

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

## Modulprüfung

Prüfungsfach: Statistik II für Betriebswirte  
 Prüfer: PD Dr. Frank Heyde  
 Tag: 3. März 2016

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

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

**1. Aufgabe:** In einer Befragung soll die Lebenssituation von ehemaligen Studenten fünf Jahre nach ihrem Studiumsabschluss erfasst werden. 25 zufällig ausgewählte ehemalige Studenten nahmen an dieser Befragung teil. In einer Frage wurde nach dem jährlichen Bruttoeinkommen gefragt. Es kann davon ausgegangen werden, dass das jährliche Bruttoeinkommen der ehemaligen Studenten näherungsweise normalverteilt ist.

Das durchschnittliche (mittlere) Bruttoeinkommen der 25 Studenten beträgt 42720 €. Die empirische Standardabweichung ermittelt aus den 25 angegebenen Bruttoeinkommen beträgt 6244 €.

Bestimmen Sie eine untere Konfidenzgrenze für das erwartete jährliche Bruttoeinkommen zum Konfidenzniveau von 95%.

**2. Aufgabe:** Die Studenten einer Universitätsstadt wurden über ihre Einschätzung des Freizeitangebots befragt. Das Ergebnis einer Befragung nach Geschlecht getrennt findet man in folgender Tabelle:

Freizeitangebot	Geschlecht	
	Männer	Frauen
eher gut	80	20
mittel	150	50
schlecht	70	30

Testen Sie zum Signifikanzniveau  $\alpha = 0,01$ , ob die Bewertung des Freizeitangebotes vom Geschlecht abhängig ist.

- 3. Aufgabe:** Die Qualität einer Lieferung soll durch einen sequentiellen Stichprobenplan gesichert werden. Eine Lieferung mit 1% Ausschuss ist eine gute Lieferung und eine mit 4% Ausschuss ist eine schlechte Lieferung.

Der Abnehmer (Konsument) und der Hersteller (Produzent) einigen sich auf das Folgende:

Eine gute Lieferung soll mit einer Wahrscheinlichkeit von höchstens 0,05 abgelehnt werden. Eine schlechte Lieferung soll mit einer Wahrscheinlichkeit von höchstens 0,01 angenommen werden.

- a) Bestimmen Sie für diese Werte die Annahme- und Ablehnungsgerade eines sequentiellen Stichprobenplanes.
- b) Nach 199 geprüften Stücken liegen genau 6 Ausschussstücke und 193 fehlerfreie Stücke vor. Welche Testentscheidungen sind möglich, nachdem man das 200. Stück geprüft hat?

**4. Aufgabe:**

Ein Gesundheitsmagazin möchte untersuchen, ob sich der Kaloriengehalt von Fetakäse, hergestellt aus Kuh-, Schafs- bzw. Ziegenmilch, unterscheidet.

Dazu wurde bei verschiedenen Produkten aus unterschiedlichen Supermärkten der Kaloriengehalt pro 100g Fetakäse ermittelt.

Die Ergebnisse für 21 verschiedene Fetakäse liegen in der folgenden Tabelle vor (Messwerte in kcal/100g). Dabei sind die Werte je Milchsorte schon sortiert.

Fetakäse aus Schafsmilch	220	227	231	246	265	268			
Fetakäse aus Kuhmilch	237	242	243	245	251	266	270	288	298
Fetakäse aus Ziegenmilch	145	207	212	218	230	235			

Die Kaloriengehalte sind **nicht** normalverteilt. Testen Sie zum Niveau  $\alpha = 0,01$ , ob die erwarteten Kaloriengehalte der Fetakäse der drei Milchsorten gleich sind oder sich signifikant voneinander unterscheiden.

5. **Aufgabe:** Es soll die Abhängigkeit der Bevölkerung (in Millionen) der USA von der Jahreszahl Jahr= 1790, 1800, ..., 2000, 2010 untersucht werden. Es liegen 23 Stichprobenpaare vor. Ein erstes Modell (Modell A) liefert das folgende Ergebnis:

### Simple Regression - Bevölkerung vs. Jahr

Dependent variable: Bevölkerung

Independent variable: Jahr

Linear model:  $Y = a + b \cdot X$

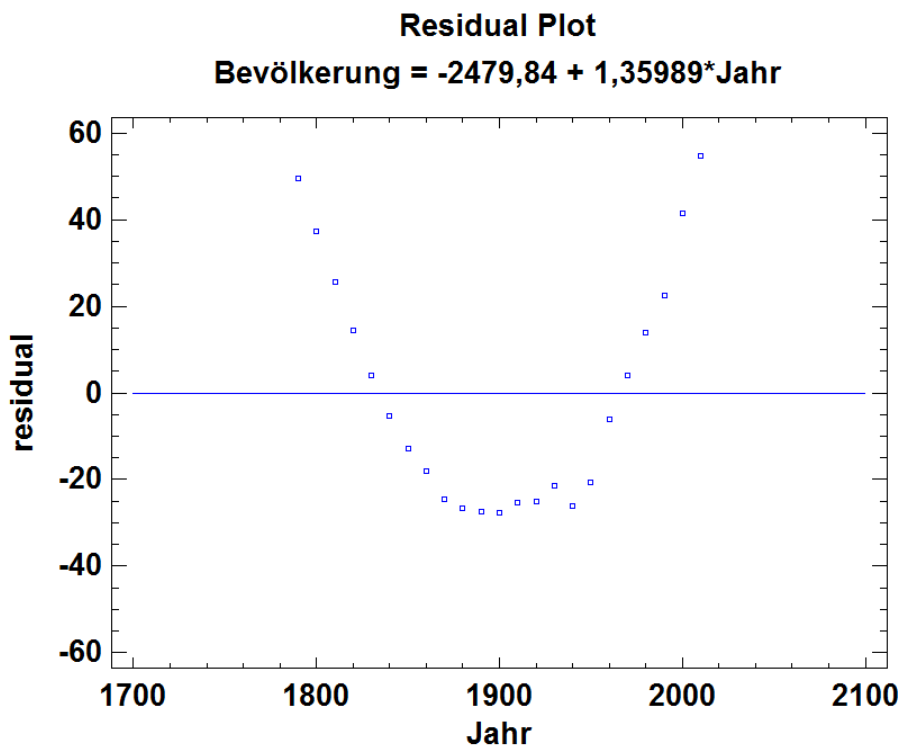
#### Coefficients

	<i>Least Squares</i>	<i>Standard</i>	<i>T</i>	
<i>Parameter</i>	<i>Estimate</i>	<i>Error</i>	<i>Statistic</i>	<i>P-Value</i>
Intercept	-2479,84	166,974	-14,8517	0,0000
Slope	1,35989	0,0878274	15,4837	0,0000

#### Analysis of Variance

<i>Source</i>	<i>Sum of Squares</i>	<i>Df</i>	<i>Mean Square</i>	<i>F-Ratio</i>	<i>P-Value</i>
Model	187150	1	187150	239,74	0,0000
Residual	16393,1	21	780,622		
Total (Corr.)	203543	22			

- Wie lautet die geschätzte Regressionsfunktion?
- Bestimmen Sie das Bestimmtheitsmaß.
- Was können Sie aus der folgenden Grafik ablesen?



Ein zweites Modell (Modell B) liefert als Ergebnis:

### Multiple Regression - Bevölkerung

Dependent variable: Bevölkerung

Independent variables:

Jahr

Jahr<sup>2</sup>

		<i>Standard</i>	<i>T</i>	
<i>Parameter</i>	<i>Estimate</i>	<i>Error</i>	<i>Statistic</i>	<i>P-Value</i>
CONSTANT	21913,8	576,209	38,0309	0,0000
Jahr	-24,3489	0,607054	-40,1099	0,0000
Jahr <sup>2</sup>	0,00676548	0,000159732	42,3552	0,0000

### **Analysis of Variance**

<i>Source</i>	<i>Sum of Squares</i>	<i>Df</i>	<i>Mean Square</i>	<i>F-Ratio</i>	<i>P-Value</i>
Model	203362	2	101681	11251,47	0,0000
Residual	180,743	20	9,03715		
Total (Corr.)	203543	22			

R-squared = 99,9112 percent

Standard Error of Est. = 3,00618

- d) Testen Sie zum Signifikanzniveau  $\alpha = 0,01$ , ob das kleinere Modell A gegenüber dem größeren Modell B genügt, oder ob das größere Modell B zu einer signifikanten Verbesserung führt.

## **6. Aufgabe:**

- a) Wie verändert sich der kritische Bereich für einen Test, wenn man das Signifikanzniveau  $\alpha$  verringert? Begründen Sie Ihre Antwort!
- b) Welchen der folgenden Tests würden Sie verwenden, um das Vorliegen einer diskreten Gleichverteilung zu überprüfen
- $\chi^2$ -Anpassungstest,
  - Kolmogorov-Smirnov-Test,
  - Shapiro-Wilk-Test?

Begründen Sie Ihre Entscheidung!

- c) Welche Kenngröße benutzt man um festzustellen, ob ein zweistufiger Stichprobenplan dem einstufigen vorzuziehen ist?
- d) Gegeben ist ein zweidimensionaler stetiger zufälliger Vektor  $(X, Y)$  mit der Dichtefunktion  $f_{(X,Y)}$ .  
Wie kann man daraus die Dichtefunktion  $f_X$  der Zufallsgröße  $X$  berechnen?
- e) Was versteht man unter den Residuen bei einer Regressionsanalyse?