Dr. Christof Luchsinger

Übungsblatt 1 zur Vorlesung

"Statistische Methoden"

Repetition WTS

Herausgabe des Übungsblattes: Woche 08, Abgabe der Lösungen: Woche 09 (bis Freitag, 1615 Uhr), Besprechung: Woche 10

Must

Aufgabe 1 [einfachste Aufgaben P, X]

a) X sei $\mathcal{N}(3,9)$-verteilt. Berechnen Sie $P[2 < X < 7]$.

b) X sei $\mathcal{N}(2,9)$-verteilt. Berechnen Sie $P[1 < X^2 < 2]$.

Aufgabe 2 [einfachste Aufgaben E, V]

a) Sei X eine $\mathcal{N}(\mu,\sigma^2)$-Zufallsgrösse und Y eine $\mathcal{N}(\mu,\sigma^2/2)$-Zufallsgrösse. Es gelte $X \coprod Y$. Wie ist die Verteilung von $X - Y$.

b) Sei X eine Zufallsgrösse mit Dichtefunktion $K \exp^{-(x-8)^{10}}$ auf $\mathbb{R}$. Geben Sie den Erwartungswert $E[X]$ an. Begründung aber ohne Beweis.

c) X sei $\mathcal{N}(3,4)$-verteilt. Berechnen Sie $E[X^2]$.

d) Sei X eine $\mathcal{N}(2,4)$-Zufallsgrösse und Y eine Po(3)-Zufallsgrösse. Berechnen Sie $E[X+Y]$, $E[X+Y+7]$ und $E[X^2+Y^2]$. Sie dürfen dazu Resultate aus der Vorlesung benutzen und müssen einzelne Erwartungswerte nicht nochmals berechnen.

Standard

Aufgabe 3 [Transformation von Zufallsgrössen] [3 Punkte]

X habe Dichte $f(x) = Kx^5$ auf dem Intervall $[0,1]$ und sei 0 sonst. Berechnen Sie

a) die Normierungskonstante K

b) $E[X]$

c) $E[1/X^2]$

d) die Verteilungsfunktion von $Y := 1/X^2$

e) die Dichte von $Y := 1/X^2$.

f) die Wahrscheinlichkeit $P[Y \in (2,3)]$.

Aufgabe 4 [Konvergenz] [2 Punkte]

Sei $X_i, i \geq 1$, eine Folge von paarweise unabhängigen, identisch verteilten Zufallsgrössen mit Erwartungswert μ und $E[|X_1|] < \infty$. Untersuchen Sie, wogegen der Ausdruck

$$\frac{1}{n^2} \sum_{i=1}^{n} X_i$$

konvergiert und beweisen Sie diese Konvergenzaussage mit Hilfe eines Satzes aus der Vorlesung WTS.

Aufgabe 5 [3 Punkte]

Sie haben eine Stichprobe $x_1 \in [0, 2]$ vom Umfang $n = 1$ (Rechnungen sind dann einfacher).

a) Testen Sie mit dem Neyman-Pearson-Lemma zum Niveau $\alpha = 0.05$, ob die Zufallsgrösse, welche diesen Wert generiert hat, eine

$\mathcal{H}_0$: Dichtefunktion der Art $f(x) = K_1 x^2$ für $x \in [0, 2]$, 0 sonst, hat, oder

$\mathcal{H}_1$: Dichtefunktion der Art $f(x) = K_2 x^3$ für $x \in [0, 2]$, 0 sonst, hat.

b) Berechnen Sie beim Test von a) das β.

c) Herr Meier weiss nichts von Neyman-Pearson und entscheidet sich mit einer einfachen Regel: Wenn $x \in [1.5, 2]$, nimmt er $\mathcal{H}_1$ an, sonst $\mathcal{H}_0$. Wie sind sein α und β?

Aufgabe 6 [4 Punkte]

Ein Hersteller behauptet, sein Starter-Gerät versagt durchschnittlich höchstens jedes 100te Mal. Eine Konsumentenschutz-Organisation bezweifelt dies uns denkt, dass es häufiger vorkommt. In 2000 unabhängigen Versuchen hat das Starter-Gerät 23 mal versagt.

Testen Sie approximativ (CLT) mit dem Neyman-Person-Lemma auf dem Niveau 0.05, ob

$\mathcal{H}_0$: die Versagenswahrscheinlichkeit $p \leq 0.01$, oder

$\mathcal{H}_1$: die Versagenswahrscheinlichkeit $p > 0.01$.

Benutzen Sie an geeigneter Stelle den CLT.

Übungsblatt 1 zur Vorlesung “Statistische Methoden”

Olivier Warin

13. Februar 2011

Aufgabe 1 [einfachste Aufgaben P,X]

a) X sei $\mathcal{N}(3,9)$-verteilt. Nun gilt

$$P[2 < X < 7] = P[X < 7] - P[X \leqslant 2] \underset{\mathrm{R}}{\doteq} \texttt{pnorm(7,3,sqrt(9)) - pnorm(2,3,sqrt(9))}$$
$$\underset{\mathrm{R}}{\doteq} 0.5393474.$$

Alternativ können wir die Z-Transformation und eine Tabelle der $\mathcal{N}(0,1)$-Verteilung benutzen und erhalten so

$$\begin{aligned} P[2 < X < 7] &= P\left[\frac{2-3}{\sqrt{9}} < \frac{X-3}{\sqrt{9}} < \frac{7-3}{\sqrt{9}}\right] = P\left[-\frac{1}{3} < \mathcal{N}(0,1) < \frac{4}{3}\right] \\ &= P\left[\mathcal{N}(0,1) < \frac{4}{3}\right] - P\left[\mathcal{N}(0,1) \leqslant -\frac{1}{3}\right] \underset{\text{Tabelle}}{\doteq} 0.9082 - (1 - 0.6293) \\ &= 0.5375. \end{aligned}$$

b) X sei $\mathcal{N}(2,9)$-verteilt. Nun gilt

$$\begin{aligned} P[1 < X^2 < 2] &= P[\sqrt{1} < X < \sqrt{2}] + P[-\sqrt{2} < X < -\sqrt{1}] \\ &= P[X < \sqrt{2}] - P[X \leqslant 1] + P[X \leqslant -1] - P[X \leqslant -\sqrt{2}] \\ &\underset{\mathrm{R}}{\doteq} \texttt{pnorm(sqrt(2),2,sqrt(9)) - pnorm(1,2,sqrt(9)) + pnorm(-1,2,sqrt(9))} \\ &\quad \texttt{- pnorm(-sqrt(2),2,sqrt(9))} \underset{\mathrm{R}}{\doteq} 0.0842624. \end{aligned}$$

Alternativ könnten wir hier natürlich, wie in Teilaufgabe a), die Z-Transformation und eine Tabelle der $\mathcal{N}(0,1)$-Verteilung benutzen.

Aufgabe 2 [einfachste Aufgaben E,V]

a) Sei X eine $\mathcal{N}(\mu,\sigma^2)$-Zufallsgrösse und Y eine $\mathcal{N}(\mu, \sigma^2/2)$-Zufallsgrösse. Es gelte $X \perp\!\!\!\perp Y$.

Nun wollen wir die Verteilung von $X - Y$ bestimmen. Dazu bemerken wir zuerst, dass aus Symmetriegründen $-Y$ offenbar eine $\mathcal{N}(-\mu, \sigma^2/2)$ hat. Mit Hilfe des Abschnittes 1.4.3 aus dem Skript schliessen wir nun

$$X - Y \;=\; X + (-Y) \;\sim\; \mathcal{N}(\mu - \mu, \sigma^2 + \sigma^2/2) = \mathcal{N}(0, 3\sigma^2/2).$$

b) Sei X eine stetige Zufallsgrösse mit Dichtefunktion $K\exp(-(x-8)^{10})$ auf $\mathbb{R}$.

Behauptung: Es gilt $E[X] = 8$.

Bemerkung: Dies ist aus Symmetriegründen eigentlich klar. Zur Vollständigkeit folgt hier noch ein Beweis.

Beweis: Nach Definition 1.14 gilt

$$\begin{aligned} E[X] &= \int_{-\infty}^{\infty} xK\exp(-(x-8)^{10})dx \overset{t\,=\,x\,-\,8}{=} \int_{-\infty}^{\infty} (t+8)K\exp(-t^{10})dt \\ &= \underbrace{K\int_{-\infty}^{\infty} \underbrace{t\exp(-t^{10})}_{\text{ungerade}}\,dt}_{=0} + 8\underbrace{\int_{-\infty}^{\infty} K\exp(-(x-8)^{10})dx}_{=1,\text{ da Dichtefunktion}} \\ &= 8. \end{aligned}$$

■

c) Sei X $\mathcal{N}(3,4)$-verteilt. Nach Lemma 1.18 b) gilt $V[X] = E[X^2] - (E[X])^2$. Wir schliessen also

$$E[X^2] = V[X] + (E[X])^2 \;=\; 4 + 3^2 \;=\; 13.$$

d) Sei X eine $\mathcal{N}(2,4)$-Zufallsgrösse und Y eine Po(3)-Zufallsgrösse. Nun gilt:

$$\begin{aligned}
E[X+Y] &\underset{1.17\text{ c)}}{\overset{\text{Lem}}{=}} E[X] + E[Y] \overset{1.4.1.5}{=} 2+3 \;=\; 5,\\
E[X+Y+7] &\underset{1.17\text{ c)}}{\overset{\text{Lem}}{=}} E[X] + E[Y] + 7 \overset{1.4.1.5}{=} 2+3+7 \;=\; 12,\\
E[X^2+Y^2] &\underset{1.17\text{ c)}}{\overset{\text{Lem}}{=}} E[X^2] + E[Y^2] \underset{1.18\text{b)}}{\overset{\text{Lem}}{=}} V[X] + (E[X])^2 + V[Y] + (E[Y])^2\\
&\overset{1.4.1.5}{=} 4 + 2^2 + 3 + 3^2 \;=\; 20.
\end{aligned}$$

Aufgabe 3 [Transformation von Zufallsgrössen]
X habe Dichte $f(x) = Kx^5$ auf dem Intervall $[0,1]$ und sei 0 sonst.

a) Da f eine Dichte ist, muss gelten

$$1 = \int_0^1 Kx^5 dx \;=\; \frac{K}{6}$$

und somit $K = 6$.

b) Nach Definition 1.14 gilt

$$E[X] = \int_0^1 x \cdot 6x^5 dx \;=\; \frac{6}{7}.$$

c) Mit Lemma 1.17 a) berechnen wir

$$E[1/X^2] = \int_0^1 \frac{1}{x^2} \cdot 6x^5 dx \;=\; \frac{3}{2}.$$

d) Nun definieren wir $Y := 1/X^2$. Wir wollen jetzt die Verteilungsfunktion F_Y von Y bestimmen. Sei dazu a eine reelle Zahl.

Falls $a < 1$ folgt sofort:

$$\begin{aligned}
F_Y(a) &\underset{1.9}{\overset{\text{Def}}{=}} P[Y \leqslant a] \;=\; P[1/X^2 \leqslant a] \;=\; P[X^2 \geqslant 1/a] \;\leqslant\; P[X^2 > 1] \;=\; P[X < -1] + P[X > 1]\\
&= \int_{-\infty}^{-1} f(x)dx + \int_1^{\infty} f(x)dx \;=\; 0
\end{aligned}$$

und damit $F_Y(a) = 0$.

Falls nun $a \geqslant 1$ folgt:

$$\begin{aligned}
F_Y(a) = P[Y \leqslant a] &\;=\; P[X^2 \geqslant 1/a] \;=\; P[X \leqslant -1/\sqrt{a}] + P[X \geqslant 1/\sqrt{a}]\\
&= \underbrace{\int_{-\infty}^{-1/\sqrt{a}} f(x)dx}_{=0} + \int_{1/\sqrt{a}}^{\infty} f(x)dx \;=\; \int_{1/\sqrt{a}}^{1} 6x^5 dx \;=\; 1 - \frac{1}{a^3}.
\end{aligned}$$

Zusammengefasst haben wir also erhalten:

$$F_Y(a) = \begin{cases} 0, & \text{falls } a < 1\\ 1 - a^{-3}, & \text{falls } a \geqslant 1. \end{cases}$$

e) Um nun eine Dichtefunktion f_Y von Y zu bestimmen, brauchen wir nur die Verteilungsfunktion aus d) abzuleiten. Wir erhalten somit

$$f_Y(a) = \begin{cases} 0, & \text{falls } a < 1 \\ 3a^{-4}, & \text{falls } a \geqslant 1. \end{cases}$$

f) Mit Hilfe von Teilaufgabe d) finden wir sofort:

$$P[Y \in (2,3)] = F_Y(3) - F_Y(2) = 1 - 3^{-3} - (1 - 2^{-3}) = \frac{19}{216} \doteq 0.08796296.$$

Aufgabe 4 [Konvergenz]
Sei X_i, $i \geqslant 1$ eine Folge von paarweise unabhängigen, identisch verteilten Zufallsgrössen mit Erwartungswert μ und $E[|X_1|] < \infty$.

Behauptung: Der Ausdruck $\frac{1}{n^2}\sum_{i=1}^n X_i$ konvergiert für n gegen unendlich in Wahrscheinlichkeit gegen 0.

Beweis: Sei n eine natürliche Zahl. Nun gilt

$$E\left[\left|\frac{1}{n^2}\sum_{i=1}^n X_i\right|\right] \leqslant E\left[\frac{1}{n^2}\sum_{i=1}^n |X_i|\right] \overset{\text{iid}}{=} \frac{1}{n^2} nE[|X_1|] = \frac{E[|X_1|]}{n} < \infty. \tag{$*$}$$

Also dürfen wir die Ungleichung von Bienayme-Tschebyschew (Satz 1.22) auf den Ausdruck $\frac{1}{n^2}\sum_{i=1}^n X_i$ anwenden. Es gilt also für jede reelle Zahl $\varepsilon > 0$:

$$0 \leqslant P\left[\left|\frac{1}{n^2}\sum_{i=1}^n X_1\right| \geqslant \varepsilon\right] \leqslant \frac{1}{\varepsilon}E\left[\left|\frac{1}{n^2}\sum_{i=1}^n X_i\right|\right] \overset{(*)}{\leqslant} \frac{E[|X_1|]}{n\varepsilon} \overset{n\to\infty}{\longrightarrow} 0.$$

Damit ist die Konvergenz in Wahrscheinlichkeit gezeigt.

■

Aufgabe 5
Wir haben eine Stichprobe $x_1 \in [0,2]$ vom Umfang $n = 1$.

a) Wir wollen nun mit dem Lemma von Neyman-Pearson zum Niveau $\alpha = 0.05$ testen, ob die Zufallsgrösse X_1, welche diesen Wert generiert hat, eine

$\mathcal{H}_0$: Dichtefunktion der Art $f(x) = K_1x^2$ für $x \in [0,2]$ und 0 sonst hat oder
$\mathcal{H}_1$: Dichtefunktion der Art $f(x) = K_2x^3$ für $x \in [0,2]$ und 0 sonst hat.

Für $i = 0,1$ bezeichnen wir die Dichtefunktion von X_1 im Fall von $\mathcal{H}_i$ mit f_i.

Nun müssen wir erst die Konstanten K_1 und K_2 bestimmen. Dazu gehen wir analog wie in Aufgabe 3 vor:

$$\begin{aligned} 1 &= \int_0^2 f_0(x)dx = \int_0^2 K_1x^2dx = \frac{8K_1}{3} \Rightarrow K_1 = \frac{3}{8} \\ 1 &= \int_0^2 f_1(x)dx = \int_0^2 K_2x^3dx = 4K_2 \Rightarrow K_2 = \frac{1}{4}. \end{aligned}$$

Nach dem Lemma von Neyman-Pearson müssen wir jetzt eine positive, reelle Zahl K finden, so dass folgendes gilt:

$$\begin{aligned} \alpha = P_0\left[\frac{f_1(X_1)}{f_0(X_1)} > K\right] &= P_0[{}^2\!/\!_3X_1 > K] = P_0[X_1 > {}^3\!/\!_2K] = \int_{\frac{3}{2}K}^\infty f_0(x)dx \\ &= \int_{\min\{\frac{3}{2}K,2\}}^2 \frac{3}{8}x^2dx = 1 - \min\{({}^3\!/\!_4K)^3, 1\}. \end{aligned}$$

Wir schliessen:

$$K = \frac{4}{3}\sqrt[3]{1-\alpha} \;=\; \frac{4}{3}\sqrt[3]{0.95}. \qquad (*)$$

Nach dem Lemma von Neyman-Pearson entscheiden wir uns genau dann für $\mathcal{H}_1$, wenn gilt

$$\frac{f_1(x_1)}{f_0(x_1)} \;>\; K \;\Leftrightarrow\; x_1 \;>\; \frac{3}{2}K.$$

(Beachte dazu, dass $\frac{f_1(x_1)}{f_0(x_1)} = \frac{2}{3}x_1$ gilt.)

Damit haben wir die folgende Entscheidungsfunktion gefunden:

$$E^{\mathrm{NP}}(x_1) = \begin{cases} 1, & \text{falls } x_1 > \frac{3}{2}K \overset{(*)}{=} 2\sqrt[3]{0.95} \doteq 1.966095 \\ 0, & \text{sonst.} \end{cases}$$

b) In dieser Teilaufgabe sollen wir noch das Risiko zweiter Art β bestimmen. Per Definition gilt

$$\beta = P_1[E(X_1)=0] \;=\; P_1[X_1 \leqslant 2\sqrt[3]{0.95}] \;=\; \int_0^{2\sqrt[3]{0.95}} f_1(x)dx \overset{\text{a)}}{=} \int_0^{2\sqrt[3]{0.95}} \frac{1}{4}x^3 dx$$
$$= 0.95^{\frac{4}{3}} \;\doteq\; 0.9338952.$$

c) Herr Meier weiss nichts von Neyman-Pearson und entscheidet sich mit einer einfachen Regel: Wenn $x \in [1.5, 2]$, nimmt er $\mathcal{H}_1$ an, sonst $\mathcal{H}_0$. Wir berechnen nun sein Risiko erster Art (α) und sein Risiko zweiter Art (β):

$$\alpha = P_0[X_1 \in [1.5,2]] \;=\; \int_{1.5}^{2} f_0(x)dx \overset{\text{a)}}{=} \int_0^{1.5} \frac{3}{8}x^2 dx \;=\; \frac{37}{64} \;\doteq\; 0.578125$$
$$\beta = P_1[X_1 \notin [1.5,2]] \;=\; \int_0^{1.5} f_1(x)dx \overset{\text{a)}}{=} \int_0^{1.5} \frac{1}{4}x^3 dx \;=\; \frac{81}{256} \;\doteq\; 0.3164062.$$

Aufgabe 6

Ein Hersteller behauptet, sein Starter-Gerät versagt durchschnittlich höchstens jedes 100te Mal. Eine Konsumentenschutz-Organisation bezweifelt dies und denkt, dass es häufiger vorkommt. In $n = 2000$ unabhängigen Versuchen hat das Starter-Gerät 23 mal versagt.

Für $i = 1, \ldots, n$ sei

$$X_i = \begin{cases} 1, & \text{falls das Starter-Gerät beim } i\text{-ten Versuch versagt hat} \\ 0, & \text{sonst.} \end{cases}$$

Die X_i's sind unabhängige Be(p)-Zufallsgrössen. Wir haben nun eine Stichprobe $(x_1, \cdots, x_n)$ mit $\sum_{i=1}^n x_i = 23$.

Wir testen die Hypothesen

$$\mathcal{H}_0\colon p \leqslant 0.01, \quad \mathcal{H}_1\colon p > 0.01$$

mit Hilfe des Neyman-Pearson-Lemmas auf dem Niveau $\alpha = 0.05$ gegeneinander. Dazu können wir genauso gut die Hypothesen

$$\mathcal{H}_0\colon p = 0.01, \quad \mathcal{H}_1\colon p > 0.01$$

gegeneinander testen.

Wie wir am Ende des Abschnitts 6.3 in der Vorlesung WTS gesehen haben, führt nun das Neyman-Pearson-Lemma darauf, dass wir eine natürliche Zahl K' finden müssen, so dass ($\overline{X} = \frac{1}{n}\sum_{i=1}^n X_i$)

$$P_0\left[n\overline{X} \leqslant K'-1\right] \;<\; 1-\alpha \;=\; 0.95 \;\;\text{und}\;\; P_0\left[n\overline{X} \leqslant K'\right] \;\geqslant\; 1-\alpha \;=\; 0.95$$

gilt. Wir benutzen nun den CLT, um dieses K' zu bestimmen: Für eine reelle Zahl L gilt nämlich

$$P_0[n\overline{X} \leqslant L] = P_0\left[\frac{n\overline{X} - nE_0[X_1]}{\sqrt{nV_0[X_1]}} \leqslant \frac{L - nE_0[X_1]}{\sqrt{nV_0[X_1]}}\right] \underset{\text{CLT}}{\doteq} P\left[\mathcal{N}(0,1) \leqslant \frac{L - 0.01n}{\sqrt{0.01 \cdot 0.99 \cdot n}}\right].$$

Da bekanntlich $P[\mathcal{N}(0,1) \leqslant 1.64] \doteq 0.95$ gilt, führt uns das zu den folgenden zwei Ungleichungen:

$$\frac{K' - 0.01n}{\sqrt{0.01 \cdot 0.99 \cdot n}} \geqslant 1.64$$
$$\frac{(K' - 1) - 0.01n}{\sqrt{0.01 \cdot 0.99 \cdot n}} < 1.64.$$

Wir schliessen

$$\begin{aligned} K' &\geqslant 1.64\sqrt{0.99 \cdot 0.01 \cdot n} + 0.01n \doteq 27.3 \\ K' - 1 &< 1.64\sqrt{0.99 \cdot 0.01 \cdot n} + 0.01n \doteq 27.3 \end{aligned}$$

Wir lesen daraus ab: $K' = 28$.

Da nun $\sum_{i=1}^{n} x_i = 23 < 28$ nehmen wir die Hypothese $\mathcal{H}_0$ an.