Bericht über die Zulassungsprüfung Stochastik im Mai 2009

Die Zulassungsprüfung Stochastik fand am 9. Mai 2009 statt. Von den maximal 90 erreichbaren Punkten waren mindestens 36 Punkte zum Bestehen notwendig. Von 53 Teilnehmern haben 44 die Prüfung bestanden. Zugelassen waren ein Taschenrechner, die ausgeteilte Formelsammlung und Wahrscheinlichkeitstafeln.

Wir gehen stets von einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, \mathbb{P})$ aus.

Aufgabe 1 (18 Punkte)
Sei $B \in \mathcal{A}$ mit $P(B) > 0$. Sei $Q : \mathcal{A} \longrightarrow \mathbb{R}$,

$$Q(A) := \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}.$$

(a) Für $A \in \mathcal{A}$ sei $1_A : \Omega \longrightarrow \{0, 1\}$ mit

$$1_A(\omega) = \begin{cases} 1 & \omega \in \mathcal{A} \\ 0 & \text{sonst} \end{cases}$$

die Indikatorfunktion von A. Zeigen Sie für alle $A \in \mathcal{A}$:

(i) $1_A \cdot 1_B = 1_{A \cap B}$

(ii) $\int_B 1_A \, d\mathbb{P} = \mathbb{P}(A \cap B)$

(b) Sei $X \geq 0$ eine Zufallsvariable. Beweisen Sie

$$\int_\Omega X dQ = \frac{1}{\mathbb{P}(B)} \int_B X \, d\mathbb{P}.$$

Hinweis: Beweisen Sie diese Gleichung erst für Treppenfunktionen.

Lösung
Zu (a)
(i) Es gilt

$$1_A(\omega) \cdot 1_B(\omega) = 1 \Longleftrightarrow \omega \in A \text{ und } \omega \in B \Longleftrightarrow \omega \in A \cap B \Longleftrightarrow 1_{A \cap B}(\omega) = 1.$$

(ii) Mit (i) gilt

$$\int_B 1_A d\mathbb{P} \quad = \quad \int_\Omega 1_B \cdot 1_A d\mathbb{P} = \int_\Omega 1_{B \cap A} d\mathbb{P} = \mathbb{P}(A \cap B).$$

Zu (b)
Sei $X = \sum_{i=1}^n \alpha_i 1_{A_i}$, $\alpha_i \geq 0$, $A_i \in \mathcal{A}$, $i = 1, \ldots, n$ eine Treppenfunktion.
Dann gilt laut Konstruktion des Integrals

$$\int_\Omega X \, dQ = \sum_{i=1}^n \alpha_i Q(A_i) = \sum_{i=1}^n \alpha_i \frac{\mathbb{P}(A_i \cap B)}{\mathbb{P}(B)}.$$

Andererseits gilt wegen (a)(ii)

$$\int_B X d\mathbb{P} \quad = \quad \int_B \sum_{i=1}^n \alpha_i 1_{A_i} d\mathbb{P} = \sum_{i=1}^n \alpha_i \int_B 1_{A_i} d\mathbb{P} \overset{(a)}{=} \sum_{i=1}^n \alpha_i \mathbb{P}(B \cap A_i)$$

und somit die Behauptung für Treppenfunktionen.
Sei nun $X \geq 0$. Es gibt eine Folge $(X_n)_{n\in\mathbb{N}}$ von Treppenfunktionen mit $X_n \geq 0$ und $X_n \nearrow X$.
Da X_n eine Treppenfunktion ist, gilt wie oben gezeigt

$$\int_\Omega X_n \, dQ = \frac{1}{\mathbb{P}(B)} \int_B X_n \, d\mathbb{P}.$$

Aus $X_n \nearrow X$ folgt $X_n 1_B \nearrow X 1_B$ und mit dem Satz von der monotonen Konvergenz folgt für die beiden Seiten

$$\begin{aligned} \lim_{n\to\infty} \int_\Omega X_n \, dQ &= \int_\Omega X \, dQ, \\ \lim_{n\to\infty} \int_B X_n \, d\mathbb{P} &= \lim_{n\to\infty} \int_\Omega 1_B X_n \, d\mathbb{P} = \int_\Omega 1_B X \, d\mathbb{P} = \int_B X d\mathbb{P}. \end{aligned}$$

und somit insgesamt:

$$\int_\Omega X \, dQ = \frac{1}{\mathbb{P}(B)} \int_B X \, d\mathbb{P}.$$

Aufgabe 2 (16 Punkte)
Sei $f : \mathbb{R} \longrightarrow [0, \infty)$ mit $f(x) = ax^2(1-x)1_{[0,1]}(x)$ die Dichte einer Zufallsvariablen X und sei $Y := 1/X$.

(a) Bestimmen Sie $a \in \mathbb{R}$.

(b) Bestimmen Sie die Verteilungsfunktion von X.

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

(d) Bestimmen Sie die Kovarianz von X und Y.

(e) Besitzt Y eine Dichte? Bestimmen Sie diese gegebenenfalls.

Lösung
Zu (a)
f ist messbar und $f \geq 0$. Der Parameter a ist so zu wählen, dass $\int_\mathbb{R} f = 1$ gilt.

$$\int_0^1 x^2(1-x) \, dx \quad = \quad \left[\frac{1}{3}x^3 - \frac{1}{4}x^4\right]_0^1 = \frac{1}{12}$$

Es folgt $a = 12$.
Zu (b)
Für $x \in (0, 1)$ gilt

$$\mathbb{P}(X \leq x) \quad = \quad \int_0^x 12 \, t^2(1-t) \, dt = 12 \left[\frac{1}{3}t^3 - \frac{1}{4}t^4\right]_0^x = 4x^3 - 3x^4.$$

Für $x \leq 0$ gilt $\mathbb{P}(X \leq x) = 0$ und für $x \geq 1$ gilt $\mathbb{P}(X \leq x) = 1$.
Zu (c)

$$\begin{aligned}
\mathbb{E}(X) &= \int_{\mathbb{R}} x f(x)dx = \int_0^1 12\, x^3(1-x)dx = 12\left[\frac{x^4}{4} - \frac{x^5}{5}\right]_0^1 = 0,6 \\
\mathbb{E}(Y) &= \int_{\mathbb{R}} \frac{1}{x} f(x)dx = \int_0^1 12\, x(1-x)dx = 12\left[\frac{x^2}{2} - \frac{x^3}{3}\right]_0^1 = 2
\end{aligned}$$

Zu (d)
Es gilt

$$\operatorname{Cov}(X,Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y) = \mathbb{E}\left(X \cdot \frac{1}{X}\right) - 0,6 \cdot 2 = 1 - 1,2 = -0,2.$$

Zu (e)
Für die Verteilungsfunktion von Y gilt für $y \geq 1$ wegen (b)

$$\mathbb{P}(Y \leq y) = \mathbb{P}\left(\frac{1}{X} \leq y\right) = \mathbb{P}\left(\frac{1}{y} \leq X\right) = 1 - \mathbb{P}\left(X \leq \frac{1}{y}\right) = 1 - \frac{4}{y^3} + \frac{3}{y^4}.$$

Für $y < 1$ gilt $\mathbb{P}(Y \leq y) = 0$.
Damit folgt aus dem Hauptsatz der Differential- und Integralrechnung, dass Y die Dichte $g : \mathbb{R} \longrightarrow \mathbb{R}$,

$$g(y) = \begin{cases} 12\left(\frac{1}{y^4} - \frac{1}{y^5}\right) & y \geq 1 \\ 0 & \text{sonst} \end{cases}$$

besitzt.
Aufgabe 3 (16 Punkte) Die Zufallsvariable (X,Y) besitze die gemeinsame Dichte $f : \mathbb{R}^2 \longrightarrow [0,\infty)$ mit

$$f(x,y) = \begin{cases} e^{-x} & \text{für } 0 \leq y \leq x \\ 0 & \text{sonst} \end{cases}$$

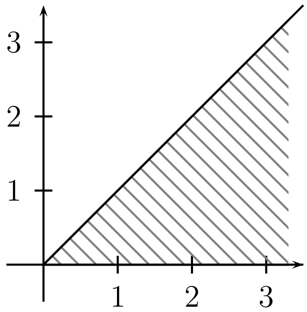
(a) Skizzieren Sie die Menge auf der $f > 0$ gilt.

(b) Bestimmen Sie die Randdichten.

(c) Bestimmen Sie $\mathbb{E}(X|Y)$.

(d) Sind X und Y unabhängig? Begründen Sie Ihre Antwort.

Lösung
Zu (a)
Es gilt $\{f > 0\} = \{(x,y) \in \mathbb{R}^2 : 0 \leq y \leq x\}$.



Zu (b)
Für $x \geq 0$ bzw. $y \geq 0$ ist

$$\begin{aligned} \int_{\mathbb{R}} f(x,y)dy &= \int_0^x e^{-x} dy = xe^{-x} \\ \int_{\mathbb{R}} f(x,y)dx &= \int_y^\infty e^{-x} dx = e^{-y}. \end{aligned}$$

Somit ergeben sich die Randdichten f_Y bzw. f_X von X bzw. Y zu

$$\begin{aligned} f_X(x) &= \begin{cases} xe^{-x} & x \geq 0 \\ 0 & x < 0 \end{cases} \\ f_Y(y) &= \begin{cases} e^{-y} & y \geq 0 \\ 0 & y < 0 \end{cases} \end{aligned}$$

Zu (c)
Für $y \geq 0$ und $0 \leq y \leq x$ ergibt sich die bedingte Dichte von X gegeben $Y = y$

$$f(x|y) = \frac{f(x,y)}{f_Y(y)} = \frac{e^{-x}}{e^{-y}} = e^{y-x},$$

also

$$\begin{aligned} \mathbb{E}(X|Y=y) &= \int_0^\infty x f(x|y)dx = \int_y^\infty xe^{y-x}dx = e^y \int_y^\infty xe^{-x}dx \\ &= e^y \left[-xe^{-x} - e^{-x}\right]_y^\infty = e^y(ye^{-y} + e^{-y}) = y + 1. \end{aligned}$$

Damit folgt $\mathbb{E}(X|Y) = Y + 1$.
Zu (d)
Wären X und Y unabhängig, dann müsste

$$f(x,y) = f_X(x)\, f_Y(y)$$

gelten. Dies steht im Widerspruch zu (b).

Aufgabe 4 (12 Punkte)
Die Dichte $f : \mathbb{R} \longrightarrow [0, \infty)$ einer Zufallsvariablen X sei gegeben durch

$$f(x) = (1 + \vartheta)x^\vartheta 1_{(0,1)}(x)$$

mit $\vartheta > 0$. Bestimmen Sie einen Maximum-Likelihood Schätzer für ϑ.

Lösung
Die gemeinsame Dichte der unabhängigen Zufallsvariablen $X_i \sim X$, $i = 1, \ldots, n$, führt zu einer Likelihood L und ihrer Loglikelihood ℓ:

$$\begin{aligned}
L(\vartheta) &= \prod_{i=1}^{n}(1+\vartheta)x_i^{\vartheta} = (1+\vartheta)^n \prod_{i=1}^{n} x_i^{\vartheta} \\
\ell(\vartheta) &= n\ln(1+\vartheta) + \vartheta \sum_{i=1}^{n} \ln x_i \\
\ell'(\vartheta) &= \frac{n}{1+\vartheta} + \sum_{i=1}^{n} \ln x_i \\
\ell''(\vartheta) &= -\frac{n}{(1+\vartheta)^2} < 0
\end{aligned}$$

Es ergibt sich für

$$\hat{\vartheta} = -\frac{n}{\sum\limits_{i=1}^{n} \ln x_i} - 1$$

das Maximum von ℓ und somit der ML-Schätzer

$$-\frac{n}{\sum\limits_{i=1}^{n} \ln X_i} - 1$$

für ϑ.

Aufgabe 5 (14 Punkte)

(a) Seien $X_1, \ldots, X_n$ unabhängig. Zeigen Sie

$$\mathbb{E}\left(\left[\sum_{i=1}^{n} X_i\right]^2\right) = \sum_{i=1}^{n} \mathbb{E}(X_i^2) + \sum_{i,j=1, i\neq j}^{n} \mathbb{E}(X_i)\mathbb{E}(X_j).$$

(b) Seien $X_1, \ldots, X_n$ unabhängig und $X_i \sim \operatorname{Exp}(\lambda)$ mit $\lambda > 0$ für alle $i = 1, \ldots, n$. Sei $T : \mathbb{R}^n \longrightarrow \mathbb{R}$ gegeben durch

$$T(x_1, \ldots, x_n) = a\left(\sum_{i=1}^{n} x_i\right)^2.$$

mit $a \in \mathbb{R}$. Bestimmen Sie a so, dass $T(X_1, \ldots, X_n)$ ein erwartungstreuer Schätzer für $1/\lambda^2$ ist.

Lösung
Zu (a)
Zunächst gilt

$$\left(\sum_{i=1}^{n} X_i\right)^2 = \sum_{i=1}^{n}\sum_{j=1}^{n} X_i X_j = \sum_{i=1}^{n} X_i^2 + \sum_{i,j=1, i\neq j}^{n} X_i X_j.$$

Für $i \neq j$ folgt wegen der Unabhängigkeit von X_i und X_j

$$\mathbb{E}(X_i X_j) = \mathbb{E}(X_i)\mathbb{E}(X_j)$$

und aufgrund der Linearität des Erwartungswerts

$$\mathbb{E}\left[\left(\sum_{i=1}^{n} X_i\right)^2\right] \quad = \quad \sum_{i=1}^{n} \mathbb{E}(X_i^2) + \sum_{i,j=1, i\neq j}^{n} \mathbb{E}(X_i)\mathbb{E}(X_j).$$

Zu (b)
Es gilt $\mathbb{E}(X) = \dfrac{1}{\lambda}$, $\mathrm{Var}(X) = \dfrac{1}{\lambda^2}$ und somit $\mathbb{E}(X^2) = \dfrac{2}{\lambda^2}$. Mit (a) erhalten wir

$$\mathbb{E}\left[\left(\sum_{i=1}^{n} X_i\right)^2\right] = \sum_{i=1}^{n} \frac{2}{\lambda^2} + \sum_{i,j=1, i\neq j}^{n} \frac{1}{\lambda}\frac{1}{\lambda} = n\,\frac{2}{\lambda^2} + (n^2 - n)\,\frac{1}{\lambda^2} = \frac{n^2+n}{\lambda^2}.$$

Damit der Schätzer erwartungstreu ist, muss man $a = \dfrac{1}{n^2+n}$ wählen.

Aufgabe 6 (14 Punkte)
In der folgenden Tabelle sind das Körpergewicht und der Blutdruck von zufällig ausgewählten Personen aufgezeichnet.

Nr.	Gewicht	Blutdruck	Nr.	Gewicht	Blutdruck
1	82,5	130	11	86	153
2	83,5	133	12	79,5	128
3	90	150	13	84	132
4	77,5	128	14	87	149
5	106	151	15	91,5	158
6	87,5	146	16	107,5	150
7	95	150	17	97,5	163
8	105	140	18	90	156
9	100	148	19	71,5	124
10	74,5	125	20	120	170

Bezeichnet man mit x_i das Gewicht und y_i den Blutdruck, so wird für die Zufallsvariable Y_i das Modell

$$Y_i = a + bx_i + \varepsilon_i,\ i = 1, \ldots, 20$$

untersucht. Die ε_i seien unabhängig und $N(0, \sigma^2)$-verteilt.

(a) Bestimmen Sie Schätzwerte für a, b und σ^2.

(b) Stützen die Daten die Behauptung, dass der Blutdruck nicht vom Gewicht abhängt? Die Aussage soll zu einem Niveau von $\alpha = 10$ % erfolgen. Sie können verwenden:

$$\begin{aligned}
\overline{x} &= 90,8 \\
\sum_{i=1}^{20}(x_i - \overline{x})^2 &= 2872,7 \\
\overline{y} &= 144,2 \\
\sum_{i=1}^{20}(y_i - \overline{y})^2 &= 3409,2 \\
\sum_{i=1}^{20}(y_i - \overline{y})(x_i - \overline{x}) &= 2392,3
\end{aligned}$$

Lösung
Zu (a)
Es ergeben sich folgende Schätzwerte:

$$\begin{aligned}
\hat{b} &= \frac{2392,3}{2872,7} = 0,833 \\
\hat{a} &= \overline{y} - \hat{b}\overline{x} = 144,2 - 0,833 \cdot 90,8 = 68,6 \\
\hat{\sigma}^2 &= \frac{1}{18}\left(3409,2 - \frac{2392,3^2}{2872,7}\right) = 78,72
\end{aligned}$$

Zu (b) Zu prüfen ist die Hypothese $b = 0$.

$$\begin{aligned}
\mathrm{se}(\hat{b}) &= \sqrt{\frac{78,72}{2872,7}} = 0,166 \\
T &= \frac{\hat{b} - 0}{\mathrm{se}(\hat{b})} = \frac{0,833}{0,166} = 5,018.
\end{aligned}$$

Die Hypothese $b = 0$ wird abgelehnt, denn $|T| > t_{18;0,95} = 1,734$. Somit wird die Hypothese, dass der Blutdruck nicht vom Alter abhängt, abgelehnt.