Grundaufgaben eine konkrete Fragestellung für das Unternehmen an, die zu dieser Aufgabe führt, und wie eine Entscheidungsfunktion dazu aussieht. Geben Sie jeweils eine sinnvoll erscheinende Entscheidungsfunktion an.

Lösung: (Induktive Statistik, S.33ff., S.40ff.)

Grundannahme 1: Der relevante Umweltzustand kann durch eine Zufallsvariable Y beschrieben werden.

Hier: Y sei die Lebensdauer einer zufällig ausgewählten Batterie.

Grundannahme 2: Die Verteilung von Y gehört einer bekannten Klasse von Wahrscheinlichkeitsverteilungen an.

Hier: Klasse der Exponentialverteilungen: $\{Exp(\lambda) | \lambda > 0\}$

Grundannahme 3: Zu Y existiert eine Stichprobe $(X_1, \ldots, X_n)$, deren Verteilung in bekannter Weise von der Verteilung von Y abhängt.

Hier: Verteilung einer einfachen Stichprobe $(X_1, \ldots, X_n)$ vom Umfang n zu Y:

$$P(X_1 \leq x_1, \ldots, X_n \leq x_n) = \prod_{i=1}^{n} P(X_i \leq x_i) = \prod_{i=1}^{n} P(Y \leq x_i)$$

Grundaufgabe 1: Parameterschätzung

Hier: Man interessiert sich für den unbekannten Parameter λ, da er die Verteilung von Y eindeutig festlegt, d.h. gesucht ist ein Schätzwert $\hat{\lambda}$ für λ anhand der Stichprobenrealisation $(x_1, \ldots, x_n)$.

Grundaufgabe 2: Intervallschätzung

Hier: Anstatt einen konkreten Schätzwert für λ anzugeben, sucht man einen Bereich, der den wahren Parameter mit vorgegebener Wahrscheinlichkeit enthält.

Grundaufgabe 3: Hypothesentests

Hier: Gegeben sind zwei Hypothesen über den unbekannten Parameter λ (z.B. $H_0 : \lambda \leq \lambda_0,\ H_1 : \lambda > \lambda_0$). Auf der Grundlage der Stichprobenrealisation ist eine Entscheidung über die Gültigkeit der Hypothesen zu treffen.

Entscheidungsfunktion zu 1):

$$\delta \ : \ \begin{array}{rcl} \mathcal{X} & \to & \Gamma \\ (x_1, \ldots, x_n) & \mapsto & \delta(x_1, \ldots, x_n) \end{array}$$

hierbei bezeichnet $\mathcal{X}$ den Stichprobenraum und Γ den Parameterraum.

Entscheidungsfunktion zu 2):

$$\delta \quad : \quad \begin{array}{rcl} \mathcal{X} & \rightarrow & \mathcal{P}(\Gamma) \\ (x_1, \ldots, x_n) & \mapsto & \delta(x_1, \ldots, x_n) = A \subseteq \Gamma \end{array}$$

wobei $P(\lambda \in A) \geq 1 - \alpha$ zu vorgegebenem α sei.

Entscheidungsfunktion zu 3)

$$\delta \quad : \quad \begin{array}{rcl} \mathcal{X} & \rightarrow & \{d_0, d_1\} \\ (x_1, \ldots, x_n) & \mapsto & \delta(x_1, \ldots, x_n) \end{array}$$

wobei d_0 und d_1 die Entscheidungen "Hypothese H_0 kann nicht abgelehnt werden bzw. H_0 wird angenommen" und "H_0 wird abgelehnt" symbolisieren.

Aufgabe 36

Ein Briefmarkensammler weiß, dass von einem bestimmten Ersttagsbrief eine gewisse Auflagenhöhe existiert. Von k dieser Ersttagsbriefe, die wie gewöhnlich mit 1 beginnend durchnumeriert wurden, kennt er per Zufall die Nummern $n_1, \ldots, n_k$. Aufgrund dieser Information möchte er einen Schätzwert für die Gesamtauflage N des Ersttagsbriefes erstellen. Begründen Sie, wie die im folgenden vorgeschlagenen Schätzer zustande gekommen sind:

(a) $N_1 = \max\limits_i n_i$

(b) $N_2 = \frac{k+1}{k} \max\limits_i n_i$

(c) $N_3 = \max\limits_i n_i + \min\limits_i n_i - 1$

(d) $N_4 = 2 \frac{1}{k} \sum\limits_{i=1}^{k} n_i$

(e) $N_5 = 2 \frac{1}{k} \sum\limits_{i=1}^{k} n_i - 1$

Lösung: (Induktive Statistik, S.40, S.61ff.)

Die Grundgesamtheit ist die Menge der Erstagsbriefe, $\Omega = \{\omega_1, \ldots, \omega_N\}$, N Auflagehöhe der Ersttagsbriefe, betrachtet wird der Laplacesche Wahrscheinlichkeitsraum mit Grundgesamtheit Ω sowie die Zufallsvariable Y, die zu jedem Erstagsbrief die zugehörige Nummer angibt: $Y(\omega_i) = i$ für $i = 1, \ldots, N$. Die bekannten Nummern $n_1, \ldots, n_k$ der Ersttagsbriefe werden als Stichprobenrealisation angesehen, mit deren Hilfe ein Schätzwert für N anzugeben sei.

(a) Schätzung von N mittels größtem Wert in der Stichprobe (ein kleinerer Wert wäre unsinnig).

(b) Hier liegt folgende Überlegung zugrunde: N_1 würde N nur treffen, wenn der größte Wert zufällig in der Stichprobe wäre, was sehr selten der Fall ist. Man wählt daher einen Korrekturfaktor "nach oben". Nachteil: i.a. keine ganze Zahl.

(c) Die Ersttagsbriefe sind folgendermaßen numeriert:
$1, \ldots, \ \min n_i, \ldots, \ \max n_i, \ldots, N$
Man nimmt an, dass der Abstand zwischen dem kleinsten Stichprobenwert und der 1 gleich groß ist, wie der Abstand zwischen der größten Stichprobennummer und dem Wert N, den man schätzen will. Hier wird sowohl der niedrigste als auch der höchste vorliegende Wert berücksichtigt, es wird also mehr Information in diesem Schätzer verarbeitet, als in (a) und (b).

(d) Ähnlich wie bei (c), nur dass jetzt die gesamte Stichprobeninformation verwendet wird. Es wird hier der Mittelwert der Stichprobe errechnet und angenommen, dass dieser auch das Mittel der Gesamtpopulation ist, d.h. dass rechts von ihm genausoviele Nummern liegen, wie links von ihm.

(e) Korrektutterm " -1 " da die Briefe von 1 und nicht von 0 an durchnummeriert sind. Ansonsten, wie (d).

Aufgabe 37

(a) Eine Zufallsvariable sei im Bereich $[0, a]$ gleichverteilt, a ist nicht bekannt und soll mit Hilfe einer Stichprobe mit Zurücklegen geschätzt werden. Überlegen Sie, was die für diese Aufgabe wesentliche Information der Stichprobenwerte $x_1, \ldots, x_n$ ist. Überprüfen Sie, ob diese Stichprobenfunktion suffizient ist.

(b) Eine Zufallsvariable X habe die Dichtefunktion

$$f(x) \quad = \quad \begin{cases} 0 & \text{für} \quad x < a \\ \\ \lambda e^{-\lambda(x-a)} & \text{für} \quad x \geq a \end{cases}$$

mit $a > 0$ und bekanntem λ. Lösen Sie die zu (a) analoge Aufgabe.

Lösung: (Induktive Statistik, S.47ff.)

(a) Wesentliche Information ist der größte Stichprobenwert. Die Dichtefunktion des Stichprobenvektors ist

$$\begin{aligned} f(x_1, \ldots, x_n) \quad &= \quad \prod_{i=1}^{n} \frac{1}{a} \mathbf{1}_{[0,a]}(x_i) = \frac{1}{a^n} \prod_{i=1}^{n} \mathbf{1}_{[0,a]}(x_i) \\ &= \quad \begin{cases} \frac{1}{a^n} & \text{für} \quad \min x_i \geq 0, \max x_i \leq a \\ \\ 0 & \text{sonst} \end{cases} \\ &= \quad \frac{1}{a^n} \mathbf{1}_{[0,a]}(\max_{i=1,\ldots,n} x_i) \cdot \mathbf{1}_{[0,\infty)}(\min_{i=1,\ldots,n} x_i) \\ &= \quad g(\max_{i=1,\ldots,n} x_i, a) \cdot h(x_1, \ldots, x_n) \end{aligned}$$

mit

$$g(\max x_i, a) = \frac{1}{a^n} \mathbf{1}_{[0,a]}(\max_{i=1,\ldots,n} x_i)$$

und

$$h(x_1, \ldots, x_n) = \mathbf{1}_{[0,\infty)}(\min_{i=1,\ldots,n} x_i)$$

oder

$$h(x_1, \ldots, x_n) = \prod_{i=1}^{n} \mathbf{1}_{[0,\infty)}(x_i)$$

$\max x_i$ ist also eine suffiziente Stichprobenfunktion.

(b) Wesentliche Information über a ist der kleinste der Stichprobenwerte. (Man beachte dabei, dass λ als bekannt vorausgesetzt wurde. Betrachten Sie den Fall "λ und a unbekannt" als Übungsaufgabe!). Die Dichtefunktion des Stichprobenvektors ist:

$$\begin{aligned} f(x_1,\ldots,x_n) &= \prod_{i=1}^{n} \lambda e^{-\lambda(x_i-a)} \mathbf{1}_{[a,\infty)}(x_i) = \lambda^n e^{-\lambda \sum\limits_{i=1}^{n}(x_i-a)} \mathbf{1}_{[a,\infty)}(\min_{i=1,\ldots n} x_i) \\ &= \lambda^n e^{-\lambda \sum\limits_{i=1}^{n} x_i} e^{n\lambda a} \mathbf{1}_{[a,\infty)}(\min x_i) = g(\min_{i=1,\ldots n} x_i, a) h(x_1,\ldots,x_n) \end{aligned}$$

mit

$$g(\min_{i=1,\ldots n} x_i, a) = e^{n\lambda a} 1_{[a,\infty)}(\min_{i=1,\ldots n} x_i)$$

und

$$h(x_1,\ldots,x_n) = \lambda^n e^{-\lambda \sum\limits_{i=1}^{n} x_i}.$$

$\min\limits_{i=1,\ldots n} x_i$ ist also eine suffiziente Statistik.

Aufgabe 38

Bestimmen Sie mit Hilfe des Faktorisierungstheorems von Neyman suffiziente Statistiken für

(a) die Klasse der Exponentialverteilungen

(b) die Klasse der Gleichverteilungen auf dem Intervall $[-\vartheta, \vartheta]$ $(\vartheta > 0)$.

(c) die Klasse der Binomialverteilungen $B(m, p)$

Lösung: (Induktive Statistik, S.47ff.)

Gegeben eine einfache Stichprobe $X = (X_1, \ldots, X_n)$ zu einer Zufallsvariable Y.

(a) Y sei λ-exponentialverteilt.

Dichtefunktion zu X ist

$$\begin{aligned}
f_X(x_1, \ldots, x_n) &= \prod_{i=1}^{n} \lambda e^{-\lambda x_i} \mathbf{1}_{[0,\infty)}(x_i) \\
&= \begin{cases} \lambda^n e^{-\lambda \sum\limits_{i=1}^{n} x_i} \cdot 1 \text{ für} & x_i > 0 \text{ für alle } i = 1, \ldots, n \\ 0 & \text{sonst} \end{cases} \\
&= g\left(\sum_{i=1}^{n} x_i, \lambda\right) \cdot h(x)
\end{aligned}$$

mit

$$\begin{aligned}
h(x_1, \ldots, x_n) &= \begin{cases} 1 & \text{für} \quad x_i > 0 \quad \text{für} \quad \text{alle } i = 1, \ldots, n \\ 0 & \text{für} \quad x_i \leq 0 \quad \text{für} \quad \text{mindestens ein } i \end{cases} \\
&= h(x_1, \ldots, x_n) = \mathbf{1}_{[0,\infty)}(\min x_i) = \prod_{i=1}^{n} \mathbf{1}_{[0,\infty)}(x_i).
\end{aligned}$$

Also ist $T(X) = \sum\limits_{i=1}^{n} x_i$ suffiziente Statistik bzgl. λ.

(b) Y gleichverteilt auf $[-\vartheta, \vartheta], \vartheta > 0$.

Dichtefunktion von X ist:

$$\begin{aligned}
f_X(x_1,\ldots,x_n) &= \prod_{i=1}^{n} \frac{1}{2\vartheta}\mathbf{1}_{[-\vartheta,\vartheta]}(x_i) \\
&= \frac{1}{(2\vartheta)^n}\prod_{i=1}^{n}\mathbf{1}_{[0,\vartheta]}(|x_i|) \\
&= \begin{cases} 0 & \max(|x_i|) > \vartheta \\ \frac{1}{(2\vartheta)^n} & \max(|x_i|) \leq \vartheta \end{cases} \\
&= \frac{1}{(2\vartheta)^n}\mathbf{1}_{[0,\vartheta]}(\max|x_i|) \\
&= g(\max|x_i|,\vartheta)\cdot\ h(x_1,\ldots,x_n) \text{ mit } h(x_1,\ldots,x_n) = 1
\end{aligned}$$

$T(X) = \max\limits_i |x_i|$ ist suffiziente Statistik bzgl ϑ.

(c) Y binomialverteilt

Wahrscheinlichkeitsverteilung von X:

$$\begin{aligned}
P(X_1 = k_1, X_2 = k_2, \cdots, X_n = k_n) &= \prod_{i=1}^{n}\binom{m}{k_i}p^{k_i}(1-p)^{m-k_i} \\
&= p^{\sum\limits_{i=1}^{n} k_i}(1-p)^{n\cdot m-\sum\limits_{i=1}^{n} k_i}\cdot\prod_{i=1}^{n}\binom{m}{k_i} \\
&= g\left(\sum_{i=1}^{n} k_i, p\right)\cdot h(k_1,\cdots,k_n)
\end{aligned}$$

$T(k_1,\cdots,k_n) = \sum\limits_{i=1}^{n} k_i$ ist suffiziente Statistik bezüglich p.

Aufgabe 39

Zeigen Sie, dass z.B. die diskrete Verteilung $B(m,p)$ eine Exponentialfamilie bildet. Wie lauten die daraus resultierenden suffizienten Statistiken?

Lösung: (Induktive Statistik, S.55ff.)

Die Wahrscheinlichkeitsverteilung der Binomialverteilung lautet:

$$
\begin{aligned}
P(Y=k) &= \binom{m}{k} p^k (1-p)^{m-k} = \binom{m}{k} (\frac{p}{1-p})^k (1-p)^m \\
&= (1-p)^m \cdot e^{k \ln \frac{p}{1-p}} \cdot \binom{m}{k} \\
&= a(p) \cdot e^{b(p) \cdot \tau(k)} \cdot h(k)
\end{aligned}
$$

Die suffiziente Statistik ist: $T(X) = \sum\limits_{i=1}^{n} k_i$ (vgl. Aufgabe 38 c), wobei $k_1, \ldots, k_n$ die Stichprobenergebnisse sind.

Aufgabe 40

(a) Der Torschützenkönig der Fußball-WM trifft an der Torwand mit Wahrscheinlichkeit p. Mit einer Stichprobe vom Umfang 2 soll p geschätzt werden. Es wird folgende Schätzfunktion vorgeschlagen:

Bei 0 Treffern lautet der Schätzwert $\hat{p} = 0$
Bei 1 Treffer lautet der Schätzwert $\hat{p} = \frac{1}{2}$
Bei 2 Treffern lautet der Schätzwert $\hat{p} = 1$

Wie lautet der Erwartungswert bei diesem Schätzverfahren? Versuchen Sie eine Schätzfunktion ausgehend von der Trefferzahl zu bestimmen, die ebenfalls diesen Erwartungswert besitzt, aber verschieden von der hier vorgeschlagenen ist.

(b) Die Zufallsvariable X sei Weibull-verteilt mit Parametern α ($\alpha > 0$) und β, d.h. ihre Dichtefunktion lautet:

$$f_X(x) = \begin{cases} \alpha\beta x^{\beta-1} e^{-\alpha x^\beta} & \text{für} \quad x > 0 \\ 0 & \quad \text{sonst} \end{cases} .$$

(i) Zeichnen Sie die Dichtefunktion für einige Parameterkombinationen (α, β). Welchen Einfluß haben die Parameter auf die Gestalt der Dichtefunktion?

(ii) Der Wert von β sei nun mit $\beta = 3$ bekannt, α soll mit Hilfe einer einfachen Stichprobe vom Umfang n geschätzt werden. Bestimmen Sie eine suffiziente Statistik für α. Ist die Statistik vollständig?

Lösung: (Induktive Statistik, S.61ff., S.47ff.)

(a) Mit

$$X_i = \begin{cases} 1 & \text{für} \quad \text{Treffer bei Schuss } i \\ \\ 0 & \text{für} \quad \text{kein Treffer bei Schuss } i \end{cases}$$

ist $T(X_1, X_2) = X_1 + X_2$ die Trefferzahl. Die vorgeschlagene Schätzfunktion lautet:

$$\delta_1(k) = \begin{cases} 0 & \text{für} \quad k = 0 \\ \\ \frac{1}{2} & \text{für} \quad k = 1 \\ \\ 1 & \text{für} \quad k = 2 \end{cases}$$

$X_1 + X_2$ ist binomialverteilt mit $n = 2$ und Parameter p. Damit ist

$$\begin{aligned} E(\delta_1(X_1 + X_2)) &= \sum_{k=0}^{2} \delta_1(k) P(X_1 + X_2 = k) \\ &= 0\binom{2}{0} p^0 (1-p)^2 + \frac{1}{2}\binom{2}{1} p(1-p) + 1\binom{2}{2} p^2 (1-p)^0 \\ &= \frac{1}{2} \cdot 2 \cdot p(1-p) + p^2 = p. \end{aligned}$$

Die Schätzfunktion hat also als Erwartungswert den wahren Wert p. Sei δ_2 eine weitere Schätzfunktion ausgehend von der Trefferzahl k, so gilt:

$$\begin{aligned}
E(\delta_2(X_1+X_2)) &= \sum_{k=0}^{2} \delta_2(k)P(X_1+X_2=k) \\
&= \delta_2(0)\binom{2}{0}p^0(1-p)^2 + \delta_2(1)\binom{2}{1}p(1-p) + \delta_2(2)\binom{2}{2}p^2(1-p)^0 \\
&= \delta_2(0)(1-p)^2 + \delta_2(1)\cdot 2\cdot p(1-p) + \delta_2(2)p^2 \\
&= p^2(\delta_2(0) - 2\delta_2(1) + \delta_2(2)) + p(-2\delta_2(0) + 2\delta_2(1)) + \delta_2(0).
\end{aligned}$$

Die Forderung nach übereinstimmendem Erwartungswert lautet damit

$$p^2(\delta_2(0) - 2\delta_2(1) + \delta_2(2)) + 2p(-\delta_2(0) + \delta_2(1)) + \delta_2(0) = p \text{ für alle } p \in [0,1].$$

Für $p = 0$ folgt daraus $\delta_2(0) = 0$. Mit $p = 1$ ergibt sich weiter $-2\delta_2(1) + \delta_2(2) + 2\delta_2(1) = 1$ und folglich $\delta_2(2) = 1$. Mit $p = \frac{1}{2}$ folgt letzendlich

$$\begin{aligned}
\frac{1}{4}(-2\delta_2(1)+1) + 2\frac{1}{2}\delta_2(1) &= \frac{1}{2} \\
\Longleftrightarrow -\frac{1}{2}\delta_2(1) + \delta_2(1) &= \frac{1}{2} - \frac{1}{4} = \frac{1}{4} \\
\Longleftrightarrow \delta_2(1) &= \frac{1}{2}.
\end{aligned}$$

Es gibt also keine weitere Schätzfunktion, die als Erwartungswert den zu Grunde liegenden Wert des Parameters liefert.

(b) Die Dichtefunktion der Weibull-Verteilung in allgemeinerer Form und anderer Parametrisierung ist:

$$f(x) = \begin{cases} \frac{c}{b}(\frac{x-a}{b})^{c-1} e^{-(\frac{x-a}{b})^c} & x \geq a \\ 0 & \text{sonst} \end{cases}$$

mit $a \in \mathbb{R}$, $b, c \in \mathbb{R}_+$. Die Parametriesierung der Aufgabenstellung erhält man hieraus für: $a = 0,\ c = \beta,\ b^{-\beta} = b^{-c} = \alpha$.

(i)

Für $\beta = 1$ und $a = 0$ ist:

$$f(x) = \begin{cases} \alpha e^{-\alpha x} & \text{für} \quad x \geq 0 \\ 0 & \text{sonst} \end{cases}$$

bzw.

$$f(x) = \begin{cases} \frac{1}{b} e^{-\frac{x}{b}} & \text{für} \quad x \geq 0 \\ 0 & \text{sonst} \end{cases}$$

die Dichtefunktion der Exponentialverteilung mit Parameter $\alpha > 0$ (bzw. $\frac{1}{b} > 0$). Eine Veränderung von β führt zu Streckung oder Stauchung bzw. von a zur Verschiebung der Dichte. Die Weibullverteilung wird z.B. in der Zuverlässigkeitstheorie zur Modellierung von Verschleißerscheinungen (genauer: Lebensdauern), verwendet[2]. Sei

[2] weitere Anwendungen gibt es unter anderem in der Warteschlangentheorie

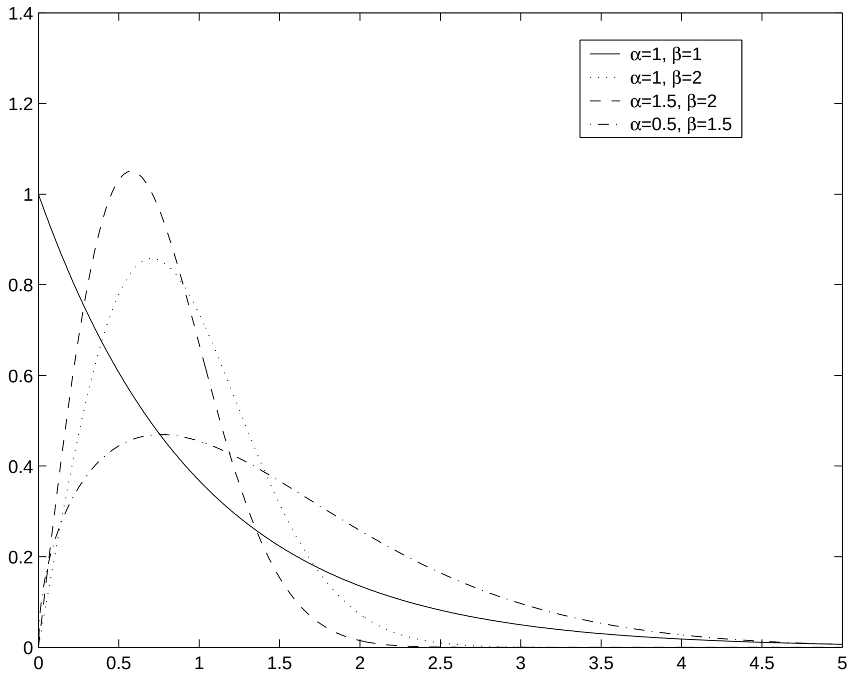


Abbildung 10: Dichten der Weibullverteilung für verschiedene Parameterwerte.

X die Lebensdauer einer Anlage, so ist $P(X \in (x, x+\Delta x] \,|\, X > x)$ die Wahrscheinlichkeit dafür, dass die Anlage in den nächsten Δx Zeiteinheiten ausfällt, wenn sie den Zeitpunkt x erreicht hat. Nach Division durch Δx und Grenzübergang $\Delta x \to 0$ kann diese Größe als Ausfallrate interpretiert werden. Es gilt:

$$\begin{aligned}\frac{1}{\Delta x} P(X \in (x; x+\Delta x] \,|\, X > x) &= \frac{F(x+\Delta x) - F(x)}{\Delta x} \cdot \frac{1}{P(X > x)} \\ &\overset{\Delta x \to 0}{\longrightarrow} \frac{f(x)}{1-F(x)} =: \underbrace{h(x)}_{\text{"Ausfallrate"}}\end{aligned}$$

Die Ausfallrate der Exponentialverteilung ist $h(x) = \lambda$. Die Ausfallrate der Weibullverteilung lautet $h(x) = \frac{c}{b}(\frac{x-a}{b})^{c-1}$ bzw. (mit $a = 0$) $h(x) = \alpha\beta x^{\beta-1}$. Folgende Fälle lassen sich hierbei unterscheiden:

$\beta \in (0,1)$: abnehmende Ausfallrate, "negative Alterung" (" Kinderkrankheiten ")
$\beta = 1$: konstante Ausfallrate, "keine Alterung" ($\Rightarrow$ Gedächtnislosigkeit der Exp.-Vert.)
$\beta > 1$: zunehmende Ausfallrate, "positive Alterung" ("Ermüdungserscheinungen")

(ii) Für $\beta = 3$ ergibt sich folgende Dichte:

$$f_X(x) = \begin{cases} 3\alpha \cdot x^2 \cdot e^{-\alpha x^3} & \text{für} \quad x > 0 \\ 0 & \text{sonst} \end{cases} = 3\alpha x^2 e^{-\alpha x^3} \mathbf{1}_{[0,\infty)}(x).$$

d.h. es liegt eine Exponentialfamilie mit

$$a(\alpha) = 3\alpha,\ h(x) = x^2 \mathbf{1}_{[0,\infty)}(x),\ b(\alpha) = -\alpha \text{ und } \tau(x) = x^3$$

vor. $T(x) = \sum\limits_{i=1}^{n} x_i^3$ ist folglich suffiziente Statistik bzgl. α. Mit $\Gamma = \{\alpha \in \mathbb{R} \,|\, \alpha > 0\}$

und $b(\alpha) = -\alpha$ gilt:

$$B = \{b(\alpha) \,|\, \alpha \in \mathbb{R}, \alpha > 0\} = \{-\alpha \,|\, \alpha > 0\} = \{x \in \mathbb{R} \,|\, x < 0\}.$$

B enthält ein offenes Intervall in $\mathbb{R}$, d.h. die oben bestimmte Statistik ist auch vollständig.

Aufgabe 41

Gegeben seien die folgenden drei anhand einer zufälligen Stichprobe mit Zurücklegen vom Umfang n gebildeten Schätzfunktionen für das arithmetische Mittel μ einer Grundgesamtheit, wobei $E(X_i) = \mu$ und $Var(X_i) = \sigma^2 < \infty$:

- $T_1(x_1, \ldots, x_n) = \frac{1}{n} \sum\limits_{i=1}^{n} x_i$
- $T_2(x_1, \ldots, x_n) = x_j$ für ein festes $j \in \{1, \ldots, n\}$
- $T_3(x_1, \ldots, x_n) = \frac{2}{3}(\frac{1}{n-1} \sum\limits_{i=1}^{n-1} x_i) + \frac{1}{3}x_n$

(a) Welche der drei Schätzfunktionen sind erwartungstreu, d.h. liefern „im Mittel“ das arithmetische Mittel μ?

(b) Von welchem dieser Schätzer würden Sie - ohne zu rechnen - annehmen, dass er die größte Varianz besitzt? Warum?

(c) Berechnen Sie die zugehörigen Varianzen, und ordnen Sie diese der Größe nach.

Zur Schätzung des arithmetischen Mittels μ seien nun alle linearen Schätzer der Form $\sum_{i=1}^{n} \alpha_i x_i$, $\alpha_i \in \mathbb{R}$, zugelassen.

(d) Zeigen Sie, dass die drei oben angegebenen Schätzfunktionen linear sind, indem Sie die jeweiligen Gewichte α_i bestimmen.

(e) Welche Bedingung an die Gewichte α_i garantiert, dass die entsprechenden Schätzer „im Mittel“ das arithmetische Mittel μ liefern?

(f) Bestimmen Sie unter allen linearen Schätzfunktionen, die die in (e) definierte Bedingung erfüllen, diejenige mit minimaler Varianz.

Lösung: (Induktive Statistik, S.55ff.)

(a)

$$
\begin{aligned}
E(T_1(X)) &= E(\frac{1}{n}\sum_{i=1}^{n} X_i) = \frac{1}{n}\sum_{i=1}^{n} E(X_i) = \frac{1}{n} \cdot n \cdot \mu = \mu \\
E(T_2(X)) &= E(X_j) = \mu \\
E(T_3(X)) &= E\left(\frac{2}{3}(\frac{1}{n-1}\sum_{i=1}^{n-1} X_i) + \frac{1}{3}X_n\right) \\
&= \frac{2}{3} \cdot \frac{1}{n-1}\sum_{i=1}^{n-1} E(X_i) + \frac{1}{3}E(X_n) \\
&= \frac{2}{3} \cdot \frac{1}{n-1} \cdot (n-1) \cdot \mu + \frac{1}{3}\mu = \mu
\end{aligned}
$$

(b) T_2, da dieser nur einen Stichprobenwert zur Schätzung heranzieht.

(c) Kriterien für die Bonität eines Schätzers sind:

- Er sollte im "Mittel" das arithmetische Mittel μ der Grundgesamtheit treffen.
- Er sollte eine möglichst kleine Varianz haben

$$
\begin{aligned}
Var(T_1(X)) &= Var\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}\sum_{i=1}^{n} Var(X_i) = \frac{1}{n^2}\cdot n\cdot\sigma^2 = \frac{\sigma^2}{n}\\
Var(T_2(X)) &= Var(X_j) = \sigma^2\\
Var(T_3(X)) &= Var\left(\frac{2}{3}\left(\frac{1}{n-1}\sum_{i=1}^{n-1} X_i\right) + \frac{1}{3}X_n\right)\\
&= \frac{4}{9}\cdot\frac{1}{(n-1)^2}\sum_{i=1}^{n-1} Var(X_i) + \frac{1}{9}Var(X_n)\\
&= \frac{4}{9}\cdot\frac{1}{n-1}\cdot\sigma^2 + \frac{1}{9}\cdot\sigma^2\\
&= \sigma^2\left(\frac{4}{9}\cdot\frac{1}{n-1} + \frac{1}{9}\right) \geq \frac{\sigma^2}{n} \text{ für alle } n \geq 2
\end{aligned}
$$

Beweis der letzten Ungleichung: Für $n \geq 2$ gilt:

$$
\begin{aligned}
\frac{1}{9}\left(\frac{4}{n-1} + \frac{n-1}{n-1}\right) = \frac{1}{9}\frac{n+3}{n-1} \geq \frac{1}{n} \quad &\Leftrightarrow \quad n(n+3) \geq 9n-9\\
&\Leftrightarrow \quad n^2 - 6n + 9 \geq 0\\
&\Leftrightarrow \quad (n-3)^2 \geq 0
\end{aligned}
$$

Im folgenden betrachten wir alle linearen Schätzfunktionen der Form $\sum_{i=1}^{n} \alpha_i X_i$.

(d)

$$
\begin{aligned}
T_1(x_1,\dots,x_n) \quad &= \quad \frac{1}{n}\sum_{i=1}^{n} x_i \text{ d.h. } \alpha_i = \frac{1}{n} \text{ für alle } i = 1,\dots n\\
T_2(x_1,\dots,x_n) \quad &= \quad x_j \text{ für ein festes } j \in \{1,2,\dots,n\}\\
\text{d.h.} \quad & \quad \alpha_j = 1 \text{ und } \alpha_i = 0 \text{ für alle } i \neq j\\
T_3(x_1,\dots,x_n) \quad &= \quad \frac{2}{3}\left(\frac{1}{n-1}\sum_{i=1}^{n-1} x_i\right) + \frac{1}{3}x_n\\
\text{d.h.} \quad & \quad \alpha_i = \frac{2}{3}\cdot\frac{1}{n-1} \text{ für alle } i = 1,\dots,n-1,\ \alpha_n = \frac{1}{3}.
\end{aligned}
$$

(e) Es soll gelten: $E(\sum_{i=1}^{n} \alpha_i X_i) \stackrel{!}{=} \mu$

$$\begin{aligned} E\left(\sum_{i=1}^n \alpha_i X_i\right) &= \sum_{i=1}^n E(\alpha_i X_i) = \sum_{i=1}^n \alpha_i E(X_i) \\ &= \sum_{i=1}^n \alpha_i \cdot \mu = \mu \text{ genau dann, wenn } \sum_{i=1}^n \alpha_i = 1 \text{ ist.} \end{aligned}$$

(f) Das Optimierungsproblem lautet: Minimiere $Var(\sum_{i=1}^n \alpha_i X_i)$ unter der Nebenbedingung $\sum_{i=1}^n \alpha_i = 1$. Es gilt:

$$Var(\sum_{i=1}^n \alpha_i X_i) = \sum_{i=1}^n \alpha_i^2 \cdot Var(X_i) = \sigma^2 \sum_{i=1}^n \alpha_i^2$$

d.h. das Minimierungsproblem lautet: Minimiere $\sum_{i=1}^n \alpha_i^2$ unter der Nebenbedingung $\sum_{i=1}^n \alpha_i = 1$.

Annahme: $\alpha_i = \frac{1}{n} + r_i$ für alle $i = 1, \ldots, n$ mit $r_i \in \mathbb{R}$. Damit gilt $\sum_{i=1}^n r_i = \sum_{i=1}^n (\alpha_i - \frac{1}{n}) = \sum_{i=1}^n \alpha_i - 1 = 0$ genau dann, wenn $\sum_{i=1}^n \alpha_i = 1$ ist. Für die Zielfunktion erhält man somit:

$$\sum_{i=1}^n \alpha_i^2 = \sum_{i=1}^n (\frac{1}{n} + r_i)^2 = \sum_{i=1}^n (\frac{1}{n})^2 + \frac{2}{n} \underbrace{\sum_{i=1}^n r_i}_{=0} + \underbrace{\sum_{i=1}^n r_i^2}_{\geq 0}$$

d.h. das Minimum der Funktion $\sum_{i=1}^n \alpha_i^2$ wird für $r_i = 0$ für alle $i = 1, \ldots, n$ angenommen.

Fazit: Unter allen erwartungstreuen linearen Schätzfunktionen ist $\frac{1}{n} \sum_{i=1}^n X_i$ varianzminimal.

Aufgabe 42

Zeigen Sie, dass die in Übung 36 untersuchte Schätzfunktion

$$N_5 = 2\frac{1}{k}\sum_{i=1}^{k} n_i - 1$$

erwartungstreu ist. Dabei gehe man davon aus, dass es sich in Aufgabe 36 um eine Stichprobe mit Zurücklegen handelt, bei der jeder Ersttagsbrief dieselbe Chance hat, gezogen zu werden.

Lösung: (Induktive Statistik, S.55ff.)

Zu zeigen ist: $E(N_5) = N$.

Beweis: Bei zufälliger Entnahme eines Ersttagsbriefes aus der von 1 bis N durchnummerierten Grundgesmtheit gilt für die Nummer n des zufällig gezogenen Briefes:

$$P(n = i) = \frac{1}{N}$$

Für den Erwartungswert der zufällig gezogenen Nummer n gilt somit:

$$E(n) = \sum_{i=1}^{N} i \cdot P(n = i) = \sum_{i=1}^{N} \frac{i}{N} = \frac{1}{N}\frac{N(N+1)}{2} = \frac{N+1}{2}$$

Für eine einfache Stichprobe mit Zurücklegen $n_1, \cdots, n_k$ vom Umfang k gilt also:

$$E(n_i) = \frac{N+1}{2} \text{ für } i = 1, \cdots, k$$

Damit gilt:

$$\begin{aligned} E(N_5) &= E\left(\frac{2}{k}\sum_{i=1}^{k} n_i - 1\right) = \frac{2}{k}\sum_{i=1}^{k} E(n_i) - 1 \\ &= \frac{2}{k} \cdot k \cdot \frac{N+1}{2} - 1 = N. \end{aligned}$$

Aufgabe 43

Die Dichtefunktion der Gammaverteilung lautet $(\alpha, \lambda > 0)$

$$f_Y(y) = \begin{cases} \dfrac{1}{\Gamma(\alpha)} \lambda^\alpha y^{\alpha-1} e^{-\lambda y} & \text{für} \quad y > 0 \\ & \\ 0 & \text{sonst} \end{cases},$$

dabei ist $\Gamma(\alpha)$ der Wert der sogenannten Gammafunktion an der Stelle α (eine explizite Darstellung dieser Funktion wird im weiteren nicht benötigt).

(a) Skizzieren Sie die Dichtefunktion für einige Parameterkombinationen (α, λ). Welchen Einfluß haben die Parameter auf die Gestalt der Dichtefunktion?

(b) Bestimmen Sie für eine einfache Stichprobe mit Zurücklegen eine suffiziente und vollständige Statistik für das Parameterpaar (α, λ).

(c) Der Erwartungswert der Gammverteilung lautet $E(Y) = \frac{\alpha}{\lambda}$. Es sei bekannt, dass $\lambda = 1$ gilt. Sei X eine einfache Stichprobe mit Zurücklegen, die zum Schätzen des Parameters α gezogen wurde. Untersuchen Sie, ob $\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$ eine gleichmäßig beste erwartungstreue Schätzfunktion für α ist.

Lösung: (Induktive Statistik, S.47ff., S.57ff.)

(a)

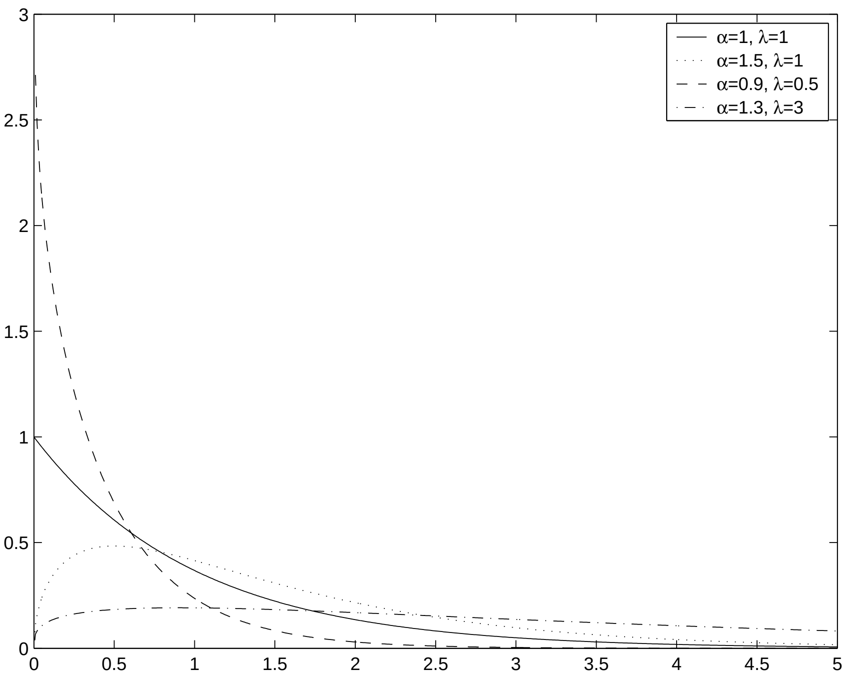


Abbildung 11: Dichten der Gammaverteilung für verschiedene Parameterwerte.

(b) Darstellung der Gammaverteilung als 2-parametrige Exponentialfamilie: Die Dichtefunktion besitzt die folgende Gestalt:

$$\begin{aligned} f_Y(y) &= \frac{1}{\Gamma(\alpha)}\lambda^\alpha \cdot \mathbf{1}_{[0,\infty)}(y) \cdot e^{(\alpha-1)\ln y-\lambda y} \\ &= a(\alpha,\lambda) \cdot h(y) \cdot e^{b_1(\alpha,\lambda)\tau_1(y)+b_2(\alpha,\lambda)\tau_2(y)} \end{aligned}$$

d.h. $T(x) = (\sum\limits_{i=1}^{n} \ln x_i; \sum\limits_{i=1}^{n} x_i)$ ist suffizient. Der Wertebereich der Funktion $b(\alpha,\lambda) = (b_1(\alpha,\lambda), b_2(\alpha,\lambda)) = (\alpha-1,-\lambda)$ ist

$$\begin{aligned} B &= \{(\alpha-1;-\lambda)|\alpha>0,\lambda>0\} \\ &= \{(\alpha^*;\lambda^*)|\alpha^*>-1,\lambda^*<0\} \end{aligned}$$

und enthält einen offenen Quader in $\mathbb{R}^2$, also ist $T(x)$ vollständig.

(c) Für $\lambda = 1$ vereinfacht sich die Dichtefunktion zu:

$$\begin{aligned} f_Y(y) &= \frac{1}{\Gamma(\alpha)} \cdot \mathbf{1}_{[0,\infty)}(y) \cdot e^{(\alpha-1)\ln y-y} \\ &= \frac{1}{\Gamma(\alpha)} \cdot \mathbf{1}_{[0,\infty)}(y) \cdot e^{-y} e^{(\alpha-1)\ln y} \\ &= a(\alpha) \cdot h(y) \cdot e^{b_1(\alpha)\tau_1(y)} \end{aligned}$$

Dies ist eine Darstellung als 1-parametrige Exponentialfamilie. Hieraus folgt, dass $T(x)$=$\sum_{i=1}^{n} \ln x_i$ eine suffiziente Statistik ist. $B = \{(\alpha-1)|\alpha>0\} = (-1,\infty)$ enthält offenes Intervall in $\mathbb{R}$ und somit ist $T(x)$ vollständig. Der Erwartungswert der Gammaverteilung ist $\frac{\alpha}{\lambda}$, mit $\lambda = 1$ damit α. $\bar{x}$ ist erwartungstreu, also muss nach dem Satz von Lehmann Scheffé der GBE-Schätzer die Gestalt $\delta(x) = u'(\sum_{i=1}^{n} \ln x_i)$ haben. u' mit $\frac{1}{n}\sum_{i=1}^{n} x_i = u'(\sum_{i=1}^{n} \ln x_i)$ existiert aber nicht. (Seien beispielsweise $n = 2$ und $x_1 = e,\ x_2 = e^3$ bzw. $x_1' = e^2 = x_2'$. Dann gilt:

$$\sum_{i=1}^{n} \ln(x_i) = 4 = \sum_{i=1}^{n} \ln(x_i') \text{ aber } \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{e+e^3}{2} \neq e^2 = \frac{1}{n}\sum_{i=1}^{n} x_i'$$

folglich kann kein u' existieren, da u' nicht den gleichen Wert auf zwei unterschiedliche Funktionswerte abbilden kann. Analoge Überlegungen funktionieren auch für beliebiges n.)

Aufgabe 44

Verdeutlichen Sie sich an folgendem Beispiel das Prinzip der *Maximum-Likelihood-Schätzung*: In einem Teich soll die Anzahl N der Fische geschätzt werden. Dazu werden M Fische gefangen und markiert. Anschließend werden sie wieder freigelassen. Nach einigen Tagen werden erneut Fische gefangen. Von den n Fischen des Fangs seien x markiert.

Für welches N hat die beobachtete Stichprobe die größtmögliche Wahrscheinlichkeit?

Lösung: (Induktive Statistik, S86ff.)

Die Anzahl markierter Fische ist mit M, die Anzahl Fische insgesamt ist mit N bezeichnet. Es handelt sich um ein "Ziehen" ohne Zurücklegen von n Fischen. Sei X die Anzahl markierter Fische der Stichprobe. X gehorcht einer hypergeometrischen Verteilung: $X \sim H(N, M, n)$, d.h.

$$P(X = x) = \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}} \qquad x = 0, 1, 2, \ldots, n.$$

Sei x eine Realisation von X, so erhalten wir den Maximum-Likelihood-Schätzwert für N durch Maximierung von $P(X = x)$ über N. Schätzwert ist die Maximalstelle N^*, d.h. ein N, für das die Wahrscheinlichkeit des Stichprobenergebnisses maximal ist.

Für die Maximalstelle N^* muss gelten:

$$(1) \qquad \frac{P_{N^*-1,M,n}(X = x)}{P_{N^*,M,n}(X = x)} \leq 1$$

und

$$(2) \qquad \frac{P_{N^*,M,n}(X = x)}{P_{N^*+1,M,n}(X = x)} \geq 1$$

Es gilt für jedes $N \geq M$:

$$\begin{aligned}
\frac{P_{N,M,n}(X = x)}{P_{N+1,M,n}(X = x)} &= \frac{\binom{M}{x}\binom{N-M}{n-x}\binom{N+1}{n}}{\binom{M}{x}\binom{N+1-M}{n-x}\binom{N}{n}} \\
&= \frac{N+1}{N+1-M}\frac{N-n+1-M+x}{N-n+1} \geq 1
\end{aligned}$$

genau dann, wenn

$$\begin{aligned}
(N+1)(N-n+1-M+x) &\geq (N+1-M)(N-n+1) \\
\Longleftrightarrow (N+1)(N+1-M) - (N+1)(n-x) &\geq (N+1-M)(N+1) - n(N+1-M) \\
\Longleftrightarrow N(x-n) + x - n &\geq nM - Nn - n \Leftrightarrow N \geq \frac{nM}{x} - 1.
\end{aligned}$$

Entsprechend ergibt sich

$$\begin{aligned}
\frac{P_{N,M,n}(X = x)}{P_{N+1,M,n}(X = x)} &\leq 1 \\
\Longleftrightarrow N &\leq \frac{nM}{x} - 1.
\end{aligned}$$

Für N^* gilt also

$$N^* - 1 \le \frac{nM}{x} - 1 \text{ bzw. } N^* \le \frac{nM}{x} \text{ und } N^* \ge \frac{nM}{x} - 1.$$

Ist $\frac{nM}{x}$ nicht ganzzahlig, so ist N^* damit eindeutig, z.B. erhalten wir für $M = 100, n = 15$ und $x = 7$:

$$\frac{15 \cdot 100}{7} - 1 = 213.286 \le N^* \le 214.286$$

Ist $\frac{nM}{x}$ ganzzahlig, ist N^* nicht eindeutig, z.B. ergeben sich für $M = 100, n = 10$ und $x = 4$ aus

$$\frac{10 \cdot 100}{4} - 1 = 249 \le N^* \le \frac{10 \cdot 100}{4} = 250$$

die Maximalstellen $N^* = 249$ und $N^* = 250$.

Aufgabe 45

Es wurden 200 Gruppen von jeweils 10 Werkstücken auf Fehler untersucht. Die Zufallsvariable X gebe jeweils die Anzahl der defekten Werkstücke an:

Anzahl defekter Werkstücke	0	1	2	3	≥ 4
beobachtete Häufigkeit	133	52	12	3	0

Aus Erfahrung weiß man, dass X binomialverteilt ist. Bestimmen Sie die Maximum-Likelihood-Schätzfunktion (ML-Schätzer) für den Parameter p dieser Verteilung. Berechnen Sie anschließend den ML-Schätzwert für die gegebenen Stichprobenergebnisse. Wie lauten unter dem geschätzten Binomialmodell die Wahrscheinlichkeiten für die Ereignisse entsprechend der Tabelle?

Lösung: (Induktive Statistik, S86ff.)

Die Binomialverteilung lautet:

$$Y \sim B(m,p): \; P(Y=y) = \binom{m}{y} p^y (1-p)^{m-y} \quad (\text{hier}: m=10)$$

Für eine Stichprobe mit Zurücklegen zu $Y : X = (X_1, \ldots, X_n),\; n = 200$, lautet die Likelihood-Funktion zum Stichprobenergebnis $x = (x_1, \ldots, x_n)$:

$$\begin{aligned} L_x(p) = P_p(X=x) &= \prod_{i=1}^{n} \binom{m}{x_i} p^{x_i}(1-p)^{m-x_i} \\ &= p^{\sum\limits_{i=1}^{n} x_i}(1-p)^{\sum\limits_{i=1}^{n}(m-x_i)} \prod_{i=1}^{n} \binom{m}{x_i} \end{aligned}$$

Der dritte Faktor ist konstant für ein gegebenes Stichprobenergebnis x. Gesucht ist die Maximalstelle von $L_x(p)$ oder äquivalent dazu die Maximalstelle von $\ln(L_x(p))$. Für $p \in (0,1)$ gilt:

$$\ln L_x(p) = \sum_{i=1}^{n} x_i \ln p + \sum_{i=1}^{n}(m-x_i)\ln(1-p) + \ln(\prod_{i=1}^{n} \binom{m}{x_i})$$

und wegen der Monotonie der Logarithmusfunktion haben $L_x(p)$ und $\ln L_x(p)$ eine übereinstimmende Maximalstelle.

$$\begin{aligned} \frac{d \ln L_x(p)}{dp} &= \frac{1}{p}\sum_{i=1}^{n} x_i + \frac{1}{1-p}(-1)\sum_{i=1}^{n}(m-x_i) \\ &\overset{!}{=} 0 \end{aligned}$$

ist gleichwertig mit

$$\begin{aligned} \sum_{i=1}^{n} x_i(1-p) &= \sum_{i=1}^{n}(m-x_i)p \\ \text{bzw.} \; \sum_{i=1}^{n} x_i - p\sum_{i=1}^{n} x_i &= nmp - p\sum_{i=1}^{n} x_i \end{aligned}$$

Lösung der Gleichung ist somit:

$$p \quad = \quad \frac{\sum\limits_{i=1}^{n} x_i}{nm} = \frac{\bar{x}}{m}$$

Das Überprüfen der hinreichenden Bedingung wird dem Leser überlassen. Damit ist $\hat{p} = \frac{\bar{x}}{m}$ Maximalstelle, falls $\sum_{i=1}^{n} x_i \neq 0$ und $\sum_{i=1}^{n} x_i \neq nm$.

Für $\sum_{i=1}^{n} x_i = 0$ ist $x_i = 0$ für $i = 1, \ldots, m$ und damit $L_x(p) = \prod_{i=1}^{n}(1-p)^m = (1-p)^{m \cdot n}$ maximal für $p = 0$. Für $\sum_{i=1}^{n} x_i = nm$ ist $x_i = m$ für $i = 1, \ldots, m$ und damit $L_x(p) = \prod_{i=1}^{n} p^m = p^{mn}$ maximal für $p = 1$.

Damit ist $\hat{p} = \frac{\bar{x}}{m}$ ML-Schätzer für den Parameter p. Für das angegebene Stichprobenergebnis gilt:

$$\hat{p} = \frac{85}{200 \cdot 10} = 0.0425.$$

Die Wahrscheinlichkeitsverteilung der Binomialverteilung mit dem Parameter $p = 0.0425$ ist

y	0	1	2	3	≥ 4
$P(Y = y)$	0.6477	0.2875	0.0574	0.0068	0.0006

Bei 200 Stichprobeneinheiten sollte die relative Häufigkeit der Werte mit $0, 1, 2, 3$ und ≥ 4 in etwa mit diesen Wahrscheinlichkeiten übereinstimmen. Die zu erwartende absolute Häufigkeit ergibt sich daraus zu

Anzahl defekter Werkstücke	0	1	2	3	≥ 4
zu erwartende Häufigkeit	129.54	57.50	11.48	1.36	0.11

Dies stimmt recht gut mit den beobachteten Häufigkeiten überein, so dass nichts gegen die Annahme einer Binomialverteilung spricht. Die Vermutung des Vorliegens einer Binomialverteilung kann seinerseits durch einen sogenannten Anpassungstest überprüft werden (für ein Beispiel s. Aufgabe 56).

Aufgabe 46

In einer automatischen Abfüllanlage wird Zucker in 1-Pfund-Tüten abgefüllt, wobei es zu Abweichungen vom Normgewicht kommen kann. Der Hersteller der Anlage gibt an, dass bei Einstellung einer bestimmten Soll-Füllmenge θ Abweichungen um maximal 5 g in beide Richtungen möglich seien. Weiterhin sollen alle Abweichungen gleich wahrscheinlich sein.

Zur Überprüfung der Einstellung der Anlage wird eine Stichprobe vom Umfang 20 gezogen. Folgende Gewichtswerte werden gemessen:

502	501	500	498	497	499	500	505	496	495,5
499	502	501	497	496	496	505	502	500	496

Bestimmen Sie auf Basis dieser Werte einen ML-Schätzwert für die tatsächlich eingestellte Abfüllmenge θ.

Lösung: (Induktive Statistik, S. 86ff.)

Das Gewicht des Inhalts einer abgefüllten Zuckertüte wird durch eine auf dem Intervall $[\theta - 5, \theta + 5]$ gleichmäßig verteilte Zufallsvariable X wiedergegeben, d.h. X besitzt die Dichte:

$$f_X(x) \quad = \quad \begin{cases} \frac{1}{10} \quad \text{für} & x \in [\theta - 5, \theta + 5] \\ & \\ 0 & \text{sonst} \end{cases}$$

Liegen nun n Realisationen $x_1, \ldots, x_n$ von X vor, so hat die Likelihoodfunktion folgende Gestalt:

$$L(\theta, x_1, \ldots, x_n) \quad = \quad \begin{cases} (\frac{1}{10})^n \quad \text{für} & x_i \in [\theta - 5, \theta + 5] \text{ für alle } i \in \{1, \ldots, n\} \\ & \\ 0 & \text{sonst} \end{cases}$$

Damit L maximal wird, muss gelten:

$$\min\{x_1, \ldots, x_n\} \geq \theta - 5 \text{ und } \max\{x_1, \ldots, x_n\} \leq \theta + 5$$

d.h. jeder Wert $\hat{\theta}$ mit

$$\max\{x_1, \ldots, x_n\} - 5 \leq \hat{\theta} \leq \min\{x_1, \ldots, x_n\} + 5$$

maximiert die Likelihoodfunktion und ist deshalb Maximumlikelihoodschätzer für θ.

Im vorliegenden Fall ist jeder Wert $\hat{\theta}$ aus $[500, 500.5]$ ein Maximumlikelihoodschätzwert für θ.

Wir sehen an diesem Beispiel, dass die Maximumlikelihoodmethode nicht notwendigerweise eindeutige Schätzer und damit auch eindeutige Schätzwerte für die zu schätzenden Parameter liefert. In unserem speziellem Beispiel der Gleichverteilung kann die Menge der Maximumlikelihoodschätzwerte ein Intervall sein. Selbstverständlich kann aber auch hier bei einem günstigen Stichprobenergebnis die Maximalstelle auch eindeutig sein.

Aufgabe 47

Eine Tankstelle veranschlagt für einen Ölwechsel mindestens α Minuten. Die tatsächlich benötigte Zeit X ($X \geq \alpha$) ist von Kunde zu Kunde verschieden, wobei man davon ausgehen kann, dass $X - \alpha$ annähernd exponentialverteilt ($\lambda = 1$) ist. Zur Schätzung von α wurden folgende Arbeitszeiten notiert:

4.2	3.1	3.6	4.5	5.1	7.6	4.4	3.5	3.8
3.9	4.1	4.3	4.4	3.7	3.5	3.8	3.9	

Bestimmen Sie

(a) einen Schätzwert für α mit Hilfe der Maximum-Likelihood-Schätzmethode.

(b) einen Schätzwert für die Zeit, die im Mittel für einen Ölwechsel benötigt wird.

Lösung: (Induktive Statistik, S. 86ff.)

(a) Dichtefunktion von X ist ($\lambda = 1$):

$$f_X(x) = \begin{cases} e^{-(x-\alpha)} & \text{für} \quad x \geq \alpha \\ 0 & \text{sonst} \end{cases}$$

Die Likelihoodfunktion hat folgende Gestalt:

$$L(\alpha, x_1, \ldots, x_n) = \begin{cases} e^{n\alpha} e^{-(x_1+\cdots+x_n)} = e^{n\alpha - n\bar{x}} & \text{für } \alpha \leq \min\{x_1, \ldots, x_n\} \\ 0 & \text{sonst.} \end{cases}$$

Da L eine streng monoton wachsende Funktion von α ist, nimmt L ein Randmaximum für den größten zulässigen $\alpha-$Wert an, und dies ist gerade $\min\{x_1, \ldots, x_n\}$, d.h. $\hat{\alpha} = \min\{x_1, \ldots, x_n\}$ ist der Maximumlikelihoodschätzer für α. Wir erhalten speziell den Schätzwert $\hat{\alpha} = 3.1$.

(b) Im Mittel werden für einen Ölwechsel $E(X) = \alpha + 1$ Minuten benötigt. Unter Ausnutzung der unten angegebenen Invarianzeigenschaft der Maximumlikelihoodmethode erhalten wir einen Schätzwert für $\mu := E(X)$ mittels Teil (a) zu $\hat{\mu} = \hat{a} + 1 = 4.1$.

Invarianzeigenschaft: Ist $\hat{\gamma}$ der ML-Schätzer für γ und $f : \Gamma \to \tilde{\Gamma}$ eine bijektive Abbildung, so ist $\hat{\eta} = f(\hat{\gamma})$ der ML-Schätzer für $\eta = f(\gamma)$ (Beweis: Übungsaufgabe).

Aufgabe 48

In einer Zufallsstichprobe von 250 Studenten einer Universität waren 22 Linkshänder.

(a) Geben Sie einen Bereich an, der mit 95% Wahrscheinlichkeit den Anteil p der Linkshänder unter den Studenten dieser Universität enthält.

(b) Berechnen Sie ein angenähertes 95%-Konfidenzintervall für p unter Verwendung des zentralen Grenzwertsatzes. Wählen Sie dabei die drei folgenden verschiedenen Vorgehensweisen:

  (i) Schätzen Sie die die unbekannte Standardabweichung durch die Stichprobenstandardabweichung.

  (ii) Schätzen Sie die unbekannte Standardabweichung durch die korrigierte Stichprobenstandardabweichung.

  (iii) Verwenden Sie den exakten (aber unbekannten) Wert für die Standardabweichung.

Lösung: (Induktive Statistik, S. 98ff.)

(a) Lösung mit Hilfe der Ungleichung von Tschebyscheff:

$$P(|X - E(X)| \geq c) \leq \frac{Var(X)}{c^2}$$

Entsprechend der Aufgabenstellung ist $Var(X)$ nicht bekannt und damit diese Ungleichung nicht direkt anwendbar. Können wir aber eine obere Schranke für die $Var(X)$ angeben, so erhalten wir eine obere Schranke für die rechte Seite der Tschebyscheffschen Ungleichung. Im Falle einer bekannten Varianz, könnte aber ein Konfidenzintervall wie folgt bestimmt werden. Sei $Y = \bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ der Stichprobenanteil der Linkshänder, so ist $E(Y) = p$ und damit

$$P(|Y - E(Y)| \geq c) = P(|\overline{X} - p| \geq c) \leq \frac{\frac{\sigma^2}{n}}{c^2},$$

wobei σ^2 die Varianz der Stichprobenelemente X_i bezeichne. Da

$$P(|\bar{X} - p| \geq c) \leq \alpha$$

gelten soll, muss

$$\frac{\sigma^2}{n \cdot c^2} = \alpha, \text{ d.h. } c = \sqrt{\frac{\sigma^2}{n \cdot \alpha}}$$

gefordert werden. Es ergibt sich, dass

$$\left[\overline{X} - \sqrt{\frac{\sigma^2}{n \cdot \alpha}}, \overline{X} + \sqrt{\frac{\sigma^2}{n \cdot \alpha}}\right]$$

ein $(1 - \alpha)$-Konfidenzintervall umfasst.

Da die einzelnen X_i Bernoulli-verteilt sind und deshalb bekanntlich die Varianz mit $Var(X_i) = p(1 - p)$ gegeben ist, können wir die Varianz sicherlich (für $p = 0.5$ wird der

Ausdruck $p(1-p)$ maximal) durch 0.25 abschätzen. Einsetzen in die oben bestimmten Formeln ergibt:

$$\overline{x} = \frac{22}{250} = 0.088, \qquad c = \sqrt{\frac{0.25}{250 \cdot 0.05}} = 0.141$$

Daraus ergibt sich die Einschließung des Konfidenzintervalls mit Hilfe von Tschebyscheff:

$$\big[\overline{x} - 0.14, \overline{x} + 0.14\big] = \big[0; 0.229\big], \qquad (p \geq 0!)$$

b) Das 95%-Konfidenzintervall auf Grund der Näherung durch die Normalverteilung bestimmt sich gemäß:

$$\left[\frac{1}{n}\sum_{i=1}^{n} x_i - \frac{\hat{\sigma}}{\sqrt{n}} u_{1-\frac{\alpha}{2}}, \frac{1}{n}\sum_{i=1}^{n} x_i + \frac{\hat{\sigma}}{\sqrt{n}} u_{1-\frac{\alpha}{2}}\right]$$

Diese Näherung beruht auf der Überlegung, dass

$$\frac{\overline{X} - p}{\sqrt{Var(\overline{X})}}$$

näherungsweise normalverteilt ist. Ein Problem stellt nun die geeignete Wahl (Schätzung) des unbekannten Terms $Var(\overline{X})$ dar. Prinzipiell kommen hierbei drei Möglichkeiten in Frage.

(i) Verwendung der Stichprobenstandardabweichung als Schätzwert: Die Stichprobenstandardabweichung berechnet sich zu:

$$\begin{aligned}
\hat{\sigma} &= \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2} = \sqrt{\left[\frac{1}{n}\sum_{i=1}^{n} x_i - \overline{x}^2\right]} = \sqrt{\bar{x}(1-\bar{x})} \\
&= \sqrt{(0.088 - (0.088)^2)} = \sqrt{0.080} \\
&= 0.283
\end{aligned}$$

Mit $u_{1-\frac{\alpha}{2}} = 1.96$ ergibt sich das folgende Konfidenzintervall:

$$I_2 = \left[0.088 - \frac{0.283}{\sqrt{250}} \cdot 1.96, 0.088 + \frac{0.283}{\sqrt{250}} \cdot 1.96\right] = [0.0529, 0.1231]$$

(ii) Verwendung der korrigierten Stichprobenstandardabweichung als Schätzwert: Die korrigierte Stichprobenstandardabweichung berechnet sich zu:

$$\begin{aligned}
\hat{\sigma} &= \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2} = \sqrt{\frac{n}{n-1}\left[\frac{1}{n}\sum_{i=1}^{n} x_i - \overline{x}^2\right]} \\
&= \sqrt{\frac{250}{249}(0.088 - (0.088)^2)} = \sqrt{0.081} \\
&= 0.284
\end{aligned}$$

Mit $u_{1-\frac{\alpha}{2}} = 1.96$ ergibt sich das folgende Konfidenzintervall:

$$I_2 = \left[0.088 - \frac{0.284}{\sqrt{250}} \cdot 1.96, 0.088 + \frac{0.284}{\sqrt{250}} \cdot 1.96\right] = [0.0527, 0.1232]$$

Bemerkung: Unter Umständen könnte man auch auf die Idee kommen, statt der Normalverteilung die t-Verteilung zu verwenden, da man einen näherungsweise normalverteilten Ausdruck mit der korrigierten Stichprobenstandardabweichung standardisiert. Dies hat aber keine Auswirkungen auf das Ergebniss, da für große n die t-Verteilung nicht mehr von der Normalverteilung zu unterscheiden ist.

(iii) Verwendung des exakten Werts $p(1-p)$ für die Standardabweichung und Lösen einer quadratischen Gleichung:

$$Y = \frac{p - \overline{X}}{\sqrt{\frac{p(1-p)}{n}}}$$

ist näherungsweise standardnormalverteilt. Folglich gilt:

$$P(-u_{1-\frac{\alpha}{2}} \leq \frac{p - \overline{X}}{\sqrt{\frac{p(1-p)}{n}}} \leq u_{1-\frac{\alpha}{2}}) = 1 - \alpha$$

Aus

$$-u_{1-\frac{\alpha}{2}} \leq \frac{p - \overline{X}}{\sqrt{\frac{p(1-p)}{n}}} \leq u_{1-\frac{\alpha}{2}}$$

erhalten wir durch quadrieren

$$\frac{(p - \overline{X})^2}{\frac{p(1-p)}{n}} \leq u^2_{1-\frac{\alpha}{2}}$$

Durch Umformen ergibt sich die quadratische Ungleichung

$$p^2(n + u^2_{1-\frac{\alpha}{2}}) - p(2a + u^2_{1-\frac{\alpha}{2}}) + \frac{a^2}{n} \leq 0, \quad a^2 := \overline{x}^2 n^2$$

Die Nullstellen dieser Parabel sind:

$$p_{1,2} = \frac{2a + u^2_{1-\frac{\alpha}{2}} \pm u_{1-\frac{\alpha}{2}}\sqrt{u^2_{1-\frac{\alpha}{2}} + 4a(1 - \frac{a}{n})}}{2(n + u^2_{1-\frac{\alpha}{2}})}$$

Damit gilt obige Ungleichung falls p zwischen diesen beiden Nullstellen liegt. Also ist $I_3 = [p_1; p_2]$ ein angenähertes 95%-Konfidenzintervall. Einsetzen ergibt:

$$p_{1,2} = \frac{44 + 1,96^2 \pm 1.96\sqrt{1.96^2 + 88(1 - 0.088)}}{2(250 + 1.96^2)} = \begin{cases} 0.1296 \\ 0.0589 \end{cases}$$

und damit das Intervall

$$I_3 = [0.0589, 0.1296]$$

als Näherung für das 95%-Konfidenzintervall.

Hinweis: Zusätzlich zu den hier aufgeführten Varianten besteht auch die Möglichkeit eine exakte Lösung für das Konfidenzintervall des Parameters p einer Bernoulliverteilung zu berechnen. Man verwendet hierfür eine Beziehung zwischen der Binomialverteilung und der F-Verteilung. (Details siehe Induktive Statistik S. 108ff) Man erhält dann folgendes exaktes 95% Konfidenzintervall (mit $c := n \cdot \bar{x}$):

$$\begin{aligned}[p_u, p_o] &= \left[\frac{c}{c + (n - c + 1) \cdot F^{-1}_{2(n-c+1),2c}(1 - \frac{\alpha}{2})}, \frac{(c+1) \cdot F^{-1}_{2(c+1),2(n-c)}(1 - \frac{\alpha}{2})}{n - c + (c + 1 \cdot F^{-1}_{2(c+1),2(n-c)}(1 - \frac{\alpha}{2})}\right] \\ &= [0.0603, 0.1233]\end{aligned}$$

Aufgabe 49

Das Gewicht von Wäscheklammern sei als näherungsweise normalverteilt angenommen. Folgende Daten wurden in einer Stichprobe gemessen (in $[g]$):

9.3	8.5	9.8	10.1	7.2	11.4	12.5	10.8
11.3	13.6	10.3	8.2	12.8	15.2	9.6	10.7

(a) Bestimmen Sie ein 95%- und ein 99%-Konfidenzintervall für den Mittelwert und die Varianz des Gewichts der Klammern.

(b) Von der Firma, die die Wäscheklammern produziert, erfahren Sie, dass das durchschnittliche Klammergewicht 10.5 g beträgt. Bestimmen Sie nun ein 95%-Konfidenzintervall für die Varianz.

(c) Bestimmen Sie ein 95%-Konfidenzintervall für den Mittelwert unter der Annahme, dass die Varianz des Klammergewichts mit $\sigma^2 = 5\ g^2$ bekannt ist.

Lösung: (Induktive Statistik, S. 98ff.)

(a) Wir beschreiben das zufällige Gewicht einer Wäscheklammer durch die Zufallsvariable X mit $X \sim N(\mu, \sigma^2)$, wobei μ, σ^2 unbekannt sind. Für das Konfidenzintervall für den unbekannten Mittelwert μ gilt:

$$\left[\overline{x} - t(n-1)_{1-\alpha/2}\frac{s^*}{\sqrt{n}}, \overline{x} + t(n-1)_{1-\alpha/2}\frac{s^*}{\sqrt{n}}\right]$$

mit

$$t(15)_{0.975} = 2.1314, \qquad t(15)_{0.995} = 2.947$$

und

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} X_i = 10.71, \qquad s^* = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2} = 2.08$$

$t(n)_\alpha$ bezeichne dabei das α-Quantil der t-Verteilung mit n Freiheitsgraden. Es ergibt sich also folgendes Konfidenzintervall für μ zum Niveau 0.95 (0.99):

$$[9.58, 11.84] \qquad ([9.17, 12.24])$$

Für das Konfidenzintervall für σ^2 bei unbekanntem μ gilt wegen $\frac{(n-1)S^{*2}}{\sigma^2} \sim \chi^2(n-1)$:

$$\left[\frac{(n-1)s^{*2}}{\chi^2(n-1)_{1-\alpha/2}}, \frac{(n-1)s^{*2}}{\chi^2(n-1)_{\alpha/2}}\right]$$

Hierbei bezeichnet $\chi^2(n)_\alpha$ das α-Quantil der Chi-Quadrat-Verteilung mit n Freiheitsgraden. Konkret ergibt sich für $1-\alpha = 0.95$ (0.99), $n = 16$, $s^{*2} = 4.334$ das Konfidenzintervall:

$$[2.365, 10.382] \qquad ([1.982, 14.13])$$

(b) Hierbei ist der Mittelwert $\mu = 10.5$ bekannt und folglich gilt, dass $\sum_{i=1}^{n}(X_i - \mu)^2/\sigma^2 \sim \chi^2(n)$ verteilt ist. Die Wahrscheinlichkeit, dass σ^2 unter- bzw. oberhalb des Konfidenzintervalls liegt soll jeweils $\alpha/2$ betragen, dies wird erreicht durch:

$$\chi^2(n)_{\alpha/2} \quad \leq \quad \frac{\sum\limits_{i=1}^{n}(X_i - \mu)^2}{\sigma^2} \quad \leq \quad \chi^2(n)_{1-\alpha/2}$$

bzw.

$$\frac{\sum\limits_{i=1}^{n}(X_i - \mu)^2}{\chi^2(n)_{1-\alpha/2}} \quad \leq \quad \sigma^2 \quad \leq \quad \frac{\sum\limits_{i=1}^{n}(X_i - \mu)^2}{\chi^2(n)_{\alpha/2}}$$

Konkret ergibt sich mit $\sum_{i=1}^{n}(x_i - \mu)^2 = 65.69$, $\chi^2(16)_{0.975} = 28.85$ und $\chi^2(16)_{0.025} = 6.91$ das 95%− Konfidenzintervall:

$$\left[2.28, 9.51\right]$$

(c) Falls σ^2 bekannt ist, kann man das Konfidenzintervall für μ analog zu a) mit Hilfe der entsprechenden Quantile der Standardnormalverteilung berechnen. Es ergibt sich dann $(1-\alpha)$-Konfidenzintervall:

$$\left[\bar{x} - u_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}, \bar{x} + u_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\right]$$

wobei $u_{1-\frac{\alpha}{2}}$ das $(1-\frac{\alpha}{2})$-Quantil der Standardnormalverteilung sei. Für α=0.05 gilt $u_{0.975} = 1.96$, d.h. das 95%-Konfidenzintervall lautet:

$$\left[10.71 - 1.96\frac{\sqrt{5}}{\sqrt{16}}, 10.71 - 1.96\frac{\sqrt{5}}{\sqrt{16}}\right] = \left[9.61, 11.81\right]$$

Aufgabe 50

Y sei Poisson-verteilt mit Parameter λ. Zu testen sei die Hypothese $H_0: \lambda = \lambda_0$ gegen $H_1: \lambda = \lambda_1$ mit $\lambda_1 > \lambda_0$. Konstruieren Sie einen einfachen besten Test zum Niveau α.

Lösung: (Induktive Statistik, S. 123ff.)

Lösung:
Für eine einfache Stichprobe mit Zurücklegen gilt:

$$P_\lambda(X = x) = \prod_{i=1}^{n} \frac{\lambda^{x_i}}{x_i!} e^{-\lambda} = e^{-n\lambda} \frac{\lambda^{\sum\limits_{i=1}^{n} x_i}}{\prod\limits_{i=1}^{n} x_i!}$$

Damit ist

$$\begin{aligned} Q(x) &= \frac{P_{\lambda_1}(X = x)}{P_{\lambda_0}(X = x)} = \frac{e^{-n\lambda_1} \lambda_1^{\sum\limits_{i=1}^{n} x_i}}{e^{-n\lambda_0} \lambda_0^{\sum\limits_{i=1}^{n} x_i}} \\ &= e^{-n(\lambda_1 - \lambda_0)} (\frac{\lambda_1}{\lambda_0})^{\sum\limits_{i=1}^{n} x_i} \end{aligned}$$

und es gilt:

$$\begin{aligned} Q(x) \le k \quad &\Leftrightarrow \quad e^{-n(\lambda_1 - \lambda_0)} (\frac{\lambda_1}{\lambda_0})^{\sum\limits_{i=1}^{n} x_i} \le k \\ &\Leftrightarrow \quad -n(\lambda_1 - \lambda_0) + \sum_{i=1}^{n} x_i (\ln(\lambda_1) - \ln(\lambda_0)) \le \ln(k) \\ &\Leftrightarrow \quad T(x) := \sum_{i=1}^{n} x_i \le \frac{\ln(k) + n(\lambda_1 - \lambda_0)}{\ln(\lambda_1) - \ln(\lambda_0)} =: c_k. \end{aligned}$$

Der Neyman/Pearson-Test lautet also:

$$\delta_k(x_1, \ldots, x_n) = \begin{cases} d_0 & \sum\limits_{i=1}^{n} x_i \le c_k \\ d_1 & \sum\limits_{i=1}^{n} x_i > c_k \end{cases}$$

Da $\sum\limits_{i=1}^{n} x_i \in \{0, 1, \ldots, \} = \mathbb{N} \cup \{0\}$ ist, kann man sich bei c_k ebenfalls auf Werte aus $\mathbb{N} \cup \{0\}$ beschränken. Für den Fehler 1. Art erhält man

$$P_I(\delta_k, \lambda_0) = P_{\lambda_0}(\sum_{i=1}^{n} X_i > c_k),$$

für den Fehler 2. Art

$$P_{II}(\delta_k, \lambda_1) = P_{\lambda_1}(\sum_{i=1}^{n} X_i \leq c_k).$$

Da $\sum\limits_{i=1}^{n} X_i$ Poisson-verteilt ist mit Parameter $n\lambda_j, j \in \{0, 1\}$, können die Fehlerwahrscheinlichkeiten bestimmt werden.

Für $\lambda_0 = 0.5, \lambda_1 = 1$ und $n = 10$ sind diese Werte in der folgenden Tabelle wiedergegeben.

c_k	$P_I(\delta_k)$	$P_{II}(\delta_k)$
0	$0,9933$	$0,0000$
1	$0,9596$	$0,0005$
2	$0,8753$	$0,0028$
3	$0,7350$	$0,0103$
4	$0,5595$	$0,0293$
5	$0,3840$	$0,0671$
6	$0,2378$	$0,1301$
7	$0,1334$	$0,2202$
8	$0,0681$	$0,3328$
9	$0,0318$	$0,4579$
10	$0,0137$	$0,5830$

Man sieht dort deutlich, wie sich bei zunehmendem c_k die Fehlerwahrscheinlichkeit 1.Art verkleinert, die Fehlerwahrscheinlichkeit 2.Art aber immer stärker anwächst. Das übliche Niveau von $\alpha = 0.05$ wird erstmals bei $c_k = 9$ unterschritten. Dann ist aber die Wahrscheinlichkeit für den Fehler 2. Art mit 0.4579 doch recht beachtlich. Eine Verbesserung lässt sich bei diesem Stichprobenumfang für das Niveau 0.0318 nach dem Satz von Neyman/Pearson nicht erreichen.

Nicht weiter überraschend ist, dass auch hier wie bei den Schätzproblemen die Entscheidung nur vom Wert $\sum\limits_{i=1}^{n} x_i$ der für λ vollständigen und suffizienten Stichprobenfunktion $T(x) = \sum\limits_{i=1}^{n} x_i$ abhängt, denn, da $T(x)$ alle Information des Stichprobenergebnisses x über λ enthält, sollte eine gute Entscheidungsfunktion δ, hier also ein Test, bei übereinstimmendem Wert von T $(T(x) = T(x'))$ auch dieselbe Entscheidung treffen $(\delta(x) = \delta(x'))$.

Aufgabe 51

Um beurteilen zu können, wie sich eine 10%-ige Preiserhöhung bei einem Artikel auf dessen Absatz auswirkt, werden Verkaufsstellen ausgewählt, in denen probeweise eine Preiserhöhung durchgeführt wird. Die Preiserhöhung soll dann generell durchgeführt werden, wenn mit einer durchschnittlichen Absatzminderung von höchstens 8% zu rechnen ist. Es werden nun zufällig n Verkaufsstellen ausgewählt und die sich dort ergebenden Absatzveränderungen bestimmt.

(a) Erläutern Sie anhand der Situationsbeschreibung kurz, was unter einem *Test* zu verstehen ist. Überlegen Sie dabei insbesondere, welches Ziel Sie beim Testen verfolgen, welche Informationen Ihnen gegeben sind und mit Hilfe welcher Schritte Sie basierend auf den Ihnen zur Verfügung stehenden Informationen Ihr Ziel erreichen können.

(b) Formulieren Sie ein geeignetes statistisches Entscheidungsproblem (Modell und Hypothesen).

(c) Welche Fehlentscheidungen können in diesem Beispiel auftreten?

(d) Welcher Zusammenhang besteht zwischen den Entscheidungen und der Wahl der Nullhypothese?

(e) Wie beeinflußt die Wahl der Nullhypothese die Entscheidung des Testverfahrens?

Lösung: (Induktive Statistik, S. 123ff.)

(a) Es besteht Unklarheit über die Auswirkung einer 10%-Preiserhöhung auf den Absatz. Eine Preiserhöhung führt nur dann zu einer Umsatzsteigerung, wenn der Absatz nicht um mehr als rund 8% zurückgeht. Wegen eventueller negativer Nebeneffekte sowie der durch die Preiserhöhung entstehenden Kosten soll diese nur durchgeführt werden, wenn die Absatzminderung nicht mehr als 8% beträgt. Es geht also um die Hypothesen

$$H_0: \text{ Absatzminderung } \geq 8\%, \quad H_1: \text{ Absatzminderung } < 8\%$$

Informationen, die einer Entscheidung über die Hypothesen zugrundgeegt werden können, liefern die probenweise durchgeführten Preiserhöhungen in einigen Verkaufstellen. Die Entscheidung wird also sinnvollerweise von den Ergebnissen in diesen Filialen abhängen. Man spricht hier von einer Testsituation.

Allgemein: Der Parameterbereich Γ wird in zwei Bereiche Γ_0, Γ_1 aufgeteilt mit $\Gamma_0 \cup \Gamma_1 = \Gamma$, $\Gamma_0 \cap \Gamma_1 = \emptyset$. (Im Beispiel ist der interessierende Parameter die prozentuale Absatzminderung, Parameterbereich ist also $\Gamma = [0, 100]$, wenn eine Absatzsteigerung ausgeschlossen wird. Ensprechend den angegebenen Hypthesen ist $\Gamma_0 = [8, 100]$ und $\Gamma_1 = [0, 8)$.) Getestet wird die Hypothese $H_0 : \gamma \in \Gamma_0$ (Nullhypothese) gegen $H_1 : \gamma \in \Gamma_1$ (Gegenhypothese). Informationen als Entscheidungsgrundlage liefern Stichprobenergebnisse. Die Entscheidung sollte dann so erfolgen, dass Fehlentscheidungen möglichst vermieden werden, d.h. mit geringer Wahrscheinlichkeit auftreten.

(b) Entsprechend den drei Grundannahmen modellieren wir die Situation wie folgt:

(i) Es sei Y die Absatzminderung einer zufällig ausgewählten Filiale.

(ii) Es ist eine Klasse von Verteilungen festzulegen, der die (unbekannte) Wahrscheinlichkeitsverteilung Y angehört.
Beispielsweise könnte hier die Klasse der Normalverteilungen gewählt werden, wobei für den Mittelwert μ und Varianz σ^2 die Einschränkungen $0 \leq \mu \leq 100$ und, um negative Absatzminderung (nahezu) anzuschließen, $\mu > 3\sigma$ gemacht werden.

(iii) Es werden die Verkaufsminderungen bei n zufällig ausgewählten Filialen beobachtet, d.h. es liegen Stichprobenergebnisse $x_1, \ldots, x_n$ eines Stichprobenvektors $X = (X_1, \ldots, X_n)$ vor. Es handelt sich um eine Stichprobe mit Zurückliegen, d.h.

$$F_X(x_1, \ldots, x_n) = \prod_{i=1}^{n} F_Y(x_i),$$

der Zusammenhang zwischen der Verteilung der Stichprobe und der Verteilung von Y ist also vollständig bekannt.

Sei $\mu = E(Y)$, so lauten die Hypothesen

$$H_0: \ \mu \geq 8\% \qquad H_1: \ \mu < 8\%$$

(c) Folgende Fehler gilt es zu unterscheiden:
Fehler 1. Art: H_0 wird abgelehnt, obwohl H_0 gilt.
Fehler 2. Art: H_0 wird nicht abgelehnt, obwohl H_0 nicht gilt.

Entscheidung für / Wirklichkeit	$\mu \geq 8\%$	$\mu < 8\%$
$\mu \geq 8\%$	richtig	falsch: Fehler 1. Art
$\mu < 8\%$	falsch: Fehler 2. Art	richtig

Diese Fehler werden beim Testen unterschiedlich gewichtet:
Der Fehler 1. Art soll mit vorgegebener Wahrscheinlichkeit vermieden werden: Hieraus ergibt sich eine entsprechende Forderung an die Wahl von H_0 und H_1. Auf das Beispiel bezogen bedeutet dies: Falls ein Verlust an Marktanteil schlimmer bewertet wird als eine entgangene Umsatzsteigerung, dass man wählt: $H_0: \ \mu \geq 8\%, \quad H_1: \ \mu < 8\%$. Dies ist implizit in der Formulierung enthalten: „Preiserhöhung nur, wenn Absatzminderung $< 8\%$ ist".
Unter der Vorgabe der Wahrscheinlichkeit für den Fehler 1. Art bestimmt man dann ein Entscheidungsverfahren, das die Wahrscheinlichkeit für den Fehler 2. Art so weit wie möglich reduziert.

(d) Die Wahl der Nullhypothese beeinflußt die Entscheidung des Testverfahrens, da man jeweils in unterschiedlicher Richtung vorsichtig ist, d.h. einen Fehler vermeiden will. Man unterscheidet zwei Arten von Test, den Signifikanztest der im Falle, dass H_0 nicht abgelehnt werden kann, keine Aussage über die Gültigkeit von H_0 und damit auch H_1 trifft, und den Alternativtest, bei dem die Schlussfolgerung entweder „H_0 gilt" oder „H_1 gilt" lautet.

Aufgabe 52

Ein Unternehmen erhält von einem Lieferanten den Hinweis, dass eine gelieferte Warenpartie von 10 000 Teilen möglicherweise nicht den gewöhnlichen Ausschußanteil von 5%, sondern aufgrund von Produktionsschwierigkeiten einen erhöhten Anteil von 15% besitzt. Zur Überprüfung dieses Hinweises wird eine Stichprobe vom Umfang $n = 30$ entnommen. Mit einem gleichmäßig besten Test zu einem Niveau von höchstens 0.1 soll „statistisch" gesichert werden, dass die Partie wie normal einen Ausschussanteil von 5% hat. Formulieren Sie den Test.

Lösung: (Induktive Statistik, S. 134ffff.)

Sei $X = (X_1, \ldots, X_n)$ der Zufallsvektor der Stichprobe, d.h. $X_1, \ldots, X_n$ sind unabhängig Bernoulli-verteilt mit Parameter p, also $P(X_i = 1) = p$, wobei p der Ausschussanteil der Warenpartie ist, d.h. für p kommen die Werte 0.05 und 0.15 in Frage.

Bei einer Stichprobe vom Umfang $n = 30$ gilt für die Wahrscheinlichkeitsverteilung von X:

$$P_p(X = x) = p^{\sum\limits_{i=1}^{30} x_i} (1-p)^{30-\sum\limits_{i=1}^{30} x_i}$$

Da ein Ausschussanteil von 5% statistisch gesichert werden soll, lauten die Hypothesen $H_0 : p = 0.15$ und $H_1 : p = 0.05$

Damit lautet mit $p_0 = 0.15$ und $p_1 = 0.05$ der Wahrscheinlichkeitsquotient:

$$\begin{aligned} Q(x) &= \frac{p_1^{\sum\limits_{i=1}^{30} x_i} (1-p_1)^{30-\sum\limits_{i=1}^{30} x_i}}{p_0^{\sum\limits_{i=1}^{30} x_i} (1-p_0)^{30-\sum\limits_{i=1}^{30} x_i}} \\ &= \underbrace{\left(\frac{p_1}{p_0}\right)}_{<1}^{\sum\limits_{i=1}^{30} x_i} \cdot \underbrace{\left(\frac{1-p_0}{1-p_1}\right)}_{<1}^{\sum\limits_{i=1}^{30} x_i} \cdot \underbrace{\left(\frac{1-p_1}{1-p_0}\right)^{30}}_{\text{const.}} \end{aligned}$$

Es gilt also, dass $Q(x)$ streng monoton in $\sum_{i=1}^{30} x_i$ fällt, d.h. der Neymann-Pearson-Test läßt sich formulieren als

$$\delta_k(x) = \begin{cases} d_1 & \text{für } \sum\limits_{i=1}^{30} x_i < c \\ & \\ d_0 & \text{für } \sum\limits_{i=1}^{30} x_i \geq c \end{cases}$$

und ist gleichmäßig bester Test zum Niveau

$$P_{p_0}(\sum_{i=1}^{30} X_i < c).$$

Aus der Bedingung an die Fehlerwahrscheinlichkeit 1. Art Testniveau $\alpha \leq 10\%$

$$P_{p_0}(\sum_{i=1}^{30} X_i < c) \leq 0.1$$

ergibt sich mit Hilfe einer Tabelle der Binomialverteilung für $p = 0.15$

$$\begin{aligned}
&P_{p_0}(\sum_{i=1}^{30} X_i < 2) = 0.048 < 0.1, \\
&P_{p_0}(\sum_{i=1}^{30} X_i < 3) = 0.1514 > 0.1 \\
&\Rightarrow c = 2
\end{aligned}$$

Für $c = 2$ ist $\delta(x)$ gleichmäßig bester Test zum Niveau $\alpha = 4.8\%$.

Aufgabe 53

Eine Firma stellt Leuchtstoffröhren her. Es wurde ein neues Produktionsverfahren eingeführt, aufgrund dessen sich die mittlere Lebensdauer der Röhren (bisher 1500 Stunden) erhöhen soll. Man möchte prüfen, ob dies tatsächlich der Fall ist und entnimmt daher eine Stichprobe vom Umfang 10 aus der laufenden Produktion.

(a) Formulieren Sie die Nullhypothese, und geben Sie einen gleichmäßig besten Test zum Niveau $\alpha = 0.05$ an, wenn angenommen wird, dass die Lebensdauer der Röhren normalverteilt ist und die Standardabweichung $\sigma = 100\ h$ beträgt.

(b) Sei $(1430, 1480, 1520, 1500, 1550, 1460, 1530, 1540, 1600, 1510)$ das Ergebnis der Stichprobe. Welche Entscheidung trifft das Unternehmen?

Lösung: (Induktive Statistik, S.151ff., S.198)

(a) Die Zufallsvariable $Y \sim N(\mu, \sigma^2)$ modelliere die Lebensdauer einer Leuchstoffröhre, wobei μ unbekannt und $\sigma = 100$ bekannt ist. Als Hypothesen wählen wir:

$$H_0 : \mu \leq 1500, \qquad H_1 : \mu > 1500$$

Es liegt eine Stichprobe X vom Umfang $n = 10$ vor, und folglich gilt:

$$\sum_{i=1}^{n} X_i \sim N(n\mu, n\sigma^2), \qquad \text{bzw. } \overline{X} \sim N(\mu, \frac{\sigma^2}{n})$$

$T(x) = \sum_{i=1}^{n} x_i$ ist suffizient und vollständig für μ mit monoton wachsendem Wahrscheinlichkeitsquotienten. Damit ist für das einseitige Testproblem

$$H_0 : \mu \leq \mu_0 \qquad \text{gegen } H_1 : \mu > \mu_0$$

die Entscheidungsfunktion

$$\delta_c(x) = \begin{cases} d_0 & \text{für } \sum_{i=1}^{n} x_i \leq c \\ d_1 & \text{für } \sum_{i=1}^{n} x_i > c \end{cases}$$

gleichmäßig bester Test zum Niveau

$$\begin{aligned} P_I(\delta_c) &= P_{\mu_0}(\sum_{i=1}^{n} X_i > c) \\ &= 1 - F_{n\mu_0, n\sigma^2}(c) = 1 - \Phi(\frac{c - n\mu_0}{\sqrt{n}\sigma}) \\ &= \alpha, \text{wenn } c = \mu n + \sigma\sqrt{n}u_{1-\alpha}. \end{aligned}$$

Die letzte Gleichheit gilt aufgrund der Beziehung $\Phi(u_{1-\alpha}) = 1 - \alpha$.

Bei der Aufgabenstellung bedeutet dies:

$$
\begin{aligned}
P_{\mu_0}(\sum_{i=1}^{n} X_i > c) &= 1 - \Phi(\frac{c - 10 \cdot 1500}{\sqrt{10} \cdot 100}) \\
&\overset{!}{=} 0.05
\end{aligned}
$$

Also ergibt sich durch Umformen:

$$c = 10 \cdot 1500 + u_{1-0.05} \cdot \sqrt{10} \cdot 100 = 15521.8$$

Hieraus ergibt sich die Entscheidungsfunktion:

$$
\delta_c(x) = \begin{cases} d_0 & \text{für } \sum_{i=1}^{n} x_i \leq 15521.8 \\ \\ d_1 & \text{für } \sum_{i=1}^{n} x_i > 15521.8 \end{cases}
$$

(b) Man erhält für die angegebenen Werte $\sum_{i=1}^{n} x_i = 15120$ und folglich gilt $\delta(x) = d_0$ und H_0 kann nicht abgelehnt werden. Die Erhöhung der mittleren Lebensdauer durch das neue Verfahren ist also <u>nicht</u> statistisch gesichert zum Niveau 0.05.

Aufgabe 54

Eine Firma hat über einen Zeitraum von 45 Wochen den Absatz eines ihrer Produkte ermittelt. Dabei ergab sich ein Stichprobenmittel $\bar{x} = 447\,000$ Stück/Woche. Die Firma weiß weiterhin aus langjähriger Erfahrung, daß der Absatz dieses Produktes normalverteilt ist mit dem Mittelwert μ und der Standardabweichung $\sigma = 20\,000$.

(a) Die Firma möchte nun testen, ob die Stichprobe der Annahme widerspricht, dass der durchschnittliche Absatz pro Woche 450 000 Stück beträgt.

- Formulieren Sie die Null- und die Gegenhypothese.
- Geben Sie einen Test zum Signifikanzniveau 0.05 an.
- Welche Entscheidung wird die Firma treffen?

(b) Die Firma hat nun eine Anzeigenkampagne gestartet. Bei einer anschließenden Untersuchung über 16 Wochen werden ein Stichprobenmittel von $\bar{x} = 460\,000$ und eine korrigierte Stichprobenstandardabweichung von $s^* = 21\,000$ festgestellt. Geben Sie einen Test zum Signifikanzniveau 0.05 an, und ermitteln Sie, ob die Anzeigenkampagne erfolgreich war. Dabei werde eine unbekannte Varianz zugrunde gelegt.

(c) Was beschreibt die Gütefunktion eines Tests allgemein? Skizzieren Sie die Gütefunktion für einen einseitigen Test beispielhaft anhand der Daten aus Aufgabenteil (a). Wie würde Ihrer Meinung nach eine ideale Gütefunktion aussehen?

Lösung: (Induktive Statistik, S.183ff., S.198)

(a) Im Unterschied zu Übung 53 interessiert hier nicht die Abweichung in eine Richtung (einseitiger Test), sondern eine Abweichung an sich (zweiseitiger Test). Die Hypothesen lauten:

$$H_0 \;:\; \mu = \mu_0 = 450\,000 \qquad H_1 \;:\; \mu \neq \mu_0$$

Eine intuitiv einleuchtende Entscheidungsfunktion ist:

$$\delta(x) \quad = \quad \begin{cases} d_1 & \text{für} \quad T(x) < c_1 \text{ oder } T(X) > c_2 \\ \\ d_0 & \qquad\;\; \text{sonst} \end{cases}$$

wobei $T(x) = \frac{1}{n} \sum\limits_{i=1}^{n} x_i$ ist.

Da die Normalverteilung symmetrisch um ihren Mittelwert ist, ist eine symmetrische Wahl von c_1 und c_2 um μ_0 sinnvoll: $c_1 = \mu_0 - d,\ c_2 = \mu_0 + d$. δ ist dann ein gleichmäßig bester unverfälschter Test zum Niveau $P_I(\delta, \mu_0) = P_{\mu_0}(\delta(x) = d_1)$.

$$\begin{aligned}
P_I(\delta, \mu_0) &= 1 - P_{\mu_0}(\mu_0 - d \leq \bar{X} \leq \mu_0 + d) \\
&= 1 - [P_{\mu_o}(\bar{X} \leq \mu_0 + d) - P_{\mu_0}(\bar{X} \leq \mu_0 - d)] \\
&= 1 - \left[\Phi\left(\frac{\mu_0 + d - \mu_0}{\sigma} \cdot \sqrt{n}\right) - \Phi\left(\frac{\mu_0 - d - \mu_0}{\sigma} \cdot \sqrt{n}\right)\right] \\
&= 1 - \Phi\left(\frac{d}{\sigma}\sqrt{n}\right) + 1 - \Phi\left(\frac{d}{\sigma}\sqrt{n}\right) \\
&= 2\left(1 - \Phi\left(\frac{d}{\sigma}\sqrt{n}\right)\right) = \alpha \\
\Leftrightarrow\ & \Phi\left(\frac{d}{\sigma}\sqrt{n}\right) = 1 - \frac{\alpha}{2}, \text{ bzw. } d = \frac{\sigma}{\sqrt{n}} u_{1-\frac{\alpha}{2}}.
\end{aligned}$$

Für n=45 und $\alpha = 0.05$ erhält man mit $\sigma = 20\,000$, $u_{0.975} = 1.96$ und $d = \frac{20\,000}{\sqrt{45}} \cdot 1.96 = 5\,843.6$ die folgende Entscheidungsfunktion:

$$\delta(x) = \begin{cases} d_1 & \text{für} \quad \bar{x} < 444\,156.4 \text{ oder } \bar{x} > 455\,843.6 \\ d_0 & \text{für} \quad 444\,156.4 \leq \bar{x} \leq 455\,843.6 \end{cases}$$

Für $\bar{x} = 447\,000$ wird H_0 also nicht abgelehnt. Es ist also nicht statistisch gesichert, dass $\mu \neq 450\,000$ ist.

(b) Es ist zu ermitteln, ob die Kampagne erfolgreich war, d.h. der Absatz angestiegen ist. Damit ergibt sich ein einseitiger Test für den Erwartungswert des Absatzes.

$$\begin{aligned} H_0 &: \mu_0 \leq 450\,000 && \text{'nicht erfolgreich'} \\ H_1 &: \mu > \mu_0 && \text{'erfolgreich'} \end{aligned}$$

Teststatistik ist $\bar{x} = \frac{1}{n}\sum x_i$ mit der Entscheidungsfunktion

$$\delta(x) = \begin{cases} d_1 & \text{für} \quad \bar{x} > c \\ d_0 & \text{für} \quad \bar{x} \leq c \end{cases}$$

Änderung gegenüber Übung 54: σ ist unbekannt und wird durch $s^* = 21\,000$ geschätzt. Die Testschranke c ergibt sich aus dem Niveau $\alpha = 0.05$ gemäß:

$$P_{\mu_0}(\bar{X} > c) = 1 - P_{\mu_0}\left(\frac{\bar{X} - \mu_0}{\frac{s^*(X)}{\sqrt{n}}} \leq \frac{c - \mu_0}{\frac{s^*(X)}{\sqrt{n}}}\right) \stackrel{!}{=} \alpha.$$

Da $\frac{\sqrt{n}(\bar{X}-\mu_0)}{s^*(X)}$ $t(n-1)-$verteilt ist, muss für das $(1-\alpha)-$Quantil $t(n-1)_{1-\alpha}$ der $t(n-1)-$Verteilung gelten:

$$t(n-1)_{1-\alpha} = \sqrt{n}\frac{c - \mu_0}{s^*}.$$

Daraus ergibt sich $c = \mu_0 + \frac{s^*}{\sqrt{n}} t(n-1)_{1-\alpha}$. Die Entscheidungsfunktion lautet damit:

$$\delta(x) \quad = \quad \begin{cases} d_0 & \text{für} \quad \bar{x} \leq \mu_0 + \frac{s^*}{\sqrt{n}} \quad t(n-1)_{1-\alpha} \\ & \\ d_1 & \qquad \text{sonst} \end{cases}$$

Mit $n = 16,\ s^* = 21\,000$ und $t(15)_{0.95} = 1.753$ ist $c = 450\,000 + \frac{21\,000}{\sqrt{16}} 1.753 = 459\,203.25$.

Wegen $\bar{x} = 460\,000$ wird H_0 abgelehnt, die Werbekampagne war erfolgreich (statistisch gesichert zum Niveau 0.05).

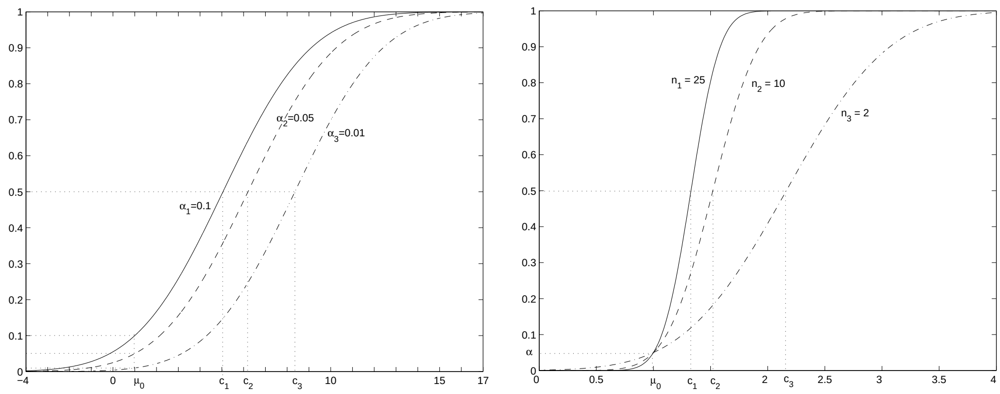
(c) Die Gütefunktion gibt die Wahrscheinlichkeit an, die Nullhypothese abzulehnen, in Abhängigkeit vom Parameter γ:

$$G(\gamma) = P_\gamma(\delta(X) = d_1)$$

Demgegenüber heißt die Wahrscheinlichkeit in Abhängigkeit von γ, die Nullhypothese nicht abzulehnen, Operations-Charakteristik (OC-Funktion):

$$L(\gamma) \quad = \quad P_\gamma(\delta(x) = d_0) = 1 - G(\gamma)$$

Abbildungen zum einseitigen Fall $H_0 : \mu \leq \mu_0$ gegen $H_1 : \mu > \mu_0$:



(a) Gütefunktionen für $\sigma^2 = 50,\ \mu_0 = 1, n{=}5$ und verschiedene α

(b) Gütefunktionen für $\alpha = 0.05, \mu_0 = 1$ und $\sigma = 1$ und verschiedene n

Abbildung 12: Gütefunktionen

Bemerkung: Da bei den in den Grafiken betrachteten Fällen die Teststatistik $\bar{x}$ symmetrisch um ihren Mittelwert μ verteilt ist, gilt für jede Testschranke c: $G(c) = P_{\mu=c}(\bar{X} > c) = \frac{1}{2}$, d.h. stimmt der wahre Mittelwert mit der Testschranke überein, so wird jeweils mit Wahrscheinlichkeit $\frac{1}{2}$ die Entscheidung d_0 bzw. d_1 getroffen.

Es folgenden einige mögliche Eigenschaften von Tests, die im Zusammenhang mit der Gütefunktion stehen.

- Ein Test ist ein *Test zum Niveau* α falls für seine Gütefunktion gilt: $G(\gamma) \leq \alpha$ für alle $\gamma \in \Gamma_0$ (hier: $\mu \leq \mu_0$).

- Ein Test ist ein *gleichmäßig bester Test zum Niveau* α falls gilt:
  - $G(\gamma) \leq \alpha$ für alle $\gamma \in \Gamma_0$ (Test zum Niveau α)
  - $G(\gamma) \geq G'(\gamma)$ für alle Gütefunktionen G' zu Tests zum Niveau α und alle $\gamma \in \Gamma_1$.
- Ein Test heißt *unverfälschter Test zum Niveau* α falls gilt:
  - $G(\gamma) \leq \alpha$ (Test zum Niveau α)
  - $G(\gamma) > \alpha$ für alle $\gamma \in \Gamma_1$ ("unverfälscht").

Der Einfluss von α und n auf die Kurve ist den Abbildungen zu entnehmen.

Aufgabe 55

Zehn Hohlkarabiner einer bestimmten Marke wurden der Produktion zufällig entnommen und einem Zerreißversuch unterzogen, bei dem die Belastung eines Karabiners jeweils so lange kontinuierlich erhöht wurde, bis er brach. Der Bruch trat bei folgenden Werten (in $[kp]$) auf:

$$2100, 2130, 2150, 2170, 2210, 2070, 2230, 2150, 2230, 2200$$

(Daten aus "Deutscher Alpenverein", 1, München 1979, Februar).

(a) Testen Sie unter der Voraussetzung, dass die Karabiner bzgl. ihrer Bruchlast X einer Normalverteilung genügen, die Hypothese

$$H_0: \ \sigma^2 = 1600 \qquad \text{gegen} \qquad H_1: \ \sigma^2 \neq 1600,$$

wobei σ^2 die Varianz der Bruchlast in der Gesamtheit aller produzierten Karabiner des untersuchten Typs bedeute, zum Niveau $\alpha = 0.05$.

(b) Es besteht die Vermutung, dass die tatsächliche Varianz größer als $\sigma^2 = 1600$ ist. Führen Sie einen geeigneten Test durch.

(c) Führen Sie den Test aus (a) durch, falls Sie den Erwartungswert mit $\mu = 2100\ kp$ kennen.

Lösung: (Induktive Statistik, S.196ff.)

(a) Situation: μ unbekannt, zweiseitiger Test.

Als Teststatistik wird $T(X) = \frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i - \bar{x})^2$ verwendet. Unter der Voraussetzung, dass $X_1, \ldots, X_n$ unabhängig und $N(\mu, \sigma^2)$-verteilt sind, ist $\frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i - \bar{X})^2 \quad \chi^2(n-1)-$verteilt.

Es sind Testschranken c_1, c_2 mit $P_{\sigma_0^2}(T(X) \in [c_1, c_2]) = 1-\alpha$ für die Entscheidungsfunktion

$$\delta(x_1, \ldots, x_n) \ = \ \begin{cases} d_1 & T(x) < c_1 \text{ oder } T(x) > c_2 \\ & \\ d_0 & c_1 \leq T(x) \leq c_2 \end{cases}$$

zu bestimmen. Mit

$$P_{\sigma_0^2}(T(X) < c_1) = P_{\sigma_0^2}(T(X) > c_2) = \frac{\alpha}{2}$$

("Equal-Tails-Test") erhält man eine Lösung, die für große n nahezu übereinstimmend mit dem gleichmäßig besten unverfälschten Test ist. Damit ist

$$c_1 = \chi^2(n-1)_{\frac{\alpha}{2}} \text{ und } c_2 = \chi^2(n-1)_{1-\frac{\alpha}{2}}.$$

Die Entscheidungsfunktion lautet damit:

$$\delta(x) \ = \ \begin{cases} d_0 & \text{für} \quad T(x) \in [\chi^2(n-1)_{\frac{\alpha}{2}}, \chi^2(n-1)_{1-\frac{\alpha}{2}}] \\ & \\ d_1 & \qquad \text{sonst} \end{cases}$$

Für $n = 10$ und $\alpha = 0.05$ erhält man

$$\chi^2(9)_{0.025} = 2.70 \quad \text{ und } \quad \chi^2(9)_{0.975} = 19.02,$$

und wegen $T(X) = \frac{26640}{1600} = 16.65$ gilt: $\delta(x) = d_0$, d.h. H_0 kann nicht abgelehnt werden.

(b) Es soll überprüft werden, ob die Varianz größer als 1600 ist, d.h. es handelt sich um einen einseitigen Test:

$$H_0 \ : \ \sigma^2 \leq \sigma_0^2 = 1600, \quad H_1 \ : \ \sigma^2 > \sigma_0^2$$

Der gleichmäßig beste (unverfälschte) Test lautet:

$$\delta(x) \quad = \quad \begin{cases} d_0 & \text{für} \quad \sum(x_i - \bar{x})^2 \frac{1}{\sigma_0^2} \leq \chi^2(n-1)_{1-\alpha} \\ \\ d_1 & \qquad \text{sonst} \end{cases}$$

Für $n = 10$ und $\alpha = 0.05$ erhält man

$$\chi^2(n-1)_{1-\alpha} = \chi^2(9)_{0.95} = 16.92$$

und wegen $\frac{\sum(x_i-\bar{x})^2}{\sigma_0^2} = 16.65 < 16.92$ gilt: $\delta(x) = d_0$, d.h. die Nullhypothese wird nicht abgelehnt.

(c) Im Gegensatz zu (a) ist jetzt μ bekannt. Die Teststatistik lautet jetzt $T(x) = \sum(x_i - \mu)^2$ und $\frac{1}{\sigma_0^2}\sum_{i=1}^{n}(X_i - \mu)^2$ ist $\chi^2(n)-$verteilt. Damit ist

$$\delta(x) \quad = \quad \begin{cases} d_0 & \text{für} \quad T(x) \in [\chi^2(n)_{\frac{\alpha}{2}}, \chi^2(n)_{1-\frac{\alpha}{2}}] \\ \\ d_1 & \qquad \text{sonst} \end{cases}$$

die Entscheidungsfunktion. Mit dem Wert $\alpha = 0.05$ ergibt sich:

$$\chi^2(10)_{0.025} = 3.25 \text{ und } \chi^2(10)_{0.975} = 20.48$$

und wegen $T(x) = \frac{67600}{1600} = 42.25 > 20.48$ gilt: $\delta(x) = d_1$, d.h. H_0 wird abgelehnt.

Bemerkung: (b) mit $\mu = 2100$ kp führt ebenfalls zur Ablehnung von H_0.

Aufgabe 56

(a) In der englischsprachigen Literatur sowie bei der Verwendung von Software-Paketen werden Sie im Zusammenhang mit Tests oftmals auf den so genannten "p-value" stossen. Erklären Sie, was man unter diesem Begriff versteht. Geben Sie an, wie sich Ihre Vorgehensweise beim Testen ändert, wenn Sie über den p-Wert argumentieren wollen.

(b) Geben Sie bei den Aufgaben 53-55 jeweils den p-Wert an.

Lösung: (Induktive Statistik, S.201ff.)

(a) Der $p-$Wert gibt das $\alpha-$Niveau an, bei dem die Nullhyphothese für eine gegebene Stichprobe gerade noch nicht abgelehnt wird (d.h. die maximale Fehlerwahrscheinlichkeit 1. Art des Tests für alle Testschranken bei denen die Nullhypothese bei dem vorliegenden Stichprobenergebnis nicht abgelehnt wird).

Bisher: Vorgabe des α-Niveaus erlaubte die Bestimmung eines Nichtablehnungsbereichs für die Teststatistik.

Jetzt: Der Wert der Statistik wird in einem p-Wert umgerechnet und dieser dann mit α verglichen:

$\alpha \leq p$-Wert : Nullhypothese nicht ablehnen
$\alpha > p$-Wert : Nullhypothese ablehnen

(b)

zu 53: Das Niveau des Tests

$$\delta_c(x) = \begin{cases} d_0 & \text{für} \quad \sum_{i=1}^{n} x_i \leq c \\ \\ d_1 & \text{für} \quad \sum_{i=1}^{n} x_i > c \end{cases}$$

nimmt ab, wenn c größer wird. Da die Nullhypothese nicht abgelehnt werden soll, wird die Testschranke möglichst klein gewählt, d.h. auf den Stichprobenwert gesetzt. Es muss also gelten:

$$\begin{aligned} 15\,120 &= \mu \cdot n + \sigma\sqrt{n} \cdot u_{1-\alpha} \\ &= 15\,000 + 100\sqrt{10}u_{1-\alpha} \\ \Rightarrow u_{1-\alpha} &= \frac{120}{100\sqrt{10}} = 0.3795, \text{ d.h. } \Phi(0.3795) = 1-\alpha. \end{aligned}$$

Daraus ergibt sich der p-Wert: $1 - \Phi(0.3795) = 0.351975$. Da $\alpha = 0.05 < 0.351975$, wird H_0 nicht abgelehnt.

zu 54 (a): Das Niveau des Tests

$$\delta(x) = \begin{cases} d_0 & \text{für} \quad \mu_0 - d \leq T(x) \leq \mu_0 + d \\ \\ d_1 & \quad\quad \text{sonst} \end{cases}$$

sinkt mit wachsendem d. Damit ist d möglichst klein zu wählen, ohne dass H_0 abgelehnt wird. $d \geq 0$ wird also so bestimmt, dass $\mu_0 - d$ oder $\mu_0 + d$ mit dem Wert t_0 der Teststatistik bei der Stichprobe übereinstimmt. Der $p-$Wert ergibt sich dann für

$$t_0 = \mu - d \quad \text{zu} \quad 2 \cdot P(T(X) < t_0)$$

und für

$$t_0 = \mu + d \quad \text{zu} \quad 2 \cdot P(T(X) > t_0)$$

Allgemein gilt beim zweiseitigen Test mit Teststatistik $T(x)$:
p-Wert= $2 \cdot \min\{P_{\gamma_0}(T(X) < t_0), P_{\gamma_0}(T(X) > t_0)\}$,
wobei t_0 hier der beobachtete Wert der Teststatistik ist. Mit $t_0 = \bar{x} = 447\,000 < \mu_0$ ist $d = 3000$ und

$$\begin{aligned} P_{\gamma_0}(T(X) < t_0) &= \Phi\left(\frac{t_0 - \mu_0}{\sigma}\sqrt{n}\right) = \Phi(-1.0062) \\ &= 1 - \Phi(1.0062) = 0.158655. \\ P_{\gamma_0}(T(X) > t_0 &= 1 - \Phi\left(\frac{t_0 - \mu_0}{\sigma}\sqrt{n}\right) = 1 - \Phi_{\mu_0}(T(X) < t_0) \\ &= 0.841345. \end{aligned}$$

Damit ist der p-Wert: $2 \cdot 0.158655 = 0.31731$

zu 54 (b): Die Entscheidungsfunktion in der Aufgabe lautet

$$\delta(x) = \begin{cases} d_0 & \text{für} \quad \bar{x} \leq \mu_0 + \frac{s^*}{\sqrt{n}} t(n-1)_{1-\alpha} \\ & \\ d_1 & \text{sonst} \end{cases}$$

Der Wert der Teststatistik ist $\bar{x} = 460\,000$. Der $p-$Wert ist dasjenige Niveau β für welches gilt

$$\begin{aligned} \bar{x} &\stackrel{!}{=} \mu_0 + \frac{s^*}{\sqrt{n}} t(n-1)_{1-\beta} \\ \Leftrightarrow \frac{460\,000 - 450\,000}{21\,000} \cdot \sqrt{16} = 1.905 &\stackrel{!}{=} t(15)_{1-\beta} \\ \stackrel{(*)}{\Rightarrow} \beta &= 0.0381. \end{aligned}$$

$(*)$ Dies kann beispielweise mit der Funktion *TINV* in *Excel* berechnet werden.

zu 55(a): Die Entscheidungsfunktion lautet:

$$\delta(x) = \begin{cases} d_0 & \text{für} \quad \frac{1}{\sigma_0^2}\sum_{i=1}^{n}(x_i - \bar{x})^2 \in [\chi^2(n-1)_{\frac{\alpha}{2}}, \chi^2(n-1)_{1-\frac{\alpha}{2}}] \\ & \\ d_1 & \text{sonst} \end{cases}$$

Der Wert der Teststatistik ist $\frac{1}{\sigma_0^2}\sum_{i=1}^{n}(x_i - \bar{x})^2 = 16.65$.

(1) $\chi^2(n-1)_{\frac{\beta}{2}} \stackrel{!}{=} 16.65 \stackrel{(*)}{\Longrightarrow} \beta = 2 \cdot 0.9455 = 1.891$

(2) $\chi^2(n-1)_{1-\frac{\beta}{2}} \stackrel{!}{=} 16.65 \stackrel{(*)}{\Longrightarrow} \beta = 2 \cdot 0.0545 = 0.109$

$(*)$ Dies kann beispielweise mit der Funktion *CHIVERT* in *Excel* berechnet werden.
Hieraus folgt: Der $p-$Wert ist 0.109.

zu 55(b): Die Entscheidungsfunktion lautet:

$$\delta(x) = \begin{cases} d_0 & \text{für } \frac{1}{\sigma_0^2}\sum_{i=1}^{n}(x_i-\bar{x})^2 \leq \chi^2(n-1)_{1-\alpha} \\ \\ d_1 & \text{sonst} \end{cases}$$

Der $p-$Wert ergibt sich für:

$$\begin{aligned} \frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i-\bar{x})^2 &\overset{!}{=} \chi^2(n-1)_{1-\beta} \\ \Leftrightarrow 16.65 &\overset{!}{=} \chi^2(n-1)_{1-\beta} \\ \Rightarrow \beta &= 0.0545 \end{aligned}$$

Der $p-$Wert ist also 0.0545.

zu 55(c): Analog zu 55(a) mit der $\chi^2(n)-$Verteilung. Es ergibt sich ein $p-$Wert von $1.354 \cdot 10^{-5}$.

Aufgabe 57

Im Rahmen einer medizinischen Untersuchung wird das Gewicht von 15-jährigen Jungen bestimmt. Es ergaben sich folgende 20 Werte:

49.1	55.0	44.9	53.8	60.4	51.6	53.2	41.2	58.3	50.4
56.1	56.5	47.8	43.6	60.5	47.3	59.7	55.2	57.1	54.5

(a) Aus medizinischer Sicht interessant ist die Frage, ob das Gewicht 15-jähriger Jungen approximativ als normalverteilt angenommen werden kann. Überlegen Sie sich, wie man eine solche Vermutung testen könnte. Gehen Sie dabei insbesondere auf die Punkte

- Hypothesenwahl
- Formulierung einer Teststatistik
- Annahme- bzw. Ablehnungsbereich

ein.

(b) Testen Sie die Hypothese, dass das Gewicht 15-jähriger Jungen normalverteilt ist mit $\mu = 50$ und $\sigma = 5$. Verwenden Sie dazu den Kolmogorov-Smirnov-Anpassungstest.

(c) Könnte man mit Hilfe des Kolmogorov-Smirnov-Anpassungstests auch überprüfen, ob die Gewichtsverteilung 15-jähriger Jungen einer Normalverteilung $N(50, \sigma^2)$ mit unbekanntem σ^2 entspricht?

Lösung: (Wahrscheinlichkeitstheorie S.189)

(a) Die neue Situation ist hier: nicht bestimmte Parameterwerte einer gegebenen Verteilung sind Gegenstand der Entscheidung, sondern 'die Verteilung selbst' soll getestet werden.

Hypothesen: Sei P die Wahrscheinlichkeitsverteilung der betrachteten Zufallsvariable Y (hier das Gewicht eines zufällig gewählten 15-jährigen Jungen). Sei $\mathcal{P}$ eine Klasse von Wahrscheinlichkeitsverteilungen, von der wir wissen möchten, ob sie die Wahrscheinlichkeitsverteilung P als Spezialfall enthält (wie also die Klasse der Normalverteilungen $N(\mu, \sigma^2)$. Dann lauten die Hypothesen: $H_0 : P \in \mathcal{P}$ $H_1 : P \notin \mathcal{P}$
Falls wie im Teil (c) die Übereinstimmung mit einer speziellen Verteilung P_0 (dort $N(50, 25)$) überprüft werden soll, vereinfachen sich die Hypothesen zu ($\mathcal{P} = \{P_0\}$) $H_0 : P = P_0$ $H_1 : P \neq P_1$.

Teststatistik: Gemessen werden könnte im einfacheren zweiten Testproblem ein 'Abstand' zwischen der Verteilung P_0 und einer mit Hilfe der Beobachtungen ermittelten Verteilung. Eine Möglichkeit: Vergleich von empirischer und theoretischer Verteilungsfunktion, entweder an bestimmten Stellen oder über den ganzen Definitionsbereich der Funktionen, also über ganz $\mathbb{R}$.
Bei dem schwierigen ersten Testproblem könnte zunächst mit den Daten eine spezielle Verteilung $\hat{P}$ in $\mathcal{P}$ geschätzt werden (z.B. durch Schätzen von μ und σ^2 bei der Klasse der Normalverteilungen) und dann ein "Abstand" von $\hat{P}$ und einer aus den Daten ermittelten Verteilung berechnet werden. Auch hier könnten die empirische Verteilungsfunktion und die Verteilungsfunktion der geschätzten Verteilung herangezogen werden.

Annahme-/Ablehnungs-bereich	Ein geringer Abstand der beiden Verteilungen wird H_0 stützen, während große Abweichungen eher gegen H_0 sprechen.

Kolmogorov-Smirnov-Anpassungstest

Voraussetzung:	P_0 hat eine stetige Verteilungsfunktion $F_0(x)$.
Hypothesen:	$H_0 : P = P_0 \quad H_1 : P \neq P_0$
Stichprobe:	$X_1, \ldots, X_n$ unabhängig identisch verteilt mit Wahrscheinlichkeitsverteilung P $(i = 1, \ldots, n)$
Vorgehen:	Es wird der maximale Abstand der empirischen Verteilungsfunktion $F_n(x)$ zur theoretischen Verteilungsfunktion $F_0(x)$ bestimmt. Überschreitet dieser eine gewisse Grenze, so wird die Nullhypothese abgelehnt.

Teststatistik ist somit die Supremums-Distanz

$$d(F_0, F_n) = \sup_{\alpha \in \mathbb{R}} |F_0(\alpha) - F_n(\alpha)|$$

kritischer Bereich:	$\{x = (x_1, \ldots, x_n) \mid z = d(F_0, F_n) > z_0\}$ Die kritische Grenze z_0 beim Kolmogorov-Smirnov-Anpassungstest ist in Abhängigkeit vom Niveau für $n < 40$ tabelliert.

(b) $P_0 = N(50, 25)$:

$$\text{empirische Verteilungsfunktion } F_n(x) \quad = \begin{cases} 0 & \text{für} \quad x < 41.2 \\ 0.05 & \text{für} \quad 41.2 < x < 43.6 \\ 0.1 & \text{für} \quad 43.6 < x < 44.9 \\ \vdots & \\ 0.95 & \text{für} \quad 60.4 \leq x < 60.5 \\ 1 & \text{für} \quad x \geq 60.5 \end{cases}$$

$F_0(x)$ ist die Verteilungsfunktion der Normalverteilung $N(50, 25)$. $F_0(x)$ ist nicht analytisch darstellbar. Es genügt jedoch, nur die Werte von $F_0(x)$ an den Sprungstellen der empirischen Verteilungsfunktion. $F_n(x)$ zu betrachten.

$$z = d\ (F_0(\alpha), F_n(\alpha)) = \sup_{\alpha \in \mathbb{R}} |F_0(\alpha) - F_n(\alpha)| = 0.3389 \text{ (bei } \alpha = 53.2)$$

Testschranke zum Niveau 0.05 ist nach Tabelle (zweiseitiger Test): 0.294. Damit ist H_0 abzulehnen.

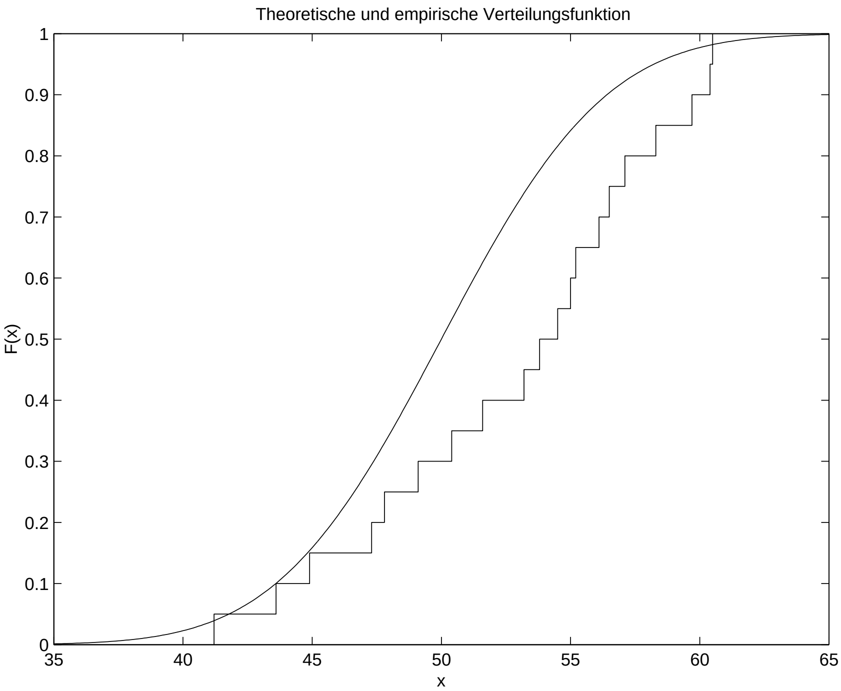


Abbildung 13: Schaubild der Verteilungsfunktionen

Anmerkung: Die senkrechten Striche bei der empirischen Verteilungsfunktion sollten weggedacht werden.

c) Die Hypthese, dass die Gewichtsverteilung einer Normalverteilung entspricht lässt sich mit Hilfe des Kolmogorov-Smirnov-Anpassungstest über einen Umweg überprüfen. Es ist bekannt, dass wenn $X_1, \ldots, X_n$ eine einfache Stichprobe zur Zufallsvariablen $X \sim N(50, \sigma^2)$ ist, dann gilt:

$$\frac{X_i - 50}{s^*} \sim t(n-1)$$

d.h. man testet die Hypothesen:

$$H_0 : \frac{X - 50}{s^*} \sim t(n-1) \quad \text{gegen} \quad H_1 : \frac{X - 50}{s^*} \nsim t(n-1)$$