Übungsbeispiele:

1. **Für einen p-Wert von 0, 03 beim t-Test für unabhängige Stichproben mit den Hypothesen: H0: $\mu 1 = \mu 2$ / HA : $\mu 1 \neq \mu 2$ gilt stets:**
   a) Das Testergebnis ist statistisch signifikant zum Signifikanzniveau $\alpha = 0{,}05$
   b) Das Testergebnis ist klinisch relevant
   c) Das Testergebnis ist statistisch signifikant zum Signifikanzniveau $\alpha = 0{,}01$
   d) Die Nullhypothese H0 kann nicht zum Signifikanzniveau $\alpha = 0{,}05$ verworfen werden
   e) Die Power beträgt genau 3 Prozent

   Antwort: a)

2. **Die 2-Jahres-Heilungsrate für Patienten mit bestimmter Diagnose ist nach einer Therapie 0, 8. Wie groß ist die Wahrscheinlichkeit, dass bei der Gruppe von 5 unabhängigen Patienten mit dieser Diagnose nach einer Therapie kein einziger nach zwei Jahren geheilt ist?**
   a) 0,84
   b) 0,672
   c) 0,328
   d) 0,16
   e) 0,00032

   Antwort: $0.2^5 = 0{,}00032$

3. **Bei einer Fall-Kontroll-Studie:**
   a) werden Patienten unterschiedlichen Therapien zufällig zugeteilt
   b) werden zwei Gruppen von Patienten über einen Zeitraum prospektiv erhoben und verglichen
   c) wird eine Stichprobe von Patienten zu einem bestimmten Zeitpunkt untersucht und die Häufigkeit eines Merkmals erhoben
   d) wird bei Patienten nach Auftreten einer Erkrankung und bei geeignet ausgewählten nicht erkrankten Personen die Risikoexposition in der Vergangenheit erhoben und verglichen
   e) wird der Therapieerfolg bei Fällen mit einer bestimmten Erkrankung nach einem vorgegebenen Zeitintervall nachkontrolliert

   Antwort: d)

4. **Bei Patienten aus Intensivstationen soll untersucht werden, ob es einen Unterschied im Logistic Organ Dysfunction Score zwischen verstorbenen Patienten und Patienten, die überlebt haben, gibt. Um diese Frage zu beantworten, wurde ein ungepaarter t-Test berechnet (Signifikanzniveau ist 0,05, zweiseitig)**

   Die Analyse wurde mit SPSS durchgeführt. Die folgenden 2 Tabellen zeigen das Ergebnis. Der Logistic Organ Dysfunction Score ist mit LOD, die Spitalsmortaliät (überlebt/verstorben) mit SM abgekürzt.

Gruppenstatistiken

	SM	N	Mittelwert	Standardabweichung	Standardfehler des Mittelwertes
LOD	überlebt	3484	2,62	2,516	,043
	verstorben	999	6,69	3,463	,110

Test bei unabhängigen Stichproben

		Levene-Test der Varianzgleichheit		T-Test für die Mittelwertgleichheit						
									95% Konfidenzintervall der Differenz	
		F	Signifikanz	T	df	Sig. (2-seitig)	Mittlere Differenz	Standardfehler der Differenz	Untere	Obere
LOD	Varianzen sind gleich	171,263	,000	-41,142	4481	,000	-4,068	,099	-4,262	-3,874
	Varianzen sind nicht gleich			-34,608	1314,500	,000	-4,068	,118	-4,299	-3,838

Welche der folgenden Aussagen ist <u>falsch</u>?

a) Es gibt einen signifikanten Unterschied im LOD Score zwischen verstorbenen Patienten und Patienten, die überlebt haben

b) Es gibt keinen signifikanten Unterschied im LOD Score zwischen verstorbenen Patienten und Patienten, die überlebt haben

c) Es haben 3484 Patienten überlebt und es sind 999 verstorben

d) Die Prüfgröße für den t-Test (bei gleichen Varianzen) beträgt −41, 142

e) Die Nullhypothese wird verworfen

Antwort: b)

5. **In einer Beobachtungsstudie soll der Zusammenhang der Gewichtszunahme mit der Ernährung untersucht werden. Dazu werden bei 100 Probanden die Gewichtszunahme innerhalb eines Monats sowie die Anzahl der Kalorien, die die Probanden zu sich genommen haben, erhoben. Welche Maßzahl ist geeignet, den Zusammenhang der beiden Variablen Gewichtszunahme und Kalorienanzahl zu beschreiben?**

a) Median

b) Varianz

c) Arithmetisches Mittel

d) Korrelationskoeffizient

e) Standardfehler des Mittelwerts

Antwort: d)

6. **Welche der folgenden ist eine Randomisierung von 20 PatientInnen in festen Zeitblöcken vom Umfang 4 in zwei Behandlungsgruppen A und B?**

a) AAAA BBBB AAAA BBBB AAAA

b) AAAB BBBA AAAB BBBA AABB

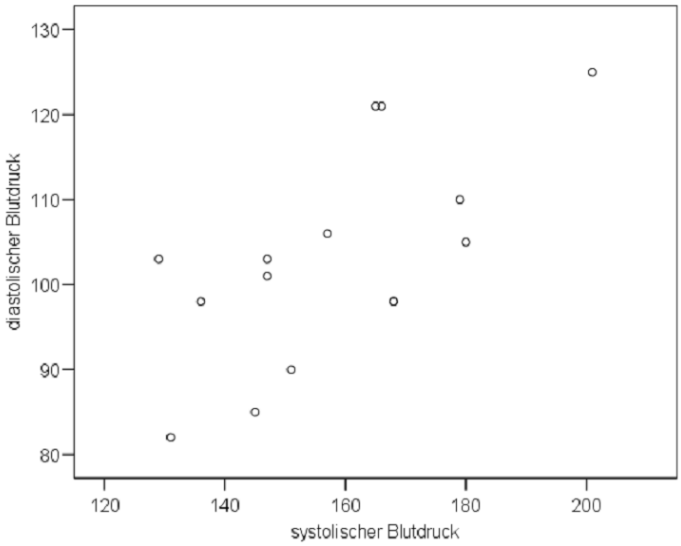
c) AAAA AAAA AABB BBBB BBBB

d) ABBA BBAA ABAB BBAA AABB

e) ABCD BCAD DABC CADB ADCB

Antwort: d)

7. Entscheiden Sie anhand des unteren Streudiagramms, welcher der anschließend angeführten Pearson'schen Korrelationskoeffizienten die Korrelation zwischen dem systolischen und diastolischen Blutdruckwert beschreibt



a) 1, 00
b) −0, 99
c) 0, 99
d) 0, 66
e) −0, 66

Antwort: d)

8. Bei einem positiven Korrelationskoeffizienten nach Pearson gilt stets:

a) die Steigung bei der linearen Regressionsfunktion ist positiv
b) die Steigung bei der linearen Regressionsfunktion ist negativ
c) die Mittelwerte beider Gruppen sind positiv
d) der Achsenabschnitt ist positiv
e) die Steigung liegt zwischen 0 und 1

Antwort: a)

9. Sie möchten den Zusammenhang des Alters mit dem diastolischen Blutdruck untersuchen und haben diese Größen von 100 zufällig ausgewählten Personen erhoben. Welches der folgenden Verfahren ist zur Analyse am besten geeignet?

a) Korrelationskoeffizient nach Pearson
b) t-Test für gepaarte Stichproben
c) t-Test für unabhängige Stichproben
d) $\chi 2$-Test
e) Histogramm

Antwort: a)

10. Das folgende Histogramm zeigt die Verteilung der Körpergröße in cm von 20 Personen:

Welche Aussage kann für dieses Histogramm zutreffen?

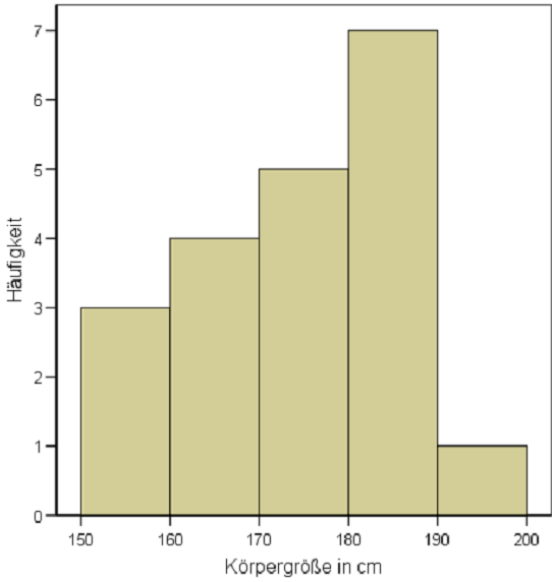
a) Der Median der Daten liegt im Intervall 160-170 cm

b) Genau 50% aller Personen sind größer als 180 cm

c) Vier Personen sind 186 cm und eine Person ist 191 cm groß

d) Eine Person ist 151 cm und drei Personen sind 159 cm groß

e) Die kleinste Person ist nur 149 cm groß



Antwort: c)

11. Bei zwei Korrelationsanalysen ergaben sich folgende Resultate:

Analyse A: Korrelationskoeffizient nach Pearson r = 0.8, p=0.07

Analyse B: Korrelationskoeffizient nach Pearson r = 0.25, p=0.001

Bei welcher der beiden Analysen war die Stichprobe vermutlich größer? Achten Sie auf die richtige Begründung!

a) Bei Analyse A, da ja der Korrelationskoeffizient größer ist als bei Analyse B und der p-Wert für die Begründung irrelevant ist

b) Bei Analyse B, da ja der Korrelationskoeffizient kleiner ist als bei Analyse A und der p-Wert für die Begründung irrelevant ist

c) Bei Analyse A, da der p-Wert größer ist

d) Bei Analyse B, da obwohl geringere Korrelation besteht, der p-Wert kleiner ist

e) Das ist auf Grund ungenügender Angaben nicht eindeutig feststellbar

Antwort: d)

12. In einem Laboratorium sind zwei Geräte für die Bestimmung eines Laborparameters in Verwendung. Das Gerät A liefert 70% der Befunde, das Gerät B 30%. Das Gerät A arbeitet mit einer Fehlerrate von 0.001, das Gerät B mit einer Fehlerrate von 0.005. Wie groß ist die Wahrscheinlichkeit, dass das Laboratorium einen fehlerhaften Befund liefert?

a) 0,006

b) 0,0022

c) 0,000005

d) 0,005

e) 0,001

Antwort: (0,7 x 0,001) + (0,3 x 0,005) = 0,0022

13. Sie haben von zwei Variablen eine Regressionsgerade bestimmt (unabhängige Variable auf der x-Achse, abhängige Variable auf der y-Achse). Was passiert mit Achsenabschnitt und Steigung, wenn Sie jetzt alle Werte der unabhängigen Variablen durch 2 dividieren?

a) Beides bleibt gleich, da die Werte von der Skalierung unabhängig sind

b) Achsenabschnitt und Steigung verdoppeln sich

c) Achsenabschnitt und Steigung halbieren sich

d) Die Steigung verdoppelt sich, der Achsenabschnitt bleibt gleich

e) Die Steigung halbiert sich, der Achsenabschnitt bleibt gleich

Antwort: $y = b_1 x + b_0$ → Wenn x halbiert wird, wird b_1 verdoppelt, b_0 verändert sich nicht (Antwort d)

14. Bei einem normalverteilten Merkmal mit dem Mittelwert μ und der Standardabweichung σ liegen zwischen μ – 2 · σ und μ + 2 · σ

a) ca. 2/3 der Werte
b) ca. 95 % der Werte
c) ca. 99 % der Werte
d) ca. 1/3 der Werte
e) ca. die Hälfte der Werte

Antwort: b)
Zwischen μ - 1 · σ und μ + 1 · σ : 68,27%
Zwischen μ - 2 · σ und μ + 2 · σ : 95,45%

15. Von einer Standardtherapie A ist bekannt, dass sie etwa eine Erfolgsrate von 40% hat. Eine neue Therapie B soll in einer randomisierten Studie mit dieser Therapie A auf Überlegenheit geprüft werden. Die Stichprobenumfänge sollen in beiden Therapiearmen gleich groß sein. Wir gehen davon aus, dass die neue Therapie B die Standardtherapie A in ihrer Erfolgswahrscheinlichkeit um 10%-Punkte übertrifft. Unter diesen Annahmen soll bei dieser Studie für den Vergleich der Erfolgswahrscheinlichkeiten der beiden Therapien mit dem zweiseitigen χ2 - Test für Vierfeldertafeln auf dem Signifikanzniveau α = 0.05 mit einer Wahrscheinlichkeit von 80% (Power) ein signifikanten Ergebnis (Verwerfung der Nullhypothese der Gleichheit der Erfolgswahrscheinlichkeiten) erzielt werden. Wie groß ist der insgesamt dafür benötigte Stichprobenumfang (beide Gruppen zusammen)?

a) 120
b) 240
c) 780
d) 1250
e) 2400

Antwort: c)

16. Ein diagnostischer Test hat einen positiven Vorhersagewert von 0.7 und eine Sensitivität von 0.9. Sie haben zwei Patienten mit positivem Testergebnis. Unter der Annahme, dass die Patienten unabhängig ausgewählt wurden (keine Verwandtschaft), wie hoch ist die Wahrscheinlichkeit, dass keiner der beiden erkrankt ist?

a) 0.09
b) 0.21
c) 0.49
d) 0.81
e) 0.001

Antwort: Wahrscheinlichkeit nicht erkrankt: 30% → 0.3 x 0.3 = 0.09

17. Eine Datentabelle mit 99 Fällen enthält ein Merkmal SCORE mit ganzzahligen Werten 1, 2,. . ., 50. Welche der folgenden Werte können unmöglich der Median sein:

a) 1
b) 16
c) 22,5
d) 50
e) 25

Antwort:

99 Fälle (ungerade) → Median ganze Zahl

Gerade Anzahl an Fälle → Median <u>kann</u> "Komma 5" am Schluss haben (z.B. wenn nach der Teilung durch 2 [da ja bei Median p = 0,5] „32" rauskommt, die nächste Zahl 33 ist, nimmt man 32,5 als Median, wäre der nächste Wert 34, wäre Median 33, deshalb <u>kann</u>!

18. Worauf bezieht sich der Begriff SEM?

a) Englisch für Standardabweichung.
b) Standardabweichung der Mittelwerte bei Ziehung wiederholter Stichproben
c) Summe der quadrierten Abweichungen vom Mittelwert
d) 95% Konfidenzintervall
e) Spezischer statistischer Terminus für ein Lagemaß

Antwort: b)

19. Es wird ein Rangsummentest (U-Test) zwischen zwei Stichproben A und B, jeweils mit dem Stichprobenumfang 4 durchgeführt. Wie groß ist die erwartete Rangsumme in jeder der beiden Stichproben unter der Nullhypothese?

a) 10
b) 26
c) 72
d) 36
e) 18

Antwort:

Minimum wäre 1+2+3+4 = 10

Maximum wäre 5+6+7+8 = 26

$$\frac{(1+2+3+4+5+6+7+8)}{2} = 18$$

Der Wert muss zwischen max. und min. sein, und da bei der Frage erwartet steht werden wir kaum die Extremen Werte nehmen, sondern 18 :)

20. Es werden zwei unabhängige, randomisierte Studien zum Vergleich zwischen den Behandlungen A und B durchgeführt. In beiden Studien wird der Unterschied des Mittelwerts der Verteilungen eines Hauptzielkriteriums in einem zweiseitigen Test zum Signifikanzniveau α = 0.05 durchgeführt

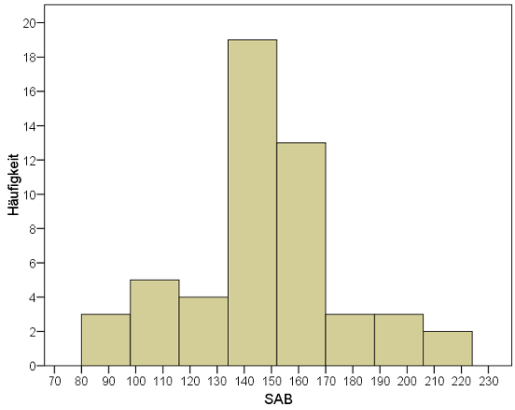
Wie groß ist die Wahrscheinlichkeit, in beiden Studien kein signifikantes Testergebnis (Ablehnung der Nullhypothese, egal in welche Richtung) zu erhalten?

a) 0.05
b) 0.9975
c) 0.9025
d) 0.95

e) kann nicht angegeben werden

Antwort: e) In der Angabe fehlt, welche Hypothese gilt!

21. Im Folgenden ist ein Histogramm für SAB für 18- bis 30-jährige Patienten der Station C gegeben.
Der Stichprobenumfang lautet 55, die erste Klasse beginnt bei 80, die Klassenbreite beträgt 18

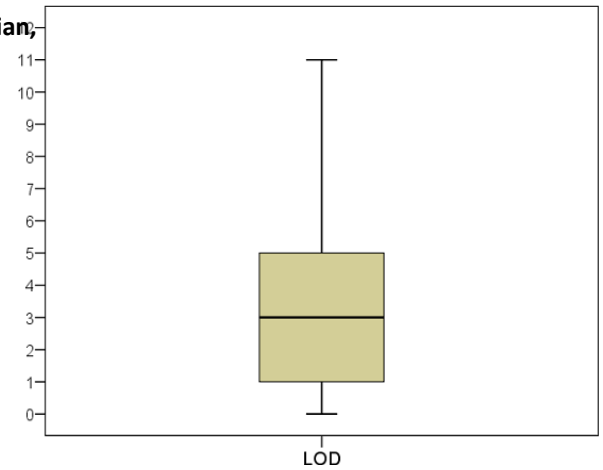


a. In welcher Klasse wird sich der Median befinden?
b. In welcher Klasse befindet sich das 1. Quartil?
c. In welcher Klasse befindet sich das 3. Quartil?
d. Welche Werte sind für das Minimum möglich?
e. Könnte die Spannweite bei diesem Histogramm 140 sein?

Antwort:

a. Median: n . p = 52 (Summe der höhen der jeweiligen Klassen; 3+5+19+...) x 0.5 = 26... Jetzt vergleichen in welcher Klasse sich die 26-27e Person befindet
   Klasse 80 - 98.....3
   Klasse 98 - 116.....3+5=8
   Klasse 116 - 134.....8+4=12
   Klasse 134 - 152.....12+19=31 =====> Der Median befindet sich in dieser Klasse
b. 1. Quartil: n.p = 52 x 0.25 = 13 ======> 1. Quartil befindet sich auch in der Klasse 134 – 152
c. 3. Quartil: n.p = 52 x 0.75 = 39
   Klasse 152 - 170.....31+13=44 =======> 3.Quartil befindet sich in dieser Klasse
d. [80-98)
e. Meiner Meinung nach ja... Auch wenn man annimmt das Spannbreite 224 - 80 = 144 beträgt, kann es ja auch sein das man die Klassenbreite 18 gewählt hat, und das max. z.B. nicht 224 sondern nur 240 wäre, und daher 240 - 80 = 140 (kommt meiner Meinung nach an wie man die Klassenbreite ausgerechnet hat)

22. Bestimmen Sie das Minimum, das 1. Quartil, den Median, das 3. Quartil und das Maximum. Berechnen Sie weiters Spannweite und den Interquartilabstand vom Boxplot LOD für 18- bis 30-jährige Patienten der Station C.



Antwort:

- 1es Quartil: 1
- Median: 3
- 3es Quartil: 5
- Maximum (sofern Ausreißer eingezeichnet): 11
- Spannweite: 11
- Interquartilabstand: 4

23. Welche Eigenschaften hat das folgende Merkmal? (Mehrfachantwort möglich)
Harnvolumen pro Tag in ml

a) dichotom
b) metrisch
c) stetig

d) nominal
e) qualitativ

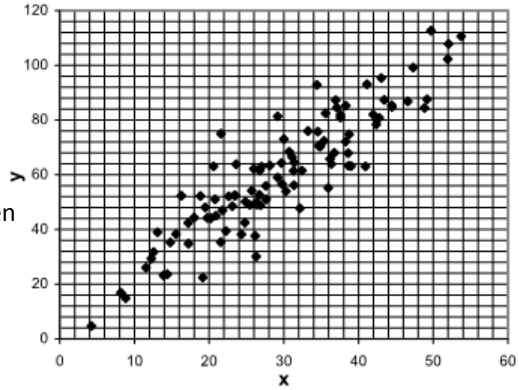
Antwort: b) & c)

24. Für welche der folgenden Merkmale ist das arithmetische Mittel ein geeignetes Lagemaß? (Mehrfachantwort möglich)

a) Güteklassen von Obst
b) Gewicht in kg
c) Geschlecht
d) Wartezeit der Patienten in der Ambulanz (Minuten)
e) Beurteilung der Patientenzufriedenheit in den Klassen: sehr unzufrieden, unzufrieden, zufrieden, sehr zufrieden

Antwort: b) & d)

25. Der Zusammenhang zweier Variablen wurde mit folgendem Streudiagramm dargestellt!
Welche der folgenden Aussagen treffen zu?



a) Es besteht ein positiver linearer Zusammenhang
b) Es besteht kein Zusammenhang
c) Ob eine Korrelation besteht, kann nicht beurteilt werden
d) Es besteht eine negative Korrelation
e) Es besteht ein nicht monotoner Zusammenhang

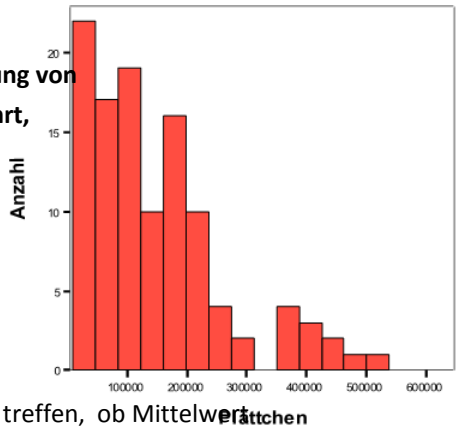
Antwort: a)

26. Das Gewicht von 5 Patienten sei 80kg, 75kg, 90kg, 60kg und 70kg. Berechnen Sie das arithmetische Mittel.

a) 74
b) 75
c) 78
d) 85
e) 90

Antwort: 75

27. Zur Untersuchung des prognostischen Faktors von der Vermehrung von Blutplättchen bei gewissen Krebsarten wurde ein Experiment durchgeführt, bei dem die Vermehrung von Blutplättchen in 112 verschiedenen Zellkulturen untersucht wurde. Betrachte das Histogramm der Variable Plättchenzahl. Welche der folgenden Aussagen ist für die Blutplättchenzahl richtig?



a) Der Median ist größer als der Mittelwert
b) Mittelwert und Median sind ungefähr gleich groß
c) Der Mittelwert ist größer als der Median
d) Man kann aufgrund des Histogramms keine Aussage darüber treffen, ob Mittelwert oder der Median größer sind
e) Alle Angaben falsch

Antwort: c)

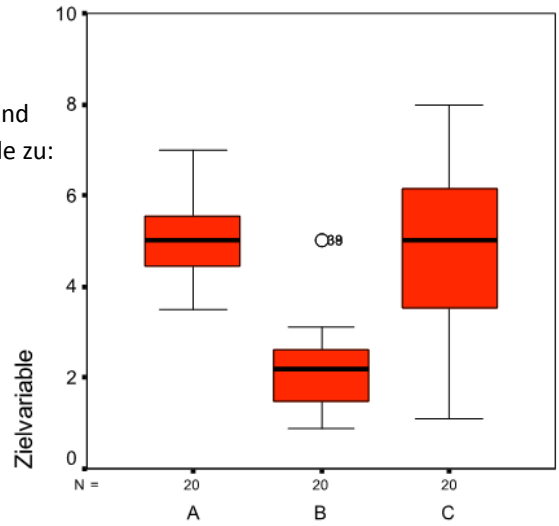
28. Betrachte die folgenden 3 Boxplots für die Behandlungsmethoden A, B und C:

Ordne dem jeweils gegebenen Median und Interquartilsabstand (IQR) der Zielvariable die jeweils richtige Behandlungsmethode zu:

a) Median = 5, IQR = 2,9
   I. Behandlung A
   II. Behandlung B
   III. Behandlung C

b) Median = 2,2, IQR = 1,15
   I. Behandlung A
   II. Behandlung B
   III. Behandlung C

c) Median = 5, IQR = 1,15
   I. Behandlung A
   II. Behandlung B
   III. Behandlung C



Antwort: a) III. / b) II. / c) I.

29. Der Behandlungserfolg einer neuen Therapie T wurde in einer randomisierten Studie mit einer Kontrollgruppe K verglichen. In der Therapiegruppe waren 100 Patienten und in der Kontrollgruppe nur 80. Innerhalb der Therapiegruppe hatten 30 einen Behandlungserfolg, in der Kontrollgruppe nur 10

Wenn Sie einen Chi 2-Test auf obige Daten anwenden, welches der genannten Hypothesen entsprechen der Nullhypothese? (Mehrfachantworten möglich)

a) T und K führen zu einer befriedigenden Erfolgsrate
b) Weder T noch K führen zu einer befriedigenden Erfolgsrate
c) Die Erfolgsrate unter T ist von der Erfolgsrate unter K verschieden
d) Die Erfolgsrate unter T ist dieselbe wie unter K
e) T führt zu einer befriedigenden Erfolgsrate ungeachtet der Erfolgsrate unter K

Antwort: d)

30. Fortsetzung Frage 29.: Welche der nachfolgenden Tabellen gibt die unter der Nullhypothese zu erwarteten Häufigkeiten wieder?

c)	Behandlungserfolg	
	Ja	nein
T	19,5	65,5
K	19,5	65,5

Antwort: e)

a)	Behandlungserfolg	
	Ja	Nein
T	12,2	87,8
K	12,8	67,2

d)	Behandlungserfolg	
	Ja	Nein
T	30	70
K	10	70

b)	Behandlungserfolg	
	Ja	Nein
T	26,7	73,3
K	13,3	66,7

e)	Behandlungserfolg	
	Ja	nein
T	22,2	77,8
K	17,8	62,2

31. Fortsetzung Frage 29.: Der Chi 2-Test liefert einen p-Wert von 0,005. Wie ist dieses Resultat zu interpretieren, wenn das Signifikanzniveau 0,05 beträgt?

a) T und K führen zu einer befriedigenden Erfolgsrate
b) Weder T noch K führen zu einer befriedigenden Erfolgsrate
c) Die Erfolgsrate unter T ist von der Erfolgsrate unter K verschieden
d) Die Erfolgsrate unter T ist dieselbe wie unter K
e) T führt zu einer befriedigenden Erfolgsrate ungeachtet der Erfolgsrate unter K

Antwort: c)

32. In einer randomisierten Doppelblindstudie erhielten 50 Patienten eine Standardtherapie A und 50 Patienten eine neue Therapie B zum Senken des Cholesterinspiegels. Um zum Niveau 5% zu testen, ob die neue Therapie zu einem signifikant geringeren Cholesterinspiegel führt als die Standardtherapie, ist

a) Der Zweistichproben t-Test durchzuführen
b) Eine 2x2 Feldertafel mit dem Chi 2-Test zu testen
c) Muss der Spearman'sche Korrelationskoeffizient berechnet werden
d) Der gepaarte t-Test durchzuführen
e) Keines der angegebenen Verfahren geeignet

Antwort: a)

33. Bei einem statistischen Prüfverfahren versteht man unter einem Fehler 2.Art:

a) Es wurde ein methodischer Fehler begangen
b) Die Nullhypothese H_0 wurde abgelehnt, obwohl sie richtig ist
c) Ein oder mehrere Patienten haben die falsche Behandlung erhalten
d) Die Nullhypothese H_0 wurde angenommen, obwohl sie falsch ist.
e) Die Alternativhypothese H_1 wurde angenommen, obwohl sie falsch ist

Antwort: d)

34. Ergänzen Sie das folgende 3x3 Lateinquadrat:

A	B	C
C		

Antwort: ABC / CAB / BCA

35. Im folgenden Beispiel wurde Mäusen ein Serum der Dosis 2, 5 und 7 verabreicht und ein Teil der Tiere verstarb an Pneumonie:

Dosis	Zahl der toten Mäuse	Zahl der überlebenden Mäuse
2	20	60
5	20	80
7	33	117

a) Wie hoch ist die Sterblichkeit der Mäuse unter Dosis 7?

a) 20%
b) 22%
c) 22.12%
d) 23,2%
e) 25%

Antwort: b) 33 / (33+117) = 0,22

Wie hoch ist die Sterblichkeit der Mäuse (in Prozent) insgesamt?

a) 20%

b) 22%
c) 22.12%
d) 23,2%
e) 25%

Antwort: c) (20+20+33) / (80+100+150) = 0,2212…

36. Alle Patienten, die in einer bestimmen geografischen Region Mitteleuropas während eines Kalenderjahres die Diagnose „Terminales Nierenversagen (End-Stage-Renal-Disease)" entwickeln, werden in die Studie aufgenommen. Sie werden bezüglich bestimmter Risikofaktoren in Ihrem zurückliegenden Leben befragt (z.B. bezüglich eines eventuellen beruflichen Kontakts mit bestimmten Chemikalien und gegebenenfalls über das Ausmaß dieses Kontakts). Parallel wird eine gleich große zufällige Stichprobe von Personen aus der gleichen Region mit ähnlicher Struktur hinsichtlich Alter und Geschlecht in analoger Weise bezüglich der gleichen zurückliegenden Risikofaktoren befragt. Bei dieser Studie handelt es sich um eine (Mehrfachantwort möglich)

a) Prospektive randomisierte Studie
b) Prospektive Kohortenstudie
c) Einzelfallbeobachtung
d) Fall-Kontroll-Studie
e) Studie mit historischer Vergleichsgruppe

Antwort: d)

37. Bei einer Fall-Kontroll-Studie

a) Werden Patienten unterschiedlichen Therapien zufällig zugeteilt
b) Werden 2 Gruppen von Patienten über einen Zeitraum prospektiv erhoben und verglichen
c) Wird eine Stichprobe von Patienten zu einem bestimmten Zeitpunkt untersucht, und die Häufigkeit einer Merkmals erhoben
d) Wird bei Patienten nach Auftreten einer Erkrankung und bei geeignet ausgewählten nicht erkrankten Personen die Risikoexposition in der Vergangenheit erhoben und verglichen
e) Wird der Therapieerfolg bei Fällen mit einer bestimmten Erkrankung nach einem vorgegebenen Zeitintervall nachkontrolliert

Antwort: d)

38. Es gibt eine annähernd lineare Beziehung zwischen der Größe von Mädchen, und Ihrem Alter (von 5 – 18 Jahren), die durch folgende Gleichung gegeben ist:

Größe = 50,3 + (6,01 x Alter)

Dabei wird die Größe in cm und das Alter in Jahren gemessen; Welche der folgenden Aussagen ist <u>falsch</u>?

a) Die geschätzte Steigung ist 6.01
b) Die geschätzte Größe eines 10 Jährigen Mädchens ist ca. 110cm
c) Der geschätzte Achsenabschnitt ist 6.01
d) Die mittlere Größe von 5 – Jährigen Mädchen ist etwa 50% der mittleren Größe von 18 Jährigen
e) Wenn ein Mädchen 8 Jahre alt ist und 115cm groß, dann ist sie größer als der Durchschnitt

Antwort: c)

39. Wofür werden Vierfeldertafeln verwendet?

a) Zur grafischen Darstellung einer metrischen Variable
b) Zur Darstellung des Zusammenhangs von 2 dichotomen Merkmalen
c) Zur grafischen Darstellung von 2 metrischen Variablen
d) Zur 2-D Darstellung einer Variable
e) Zur Berechnung des Korrelationskoeffizienten nach Spearman

Antwort: b)

40. Welche der folgenden Merkmale ist qualitativ?

a) Körpergröße
b) Brustumfang
c) Taillenweite
d) Blutgruppe
e) Körpergewicht

Antwort: d)

41. Bei einer Untersuchung an 800 Patienten wurde unter anderem das Körpergewicht in kg erhoben. Die Daten wurden in Klassen eingeteilt. In der Klasse (50,65] fielen 325 Patienten. Die relative Häufigkeit für diese Klasse ist:

a) 0,59
b) 2,64
c) 2,46
d) 0,41
e) 0,24

Antwort: d)

42. Gegeben sei die Stichprobe eines metrischen Merkmals:

4 7 2 9 6 7 4

Welche Aussage ist richtig?

a) Der Mittelwert ist grösser als der Median
b) Der Mittelwert ist kleiner als der Median
c) Der Mittelwert ist genau gleich so groß wie der Median
d) Mittelwert und Median lassen sich nicht vergleichen
e) Die Differenz zwischen Mittelwert und Median ist nicht berechenbar

Antwort: Mittelwert: 5,57 Median: 6 → b)

43. Aus welcher der unten genannten Grafiken können Sie für eine metrische Variable stets den Median ablesen?

a) Histogramm
b) Kreisdiagramm
c) Boxplot
d) Stabdiagramm
e) Verlaufskurve für den Mittelwert

Antwort: c)

Beantworten Sie bitte Fragen 44 – 47 Anhand folgenden Texts:

Eine neue Therapie (T) soll gegen die Standarttherapie (S) getestet werden. Von S nimmt man aus Langzeiterfahrung an, dass sie eine Heilungsrate von 60% hat. Man will in dem Versuch einen Unterschied von 15% (Daher Therapie T mit Heilungsrate 75%) zwischen den Therapien entdecken. Bei dem üblichen zweiseitigen Signifikanzniveau von $\alpha = 0{,}05$ soll der Nachweis einer derartigen Überlegenheit von T gegenüber S mit einer Wahrscheinlichkeit von 0.8 gelingen (Power $1 - \beta = 0.8$). Dazu werden 165 Patienten pro Behandlungsgruppe benötigt.

44. Wie verändert sich der benötigte Stichprobenumfang, wenn bei sonst gleichbleibenden Planungsvoraussetzungen das *2-Seitige Signifikanzniveau* von *0,5 auf 0,025 gesenkt* wird?

a) Wird größer
b) Unverändert
c) Nicht zu beantworten
d) Wird kleiner
e) Wird kleiner oder größer

Antwort: a)

45. Wie verändert sich der benötigte Stichprobenumfang, wenn bei sonst gleichbleibenden Planungsvoraussetzungen statt des 2-Seitigen Signifikanzniveaus von 0,05 das 1-Seitige Signifikanzniveau von 0,025 verwendet wird?

a) Wird größer
b) Unverändert
c) Nicht zu beantworten
d) Wird kleiner
e) Wird kleiner oder größer

Antwort: d)

46. Wie ändert sich die Power, wenn bei sonst gleich bleibenden Voraussetzungen nur 120 Patienten/Gruppe rekrutiert werden?

a) Wird größer
b) Unverändert
c) Nicht zu beantworten
d) Wird kleiner
e) Wird kleiner oder größer

Antwort: d)

47. Wie ändert sich die Power, wenn bei sonst gleich bleibenden Voraussetzungen der Unterschied zwischen den Therapien überschätzt wurde und tatsächlich nur 10% (Heilungsrate von 70%) ausmacht?

a) Wird größer
b) Unverändert
c) Nicht zu beantworten
d) Wird kleiner
e) Wird kleiner oder größer

Antwort: d)

In einer Querschnittsstudie wird eine erhöhte Plasmakonzentration von BNP (brain natriuretic peptide) zur Diagnose von linksventrikulärer Dysfunktion (LVD) bei 126 Verdachtsfällen untersucht (BMJ 2000). Der medizinische Hintergrund ist, dass das Hormon BNP hauptsächlich von den Kardiozyten als Antwort auf eine starke Dehnung ausgeschüttet wird. Die gemessenen Messergebnisse sind in der folgenden Tabelle wiedergegeben. Ein Test wurde als positiv gewertet, falls BNP > 17,9 pg/ml. Als negativ < 17,9 pg/ml.

Krankheit * Test-Kreuztabelle

		Test positiv	Test negativ	Gesamt
Krankheit: LVD:	Anzahl	35	5	40
	% von Krankheit	87,5 %	12,5 %	100,0 %
	% von Test	38,0 %	14,7 %	31,7 %
	% der Gesamtzahl	27,8 %	4,0 %	31,7 %
Ø LVD:	Anzahl	57	29	86
	% von Krankheit	66,3 %	33,7 %	100,0 %
	% von Test	62,0 %	85,3 %	88,3 %
	% der Gesamtzahl	45,2 %	23,0 %	88,3 %
Gesamt	Anzahl	92	34	126
	% von Krankheit	73,0 %	27,0 %	100,0 %
	% von Test	100,0 %	100,0 %	100,0 %
	% der Gesamtzahl	73,0 %	27,0 %	100,0 %

48. Wie hoch ist die Sensitivität?

a) 12,5
b) 33,7
c) 38,0
d) 66,3
e) 87,5

Antwort: (35 / 40) x 100 = 87,5

49. Wie hoch ist die Spezifität?

a) 12,5
b) 33,7
c) 85,2
d) 66,3
e) 87,5

Antwort: (29 / 86) x 100 = 33,6

50. Wie groß ist die Wahrscheinlichkeit für ein falsch positives Testergebnis zugunsten von A in einem zweiseitigen Test zum Signifikanzniveau 0.05?

a) 0,025
b) 0,5
c) 0,125
d) 0,875
e) 0,975

Antwort: a)

51. Wie groß ist die Wahrscheinlichkeit, unter fünf unabhängigen Studien kein einzig signifikantes Ergebnis zugunsten von A zu beobachten? (Hier fehlt die Angabe: „bei Zutreffen der Nullhypothese, bei einem zweiseitigen Test zum Signifikanzniveau 0,05“)

a) 0,12
b) 0,2
c) 0,66
d) 0,88
e) 0,16

Antwort: d) $1 - 0{,}025 = 0{,}975$
$0{,}975^5 = 0{,}88$

52. Wie groß ist die Wahrscheinlichkeit mindestens ein signifikantes Ergebnis zugunsten von A zu erhalten? (Hier fehlt die Angabe: „bei Zutreffen der Nullhypothese, bei einem zweiseitigen Test zum Signifikanzniveau 0,05“)

a) 0,08
b) 0,10
c) 0,12
d) 0,14
e) 0,16

Antwort: 1 - 0.88 = 0.12 →c)

53. Bei einem Doppelblindversuch:

a) Wissen die Angehörigen der Patienten nicht, welche Therapie die Patienten erhalten wohl aber die Patienten selbst
b) Wissen weder die Patienten noch das medizinische Personal, welche Therapie die Patienten erhalten wohl aber die Ärzte
c) Wissen weder die Patienten noch die Angehörigen, welche Therapie die Patienten erhalten wohl aber die Ärzte
d) Wissen weder die Patienten noch die Ärzte oder das medizinische Personal, welche Therapie die Patienten erhalten
e) Wissen weder die Ärzte noch das medizinische Personal, welche Therapie die Patienten erhalten wohl aber die Patienten selbst

Antwort: d)

54. Ein p-Wert von p = 0,028 bedeutet, dass

a) Die Nullhypothese zur Signifikanz von 0,028 abgelehnt wird
b) Keine Testentscheidung möglich
c) Die Nullhypothese zur Signifikanz von 0,028 abgelehnt werden kann
d) Die Nullhypothese mit einer Wahrscheinlichkeit von 0,028 richtig verworfen wird
e) Die Alternativhypothese mit einer Wahrscheinlichkeit von 0,028 zutrifft

Antwort: c)

55. Ein Streudiagramm dient zur Darstellung

a) Der Variabilität eines metrisch skalierten Merkmals
b) Eines Ordinal skalierten Merkmals
c) Des Zusammenhangs zweier metrisch skalierter Merkmale
d) Einer Kreuztabelle
e) Des Zusammenhangs zweier nominal skalierter Merkmale

Antwort: c)

56. Die Regressionsgerade geht stets durch:

a) Mindestens zwei Punkte des Streudiagramms
b) Den Nullpunkt des Koordinatensystems
c) Den Nullpunkt und den Schwerpunkt
d) Keinen Punkt des Streudiagramms
e) Durch beide Mittelwerte (den Schwerpunkt) x,y

Antwort: e)

57. Welches der folgenden Merkmale ist quantitativ skalierbar?

a) Augenfarbe
b) Blutgruppe
c) Brustumfang
d) Haarfarbe
e) Rhesusfaktor

Antwort: c)

58. Es wird 15 Patienten, die an essentieller Hypertonie leiden, Captopril verabreicht. Die folgende Tabelle enthält den systolischen Blutdruck vor der Einnahme des Medikaments bzw. zwei Stunden danach. Angenommen die Blutdruckwerte sind normal verteilt, welches statistische Verfahren kann verwendet werden, um zu überprüfen, ob nach Gabe des Medikamentes der mittlere systolische Blutdruck gesenkt ist?

Patientennummer	Systolische RR: vorher	Systolischer RR: nachher
1	210	201
2	169	165
3	187	166
4	160	157
5	167	147
6	176	145
7	185	168
8	206	180
9	173	147
10	146	136
11	174	151
12	201	168

a) Korrelationskoeffezient nach Spearman
b) Gepaarter t-Test
c) $\chi 2$ - Test
d) Korrelationskoeffizient nach Pearson
e) t Test für unabhängige Stichproben

Antwort: b)

59. In einer klinischen Studie erhielten 100 Personen ein Vitaminpräparat und 100 Personen ein Placebo. Innerhalb eines Monats haben 30 Personen in der Vitamingruppe eine Grippe bekommen und 40 Personen in der Placebo-Gruppe. Mit welchem Test wird die Wirksamkeit des Vitaminpräparats überprüft?

a) t-Test (gepaart)
b) $\chi 2$-Test
c) t-Test für unabhängige Gruppen
d) Korrelationstest
e) Kann nicht durchgeführt werden

Antwort: b)

60. Welche der Aussagen zum Korrelationskoeffizienten nach Pearson ist falsch?

a) Ein einzelner extremer Datenpunkt kann den Korrelationskoeffizienten nach Pearson stark beeinflussen
b) Ist die Steigung der Regressionsgerade positiv, so muss der Korrelationskoeffizient nach Pearson ebenfalls positiv sein
c) Der Korrelationskoeffizient liegt immer im Intervall [–1,1]
d) Der Korrelationskoeffizient nach Pearson beschreibt den linearen Zusammenhang zweier Variablen
e) Eine stark positive Korrelation ist Beweis für einen kausalen Zusammenhang zweier Variablen

Antwort: e)

61. Wie interpretieren Sie eine Korrelationskoeffizienten von -0,8?

a) Als deutlichen positiven Zusammenhang zwischen den betrachteten
b) Ein Anstieg der einen Variablen bedeutet im Mittel einen Abfall der anderen.
c) Kein Zusammenhang zwischen den Variablen
d) Die betrachteten Variablen werden von einer dritten negativ beeinflusst
e) Die Variablen liegen perfekt auf einer geraden

Antwort: b)

62. Gegeben sei folgende Häufigkeitstabelle für ein ordinales Merkmal:

BMI

	Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig:				
Untergewicht	5	20,0	20,0	20,0
Idealgewicht	17	68,0	68,0	88,0
Übergewicht	3	12,0	12,0	100,0
Gesamt	25	100,0	100,0	

Wie groß ist die kumulierte absolute Häufigkeit von Idealgewicht?

a) 22
b) 8
c) 25
d) 3
e) 20

Antwort: 5 + 17 = 22

Der Behandlungserfolg einer neuen Therapie (T) wurde in einer randomisierten Studie mit einer Kontrollgruppe (K) verglichen. Insgesamt wurden 103 Patienten in der Studie untersucht, davon fielen 52 in die Kontrollgruppe. Nachfolgend werden die erhaltenen Ergebnisse dargestellt:

Gruppe * Erfolg-Kreuztabelle

		Erfolg : Ja	Erfolg : Nein	Gesamt
Gruppe: Kontrolle (K):	Anzahl	40	12	52
	% von Gruppe	78,9 %	23,1 %	100,0 %
	% von Erfolg	58,0 %	35,3 %	50,5 %
	% der Gesamzahl	38,8 %	11,7 %	50,5 %
Therapie (T):	Anzahl	29	22	51
	% von Gruppe	56,9 %	43,1 %	100,0 %
	% von Erfolg	42,0 %	64,7 %	49,5 %
	% der Gesamzahl	28,2 %	21,4 %	49,5 %
Gesamt:	Anzahl	69	34	103
	% von Gruppe	67,0 %	33,0 %	100,0 %
	% von Erfolg	100,0 %	100,0 %	100,0 %
	% der Gesamzahl	67,0 %	33,0 %	100,0 %

	Wert	df	Asymptomatische Signifikanz (2-seitig)	Exakte Signifikanz (2-seitig)	Exakte Signifikanz (1-seitig)
Chi^2 nach Pearson	4,686[b]	1	0,030		
Anzahl der gültien Fälle	103				

63. Wie hoch war der prozentuelle Behandlungserfolg in der Kontrollgruppe (K)?

a) 11,7%
b) 23,1%
c) 35,3%
d) 50,5%
e) 76,9%

Antwort: e) 40 / 52 = 0,769

64. Betrachten Sie die Prüfgröße des χ2-Tests nach Pearson. Wie ist dieses Resultat zu interpretieren, wenn das Signifikanzniveau 0,05 (zweiseitig) beträgt?

a) Es kann kein Zusammenhang zwischen Behandlung und Erfolgsrate nachgewiesen werden
b) Der Unterschied der Erfolgsraten zwischen (T) und (K) beträgt in Wahrheit 3%
c) (T) und (K) führen zu einer befriedigenden Erfolgsrate
d) Es kann ein Zusammenhang zwischen Behandlung und Erfolgsrate nachgewiesen werden
e) (T) führt zu einer befriedigenden Erfolgsrate ungeachtet der Erfolgsrate unter (K)

Antwort: d)

65. Bei wie vielen Patienten konnte in dieser Studie insgesamt ein Erfolg erzielt werden?

a) 103
b) 69
c) 34
d) 22
e) 12

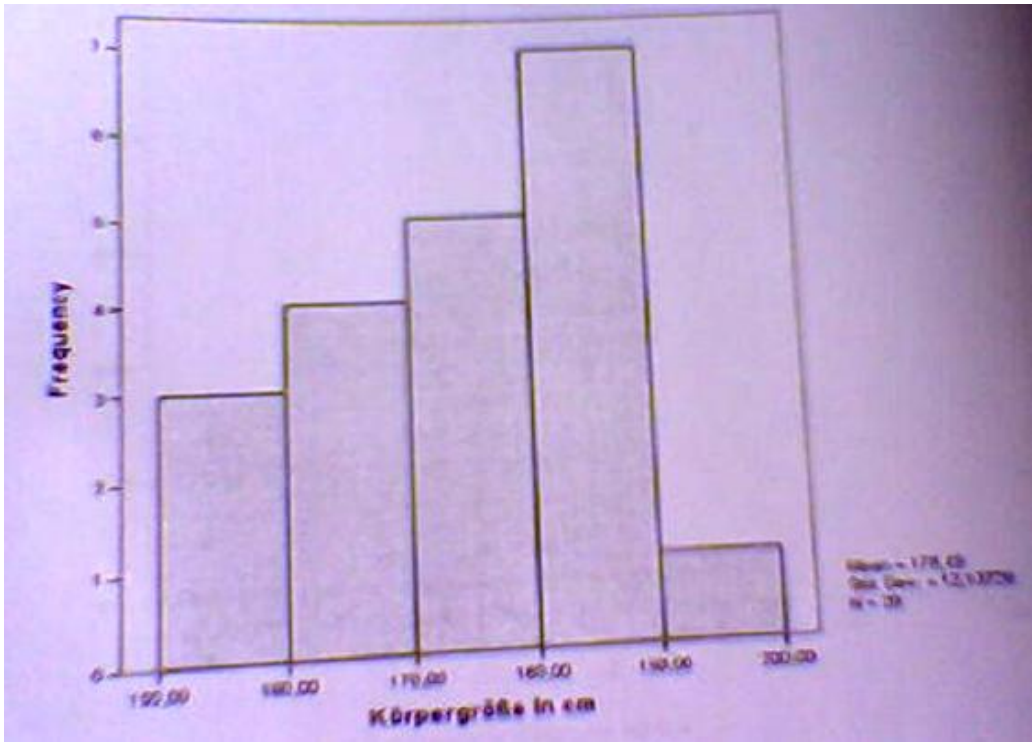
Antwort: b)

66. Für das obige Beispiel ist die erwartete Häufigkeit unter H_0 für die Zelle Gruppe, Therapie (T) und Erfolg = Nein:

a) 40,3
b) 34,2
c) 17,2
d) 16,8
e) 14,4

Antwort: d) (34/103) = 0,33
0,33 * 51 = 16,8

67. Das folgende Histogramm zeigt eine Verteilung der Körpergröße in cm von 20 Personen (Leider kein besseres Bild vorhanden; Frequency: Je eine Person; Klassenbreite: 10cm, erste Klasse: 150 – 160; …)



Welche Daten gehören zu diesem Histogramm?

a) 153, 155, 156, 165, 166, 168, 169, 171, 175, 177, 178, 178, 182, 185, 187, 187, 188, 189, 189, 191
b) 153, 155, 156, 156, 156, 168, 169, 171, 175 177 178, 178, 182, 185, 187, 187, 188, 189, 189, 191

c) 145, 155, 156, 165, 166, 168, 169, 171, 175, 177, 178, 178, 182, 185, 187, 187, 188, 189,189, 191
d) 153, 155, 156, 165, 166, 168, 169, 171, 175, 177, 178, 178, 182, 185, 187, 187, 188, 189, 189, 198, 205
e) 153, 155, 156, 165, 165, 166, 168, 169, 169, 171, 175, 177, 178, 173, 182, 185, 187, 187, 188, 189, 189, 189

Antwort: a)

68. Welche der folgenden Versuchsanordnung ist ein lateinisches Quadrat?

a)

A	B	C
B	C	A
C	A	B

b)

A	B	C
A	B	C
A	B	C

c)

A	B	C
B	B	A
A	B	C

d)

A	B	A
B	C	B
C	A	C

e)

A	B	C
C	A	B
C	B	A

Antwort: a)

69. Gegeben sei die folgende Stichprobe eines metrischen Merkmals: 10, 7, 1, 2, 9, 3, 4, 5 Welche Aussage ist richtig?

a) Mittelwert und Median lassen sich nicht vergleichen
b) Die Differenz zwischen Mittelwert und Median ist nicht berechenbar
c) Der Mittelwert ist genau gleich dem Median
d) Der Mittelwert ist kleiner als der Median
e) Der Mittelwert ist größer als der Median

Antwort: e) Mittelwert = 5,1
Median = 4,5

70. Bei 50 erstgebärenden Frauen wird die Wirksamkeit eines Medikaments untersucht, das bei Wehenschwäche den Muttermund öffnen soll. Die Beobachtungseinheiten sind:

a) Die Hebammen, die die Frauen während der Geburt beobachten
b) Die gemessenen Werte für die Geschwindigkeit, mit der sich nach Verabreichung des Medikaments der Muttermund öffnet
c) Die Ärzte, die die Daten auswerten
d) Die 50 Frauen, die das Medikament erhalten
e) Alle Frauen, bei denen dieses Medikament theoretisch hilfreich sein kann

Antwort: d)

71. Jeder Patient, der sich im Klinikum M. einer Operation unterzieht, wird bzgl. des Risikos eingestuft nach ASA I (geringes Risiko), ASA II bis ASA V (sehr schweres Risiko). Welche Eigenschaften hat dieses Merkmal?

a) Alternativmerkmal
b) qualitativ, nur nominalskaliert
c) qualitativ, ordinalskaliert

d) quantitativ, diskret
e) quantitativ, stetig

Antwort: c)

72. In einer gynäkologischen Klinik wird untersucht, wie sich die Rauchgewohnheiten schwangerer Frauen auf das Geburtsgewicht ihrer Kinder auswirken. Es werden erfaßt:

I. die durchschnittliche Anzahl der Zigaretten, die die Mutter pro Tag raucht,
II. das Geburtsgewicht des Kindes,
III. das Alter der Mutter,
IV. das Körpergewicht der Mutter vor der Schwangerschaft

(der Einfluß der beiden letzten Merkmale wird aber nicht ausgewertet). Wie lassen sich die Merkmale einordnen?

I. Anzahl der pro Tag gerauchten Zigaretten
II. Alter der Mutter
III. Gewicht der Mutter vor der Schwangerschaft
IV. psychische oder soziale Belastungen der Mutter
V. Ernährungsweise der Mutter
VI. Geburtsgewicht des Kindes

a. Faktor
b. Begleitmerkmal
c. Störgröße
d. Zielgröße

a) 1a, 2b, 3b, 4c, 5c, 6d
b) 1a, 2a, 3a, 4c, 5c, 6d
c) 1c, 2a, 3a, 4c, 5b, 6d
d) 1a, 2b, 3b, 4a, 5a, 6d
e) 1a, 2a, 3a, 4c, 5c, 6a

Antwort: a)

73. Welche Aussage trifft zu?

a) Eine Ordinalskala kann auf eine Intervallskala transformiert werden, wenn alle Ausprägungen numerisch codiert sind
b) Ein qualitatives Merkmal mit sehr zahlreichen Ausprägungen kann als ein stetiges angesehen werden
c) Die Transformation auf ein anderes Skalenniveau ist generell nicht möglich
d) Die Transformation auf ein anderes Niveau ist immer möglich, aber niemals sinnvoll
e) Eine Verhältnisskala kann auf eine Ordinalskala transformiert werden

Antwort: e)

74. Der Eiweißgehalt im Urin lässt sich exakt in mg/dl messen. Bei einer Routineuntersuchung werden Teststreifen verwendet, mit denen sich lediglich feststellen lässt, ob der Eiweißgehalt im pathologischen Bereich liegt. Welche Aussage bzgl. dieses Meßverfahrens trifft <u>nicht</u> zu?

a) Theoretisch wird eine Reduktion von einer metrischen Skala auf eine Nominalskala durchgeführt
b) Der Eiweißgehalt wird bei dieser Meßmethode als ein Alternativmerkmal erfasst
c) Die Meßmethode ist in jedem Fall sinnlos, da sehr viel Information verloren geht

d) Das Meßverfahren ist einfacher durchzuführen als die exakte Messung in mg/dl
e) Die Ergebnisse dieses Meßverfahrens ermöglichen weniger differenzierte Auswertungen als die exakten Meßwerte in mg/dl

Antwort: c)

75. Das Merkmal X „Körpergröße“ soll bei Studenten grob erfasst werden. Dazu werden Klassen mit der Breite 10 cm gebildet. Die Größe wird folgendermaßen erfasst:

0: $X \leq 150$ cm
1: 150 cm $\leq X \leq$ 160 cm
2: 160 cm $\leq X \leq$ 170 cm
3: 170 cm $\leq X \leq$ 180 cm
4: 180 cm $\leq X \leq$ 190 cm
5: 190 cm $\leq X \leq$ 200 cm
6: $X \geq 200$ cm

Welche Aussage trifft zu?

a) Diese Codierung ist für praktische Untersuchungen zu grob und deshalb unbrauchbar
b) Die Ausprägungsliste ist vollständig und disjunkt
c) Die Ausprägungsliste ist nicht vollständig
d) Die Ausprägungsliste ist nicht disjunkt
e) Die Codierung ermöglicht eine übersichtliche Darstellung der Körpergrößen ohne Informationsverlust

Antwort: d) „nicht disjunkt“ bedeutet: Wenn eine Person z.B. genau 170 cm groß ist, dann kann sie hier entweder noch zu Klasse 2, oder aber schon zu Klasse 3 gezählt werden! Die Klassen überschneiden sich leicht

76. Beurteilen Sie die folgende Aussage:

1. Der Glukosegehalt im Blut ist ein stetiges Merkmal,

denn

2. dieses Merkmal ist nicht qualitativ

	Aussage 1	Aussage 2	Verknüpfung
a)	Richtig	Richtig	Richtig
b)	Richtig	Richtig	Falsch
c)	Richtig	Falsch	Falsch
d)	Falsch	Richtig	Richtig
e)	Falsch	Falsch	Richtig

Antwort: b)

77. Mit einem Balkendiagramm lassen sich Häufigkeiten graphisch darstellen. Für welche Merkmale ist diese Darstellungsform geeignet?

a) generell für alle Merkmale
b) für alle diskreten Merkmale sowie für stetige Merkmale, wenn die Daten in Klassen eingeteilt sind
c) nur für nominalskalierte Merkmale
d) nur für qualitative Merkmale
e) für alle diskreten Merkmale

Antwort: e)

78. Im Rahmen einer klinischen Studie wird die Wirksamkeit einer therapeutischen Maßnahme an 22 Patienten untersucht. Bei 14 1 = n Patienten ist die Therapie erfolgreich.

Welche Darstellung der entsprechenden relativen Häufigkeit ist am sinnvollsten?

a) $h_1 = 64\%$
b) $h_1 = 0{,}6363\%$
c) $h_1 = 14 / 22$
d) $h_1 > 50\%$
e) $60\% < h_1 < 70\%$

Antwort: c)

79. In einer Stichprobe von 200 = n Personen werden die in der Tabelle aufgelisteten Blutdruckwerte (in mm Hg) gemessen

diastolischer Blutdruck	systolischer Blutdruck ≤ 120	121 – 150	> 150	Summe
≤ 80	27	13	3	43
81 – 100	19	102	20	141
> 100	0	1	15	16
Summe	46	116	38	200

Hypertonie liege vor, wenn der systolische Blutdruck mehr als 150 mmHg oder der diastolische mehr als 100 mmHg beträgt. Wie groß ist dann die relative Häufigkeit der Patienten, die an Hypertonie leiden?

a) 15/200
b) 15/38
c) 15/16
d) 54/200
e) 39/200

Antwort: e)

80. Welche Aussage bzgl. klassierter Daten ist <u>falsch</u>?

a) Die Klassenbildung setzt ein quantitatives Merkmal voraus
b) Die optimale Klassenanzahl ist abhängig vom Stichprobenumfang
c) Durch die Klassenbildung geht Information verloren, dafür ist die Darstellung der Häufigkeitsverteilung übersichtlicher
d) Die Klassen müssen immer gleich breit sein
e) An einem Histogramm sind charakteristische Eigenschaften der Merkmalsverteilung (Lage, Streuung, Verteilungsform) erkennbar

Antwort: d) (sogar innerhalb eines Histogramms können sich die Klassenbreiten unterscheiden, denn es zählt nur die Fläche!)

81. 69 Studenten schreiben eine Klausur, bei der maximal 10 Punkte zu erreichen sind. Die Klausur gilt als bestanden, wenn mindestens 5 Punkte erreicht werden. Es ergibt sich folgende Häufigkeitstabelle:

I.	**Anzahl Punkte**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
II.	**Häufigkeit**	**12**	**9**	**14**	**18**	**10**	**4**	**2**

Der Median beträgt:

a) 6
b) 6,5
c) 7
d) 14
e) nicht bestimmbar

Antwort: a)

82. Welche Aussage ist richtig? Der Median bleibt in jedem Fall unverändert, wenn

a) alle Werte außerhalb des Intervalls s x 2 ± aus der Stichprobe entfernt werden
b) zum größten Wert eine positive Zahl addiert wird
c) alle Werte mit der gleichen Zahl multipliziert werden
d) zu allen Werten eine Konstante addiert wird
e) man einen Ausreißer weglässt

Antwort: b) Vorsicht, Median ist nicht ganz gegen Ausreißer geschützt, sondern nur nicht so sensibel wie der Mittelwert (n ändert sich trotzdem) daher nicht e)

83. Jeder Patient, der sich im Klinikum M. einer Operation unterzieht, wird bezüglich des Risikos eingestuft nach ASA I (geringes Risiko) bis ASA V (sehr schweres Risiko). Welche Maßzahlen lassen sich bei diesem Merkmal berechnen?

a. Mittelwert
b. Median
c. Varianz
d. Standardabweichung
e. Spannweite

a) alle angegebenen Maßzahlen können berechnet werden
b) nur a, c und e
c) nur c
d) nur b und e
e) nur a und b

Antwort: d) Mittelwert, Varianz und Standartabweichung kann man eigentlich nur bei einer Normalverteilung (Glockenverteilung) angeben!

84. Der Zusammenhang 2er metrischer Merkmale lässt sich durch eine Punktwolke graphisch darstellen. Welche Informationen lassen sich nicht der Punktwolke entnehmen?

a) ob der Zusammenhang annähernd linear ist
b) ob Ausreißer vorhanden sind
c) ob der Zusammenhang stark oder eher schwach ist
d) ob der Zusammenhang gleich- oder gegensinnig ist
e) ob die beiden Merkmale in einem kausalen Zusammenhang stehen

Antwort: e)

85. Welches Intervall umfasst alle Zahlenwerte, die ein empirischer Korrelationskoeffizient r annehmen kann?

a) $]-\infty, +\infty[$
b) $[0, 1]$
c) $[0, +\infty[$
d) $]-1, +1[$
e) $[-1, +1]$

Antwort: e)

86. Welcher empirische Korrelationskoeffizient nach Pearson bezeichnet den stärksten (linearen) Zusammenhang?

a) r = 0
b) r = 0,8
c) r = - 0,95
d) r = 1,2
e) Dieses Maß ist nicht geeignet, um die Stärke eines Zusammenhangs zu quantifizieren

Antwort: c)

87. Eine Regressionsgerade habe die Form: y = -0,8 + 0,3 x Was folgt daraus für den dargestellten Zusammenhang?

a) Der Zusammenhang ist gegensinnig.
b) Der Zusammenhang ist gleichsinnig.
c) Der Korrelationskoeffizient beträgt r= - 0,8
d) Der Korrelationskoeffizient beträgt r= + 0,3
e) Keine dieser Aussagen lässt sich herleiten

Antwort: b)

88. Ein Korrelationskoeffizient betrage r = 0,2. Was folgt daraus für die Regressionsgerade?

a) Die Steigung der Regressionsgeraden ist positiv
b) Die Steigung der Regressionsgeraden ist negativ
c) Die Steigung der Regressionsgeraden beträgt 0,2
d) Der y-Achsenabschnitt beträgt 0,2
e) Da der Zusammenhang sehr schwach ist, ist die Darstellung durch eine Regressionsgerade nicht erlaubt

Antwort: a)

89. Ein Patient mit Bluthochdruck erhält eine Therapie über 15 Tage, die danach abgebrochen wird. Bezüglich der Änderung des systolischen Blutdrucks über die Zeit wird ein Korrelationskoeffizient r = - 0,89 und die Regressionsgerade y = 180 – 4.x ermittelt (wobei x die Behandlungstage und y die Blutdruckwerte in mmHg sind). Welche der folgenden Aussagen lassen sich aus diesen Informationen schlussfolgern?

I. Der Blutdruck sinkt während der Therapie um durchschnittlich 0,89 mmHg pro Tag.
II. Der Blutdruck sinkt während der Therapie um durchschnittlich 4 mmHg pro Tag.
III. Der Schätzwert für den letzten Tag der Therapie beträgt 120 mmHg.
IV. Zu Beginn der Therapie hatte der Patient einen Blutdruck von etwa 180 mmHg.
V. Am 20. Tag nach Beginn der Therapie ist bei dem Patienten ein Blutdruck von 100 mmHg zu erwarten

a) nur die Aussage 2 ist herleitbar
b) nur die Aussage 1 ist herleitbar
c) nur die Aussagen 2, 3 und 4 sind herleitbar
d) die Aussagen 2, 3, 4 und 5 sind herleitbar
e) die Aussagen 1, 3, 4 und 5 sind herleitbar

Antwort: c)

90. Von zwei metrisch skalierten Merkmalen ist nur bekannt, dass der Zusammenhang monoton steigend ist. Welches Maß eignet sich zur Quantifizierung der Stärke dieses Zusammenhangs?

a) die Korrelation
b) der Korrelationskoeffizient nach Pearson
c) der Korrelationskoeffizient nach Spearman
d) das Produkt der beiden Standardabweichungen
e) keine der Angaben aus A–D

Antwort: c)

91. Wie viele der folgenden 10 Maßzahlen können niemals negative Werte annehmen?

Spannweite – Varianz – Standardabweichung – Minimum – Maximum – Modus – Median – Korrelationskoeffizient – Kovarianz – Bestimmtheitsmaß

a) alle 10
b) nur 7
c) nur 5
d) nur 4
e) nur 2

Antwort: d) (Spannweite, Varianz, Standardabweichung und Bestimmtheitsmaß)

92. Beurteilen Sie die folgende Aussagen:

1. Der Korrelationskoeffizient, der den Zusammenhang zwischen Körpergröße und Körpergewicht von männlichen Erwachsenen quantifiziert, ist positive

denn

2. diese beiden Merkmale können nur positive Werte annehmen

	Aussage 1	Aussage 2	Verknüpfung
a)	Richtig	Richtig	Richtig
b)	Richtig	Richtig	Falsch
c)	Richtig	Falsch	Falsch
d)	Falsch	Richtig	Richtig
e)	Falsch	Falsch	Richtig

Antwort: b)

93. In einer Klinik betrage der Anteil der an Diabetes mellitus erkrankten Patienten 20%. 30% der Patienten leiden an einer Herzerkrankung; 5% haben beide Krankheiten. Wie groß ist der Anteil der Patienten, die entweder Diabetes- oder herzkrank sind?

a) 20%
b) 30%
c) 45%
d) 40%
e) 50%

Antwort: d)

94. Ignaz Semmelweis ermittelte für einen Monat des Jahres 1846, dass in einer Abteilung des Wiener Gebärhauses 24% der gebärenden Frauen an Kindbettfieber erkrankten. Die Wahrscheinlichkeit, an Kindbettfieber zu sterben, betrug damals 80%. Wie groß war dann die Wahrscheinlichkeit für eine Frau, an Kindbettfieber zu erkranken und daran zu sterben?

a) Um diese Frage zu beantworten, bedarf es weiterer Informationen.
b) 104%
c) 80%
d) 24%
e) etwa 19%

Antwort: e)

95. Bei einem Versuchstier werden 2 Stellen am Rücken mit 2 unterschiedlichen Karzinogenen bepinselt. Die Wahrscheinlichkeit, ein Karzinom zu erzeugen, betrage 0,3 bzw. 0,8. Die Ereignisse seien unabhängig. Wie groß ist die Wahrscheinlichkeit, daß mindestens ein Karzinom entsteht?

a) 1,1
b) 0,8
c) 0,3
d) 0,24
e) 0,86

Antwort: e) Einfacher Trick: man rechnet sich die Gegenwahrscheinlichkeit aus, nämlich:
wie wahrscheinlich ist es, dass eine Maus GAR KEIN Karzinom entwickelt? = 0,7*0,2
das zieht man dann von 1 ab —> 1-(0,7*0,2) = 0,86

96. Von allen Frauen, die nicht familiär vorbelastet sind, erkrankt ungefähr jede 10. Im Laufe ihres Lebens an einem Mammakarzinom. Falls eine nahe weibliche Verwandte bereits an einem Mammakarzinom erkrankt war oder ist, steigt der Wert für die Wahrscheinlichkeit, zu erkranken, auf das 3,5fache. Wie hoch ist dann das zuschreibbare Risiko für den Faktor „familiär belastet"?

a) das zuschreibbare Risiko ist nicht berechenbar
b) 3,5
c) 0,10
d) 0,35
e) 0,25

Antwort: e) 0,35 – 0, 1 = 0,25

97. Was ist bei einem diagnostischen Test, der in der Notfallmedizin bei lebensbedrohlichen Erkrankungen mit guter therapeutischer Beeinflussbarkeit angewandt wird, vor allem anzustreben?

a) eine hohe Spezifität
b) eine hohe Sensitivität
c) eine niedrige Spezifität
d) eine niedrige Sensitivität
e) Spezifität und Sensitivität sollten gleich groß sein

Antwort: b)

98. Ein diagnostischer Test habe eine Sensitivität von 95%; die Prävalenz betrage 0,1%. Was folgt aus diesen Angaben?

a) Die Wahrscheinlichkeit für ein falsch positives Testergebnis beträgt 5%

b) Die Wahrscheinlichkeit für ein falsch negatives Testergebnis beträgt 5%
c) Der positive Vorhersagewert beträgt 5%
d) Die Spezifität beträgt mindestens 95%
e) Keine dieser Aussagen ist herleitbar

Antwort: b) Sensitivität ist die Wahrscheinlichkeit, daß der Test richtig positiv erkennt. Das dazu komplementäre Ereignis ist falsch negativ. Die Wahrscheinlichkeit berechnet sich dann als 05

99. Die Prävalenz einer Krankheit sei 0,0001. Ein diagnostischer Test habe eine Spezifität von 0,995 und eine Sensitivität von 0,98. Wie groß ist etwa der positive Vorhersagewert?

a) 0,02
b) 0,98
c) 0,50
d) 0,999
e) Der positive Vorhersagewert lässt sich aus den Angaben nicht ermitteln

Antwort: a)

100. Bei vielen Tests ist das Ergebnis auch abhängig von einem Schwellenwert. Wie ändern sich die Sensitivität und die Spezifität, wenn der Schwellenwert herabgesetzt wird?

a) Die Sensitivität und die Spezifität bleiben unverändert.
b) Die Sensitivität sinkt, die Spezifität steigt.
c) Die Spezifität sinkt, die Sensitivität steigt.
d) Die Sensitivität und die Spezifität werden größer.
e) Die Sensitivität und die Spezifität werden kleiner.

Antwort: c)

101. Bei einem jungen Mann, der drogenabhängig ist, wird ein HIV-Test durchgeführt (die Prävalenz bei Drogenabhängigen beträgt 10%). Die Sensitivität des Tests liegt bei 98%. Das Testergebnis ist positiv. Wie ist dieses Ergebnis zu interpretieren?

a) Das Ergebnis belegt eindeutig, dass der Patient infiziert ist.
b) Das Ergebnis belegt eindeutig, dass der Patient nicht infiziert ist.
c) Mit 98%-iger Wahrscheinlichkeit ist der Patient infiziert.
d) Die Angabe, dass der Patient drogenabhängig ist, ist irrelevant für die Interpretation des Ergebnisses.
e) Aufgrund des Testergebnisses ist eine HIV-Infektion nicht auszuschließen. Für eine sichere Diagnose ist das Ergebnis jedoch unzureichend.

Antwort: e)

102. Beide Partner eines Elternpaares haben die Anlage für ein rezessives Erbleiden im heterozygoten Zustand. Sie haben n Kinder; X sei die Anzahl der Kinder mit heterozygoten Erbanlagen. Welche Aussage trifft <u>nicht</u> zu?

a) X folgt einer Binomialverteilung.
b) Der Erwartungswert von X beträgt n / 2
c) Die Varianz von X beträgt n / 4
d) Die Verteilung von X ist symmetrisch.
e) X hat denselben Erwartungswert wie die Zufallsvariable, die die Anzahl der homozygot erkrankten Kinder beschreibt

Antwort: c)

103. Wie groß ist die Wahrscheinlichkeit, dass in einer Übungsgruppe mit 30 Studenten am Tag der Klausur ein Student Geburtstag hat?

a) 0,0657
b) 0,0789
c) 0,0822
d) 0,0027
e) Keiner der Antwortmöglichkeiten

Antwort: b) a) wäre wenn die Wahrscheinlichkeit das genau 1 Student Geburtstag hat

104. Welche Aussage ist richtig?

a) Bei einer Normalverteilung liegen alle Werte zwischen $\mu - 3\sigma$ und $\mu + 3\sigma$
b) Je größer die Varianz, desto flacher verläuft die Glockenkurve.
c) In der Regel stimmen der Erwartungswert und der Median nicht überein.
d) Der Erwartungswert ist immer gleich 0.
e) Der Erwartungswert kann keine negativen Werte annehmen

Antwort: b)

105. Die Körpergröße männlicher Studenten sei normalverteilt mit einem Erwartungswert von 180 = m cm und einer Standardabweichung von 6 = s cm. Wieviel Prozent der Studenten sind dann größer als 186 cm oder kleiner als 168 cm?

a) etwa 32 %
b) etwa 5 %
c) etwa 18,5 %
d) etwa 34,5 %
e) etwa 95 %

Antwort: c) Erklärung: 186 ist $\mu + 1\sigma$ und 168 ist $\mu - 2\sigma$

106. Herr Fröhlich erwartet sehnsüchtig den Besuch seiner Schwiegermutter. Diese muss auf dem Weg zu seiner Wohnung Autobus und U-Bahn benützen. Erfahrungsgemäß beträgt die Wahrscheinlichkeit, dass die Schwiegermutter das jeweilige Verkehrsmittel planmäßig erreicht, beim Autobus 7/8, und bei der U-Bahn 4/5. Die Schwiegermutter trifft pünktlich in Herrn Fröhlichs Wohnung ein, wenn sie bei zumindest einem Verkehrsmittel rechtzeitig ankommt.

Berechnen Sie die Wahrscheinlichkeit, dass sie pünktlich in Herrn Fröhlichs Wohnung eintrifft.

a) 0,0875
b) 0,6
c) 0,0097
d) 0,0975
e) 0,975

Antwort: e)

107. Gegeben sei die folgende Stichprobe eines metrischen Merkmals;

5 1 100 25 125 10

Wie groß ist die Summe des Mittelwerts, Median, 1.es und 3.es Quartil?

a) 154
b) 145
c) 145,5
d) 166,8
e) 192,4

Antwort: d) 44,3 + 17,5 + 5 + 100 = 166,8

108. Ein Student arbeitet an einer Studie mit. Er erhält einen Datensatz zur Statistischen Analyse mit Messwerten eines medizinischen Indikators für Stoffwechselbedingtes Augenleiden. Der Datensatz enthält Messwerte von 173 Patienten. Dieser Indikator ist metrisch und sei weiters für das linke und das rechte Auge jedes Patienten gemessen worden. Die Fragestellung lautet, ob es Rechts-Links Unterschiede bezüglich der Indikatorwerte gibt;

a) Regressionsanalyse, speziell Prognose
b) Chi^2 Test
c) T-Test für abhängige (verbundene) Stichproben
d) T-Test für unabhängige Stichproben
e) Berechnung des Korrelationskoeffizienten nach Pearson

Antwort: c)