Musterlösung zur Musterklausur WS 05/06
Elementare Wahrscheinlichkeitsrechnung (L)

1) X_1 – Anzahl der Treffer der ersten Schützin

X_2 – Anzahl der Treffer der zweiten Schützin

X_3 – Anzahl der Treffer der dritten Schützin

X_4 – Anzahl der Treffer der vierten Schützin Annahme: $X_i \sim B(5, p_i)$

1a) Gesucht: $P(X_1 = 5, X_2 = 5, X_3 = 3, X_4 = 2)$

$$P(X_1 = 5, X_2 = 5, X_3 = 3, X_4 = 2)$$

$$= P(X_1 = 5) \cdot P(X_2 = 5) \cdot P(X_3 = 3) \cdot P(X_4 = 2)$$

$$= p_1^5 \cdot p_2^5 \cdot \binom{5}{3} \cdot p_3^3 \cdot (1 - p_3)^2 \cdot \binom{5}{2} \cdot p_4^2 \cdot (1 - p_4)^3$$

$$= p_1^5 \cdot p_2^5 \cdot 10 \cdot p_3^3 \cdot (1 - p_3)^2 \cdot 10 \cdot p_4^2 \cdot (1 - p_4)^3 .$$

1b) A – Eine Starterin trifft immer und die anderen machen mind. einen Fehler

$$P(A) = P(X_1 = 5, X_i < 5;\ i = 2..4) + P(X_2 = 5, X_i < 5;\ i = 1,3,4) +$$

$$P(X_3 = 5, X_i < 5;\ i = 1,2,4) + P(X_4 = 5, X_i < 5;\ i = 1..3)$$

$$= p_1^5 (1 - p_2^5)(1 - p_3^5)(1 - p_4^5) + (1 - p_1^5)\, p_2^5 (1 - p_3^5)(1 - p_4^5) +$$

$$(1 - p_1^5)(1 - p_2^5)\, p_3^5 (1 - p_4^5) + (1 - p_1^5)(1 - p_2^5)(1 - p_3^5)\, p_4^5 .$$

1c) Z – Anzahl der Treffer aller Schützinnen

$Z = X_1 + X_2 + X_3 + X_4$

$$EZ = E(X_1 + X_2 + X_3 + X_4) = EX_1 + EX_2 + EX_3 + EX_4$$

$$= 5p_1 + 5p_2 + 5p_3 + 5p_4 = 5(p_1 + p_2 + p_3 + p_4) .$$

$$\text{Var } Z = \text{Var } (X1 + X2 + X3 + X4) = \text{Var } X1 + \text{Var } X2 + \text{Var } X3 + \text{Var } X4$$

$$= 5p_1(1 - p_1) + 5p_2(1 - p_2) + 5p_3(1 - p_3) + 5p_4(1 - p_4)$$

$$= 5(p_1 - p_1^2 + p_2 - p_2^2 + p_3 - p_3^3 + p_4 - p_4^4) .$$

1d) Es seien N – Nettogewinn, A – alle ohne Fehler, $\overline{A}$ – ein Fehlschuss, E – Einsatz

ω	A	$\overline{A}$
$N(\omega)$	100 – E	– E
$P(\omega)$	$p_1^5 p_2^5 p_3^5 p_4^5$	$1 - p_1^5 p_2^5 p_3^5 p_4^5$

Fairnessbedingung: EN = 0.

D.h.: $(100 - E) \cdot p_1^5 p_2^5 p_3^5 p_4^5 + (-E) \cdot (1 - p_1^5 p_2^5 p_3^5 p_4^5) = 0$

$\Leftrightarrow$ $E = 100 \cdot p_1^5 p_2^5 p_3^5 p_4^5 = 100 \cdot 0{,}8^{5 \cdot 4} \approx 1{,}15$.

D.h. der faire Einsatz liegt bei 1,15 €.

2) Sei $S = 40e^R$, $R \sim N(0{,}08;\ 0{,}04)$.

2a) Gesucht: $P(S > 50)$

$$P(S > 50) = P(40e^R > 50) = P(e^R > \frac{50}{40}) = P(R > \ln\frac{5}{4})$$

$$= 1 - P(R \le 0{,}223) = 1 - \Phi(\frac{0{,}223 - 0{,}08}{0{,}2}) = 1 - \Phi(0{,}715)$$

$$\approx 1 - 0{,}7580 = 0{,}242.$$

Die Wahrscheinlichkeit liegt bei rund 24%.

2b)

Bei einer normalverteilten Zufallsgröße liegt der Wert mit einer Wahrscheinlichkeit von 95% in einem 2σ-Intervall, d.h. in $I_R = [\mu - 2\sigma;\ \mu + 2\sigma] = [-0{,}32;\ 0{,}48]$

$$\Rightarrow I_S = [40e^{-0{,}32};\ 40e^{0{,}48}] = [29{,}05;\ 64{,}64].$$

3) Es seien

H – Der Säugling hat die Krankheit $\qquad \overline{H}$ – Der Säugling hat sie nicht

E – Die Krankheit wird erkannt $\qquad \overline{E}$ – Die Krankheit wird nicht erkannt

Gegeben: $P(H) = \dfrac{100}{1{,}1 \cdot 10^6}$, $P(\overline{E} \mid H) = 0{,}0001$, $P(E \mid \overline{H}) = 0{,}001$

Gesucht: Die Wkeit, dass ein Säugling die Krankheit hat, wenn er sie erkannt wurde,

d.h. $P(H \mid E)$.

$P(H \mid E)$

$$= \frac{P(H \cap E)}{P(E)} = \frac{P(H) \cdot P(E \mid H)}{P(E \cap H) + P(E \cap \overline{H})} = \frac{P(H) \cdot (1 - P(\overline{E} \mid H))}{P(H) \cdot P(E \mid H) + P(\overline{H}) \cdot P(E \mid \overline{H})}$$

$$= \frac{\frac{100}{1{,}1 \cdot 10^6} \cdot 0{,}9999}{P(H) \cdot (1 - P(\overline{E} \mid H)) + (1 - P(H)) \cdot 0{,}001} = \frac{\frac{100}{1{,}1 \cdot 10^6} \cdot 0{,}9999}{\frac{100}{1{,}1 \cdot 10^6} \cdot 0{,}9999 + (1 - \frac{100}{1{,}1 \cdot 10^6}) \cdot 0{,}001}$$

$= 0{,}083.$

D.h. die Wkeit liegt bei 8,3%.

4) Es seien

H : $p = 0{,}2$ A : $p > 0{,}2$ $n = 2000$ X – Anteil der ABC-Wähler, Annahme: $X \sim B(n, p)$

4a) $E_H X = np = 2000 \cdot 0{,}2 = 400$.

4b)

Kritischer Bereich $K = \{X > c\}$, Signifikanzniveau $\alpha = 0{,}05$. Gesucht ist c.

Es gilt: $P_H(K) \le 0{,}05 \Leftrightarrow P_H(\{X > c\}) \le 0{,}05 \Leftrightarrow 1 - P_H(\{X \le c\}) \le 0{,}05$

Anwendung des Grenzwertsatzes von deMoivre-Laplace:

Prüfen der Faustregel: $\sigma = \sqrt{np(1-p)} = \sqrt{400 \cdot 0{,}8} = \sqrt{320} \approx 18 > 9$

Also: $1 - P_H(\{X \le c\}) \le 0{,}05$

$$\Leftrightarrow \quad 1 - \lim_{n\to\infty} P_H(-\infty < X \le \frac{c-400}{\sqrt{320}}) \le 0{,}05$$

$$\Leftrightarrow \quad 1 - \Phi(\frac{c-400}{\sqrt{320}}) \le 0{,}05$$

$$\Leftrightarrow \quad \Phi(\frac{c-400}{\sqrt{320}}) \ge 0{,}95$$

$$\Rightarrow \quad \frac{c-400}{\sqrt{320}} = 1{,}7 \qquad (\text{da } \Phi(1{,}7) = 0{,}9554)$$

$$\Leftrightarrow \quad c = 1{,}7 \cdot \sqrt{320} + 400 = 430{,}41.$$

D.h. $K = \{X > 431\}$.

4c)

Gesucht ist die Wkeit, H anzunehmen ($X \notin K$), wenn H falsch ist, und

es gelte eine Alternative $p = 0{,}22$. Gesucht ist also $P_{0,22}(\{X \le 430\})$.

Für $p = 0{,}22$ ist $E_{0,22} X = 440$ und $\sigma_{0,22} = \sqrt{343{,}2}$ ($\approx 18{,}5 > 9$, d.h. Grenzwertsatz gilt).

$$P_{0,22}(\{X \le 430\}) = \lim_{n\to\infty} P_{0,22}(-\infty < X \le \frac{430-440}{\sqrt{343{,}2}})$$

$$= \Phi(-0{,}54) = 1 - \Phi(0{,}54) = 1 - 0{,}6915 = 0{,}3085.$$

D.h. die Wkeit des Fehler 2. Art liegt bei 30%.

4d)

Da 440 Wähler beobachtet wurden, gilt für den kritischen Bereich K_1, für den dieses Ergebnis signifikant sein soll: $K_1 = \{X > 440\}$. Wir wissen, dass das gesuchte Signifikanzniveau α_1 kleiner als 0,05 sein muss, da 440 > 431 ($K_1 < K$). Gesucht ist also das größte α_1, das unter 0,05 liegt und für das $P_H(\{X > 440\}) \geq \alpha_1$ gilt:

$$
\begin{aligned}
& P_H(\{X > 440\}) \geq \alpha_1 \\
\Leftrightarrow \quad & 1 - P_H(\{X \leq 440\}) \geq \alpha_1 \\
\Leftrightarrow \quad & 1 - \lim_{n\to\infty} P_H(-\infty < X \leq \frac{440-400}{\sqrt{320}}) \geq \alpha_1 \\
\Leftrightarrow \quad & 1 - \Phi(\frac{40}{\sqrt{320}}) = 1 - \Phi(2{,}24) \geq \alpha_1 \\
\Rightarrow \quad & 0{,}0139 \geq \alpha_1 \qquad\qquad (\text{da } \Phi(2{,}2) = 0{,}9861)
\end{aligned}
$$

D.h. $\alpha_1 \leq 0{,}0139$ ist das gesuchte Signifikanzniveau.

5a)

Y \ X	**1**	**2**	**3**	
1	**0,35**	0,15	**0,2**	**0,7**
2	0,14	**0,06**	0	**0,2**
3	0,01	0,09	0	0,1
	0,5	**0,3**	0,2	**1**

5b) X und Y sind nicht unabhängig, weil die Produktformel nicht gilt; so ist z.B. $0{,}2 \cdot 0{,}7 \neq 0{,}2$

5c) Für den Grad des linearen Zusammenhangs ist der Korrelationskoeffizient von X und Y entscheidend: $\rho(X,Y) = \frac{\operatorname{cov}(X,Y)}{\sqrt{VarX}\sqrt{VarY}}$

$\operatorname{cov}(X,Y) = EXY - EYEX =$

$$1\cdot1\cdot0{,}35 + 2\cdot0{,}15 + 3\cdot0{,}2 + 2\cdot0{,}14 + 4\cdot0{,}06 + 3\cdot0 + 3\cdot0{,}01 + 6\cdot0{,}09 + 9\cdot0$$
$$-(1\cdot0{,}7 + 2\cdot0{,}2 + 3\cdot0{,}1)(1\cdot0{,}5 + 2\cdot0{,}3 + 3\cdot0{,}2) = -0{,}04$$

$VarX = EX^2 - (EX)^2 = (1\cdot0{,}5 + 4\cdot0{,}3 + 9\cdot0{,}2) - (1{,}7)^2 = 0{,}61$

$VarY = EY^2 - (EY)^2 = (1\cdot0{,}7 + 4\cdot0{,}2 + 9\cdot0{,}1) - (1{,}4)^2 = 0{,}44$

$$\rho(X,Y) = \frac{\operatorname{cov}(X,Y)}{\sqrt{VarX}\sqrt{VarY}} = \frac{-0{,}04}{\sqrt{0{,}61}\sqrt{0{,}44}} = -0{,}0772.$$