

# Stochastik - Übungsklausur

Für die Klausur sind keinerlei Hilfsmittel wie Skript, Bücher oder Taschenrechner zulässig. Die Aufgaben sind an Übungen und Vorlesung orientiert. Für Richtigkeit der Lösungen besteht keine Garantie.

## Aufgabe 1: Laplace-Verteilung

Sie haben zwei Tetraeder zur Verfügung, deren Flächen jeweils mit den Augenzahlen 1, 2, 3 und 4 beschriftet sind. Beide Tetraeder werden geworfen und aus den beiden geworfenen Augenzahlen wird der Absolutbetrag ihrer Differenz, diesen nennen wir  $D$ , ermittelt.  $D$  kann also die Werte 0, 1, 2 und 3 annehmen.

- a) Geben Sie einen Wahrscheinlichkeitsraum für ein geeignetes Laplace-Experiment an und definieren Sie auf diesem Wahrscheinlichkeitsraum  $D$  durch Angabe der Abbildungsvorschrift als Zufallsvariable. Vergessen Sie nicht zu begründen, warum es sich bei Ihrem gewählten Experiment um ein Laplace-Experiment handelt. (5 Punkte)

Es handelt sich um ein Laplace Experiment, da jedes Elementarereignis die gleiche Wahrscheinlichkeit  $P(E) = \frac{|E|}{|D|}$  hat, mit Wahrscheinlichkeitsraum  $(\Omega, \Sigma, P)$ ,  $\Sigma = P(\Omega)$ ,  $\Omega = \{1, 2, 3, 4\}^2$  und  $D(\{\omega_1, \omega_2\}) = |\omega_1 - \omega_2|$

- b) Berechnen Sie die Verteilung von  $D$ . Ist  $D$  Laplaceverteilt? Begründen Sie Ihre Antwort. (3 Punkte)

$$D(\omega) = \begin{cases} 0 & \text{für } \{1, 1\}, \{2, 2\}, \{3, 3\}, \{4, 4\} & P_D(0) = \frac{4}{16} \\ 1 & \text{für } \{1, 2\}, \{2, 3\}, \{3, 4\}, \{4, 3\}, \{3, 2\}, \{2, 1\} & P_D(1) = \frac{6}{16} \\ 2 & \text{für } \{1, 3\}, \{2, 4\}, \{4, 2\}, \{3, 1\} & P_D(2) = \frac{4}{16} \\ 3 & \text{für } \{1, 4\}, \{4, 1\} & P_D(3) = \frac{2}{16} \end{cases}$$

Durch die unterschiedliche Wahrscheinlichkeitsverteilung der Ergebnisse ist  $D$  nicht Laplaceverteilt.

- c) Berechnen Sie den Erwartungswert von  $D$ . (2 Punkte)

$$\mu_D = E(X) = \sum_i D_i * P(X = D_i) = 0 * \frac{4}{16} + 1 * \frac{6}{16} + 2 * \frac{4}{16} + 3 * \frac{2}{16} = 1,25$$

## Aufgabe 2: Binomial-Verteilung

- a) Geben Sie mit Hilfe Bernoullivertilter Zufallsvariablen  $Z_1, Z_2, \dots$  eine Zufallsvariable  $X$  an, welche Binom  $(n, p)$ -verteilt ist. Welche Voraussetzungen müssen  $Z_1, Z_2, \dots$  erfüllen? Was modelliert Binom  $(n, p)$  anschaulich? (3 Punkte)

Eine Binominalverteilung mit Parametern  $n, p$  gilt bei  $P(n, p, k) = \binom{n}{k} * p^k * (1-p)^{n-k}$  mit  $k = Z_1, Z_2, \dots$ . Die Zufallsvariablen  $Z_1, Z_2, \dots$  müssen dafür Ganzzahlig und Positiv sein.

Ein anschauliches Binom ist das  $(n-)$  mehrmalige Werfen einer Münze, mit Ergebnis Erfolg ( $p = 0,5$ ) (Wappen) oder Misserfolg (Zahl).

- b) Bestimmen Sie basierend auf  $X$  den Maximum-Likelihood-Schätzer  $\hat{p}$  für  $p$ . Existiert dieser stets eindeutig? (5 Punkte)

$L : \Theta \rightarrow [0; 1], v \rightarrow p(x|v)$  wobei  $\Theta$  der Parameterraum ist.

$$\hat{p} = L(X) = \frac{1}{2}^k * (-\frac{1}{2})^{n-k}$$

Da der Schätzer eine Wahrscheinlichkeits-parabel abzeichnet, ist immer ein Wert als Maximum möglich.

c) Berechnen Sie den MSE (mean squared error) von  $\hat{p}$ . Ist  $\hat{p}$  unverzerrt? (2 Punkte)

$$MSE(T, v) = E_v((T - g(v))^2) = Var_v(T) + (B_T(v))^2$$

Da die Zufallsvariablen bernoulliverteilt sind, ist die Verzerrung  $B_T = 0$  und die Varianz 1:  $MSE(\hat{p}) = \frac{1}{n}$

### Aufgabe 3: Geometrische Verteilung

a) Geben Sie die Wahrscheinlichkeitsfunktion zur geometrischen Verteilung mit Parameter  $p$  an. Welche Werte darf  $p$  annehmen? (2 Punkte)

$$m_X(s) = \frac{pe^s}{1 - (1-p)e^s}$$

mit  $p = [0, \infty]$

b) Was modelliert die geometrische Verteilung mit Parameter  $p$ ? (1 Punkt)

die Wahrscheinlichkeitsverteilung der Anzahl  $X$  der Bernoulli-Versuche, die notwendig sind, um einen Erfolg zu haben. Diese Verteilung ist auf der Menge  $\mathbb{N}$  definiert.

c) Sie beobachten  $X_1, X_2, \dots, X_n$  unabhängig und identisch verteilte Zufallsgrößen, die jeweils eine  $\text{Geom}(p)$ -Verteilung besitzen. Bestimmen Sie den Momentenschätzer  $\hat{p}$  für  $p$ . (3 Punkte)

geometrische Verteilung:  $f(x) = (1-p)^{x-1} * p$

$$m_1 = \hat{m}_1 : \hat{\mu} = \frac{1}{n} \sum_{i=1}^n x_i \text{ und } m_2 = \hat{m}_2 : \hat{\sigma}^2 + \hat{\mu}^2 = \frac{1}{n} \sum_{i=1}^n x_i^2$$

$$\hat{p} = \hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$$

d) Ist  $\hat{p}$  unverzerrt? (2 Punkte)

Eine Verteilung wird als verzerrt bezeichnet, wenn sich die Datenpunkte mehr zu einer Seite der Skala als zur anderen gruppieren und eine Kurve erzeugen, die nicht symmetrisch ist. Mit anderen Worten, die rechte und die linke Seite der Verteilung sind unterschiedlich geformt. Die vorliegende Verteilung ist unverzerrt.

e) Ist  $\hat{p}$  konsistent? (2 Punkte)

In der Statistik ist ein konsistenter Schätzer eine Regel zum Berechnen von Schätzungen eines Parameters  $p$  mit der Eigenschaft, dass die resultierende Folge von Schätzungen mit zunehmender Wahrscheinlichkeit der Anzahl der verwendeten Datenpunkte in der Wahrscheinlichkeit gegen  $p$  konvergiert.

Dies ist nicht der Fall.

### Aufgabe 4: Exponentialverteilung

Gegeben sei eine Zufallsvariable  $X \sim \text{Exp}(2)$ .

a) Bestimmen Sie den Median und geben Sie den Erwartungswert von  $X$  an. Vergleichen Sie die beiden Werte und erklären Sie einen eventuell vorhandenen Unterschied zwischen den beiden Ergebnissen. (3 Punkte)

$$\text{Median } x_{med} = \frac{\ln 2}{\lambda} = \frac{\ln 2}{2} = 0,346574$$

$$\text{Erwartungswert } E(x) = \frac{1}{\lambda} = \frac{1}{2} = 0,5$$

Für die weitere Rechnung dürfen Sie ohne Nachweis benutzen, dass eine Zufallsvariable  $V \sim \text{Exp}(p)$  die Varianz  $Var(V) = \frac{1}{p^2}$  besitzt. Seien nun  $X_{i,i} \in N$  unabhängig und identisch  $\text{Exp}(2)$ -verteilte Zufallsgrößen und  $\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_{i,n} \in N$  die zugehörige Folge der Mittelwerte.

b) Bestimmen Sie  $E(\bar{X}_n)$  und  $Var(\bar{X}_n)$ . (2 Punkte)

$$E(\bar{X}_n) = \frac{1}{\bar{X}}; \quad Var(\bar{X}_n) = \frac{1}{\bar{X}^2}$$

c) Wie ist  $Z_n = \cos(\pi * \bar{X}_n)$  für großes  $n$  approximativ verteilt? (5 Punkte)

der  $\cos$  verteilt alle Ereignisse auf  $[-1, 1]$  und damit approximativ auf 0 für große  $n$

## Aufgabe 5: Uniforme Verteilung

Gegeben sei eine Zufallsvariable  $X \sim Unif(-0.5, 0.5)$ .

a) Geben Sie die Wahrscheinlichkeitsdichte und die Verteilungsfunktion von  $X$  an. (2 Punkte)

Wahrscheinlichkeitsdichte	$f(x) = \frac{1}{b-a} = \frac{1}{-0,5 - 0,5} = -1$
Verteilungsfunktion	$F(x) = \frac{x-a}{b-a} = -x + 0,5$

b) Geben Sie den Median, den Erwartungswert und die Varianz von  $X$  an. (3 Punkte)

Median	$Median(X) = F^{-1}(X) = \frac{a+b}{2} = 0/2 = 0$
Erwartungswert	$E(X) = \int_{-\infty}^{\infty} xf(x)dx = \frac{a+b}{2} = 0/2 = 0$
Varianz	$Var(X) = E(X^2) - (E(X))^2 = \frac{1}{12}(b-a)^2 = \frac{1}{12}$

Berechnungen im Rahmen von Bankgeschäften ergeben oft Ergebnisse mit gebrochenen Centanteilen, welche für Buchungsvorgänge gerundet werden müssen. Zum Beispiel würde man einen Betrag von 7,35... Cent auf 7 Cent abrunden und einen Betrag von 15,87... Cent auf 16 Cent aufrunden. Den Rundungsfehler kann man als uniformverteilt auf  $(-0.5, 0.5)$  (eigentlich  $(-0.5, 0.5]$ ) modellieren. Nehmen Sie nun an, dass in einer Bank 106 unabhängig und identisch  $Unif(-0.5, 0.5)$ -verteilte Rundungsvorgänge stattfinden.

c) Geben Sie mit Hilfe der Ungleichung von Tschebyscheff eine sinnvolle obere Abschätzung für die Wahrscheinlichkeit an, dass der Absolutbetrag der Summe der Rundungsfehler mindestens 10 Euro (1000 Cent) beträgt. (2 Punkte)

$1000Cent/106 \approx 9,4Cent \rightarrow \pm 4,7Cent$ $\mu = 0$ $\sigma = \sqrt{\frac{1}{12}} = 0,289$ $k = 4,7Cent/\sigma = \frac{4,7}{0,289} = 16,26$ $P( X - \mu  \geq k * \sigma) \leq \frac{1}{k^2} \Rightarrow P( X  \geq 16,26 * 0,289) \leq \frac{1}{16,26^2} = P( X  \geq 4,7) \leq 0.061$ ...?
--

d) Mit welcher Wahrscheinlichkeit verliert die Bank mindestens einen Euro (100 Cent)? Nutzen Sie den zentralen Grenzwertsatz, um diese Wahrscheinlichkeit geeignet zu approximieren. Runden Sie Ihr Ergebnis mit Hilfe der unten angegebenen Tabelle auf volle 10%. (3 Punkte)

Grenzwertsatz	$\lim_{n \rightarrow \infty} P(\sqrt{n} * \frac{\bar{X}_n - \mu}{\sigma} \leq x) = \phi(x)$
---------------	---

Quantile der Standardnormalverteilung

$\alpha$	10%	20%	30%	40%	50%	60%	70%	80%	90%
$q_\alpha$	-1,28	-0,84	-0,52	-0,25	0	0,25	0,52	0,84	1,28

## Aufgabe 6: Normalverteilung

Gegeben seien die unabhängigen Zufallsvariablen  $X_1, X_2, \dots, X_n$  mit den Verteilungen  $X_i \sim \mathcal{N}(\mu, 9)$  für  $i = 1, 2, \dots, n$  und einem reellen Parameter  $\mu$ .

a) Geben Sie einen erwartungstreuen Schätzer  $\hat{\mu}$  für  $\mu$  an. (1 Punkt)

b) Wie ist  $\hat{\mu}$  verteilt? (2 Punkte)

c) Geben Sie mit Hilfe von  $\hat{\mu}$  und der Faustformel ein 95%-Konfidenzintervall für  $\mu$  an. (2 Punkte)

d) Wie würden Sie die Hypothese  $\mu = 1$  zum Niveau 5% testen? (2 Punkte)

e) Berechnen Sie  $Cov(X_1 + X_2, X_2 + X_3)$ . (2 Punkte)

f) Sind  $X_1 + X_2$  und  $X_2 + X_3$  stochastisch unabhängig? (1 Punkt)

## Aufgabe 7: Sinus-Verteilung

a) Wann ist eine Funktion  $f : R \rightarrow R$  eine Wahrscheinlichkeitsdichte? (2 Punkte)

Gegeben sei  $f : R \rightarrow R$  mit  $f(x) = a * \sin(x) * 1_{(0,\pi)}(x)$ , wobei  $a$  ein reeller Parameter ist.

b) Bestimmen Sie  $a$  so, dass  $f$  eine Wahrscheinlichkeitsdichte ist. (1 Punkt)

Sei nun  $X$  eine Zufallsvariable, die die Wahrscheinlichkeitsdichte  $f$  besitzt.

c) Berechnen Sie  $P(X > \frac{\pi}{2})$ . (2 Punkte)

d) Berechnen Sie  $EX$ . (2 Punkte)

e) Begründen Sie, dass hier die Markoff-Ungleichung angewendet werden kann und verwenden Sie sie, um  $P(X > \frac{\pi}{2})$  abzuschätzen. Wie erklären Sie den Unterschied zu c)? (3 Punkte)

## Aufgabe 8: Deskriptive Statistik

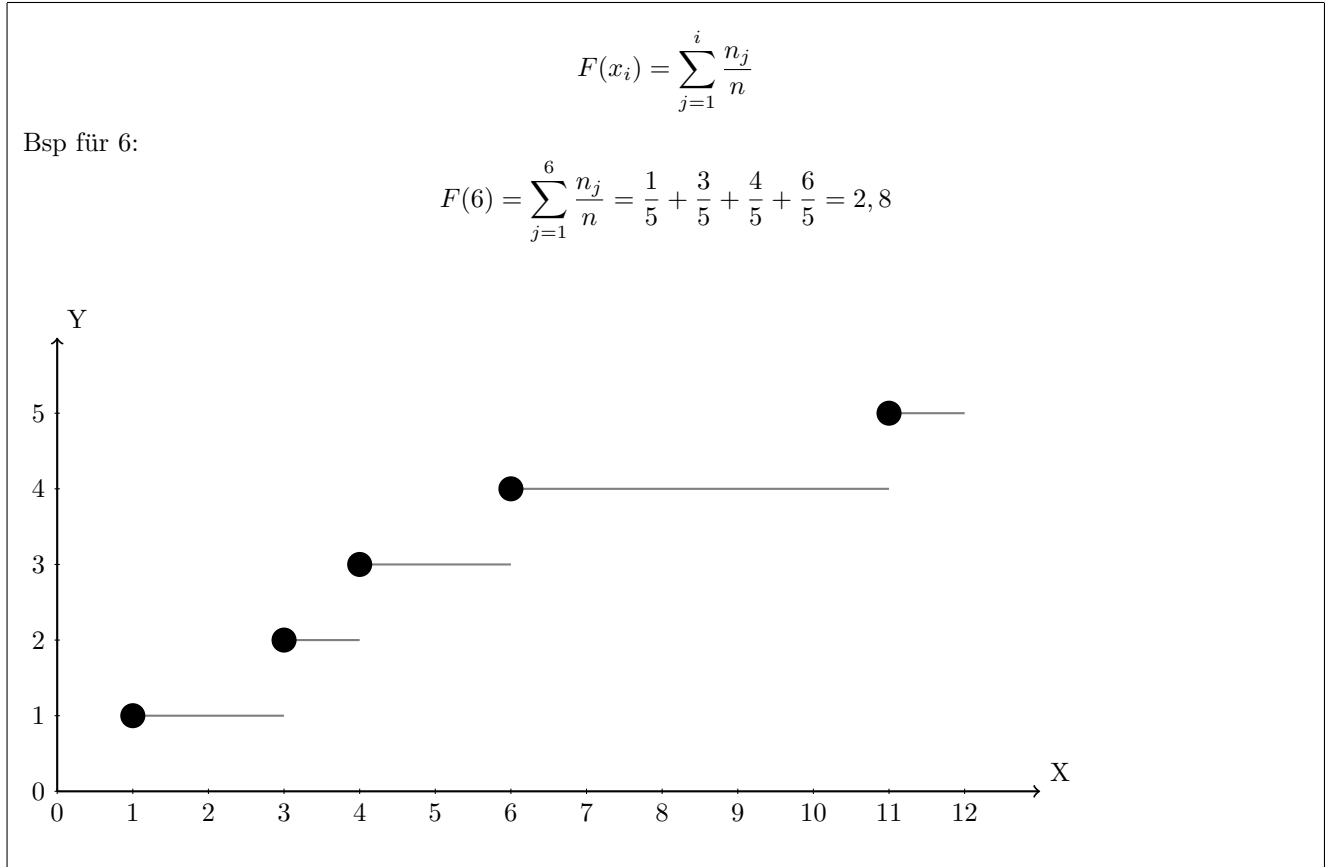
Aus einer Charge von Fäden werden 5 Stück entnommen, um sie auf Reißfestigkeit zu testen. Notiert werden die erreichten Dehnungslängen  $L_i, i = 1, \dots, 5$  in cm zum Zeitpunkt des Reißens. Die Ergebnisse lauten:

$L_1$	$L_2$	$L_3$	$L_4$	$L_5$
4	11	1	6	3

a) Befinden sich diese Daten auf einer Verhältnisskala? (1 Punkt)

Die Daten befinden sich auf der Ordinal-Skala, da diese benannt und in einer natürlichen Ordnung existieren.

b) Skizzieren Sie die empirische Verteilungsfunktion zu diesem Datensatz. (1 Punkt)



c) Bestimmen Sie den Mittelwert, den Median, die Quartile und den Interquartilsabstand der Daten. Wie erklärt sich der Unterschied zwischen Median und Mittelwert? (4 Punkte)

Mittelwert:

$$x_{mit} = \frac{1}{n} \sum_{i=1}^n x_i = 5$$

Median:  $x_{med} \geq 50\%$  aller Werte  $\Rightarrow x_{med} = 4$

Quartile:

$$Q_{0,75} = 1,75 * x_{mit} = 8,75; \quad Q_{0,25} = 0,25 * x_{mit} = 1,25$$

Interquartilsabstand:  $IQS = Q_{0,75} - Q_{0,25} = 8,75 - 1,25 = 7,5$

d) Skizzieren Sie einen Boxplot zu diesem Datensatz. (3 Punkte)

