

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

Modulprüfung

Prüfungsfach: Statistik II für Betriebswirte
 Prüfer: PD Dr. Frank Heyde
 Tag: 24. Juli 2014

Aufgabe	1	2	3	4	5	6
erreichbare Punkte	4	5	5	6	4	6
erreichte Punkte						

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

- 1. Aufgabe:** Bei einem Würfelspiel kommt einem Spieler der Verdacht, dass sein Mitspieler möglicherweise einen gezinkten Würfel verwendet, der in mehr als $1/6$ der Fälle eine Sechs würfelt. Er lässt sich den Würfel des Mitspielers geben und würfelt 50 mal. Unter seinen Ergebnissen befinden sich 18 Sechsen.

Bestimmen Sie zum Konfidenzniveau von 0,95 eine untere Konfidenzgrenze (und damit ein einseitiges Konfidenzintervall) für die Wahrscheinlichkeit, dass eine Sechs gewürfelt wird (d.h. für den Anteil der Sechsen).

- 2. Aufgabe:** Im Rahmen einer Studie wird diskutiert, ob eher Frauen oder Männer eine höhere Handynutzung haben. Dazu wurden zufällig sieben Jungen und sieben Mädchen ausgewählt und deren durchschnittliche Rechnungshöhe (in €) der letzten 4 Monate erfasst.

Jungen	77,45	32,18	15,43	3,03	82,58	27,31	52,39
Mädchen	26,37	62,75	22,81	71,56	51,87	73,74	61,54

Es wird vermutet, da Mädchen kommunikativer sind, dass sich das bei der Handynutzung widerspiegelt. Testen Sie zum Niveau $\alpha = 0,01$, ob die erwartete durchschnittliche Rechnungshöhe der Mädchen signifikant größer ist als die der Jungen. Die durchschnittlichen Rechnungshöhen sind **nicht** normalverteilt.

- 3. Aufgabe:** Eine sächsische Molkerei füllt Milch in Tetrapacks ab. Es wird vermutet, dass die Füllmenge normalverteilt ist.

In einer Stichprobe von 1000 Tetrapacks wurden die Füllmengen gemessen und damit ein Test mit folgendem Testergebnis durchgeführt:

Goodness-of-Fit Tests for Füllmenge

Chi-Square Test

	<i>Lower</i>	<i>Upper</i>	<i>Observed</i>	<i>Expected</i>	
	<i>Limit</i>	<i>Limit</i>	<i>Frequency</i>	<i>Frequency</i>	<i>Chi-Square</i>
at or below		497,333	4	3,96	0,00
	497,333	497,667	4	5,98	0,66
	497,667	498,0	18	12,74	
	498,0	498,333	19	24,37	1,19
	498,333	498,667	44	41,95	0,10
	498,667	499,0	60	64,93	0,37
	499,0	499,333	100	90,38	1,02
	499,333	499,667	105	113,14	0,59
	499,667	500,0	124	127,37	0,09
	500,0	500,333	141	128,95	1,13
	500,333	500,667	120	117,41	0,06
	500,667	501,0	100	96,13	0,16
	501,0	501,333	65	70,79	0,47
	501,333	501,667	46	46,88	0,02
	501,667	502,0	25	27,92	0,31
	502,0	502,333	10	14,95	
	502,333	502,667	6	7,20	0,20
above	502,667		9	4,94	3,33

Chi-Square = 13,4925

- Bestimmen Sie die beiden fehlenden Werte in der Tabelle (Spalte Chi-Square).
- Geben Sie die Null- und die Alternativhypothese des Testes an, führen Sie den Test zu Ende und treffen Sie die Testentscheidung zum Niveau $\alpha = 0,05$.

4. **Aufgabe:** Der Produzent und der Konsument einigen sich auf Folgendes. Ein Posten mit $p \leq 0,02 = p_\alpha$ wird als gut und ein Posten mit $p \geq 0,04 = p_\beta$ als schlecht angesehen. Das Risiko des Produzenten betrage 3%. Das Risiko des Konsumenten betrage 5%. Zur Stichprobenkontrolle soll ein (n,c) -Stichprobenplan verwendet werden.

- Bestimmen Sie eine untere Grenze für den Stichprobenumfang n .
- Bestimmen Sie, falls möglich, für $n = 1025$ alle (n, c) -Stichprobenpläne, welche die Bedingungen erfüllen!
Hinweis: Nutzen Sie die Normalverteilungs-Approximation.
- Für den Stichprobenumfang $n = 1110$ gibt es folgende 3 Pläne, welche die Bedingungen erfüllen:

$$(n, c) = (1110, 31)$$

$$(n, c) = (1110, 32)$$

$$(n, c) = (1110, 33)$$

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

5. **Aufgabe:** Für 29 PKWs wurden die Merkmale

X_1 - Alter,

X_2 - Leistung und

Y - Verbrauch erfasst.

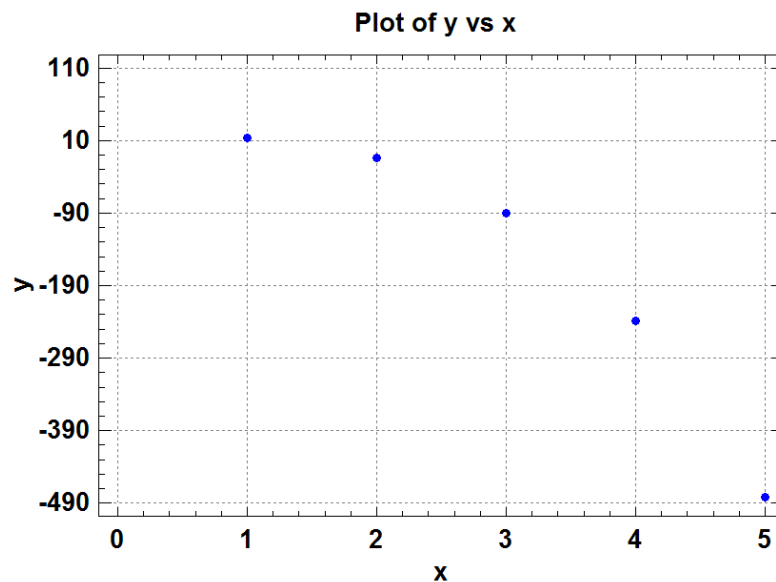
Aus der Stichprobe erhält man folgende Schätzung der Korrelationen:

$$r_{X_1, X_2} = -0,18 \quad ; \quad r_{Y, X_1} = 0,39 \quad \text{und} \quad r_{Y, X_2} = 0,51.$$

- Schätzen Sie die multiple Korrelation $\rho_{Y, (X_1, X_2)}$ zwischen dem Verbrauch einerseits und dem Alter und der Leistung andererseits.
- Testen Sie (unter der Annahme, dass die Merkmale normalverteilt sind), ob die multiple Korrelation $\rho_{Y, (X_1, X_2)}$ signifikant ($\alpha = 0,01$) größer als 0 ist.

6. **Aufgabe:**

- Im Modell der einfachen linearen Regression wurde aus den Daten (x_i, y_i) , $i = 1, \dots, 100$, die Regressionsgerade geschätzt und dabei $SST = 1000$ und $r_{x,y} = 0,9$ bestimmt.
 - Wie groß ist das Bestimmtheitsmaß?
 - Wie groß ist SSR?
- Warum ist der 2-stufige Stichprobenplan mit: $n_1 = 5$, $n_2 = 5$, $c_1 = 0$, $c_2 = 4$ und $c_3 = 10$ praktisch gleich einem einstufigen Plan?
- Wie groß ist für die folgenden Datenpaare (x_i, y_i) , $i = 1, \dots, 5$,



die Rangkorrelation von Spearman? Kurze Begründung!

d) Mit den Daten aus der 3. Aufgabe wurde getestet:

H_0 : „ Die Füllmenge ist exponentialverteilt. “

H_A : „ Die Füllmenge ist nicht exponentialverteilt. “

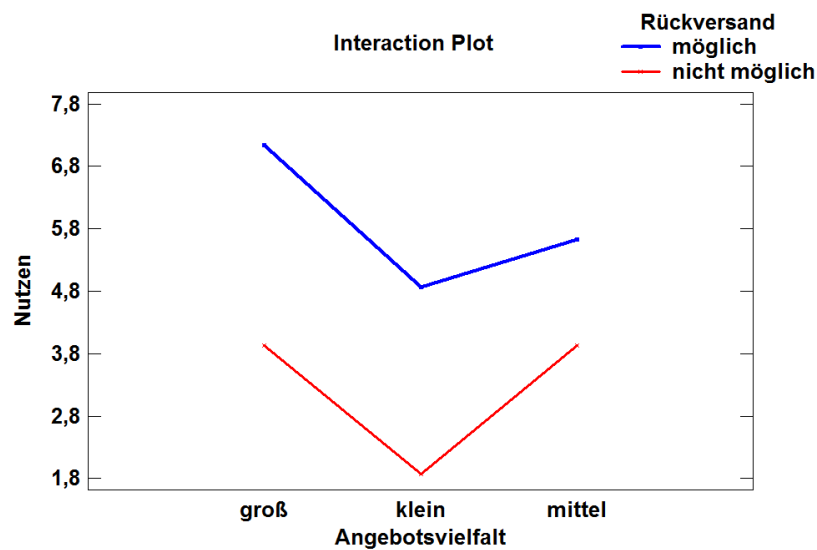
Kolmogorov-Smirnov Test

	<i>Exponential</i>
DPLUS	0,365233
DMINUS	0,629406
DN	0,629406
P-Value	0,0

Wie lautet die Testentscheidung bei $\alpha = 0,05$?

e) Für Online-Shops wurde der Nutzen in Abhängigkeit der Angebotsvielfalt und der Möglichkeit eines Rückversandes untersucht.

Im Modell der Varianzanalyse mit 2 Faktoren und Wechselwirkungen erhält man folgenden Wechselwirkungsplot:



Was vermuten Sie, gibt es in diesem Modell signifikante Wechselwirkungen oder nicht? Begründen Sie kurz!