

Statistik

Aufgabe 1: [Merkmalstypen]

Welchen Merkmalstypen (nach den Einteilungskriterien "metrisch - ordinal - nominal" beziehungsweise "diskret- stetig") gehören die folgenden Merkmale an?

- a) Länge von Videobändern einer Produktion in cm
- b) Reiseziel von befragten Urlaubsbuchern
- c) Güteklassen von Obst am Markt
- d) Inflationsrate verschiedener Länder
- e) Religion von Befragten
- f) Heizungsart von Mietwohnungen in einer Stadt
- g) Anzahl an Kinobesuchen von Schülern einer Schule in den Ferien
- h) Einstellung von Befragten zur Einführung einer Berufsarmee (Antwortalternativen: sehr skeptisch, eher skeptisch, unentschieden, eher positiv, sehr positiv)

Aufgabe 2: [Merkmalstypen]

Welchen Merkmalstypen gehören die folgenden Merkmale an und wie sind eventuelle Kodierungen der Merkmalsausprägungen vorzunehmen?

- a) Fußballinteresse von Befragten (Merkmalsausprägungen: sehr groß, groß, mittel, schwach, gar keines)
- b) Einkommen von Erwerbstätigen in ganzen EURO
- c) Einstellung der Bevölkerung zu einem EU-Beitritt eines Kandidatenlandes (Merkmalsausprägungen: dafür, teils-teils, dagegen)
- d) Temperatur um 12 Uhr
- e) Anzahl an in einem Monat in einem Konzern produzierten Autos
- f) Militärische Dienstgrade in der deutschen Bundeswehr
- g) Lieblingsschauspieler/in aus einer Liste von 20 vorgeschlagenen Personen
- h) Meereshöhe

Aufgabe 3: [Tabellarische und grafische Darstellung von Häufigkeitsverteilungen]

Am 18. März 2012 wählte die deutsche Bundesversammlung einen Bundespräsidenten. Recherchieren Sie im Internet, wie viele Stimmen der Delegierten auf die einzelnen Kandidaten gefallen sind und wie viele ungültig gewählt beziehungsweise sich enthalten haben. Berechnen Sie danach mit dem Taschenrechner die relativen Häufigkeiten und Prozentzahlen des Stimmverhaltens der **anwesenden** Delegierten.

Aufgabe 4*: [Tabellarische und grafische Darstellung von Häufigkeitsverteilungen]

An N= 100 Kindern derselben Schulstufe wird in einer Schule im Rahmen einer Studie ihre Fähigkeit im Verstehen eines einfachen Textes durch sechs sich auf den Text beziehende Fragen überprüft. Für jedes Kind liegt die Anzahl der falsch beantworteten Fragen vor:

0	2	0	2	0	0	1	2	0	0	2	1	2	1	2	1	1	1	1	6
1	2	0	2	0	0	1	0	0	1	0	0	1	0	0	1	1	0	2	0
1	1	1	1	2	3	2	3	2	0	0	2	1	2	5	1	1	2	4	5
0	1	1	1	2	3	2	3	0	3	5	1	3	3	2	2	1	1	2	1
0	4	1	0	2	3	0	3	1	0	0	0	3	1	0	0	1	2	1	0

Verwenden Sie die im Internet bereitstehende Excel-Lerndatei und stellen Sie darin der Anleitung folgend

- die Häufigkeiten,
- die relativen Häufigkeiten,
- die Prozentzahlen,
- die relativen Summenhäufigkeiten tabellarisch dar.

Aufgabe 5: [Tabellarische und grafische Darstellung von Häufigkeitsverteilungen]

Die erste Nationalratswahl der 2. Republik in Österreich am 25.11.1945 ergab folgende Verteilung der gültigen Stimmen auf die damals zur Wahl stehenden Parteien:

Partei	Stimmen
ÖVP	1.602.227
SPÖ	1.434.898
KPÖ	174.257
Sonstige	5.972

Berechnen Sie mit dem Taschenrechner und zur Kontrolle Ihrer Ergebnisse auch in Excel die relativen Häufigkeiten und die Prozentzahlen der einzelnen Parteien.

Aufgabe 6: [Tabellarische und grafische Darstellung von Häufigkeitsverteilungen]

Auf der Internetseite www.welt-in-zahlen.de/laenderinformation.phtml/ finden sich folgende Angaben über die Häufigkeitsverteilung von Staaten und Territorien bezüglich des Merkmals Kontinent [Stand Januar 2014].

Kontinent	Anzahl der Staaten
Afrika	58
Asien	50
Australien und Ozeanien	24
Europa	51
Nord-, Mittel- und Südamerika	51

Berechnen Sie mit dem Taschenrechner die relativen Häufigkeiten und die Prozentzahlen der Aufteilung aller Staaten und Territorien auf die einzelnen Kontinente.

Aufgabe 7*: [Tabellarische und grafische Darstellung von Häufigkeitsverteilungen]

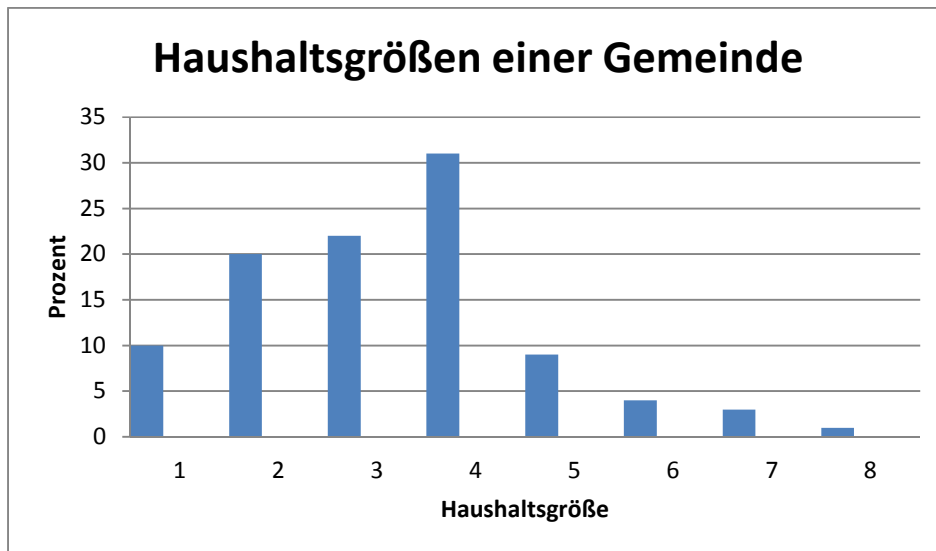
Lösen Sie folgende Aufgaben unter Verwendung der dafür im Internet bereitstehenden Excel-Lerndatei: Stellen Sie die Häufigkeitsverteilung aus Aufgabe 3 in Excel in einem Kreisdiagramm dar. Verwenden Sie die Daten aus Aufgabe 4 und stellen Sie diese Häufigkeitsverteilung in Excel in einem Säulendiagramm dar.

Aufgabe 8*: [Tabellarische und grafische Darstellung von Häufigkeitsverteilungen]

Verwenden Sie die Angaben aus Aufgabe 5 und stellen Sie diese Häufigkeitsverteilung in Excel in einem Säulendiagramm und einem Kreisdiagramm dar.

Aufgabe 9: [Tabellarische und grafische Darstellung von Häufigkeitsverteilungen]

Das unten folgende Säulendiagramm beschreibt die Verteilung des Merkmals Anzahl der Personen in einem Haushalt (= Haushaltsgröße) in einer Gemeinde:



Lesen Sie die relativen Häufigkeiten der einzelnen Merkmalsausprägungen (ungefähr) ab und erstellen Sie damit in Excel eine Tabelle mit den Prozentzahlen und den relativen Summenhäufigkeiten. Zeichnen Sie ferner in Excel das dazu gehörende Kreisdiagramm und versuchen Sie darin die relativen Summenhäufigkeiten abzulesen.

Aufgabe 10*: [Tabellarische und grafische Darstellung von Häufigkeitsverteilungen]

Lösen Sie folgende Aufgabe unter Verwendung der dafür im Internet bereit stehenden Excel-Lerndatei. In dieser Datei befinden sich die durch einen Arzt aufgezeichneten Ergebnisse von Messungen des Cholesterinwertes (in mg/dl) von Personen bei Gesundenuntersuchungen in einem bestimmten Monat. Erstellen Sie eine Tabelle mit den Häufigkeiten, den relativen Häufigkeiten, den Prozentzahlen und den relativen Summenhäufigkeiten des Merkmals Cholesterin, wobei der Arzt die Merkmalsausprägungen in folgende Intervalle einteilen möchte: "unter 220" (normal), „220 bis unter 250" (unter Beobachtung) und „250 und höher" (erhöht). Stellen Sie diese Häufigkeitsverteilung auch grafisch in einem Säulen- und in einem Kreisdiagramm dar (Anleitungen dafür entnehmen Sie den Tabellenblättern der EXCEL-Lerndatei zu Aufgabe 7).

Aufgabe 11: [Tabellarische und grafische Darstellung von Häufigkeitsverteilungen]

Bei einer Befragung von 794 zufällig ausgewählten Personen gaben 241 Befragte auf dem Fragebogen an, dass Sie männlich und mit der EURO-Währung zufrieden sind, 198, dass Sie weiblich und zufrieden sind, 122, dass Sie männlich und unzufrieden sind, 173, dass Sie weiblich und unzufrieden sind. Von den restlichen 60 Personen waren 40 männlich und in der Euro-Frage neutral und 20 weiblich und neutral.

- Erstellen Sie je eine Tabelle mit den Häufigkeiten und mit den relativen Häufigkeiten der gemeinsamen Verteilung auf dem Merkmal Geschlecht und Euro-Einstellung.
- Berechnen Sie die bedingten relativen Häufigkeiten der Verteilungen des Merkmals Euro-Einstellung unter den Männern und unter den Frauen.

Aufgabe 12 [Tabellarische und grafische Darstellung von Häufigkeitsverteilungen]

Vergleichen Sie nun die Ergebnisse der beiden Verteilungen aus Aufgabe 11 b) und beantworten Sie damit die Frage: Sind die Frauen in dieser Erhebung zufriedener mit der Euro-Währung als die Männer oder verhält es sich umgekehrt?

Aufgabe 13* [Tabellarische und grafische Darstellung von Häufigkeitsverteilungen]

Seit Jahrzehnten verschiebt sich in den Industrieländern das Geburtenverhältnis von Mädchen und Jungen. Dass heute mehr Mädchen geboren werden als früher, soll auch auf den Einfluss von Giftstoffen aus der Umwelt auf das männliche Fortpflanzungssystem zurückzuführen sein. In einer wissenschaftlichen Studie an 400 Kindern wurden das Geschlecht der Kinder (Kodierung: männlich = 1; weiblich = 2) und das Rauchverhalten der Eltern in den drei Monaten vor der Zeugung (beide Nichtraucher =1; ein Elternteil Raucher, der andere Nichtraucher =2; beide Raucher =3) erhoben. Verwenden Sie die im Internet bereitstehende Excel-Lerndatei und berechnen Sie darin den Anweisungen folgend die relativen Häufigkeiten

- der gemeinsamen Verteilung der 400 Erhebungseinheiten auf den Merkmalen Geschlecht
- und Rauchverhalten der Eltern,
- der bedingten Verteilungen des Geschlechts der Kinder innerhalb der drei verschiedenen Rauchergruppen !

Aufgabe 14 [Tabellarische und grafische Darstellung von Häufigkeitsverteilungen]

In einer Befragung unter 1.000 Personen wurden in diesem Jahr die Merkmale Geschlecht und Einstellung zum Ankauf neuer Abfangjäger für die Landesverteidigung erhoben.

Ankauf von Abfangjägern				
Geschlecht	dafür	unentschlossen	dagegen	Summe
Weiblich	10	40	250	300
Männlich	190	160	350	700
Summe	200	200	600	1000

Berechnen Sie mit dem Taschenrechner die relativen Häufigkeiten in dieser Tabelle.

Aufgabe 15 [Tabellarische und grafische Darstellung von Häufigkeitsverteilungen]

Berechnen Sie mit dem Taschenrechner die bedingten relativen Häufigkeiten des Merkmals Ankauf von Abfangjägern aus Aufgabe 14 unter den Frauen und auch jene unter den Männern.

Aufgabe 16* [Kennzahlen der Lage]

Berechnen Sie mit dem Taschenrechner mit Hilfe der in Aufgabe 4 erzeugten Tabelle und den dort eingetragenen Häufigkeiten und in Excel unter Verwendung der Anweisungen in der dafür im Internet bereitstehenden Excel-Lerndatei Modus, Median, unteres und oberes Quartil sowie den Mittelwert für die Häufigkeitsverteilung von Aufgabe 4. Lesen Sie ferner Median, unteres und oberes Quartil auch in Excel aus einem Kreisdiagramm ab.

Aufgabe 17* [Kennzahlen der Lage]

Die 28 alphabetisch geordneten EU-Länder des Jahres 2014 besaßen zu Beginn des Jahres 2013 folgende Ausprägungen beim Merkmal Einwohnerzahl in Millionen (*Quelle: Eurostat*):

Staat	Einwohner (in Mio.)	Staat	Einwohner (in Mio.)
Belgien	11,2	Luxemburg	0,5
Bulgarien	7,3	Malta	0,4
Dänemark	5,64	Niederlande	16,8
Deutschland	80,5	Österreich	8,5
Estland	1,3	Polen	38,5
Finnland	5,4	Portugal	10,5
Frankreich	65,6	Rumänien	20,0
Griechenland	11,1	Schweden	9,6
Großbritannien	63,9	Slowakei	5,4
Irland	4,6	Slowenien	2,1
Italien	59,7	Spanien	46,7
Kroatien	4,3	Tschechien	10,5
Lettland	2,0	Ungarn	9,9
Litauen	3,0	Zypern	0,9

- Berechnen Sie in Excel Median und Mittelwert sowie das obere und das untere Quartil für dieses Merkmals.
- Tippen Sie nun in den Daten Ihrer Excel-Datei statt des korrekten Werts für Deutschland die Zahl 805 (Mio.) ein und verfolgen Sie, wie unterschiedlich sich dieser "Tippfehler" auf Ihre Ergebnisse für Median und Mittelwert auswirkt.
- Ausgehend von a) verändern Sie den Wert für Luxemburg auf 50 Mio. Verfolgen Sie die Auswirkungen diese Tippfehlers.
- Erstellen Sie ferner für dieses Merkmal einen Boxplot, indem Sie folgendermaßen vorgehen: Verwenden Sie die Anweisungen in der Excel-Lerndatei, um den Boxplot auf der Internetseite

http://www.wessa.net/rwasp_notchedbox1.wasp zu zeichnen. Versuchen Sie in diesem Boxplot den Median, das untere und obere Quartil und die Whiskers ungefähr abzulesen.

Aufgabe 18 [Kennzahlen der Lage]

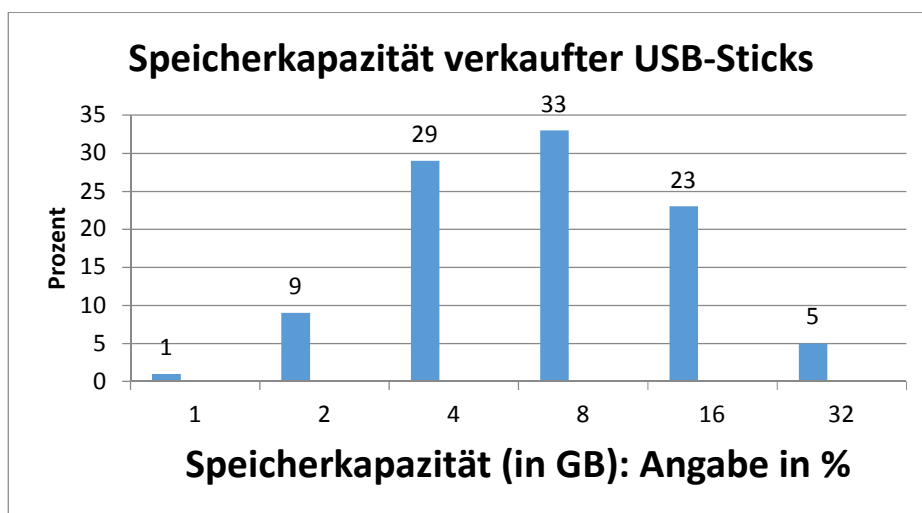
Bei einer Notierte einer der Teilnehmer die täglich am Tachometer angezeigte Kilometerleistung:

66	85	55	32	73	55	42
30	102	48	53	75	60	64

- Berechnen Sie Mittelwert und Median dieses Merkmals.
- Berechnen Sie $Q_{0,25}$ und $Q_{0,75}$ für dieses Merkmal.

Aufgabe 19 [Kennzahlen der Lage]

Das nachfolgende Säulendiagramm beschreibt die relative Häufigkeitsverteilung des Merkmals *Speicherkapazität von verkauften USB-Sticks* (weltweit im Jahr 2009; aus: *CHIP Test & Kauf*, Juni 2010, S. 8):



Berechnen Sie Median, 20%-Perzentil, Mittelwert und Modus für dieses Merkmal.

Aufgabe 20* [Kennzahlen der Lage]

Die Aktie eines Unternehmens wuchs innerhalb von fünf Jahren im 1. Jahr um 1,5 Prozent, im 2. um 28,2 Prozent, im 3. um 80,7 Prozent, im 4. um 14,5 Prozent und fiel im 5. Jahr um 0,7 Prozent. Berechnen Sie

- die fünf Wachstumsfaktoren und damit ausgehend von einem Basiskurs von 100 Euro vor Beginn dieser Zeitspanne den Aktienkurs am Ende der fünf Jahre unter Berücksichtigung des jährlichen Wachstums,

- b) mit den fünf Wachstumsfaktoren das durchschnittliche jährliche prozentuelle Kurswachstum mit dem Taschenrechner und überprüfen Sie Ihr Ergebnis in Excel mit Hilfe der im Internet bereitstehenden Excel-Lerndatei.
- c) Wie hoch hätte demnach der Prozentsatz eines festverzinslichen Wertpapiers sein müssen, um einen höheren Gewinn als die Aktie zu erzielen?

Aufgabe 21 [Kennzahlen der Lage]

Der Preisindex für die Lebenshaltung lag vor genau drei Jahren bei einem Wert von 204. Nach diesen drei Jahren beträgt er 262. Wie groß war

- a) der gesamte prozentuelle Zuwachs in diesen drei Jahren?
- b) die durchschnittliche jährliche Inflationsrate?

Aufgabe 22 [Kennzahlen der Streuung]

Berechnen Sie mit dem Taschenrechner Varianz, Standardabweichung und Variationskoeffizient des Merkmals mit folgenden Ausprägungen 66, 85 und 55.

Aufgabe 23* [Kennzahlen der Streuung]

Berechnen Sie für das Merkmal aus Aufgabe 4 unter Verwendung der im Internet bereitstehenden Excel-Lerndatei und der darin befindlichen Anweisungen Varianz, Standardabweichung und Variationskoeffizient.

Aufgabe 24* [Kennzahlen der Streuung]

Berechnen Sie für das Merkmal Einwohner aus Aufgabe 17 in Excel Varianz, Standardabweichung und Variationskoeffizient. Halbieren Sie nun die Einwohnerzahl jedes Landes aus Aufgabe 17 und berechnen Sie in Excel (Excelblatt „Ü17“) neuerlich Varianz, Standardabweichung und Variationskoeffizient. Wie wirkt sich diese "Transformation" auf die Kennzahlen aus?

Aufgabe 25* [Konzentration]

Erstellen Sie unter Verwendung der im Internet bereitstehenden Excel-Lerndatei bei Einkommen von fünf Personen in der Höhe von 11.000, 13.000, 14.000, 11.000 und nochmals 11.000 Euro in Excel die für das Zeichnen der Lorenzkurve der Konzentration der Merkmalssumme auf die einzelnen Erhebungseinheiten nötige Tabelle und zeichnen Sie diese anschließend der Anleitung folgend. Berechnen Sie ferner mit dem Taschenrechner den dazugehörigen normierten Ginikoeffizienten.

Aufgabe 26 [Konzentration]

Laut Statistik Austria (Lohnsteuerstatistik 2005) verdienen die 25 Prozent der am schlechtesten verdienenden österreichischen Arbeitnehmer 3,9 Prozent des Gesamtbruttolohns, die 50 Prozent der mittleren Verdienenden 42,6 Prozent des Gesamtbruttolohns und die 25 Prozent der am besten Verdienenden insgesamt 53,5 Prozent des Gesamten. Skizzieren Sie die Lorenzkurve der Konzentration des Gesamtbruttolohns auf die drei Gruppen von Arbeitnehmer und berechnen Sie den normierten Ginikoeffizienten. Die Anzahl der österreichischen Arbeitnehmer beträgt 4 Mio.

Aufgabe 27 [statistischer Zusammenhang]

Gemessen wurde der Zusammenhang zwischen Prüfungserfolg (Merkmalsausprägungen: bestanden, nicht bestanden) und selbstständiger Lernfrequenz (regelmäßig, nicht regelmäßig) bei 200 Studierenden:

Lernfrequenz		
Erfolg	regelmäßig	Nicht regelmäßig
Bestanden	152	8
Nicht bestanden	8	32

Berechnen Sie mit dem Taschenrechner eine geeignete Kennzahl zur Messung des statistischen Zusammenhanges zwischen diesen beiden Merkmalen.

Aufgabe 28* [statistischer Zusammenhang]

Verwenden Sie die Angaben aus Aufgabe 11 und berechnen Sie mit Hilfe der im Internet bereitgestellten Excel-Lerndatei die Tabelle der bei Fehlen eines statistischen Zusammenhangs erwarteten relativen Häufigkeiten. Darauf basierend berechnen Sie eine geeignete Kennzahl, um festzustellen, wie stark der statistische Zusammenhang zwischen den beiden Merkmalen in dieser Grundgesamtheit ist.

Aufgabe 29* [statistischer Zusammenhang]

Fortsetzung von Aufgabe 13: In einer Studie dänischer und japanischer Wissenschaftler an 400 Kindern wurden das Geschlecht der Kinder und das Rauchverhalten der Eltern in den drei Monaten vor der Zeugung erhoben. Berechnen Sie in Excel die Stärke des statistischen Zusammenhanges zwischen den beiden Merkmalen.

Aufgabe 30 [statistischer Zusammenhang]

"Sie werden als Erste entlassen, aber offenbar auch als Erste wieder eingestellt: Die Zahl der Zeitarbeitnehmer in Deutschland steigt nach einem Einbruch Anfang des Jahres 2009 kontinuierlich an.... Die Kunden der Zeitarbeitsfirmen nutzen die Möglichkeit, Arbeiter zu beschäftigen, von denen sie sich auch schnell wieder trennen können" (*DIE ZEIT*, Nr.51, 10. Dezember 2009, S.35):

Anfang des Quartals	Arbeitslose (in Mio.)	Zeitarbeitnehmer (in T.)
I	3,49	590
II	3,59	520
III	3,46	550
IV	3,23	600

Skizzieren Sie per Hand das Streudiagramm dieser Verteilung und berechnen Sie mit dem Taschenrechner den Korrelationskoeffizienten der beiden Merkmale.

Aufgabe 31* [statistischer Zusammenhang]

Bei der Erhebung des statistischen Zusammenhangs zwischen der Anzahl an vorbereiteten Hausübungsbeispielen in einem Mathematikurs (x : 0- 50 Beispiele) und den Punkten bei der Abschlussklausur (y : 0 - 80 Punkte) ergab sich für 155 Prüflinge eine Datenreihe, die im Internet als Excel-Lerndatei abgespeichert ist. Folgen Sie dort den Anweisungen und

- stellen Sie die Daten in Excel in einem Streudiagramm dar,
- berechnen Sie in Excel zuerst die beiden Mittelwerte, dann die beiden Standardabweichungen und die Kovarianz und zuletzt den Korrelationskoeffizienten zur Messung des statistischen Zusammenhangs zwischen den beiden Merkmalen.

Aufgabe 32 [statistischer Zusammenhang]

Verwenden Sie Ihre Ergebnisse von Aufgabe 31 b) und

- berechnen Sie mit dem Taschenrechner a) die Gleichung der Regressionsgeraden,
- mit dieser Regressionsgeraden eine Schätzung für die Punktezahl bei der Klausur bei $x = 40$ Hausübungsbeispielen,
- das Bestimmtheitsmaß zur Messung der Güte dieses Regressionsmodells.

Die zusammengefassten Ergebnisse aus Aufgabe 31 sind:

$$\bar{x} = 32,8; \bar{y} = 48,6; s_x = 10,103; s_y = 17,912; s_{xy} = 129,46; r = 0,716$$

Aufgabe 33 [statistischer Zusammenhang]

Berechnen Sie mit den Daten aus Aufgabe 30 mit dem Taschenrechner

- die Gleichung der Regressionsgeraden,
- die Schätzung für die Zahl der Zeitarbeitnehmer bei einer Arbeitslosenzahl von 3,5 Millionen,
- das Bestimmtheitsmaß dieses Regressionsmodells.

Die Daten aus Aufgabe 30 zusammengefasst sind:

$$\bar{x} = 3,44; \bar{y} = 565; s_x = 0,132; s_y = 32,016; s_{xy} = -3,288; r = -0,779$$

Aufgabe 34 [statistischer Zusammenhang]

Bei 4 Schülern (A-D) wurden die Mathematiknoten und die Ränge bei einem Sportwettbewerb gemessen:

Schüler	Mathematiknote	Rang im Sport
A	4	1
B	1	2
C	4	4
D	5	3

Berechnen Sie in Excel den Rangkorrelationskoeffizienten. Dazu müssen Sie zuerst auch aus den Mathematiknoten der 4 Schüler die *Ränge* in Mathematik bestimmen.

Aufgabe 35* [statistischer Zusammenhang]

Bei 4 Studierenden (A-D) wurden ihre Beurteilung der Qualität einer Lehrveranstaltung (nach Schulnoten) und ihre Klausurnoten in dieser Lehrveranstaltung beobachtet:

Studierende	Lehrveranstaltung	Klausur
A	1	1
B	2	2
C	2	1
D	4	4

Berechnen Sie in Excel den Rangkorrelationskoeffizienten. Dazu müssen Sie zuerst auch aus beiden Benotungen der 4 Studierenden die *Ränge* bestimmen.

Aufgabe 36 [Wahrscheinlichkeitsrechnung: Grundbegriffe]

Bei einem Rouletteversuch wird zufällig eine der ganzen Zahlen von 0 bis 36 realisiert. Wie groß sind die Wahrscheinlichkeiten dafür, dass die Kugel

- a) auf die Zahl 22
- b) auf eine gerade Zahl
- c) eine Zahl zwischen 1 und 12 fällt (jeweils einschließlich),
- d) eine Zahl zwischen 1 und 12 (jeweils einschließlich), gefallen ist, wenn man weiß, dass eine Zahl zwischen 1 und 18 gekommen ist (jeweils einschließlich)?

Aufgabe 37 [Wahrscheinlichkeitsrechnung: Grundbegriffe]

Wie viele gleich wahrscheinliche Gruppen zu 5 Personen lassen sich aus 20 Personen (A-T) ziehen und wie wahrscheinlich ist es, dass

- a) die Gruppe {A, B, C, D, E}
- b) die Gruppe {A, C, F, H, M} gezogen wird,
- c) die Person A in der 5-Personen-Gruppe vorkommt?

Aufgabe 38 [Wahrscheinlichkeitsrechnung: Anwendung]

In einem Vorrat von neun Glühlampen befinden sich vier mit einer Leistung von 40 Watt und fünf mit einer Leistung von 60 Watt. Sie wählen zufällig drei Stück aus. Berechnen Sie mit dem Taschenrechner, wie wahrscheinlich es ist, dass sich darunter

- a) keine,
- b) eine,
- c) mehr als eine

40-Watt-Glühlampe befindet.

Aufgabe 39 [Wahrscheinlichkeitsrechnung: Anwendung]

Alle vier Jahre werden "seit Generationen" Abziehbilder für das Fußballer-WM-Sammelalbum der Firma Panini gesammelt. Nehmen Sie an, dass jedes Paket mit fünf Abziehbilder bei der Produktion im Jahr 2010 völlig zufällig mit fünf verschiedenen aus den insgesamt 640 Fotos gefüllt wurde. Wie wahrscheinlich war es dann, dass man durch Kauf eines Paketes sein Album vervollständigt hat, wenn nur noch 5 Abziehbilder gefehlt haben?

Aufgabe 40 [Wahrscheinlichkeitsrechnung: Anwendung]

Aus einer Gruppe von 20 Personen mit acht Frauen und zwölf Männern werden fünf Personen zufällig ausgewählt. Wie wahrscheinlich ist es, dass in einer so ausgewählten Gruppe mehr Frauen als Männer sind?

Aufgabe 41 [Wahrscheinlichkeitsrechnung: Anwendung]

Beim Elfmeterschießen treten im Fußball 5 Schützen je Mannschaft an. Angenommen die Trefferwahrscheinlichkeit aller Schützen einer Mannschaft ist jeweils 70%: Wie wahrscheinlich ist es, dass vier Schützen einer Mannschaft treffen?

Aufgabe 42 [Wahrscheinlichkeitsrechnung: Anwendung]

Erfahrungsgemäß bedarf ein neuer "Tablet - PC" einer bestimmten Marke innerhalb der ersten beiden Jahre mit einer Wahrscheinlichkeit von 0,2 einer Reparatur. Wie wahrscheinlich ist es, dass von 5 lagernden Geräten

- a) keines
- b) eines
- c) mindestens eines

in den ersten beiden Jahren nach dem Verkauf repariert werden muss?

Aufgabe 43 [Wahrscheinlichkeitsrechnung: Anwendung]

In einer Bevölkerung vom Umfang $N=2$ Millionen besitzen 40% eine bestimmte Eigenschaft. Es werden 40 Erhebungseinheiten zufällig (und natürlich ohne Zurücklegen) gezogen. Wie wahrscheinlich ist es, dass 16 davon diese Eigenschaft besitzen, dass also in der Stichprobe derselbe Anteil dieser Eigenschaft vorkommt, wie in der Grundgesamtheit? Sollten Sie Probleme mit den großen Zahlen bei der Hypergeometrischen Verteilung bekommen, versuchen Sie es mit der Binomialverteilung. Warum ist dies hier erlaubt?

Aufgabe 44 [Normalverteilung]

Ein Merkmal sei standardnormalverteilt. Wie wahrscheinlich ist es, dass ein Messwert von

- a) höchstens 1,65,
- b) höchstens 1,96,
- c) mindestens 1,65,
- d) größer als - 1,65,
- e) kleiner als -1,96 auftritt?

Aufgabe 45 [Normalverteilung]

Die Länge von Schrauben einer Produktion ist normalverteilt mit Erwartungswert $\mu = 6$ cm und theoretische Varianz $\sigma^2 = 0,01$ (cm²). Wie wahrscheinlich ist es, dass eine Schraube

- a) höchstens 6,15 cm,
- b) höchstens 6,196 cm lang ist?
- c) mindestens 5,85 cm,
- d) höchstens 5,83 cm lang,
- e) mindestens 5,9 cm und höchstens 6,1 cm lang ist?

Aufgabe 46 [Normalverteilung]

Fortsetzung von Aufgabe 45: Berechnen Sie nun jene Länge der Schrauben, die mit einer Wahrscheinlichkeit von 0,95 unterschritten wird.

Aufgabe 47 [Normalverteilung]

Fortsetzung von Aufgabe 45 bis Aufgabe 46: Berechnen Sie das symmetrische Intervall, in welchem die Länge der Schrauben mit einer Wahrscheinlichkeit von 0,95 liegt.

Aufgabe 48 [Normalverteilung]

Die Länge von Drehbleistiftminen (in cm) verteilt sich annähernd normal mit Erwartungswert $\mu = 5$ und theoretische Varianz $\sigma^2 = 0,25$.

Wie wahrscheinlich ist es dann, dass eine solche Mine

- a) kürzer als 5,2 cm,
- b) länger als 5,98 cm ist?

Aufgabe 49 [Normalverteilung]

Fortsetzung von Aufgabe 48: Wie wahrscheinlich ist es, dass eine solche Mine zwischen 4,75 und 5,5 cm lang ist?

Aufgabe 50 [Normalverteilung]

Die Psychologen Stanford, Binet und Wechsler haben festgestellt, dass die Intelligenz – definiert als geistige Leistungsfähigkeit bei der Lösung von Testaufgaben - unter der Bevölkerung normalverteilt ist. Für Mittelwert und Standardabweichung des Intelligenzquotienten (= IQ) gilt:

$\mu = 100$ und $\sigma = 15$.

In welchem symmetrischen Intervall liegt der IQ einer zufällig ausgewählten Testperson mit einer Wahrscheinlichkeit von

- a) von 0,95,
- b) von 0,5?

Aufgabe 51* [Die Handlungslogik der schließenden Statistik]

Im Internet finden Sie in der Excel-Datei zum Buch eine Grundgesamtheit mit allen 2.500 an einer Universität inskribierten Studentinnen und eine Anleitung zur Ziehung einer uneingeschränkten Zufallsauswahl. Die einzelnen Erhebungseinheiten sind durch eine Identifikationsnummer charakterisiert (und sie wären darüber kontaktierbar). Sie sollen aus dieser Grundgesamtheit eine Zufallsstichprobe vom Umfang $n = 250$ ziehen, in die jede dieser Studentinnen mit gleicher Wahrscheinlichkeit gelangen soll.

Aufgabe 52* [Die Handlungslogik der schließenden Statistik]

Berechnen Sie in Ihrer in Aufgabe 51 gezogenen Stichprobe vom Umfang $n = 250$ den Punktschätzer für

- a) die relative Häufigkeit der Studentinnen, die noch kinderlos sind,
- b) den Mittelwert der Kinderzahl,
- c) den Korrelationskoeffizienten zur Messung des statistischen Zusammenhangs zwischen den beiden Merkmalen Kinderzahl und Anzahl belegter Kurse und vergleichen Sie die drei Stichprobenergebnisse mit den Parametern. Letztere lassen sich aus der ursprünglichen Grundgesamtheit von 2.500 Studentinnen aus Aufgabe 51 errechnen.

Aufgabe 53* [Die Handlungslogik der schließenden Statistik]

Berechnen Sie in Ihrer in Aufgabe 51 gezogenen Stichprobe den Punktschätzer für die relative Häufigkeit der Studentinnen, die noch kinderlos sind. Verwenden Sie dazu jedoch nur die antwortenden Studentinnen, die Sie daran erkennen, dass bei Ihnen in der Spalte "Responseverhalten" eine "1" steht. Vergleichen Sie die relative Häufigkeit an Kinderlosen unter den Antwortenden dieser Stichprobe mit der relativen Häufigkeit unter allen 2.500 Studentinnen der Grundgesamtheit aus Aufgabe 51 und mit jener unter den 250 Studentinnen aus Aufgabe 52.

Aufgabe 54 [Schätzen einer relativen Häufigkeit]

Bei der Abstimmung über den EU-Beitritt Österreichs stimmten im Jahr 1994 66,6 Prozent der Bevölkerung dafür. In welchem Bereich wäre an diesem Tag mit einer Wahrscheinlichkeit von 0,95 ein diesbezüglicher Anteil in einer zufällig aus dieser Grundgesamtheit gezogenen uneingeschränkten Zufallsauswahl gelegen, wenn man dazu

- a) 100,
- b) 400,
- c) 800

Personen ausgewählt hätte? Was ist hier gegeben: die relative Häufigkeit in der Grundgesamtheit oder jene in der Stichprobe?

Aufgabe 55* [Schätzen einer relativen Häufigkeit]

Eine Woche vor der letzten Wahl wurden zufällig ausgewählte wahlberechtigte Staatsbürgerinnen und Staatsbürger gefragt, welche Partei sie wählen würden. Die kodierten Antworten finden Sie in einer zu diesem Beispiel im Internet bereitstehenden Excel-Datei. Bestimmen Sie die relative Häufigkeit der Partei A in der Stichprobe mit Hilfe der dafür vorgesehenen Excel-Funktion (siehe Aufgabe 4) und berechnen Sie davon ausgehend mit dem Taschenrechner das Konfidenzintervall zur Sicherheit $1 - \alpha = 0,95$ für die relative Häufigkeit an Wählern dieser Partei in der Gesamtbevölkerung.

Aufgabe 56 [Schätzen einer relativen Häufigkeit]

Ein Meinungsforschungsinstitut veröffentlichte das Umfrageergebnis einer Zufallsstichprobe von $n = 500$ Personen aus der Grundgesamtheit aller Wahlberechtigten. Es gaben 80 Prozent der Befragten an, gegen den Einkauf von neuen Flugzeugen für die Landesverteidigung zu sein. Was ist hier gegeben: die relative Häufigkeit in der Grundgesamtheit oder jene in der Stichprobe? Berechnen Sie

- a) das Konfidenzintervall zur Sicherheit $1 - \alpha = 0,95$,
- b) das Konfidenzintervall zur Sicherheit $1 - \alpha = 0,9$,
- c) das Konfidenzintervall zur Sicherheit $1 - \alpha = 0,99$

für diese relative Häufigkeit in der Gesamtheit aller Wahlberechtigten.

Aufgabe 57 [Schätzen einer relativen Häufigkeit]

Eine Zufallsstichprobe von $n = 300$ Werkstücken aus einer Produktion ergab einen Ausschussanteil $p = 0,14$. Berechnen Sie daraus das Konfidenzintervall zur Sicherheit $1 - \alpha = 0,95$ für den Ausschussanteil π in der gesamten Produktion.

Aufgabe 58 [Schätzen einer relativen Häufigkeit]

In einer Meinungsumfrage soll festgestellt werden, wie hoch der derzeitige Stimmenanteil einer bestimmten Partei wäre. Wie viele Wahlberechtigte sind in einer uneingeschränkten Zufallsauswahl aus der Grundgesamtheit mindestens zu befragen, wenn das Konfidenzintervall zur Sicherheit $1 - \alpha = 0,95$ eine Schwankungsbreite ϵ von 0,03 besitzen soll und man davon ausgehen kann, dass diese relative Häufigkeit

- a) bei etwa 42 Prozent,
- b) zwischen 5 und 15 Prozent,
- c) bei mindestens 30 Prozent liegen wird?

Aufgabe 59 [Schätzen einer relativen Häufigkeit]

Die relative Häufigkeit des Auftretens einer bestimmten Eigenschaft in der Gesamtbevölkerung soll in einer Stichprobe geschätzt werden. Welcher Stichprobenumfang ist zu wählen, wenn keinerlei Abschätzung des tatsächlichen Anteils existiert und das Konfidenzintervall zur Sicherheit $1 - \alpha = 0,95$ eine von Ihnen festzulegende Schwankungsbreite aufweisen soll?

Aufgabe 60 [Testen von Hypothesen über eine relative Häufigkeit]

Eine politische Partei will feststellen, ob sich die relative Häufigkeit an österreichischen EU-Beitritts-Befürwortern gegenüber dem Ergebnis bei der Volksabstimmung im Jahr 1994 inzwischen verändert hat. Bei der Volksabstimmung hatten 66,6 Prozent zugestimmt.

- a) Formulieren Sie für dieses Problem geeignete statistische Hypothesen.
- b) In einer aktuellen Umfrage unter $n = 800$ zufällig ausgewählten Wahlberechtigten befürworten die EU 43,6%. Entscheiden Sie sich auf Basis dieses Stichprobenergebnisses auf einem Signifikanzniveau $\alpha = 0,05$ für eine der beiden Hypothesen.
- c) Wie würde Ihre Entscheidung bei $p = 0,536$ lauten?
- d) Wie würde Ihre Entscheidung bei $p = 0,636$ lauten?

Aufgabe 61 [Testen von Hypothesen über eine relative Häufigkeit]

Ein TV-Sender möchte feststellen, ob die Einschaltquote (= relative Häufigkeit der zugeschalteten Haushalte von allen Haushalten mit TV-Anschluss) einer TV-Show unter 10 Prozent gefallen ist.

- a) Formulieren Sie für dieses Problem geeignete statistische Hypothesen. In einer Stichprobe unter $n = 1.200$ Haushalten hatten sich 102 Haushalte zugeschaltet.
- b) Entscheiden Sie sich auf Basis dieses Stichprobenergebnisses auf einem Signifikanzniveau $\alpha = 0,05$ für eine der beiden Hypothesen.
- c) Wie würde Ihre Entscheidung bei $p = 0,09$ lauten?
- d) Wie würde Ihre Entscheidung bei $p = 0,12$ lauten?

Aufgabe 62 [Testen von Hypothesen über eine relative Häufigkeit]

In einer Zufallsstichprobe vom Umfang $n = 350$ aus der wahlberechtigten Bevölkerung stellte man Ende des vergangenen Jahres fest, dass 192 der Befragten einem EU-Beitritt eines bestimmten Kandidatenlandes skeptisch gegenüberstanden. Konnte man aus dem Stichprobenergebnis auf dem Signifikanzniveau $\alpha = 0,05$ folgern, dass eine Mehrheit der Gesamtbevölkerung in dieser Frage skeptisch eingestellt war?

Aufgabe 63 [Testen von Hypothesen über eine relative Häufigkeit]

Von Werkstücken, die ein Jahr lang gelagert wurden, sind 40 Prozent unbrauchbar. Nach einer Änderung der Lagerbedingungen wird überprüft, ob sich die relative Häufigkeit an unbrauchbaren Werkstücken verringert hat.

- a) Formulieren Sie für dieses Problem geeignete statistische Hypothesen.
- b) Entscheiden Sie sich auf dem Signifikanzniveau $\alpha = 0,05$ für eine der beiden Hypothesen, wenn unter 100 zufällig ausgewählten Stücken nunmehr 36 Prozent unbrauchbar waren.

Aufgabe 64 [Testen von Hypothesen über eine relative Häufigkeit]

Entscheiden Sie sich bei den Hypothesen in Aufgabe 60 bis Aufgabe 63 auf dem Signifikanzniveau $\alpha = 0,05$, wenn Sie erfahren, dass der zweiseitige p-Wert des Tests

- a) 0,225,
- b) 0,064,
- c) 0,021 betragen hat.

Zur Erinnerung: Einseitige Tests wurden in Aufgabe 61-63 durchgeführt. Nur in Aufgabe 60 wurde ein zweiseitiger Test durchgeführt. Weitere Infos aus diesen Aufgaben benötigen Sie für die Aufgabe 64 nicht.

Aufgabe 65 [Schätzen eines Mittelwertes]

- a) Zur Erprobung der Wirksamkeit eines Schlafmittels wurde bei allen 20.000 Angestellten (große Grundgesamtheit) eines sehr großen pharmazeutischen Unternehmens die Schlafdauer gemessen. Der Mittelwert betrug 6,36 Stunden mit einer Varianz von 1,561. Wenn man 200 Personen dieser Testgruppe per Zufallsstichprobe auswählen würde, in welchem Bereich wäre mit einer Wahrscheinlichkeit von 0,99 der Mittelwert der Schlafdauer für diese Personengruppe?
- b) Die Ergebnisse der Tests mit allen Angestellten sollen nun auf die gesamte deutsche Bevölkerung hochgerechnet werden. Berechnen Sie für die Schlafdauer das Konfidenzintervall zur Sicherheit von $1 - \alpha = 0,99$. Achtung: Die 2.000 Angestellte sind nun als Stichprobe aus der Grundgesamtheit aller Deutschen zu behandeln.
- c) Das Gesundheitsministerium ist unzufrieden mit der großen Schwankungsbreite des Konfidenzintervalls aus Aufgabe b). Wie hoch müsste die Stichprobe insgesamt sein um die Schwankungsbreite des Konfidenzintervalls um den Mittelwert herum auf 0,05 zu reduzieren bei sonst gleichbleibenden Daten (Mittelwert, Varianz, Sicherheitsniveau)

Aufgabe 66 [Schätzen eines Mittelwertes]

Bei der Abfüllung von Mineralwasser in Literflaschen wird der Magnesiumgehalt je Liter gemessen. $n = 116$ Kontrollmessungen ergaben folgende Werte für den Gehalt an Magnesium (in mg/l): $\bar{x} = 25,452$, $s^2 = 0,850$. Berechnen Sie das Konfidenzintervall zur Sicherheit $1 - \alpha = 0,95$ für den mittleren Magnesiumgehalt in der Gesamtproduktion.

Aufgabe 67 [Schätzen eines Mittelwertes]

Eine Zufallsstichprobe vom Umfang $n = 836$ ergab hinsichtlich eines Merkmals x einen Mittelwert $\bar{x} = 22,5$ bei einer Standardabweichung von $s = 3,2$. Bestimmen Sie das Konfidenzintervall zur Sicherheit $1 - \alpha = 0,95$ für den wahren Mittelwert von x .

Aufgabe 68* [Testen von Hypothesen über einen Mittelwert]

Eine Zufallsstichprobe vom Umfang $n = 86$ ergab hinsichtlich eines normalverteilten Merkmals x Messergebnisse, die in der im Internet bereitgestellten Excel-Datei zu finden sind. Testen Sie zum Signifikanzniveau $\alpha = 0,05$ folgende Hypothesen:

- a) $H_0: \mu = 4500$ gegen $H_1: \mu \neq 4500$.
- b) $H_0: \mu \geq 4500$ gegen $H_1: \mu < 4500$.

Aufgabe 69 [Testen von Hypothesen über einen Mittelwert]

Die (stetige) Punktezahl bei einem Aufnahmetest sei annähernd normalverteilt mit $\mu = 75$. Nach Einführung verpflichtender vorbereitender Kurse soll überprüft werden, ob sich der Mittelwert der Punktezahlen erhöht hat.

- a) Formulieren Sie für dieses Problem geeignete statistische Hypothesen.
- b) Testen Sie diese Hypothesen auf einem Signifikanzniveau von $\alpha = 0,05$, wenn in einer Zufallsstichprobe vom Umfang $n = 120$ nach Einführung der Kurse ein Mittelwert von 78,4 Punkten erzielt wird und in dieser Stichprobe eine Standardabweichung von $s = 6$ gemessen wird.

Aufgabe 70 [Testen von Hypothesen über einen Mittelwert]

Die Psychologen Stanford, Binet und Wechsler haben festgestellt, dass der Intelligenzquotient in der Bevölkerung normalverteilt ist. In der Bevölkerung besitzt der IQ einen Mittelwert von 100. Überprüfen Sie auf einem Signifikanzniveau von $\alpha = 0,05$, ob dieser Mittelwert unter Studierenden höher ist. Dazu steht eine Zufallsstichprobe vom Umfang $n = 100$ zur Verfügung, in der ein durchschnittlicher IQ von 108 bei einer Standardabweichung von $s = 19$ gemessen wurde.

Aufgabe 71* [einfache Regressionsanalyse]

Mit den Daten der im Internet bereitgestellten Excel-Datei aus Aufgabe 87 soll auf einem Signifikanzniveau $\alpha = 0,05$ überprüft werden, ob die Punkte in der Hausübung einen Einfluss auf die Punkte in der Statistiklausur haben.

Aufgabe 72* [einfache Regressionsanalyse]

Mit den Daten der im Internet bereitgestellten Excel-Datei aus Aufgabe 87 soll auf einem Signifikanzniveau $\alpha = 0,05$ überprüft werden, ob der Anstieg der Regressionsgerade ($UV = \text{Punkte in Übung}$, $AV = \text{Punkte in der Klausur}$) größer als 0 ist.

Aufgabe 73 [einfache Regressionsanalyse]

Nehmen Sie an, es wurden die Tests von Aufgabe 71 bis Aufgabe 72 durchgeführt. Sie erfahren als Testergebnis von jedem dieser Tests lediglich, dass der zweiseitige p – Wert 0,075 beträgt. Für welche der beiden Hypothesen aus Aufgabe 71 bis Aufgabe 72 sollten Sie sich demnach jeweils entscheiden?

Aufgabe 74 [multivariate Regressionsanalyse]

Interpretieren Sie bitte folgendes Ergebnistableau der Regressionsanalyse mit $y = \beta_1 \cdot x_1 + \beta_2 \cdot x_2 + \beta_3 \cdot x_3 + \beta_4$ mit

y = Arbeitseinkommen je Monat

x_1 = Alter in Jahren

x_2 = Dummy Humankapital (1=hat mindestens Bachelor, 0=hat keinen Bachelor)

x_3 = Größe der Firma in Anzahl Mitarbeiter

wenn ihre zugestandene Irrtumswahrscheinlichkeit $\alpha=0,05$ beträgt und zweiseitige Hypothesentests auf $H_0: \beta_i = 0$ durchgeführt wurden.

Unabhängige Variable UV	Regressionskoeffizient	p-Wert (2 Stellen gerundet)
Alter in Jahren	100	0,07
Dummy Humankapital	500	0,00
Größe der Firma	0,5	0,23
Achsenabschnitt β_4	300	0,00