Universität des Saarlandes
Lehrstab Statistik
PD Dr. Martin Becker

11. Übungsblatt zur Vorlesung
Schließende Statistik WS 2019/20

Aufgabe 42

Anhand der Ergebnisse der Klausur zur Veranstaltung „Deskriptive Statistik und Wahrscheinlichkeitsrechnung" des Sommersemesters 2010 soll mit Hilfe einer einfachen Varianzanalyse untersucht werden, ob die Verteilung der von den Studierenden im Fach „Deskriptive Statistik und Wahrscheinlichkeitsrechnung" erreichten Punktzahlen abhängig davon ist, in welchem Prüfungsversuch (1. Versuch – 3. Versuch) sich der Studierende befindet. Hierzu wurden folgende Daten zu den erreichten Punktzahlen erhoben:

j (Versuch)	n_j	$\overline{x}_j = \frac{1}{n_j}\sum_{i=1}^{n_j} x_{j,i}$	$\sum_{i=1}^{n_j} x_{j,i}^2$	s_j^2
1	312	59.431	1237338.75	435.180
2	41	51.841	121766.75	289.492
3	15	62.633	64820.25	426.919

Für die Durchführung der einfachen Varianzanalyse wurde hieraus bereits die Größe $SW = 152897.65$ berechnet.

(a) Überprüfen Sie mit einer einfachen Varianzanalyse (Signifikanzniveau $\alpha = 0.05$) unter der Annahme, dass die erreichten Punktzahlen Realisierungen von unabhängig $N(\mu_j, \sigma^2)$-verteilten Zufallsvariablen $X_{j,i} (1 \leq j \leq 3, 1 \leq i \leq n_j)$ sind, ob die Ausprägung des Merkmals „Prüfungsversuch" einen Einfluss auf die erreichte Punktzahl hat. Fassen Sie das Ergebnis der Varianzanalyse auch in einem Antwortsatz zusammen.

(b) Geben Sie zwei unterschiedliche Formeln an, mit denen die (vorgegebene) Größe SW auch aus den Angaben der obigen Tabelle berechnet werden könnte.

Hinweis: Verwenden Sie für Aufgabenteil (a) den folgenden Tabellenausschnitt mit 0.95*-Quantilen von* $F(m, n)$*-Verteilungen:*

$n\backslash m$	1	2	3	4	5	364	365	366	367	368
1	161.448	199.500	215.707	224.583	230.162	253.964	253.965	253.966	253.967	253.968
2	18.513	19.000	19.164	19.247	19.296	19.493	19.493	19.493	19.493	19.493
3	10.128	9.552	9.277	9.117	9.013	8.534	8.534	8.534	8.534	8.534
4	7.709	6.944	6.591	6.388	6.256	5.638	5.638	5.638	5.638	5.638
5	6.608	5.786	5.409	5.192	5.050	4.376	4.376	4.376	4.376	4.376
364	3.867	3.021	2.629	2.396	2.239	1.188	1.188	1.188	1.188	1.188
365	3.867	3.020	2.629	2.396	2.239	1.188	1.188	1.188	1.188	1.188
366	3.867	3.020	2.629	2.396	2.239	1.188	1.188	1.188	1.188	1.188
367	3.867	3.020	2.629	2.396	2.239	1.188	1.188	1.188	1.188	1.187
368	3.867	3.020	2.629	2.396	2.239	1.188	1.188	1.188	1.187	1.187

Aufgabe 43

Je 5 gleichaltrige Schüler zweier Volksschulklassen (eine Mädchen- und eine Jungenklasse) und zweier Klassen aus Gymnasien (ebenfalls eine Mädchen- und Jungenklasse) sollen auf unterschiedliches technisches Verständnis untersucht werden. Bei dem Versuch hat jedes Kind einige einfache Apparaturen zusammenzusetzen. Es wurde für jedes Kind die Zeit (in Minuten) gemessen bis die Aufgabe gelöst war. Die Ergebnisse der Messung sind in folgender Tabelle zusammengefasst:

Klasse	Schüler				
1	12	13	11	13	11
2	13	14	12	15	11
3	15	15	15	13	12
4	16	18	17	14	20

Es bezeichne $x_{j,i}$ die Zeit des i-ten Schülers bzw. der i-ten Schülerin der j-ten Klasse ($1 \leq j \leq 4, 1 \leq i \leq n_j$). Unter der Annahme, dass die einzelnen Zeiten $x_{j,i}$ Realisierungen von unabhängigen identisch $N(\mu_j, \sigma^2)$-verteilten Zufallsvariablen $X_{j,i}$ ($1 \leq j \leq 4, 1 \leq i \leq n_j$) sind, teste man anhand dieser Daten, ob die Schüler bzw. Schülerinnen der vier Klassen im Mittel gleich geschickt sind. (Signifikanzniveau $\alpha = 0.05$)

Aufgabe 44

Beweisen Sie die Streuungszerlegung

$$SS = SB + SW$$

der einfachen Varianzanalyse.

Aufgabe 45

Gegeben seien n Punkte (y_i, x_i), durch die eine Gerade $y = a + b \cdot x$ gelegt werden soll.

(a) Wie lautet das Kleinst-Quadrate-Prinzip?

(b) Zeigen Sie, dass das Kleinst-Quadrate-Prinzip auf die beiden Normalgleichungen

(1) $na + b\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} y_i$

(2) $a\sum_{i=1}^{n} x_i + b\sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} x_i \cdot y_i$

führt.

(c) Lösen Sie das Gleichungssystem aus den beiden Normalgleichungen, um die Kleinst-Quadrate-Lösungen $\widehat{a}$ und $\widehat{b}$ zu erhalten.