ETH Zürich SS 2016
Prof. Dr. P. Embrechts **August 2016**

Prüfung

Wahrscheinlichkeit und Statistik

BSc INFK

Korrekturschema

Aufgabe 1: (10 pts)
First/Second Zivoi/Gonon

(a)
- **1pt**: (1),
- **1pt**: (2),
- **1pt**: conclusion.

(b)
- **1pt**: (3),
- **1pt**: (4).

(c)
- **1pt**: (5),
- **1pt**: (6),
- **1pt**: (7).

(d)
- (**1pt**: SGGZ[1])
- **2pts/1pt**: (8).

Aufgabe 2: (10 pts)
First/Second Gonon/Liu

(a)
- **1pt**: (9),
- **1pt**: (10) *and* $F_U(u) = 0$, *else* statement,
- **1pt**: (11).

(b)
- **1pt**: (12) or correct $\mathrm{Cov}(X, U)$,
- **1pt**: (13).

(c)
- **1pt**: (14),
- **1pt**: stating *Chebychev* (*inequality*) and applied to problem,
- **1pt**: (15).

(d)
- **1pt**: (16),
- **1pt**: (17).

Aufgabe 3 : (10 pts)
First/Second Liu/Zivoi

(a)
- **1pt**: (18),
- **1pt**: $E[\xi]$,
- **1pt**: (19).

(b)
- **1pt** (20),
- **1pt**: $\sigma = 3$,
- **1pt**: (21),
- **1pt**: (22),
- **1pt**: correct conclusion.

(c)
- **1pt**: $P_{H_0}[T < t]$,
- **1pt**: solution.

[1] starkes Gesetz der grossen Zahl

ETH Zürich SS 2016
Prof. Dr. P. Embrechts **August 2016**

Prüfung

Wahrscheinlichkeit und Statistik

BSc INFK

Solutions

Lösung 1. (a) Wir verwenden folgende Notation:

X	bezeichne die Kugel, die zu Beginn entfernt wurde (B blau und R rot).
Y^{PE}	bezeichne die Anzahl roter Kugeln von 6 Ziehungen mit Wiederholung.
Y^{PA}	bezeichne die Anzahl roter Kugeln von 600 Ziehungen mit Wiederholung.

Es müssen die folgenden zwei bedingten Wahrscheinlichkeiten verglichen werden:

$$P[X = \text{B}|Y^{\text{PE}} = 6] \qquad \text{und} \qquad P[X = \text{B}|Y^{\text{PA}} = 303].$$

Dazu werden wir die Formel von Bayes verwenden.

Peter's Experiment $P[X = \text{B}|Y^{\text{PE}} = 6]$**:** Hier gilt

$$P[X = \text{B}|Y^{\text{PE}} = 6]$$

(1)

$$= \frac{P[Y^{\text{PE}} = 6|X = \text{B}]P[X = \text{B}]}{P[Y^{\text{PE}} = 6|X = \text{B}]P[X = \text{B}] + P[Y^{\text{PE}} = 6|X = \text{R}]P[X = \text{R}]}$$

$$= \frac{1}{1 + \frac{P[Y^{\text{PE}}=6|X=\text{R}]}{P[Y^{\text{PE}}=6|X=\text{B}]}},$$

da $P[X = \text{B}] = P[X = \text{R}] = \frac{1}{2}$. Desweiteren ist klar, dass

$$P[Y^{\text{PE}} = 6|X = \text{B}] = \left(\frac{3}{5}\right)^6 \qquad \text{und} \qquad P[Y^{\text{PE}} = 6|X = \text{R}] = \left(\frac{2}{5}\right)^6.$$

Damit erhalten wir

$$P[X = \text{B}|Y^{\text{PE}} = 6] = \frac{1}{1 + \frac{P[Y^{\text{PE}}=6|X=\text{R}]}{P[Y^{\text{PE}}=6|X=\text{B}]}} = \frac{1}{1 + \left(\frac{2}{3}\right)^6} \quad (\approx 91,9\%).$$

Paula's Experiment $P[X = \text{B}|Y^{\text{PA}} = 303]$**:** Die Zufallsvariable Y^{PA}, gegeben, dass $\{X = \text{B}\}$, ist Binomomialverteilt mit Parameter $n = 600$ und $p = \frac{3}{5}$, also

$$P[Y^{\text{PA}} = 303|X = \text{B}] = \binom{600}{303}\left(\frac{3}{5}\right)^{303}\left(\frac{2}{5}\right)^{297}.$$

Mittels analoger Argumentation gilt

$$P[Y^{\text{PA}} = 303|X = \text{R}] = \binom{600}{303}\left(\frac{2}{5}\right)^{303}\left(\frac{3}{5}\right)^{297}.$$

Wieder mittels Bayes' Formel folgt

$$P[X = \text{B}|Y^{\text{PA}} = 303]$$

Siehe nächste Seite!

(2)
$$= \frac{P[Y^{\mathrm{PA}} = 303|X = \mathrm{B}]P[X = \mathrm{B}]}{P[Y^{\mathrm{PA}} = 303|X = \mathrm{B}]P[X = \mathrm{B}] + P[Y^{\mathrm{PA}} = 303|X = \mathrm{R}]P[X = \mathrm{R}]}$$

$$= \frac{1}{1 + \frac{P[Y^{\mathrm{PA}}=303|X=\mathrm{R}]}{P[Y^{\mathrm{PA}}=303|X=\mathrm{B}]}}$$

$$= \frac{1}{1 + \left(\frac{2}{3}\right)^6} \quad (\approx 91,9\%).$$

Schlussfolgerung: Peter und Paula haben genau die gleiche (bedingte) Wahrscheinlichkeit, dass am Anfang eine blaue Kugel gezogen wurde.

(b) Der erste Term:

$$\begin{aligned} \mathrm{Corr}(X,Y) &= \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)}\sqrt{\mathrm{Var}(Y)}} \\ &= \frac{E[XY] - E[X]E[Y]}{\sqrt{\mathrm{Var}(X)}\sqrt{\mathrm{Var}(Y)}} \\ &= \frac{-1+2}{\sqrt{4}\sqrt{9}} \\ &= \frac{1}{6}. \end{aligned} \tag{3}$$

Der zweite Term:

$$\begin{aligned} \mathrm{Var}\left(\frac{1}{3}(X - 5Y)\right) &= \frac{1}{9}\left(\mathrm{Var}(X) + 25\mathrm{Var}(Y) - 10\mathrm{Cov}(X,Y)\right) \\ &= \frac{1}{9}\left(4 + 25 \cdot 9 - 10\,(-1+2)\right) \\ &= \frac{219}{9}. \end{aligned} \tag{4}$$

(c) Die Auszahlung ist per Annahme gleich $c\mathbb{1}_{\{Z>c\}}$. Die erwartete Auszahlung ist demzufolge gleich

$$G(c) := E\left[c\mathbb{1}_{\{Z>c\}}\right] = cP[Z > c] = ce^{-\lambda c}, \tag{5}$$

wobei die zweite Gleichung per Annahme an Z folgt. Leiten wir (5) nach c ab und setzten es gleich 0, so erhalten wir

$$\frac{\mathrm{d}}{\mathrm{d}c}G(c) = (1 - \lambda c)\,e^{-\lambda c} = 0$$

und damit

$$c = \frac{1}{\lambda}. \tag{6}$$

Da $\frac{\mathrm{d}^2}{\mathrm{d}c^2} G\left(\frac{1}{\lambda}\right) = -\lambda e^{-1} < 0$, schlussfolgern wir, dass $\frac{1}{\lambda}$ ein Maximum von G ist. Der erwartete maximale Gewinn ist demzufolge

$$G\left(\frac{1}{\lambda}\right) = \frac{1}{\lambda e}. \tag{7}$$

(d) Für $X_1 \sim \mathrm{Exp}(2)$ gilt nach dem *starken Gesetz der grossen Zahl*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} e^{X_k} = E[e^{X_1}] = 2 \int_0^\infty e^x e^{-2x} \,\mathrm{d}x = 2 \int_0^\infty e^{-x} \,\mathrm{d}x = 2. \tag{8}$$

Lösung 2. (a) Wir rechnen, für $u \geqslant 0$,

$$\begin{aligned} F_U(u) &= P\left[U \leqslant u\right] \\ &= \int_0^u \left(\int_0^{u-y} e^{-x}\,\mathrm{d}x\right) 2e^{-2y}\,\mathrm{d}y \qquad (9) \\ &= \int_0^u 2e^{-2y}\left(1-e^{-u+y}\right)\,\mathrm{d}y \\ &= \int_0^u 2e^{-2y}\,\mathrm{d}y - 2e^{-u}\int_0^u e^{-y}\,\mathrm{d}y \\ &= 1-e^{-2u}-2e^{-u}(1-e^{-u}) \\ &= 1-2e^{-u}+e^{-2u} \\ &= (1-e^{-u})^2, \qquad (10) \end{aligned}$$

und $F_U(u) = 0$, sonst. Wir erhalten

$$f_U(u) = F_U'(u) = \begin{cases} 2e^{-u}-2e^{-2u} = 2e^{-u}\left(1-e^{-u}\right), & u \geq 0 \\ 0, & \text{sonst.} \end{cases} \qquad (11)$$

(b) Die Korrelation von X und U ist gleich

$$\begin{aligned} \mathrm{Corr}(X,U) &= \frac{\mathrm{Cov}(X,U)}{\sqrt{\mathrm{Var}(X)}\sqrt{\mathrm{Var}(U)}} \\ &= \frac{\mathrm{Cov}(X,X)+\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)}\sqrt{\mathrm{Var}(U)}} \\ &= \frac{\mathrm{Var}(X)+\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)}\sqrt{\mathrm{Var}(U)}}. \end{aligned}$$

Wegen der Unabhängigkeit von X und Y, gilt $\mathrm{Cov}(X,Y) = 0$ und somit

$$\mathrm{Var}(U) = \mathrm{Var}(X) + \mathrm{Var}(Y). \qquad (12)$$

Einsetzen liefert nun

$$\mathrm{Corr}(X,U) = \frac{\mathrm{Var}(X)}{\sqrt{\mathrm{Var}(X)}\sqrt{\mathrm{Var}(X)+\mathrm{Var}(Y)}} = \sqrt{\frac{\mathrm{Var}(X)}{\mathrm{Var}(X)+\mathrm{Var}(Y)}}. \qquad (13)$$

Es gilt, für $Z \sim \mathrm{Exp}(\lambda)$, die Formel $\mathrm{Var}(Z) = \frac{1}{\lambda^2}$. Einsetzen ($\lambda = 1, 2$) liefert schliesslich $\mathrm{Corr}(X,U) = \frac{2}{\sqrt{5}}$.

(c) Zunächst gilt

$$E[U] = E[X] + E[Y] = 1 + \frac{1}{2} = 1.5$$

und wegen $\mathrm{Cov}(X,Y) = 0$ (Unabhängigkeit von X und Y)

$$\mathrm{Var}(U) = \mathrm{Var}(X) + \mathrm{Var}(Y) = 1 + \frac{1}{4} = 1.25.$$

Demzufolge erhalten wir

$$E[S_{20}] = 20E[U_1] = 20 \cdot 1.5 = 30 \tag{14}$$

und

$$\mathrm{Var}(S_{20}) = 20\mathrm{Var}(U_1) = 20 \cdot 1.25 = 25.$$

Mit der Chebyshev-Ungleichung erhalten wir

$$P\left[S_{20} \notin (20,40)\right] = P\left[|S_{20} - 30| \geqslant 10\right] \leqslant \frac{25}{10^2} = 0.25, \tag{15}$$

und somit $P\left[S_{20} \in (20,40)\right] \geqslant 0.75$.

(d) Nach dem zentralen Grenzwertsatz gilt

$$\lim_{n\to\infty} P\left[\frac{S_n - nE[U_1]}{\sqrt{n\mathrm{Var}(U_1)}} \leqslant x\right] = \Phi(x). \tag{16}$$

Da $E[U] = 1.5$ und $\mathrm{Var}(U) = 1.25$ und mittels Normalisierung ergibt sich

$$\begin{aligned} P\left[S_{20} \in (25,35)\right] &= P\left[-5 < S_{20} - 30 < 5\right] \\ &= P\left[-1 < \frac{S_{20} - 30}{5} < 1\right] \\ &\approx \Phi(1) - \Phi(-1) = 2\Phi(1) - 1 \\ &= 0.6827. \end{aligned} \tag{17}$$

Lösung 3. (a) Seien X, Y, Z unabhängig mit $X, Y \sim \mathcal{N}(\mu_A, \sigma_A^2)$ und $Z \sim \mathcal{N}(\mu_C, \sigma_C^2)$. Für die Gesamtarbeitszeit ξ gilt dann

$$\xi \overset{d}{=} 2X + 2Y + Z \sim \mathcal{N}(4\mu_A + \mu_C, 8\sigma_A^2 + \sigma_C^2) = \mathcal{N}(110, 27). \tag{18}$$

Daraus ergibt sich erstens $E[\xi] = 110$ Minuten und zweitens, mittels Normalisation,

$$P[\xi \leq 120] = P\left[\frac{\xi - 110}{\sqrt{27}} \leq \frac{120 - 110}{\sqrt{27}}\right] = \Phi\left(\frac{10}{3\sqrt{3}}\right) \approx 0.9726. \tag{19}$$

(b) Aufgrund der angegebenen Daten kann nur getestet werden, ob der spezialisierte Monteur ebenfalls mit dem Mittelwert $\mu = 38$ Minuten arbeitet. Also ist über den Mittelwert μ einer normalverteilten Zufallsgrösse die Nullhypothese

$$H_0 : \mu = 38 \qquad \text{gegenüber} \qquad H_A : \mu < 38 \tag{20}$$

bei bekannter Standardabweichung $\sigma = 3$ Minuten zu testen.

Bemerkung. Die einseitige Fragestellung erscheint hier desshalb sinnvoll, weil man wohl davon ausgehen kann, dass der spezialisierte Monteur nicht länger braucht als ein anderer. Sollte diese Annahme in einem konkreten Fall nicht gerechtfertigt sein, so müsste man natürlich $\mu = 38$ gegen $\mu \neq 38$ testen.

Dies ist ein z-Test mit Signifikanzniveau $\alpha = 5\%$. Die Teststatistik ist gleich

$$T = \sqrt{n}\frac{\overline{X} - \mu_0}{\sigma} \overset{H_0}{\sim} \mathcal{N}(0,1) \quad \text{mit} \quad K = (-\infty, -z_{0.95}). \tag{21}$$

Nach der Tabelle ist $-z_{0.95} = -1.64$. D.h. wir verwerfen H_0 wenn $t = T(\omega) < -1.64$. Einsetzen liefert nun

$$t = T(\omega) = \frac{36 - 38}{3}\sqrt{10} \approx -2.11 < -1.64. \tag{22}$$

Das bedeutet die Nullhypothese kann zugunsten der Alternative zum Niveau 5% abgelehnt werden.

Bemerkung. Im zweiseitigen Fall $H_0 : \mu = 38$ gegen $H_A : \mu \neq 38$ wäre

$$T = \sqrt{n}\frac{\overline{X} - \mu_0}{\sigma} \quad \text{mit} \quad K = (-\infty, -z_{0.975}) \cup (z_{0.975}, \infty),$$

mit $z_{0.975} = 1.96$. Nach (22) gilt $t < -1.96$, sodass H_0 zugunsten H_A zum Niveau 5% abgelehnt wird.

(c) $P_{H_0}[T < t] = P_{H_0}[T < -2.11] = 1 - \Phi(2.11) \approx 0.0175$.

Bemerkung. Im zweiseitigen Fall gilt $P_{H_0}[|T| > 2.1] = 2(1 - \Phi(2.11)) \approx 0.0350$.