

Name:

Matrikelnummer:

Vorname:

Studiengang:

Klausur zur Vorlesung
Wahrscheinlichkeitstheorie und Statistik
für Studierende des Maschinenbaus

Datum: 22. Juli 2016

Dauer: 180 Minuten

Aufgabe	1	2	3	4	5	6	7	Σ
Punkte								

ACHTUNG:

- Die Antworten brauchen *nicht* begründet werden!
Vereinfachen Sie die Ergebnisse soweit wie möglich. Geben Sie die Ergebnisse so exakt wie möglich an (z.B. als Bruch), runden Sie Endergebnisse ggf. auf 3 Nachkommastellen.
- Als **Hilfsmittel** sind zugelassen: Skriptum zur Vorlesung, Vorlesungsfolien, Taschenrechner (nicht vernetzbar), Wörterbuch
- Schreiben Sie nicht mit radierfähigen Stiften. Die Farben rot und grün sind nicht zugelassen, da diese bei der Korrektur verwendet werden.

Hinreichend zum Bestehen der Klausur sind 17 von 49 Punkten.

VIEL ERFOLG!

Aufgabe 1 (8 Punkte)

Vorgegeben sei die folgende 2-dimensionale Stichprobe vom Umfang $n = 17$.

j	1	2	3	4	5	6	7	8	9
x_j	2.2	2.3	2.3	2.3	2.4	2.4	2.5	2.5	2.7
y_j	3.2	3.3	3.1	2.9	2.4	2.9	2.7	2.4	2.6
j	10	11	12	13	14	15	16	17	
x_j	2.8	2.8	2.9	2.9	2.9	3.2	3.4	3.4	
y_j	2.4	3.1	2.2	2.1	2.3	1.7	1.6	1.6	

Rechenhilfe: Es gilt

$$\sum_{j=1}^{17} x_j = 45.9, \quad \sum_{j=1}^{17} x_j^2 = 126.29, \quad \sum_{j=1}^{17} y_j = 42.5, \quad \sum_{j=1}^{17} y_j^2 = 111.05, \quad \sum_{j=1}^{17} x_j \cdot y_j = 111.78.$$

1. Bestimmen Sie die arithmetischen Mittel $\bar{x}$ und $\bar{y}$ sowie die Stichprobenvarianzen s_x^2 und s_y^2 .

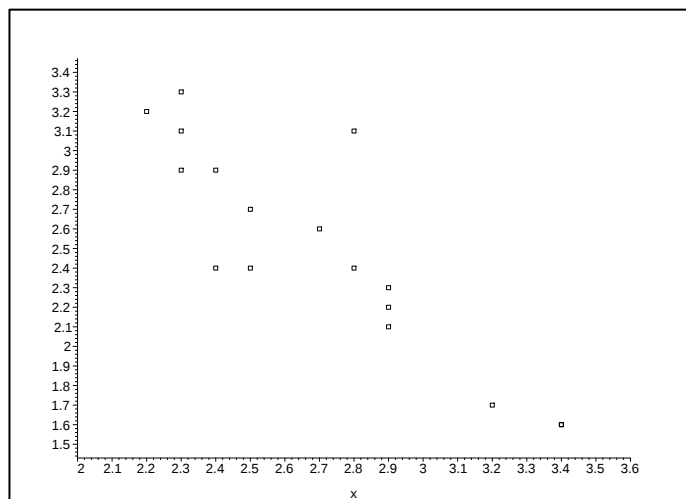
$\bar{x} = 2.7$	$\bar{y} = 2.5$
$s_x^2 = 0.1475$	$s_y^2 = 0.3$

2. Bestimmen Sie den Pearson-Korrelationskoeffizienten r_{xy} .

$$r_{xy} = -0.8824$$

3. Bestimmen Sie die Regressionsgerade $y = a^* + b^* \cdot x$, und zeichnen Sie die Regressionsgerade in das folgende Punktediagramm ein.

$a^* = 5.8979$	$b^* = -1.2585$
----------------	-----------------



Punktwolke zum obigen Datensatz.

Lösung:

1. Es gilt (vergleiche Rechenhilfe)

$$\bar{x} = \frac{1}{17} \cdot \sum_{j=1}^{17} x_j = \frac{1}{17} \cdot 45.9 = 2.7$$

und

$$\bar{y} = \frac{1}{17} \cdot \sum_{j=1}^{17} y_j = \frac{1}{17} \cdot 42.5 = 2.5.$$

Weiter gilt

$$s_x^2 = \frac{1}{16} \cdot \left(\sum_{j=1}^{17} x_j^2 - 17 \cdot \bar{x}^2 \right) = \frac{1}{16} \cdot (126.29 - 17 \cdot (2.7)^2) = 0.1475$$

und

$$s_y^2 = \frac{1}{16} \cdot \left(\sum_{j=1}^{17} y_j^2 - 17 \cdot \bar{y}^2 \right) = \frac{1}{16} \cdot (111.05 - 17 \cdot (2.5)^2) = 0.3.$$

2. Der Pearson-Korrelationskoeffizient r_{xy} ergibt sich zu

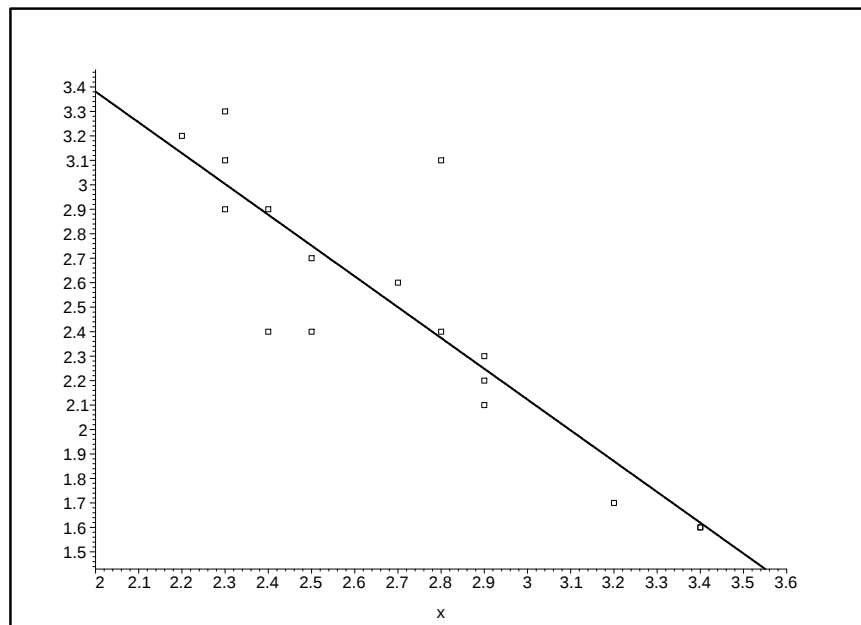
$$r_{xy} = \frac{\frac{1}{16} \cdot \left(\sum_{j=1}^{17} x_j \cdot y_j - 17 \cdot \bar{x} \cdot \bar{y} \right)}{\sqrt{s_x^2} \cdot \sqrt{s_y^2}} = \frac{\frac{1}{16} \cdot (111.78 - 17 \cdot 2.7 \cdot 2.5)}{\sqrt{0.1475} \cdot \sqrt{0.3}} \approx -0.8824.$$

3. Für die Koeffizienten a^* und b^* der Regressionsgeraden ergibt sich

$$b^* = r_{xy} \cdot \frac{\sqrt{s_y^2}}{\sqrt{s_x^2}} \approx -0.8824 \cdot \frac{\sqrt{0.3}}{\sqrt{0.1475}} \approx -1.2585$$

und

$$a^* = \bar{y} - b^* \cdot \bar{x} \approx 2.5 + 1.2585 \cdot 2.7 \approx 5.8979.$$



Punktwolke zum obigen Datensatz mit eingezeichneter Regressionsgerade.

Aufgabe 2 (6 Punkte)

Auf Grund langjähriger Beobachtungen kann man davon ausgehen, dass die mittlere Lufttemperatur (in Grad Celsius) im Monat Mai in Karlsruhe näherungsweise $\mathcal{N}(14, 4)$ -verteilt ist. X_1, X_2, X_3, X_4 bezeichnen die mittleren Lufttemperaturen des Monats Mai in den nächsten vier Jahren. Sie können dabei annehmen, dass die Temperaturen in den einzelnen Jahren unabhängig voneinander sind.

1. Berechnen Sie die Wahrscheinlichkeit p_1 , dass die mittlere Lufttemperatur im Mai nächsten Jahres höchstens 16 Grad Celsius beträgt.

$$p_1 = 0.8413$$

2. Berechnen Sie die Wahrscheinlichkeit p_2 , dass die mittlere Lufttemperatur im Mai nächsten Jahres mindestens 15.2 Grad Celsius beträgt.

$$p_2 = 0.2743$$

3. Berechnen Sie die Wahrscheinlichkeit p_3 , dass die mittlere Lufttemperatur im Mai nächsten Jahres höchstens 11.4 Grad Celsius beträgt.

$$p_3 = 0.0968$$

4. Bestimmen Sie die Verteilung der Zufallsvariablen $Y := X_1 + X_2 + X_3 + X_4$.

$$Y \sim \mathcal{N}(56, 16)$$

5. Wie groß ist die Wahrscheinlichkeit p_4 , dass der Durchschnitt der Mai-Temperaturen in den nächsten vier Jahren (also $\frac{1}{4} \cdot (X_1 + X_2 + X_3 + X_4)$) höchstens 16 Grad Celsius beträgt?

$$p_4 = 0.9772$$

6. Bestimmen Sie die Verteilung der Zufallsvariablen $Z := 1 + 2 \cdot X_1 - 3 \cdot X_4$

$$Z \sim \mathcal{N}(-13, 52)$$

Lösung:

Nach Voraussetzung sind die Zufallsvariablen X_1, X_2, X_3, X_4 stochastisch unabhängig, und es gilt

$$X_1, X_2, X_3, X_4 \sim \mathcal{N}(14, 4).$$

1. Es gilt

$$p_1 = \mathbb{P}(X_1 \leq 16) = \mathbb{P}\left(\frac{X_1 - 14}{2} \leq \frac{16 - 14}{2}\right) = \Phi_{0,1}(1) = 0.8413.$$

2. Es gilt

$$\begin{aligned} p_2 &= \mathbb{P}(X_1 \geq 15.2) = 1 - \mathbb{P}(X_1 \leq 15.2) = 1 - \mathbb{P}\left(\frac{X_1 - 14}{2} \leq \frac{15.2 - 14}{2}\right) \\ &= 1 - \Phi_{0,1}(0.6) = 1 - 0.7257 = 0.2743. \end{aligned}$$

3. Es gilt

$$\begin{aligned} p_3 &= \mathbb{P}(X_1 \leq 11.4) = \mathbb{P}\left(\frac{X_1 - 14}{2} \leq \frac{11.4 - 14}{2}\right) \\ &= \Phi_{0,1}(-1.3) = 1 - \Phi_{0,1}(1.3) = 1 - 0.9032 = 0.0968. \end{aligned}$$

4. Da die Zufallsvariablen X_1, X_2, X_3, X_4 stochastisch unabhängig sind, folgt

$$Y = X_1 + X_2 + X_3 + X_4 \sim \mathcal{N}(4 \cdot 14, 4 \cdot 4) = \mathcal{N}(56, 16).$$

5. Es folgt

$$Y = \frac{1}{4} \cdot (X_1 + X_2 + X_3 + X_4) \sim \mathcal{N}(14, 1),$$

und damit ergibt sich

$$p_4 = \mathbb{P}(Y \leq 16) = \mathbb{P}\left(\frac{Y - 14}{1} \leq \frac{16 - 14}{1}\right) = \Phi_{0,1}(2) = 0.9772.$$

6. Es gilt

$$Z = 1 + 2 \cdot X_1 - 3 \cdot X_4 \sim \mathcal{N}(1 + 2 \cdot 14 - 3 \cdot 14, 4 \cdot 4 + 9 \cdot 4) = \mathcal{N}(-13, 52).$$

Aufgabe 3 (7 Punkte)

Die Reparatur eines bestimmten Bauteils einer Maschine erfolgt in zwei Stufen. Die zufälligen Reparaturzeiten X und Y für die beiden Stufen seien $\exp(\frac{1}{2})$ - bzw. $\exp(\frac{1}{3})$ -verteilt und stochastisch unabhängig. Die Zufallsvariablen U und Z seien definiert durch

$$U := Y + X \text{ und } Z = Y - X.$$

1. Bestimmen Sie die folgenden Erwartungswerte.

$\mathbb{E}(U) = 5$	$\mathbb{E}(Z) = 1$
---------------------	---------------------

2. Welche der folgenden Aussagen ist wahr? Kreuzen Sie die wahre Aussage an.

$$\mathbb{V}(U) > \mathbb{V}(Z) \quad \square \quad \mathbb{V}(U) = \mathbb{V}(Z) \quad \boxed{\times} \quad \mathbb{V}(U) < \mathbb{V}(Z) \quad \square$$

3. Bestimmen Sie die folgende Wahrscheinlichkeit.

$\mathbb{P}(U > 4 + Z) = e^{-1}$

4. Bestimmen Sie die beiden folgenden Kovarianzen.

$C(U, X) = 4$	$C(U, Z) = 5$
---------------	---------------

5. Bestimmen Sie den folgenden Erwartungswert.

$\mathbb{E}(Z \cdot U) = 10$

Lösung:

Ist eine Zufallsvariable exponentialverteilt mit dem Parameter $\lambda > 0$, gilt also

$$X \sim \exp(\lambda),$$

so gilt

$$\mathbb{E}(X) = \frac{1}{\lambda} \quad \text{und} \quad \mathbb{V}(X) = \frac{1}{\lambda^2}.$$

1. Es gilt

$$\mathbb{E}(U) = \mathbb{E}(Y + X) = \mathbb{E}(Y) + \mathbb{E}(X) = 3 + 2 = 5$$

und

$$\mathbb{E}(Z) = \mathbb{E}(Y - X) = \mathbb{E}(Y) - \mathbb{E}(X) = 3 - 2 = 1.$$

2. Da die beiden Zufallsvariablen X und Y stochastisch unabhängig sind, gilt

$$\mathbb{V}(U) = \mathbb{V}(Y + X) = \mathbb{V}(Y) + \mathbb{V}(X) = 9 + 4 = 13$$

und

$$\mathbb{V}(Z) = \mathbb{V}(Y - X) = \mathbb{V}(Y) + \mathbb{V}(-X) = \mathbb{V}(Y) + \mathbb{V}(X) = 13.$$

3. Es gilt

$$\begin{aligned} \mathbb{P}(U > 4 + Z) &= \mathbb{P}(Y + X > 4 + Y - X) = \mathbb{P}(2 \cdot X > 4) \\ &= \mathbb{P}(X > 2) = 1 - \mathbb{P}(X \leq 2) \\ &= 1 - \left(1 - e^{-\frac{1}{2} \cdot 2}\right) = e^{-1}. \end{aligned}$$

4. Da die beiden Zufallsvariablen X und Y stochastisch unabhängig sind, gilt

$$C(U, X) = C(Y + X, X) = C(Y, X) + \mathbb{V}(X) = \mathbb{V}(X) = 4$$

und

$$C(U, Z) = C(Y + X, Y - X) = \mathbb{V}(Y) - \mathbb{V}(X) = 9 - 4 = 5.$$

5. Es gilt

$$\begin{aligned} \mathbb{E}(Z \cdot U) &= \mathbb{E}((Y - X) \cdot (Y + X)) = \mathbb{E}(Y^2) - \mathbb{E}(X^2) \\ &= \mathbb{V}(Y) + (\mathbb{E}(Y))^2 - \mathbb{V}(X) - (\mathbb{E}(X))^2 \\ &= 9 + 9 - 4 - 4 = 10. \end{aligned}$$

Aufgabe 4 (7 Punkte)

Ein Bedienungssystem bestehe aus einem Schalter (mit einem Platz) und einem Warteraum mit 3 Plätzen. Ein ankommender Kunde findet X Kunden im System ($:=$ Schalter und Warteraum). Dabei gelte

$$\mathbb{P}(X = k) = c \cdot p^k, \quad k = 0, 1, 2, 3, 4,$$

mit $p \in (0, 1)$ und einer Konstanten $c \in \mathbb{R}$. Im Fall $X = 4$ wird der ankommende Kunde abgewiesen.

1. Bestimmen Sie (in Abhängigkeit von $p \in (0, 1)$) die Konstante $c \in \mathbb{R}$.

$$c = \frac{1 - p}{1 - p^5}$$

2. Vorgegeben sei das Ereignis

$A := \{\text{Ankommender Kunde wird nicht abgewiesen, muss aber auf die Bedienung warten.}\}$

- (a) Stellen Sie das Ereignis A mit Hilfe der Zufallsvariablen X dar.

$$A = \{1 \leq X \leq 3\}$$

- (b) Berechnen Sie die folgende Wahrscheinlichkeit

$$\mathbb{P}(A) = c \cdot (p + p^2 + p^3)$$

- (c) Berechnen Sie für $0 \leq k \leq 4$ die folgende bedingte Wahrscheinlichkeit.

$$\mathbb{P}(X = k \mid A) = \begin{cases} 0, & \text{falls } k = 0 \text{ oder } k = 4, \\ \frac{p^{k-1}}{1 + p + p^2}, & \text{falls } 1 \leq k \leq 3. \end{cases}$$

3. Findet ein Kunde bei seiner Ankunft X Kunden im System vor, so entstehen Kosten der Höhe

$$K := \begin{cases} 2 \cdot X, & \text{falls } 0 \leq X \leq 3, \\ 10, & \text{falls } X = 4. \end{cases}$$

- (a) Berechnen Sie die folgende bedingte Wahrscheinlichkeit.

$$\mathbb{P}(K \geq 4 \mid A) = \frac{p + p^2}{1 + p + p^2}$$

- (b) Berechnen Sie den Erwartungswert der Zufallsvariablen K .

$$\mathbb{E}(K) = 2 \cdot c \cdot (p + 2 \cdot p^2 + 3 \cdot p^3 + 5 \cdot p^4)$$

Lösung:

1. Es gilt

$$\sum_{k=0}^4 P(X = k) = c \cdot \sum_{k=0}^4 p^k = c \cdot \frac{1-p^5}{1-p} \stackrel{!}{=} 1 \iff c = \frac{1-p}{1-p^5}.$$

2.(a) Es gilt

$$A = [1 \leq X \leq 3].$$

2.(b) Weiter gilt

$$\begin{aligned} \mathbb{P}(A) &= \mathbb{P}(1 \leq X \leq 3) = \mathbb{P}(X = 1) + \mathbb{P}(X = 2) + \mathbb{P}(X = 3) \\ &= c \cdot (p + p^2 + p^3) = \frac{1-p}{1-p^5} \cdot (p + p^2 + p^3) \\ &= \frac{p - p^4}{1 - p^5} \end{aligned}$$

2.(c) und

$$\begin{aligned} \mathbb{P}(X = k \mid A) &= \frac{\mathbb{P}(X = k, 1 \leq X \leq 3)}{\mathbb{P}(A)} \\ &= \begin{cases} 0, & \text{falls } k = 0 \text{ oder } k = 4, \\ \frac{p^k}{p + p^2 + p^3} = \frac{p^{k-1}}{1 + p + p^2}, & \text{falls } 1 \leq k \leq 3. \end{cases} \end{aligned}$$

3.(a) Es gilt

$$\begin{aligned} \mathbb{P}(K \geq 4 \mid A) &= \mathbb{P}(X \geq 2 \mid A) \\ &= \frac{\mathbb{P}(X \geq 2, 1 \leq X \leq 3)}{\mathbb{P}(A)} = \frac{\mathbb{P}(X = 2) + \mathbb{P}(X = 3)}{\mathbb{P}(A)} \\ &= \frac{p^2 + p^3}{p + p^2 + p^3} = \frac{p + p^2}{1 + p + p^2}. \end{aligned}$$

3.(b) Weiter gilt

$$\begin{aligned} \mathbb{E}(K) &= \sum_{j=0}^3 2 \cdot j \cdot \mathbb{P}(X = j) + 10 \cdot \mathbb{P}(X = 4) \\ &= 2 \cdot c \cdot (p + 2 \cdot p^2 + 3 \cdot p^3 + 5 \cdot p^4). \end{aligned}$$

Aufgabe 5 (6 Punkte)

Gegeben sei eine Urne mit 4 roten, 4 schwarzen und 4 weißen Kugeln. Der Urne werden nacheinander und ohne Zurücklegen Kugeln entnommen, in jedem Zug jeweils genau eine.

1. Wie groß ist die Wahrscheinlichkeit p_1 , dass nach $n = 3$ Ziehungen alle gezogenen Kugeln rot sind ?

$$p_1 = \frac{1}{55}$$

2. Wie groß ist die Wahrscheinlichkeit p_2 , dass nach $n = 3$ Ziehungen alle gezogenen Kugeln dieselbe Farbe haben ?

$$p_2 = \frac{3}{55}$$

3. Es werden $n = 5$ Ziehungen vorgenommen. Weiter seien die drei folgenden Ereignisse definiert:

$A_1 :=$ „es wird in 5 Durchgängen keine rote Kugel gezogen“,
 $A_2 :=$ „es wird in 5 Durchgängen keine schwarze Kugel gezogen“,
 $A_3 :=$ „es wird in 5 Durchgängen keine weiße Kugel gezogen“.

- (a) Sind die Ereignisse A_1, A_2 und A_3 paarweise disjunkt?

Ja

- (b) Bestimmen Sie die folgende Wahrscheinlichkeit.

$$\mathbb{P}(A_1 \cup A_2) = \frac{14}{99}$$

- (c) Kann das Ereignis $B :=$ „Nach fünf Ziehungen wurden alle Farben mindestens einmal gezogen.“ mit Hilfe der Ereignisse A_1, A_2 und A_3 ausgedrückt werden, wenn ja, wie?

Ja, als $A_1^c \cap A_2^c \cap A_3^c$.

- (d) Wie groß ist die Wahrscheinlichkeit p_3 , dass das Ereignis B eintritt?

$$p_3 = \frac{26}{33}$$

Lösung:

1. Markiere die roten Kugeln mit einem „+“ und die anderen Kugel mit einem „−“, es liegt die Situation der hypergeometrischen Verteilung $\text{hyp}(n, r, s)$ vor mit den Parametern

$$n = 3, r = 4, s = 8.$$

Es folgt

$$p_1 = \frac{\binom{4}{3} \cdot \binom{8}{0}}{\binom{12}{3}} = \frac{4 \cdot 6}{12 \cdot 11 \cdot 10} = \frac{1}{55}.$$

2. Mit derselben Modellierung wie in 1. ergibt sich

$$p_2 = 3 \cdot \frac{1}{55}.$$

3. (a) Es ist nicht möglich, dass in 5 Ziehungen keine rote und keine schwarze Kugel enthalten sind, weil es nur 4 andere Kugeln gibt. Daher sind A_1 und A_2 disjunkt. Analog erhalten wir den Rest der Behauptung.
(b) Damit ergibt sich

$$\begin{aligned} \mathbb{P}(A_1 \cup A_2) &= \mathbb{P}(A_1) + \mathbb{P}(A_2) - \mathbb{P}(A_1 \cap A_2) \\ &= \mathbb{P}(A_1) + \mathbb{P}(A_2) \\ &= 2 \cdot \frac{7}{99} = \frac{14}{99}. \end{aligned}$$

- (c) Rot mindestens einmal zu ziehen wird durch das Ereignis A_1^c beschrieben. Daher ist $A_1^c \cap A_2^c \cap A_3^c$ das Ereignis jede Farbe mindestens einmal zu ziehen.

- (d) Es folgt

$$p_3 = \mathbb{P}(A_1^c \cap A_2^c \cap A_3^c) = 1 - \mathbb{P}(A_1 \cup A_2 \cup A_3) = 1 - 3 \cdot \frac{7}{99} = \frac{26}{33}.$$

Aufgabe 6 (7 Punkte)

Wir betrachten die Partie Frankreich gegen Deutschland der EM2020. Die Anzahl der Tore T_D die Deutschland bzw. T_F die Frankreich schießt, seien $\text{Po}(\lambda_D)$ bzw. $\text{Po}(\lambda_F)$ verteilte Zufallsvariablen.

1. Wie ist die Gesamtanzahl der Tore verteilt, wenn wir annehmen, dass T_D und T_F unabhängig sind?

Sie ist $\text{Po}(\lambda_D + \lambda_F)$ verteilt.

2. Berechnen Sie die Verteilung der Anzahl der Tore die Deutschland schießt unter der Bedingung, dass in der Partie genau 3 Tore fallen.

Hinweis: es gilt $(\lambda_D + \lambda_F)^3 = (\lambda_D + \lambda_F)^k \cdot (\lambda_D + \lambda_F)^{3-k}$ für $k = 0, \dots, 3$

$$\mathbb{P}(T_D = k \mid T_D + T_F = 3) = \binom{3}{k} \left(\frac{\lambda_D}{\lambda_D + \lambda_F} \right)^k \left(\frac{\lambda_F}{\lambda_D + \lambda_F} \right)^{3-k}$$

3. Geben Sie den Namen und die Parameter dieser Verteilung an.

$\text{Bin}(3, \frac{\lambda_D}{\lambda_D + \lambda_F})$

4. Wir möchten eine Langzeitprognose aufgrund der aktuellen Form der deutschen Mannschaft erstellen. Dazu nehmen wir an, dass auch in den nächsten n Begegnungen der deutschen Mannschaft mit Frankreich die Anzahl der deutschen Tore durch die Zufallsvariablen $T_1, T_2, \dots, T_n$ beschrieben wird. Diese seien alle unabhängig voneinander und $\text{Po}(2)$ verteilt.

Geben Sie den Erwartungswert und die Varianz von $S_n := \sum_{k=1}^n T_k$ an.

$$\mathbb{E}(S_n) = 2n$$

$$\mathbb{V}(S_n) = 2n$$

5. Berechnen Sie nun den Limes für $n \rightarrow \infty$ von der Wahrscheinlichkeit, dass in diesen n Partien insgesamt mehr als $2n + \sqrt{n}$ bzw. mehr als $3n$ Tore fallen.

$$\lim_{n \rightarrow \infty} \mathbb{P}(S_n > 2n + \sqrt{n}) = 1 - \Phi\left(\frac{1}{\sqrt{2}}\right) = 0.2389$$

$$\lim_{n \rightarrow \infty} \mathbb{P}(S_n > 3n) = 0$$

Lösung:

1. Folgt direkt aus dem Additionssatz für Poisson-verteilte Zufallsvariablen.
2. Es gilt

$$\begin{aligned}\mathbb{P}(T_D = k \mid T_D + T_F = 3) &= \frac{\mathbb{P}(T_D = k, T_F = 3 - k)}{\mathbb{P}(T_D + T_F = 3)} \\ &= \frac{\frac{\lambda_D^k}{k!} e^{-\lambda_D} \frac{\lambda_F^{3-k}}{(3-k)!} e^{-\lambda_F}}{\frac{(\lambda_D + \lambda_F)^3}{3!} e^{-(\lambda_D + \lambda_F)}} \\ &= \binom{3}{k} \left(\frac{\lambda_D}{\lambda_D + \lambda_F} \right)^k \left(\frac{\lambda_F}{\lambda_D + \lambda_F} \right)^{3-k}.\end{aligned}$$

3. Wir können direkt aus der Formel ablesen, dass es sich um eine Binomial-verteilte Zufallsgröße mit Parametern 3 und $\frac{\lambda_D}{\lambda_D + \lambda_F}$ handelt.
4. Aus dem Additionssatz für Poisson-verteilte Zufallsvariablen folgt, dass $S_n \sim Po(2n)$ ist. Erwartungswert und Varianz stimmen bei der Poissonverteilung mit dem Parameter überein.
5. Der Zentrale Grenzwertsatz besagt, dass

$$\lim_{n \rightarrow \infty} \mathbb{P} \left(\frac{S_n - \mathbb{E}(S_n)}{\sqrt{\mathbb{V}(S_n)}} \leq t \right) = \Phi(t)$$

gilt. Daraus folgt $\lim_{n \rightarrow \infty} \mathbb{P}(S_n \leq 2n + \sqrt{2nt}) = \Phi(t)$. Für $t = \frac{1}{\sqrt{2}}$ ergibt sich der gesuchte Ausdruck.

Dass der zweite Limes gleich null ist, folgt unmittelbar aus dem schwachen Gesetz der großen Zahlen, wenn man $\varepsilon = 0.5$ wählt.

Aufgabe 7 (8 Punkte)

Ein Merkmal besitze die Zähldichte

$$f_{\vartheta}(k) = \frac{1}{(k - \vartheta)^2} - \frac{1}{(k + 1 - \vartheta)^2}, \quad k \in \mathbb{N}, \quad k > \vartheta.$$

mit unbekanntem Parameter $\vartheta \in \mathbb{N}$.

1. Es besitze X die Zähldichte f_{ϑ} . Geben Sie $\mathbb{E}_{\vartheta}(X)$ an. *Hinweis: Es gilt*

$$\sum_{k=1}^{\infty} \left(\frac{1}{k} - \frac{k}{(k+1)^2} \right) = \frac{\pi^2}{6}.$$

$$\mathbb{E}_{\vartheta}(X) = \vartheta + \frac{\pi^2}{6}$$

2. Bestimmen Sie den Momentenschätzer $\hat{\vartheta}_1(x)$ zur Stichprobe x .

$$\hat{\vartheta}_1(x) = \bar{x} - \frac{\pi^2}{6}$$

3. Ist $\hat{\vartheta}_1(x)$: (a) erwartungstreu, (b) nicht erwartungstreu aber asymptotisch erwartungstreu oder (c) keines von beidem?

(a) erwartungstreu

4. Geben Sie zur Stichprobe $x = (x_1, \dots, x_n)$ die Likelihood-Funktion $L_x(\vartheta)$ an.

$$L_x(\vartheta) = \begin{cases} \prod_{i=1}^n \left(\frac{1}{(x_i - \vartheta)^2} - \frac{1}{(x_i + 1 - \vartheta)^2} \right), & \text{falls } x_i > \vartheta \text{ für alle } i \in \{1, \dots, n\}, \\ 0, & \text{sonst.} \end{cases}$$

5. Geben Sie den größtmöglichen Bereich $A \subset \mathbb{N}$ an, auf dem die Abbildung $k \mapsto f_{\vartheta}(k)$ streng monoton fallend ist.

Hinweis: Betrachten Sie $\frac{1}{(x-\vartheta)^2} - \frac{1}{(x+1-\vartheta)^2}$ als Funktion von x und leiten Sie ab.

$$A = \{k \in \mathbb{N} \mid k > \vartheta\}$$

6. Geben Sie den Maximum-Likelihood-Schätzer $\hat{\vartheta}(x)$ zur Stichprobe x an.

$$\hat{\vartheta}(x) = \min_{i \in \{1, \dots, n\}} x_i - 1$$

Lösung:

1. Anschaulich ist klar, wenn die Dichte einer Zufallsvariablen um m verschoben wird, vergrößert sich der Erwartungswert ebenfalls um m . Es gilt $f_{\vartheta}(k) = f_0(k - \vartheta)$ und wegen des Hinweises ist

$$\mathbb{E}_0(X) = \sum_{k=1}^{\infty} k \left(\frac{1}{k^2} - \frac{1}{(k+1)^2} \right) = \frac{\pi^2}{6}.$$

Diese Lösung erhalten wir alternativ auch durch direkte Rechnung:

$$\begin{aligned} \mathbb{E}_{\vartheta}(X) &= \sum_{k=\vartheta+1}^{\infty} k \left(\frac{1}{(k-\vartheta)^2} - \frac{1}{(k+1-\vartheta)^2} \right) \\ &= \sum_{k=1}^{\infty} (k+\vartheta) \left(\frac{1}{k^2} - \frac{1}{(k+1)^2} \right) \\ &= \vartheta \sum_{k=1}^{\infty} \left(\frac{1}{k^2} - \frac{1}{(k+1)^2} \right) + \sum_{k=1}^{\infty} k \left(\frac{1}{k^2} - \frac{1}{(k+1)^2} \right) \\ &= \vartheta + \frac{\pi^2}{6} \end{aligned}$$

2. Es ergibt sich der Momentenschätzer als $\bar{x} - \frac{\pi^2}{6}$.
3. Es gilt

$$\mathbb{E}_{\vartheta} \left(\frac{1}{n} \sum_{k=1}^n X_k - \frac{\pi^2}{6} \right) = \mathbb{E}_{\vartheta}(X_1) - \frac{\pi^2}{6} = \vartheta + \frac{\pi^2}{6} - \frac{\pi^2}{6} = \vartheta,$$

woraus die erwartungstreue folgt.

4. Es ist lediglich zu beachten, wann das Produkt null wird.
5. Die Zähldichte ist für $k > \vartheta$ monoton fallend, denn

$$\left(\frac{1}{(x-\vartheta)^2} - \frac{1}{(x+1-\vartheta)^2} \right)' = \frac{-2}{(x-\vartheta)^3} - \frac{-2}{(x+1-\vartheta)^3}$$

ist für $x > \vartheta$ offensichtlich kleiner als null. Daher muss $f_{\vartheta}(k)$ wenigstens für $k > \vartheta$ streng monoton fallend sein. Für alle anderen Werte ist die Dichte jedoch gleich null, daher ist dies bereits der gesuchte Bereich.

6. Aus 4. folgt, dass jeder Faktor des Produkts für $\vartheta < k$ streng monoton wachsend in ϑ ist. Damit wird die Likelihood-Funktion maximal, wenn ϑ größtmöglich gewählt wird unter der Bedingung $\vartheta < x_i$ für $i = 1, \dots, n$.