

Matrikel-Nr.					
--------------	--	--	--	--	--

Modulprüfung

Prüfungsfach: Statistik II für Betriebswirte
 Prüfer: Prof. Hans-Jörg Starkloff
 Tag: 8. März 2018

Aufgabe	1	2	3	4	5	6	Σ
erreichbare Punkte	6	6	8	6	6	8	40
erreichte Punkte							

**Geben Sie zu allen Aufgaben einen nachvollziehbaren Lösungsweg an!
 Beantworten Sie die Fragen möglichst kurz und mit eigenen Worten!**

- 1. Aufgabe:** Eine noch junge Firma aus dem IT-Bereich möchte ihre Position in der Branche festigen, indem sie ihren Marktanteil und den damit verbundenen Umsatz steigert.

Basierend auf den bisherigen Erfahrungen und den Umsatzzahlen beschließt das Unternehmen, zwei favorisierte Verkaufsstrategien auszuprobieren.

Dazu hat die Firma die Zeit in Minuten bis zur Abhandlung der ersten 50 Verkäufe gemessen. Diese Zeiten sind in folgender Tabelle zu finden.

Strategie 1	58	92	86	78	122	94
Strategie 2	103	233	205	151	114	

Testen Sie zum Signifikanzniveau von 5%, ob die erwartete Zeit bei Strategie 1 kleiner ist als bei Strategie 2, ob Strategie 1 also die bessere ist.

Hinweis: Die Zeit ist **nicht** normalverteilt.

- 2. Aufgabe:** Ein neues Geschäft in der Innenstadt analysiert die aktuelle Kundschaft. Dazu wird eine Befragung in der Fußgängerzone durchgeführt. Die Frage, ob man schon Kunde des Geschäftes ist, kann mit „Ja“ oder „Nein“ beantwortet werden.

Das Ergebnis der Befragung, getrennt nach drei Altersklassen, findet man in folgender Tabelle:

Alter	Kunde	
	Ja	Nein
unter 18 Jahre	3	17
über 18 Jahre und unter 30 Jahre	9	11
über 30 Jahre	18	2

Testen Sie zum Signifikanzniveau $\alpha = 0,05$, ob die Eigenschaft, Kunde des Geschäftes zu sein, vom Alter abhängig ist.

3. **Aufgabe:** Für drei Weizensorten (Sorte 1 bis Sorte 3) wurden drei verschiedene Standorte (Standort 1 bis Standort 3) hinsichtlich der Erträge untersucht. Die Erträge wurden in immer der gleichen Einheit gemessen und wie folgt ausgewertet.

Analysis of Variance for Ertrag - Type I Sums of Squares

<i>Source</i>	<i>Sum of Squares</i>	<i>Df</i>	<i>Mean Square</i>	<i>F-Ratio</i>	<i>P-Value</i>
MAIN EFFECTS					
A:Standort	604,627	2	302,314	164,37	0,0000
B:Weizensorte	287,832	2	143,916	78,25	0,0000
INTERACTIONS					
AB	61,5726	4	15,3931	8,37	0,0005
RESIDUAL	33,1067	18	1,83926		
TOTAL (CORRECTED)	987,139	26			

All F-ratios are based on the residual mean square error.

- a) Die Modellgleichung lautet:

$$X_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + \varepsilon_{ijk}, \quad i = 1, \dots, 3; \quad j = 1, \dots, 3; \quad k = 1, \dots, 3.$$

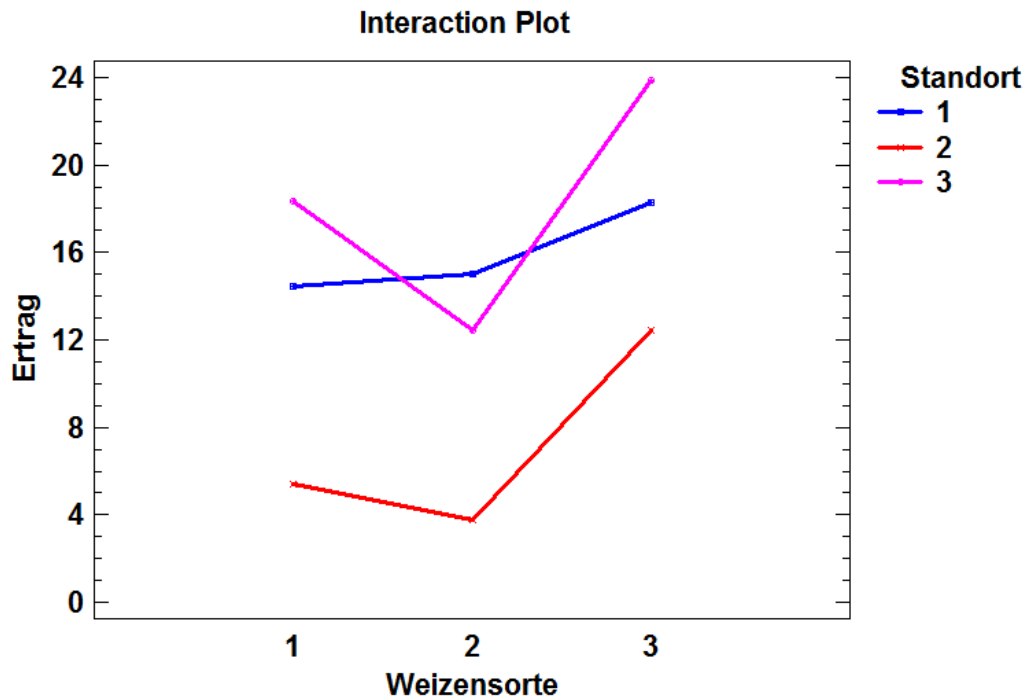
Was sind in dieser Gleichung die Größen X , α_i , β_j , und γ_{ij} ?

- b) Welche Hypothesen werden in den Zeilen B:Weizensorte und AB der obigen Tabelle getestet und wie lauten die Testentscheidungen und die Testergebnisse bei einem Signifikanzniveau von $\alpha = 0,01$?
- c) Schätzen Sie den Wechselwirkungseffekt zwischen dem ersten Standort (Standort 1) und der zweiten Weizensorte (Sorte 2)!

Table of Least Squares Means for Ertrag

<i>Level</i>	<i>Count</i>	<i>Mean</i>
GRAND MEAN	27	13,8
Standort		
1	9	15,9
2	9	7,2
3	9	18,2
Weizensorte		
1	9	12,7
2	9	10,4
3	9	18,2
Standort by Weizensorte		
1,1	3	14,4
1,2	3	15,0
1,3	3	18,3
2,1	3	5,4
2,2	3	3,8
2,3	3	12,5
3,1	3	18,4
3,2	3	12,4
3,3	3	23,9

d) Welches Testergebnis aus b) legt der folgende Plot nahe und warum?



4. Aufgabe: In einer Untersuchung wurden die Merkmale

X_1 - Alter,

X_2 - Merkfähigkeit und

X_3 - intellektuelle Aktivität
erfasst.

Aus der Stichprobe von 20 Personen erhält man folgende Schätzung der Korrelationen:

$$r_{X_1, X_2} = -0,65 \quad ; \quad r_{X_1, X_3} = -0,69 \quad \text{und} \quad r_{X_2, X_3} = 0,82.$$

- Schätzen Sie die partielle Korrelation ($\rho_{X_1, X_2 | X_3}$) zwischen Alter und Merkfähigkeit unter Partialisierung der intellektuellen Aktivität.
- Testen Sie zum Signifikanzniveau $\alpha = 0,05$, ob die partielle Korrelation zwischen Alter und Merkfähigkeit unter Partialisierung der intellektuellen Aktivität signifikant kleiner als 0 ist.

- 5. Aufgabe:** Die im Liefervertrag vereinbarten Qualitätsnormen sollen durch eine Stichprobe überprüft werden. Dafür haben sich der Produzent und der Konsument auf Folgendes geeinigt.

Eine Lieferung mit höchstens 1% Ausschuss ist eine gute Lieferung und eine mit mindestens 5% Ausschuss ist eine schlechte Lieferung. Eine gute Lieferung soll mit einer Wahrscheinlichkeit von höchstens 0,03 abgelehnt und eine schlechte mit einer Wahrscheinlichkeit von höchstens 0,02 angenommen werden.

- Bestimmen Sie eine untere Grenze für den Stichprobenumfang n .
- Bestimmen Sie, falls möglich, für $n = 300$ alle (n, c) -Stichprobenpläne, welche die Bedingungen erfüllen!
- Für den Stichprobenumfang $n = 400$ gibt es folgende drei Pläne, welche die Bedingungen erfüllen:

$$(n, c) = (400, 8)$$

$$(n, c) = (400, 9)$$

$$(n, c) = (400, 10)$$

Welchen dieser Pläne würde der Konsument bevorzugen? Begründen Sie kurz!

6. Aufgabe:

- Warum gilt für die Quantile der t -Verteilung mit n Freiheitsgraden und $p \in (0, 1)$: $t_{n, 1-p} = -t_{n, p}$?
- Vervollständigen Sie die folgende Korrelationsmatrix dreier Merkmale:

$$\begin{pmatrix} 1 & 0.7 & \\ & & 0.8 \\ 0.2 & & \end{pmatrix}$$

- Zeichnen Sie drei Punkte (x_i, y_i) $i = 1, \dots, 3$, so dass zwei Punktepaare diskordant sind und ein Punktepaar konkordant ist.
- Warum ist im Modell der multiplen parameterlinearen Regression das Prognoseintervall für $Y(\underline{x})$ größer als das Konfidenzintervall für $\mathbf{E}(Y(\underline{x}))$? Kurze inhaltliche Begründung!
- Wie lauten im Modell mit zwei Einflussgrößen x_1 und x_2 und den Parametern a , b und c

$$Y(x_1, x_2) = a \cdot \sin(2x_1) - b \cdot \cos(-4x_2) + c \cdot (x_1 + x_2) + \varepsilon$$

die bekannten Funktionen f_1, \dots, f_r des allgemeinen parameterlinearen Ansatzes:

$$Y(x_1, x_2) = a_1 \cdot f_1(x_1, x_2) + \dots + a_r \cdot f_r(x_1, x_2) + \varepsilon$$

- Ein Posten vom Umfang N mit M Ausschussteilen wird durch eine zweistufige Stichprobe überprüft. Der zweistufige Stichprobenplan hat als Parameter n_1, n_2 und c_1, c_2, c_3 . Weiter sei X_1 die zufällige Anzahl der Ausschussteile in der ersten Stichprobe und X_2 die in der zweiten Stichprobe.

Wie ist X_2 bedingt $X_1 = x_1$ mit $c_1 < x_1 \leq c_2$ verteilt?

Oder, anders gefragt, wie groß ist $P(X_2 = x_2 | X_1 = x_1)$?