1. Gegeben seien folgende Daten einer statistischen Erhebung, bereits nach Größe sortiert (Rangliste):

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 10 | 13 | 14 | 14 | 15 | 16 | 17 | 17 | 18 | 18 |
| 18 | 19 | 19 | 20 | 20 | 20 | 20 | 20 | 21 | 21 |
| 21 | 21 | 21 | 21 | 21 | 21 | 22 | 22 | 22 | 22 |
| 22 | 22 | 23 | 23 | 23 | 23 | 24 | 24 | 25 | 25 |
| 25 | 25 | 25 | 26 | 26 | 26 | 27 | 27 | 28 | 30 |

Erstellen Sie eine Tabelle, in der die Merkmalsausprägungen übersichtlicher dargestellt werden, und die auch die Summe der Merkmalsausprägungen (Merkmalssumme) enthält.

1. Stellen Sie die in 1. entwickelte Tabelle grafisch dar.
2. Ihrem Auftraggeber ist sowohl die Tabelle als auch die Zeichnung zu groß und zu unübersichtlich. Fassen Sie deswegen die Merkmalsausprägungen in Klassen zusammen und erstellen Sie die Tabelle neu. Die Klassenbreite soll 2 betragen, die erste Klasse geht von der Merkmalsaus­prägung 9 bis unter 11.
3. Wieso stimmt die Merkmalssumme der Tabelle aus 3. nicht mit der Merkmalssumme der ursprünglichen Tabelle überein?
4. Wie lässt sich die Tabelle aus 3. grafisch darstellen?
5. Die grafische Darstellung der Tabelle aus 3. nennt man ein Histogramm. Lassen Sie dieses von Mathcad oder Excel zeichnen. Ein spezielles Statistikprogramm ist dafür nicht notwendig, jedoch folgender Hinweis:

In Mathcad muss das *Data Analysis Extension Pack* installiert sein. Hier gibt es die Funktion *Histogramm(n,M)*, wobei n ein Vektor der Klassengrenzen ist und M eine Matrix der Merkmals­ausprägungen. In Excel gehört die Histogrammfunktion nicht zur Standardinstallation, sondern zu den Analyse-Funktionen, die nachträglich zu installieren sind. In Excel 2003 wählt man den Befehl *Extras* > *Add-Ins* und aktiviert das Kontrollkästchen *Analyse-Funktionen*. In Excel 2007: *Office-Schaltfläche* > *Excel-Optionen* > *Add-Ins* > Im Listenfeld *Verwalten* den Eintrag *Excel-Add-Ins* wählen > Schaltfläche *Gehe zu…* > Im Fenster *Add-Ins* das Kontrollkästchen *Analyse-Funktionen* aktivieren > *OK*. In Excel 2010: *Datei* > *Optionen* > *Add-Ins* > Im Listenfeld *Verwalten* den Eintrag *Excel-Add-Ins* wählen > Schaltfläche *Gehe zu…* > Im Fenster *Add-Ins* das Kontroll­kästchen *Analyse-Funktionen* aktivieren > *OK*.
6. Es gebe n Ausprägungen des Merkmals x. Die Merkmalsausprägungen werden mit i = 1…n indiziert. Wie hoch ist die Merkmalssumme?
7. Es gebe m unterschiedliche Ausprägungen des Merkmals x, die mit der Häufigkeit h vorkommen. Die unterschiedlichen Merkmalsausprägungen und die Häufigkeiten werden mit j = 1…m indiziert. Wie hoch ist die Merkmalssumme?
8. Die relative Häufigkeit ist die absolute Häufigkeit, geteilt durch die Anzahl *aller* Merkmalsaus­prägungen. Wie groß ist die Summe aller relativen Häufigkeiten?
9. Wie lässt sich der Ausdruck  kürzer darstellen?
10. Gegeben ist folgende Häufigkeitsverteilung:



Ermitteln Sie die relative Häufigkeit und die kumulierte relative Häufigkeit der Merkmalsaus­prägungen. Stellen Sie die relative Häufigkeit in Abhängigkeit von den Merkmalsausprägungen (Häufigkeitsfunktion) sowie die kumulierte relative Häufigkeit in Abhängigkeit von den Merkmals­ausprägungen grafisch dar (Verteilungsfunktion).
11. Gegeben sind folgende Daten:



Ermitteln Sie den kumulierten Anteil der Merkmalsträger an der Gesamtzahl der Merkmalsträger und den kumulierten Anteil der Merkmalsausprägungen an der Merkmalssumme, jeweils beginnend bei null, und stellen sie dies in einem Koordinatensystem dar. Die kumulierten Anteile der Merkmalsträger werden dabei auf der Abszisse abgetragen und die kumulierten Anteile der Merkmalsausprägungen auf der Ordinate (Lorenzkurve).

Zeichnen Sie die Lorenzkurve (oder lassen Sie sie von Excel zeichnen) auch für folgende Daten:



Die Lorenzkurve bildet nun die sogenannte Gleichverteilungsgerade.

Verwenden Sie auch folgende Daten:



Die Lorenzkurve zeigt nun eine extreme Ungleichverteilung.

Wer sich daran stört, dass es „Merkmalsträger“ mit einer Merkmalsausprägung von null geben soll, lasse diese Kurve zeichnen:


12. Die Beitragseinnahmen der 10 größten Versicherungsunternehmen Deutschlands betrugen im Jahre 2009[[1]](#footnote-1):



Ermitteln Sie das 10%-Quantil, das 20%-Quantil, das untere Quartil, das 30%-Quantil, das 40%-Quantil, den Median, das 60%-Quantil, das 70%-Quantil, das obere Quartil, das 80%-Quantil und das 90%-Quantil.
13. Entwerfen Sie eine Grafik, in der die Zusammenhänge aus Aufgabe 13 sinnfällig dargestellt werden.
14. Ein Unternehmen hat folgende Daten ermittelt:



Wie hoch sind die Lohnkosten pro Arbeitsstunde im Jahresdurchschnitt?
15. Beweisen Sie: Wenn das gewogene arithmetische Mittel mithilfe von relativen Häufigkeiten ermittelt wird, muss nicht durch die Summe der Gewichte geteilt werden.
16. Beweisen Sie: Die Summe der Abweichungen vom arithmetischen Mittel ist gleich null.
17. Es sei:

n = Anzahl der Merkmalsausprägungen
i = 1…n Index der Merkmalsausprägungen
xi = Merkmalsausprägung i

Das arithmetische Mittel ist .

Beweisen Sie: Die Funktion  hat ein Minimum, wo .
18. Wie hoch ist für die Daten aus Aufgabe 13 die durchschnittliche Merkmalsausprägung?
19. Wie hoch ist für die Daten aus Aufgabe 13 die mittlere absolute Abweichung?
20. Wie hoch ist für die Daten aus Aufgabe 13 die Varianz und die Standardabweichung?
21. Gegeben sind folgende Daten:





Wie hoch ist die Merkmalssumme?
22. Es gelte folgende Ungleichung:



wobei

n = Anzahl der Merkmalsausprägungen
i = 1...n Index der Merkmalsausprägungen
ai = Absolute Abweichung der Merkmalsausprägung i vom arithmetischen Mittel

Beweisen Sie folgende Behauptung:

Wenn alle Abweichungen einander gleich sind, wenn also gilt ai = a, dann wird aus der Ungleichung eine Gleichung.
23. Beweisen Sie:


24. Beweisen Sie:

Aus dem Steinerschen Verschiebungssatz folgt für c = 0:


25. In einem bestimmten Zeitraum war der Durchschnittskurs von Aktie A 35,00 mit einer Standard­abweichung von 5,00. Bei Aktie B war der Durchschnittskurs im gleichen Zeitraum 9,50 mit einer Standardabweichung von 2,60. Bei welcher Aktie waren die Kursschwankungen größer? Welche Aktie war – ohne Berücksichtigung der Dividenden – volatiler?[[2]](#footnote-2)
26. Beweisen Sie:

Ein Rechteck mit der Breite B und der Höhe H, für welches gilt B + H = C = const, hat dann die größte Fläche, wenn B = H.
27. Stimmen die folgenden Gleichungen?




28. Wie hoch ist der Korrelationskoeffizient, wenn für jede Merkmalsausprägung gilt, dass
?
29. Wie hoch ist der Korrelationskoeffizient, wenn für jede Merkmalsausprägung gilt, dass
?
30. Wie hoch ist der Korrelationskoeffizient, wenn für jede Merkmalsausprägung gilt, dass
? Hierbei ist a eine Konstante.
31. Gegeben ist die Funktion . Wie lauten die ersten und zweiten Ableitungen der Funktion  nach a und b?
32. Gegeben ist folgende Aufstellung des Alters und des Einkommens von 5 Arbeitnehmern[[3]](#footnote-3):



Vermutlich besteht ein Zusammenhang zwischen dem Alter und dem Einkommen. Wie stark ist die Korrelation?
33. Es hat sich bestätigt, dass bei den Daten aus Aufgabe 33 eine Korrelation besteht. Deswegen wird eine Regressionsfunktion ermittelt, welche diesen Zusammenhang beschreibt. Da der Zusammenhang linear zu sein scheint, wird eine lineare Regressionsfunktion vom Typ gewählt. Welche konkreten Werte sind für a und b einzusetzen?

Zur Ermittlung der Parameter a und b sind folgende Formeln erforderlich:





wobei mit i = 1…n und xi und yi für die Merkmalsausprägungen:





Alternativ kann b auch nach der Formel  errechnet werden, wobei






34. Mit welchem Einkommen kann ein Beschäftigter unter Zugrundelegung der Daten aus Aufgabe 33 im Alter von 60, 65 und 67 Jahren rechnen?
35. Gegeben sind folgende Merkmalsausprägungen



Wie hoch ist der Korrelationskoeffizient?
36. Durch welche Funktion wird die abhängige Variable in Aufgabe 36 vollständig erklärt? Wenn diese Funktion als Regressionsfunktion angenommen wird, wie hoch ist dann der Korrelations­koeffizient zwischen den Werten der Regressionsfunktion und y?
37. Gegeben sind folgende Merkmalsausprägungen



Wie hoch ist der Korrelationskoeffizient?
38. Durch welche Funktion wird die abhängige Variable in Aufgabe 38 vollständig erklärt? Wenn diese Funktion als Regressionsfunktion angenommen wird, wie hoch ist dann der Korrelations­koeffizient zwischen den Werten der Regressionsfunktion und y?
39. Eine Regressionsgerade muss die Bedingung  erfüllen. Wieso folgt daraus, dass der Korrelationskoeffizient r nicht größer als |1| sein darf?
40. Die folgende Tabelle zeigt den durchschnittlichen Bruttomonatsverdienst der Arbeitnehmer im Kredit- und Versicherungsgewerbe in den Jahren 2001 bis 2010[[4]](#footnote-4). Durch welche lineare Funktion lässt sich diese Entwicklung annähern?


41. Der Verbraucherpreisindex insgesamt ist von 94,5 im Jahre 2001 auf 108,2 im Jahre 2010 gestiegen (Quelle: Statistisches Bundesamt). Wie hoch war die durchschnittliche jährliche Preissteigerungsrate?
42. Wie hoch ist die durchschnittliche jährliche Wachstumsrate der nominalen Gehälter aus Aufgabe 41 zwischen 2001 und 2010?
43. Wie hoch ist die reale Wachstumsrate der Gehälter aus Aufgabe 41 zwischen 2001 und 2010, wenn man die Preisentwicklung aus Aufgabe 42 zugrunde legt?
44. Ein Gehalt steigt von 3.321 im Jahre 2001 auf 4.258 im Jahre 2010. Im gleichen Zeitraum stieg der Index der Verbraucherpreise von 94,5 auf 108,2. Wie hoch ist das reale Einkommen des Jahres 2010 zu Preisen von 2001?
45. Eine Kapitalanlage verzinst sich mit 1,75 % pro Jahr. Die Inflationsrate beträgt 2,5 %. Wie hoch ist die reale Verzinsung?
46. Sind die Schadenseintritte bei Versicherungsnehmern disjunkte Ereignisse?
47. Wie lässt sich mithilfe der folgenden Zeichnung die Wahrscheinlichkeit dafür ableiten, dass entweder ein beliebiges Ereignis A oder ein anderes beliebiges Ereignis B eintritt?


48. Die Wahrscheinlichkeit für das Eintreten des Ereignisses A sei 0,1 und die Wahrscheinlichkeit für das Eintreten des Ereignisses B sei 0,2. Die Wahrscheinlichkeit für das gemeinsame Eintreten der Ereignisse sei 0. Wie hoch ist die Wahrscheinlichkeit, dass entweder A oder B eintritt?
49. Die Wahrscheinlichkeit für das Eintreten des Ereignisses A sei 0,1 und die Wahrscheinlichkeit für das Eintreten des Ereignisses B sei 0,2. Die Wahrscheinlichkeit für das gemeinsame Eintreten der Ereignisse sei 0,02. Wie hoch ist die Wahrscheinlichkeit, dass A oder B eintritt?
50. Die Wahrscheinlichkeit für das Eintreten des Ereignisses A sei 0,9 und die Wahrscheinlichkeit für das Eintreten des Ereignisses B sei 0,9. Die Wahrscheinlichkeit für das gemeinsame Eintreten der Ereignisse sei 0,81. Wie hoch ist die Wahrscheinlichkeit, dass A oder B eintritt?
51. Die Wahrscheinlichkeit für das Eintreten des Ereignisses A sei 0,9 und die Wahrscheinlichkeit für das Eintreten des Ereignisses B sei 0,9. Die Wahrscheinlichkeit für das gemeinsame Eintreten der Ereignisse sei 0,81. Wie hoch ist die Wahrscheinlichkeit, dass
nur A eintritt
nur B eintritt
nur A oder nur B eintritt
A und B gemeinsam eintreten
entweder nur A oder nur B oder beide Ereignisse gemeinsam eintreten?
52. Sind disjunkte Ereignisse abhängige oder unabhängige Ereignisse?
53. Die Wahrscheinlichkeit, dass ein Versicherungsnehmer einen Schaden erleidet, sei 0,1. Die Wahr­scheinlichkeit, dass ein anderer Versicherungsnehmer einen Schaden erleidet, sei eben­falls 0,1. Ein Kumulrisiko wird nicht in Betracht gezogen, d.h. die Wahrscheinlichkeit für den Schaden­eintritt bei einem Versicherungsnehmer wird als unabhängig vom Schadeneintritt bei einem anderen Versicherungsnehmer angesehen. Wie groß ist die Wahrscheinlichkeit, dass beide Versicherungs­nehmer einen Schaden haben?
54. Die Wahrscheinlichkeit, dass bei einem Versicherungsnehmer der Schadenfall eintritt, sei ws. Wie wahrscheinlich ist es, dass der Schadenfall bei i Versicherungsnehmern derselben Risikogruppe gemeinsam eintritt?
55. Die Wahrscheinlichkeit, dass bei einem Versicherungsnehmer der Schadenfall nicht eintritt, sei 1 – ws. Wie hoch ist die Wahrscheinlichkeit, dass bei n – i Versicherungsnehmern kein Schaden eintritt?
56. Ein Kaufmann schätzt die Gewinnaussichten von zwei verschiedenen Geschäftsmöglichkeiten folgendermaßen ein:



Wie hoch ist der Gewinnerwartungswert und die Standardabweichung beider Geschäfte?
57. Die Wahrscheinlichkeit für den Eintritt des Schadens s bei einem Versicherungsnehmer sei ws. Die Wahrscheinlichkeit, dass der Schaden nicht eintritt, sei 1 – ws. Wie hoch ist der Erwartungs­wert, die Varianz und die Standardabweichung des Schadens?
58. Die Erwartungswerte der Schäden von Versicherungsnehmern können zum Erwartungswert des Gesamtschadens der Versicherungsunternehmung addiert werden. Die Varianzen der Schäden können zur Varianz des Gesamt­schadens addiert werden, wenn die Schadenereignisse unab­hängig voneinander sind. Dies vorausgesetzt, wie hoch ist der Erwartungswert und die Standard­abweichung des Gesamt­schadens einer Versicherungsunternehmung mit n gleichartigen Ver­siche­rungsnehmern wie in Aufgabe 58 beschrieben?
59. Inwiefern ist der Value at Risk für den Statistiker nichts anderes als ein Quantil?
1. Quelle: Institut der deutschen Wirtschaft Köln (Hrsg.), Deutschland in Zahlen - Ausgabe 2011 -, Köln 2011, S.52 [↑](#footnote-ref-1)
2. Quelle für das Beispiel: M. Piazolo, Statistik für Wirtschaftswissenschaftler – Daten sinnvoll aufbereiten, analysieren und interpretieren, Karlsruhe 2007, S. 95 [↑](#footnote-ref-2)
3. Quelle für das Beispiel: A. Quatember, Statistik ohne Angst vor Formeln – Das Studienbuch für Wirtschafts- und Sozial­wissen­schaftler, 2. Aufl. München 2008, S. 67 [↑](#footnote-ref-3)
4. Quelle: Institut der deutschen Wirtschaft Köln (Hrsg.), Deutschland in Zahlen - Ausgabe 2011 -, Köln 2011, S. 56 [↑](#footnote-ref-4)