

Aufgabe 1

(10 Punkte)

Eine Kleinstadt des Kantons Bern mit 20'176 Einwohnern weist die in der nachfolgenden Tabelle dargestellte Altersstruktur auf. Um diese Altersstruktur zu ermitteln, wurden Daten von genau 10'000 Einwohnern im November 2015 ausgewertet.

Alter	Relative Häufigkeit
0 - 14	14.60%
15 - 24	10.40%
25 - 34	11.20%
35 - 44	13.80%
45 - 54	10.70%
55 - 64	14.30%
65 - 74	14.50%
75 - 100	10.50%

Markieren Sie die jeweils einzig korrekte Antwort.

- a) Welche Aussage ist korrekt? (2P)
- Die Erhebung sagt aus, dass 1'040 der Einwohner der Kleinstadt zwischen 15 und 24 Jahren alt sind.
 - Die Erhebung besagt, dass ca. 2'260 der Einwohner der Kleinstadt zwischen 25 und 34 Jahre alt sind.
 - 1'070 der Einwohner der Kleinstadt sind zwischen 45 und 54 Jahre alt.
 - Jünger als 24 Jahre sind 2'500 der Einwohner der Kleinstadt.
 - Es gibt 105 75-jährige Einwohner in der Kleinstadt.
- b) Für die Stichprobe gilt: (2P)
- Das Durchschnittsalter ist 25 Jahre.
 - Die Modusklasse ist die Klasse 0 – 14 Jahre.
 - Der Median der Verteilung ist 35.
 - Der Median der Verteilung ist 54.
 - Die Modusklasse ist die Klasse 65 – 74.

c) Welche Aussage ist korrekt? (2P)

- Da es sich um eine grosse Stichprobe handelt, gilt mit Sicherheit $s = \sigma$.
- Bei grossen Stichproben werden sämtliche Ausreisser erfasst.
- Bei der vorliegenden Tabelle handelt es sich um eine deskriptive Statistik.
- Da es sich um eine Stichprobe handelt, wird der Erwartungswert $E(X)$ mit folgender Formel bestimmt:

$$E(X) = \frac{1}{9'999} \cdot \sum_{i=1}^{10'000} x_i$$

- Für die Bestimmung der Standardabweichung s wird folgende Formel verwendet:

$$s = \frac{1}{9'999} \cdot \sum_{i=1}^{10'000} (x_i - \bar{x})^2$$

d) Für obige Stichprobe gilt: (2P)

- Der Interquartilsabstand beträgt ca. 40 Jahre.
- Die Spannweite beträgt 60 Jahre.
- Da es sich bei der angegebenen Verteilung um eine Normalverteilung handelt ist der Mittelwert gleich dem Median.
- Der Interquartilsabstand ist immer grösser als die Spannweite.
- Der Interquartilsabstand beträgt 10.40 %.

e) Markieren Sie die jeweils einzig korrekte Antwort: (2P)

- Wenn das arithmetische Mittel grösser ist als der Median, handelt es sich um eine linksschiefe Verteilung.
- In einem Boxplot wird der Mittelwert häufig als Strich eingezeichnet.
- Bei einer Normalverteilung ist der Mittelwert immer grösser als der Median.
- Das 50. Perzentil entspricht dem Mittelwert.
- Der Interquartilsabstand ist ein Streuungsmass.

Aufgabe 2

(10 Punkte)

Markieren Sie die jeweils einzig korrekte Antwort.

- a) Die Ausfallwahrscheinlichkeit einer Obligation A betrage 10%, die einer Obligation B betrage 15%. Die Wahrscheinlichkeit, dass (mindestens) eine dieser beiden Obligationen ausfällt betrage 22.5%. Wie gross ist die Wahrscheinlichkeit, dass beide Obligationen (gleichzeitig) ausfallen? (2P)
- 25%
 - 1.5%
 - 2.5%
 - 47.5%
 - 22.5%
- b) Für zwei andere Bonds C und D seien folgende Ausfallwahrscheinlichkeiten in Form einer Kreuztabelle gegeben:

		Obligation C	
		fällt aus	fällt nicht aus
Obligation D	fällt aus	0.05	0.15
	fällt nicht aus	0.10	0.70

- Welche (einzige) der folgenden Aussagen ist richtig? (3P)
- Die Ausfallwahrscheinlichkeit für Obligation C beträgt 0.05.
 - Die Ausfallwahrscheinlichkeit für Obligation C beträgt 0.2.
 - Man weiss, dass Obligation D ausgefallen ist. Die Wahrscheinlichkeit, dass Obligation C auch ausfällt, beträgt 0.25.
 - Die beiden Ereignisse, Obligation C fällt aus und Obligation D fällt aus, sind stochastisch unabhängig.
 - Die Wahrscheinlichkeit, dass mindestens eine der beiden Obligationen ausfällt, beträgt 0.2.

- c) Acht Personen sitzen um einen Tisch und essen Fondue. Für jede dieser acht Personen sei die Wahrscheinlichkeit, ihr Brotstück beim Rühren zu verlieren, gleich gross und betrage $p = 0.05$. Weiter seien die Ereignisse, dass zwei verschiedene Personen ihr Brotstück verlieren, voneinander unabhängig.

Wie gross ist die (gerundete) Wahrscheinlichkeit, dass mehr als 1 Person beim Rühren ihr Brotstück verliert? (2P)

- 0.9428
- 0.0352
- 0.95
- 0.05
- 0.0572

- d) Wie viele dreistellige Zahlen bestehen aus 3 *verschiedenen* Ziffern? (3P)

- 1000
- 990
- 729
- 720
- 648

Platz für Notizen (ohne Bewertung)

Aufgabe 3

(10 Punkte)

Eine Maschine produziert Schrauben. Das Gewicht X (gemessen in Gramm) der produzierten Schrauben hat die in Abbildung 1 gezeigte symmetrische Dichte (mit $f(x)$ bezeichnet, dick ausgezogene Linie). Die Varianz von X ist 1.8.

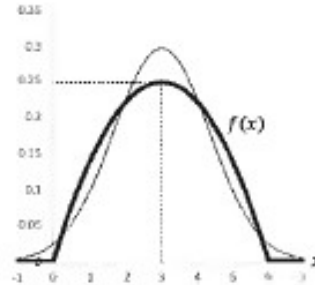
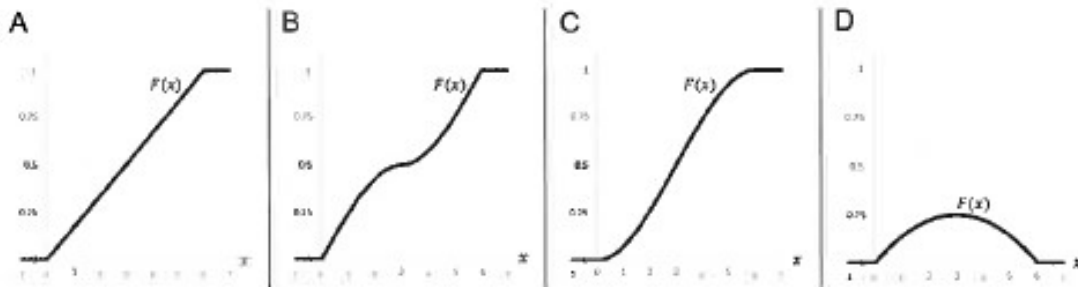


Abbildung 1

Markieren Sie die jeweils einzig korrekte Antwort.

a) Die Verteilungsfunktion $F(x)$ des Schraubengewichts hat die Form (2P)

- A.
- B.
- C.
- D.
- keine davon.



b) Das Gewicht X (2P)

- ist eine diskrete Zufallsvariable.
- hat den Median 3 Gramm.
- ist linksschief verteilt.
- erfüllt $F(3) = 0.25$.
- erfüllt $P(X = 3) = 0.25$.

- c) Die Verteilungsfunktion hat die Eigenschaft $F(5) = \frac{25}{27}$.
 $P(1 \leq X \leq 5)$ ist gleich (gerundet) (2P)
- 0.852.
 - 0.926.
 - 0.842.
 - 0.750.
 - 0.148.
- d) Nehmen Sie an, das Schraubengewicht $X \sim N(\mu, \sigma^2)$ sei normalverteilt wie in Abbildung 1 dargestellt (dünn ausgezogene Linie). Die Parameter μ und σ sind dieselben wie die Parameter der Verteilung mit Dichtefunktion $f(x)$.
Die Wahrscheinlichkeit, dass das Gewicht einer zufällig ausgesuchten Schraube kleiner oder gleich 5 ist, beträgt (gerundet) (2P)
- 0.852.
 - 0.867.
 - 0.997.
 - 0.932.
 - 0.777.
- e) Es werden Schrauben aus einer sehr grossen Produktion zufällig ausgesucht. Wie viele Schrauben müssen mindestens ausgesucht werden, damit die Wahrscheinlichkeit, dass sich der Stichprobenmittelwert des Gewichts mehr als 0.1 Gramm vom Populationsmittelwert unterscheidet, höchstens 5% beträgt? (2P)
- 35
 - 692
 - 978
 - 459
 - 873

Aufgabe 4
(10 Punkte)

Eine Zusammenfassung von Kennzahlen zweier Stichproben der Grösse $n = 20$ zu zwei Variablen x und y ist in Tabelle 4.1 gegeben.

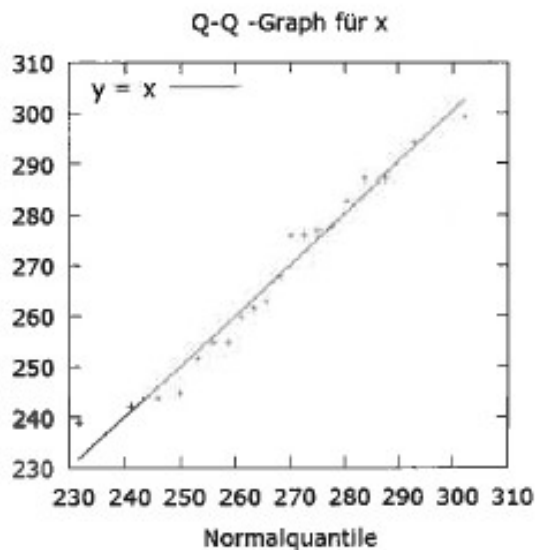
Tabelle 4.1:

	Stichproben- grösse	Arithmetisches Mittel	Standard- abweichung	Variations- koeffizient
x	20	267		0.07116
y	20	291	19	

- a) Welche Aussage zu den Stichprobenkennzahlen trifft am ehesten zu? (2P)
- Der Variationskoeffizient von x ist grösser als derjenige von y .
 - Die Standardabweichung von y ist deutlich grösser als diejenige von x .
 - Der Variationskoeffizient von x ist kleiner als derjenige von y .
 - Die Variationskoeffizienten der beiden Variablen sind identisch.
 - Die Standardabweichung von x ist deutlich grösser als diejenige von y .
- b) Das 95%-Konfidenzintervall für den (unbekannten) Mittelwert der Grundgesamtheit μ_x lautet: (3P)
- [258.108 ; 275.892].
 - [259.671 ; 274.329].
 - [258.673 ; 275.327].
 - [259.654 ; 274.346].
 - [258.138 ; 275.862].
- c) Der Durchschnitt $\bar{x} = 267$ wird zweiseitig bezüglich $\mu = 279$ getestet und verworfen. Die korrespondierende Null- und Alternativhypothese lautet: (2P)
- $H_0: \bar{x} = 267$; $H_1: \bar{x} \neq 267$.
 - $H_0: \bar{x} \neq 279$; $H_1: \bar{x} = 279$.
 - $H_0: \bar{x} \neq 267$; $H_1: \bar{x} = 267$.
 - $H_0: \mu \neq 279$; $H_1: \mu = 279$.
 - $H_0: \mu = 279$; $H_1: \mu \neq 279$.

- d) Die Normalverteilungsannahme für obigen Hypothesentest wird mit einem QQ-Graphen (Abbildung 4.1) überprüft. Welche Aussage trifft am ehesten zu? (3P)

Abbildung 4.1:



Die Konklusion des Hypothesentests ist ...

- gültig, da die Variable x approximativ uniformverteilt ist.
- gültig, da die Variable x approximativ normalverteilt ist.
- nicht gültig, da die Variable x approximativ normalverteilt ist.
- nicht gültig, da die Variable x approximativ uniformverteilt ist.
- gültig, da die Punkte nahe der Regressionsgeraden $y = x$ liegen.

Aufgabe 5
(10 Punkte)

Folgende lineare Einfachregression wird mit Daten einer Stichprobe der Grösse $n = 32$ in Gretl geschätzt.

$$y_i = \beta_0 + \beta_1 \cdot x_i + \varepsilon_i$$

Tabelle 5.1 zeigt das Resultat der Schätzung.

Tabelle 5.1:

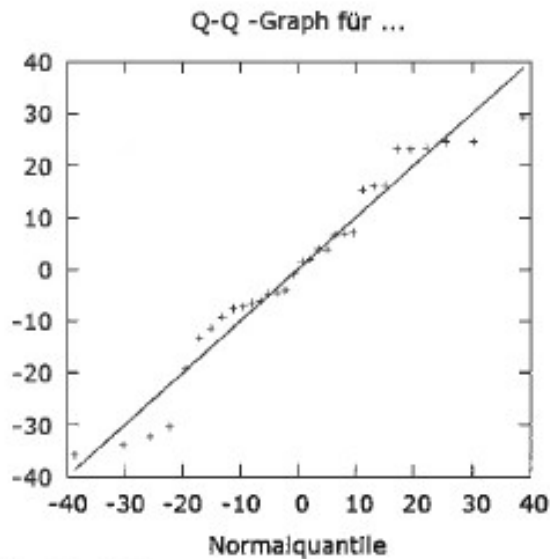
Modell 1: KQ, benutze die Beobachtungen 1-32				
Abhängige Variable: y				
	Koeffizient	Std.-fehler	t-Quotient	p-Wert
const		35.3404	3.124	0.0039 ***
x	0.854003	0.178103	4.795	4.15e-05 ***
Mittel d. abh. Var.		279.1562	Stdabw. d. abh. Var.	23.92880
Summe d. quad. Res.		10048.79	Stdfehler d. Regress.	18.30191
R-Quadrat		0.433878	Korrigiertes R-Quadrat	0.415007

a) Die folgenden Teilaufgaben beziehen sich auf Tabelle 5.1.

- i. Welche Hypothese wird in der Tabelle (implizit) getestet? (2P)
- $H_0: \beta_1 = 0.854003; \quad H_1: \beta_1 \neq 0.854003$
 $H_0: \beta_1 \neq 0; \quad H_1: \beta_1 = 0$
 $H_0: \beta_1 \neq 0.854003; \quad H_1: \beta_1 = 0.854003$
 $H_0: \beta_0 = 0; \quad H_1: \beta_0 \neq 0$
 $H_0: \beta_0 \neq 0; \quad H_1: \beta_0 = 0$
- ii. Wie gross ist die Korrelation zwischen den Variablen x und y ? (2P)
- $r = -0.658942$
 $r = 0.658942$
 $r = 0.433878$
 $r = -0.433878$
 $r \approx 0$
- iii. Welches Intervall ist das (gerundete) 95%-Konfidenz-Intervall für die Steigung in der linearen Einfachregression? (2P)
- [0.49 ; 1.22]
 [0.59 ; 1.12]
 [0.39 ; 1.32]
 [-0.32 ; 0.32]
 [-0.59 ; 1.12]

- b) Für die Überprüfung der Annahmen der linearen Einfachregression wird ein Q-Q-Graph erstellt, der in Abbildung 5.1 zu finden ist. Zu welcher Grösse der obigen linearen Einfachregression wurde der Q-Q-Graph in Abbildung 5.1 erzeugt? (2P)

Abbildung 5.1:



- Variable y
 - Störterme
 - Achsenabschnitt und Steigung
 - Variable x
 - Residuen
- c) Der fehlende Schätzwert für den Achsenabschnitt β_0 in Tabelle 5.1 kann mit den übrigen Resultaten der Tabelle bestimmt werden. Der (gerundete) Schätzwert ist (2P)
- 239
 - 110
 - 178
 - 0
 - 379

ENDE DER PRÜFUNG