

Aufgabenserie 12 zur Vorlesung "Statistik für Betriebswirte"

1. Inspektoren der Stadt Spassing haben über ein Jahr nach Schwarzarbeitern gefahndet. Die Untersuchung lieferte folgende Werte für die Anzahl der aufgedeckten Fälle von illegaler Beschäftigung in den einzelnen Quartalen:

	Anzahl der Schwarzarbeiter
1. Quartal	96
2. Quartal	109
3. Quartal	78
4. Quartal	81

Überprüfen Sie die Hypothese, dass sich die aufgedeckten Fälle von Schwarzarbeit gleichmäßig auf die einzelnen Quartale verteilen (Signifikanzniveau 0.05).

2. Onlineshops erfreuen sich großer Beliebtheit bei den Kunden. Bei einer Untersuchung zur Beliebtheit von 4 etwa vergleichbaren Versandhandelunternehmen in der Stadt Kaufrauschershagen ergaben sich folgende Werte für die Anzahl der Kaufvorgänge innerhalb eines Tages

Versandhandelsunternehmen	Anzahl Kaufvorgänge
Rammerzon	166
Otter	182
Qualle	134
Witter	148

a) Man untersuche, ob sich das Kaufinteresse bei den Unternehmen in der betrachteten Stadt gleichmäßig verteilt, d.h. der Anteil der jeweiligen Kaufvorgänge gleich ist. (Signifikanzniveau 0.01).

b) Klären Sie an Hand der Definition und an Hand des Tests in a) die Begriffe: Nullhypothese, Alternativhypothese, Fehler 1. Art, Fehler 2. Art, Signifikanzniveau, Verteilung der Testgröße, asymptotischer Test, kritischer Bereich.

c) Wie wird der kritische Bereich bei einfachen Hypothesen festgelegt? Warum wird der kritische Bereich in a) einseitig an den rechten Rand gelegt?

3. Nach dem Hardy-Weinberg-Gesetz kommen in einer Population die zwei Allele A

und a eines Gens mit den Wahrscheinlichkeiten p und $q = 1 - p$ vor. Bei zufälliger Durchmischung treten die drei Genotypen AA, Aa und aa mit den Wahrscheinlichkeiten p^2 , $2pq$ und q^2 auf. Bei einer vorliegenden Stichprobe zu einem Gen wurde die Häufigkeit der drei Genotypen AA, Aa und aa ermittelt. Dabei ergaben sich folgende Werte:

Genotyp	AA	Aa	aa
Häufigkeit	17	23	6

Man prüfe mit Hilfe eines Tests, ob die drei Genotypen AA, Aa mit den Wahrscheinlichkeiten $\frac{4}{9}$, $\frac{4}{9}$, $\frac{1}{9}$ auftreten, d.h. $p = \frac{2}{3}$ ist (Signifikanzniveau 0.05).

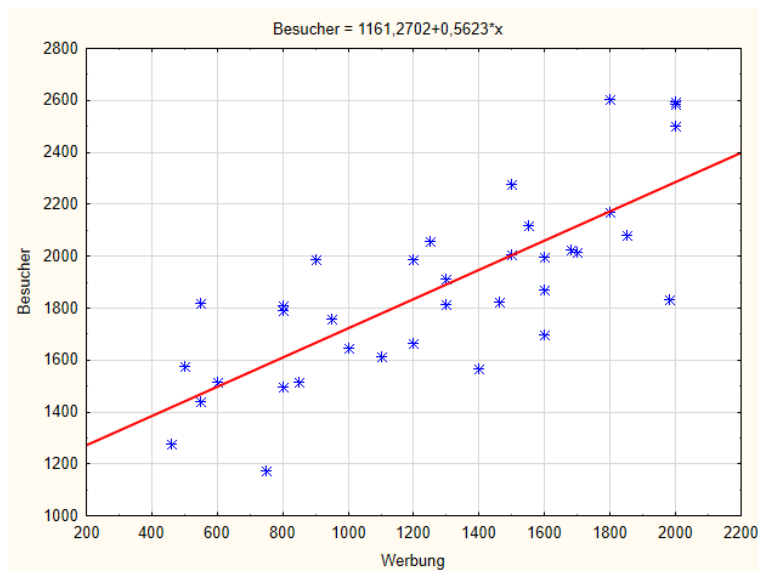
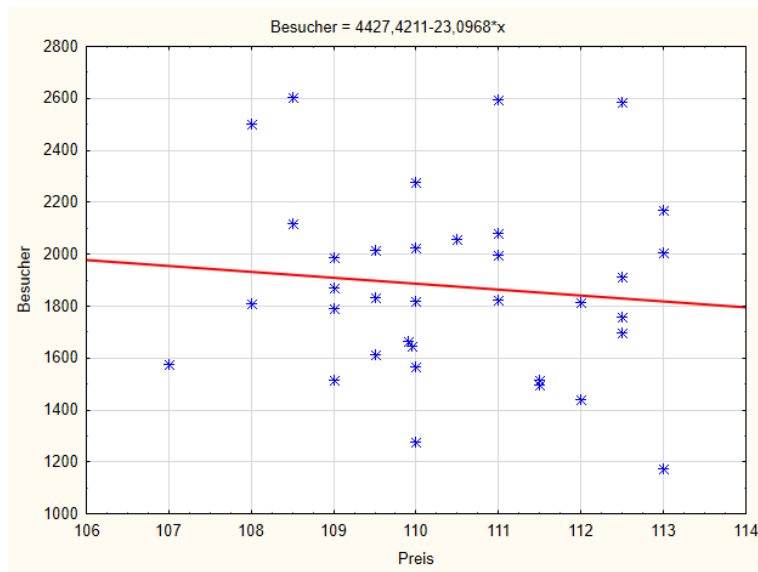
4. In einem Konzern mit 14 Tochterunternehmen werden ähnliche Produkte hergestellt. Betrachtet wird die Abhängigkeit der Gesamtkosten (Variable Y in Tausend Euro) von der Produktionsmenge (Variable x in Tausend Stück). Stellen Sie das lineare Regressionsmodell auf. Berechnen Sie dazu die Schätzer der Parameter, die Konfidenzintervalle und den Korrelationskoeffizienten. Interpretieren Sie die Ergebnisse.

i	x_i	Y_i
1	45	205
2	31	128
3	36	165
4	40	175
5	19	104
6	56	240
7	64	275
8	58	249
9	28	137
10	61	266
11	25	112
12	49	209
13	33	138
14	24	117

Hinweise: Nutzen Sie für die Rechnung Excel (Daten \rightarrow Datenanalyse \rightarrow Regression), ggf. muss vorher das Add-In Analysefunktionen installiert werden. Interessant ist auch die Webseite <https://www.regression-online.de/rechner/> mit einer interaktiven Web-Lösung.

5. Der Konzertveranstalter Bietrokk untersuchte die Abhängigkeit der Besucherzahl vom Preis des Konzerts in Euro und vom Werbeaufwand in Geldeinheiten.

a) Bei der Untersuchung der bivariaten Abhängigkeit der Besucherzahl von den anderen Variablen ergaben sich die folgenden 2 Streubilder:



Interpretieren Sie die beiden Grafiken inklusive der Regressionsfunktion. Ordnen Sie die Korrelationskoeffizienten $R_1 = 0.772$ und $R_2 = -0.106$ den Bildern zu. Wie erkennt man an den Bildern, dass Korrelationen negativ sind und was bedeutet das?

b) Die Regressionsanalyse zur Abhängigkeit der Besucheranzahl vom Werbeaufwand brachte folgendes Ergebnis:

Regression Zusammenf. für abh. Variable: Besucher (konzerte.sta)			
R= ,77241329 R²= ,59662230 korr. R²= ,58475824			
F(1,34)=50,288 p<,00000 Stdf. der Schätzung: 225,87			
N=36	b	t(34)	p-Wert
Konstante	1161,270	10,76860	0,000000
Werbung	0,562	7,09142	0,000000

Geben Sie das Regressionsmodell inklusive der Voraussetzungen und der Zuordnung der Variablen an. Notieren Sie die geschätzte Regressionsfunktion. Wozu kann diese Funktion verwendet werden? Interpretieren Sie den multiplen Korrelationskoeffizienten.

c) Notieren Sie die Signifikanztests zu den Modellparametern für Aufgabenteil b) und die entsprechenden Schlussfolgerungen. (Signifikanzniveau 0.05)

6*. In der Region Teurenbriezen wurden die Preise in Euro für 1l Superbenzin an 27 Tankstellen innerhalb eines Tages erfasst. Dabei ergab sich ein Mittelwert von 1.907 bei einer Stichprobenvarianz von 0.002945. Wir nehmen an, dass die Preise normalverteilt sind.

a) Bestimmen Sie ein Konfidenzintervall für den erwarteten Preis für 1l Superbenzin zum Niveau 0.95. Interpretieren Sie das Intervall. Wie kann man die Aussage des Intervalls in Bezug auf den Erwartungswert verbessern?

b) Ein Fachjournal behauptet, dass der erwartete Preis 1.90 beträgt. Man prüfe diese Aussage mit Hilfe der Daten zum Signifikanzniveau 0.05. Kann man das Resultat auch aus dem Konfidenzintervall a) ableiten? Welche Beziehung besteht zwischen statistischem Test und Konfidenzintervall?

c) Prüfen Sie die Hypothese, dass der erwartete Preis höchstens 1.90 beträgt (Signifikanzniveau 0.05). Was ändert sich im Vergleich zu Aufgabenteil b)?

d) Herr Glukner möchte gern, dass die Nullhypothese in b) angenommen wird. Er macht den Vorschlag, das das Signifikanzniveau so groß zu wählen ist, dass die Nullhypothese tatsächlich angenommen wird. Ist das sinnvoll?