Universität des Saarlandes
Lehrstab Statistik
PD Dr. Martin Becker

8. Übungsblatt zur Vorlesung
Schließende Statistik WS 2019/20

Aufgabe 28

Das Reiseunternehmen „Schöne alte Welt“ führt regelmäßig Buch über den Auslastungsgrad Y (in %) seiner Reisebusse. Erfahrungsgemäß sollte es sich bei Y um eine normalverteilte Zufallsvariable mit dem Erwartungswert μ und der Varianz σ^2 handeln (wobei man annimmt: $\sigma_0^2 = 49$). Nach einer Werbeaktion, mit der man den Auslastungsgrad der Busse steigern wollte, hatte man den Verdacht, dass dieser wesentlich stärker streute als angenommen. Eine einfache Stichprobe zu Y ergab die Auslastungsgrade (in %)

$$91, 86, 54, 65, 70, 66, 70, 82 \; .$$

Kann man von einer signifikanten Vergrößerung der Streuung des Auslastungsgrades der Busse sprechen ($\alpha = 0.01$)?

Aufgabe 29

Der Ertrag einer neuen Getreidesorte sei normalverteilt mit bekanntem Erwartungswert $\mu = 3.5$ und unbekannter Varianz σ^2. Vor Vermarktungsstart der Getreidesorte wurden 14 Flächen gleicher Größe mit dieser Sorte bestellt und die Erträge (in t/ha) ermittelt. Man erhielt dabei folgende Werte als Realisation einer einfachen Stichprobe zum Ertrag Y:

$$3.06, 3.63, 2.91, 4.63, 3.73, 2.92, 3.84, 4.02, 3.91, 3.28, 4.57, 3.78, 3.06, 1.93$$

Kann auf Grundlage eines Hypothesentests zum Signifikanzniveau $\alpha = 0.05$ davon ausgegangen werden, dass für die Varianz des Ertrags $\sigma^2 = 0.5$ $[\text{t/ha}]^2$ gilt?
(Hinweis: es gilt $\frac{1}{14}\sum_{i=1}^{14}(x_i - 3.5)^2 = 0.479$)

Aufgabe 30

Zeigen Sie:
Für die Testgröße χ^2 im Chi-Quadrat-Anpassungstest gilt mit den üblichen Bezeichnungen

$$\chi^2 = \sum_{i=1}^{k} \frac{(n_i - np_i^0)^2}{np_i^0} \stackrel{!}{=} \left(\frac{1}{n}\sum_{i=1}^{k}\frac{n_i^2}{p_i^0}\right) - n \; .$$

Aufgabe 31

Es werde vermutet, dass die Zeit Y (in Sekunden) zwischen zwei Telefonanrufen in einer Telefonzentrale einer Exp(0.5)-Verteilung genügt. Die Ziehung einer einfachen Stichprobe vom Umfang $n = 200$ zu Y lieferte das folgende Stichprobenergebnis:

i	1	2	3	4	5
K_i	$(-\infty, 1]$	$(1, 2]$	$(2, 3]$	$(3, 4]$	$(4, \infty)$
n_i	82	55	31	19	13

Testen Sie zum Signifikanzniveau $\alpha = 0.05$, ob die Zeit zwischen zwei Telefonanrufen Exp(0.5)-verteilt ist.
Hinweis: Die Verteilungsfunktion einer Exp(0.5)*-verteilten Zufallsvariable ist bekanntlich gegeben durch*

$$F_{\text{Exp}(0.5)} : \mathbb{R} \to \mathbb{R}; F_{\text{Exp}(0.5)}(x) = \begin{cases} 0 & \textit{für } x < 0 \\ 1 - e^{-0.5 \cdot x} & \textit{für } x \geq 0 \end{cases} .$$

Aufgabe 32

In einer bekannten Quiz-Show werden einem Kandidaten nacheinander mehrere Fragen sowie zu jeder Frage jeweils vier Antwortmöglichkeiten mit den Bezeichnungen „A“, „B“, „C“ und „D“ präsentiert, von denen genau eine richtig ist. Der Kandidat kann zu seiner Unterstützung bei einer Frage seiner Wahl den sogenannten „Publikums-Joker“ einsetzen. Hierbei haben die Zuschauer im Fernsehstudio ebenfalls die Gelegenheit, sich für eine der vier Antworten zu entscheiden, sie können sich aber auch enthalten. Dem Kandidaten werden anschließend die relativen Häufigkeiten der abgegebenen Zuschauerantworten präsentiert.
In einer bestimmten Ausgabe der Show erhält der Kandidat das folgende Ergebnis:

Antwort	A	B	C	D
Anteil in Prozent	34	16	20	30

Der Kandidat ist sich unsicher, ob er bei diesem Ergebnis davon ausgehen soll, dass sich die teilnehmenden Zuschauer unabhängig voneinander rein zufällig für eine der vier Antworten entschieden haben, oder eher davon, dass ein gewisses „Wissen“ im Publikum vorhanden ist.

(a) Der Show-Master teilt dem Kandidaten mit, dass sich genau 150 Zuschauer an der Abstimmung beteiligt haben. Überprüfen Sie unter der Annahme, dass sich die abgegebenen Zuschauerstimmen als einfache Stichprobe auffassen lassen, mit einem geeigneten statistischen Test, ob sich die beteiligten Zuschauer rein zufällig für eine der vier Antworten entschieden haben oder nicht (Signifikanzniveau $\alpha = 0.05$).

(b) Würde sich das Ergebnis des Tests aus Teil (a) ändern, wenn nur 50 Zuschauer an der Abstimmung teilgenommen hätten? Begründen Sie Ihre Antwort! (*Beachten Sie, dass Sie den Test nicht komplett neu durchführen müssen und insbesondere die realisierte Teststatistik relativ leicht aus dem Ergebnis des vorherigen Aufgabenteils gewinnen können!*)

(c) Welche Anzahl an der Abstimmung beteiligter Zuschauer müsste der Show-Master mindestens nennen, um nach dem in den Teilen (a) und (b) verwendeten Test davon auszugehen, dass sich die beteiligten Zuschauer *nicht* rein zufällig für eine der vier Antwortmöglichkeiten entschieden haben?

Hinweis: Verwenden Sie den folgenden Tabellenausschnitt mit p-Quantilen von $\chi^2(n)$-Verteilungen:

$n \backslash p$	0.01	0.025	0.05	0.50	0.90	0.95	0.975	0.99
1	0.000	0.001	0.004	0.455	2.706	3.841	5.024	6.635
2	0.020	0.051	0.103	1.386	4.605	5.991	7.378	9.210
3	0.115	0.216	0.352	2.366	6.251	7.815	9.348	11.345
4	0.297	0.484	0.711	3.357	7.779	9.488	11.143	13.277
5	0.554	0.831	1.145	4.351	9.236	11.070	12.833	15.086