Eingangsprüfung Stochastik, 06.05.2017

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

Aufgabe 1 (18 Punkte)
Sei $\mathcal{C} \subset \mathcal{A}$ eine Unter-σ-Algebra von $\mathcal{A}$. Beweisen Sie:

(a) $Y : \Omega \longrightarrow [0, \infty)$ sei $\mathcal{C}$-messbar. Dann ist Y $\mathcal{A}$-messbar.

(b) Gilt die Umkehrung von (a)? Begründen Sie Ihre Antwort.

(c) $Y : \Omega \longrightarrow [0, \infty)$ sei $\mathcal{C}$-messbar, $X : \Omega \longrightarrow [0, \infty)$ sei $\mathcal{A}$-messbar und es gelte für alle $C \in \mathcal{C}$

$$\int_C Y \, dP = \int_C X \, dP. \tag{1}$$

Dann folgt für alle $\mathcal{C}$-messbaren $Z : \Omega \longrightarrow [0, \infty)$:

$$\int_\Omega ZY \, dP = \int_\Omega ZX \, dP. \tag{2}$$

Aufgabe 2 (18 Punkte)
Eine Zufallsvariable $Z : \Omega \longrightarrow (0, \infty)$ heißt lognormalverteilt $\mathcal{L}N(\mu, \sigma^2)$, wenn $\ln Z \sim N(\mu, \sigma^2)$ gilt.

(a) Sei $Z \sim \mathcal{L}N(\mu, \sigma^2)$. Beweisen Sie, dass $\alpha Z^\beta \sim \mathcal{L}N(\beta\mu + \ln \alpha, \beta^2\sigma^2)$ für alle $\alpha > 0$, $\beta \neq 0$ gilt.

(b) Ein Versicherungsunternehmen mit Gesamtschaden X und Prämienvolumen v modelliert die Schadenquote $Z := \dfrac{X}{v}$ mit $Z \sim \mathcal{L}N\left(-\dfrac{36}{100}, \dfrac{5}{1000}\right)$.

  (i) Bestimmen Sie die Wahrscheinlichkeit, dass die Schadenquote zwischen 60% und 80 % liegt.

  (ii) Angenommen, das Prämienvolumen ist um 10 % höher als v bei unverändertem Gesamtschaden X. Bestimmen Sie nun die Wahrscheinlichkeit, dass die Schadenquote zwischen 60% und 80 % beträgt.

  (iii) Interpretieren Sie das Ergebnis in (ii) qualitativ.

  (iv) Seien Z_1, Z_2 die Schadenquoten zweier Jahre, Z_1 und Z_2 seien unabhängig und identisch $\mathcal{L}N\left(-\dfrac{36}{100}, \dfrac{5}{1000}\right)$ verteilt. Bestimmen Sie die Wahrscheinlichkeit, dass $Y := \max(Z_1, Z_2)$ größer als 80 % ist.

Aufgabe 3 (18 Punkte)
Gegeben seien Zufallsvariablen $X, V, Z : \Omega \longrightarrow [0, 1]$, $E(V), E(Z) < \infty$. Die Dichte von Z sei f_Z. Die bedingte Dichte von X gegeben $V = v$ sei

$$f(x|v) = \begin{cases} \dfrac{1}{v} f_Z\left(\dfrac{x}{v}\right) & \text{für } v \in (0,1),\ x \in (0, v) \\ 0 & \text{sonst.} \end{cases}$$

(a) Beweisen Sie für $x > 0$

$$P(X \leq x|V = v) = \begin{cases} P\left(Z \leq \frac{x}{v}\right) & \text{für } x < v, \\ 1 & \text{sonst.} \end{cases}$$

(b) Zeigen Sie $E(X|V) = VE(Z)$.

(c) Bestimmen Sie $E(X)$.

(d) Sei $V \sim U(0,1)$, $P(Z \leq z) = e^{1-\frac{1}{z}}$, $z \in (0,1)$. Bestimmen Sie $P(X \leq x)$ für $x \in (0,1)$.

Aufgabe 4 (18 Punkte)
Sei $X : \Omega \longrightarrow (0, \infty)$ eine Zufallsvariable

$$F(x) = \begin{cases} 1 - (1+x)^{-\alpha} & x > 0 \\ 0 & \text{sonst.} \end{cases}$$

die Verteilungsfunktion von X mit Parameter $\alpha > 0$.

(a) Zeigen Sie, dass $Y := \ln(1+X) \sim \text{Exp}(\alpha)$ gilt, und bestimmen Sie $E(Y)$.

(b) Gegeben sei eine Stichprobe $x_1, \ldots, x_n$ aus unabhängigen Realisierungen $X_1, \ldots, X_n \overset{\text{iid}}{\sim} X$.

   (i) Beweisen Sie, dass $\hat{\alpha} = n \left(\prod_{k=1}^{n} \ln(1+X_i) \right)^{-1}$ ML-Schätzer von α ist.

   (ii) Gegeben sei eine Stichprobe $x_1, \ldots, x_{20}$ mit $\prod_{k=1}^{20} (1+x_i) = e^{8,40}$. Bestimmen Sie den ML-Schätzwert.

Aufgabe 5 (18 Punkte)
Der Schadengrad z eines Schadens s bei einem Risiko mit Versicherungssumme v ist als $z = \frac{s}{v}$ definiert.
In der Feuerversicherung wurden die folgenden logarithmierte Schadengrade $y_i := \ln z_i$ und die logarithmierten Versicherungssummen der betroffenen Risiken $x_i := \ln v_i$ beobachtet:

$x_i = \ln v_i$	17,85	16,58	17,72	17,46	17,52	18,27	16,9	18,24	16,7	17,98
$y_i = \ln z_i$	-6,23	-5,95	-6,18	-6,02	-6,13	-6,47	-5,86	-6,42	-5,87	-6,45

Eine graphische Darstellung finden Sie in Abbildung 1.
Sie können für diese Aufgabe die folgenden Kennzahlen verwenden:

$$\overline{x} = 17,52, \ \overline{y} = -6,16, \ \sum_{i=1}^{10} (x_i - \overline{x})^2 = 3,39, \ \sum_{i=1}^{10} (y_i - \overline{y})(x_i - \overline{x}) = -1,21$$

$$\sum_{i=1}^{10} (y_i - \overline{y})^2 - \frac{\left(\sum_{i=1}^{10} (y_i - \overline{y})(x_i - \overline{x})\right)^2}{\sum_{i=1}^{10} (x_i - \overline{x})^2} = 0,06.$$

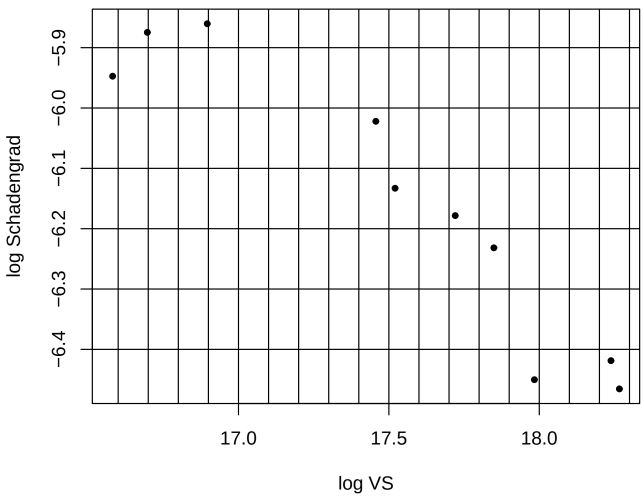


Abbildung 1:*Beobachtete logarithmierte Schadengrade und logarithmierte Versicherungssummen. Beobachtungen sind mit Punkten • eingetragen, senkrechte bzw. waagerechte Linien sind in Intervallen von* $0,1$ *eingezeichnet.*

Für die Zufallsvariable Y_i wird das folgende Modell untersucht:

$$Y_i = a + bx_i + \varepsilon_i, \text{ mit } \varepsilon_i \overset{\text{iid}}{\sim} N(0, \sigma^2),\ i = 1, \ldots, n.$$

(a) Bestimmen Sie Schätzwerte $\hat{a}$, $\hat{b}$ von a, b.

(b) Bestimmen Sie graphisch mit den Ergebnissen aus (a) und Abbildung 1 einen Schätzwert für den Erwartungswert des logarithmierten Schadengrads bei einer Versicherungssumme von 30 Millionen.

(c) (i) Untersuchen Sie die drei Hypothesen, $b = 0$, $b = -0,4$ und $a = 0$. Geben Sie eine Interpretation Ihrer Ergebnisse. Die Aussagen sollen zu einem Niveau von $\alpha = 5$ % erfolgen.

(ii) Welche Verteilung des Schadengrads Z für ein Risiko mit Versicherungssumme v ergibt sich für $a = 0$ und $b = -0,4$?

Lösungsvorschläge

Aufgabe 1 [(a) 3 (b) 3 (c) 12 Punkte]
Zu (a)
Sei $B \in \mathcal{B}^1$. Da Y $\mathcal{C}$-messbar ist, folgt $Y^{-1}(B) \in \mathcal{C}$, also wegen $\mathcal{C} \subset \mathcal{A}$ auch $Y^{-1}(B) \in \mathcal{A}$.

Zu (b)
Sei $\mathcal{C} = \{\emptyset, \Omega\} \neq \mathcal{A}$ und $A \in \mathcal{A} \setminus \mathcal{C}$. Dann ist die Indikatorfunktion 1_A offensichtlich $\mathcal{A}$-messbar, aber wegen $1_A^{-1}(\{1\}) = A \notin \mathcal{C}$ ist 1_A nicht $\mathcal{C}$-messbar.

Zu (c)
Sei Z eine Treppenfunktion, also $Z = \sum_{i=1}^n \alpha_i 1_{C_i}$, wobei $C_i \in \mathcal{C}$ und $\alpha_i > 0$. Dann folgt laut Voraussetzung

$$\begin{aligned}\int_\Omega ZY\,dP &= \int_\Omega \sum_{i=1}^n \alpha_i 1_{C_i} Y\,dP = \sum_{i=1}^n \alpha_i \int_{C_i} Y\,dP \overset{(1)}{=} \sum_{i=1}^n \alpha_i \int_{C_i} X\,dP \\ &= \int_\Omega \sum_{i=1}^n \alpha_i 1_{C_i} X\,dP = \int_\Omega ZX\,dP.\end{aligned}$$

Somit gilt (2) für $\mathcal{C}$-messbare Treppenfunktionen Z.
Sei Z $\mathcal{C}$-messbar. Dann gibt es eine Folge $(Z_n)_{n\in\mathbb{N}}$ von $\mathcal{C}$-messbaren Treppenfunktionen mit $Z_n \geq 0$ und $Z_n \uparrow Z$. Dann folgt $0 \leq Z_nY \uparrow ZY$ und $0 \leq Z_nX \uparrow ZX$. Ferner gilt wegen des oben Gezeigten auch

$$\forall n \in \mathbb{N}: \quad \int_\Omega Z_nY\,dP = \int_\Omega Z_nX\,dP.$$

Bildet man auf beiden Seiten den Grenzübergang $n \to \infty$, so folgt mit dem Satz von der monotonen Konvergenz die Behauptung.

Aufgabe 2 [(a) 4, (b) 4+4+2+4]
Zu (a)
Sei $\alpha > 0$ und $\beta \neq 0$. Dann gilt

$$\ln(\alpha Z^\beta) = \ln\alpha + \beta \ln Z.$$

Da $\ln Z \sim N(\mu, \sigma^2)$ folgt die Behauptung aus der Reproduktivität der Normalverteilung.
Zu (b)
(i) Sei Φ die Verteilungsfunktion der Standardnormalverteilung. Es gilt

$$\begin{aligned}P(0,6 < Z \leq 0,8) &= P(Z \leq 0,8) - P(Z \leq 0,6) \\ &= P(\ln Z \leq \ln 0,8) - P(\ln Z \leq \ln 0,6) \\ &= \Phi\left(\frac{\ln 0,8 + 0,36}{\sqrt{0,005}}\right) - \Phi\left(\frac{\ln 0,6 + 0,36}{\sqrt{0,005}}\right) \approx \Phi(1,94) - \Phi(-2,13) \\ &= 0,9738 - (1 - \Phi(2,13)) = 0,9738 - (1 - 0,9834) = 0,9572\end{aligned}$$

(ii) Wegen $Z' = \dfrac{X}{1,1v} = Z\dfrac{1}{1,1}$ folgt $\ln Z' = \ln Z - \ln 1,1$ und somit

$$\begin{aligned} P\left(0,6 < Z' \leq 0,8\right) &= P\left(\ln 0,6 < \ln Z' \leq \ln 0,8\right) \\ &= P\left(\ln 0,6 < \ln Z - \ln 1,1 \leq \ln 0,8\right) \\ &= P\left(\ln 0,66 < \ln Z \leq \ln 0,88\right) \\ &= \Phi\left(\frac{\ln 0,88 + 0,36}{\sqrt{0,005}}\right) - \Phi\left(\frac{\ln 0,66 + 0,36}{\sqrt{0,005}}\right) \\ &\approx \Phi\left(3,28\right) - \Phi\left(-0,79\right) = 0,9995 - 1 + 0,7852 = 0,7802. \end{aligned}$$

(iii) Erhöht man die Prämie um 10 %, dann werden die Schadenquoten kleiner, also steigt die Wahrscheinlichkeit kleinerer Schadenquoten.
(iv) Es gilt wegen der Unabhängigkeit von Z_1 und Z_2 für alle z

$$P(\max(Z_1, Z_2) \leq z) = P(Z_1 \leq z, Z_2 \leq z) = P(Z_1 \leq z)P(Z_2 \leq z)$$

und wegen der Verteilungsannahme

$$P(\max(Z_1, Z_2) \leq 0,8) = P(Z_1 \leq 0,8)^2 = 0,9738^2 = 0,9483.$$

Damit ergibt sich

$$P(\max(Z_1, Z_2) > 0,8) = 1 - P(Z_1 \leq 0,8)^2 = 0,0517.$$

Aufgabe 3 [(a) 5, (b) 5, (c) 3, (d) 5]
Zu (a) Sei $x, v \in (0,1)$.
1. Fall: $x > v$.

$$P(X \leq x|V = v) = \int_0^x f(t|v)\,dt = \int_0^v \frac{1}{v} f_Z\left(\frac{t}{v}\right)\,dt \overset{\tau = \frac{t}{v}}{=} \int_0^1 f_Z\left(\tau\right)\,d\tau = 1.$$

Hierbei wurde die Sutstitution $\tau = \dfrac{t}{v}, dt = v d\tau$ verwendet.
2. Fall: $x \leq v$.

$$\begin{aligned} P(X \leq x|V = v) &= \int_0^x f(t|v)\,dt = \int_0^x \frac{1}{v} f_Z\left(\frac{t}{v}\right)\,dt \overset{\tau = \frac{t}{v}}{=} \int_0^{x/v} f_Z\left(\tau\right)\,d\tau \\ &= P\left(Z \leq \frac{x}{v}\right) \end{aligned}$$

Zu (b)
Mit der Substitution $y = \dfrac{x}{v}$, $dx = v dy$ ergibt sich

$$\begin{aligned} E(X|V = v) &= \int_0^v x f(x|v)\,dx = \int_0^v \frac{x}{v} f_Z\left(\frac{x}{v}\right)\,dx \overset{y = \frac{x}{v}}{=} \\ &= \int_0^1 y f_Z(y) v\,dy = v \int_0^1 y f_Z(y)\,dy = vE(Z) \end{aligned}$$

Die Behauptung ergibt sich, in dem man v durch V ersetzt.
Zu (c)

Mit der iterierten Erwartung erhalten wir

$$E(X) = E(E(X|V)) = E(E(Z)V) = E(Z)E(V).$$

Zu (d)
Mit (a) ergibt sich für $x \in (0, 1)$

$$\begin{aligned} P(X \leq x) &= \int_0^1 P(X \leq v|V = v) f_V(v)\, dv = \int_0^1 P(X \leq v|V = v)\, dv \\ &= \int_0^x P(X \leq v|V = v)\, dv + \int_x^1 P(X \leq v|V = v)\, dv \\ &\overset{(a)}{=} x + \int_x^1 P\left(Z \leq \frac{x}{v}\right)\, dv = x + \int_x^1 \exp\left(1 - \frac{v}{x}\right)\, dv \\ &= x - x \exp\left(1 - \frac{v}{x}\right)\Bigg|_{v=x}^{v=1} = x - x\exp\left(1 - \frac{1}{x}\right) + x \\ &= 2x - \exp\left(1 - \frac{1}{x}\right). \end{aligned}$$

Aufgabe 4[(a) 3 (b) 13 + 2]
Zu (a)
Es gilt für $y > 0$

$$P(Y \leq y) = P(\ln(X + 1) \leq y) = P(X \leq -1 + e^y) = 1 - \left(\frac{1}{e^y}\right)^\alpha = 1 - e^{-\alpha y}$$

und somit folgt die Behauptung.
Zu (b)
(i)
Die Dichte f von X ergibt sich für durch differenzieren nach x. Für $x < 0$ ergibt sich $f(x) = 0$ und für $x > 0$

$$f(x) = \frac{d}{dx} F(x) = \alpha(1 + x)^{-(1+\alpha)}$$

Die X_i sind unbhängig, die gemeinsame Dichte ist das Produkt der Dichten der X_i, es ergibt sich also die Likelihood

$$L(\alpha; x_1, \ldots, x_n) = \prod_{i=1}^n \left(\alpha(1 + x_i)^{-1-\alpha}\right).$$

Logarithmiert man beide Seiten folgt

$$\begin{aligned} \ell(\alpha) &= n \ln \alpha - (\alpha + 1) \sum_{i=1}^n \ln(1 + x_i), \\ \frac{d}{d\alpha}\ell(\alpha) &= \frac{n}{\alpha} - \sum_{i=1}^n \ln(1 + x_i), \\ \frac{d^2}{d\alpha^2}\ell(\alpha) &= -\frac{n}{\alpha^2} < 0. \end{aligned}$$

Einzige Nullstelle der Ableitung von ℓ liegt bei

$$\hat{\alpha} = \frac{n}{\sum_{i=1}^{n} \ln(1 + x_i)},$$

es liegt ein lokales Maximum vor. Wegen des Vorzeichenwechsels von ℓ' in $\hat{\alpha}$ liegt ein globales Maximum vor. Der ML-Schätzer von α ist

$$\hat{\alpha} = \frac{n}{\sum_{i=1}^{n} \ln(1 + X_i)}.$$

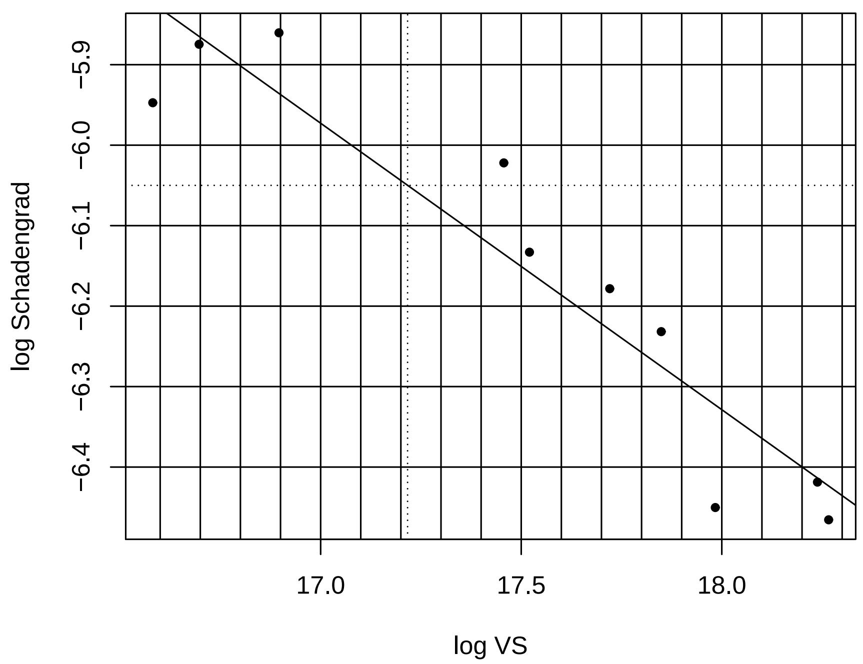
(ii)

$$\hat{\alpha} = \frac{n}{\ln\left(\prod_{i=1}^{20}(1 + x_i)\right)} = \frac{20}{8,4} = 2,38.$$

Aufgabe 5 [(a) 4, (b) 3, (c) 8 + 3]
Zu (a)

$$\hat{b} = \frac{-1,21}{3,39} = -0,36,\ \hat{a} = -6,16 - (-0,36) \cdot 17,52 = 0,15.$$

Zu (b)
Man liest auf der Ausgleichsgeraden bei $\ln(3 \cdot 10^7) \approx 17,2$ den Wert von ca. -6,05 ab.



Zu (c) (i) Es gilt

$$\mathrm{se}(\hat{a}) = \sqrt{\frac{0,06}{8}\left(\frac{1}{10} + \frac{17,52^2}{3,39}\right)} = 0,82 \qquad \mathrm{se}(\hat{b}) = \sqrt{\frac{0,06}{8}\frac{1}{3,39}} = 0,05$$

Damit ergeben sich folgende Testgrößen bei den durchzuführenden Hypothesentests:

$$\begin{array}{ll} H_0 : b = 0 & T = \dfrac{-0,36}{0,05} = -7,2 \\ H_0 : b = -0,4 & T = \dfrac{-0,36 + 0,4}{0,05} = 0,8 \\ H_0 : a = 0 & T = \dfrac{0,15}{0,82} = 0,18 \end{array}$$

H_0 wird abgelehnt, wenn $|T| > t_{8;0,975} = 2,306$ gilt. Somit wird die Hypothese $b = 0$ verworfen, die beiden Hypothesen $a = 0$ und $b = -0,4$ können nicht verworfen werden. Somit wird verworfen, dass der logarithmierte Schadengrad nicht von der logarithmierten Versicherungssumme abhängt.
(ii) Nach Modellannahme gilt für den Schadengrad mit Versicherungssumme v,

$$Z = \exp(Y) = \exp\left(-0,4 \ln v + \varepsilon\right) = v^{0,4} e^{\varepsilon},\ \varepsilon \sim N(0, \sigma^2)$$

also $Z \sim \mathcal{L}N\left(-\frac{2}{5}\ln v, \sigma^2\right)$.