

Aufgabe 1*(10 Punkte)***Teil A**

Sei die Zufallsvariable X gemäß einer Binomialverteilung mit $n = 25$ und $p = 0.2$ verteilt. Bestimmen Sie $P(X < \mu_X - 2\sigma_X)$. Dabei bezeichnet μ_X den Erwartungswert von X und σ_X die Standardabweichung von X .

Lösung: Die Zufallsvariable X ist binomialverteilt mit $n = 25$ und $p = 0.2$. Damit gilt für den Erwartungswert μ_X und die Varianz σ_X^2

$$\begin{aligned}\mu_X = E(X) &= np = 5, \\ \sigma_X^2 = \text{Var}(X) &= np(1-p) = 4.\end{aligned}$$

Damit folgt

$$\begin{aligned}P(X < \mu_X - 2\sigma_X) &= P(X < 5 - 2 \cdot 2) \\ &= P(X < 1) \\ &= P(X = 0) \\ &= \binom{n}{0} p^0 (1-p)^n = 0.8^{25} \approx 0.0038.\end{aligned}$$

Teil B

Sei die Zufallsvariable X normalverteilt mit Erwartungswert $\mu_X = 2$ und Varianz $\sigma_X^2 = 1$. Berechnen Sie die Wahrscheinlichkeit für das Ereignis $[|X - 2| < 1]$.

Lösung: Zu berechnen ist $P(|X - 2| < 1)$, wobei $X \sim N(2, 1)$.

$$\begin{aligned}P(|X - 2| < 1) &= P(1 < X < 3) \\ &= P(X < 3) - P(X < 1) \\ &= P\left(X < \frac{3-2}{1}\right) - P\left(X < \frac{1-2}{1}\right) \\ &= \phi(1) - \phi(-1) \\ &\approx 0.841 - 0.159 \\ &= 0.682\end{aligned}$$

Aufgabe 2*(10)*

Beweisen Sie

$$P(A|B) \geq P(A) \Rightarrow P(B|A) \geq P(B).$$

Lösung: Ein mögliche Lösung ist über den Satz der totalen Wahrscheinlichkeit herzuleiten

$$\begin{aligned}
 P(A|B) &= \frac{P(B|A)P(A)}{P(B)} \\
 \Leftrightarrow \underbrace{\frac{P(A|B)}{P(A)}}_{\geq 1} P(B) &= P(B|A) \\
 \Rightarrow P(B) &\leq P(B|A).
 \end{aligned}$$

Aufgabe 3

(15 Punkte)

Beantworten Sie die folgenden Fragen eindeutig mit JA oder NEIN. Begründen Sie die Antworten! (Für unbegründete richtige Antworten gibt es keine Punkte)

- Das Prognose-Intervall für die Anzahl der Erfolge bei n unabhängigen Wiederholungen eines Versuchs wird umso breiter, je größer n wird.

Lösung: JA

k : Anzahl der Erfolge

$P(a \leq k \leq b) \geq 1 - \alpha$ ist $(1 - \alpha)$ Prognoseintervall.

Approximatives Prognoseintervall, aus $k \sim N(n\pi, n\pi(1-\pi))$

Anteil:

$$\pi - \tau_{\frac{\alpha}{2}} \sqrt{\frac{\pi(1-\pi)}{n}} \leq p \leq \pi + \tau_{1-\frac{\alpha}{2}} \sqrt{\frac{\pi(1-\pi)}{n}}$$

Anzahl:

$$n\pi - \tau_{\frac{\alpha}{2}} \sqrt{n\pi(1-\pi)} \leq k \leq n\pi + \tau_{1-\frac{\alpha}{2}} \sqrt{n\pi(1-\pi)}$$

Das Prognoseintervall wird mit zunehmendem n immer größer.

- Die Lebensdauer Y eines Gerätes (in Tagen) sei exponentialverteilt mit dem Erwartungswert $EY = 20$. Es sei X die Lebensdauer der Geräte, die bereits 10 Tage hinter sich haben. Dann ist auch X exponentialverteilt mit dem Erwartungswert $EX = 20$.

Lösung: JA

Die Exponentialverteilung ist eine Verteilung "ohne Gedächtnis". Es gilt:

$$P(X \geq T + t | X \geq T) = P(X \geq t).$$

- Damit Erwartungswert und Median einer Zufallsvariable identisch sind, muß die Zufallsvariable symmetrisch verteilt sein.

Lösung: NEIN

Gegenbeispiel: $P(X = -3) = 1/6$, $P(X = -1) = 1/6$, $P(X = 0) = 1/3$, $P(X = 2) = 1/3$. Damit hat X den Erwartungswert $E(X) = 0$ und den Median $Median(X) = 0$, aber ist offensichtlich nicht symmetrisch.

4. Die Merkmale X und Y sollen gemeinsam von der Zeit T abhängen: Zu jedem Zeitpunkt $T = t$ sei die Korrelation zwischen $X(t)$ und $Y(t)$ negativ

$$\rho(X(t); Y(t)) < 0.$$

Muß dann bei Berücksichtigung aller Zeitpunkte auch die Korrelation

$$\rho(X; Y)$$

zwischen X und Y negativ sein? Begründen Sie die Antwort an einer Skizze.

Lösung: NEIN

Passende Skizze:

		$X_1^{(3)}$	$Y_2^{(3)}$
		$Y_1^{(3)}$	$X_2^{(3)}$
	$X_1^{(2)}$	$Y_2^{(2)}$	
	$Y_1^{(2)}$	$X_2^{(2)}$	
$t =$	1	2	3

5. Die folgenden drei Teilfrage sind nur mit JA oder NEIN zu beantworten. Eine Begründung ist nicht anzugeben. Eine richtige Antwort gibt +1 Punkte, eine falsche Antwort -1.

- (a) Ein erwartungstreuer Schätzer der Varianz ist gegeben durch

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

Lösung: JA.

- (b) Ein konsistenter Schätzer der Varianz ist gegeben durch

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

Lösung: NEIN.

- (c) Ein asymptotisch erwartungstreuer Schätzer ist auch konsistent.

Lösung: NEIN

Aufgabe 4*(10 Punkte)*

Heutzutage müssen Brötchen gewissen Gewichtsanforderungen entsprechen. Sie müssen mehr als $\mu_1 = 48.5$ Gramm wiegen, aber dürfen nicht schwerer als $\mu_2 = 52.5$ Gramm sein. Nun gibt es zwei Interessensgruppen, die an der Einhaltung dieser Richtlinien interessiert sind:

- die Verbraucher, die sicher sein wollen, daß das Mindestgewicht gegeben ist (Wenn die Brötchen zu leicht sind, werden die Verbraucher den Bäcker verklagen), und
- die Konkurrenz, die kontrolliert, daß die Brötchen nicht zu schwer werden (Wenn die Brötchen zu schwer sind, wird ein Bußgeld fällig).

Beide Interessengruppen überprüfen eine Bäckerei. Sie nehmen dabei an, daß das Gewicht der Brötchen normalverteilt ist mit Erwartungswert μ_0 und bekannter Varianz $\sigma^2 = 2$.

1. Wie werden die Hypothesen der beiden Interessengruppen (Verbraucher; Konkurrenz) lauten? Begründen Sie die Hypothesenwahl.

Lösung: Die Verbraucher wollen sicher sein, dass die Brötchen mindestens 48.5 Gramm wiegen. Sie wollen also mit einer vorbestimmten Wahrscheinlichkeit sagen können, daß die Brötchen nicht leichter sind als 48.5 Gramm. Daher die Nullhypothese $H_0 : \mu \geq \mu_1$ gegen die Alternative $H_1 : \mu < \mu_1$. Ein Ablehnung der Nullhypothese für zur Annahme der Alternative mit der Gewissheit nur einen Fehler mit Wahrscheinlichkeit α zu begehen. Die Vertauschung der Nullhypothese und der Alternative würde keine gesicherte Aussage über die Fragestellung ermöglichen.

Im Falle der Konkurrenz besteht das Interesse den Verstoß nachzuweisen. Die zu wählende Nullhypothese lautet $H_0 : \mu \leq \mu_2$ gegen die Alternative $H_1 : \mu > \mu_2$.

2. Beide Interessengruppen betrachten die gleiche Stichproben von 10 Brötchen. Dabei ergab sich ein mittleres Gewicht von 51.9 Gramm. Überprüfen Sie anhand dieser Stichprobe die Hypothesen der beiden Interessengruppen zu einem Signifikanzniveau von $\alpha = 0.05$. Formulieren Sie die Aussagen der Testergebnisse genau.

Lösung: Prüfgröße in beiden Fällen

$$T_i(X_1, \dots, X_n) = \sqrt{n} \frac{\bar{X} - \mu_i}{\sigma}$$

die unter der Nullhypothese $N(0, 1)$ verteilt ist.

Die kritische Region im Konkurrenzfall ist $KR = (u_{1-\alpha}; \infty)$ und im Verbraucherfall $KR =$

von ... bis ...	h_i	$f_i = \frac{h_i}{n}$
1 - 20	7	0.23
21 - 40	5	0.17
41 - 60	6	0.20
61 - 80	6	0.20
81 - 100	6	0.20
Summe	30	1.00

$(-\infty, -u_{1-\alpha})$. Für $\alpha = 0.05$ ist $u_{0.95} = 1.6445$.

Die Realisationen der Teststatistiken und die Entscheidungen sind

$$T_1(x_1, \dots, x_{10}) = \sqrt{10} \frac{51.9 - 48.5}{\sqrt{2}} = 7.60 \notin KR = (-\infty; -1, 6445) \Rightarrow H_0 \text{ wird nicht abgelehnt,}$$

$$T_1(x_1, \dots, x_{10}) = \sqrt{10} \frac{51.9 - 52.5}{\sqrt{2}} = -1.34 \notin KR = (1.6445; \infty) \Rightarrow H_0 \text{ wird nicht abgelehnt.}$$

Damit können die Verbraucher nicht davon ausgehen, dass die Brötchen das Mindestgewicht unterschreiten.

Die Konkurrenz kann nicht klagen, denn es ist nicht davon auszugehen, dass die Brötchen zu schwer sind.

Aufgabe 5

(15 Punkte)

Gegeben seien 30 Beobachtungen der Zufallsvariablen X , deren Realisationen die Werte $1, 2, 3, \dots, 100$ annehmen können. Die 30 Beobachtungen wurden zu folgender Tabelle zusammengefaßt:

von ... bis ...	1-20	21-40	41-60	61-80	81-100
Anzahl	7	5	6	6	6

Teil A

1. Stellen Sie die relative Klassenhäufigkeit in einem Histogramm dar.

Lösung: Berechnung der relative Klassenhäufigkeiten:

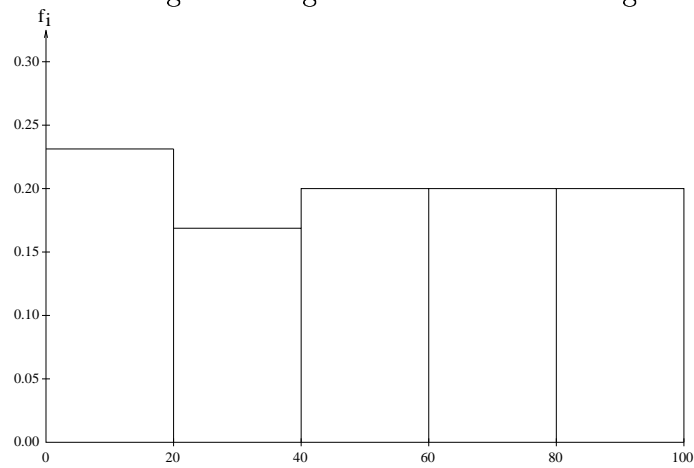
Das Histogramm ist in Abbildung 1 auf Seite 6 zu sehen.

2. Zeichnen Sie die empirische Verteilungsfunktion der Daten.

Lösung: Berechnung der Werte für die Verteilungsfunktion:

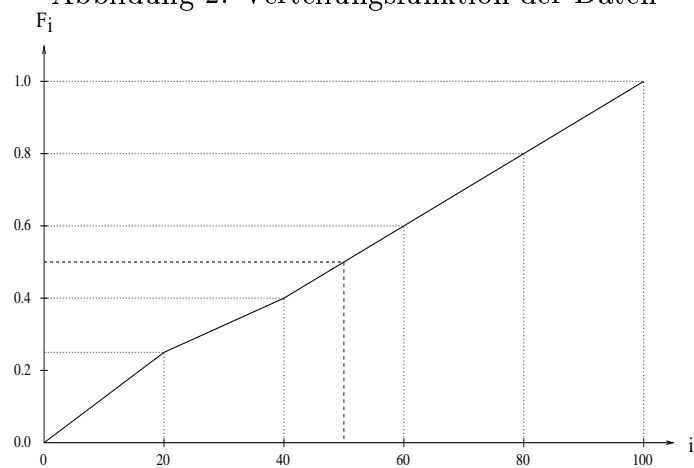
Die Verteilungsfunktion ist in Abbildung 2 auf der Seite 6 zu sehen.

Abbildung 1: Histogramm der Beobachtungen



i	f_i	F_i
1 - 20	0.23	0.23
21 - 40	0.17	0.40
41 - 60	0.20	0.60
61 - 80	0.20	0.80
81 - 100	0.20	1.00

Abbildung 2: Verteilungsfunktion der Daten



3. Bestimmen Sie graphisch den empirischen Median der gruppierten Daten.

Lösung: Aus der Grafik kann man den Median bei 50 ablesen.

4. Berechnen Sie das arithmetische Mittel und die empirische Varianz der klassierten Beobachtungen.

Lösung: Arithmetisches Mittel bei gruppierten Daten:

Die Formel hierfür lautet: $\bar{x} = \frac{1}{n} \sum_{j=1}^k m_j \cdot h_j$ wobei $m_j = \frac{1}{2}(g_i + g_{i+1})$ die Gruppenmitte ist.

$$\begin{aligned}\bar{x} &= \frac{1}{30} (10.5 \cdot 7 + 30.5 \cdot 5 + 50.5 \cdot 6 + 70.5 \cdot 6 + 90.5 \cdot 6) \\ &= \frac{1}{30} (73.5 + 152.5 + 303 + 423 + 543) \\ &= \frac{1}{30} \cdot 1495 \\ \bar{x} &= 49.8\bar{3}\end{aligned}$$

Das arithmetische Mittel beträgt 49.83.

Varianz bei gruppierten Daten:

Die Formel hierfür lautet: $var x = \frac{1}{n} \sum_{j=1}^k (m_j - \bar{x})^2 \cdot h_j$

$$\begin{aligned}var x &= \frac{1}{30} ((10.5 - 49.83)^2 \cdot 7 + (30.5 - 49.83)^2 \cdot 5 + (50.5 - 49.83)^2 \cdot 6 + (70.5 - 49.83)^2 \cdot 6 + \\ &= \frac{1}{30} (10\,827.94 + 1\,868.24 + 2.69 + 2\,563.49 + 9\,924.29) \\ &= \frac{1}{30} \cdot 25\,186.67 \\ var x &= 839.59\end{aligned}$$

Die Varianz beträgt 839.59.

Teil B

1. Skizzieren Sie eine Testidee für die Hypothese "Die Beobachtungen stammen aus einer Gleichverteilung". Wie lautet eine Prüfgröße?

Lösung: Es handelt sich hierbei um einen χ^2 -Anpassungstest. Die Hypothesen lauten:

$$H_0 : X \sim GL(a, b)$$

$H_1 : X$ ist nicht gleichverteilt

Berechnung der Prüfgröße:

$$PG: \chi_{PG}^2 = \sum \frac{(n f_i - n \pi)^2}{n \pi}$$

2. Unter welchen Voraussetzungen kann die Verteilung der Prüfgröße approximativ angegeben werden?

Lösung: Das gilt wenn n groß ist und wenn alle $n \pi_i \geq 1$ und die meisten $n \pi_i \geq 5$ sind. Die Klassen müssen so gewählt werden, daß die genannten Bedingungen gelten. In unserem Fall sind alle Bedingungen gewährleistet.

3. Geben Sie nun den Ablehnbereich für den Test zum Signifikanzniveau von $\alpha = 0.05$ an.

Lösung: Ablehnbereich: $\chi_{PG}^2 > \chi_{k-1; 1-\alpha}^2$ wobei $\chi_{4; 0.95}^2 = 9.4877$ ist.

Name:

Matr.-Nr.

Platznr.

4. Die berechnete Prüfgröße ist kleiner als 2. Zu welchem Schluß kommen Sie?

Lösung: Da die berechnete Prüfgröße kleiner als 9.4877 ist, kann angenommen werden, daß die Beobachtungen aus einer Gleichverteilung stammen.