Lösung der Nachklausur

1. Wir verwenden die Bayes-Formel. Der Nenner der Bayesformel wird

$$0.96 \cdot 0.05 + 0.022 \cdot 0.3 + 0.015 \cdot 0.2 + 0.003 \cdot 0.6 = 0.0594\ .$$

Somit erhalten wir aus der Bayes-Formel die Wahrscheinlichkeiten

Krankheit	0	A	B	A B
Wahrscheinlichkeit	0.8081	0.1111	0.0505	0.0303

2. **a)** Dies folgt aus Proposition 2.7, oder direkt aus dem Borel–Cantelli-Lemma, da $A_k = \{X_{2k-1} = 6, X_{2k} = 6\}$ unabhängig sind und Wahrscheinlichkeit 1/36 haben.

b) Um nach 5 Würfen zu enden, müssen Wurf 4 und 5 eine sechs zeigen. Der dritte Wurf darf keine Sechs zeigen, sonst wäre nach 4 Würfen fertig. Die ersten beiden Würfe dürfen nicht beide eine Sechs sein. Also ist die gesuchte Wahrscheinlichkeit

$$[1 - (\tfrac{1}{6})^2]\tfrac{5}{6}(\tfrac{1}{6})^2 = \tfrac{175}{7776} \approx 0.0225051\ .$$

c) Sei p die Wahrscheinlichkeit, dass das Experiment mit einer geraden Anzahl Würfen endet. Dann ist $1 - p$ die Wahrscheinlichkeit, dass das Experiment mit einer ungeraden Anzahl Würfen endet. Ist der erste Wurf eine 6, so muss entweder der zweite Wurf eine 6 sein, oder der zweite Wurf keine 6 sein, und danach das Experiment mit einer geraden Anzahl weiterer Würfe enden. Ist der erste Wurf keine 6, so muss das Experiment mit einer ungeraden Anzahl weiterer Würfe enden. Also

$$p = (\tfrac{1}{6})^2 + \tfrac{1}{6}\tfrac{5}{6}p + \tfrac{5}{6}(1-p) = \tfrac{31}{36} - \tfrac{25}{36}p\ .$$

Daraus folgt

$$p = \frac{\frac{31}{36}}{\frac{61}{36}} = \tfrac{31}{61} \approx 0.508197\ .$$

3. a) Wir wollen $\mathbb{P}[Z \le z]$ berechnen. Wegen der Symmetrie der Variablen ist die Verteilung von Z symmetrisch. Weiter ist für $z > 0$, $\mathbb{P}[0 < Z \le z] = 2\mathbb{P}[Z \le z, X > 0, Y > 0]$. Sei nun $z > 0$. Ist $Y = y > 0$ bekannt, so muss $X \le zy$ sein. Also

$$\mathbb{P}[0 < Z \le z] = 2\int_0^\infty \int_0^{yz} \frac{1}{2\pi} \mathrm{e}^{-(x^2+y^2)/2} \, \mathrm{d}x \, \mathrm{d}y \, .$$

Leiten wir nach z ab, erhalten wir die Dichte

$$\frac{1}{\pi}\int_0^\infty y\mathrm{e}^{-(1+z^2)y^2/2} \, \mathrm{d}y = \frac{1}{\pi}\frac{1}{1+z^2}\int_0^\infty (1+z^2)y\mathrm{e}^{-(1+z^2)y^2/2} \, \mathrm{d}y = \frac{1}{\pi}\frac{1}{1+z^2} \, .$$

Dies ist die gesuchte Dichte.

b) Der Erwartungswert wäre wegen der Symmetrie 0, sofern das Integral für den Positivteil endlich ist. Wir erhalten

$$\int_0^\infty \frac{y}{(1+y^2)\pi} \, \mathrm{d}y = \frac{1}{2\pi}\int_0^\infty \frac{2y}{1+y^2} \, \mathrm{d}y = \infty \, ,$$

da $\log(1+y^2)$ für $y \to \infty$ nach Unendlich konvergiert. Somit kann der Erwartungswert nicht definiert werden.

4. a) Leiten wir ab, erhalten wir

$$\begin{aligned}\frac{\partial}{\partial x}\frac{\partial}{\partial y}F(x,y) &= \frac{\partial}{\partial x}\frac{x(x+y-xy) - xy(1-x)}{(x+y-xy)^2} = \frac{\partial}{\partial x}\frac{x^2}{(x+y-xy)^2} \\ &= 2\frac{x(x+y-xy) - x^2(1-y)}{(x+y-xy)^3} = \frac{2xy}{(x+y-xy)^3}\end{aligned}$$

Es handelt sich um ein absolutstetige Verteilung, da $x+y-xy = x(1-y)+y \ge 0$. Somit ist die Dichte positiv.

b) Setzen wir $y = 1$, so erhalten wir $F(x,1) = x$. Somit ist die Randverteilung die Gleichverteilung auf $(0,1)$. Wegen der Symmetrie ist auch Y auf $(0,1)$ gleichverteilt.

c) Die beste Prognose ist die bedingte Erwartung von X gegeben $\{Y = y\}$. Die bedingte Verteilung hat die Dichte $f_{X|y}(x) = \frac{2xy}{(x+y-xy)^3}$, da die Randdichte von Y 1 ist. Der begingte Erwartungswert wird

$$\int_0^1 \frac{2x^2y}{(x+y-xy)^3} \, \mathrm{d}x \quad \left(= -\frac{y(y^2-4y+3+2\log y)}{(1-y)^3} \right) .$$

5. **a)** Wir haben $\mathbb{E}[X_k] = \gamma$. Somit ist der Schätzer

$$\hat{\gamma} = \frac{1}{n} \sum_{k=1}^{n} X_k \; .$$

b) Der Log-Likelihood ist

$$(\gamma - 1) \sum_{k=1}^{n} \log X_k - n \log(\Gamma(\gamma)) - \sum_{k=1}^{n} X_k \; .$$

Ableiten nach γ und Nullsetzen ergibt

$$\sum_{k=1}^{n} \log X_k - n \frac{\Gamma'(\gamma)}{\Gamma(\gamma)} = 0 \; .$$

Dies ist die zu lösende Gleichung. Dass es sich um ein Maximum handelt, folgt aus der Tatsache, dass $\log \Gamma(x)$ eine konvexe Funktion ist.

c) Die gemeinsame Dichte der Daten und des Parameters γ ist

$$\frac{1}{g} \mathbb{I}_{g \in (1,\mathrm{e})} \frac{1}{(\Gamma(g))^n} \exp\Big\{ - \sum_{k=1}^{n} x_k \Big\} \prod_{k=1}^{n} x_k^{g-1} \; .$$

Somit wird der bedingte Erwartungswert gegeben X_k (Bayes-Schätzer)

$$\frac{\int_1^{\mathrm{e}} g \frac{1}{g} \frac{1}{(\Gamma(g))^n} \exp\{-\sum_{k=1}^n X_k\} \prod_{k=1}^n X_k^{g-1} \,\mathrm{d}g}{\int_1^{\mathrm{e}} \frac{1}{g} \frac{1}{(\Gamma(g))^n} \exp\{-\sum_{k=1}^n X_k\} \prod_{k=1}^n X_k^{g-1} \,\mathrm{d}g} = \frac{\int_1^{\mathrm{e}} \frac{1}{(\Gamma(g))^n} \prod_{k=1}^n X_k^{g-1} \,\mathrm{d}g}{\int_1^{\mathrm{e}} \frac{1}{g} \frac{1}{(\Gamma(g))^n} \prod_{k=1}^n X_k^{g-1} \,\mathrm{d}g} \; .$$

6. Wir verwenden den χ^2-Anpassungstest. Die Wahrscheinlichkeit der Werte ist $p_k = \binom{4}{k}(1/3)^k(2/3)^{4-k}$. Wir fassen 3 und 4 zusammen, da für die Normalapproximation 5 Werte pro Zelle auftreten sollten.

Wert	0	1	2	3 & 4
p_k	0.1975	0.3950	0.2963	0.1111
$\bar{n}_k$	19.75	39.50	29.63	11.11

Die Teststatistik wird

$$\begin{aligned} T &= \frac{(26-19.75)^2}{19.75} + \frac{(49-39.5)^2}{39.5} + \frac{(17-29.63)^2}{29.63} + \frac{(8-11.11)^2}{11.11} \\ &= 10.5169 \; . \end{aligned}$$

T ist näherungsweise ξ^2 verteilt mit 3 Freiheitsgraden. Der Wert sollte unter der Hypothese im Intervall $(0.2158, 9.3484)$ liegen. Somit wird die Hypothese verworfen.

Falls die Zellen 3 und 4 nicht getrennt werden,

Wert	0	1	2	3	4
p_k	0.1975	0.3950	0.2963	0.09877	0.01234
$\bar{n}_k$	19.75	39.50	29.63	9.877	1.234

Die Teststatistik wird

$$T = \frac{(26-19.75)^2}{19.75} + \frac{(49-39.5)^2}{39.5} + \frac{(17-29.63)^2}{29.63} + \frac{(7-9.877)^2}{9.877} + \frac{(1-1.234)^2}{1.234}$$
$$= 10.5287\ .$$

Aus der Tabelle der χ^2-Verteilung mit 4 Freiheitsgraden sollte der Wert im Intervall $(0.4844, 11.143)$ liegen, was der Fall ist, Somit würde die Hypothese beibehalten.