Vorwort

Im folgenden werden einige Aufgaben zum Lehrbuch „Basiswissen – Medizinische Statistik" zusammen mit kommentierten Lösungen präsentiert. Der Leser soll damit die Möglichkeit haben, sein Verständnis zu überprüfen und sein Wissen zu vertiefen. Außerdem kann er (oder sie) erkennen, dass die erworbenen Kenntnisse keineswegs nur theoretischer Natur sind, sondern sich vor allem bei praktischen Fragestellungen anwenden lassen. Der Autorin bietet sich dabei die Gelegenheit, ergänzend zum Buch auf wichtige Zusammenhänge, Besonderheiten, beliebte Irrtümer etc. hinzuweisen.

Die Aufgaben sind wie bei einer Biomathe-Klausur gestaltet: es werden 5 Antworten vorgegeben, von denen genau eine richtig ist. Für jedes Buchkapitel 2 bis 11 werden jeweils 10 bis 15 Aufgaben aufgelistet, daran anschließend (farblich hervorgehoben) die dazugehörenden Lösungen. Leider kann zwar nicht garantiert werden, dass ein Student, nachdem er die Aufgaben gelöst und inhaltlich verstanden hat, die Klausur mit Bravour besteht – er hat aber mit Sicherheit eine gute Ausgangsposition. Die mit einem * gekennzeichneten Aufgaben beinhalten Themen, die über die Anforderungen des Gegenstandskatalogs hinausgehen.

Übrigens: für Hinweise, Kritik und Verbesserungsvorschläge bin ich jederzeit dankbar. Sie können mich erreichen über meine Adresse (Klinikum Mannheim, Institut für Medizinische Statistik, 68135 Mannheim) oder per E-Mail (christel.weiss@urz.uni-heidelberg.de).

Mannheim, im April 2000

Christel Weiß

1 Inhaltsverzeichnis der Aufgabenseiten

Teil I: Deskriptive Statistik

2 Theoretische Grundlagen

Aufgabe 2.1: Beobachtungseinheiten

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

A. die Hebammen, die den Geburtsvorgang 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

Aufgabe 2.2: Merkmalseigenschaften ASA-Gruppe

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

A. Alternativmerkmal
B. qualitativ, nur nominal skaliert
C. qualitativ, ordinal skaliert
D. quantitativ, diskret
E. quantitativ, stetig

Aufgabe 2.3: Einfluss- und Zielgrößen

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

das Geburtsgewicht des Kindes,
das Alter der Mutter,
das Körpergewicht der Mutter vor der Schwangerschaft

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

1 Anzahl der pro Tag gerauchten Zigaretten	
2 Alter der Mutter	a Faktor
3 Gewicht der Mutter vor der Schwangerschaft	b Begleitmerkmal
4 psychische oder soziale Belastungen der Mutter	c Störgröße

5 Ernährungsweise der Mutter d Zielgröße
6 Geburtsgewicht des Kindes

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

Aufgabe 2.4: Merkmale des Blutes und deren Skalenniveaus

Bitte ordnen Sie den folgenden Merkmalen, die das Blut betreffen, die dazugehörenden Skalen (mit dem jeweils höchsten Niveau) zu.

1. spezifisches Gewicht	a. Nominalskala
2. Senkungsgeschwindigkeit	b. Ordinalskala
3. Anzahl der Thrombozyten pro µl	c. Intervallskala
4. Blutgruppe	d. Verhältnisskala
5. Temperatur in Celsius-Graden	

A. 1d, 2d, 3d, 4a, 5d
B. 1d, 2d, 3d, 4a, 5c
C. 1d, 2c, 3b, 4a, 5a
D. 1d, 2d, 3c, 4b, 5c
E. 1c, 2c, 3b, 4c, 5a

Aufgabe 2.5: Diskrete und stetige Merkmale des Blutes

Welche der im folgenden aufgelisteten Merkmale sind diskret ?

1. pH-Wert
2. Rhesusfaktor
3. Blutgruppe
4. Hämatokrit
5. Anzahl der Leukozyten pro µl Blut
6. Erkrankung an Leukämie (mit den Ausprägungen ja / nein)

A. kein Merkmal ist diskret
B. alle Merkmale sind diskret
C. nur 2 und 3 sind diskret
D. nur 2, 3, 5 und 6 sind diskret
E. alle außer 4 sind diskret

Aufgabe 2.6: Eigenschaften stetiger Merkmale

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	——
D.	falsch	richtig	——
E.	falsch	falsch	

Aufgabe 2.7: Skalentransformationen allgemein

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.

Aufgabe 2.8: Skalentransformation beim Eiweißgehalt

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 ***nicht*** zu?

A. Theoretisch wird eine Reduktion von einer metrischen Skala auf eine Nominalskala durchgeführt.
B. Der Eiweißgehalt wird als Alternativmerkmal erfasst.
C. Die Messmethode ist in jedem Fall sinnlos, da sehr viel Information verloren geht.
D. Das Messverfahren ist einfacher als die exakte Messung in mg/dl.
E. Die Ergebnisse dieses Messverfahrens ermöglichen weniger differenzierte Auswertungen als die exakten Messwerte in mg/dl.

Aufgabe 2.9: Ausprägungsliste mit Laborwerten

In einer Klinik werden bei Patienten, die planmäßig operiert werden, präoperativ Laborwerte erfasst und in einem EDV-System wie folgt dokumentiert:

0: es liegen keine Laborwerte vor
1: alle Werte sind normal
2: Blutwerte pathologisch
3: Gerinnungswerte pathologisch
4: Säure-Basen-Haushalt pathologisch
5: andere Werte pathologisch

Falls mehrere Ausprägungen zutreffen, wird die Summe notiert. Diese Ausprägungsliste ist:

A. unzulässig, da qualitative Merkmale nicht numerisch codiert werden dürfen
B. weder vollständig noch disjunkt
C. vollständig und disjunkt
D. disjunkt, aber nicht vollständig
E. vollständig, aber nicht disjunkt

Aufgabe 2.10: Ausprägungsliste mit klassierten Daten

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: $X \leq 150\,\text{cm}$
1: $150\,\text{cm} \leq X \leq 160\,\text{cm}$
2: $160\,\text{cm} \leq X \leq 170\,\text{cm}$
3: $170\,\text{cm} \leq X \leq 180\,\text{cm}$
4: $180\,\text{cm} \leq X \leq 190\,\text{cm}$
5: $190\,\text{cm} \leq X \leq 200\,\text{cm}$
6: $X \geq 200\,\text{cm}$

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.

Aufgabe 2.1: **Beobachtungseinheiten**

Lösung: ***D*** (die Frauen, die das Medikament erhalten)

Beobachtungseinheiten sind die Personen oder Objekte, die ***beobachtet werden*** (also die 50 Frauen; sie bilden die Stichprobe) und ***nicht*** die Personen, die beobachten (also die Hebammen oder Ärzte, Antworten A und C sind deshalb falsch). Die Messwerte für die Geschwindigkeit, mit denen sich der Muttermund öffnet, sind die Merkmalsausprägungen (Antwort B). Alle Frauen, denen das Medikament theoretisch helfen könnte, bilden die Grundgesamtheit (Antwort E).

Aufgabe 2.2: **Merkmalseigenschaften ASA-Gruppe**

Lösung: ***C*** (qualitativ, ordinal skaliert)

Die Ausprägungen ASA I bis ASA V lassen sich in sinnvoller Weise in einer Reihenfolge anordnen; deshalb ist dieses Merkmal ordinal skaliert. Dieses Merkmal hat 5 Ausprägungen; es ist also kein Alternativmerkmal (das hat nur 2 Ausprägungen). Antwort A ist deshalb falsch. Da die Ordinalskala ein höheres Niveau hat als die Nominalskala, ist auch die Antwort B falsch. – Das ASA-Merkmal ist zwar diskret, aber nicht quantitativ (der Abstand zwischen 2 Ausprägungen ist nicht definiert). Daher sind die Antwort D und (erst recht) E falsch.

Aufgabe 2.3: **Einfluss- und Zielgrößen**

Lösung: ***A***

Es soll der Einfluss des Zigarettenkonsums auf das Geburtsgewicht der Babys untersucht werden – also ist das Merkmal 1 ein Faktor und das Merkmal 6 eine Zielgröße (1a und 6d). Möglicherweise beeinflussen auch das Alter der Mutter und deren Gewicht vor der Schwangerschaft das Geburtsgewicht des Kindes – da diese beiden Merkmale zwar erfasst, aber nicht ausgewertet werden, handelt es sich um Begleitmerkmale (2b und 3b). Die Merkmale 4 und 5 beeinflussen auch die Zielgröße; sie werden jedoch hier nicht berücksichtigt und sind deshalb als unverzerrende Störgrößen aufzufassen (4c und 5c).

Aufgabe 2.4: **Merkmale des Blutes und deren Skalenniveaus**

Lösung: ***B***

Die ersten 3 Merkmale sind verhältnisskaliert (1d, 2d, 3d); man kann nämlich 2 Ausprägungen eines Merkmals in ein Verhältnis setzen wie etwa: 100.000 Thrombozyten pro µl Blut sind doppelt so viele wie 50.000. Die Blutgruppe ist nur nominalskaliert (4a). Die Temperatur in Celius-Graden ist nur intervallskaliert (5c). Siehe auch Beispiel 2.2 im Buch auf Seite 22.

Aufgabe 2.5: **Diskrete und stetige Merkmale des Blutes**

Lösung: ***D***

Diskret heißt: das Merkmal hat nur abzählbar viele Ausprägungen. Dazu zählen alle qualitativen Merkmale; also hier: Rhesusfaktor (2), Blutgruppe (3), Erkrankung an Leukämie (6). Das Merkmal „Anzahl der Leukozyten pro µl Blut" hat zwar sehr viele Ausprägungen, aber nur ganzzahlige und ist deshalb ebenfalls diskret. – Die beiden anderen Merkmale (pH-Wert, Hämatokrit) können innerhalb eines bestimmten Bereichs theoretisch jeden Wert annehmen und sind deshalb quantitativ stetig.

Aufgabe 2.6: **Eigenschaften stetiger Merkmale**

Lösung: ***B*** (beide Aussagen richtig, Verknüpfung falsch)

Der Glukosegehalt ist ein quantitativ-stetiges Merkmal (also nicht qualitativ). Demnach sind die Aussagen 1 und 2 korrekt. Die Verknüpfung ist allerdings falsch. Aus der Tatsache, dass ein Merkmal nicht qualitativ (also quantitativ) ist, folgt nicht automatisch, dass es auch stetig ist. Es gibt auch quantitativ-diskrete Merkmale.

Aufgabe 2.7: **Skalentransformationen allgemein**

Lösung: ***E*** (Verhältnisskala → Ordinalskala)

Generell kann eine Skala nur auf eine andere Skala mit niedrigerem Niveau transformiert werden. Daher sind die Antworten A und C falsch, und Antwort E ist richtig (die Ordinalskala hat ein niedrigeres Niveau als die Verhältnisskala). – Ein stetiges Merkmal ist immer quantitativ; also ist Antwort B falsch. Antwort D ist ebenfalls falsch: die Transformation ist nicht immer möglich (eine Nominalskala kann nicht transformiert werden). Ob sie sinnvoll ist, hängt von den speziellen Rahmenbedingungen ab (siehe auch Seite 25 oben im Buch).

Aufgabe 2.8: **Skalentransformation beim Eiweißgehalt**

Lösung: ***C*** (die Aussage trifft ***nicht*** zu)

Vorsicht ! Gesucht ist hier die Aussage, die ***falsch*** ist. Das ist etwas verwirrend, aber solche Aufgaben findet man hin und wieder in Klausuren. – Wenn man das Eiweiß in mg/dl missst, hat man ein verhältnisskaliertes Merkmal. Wenn man es nur mit den Ausprägungen pathologisch (ja/nein) erfasst, hat man ein Alternativmerkmal (also ein nominalskaliertes). Antworten A und B sind also richtig. Auch die Aussage D ist leicht nachvollziehbar. Da bei der einfachen Messmethode sehr viel Information verloren geht, kann man die Daten natürlich auch nicht so gut analysieren – deshalb ist die Antwort E korrekt. Falsch ist dagegen Antwort C. Ein einfacheres Messverfahren kann durchaus einen Informationsverlust rechtfertigen; im konkreten Einzelfall ist abzuwägen, was wichtiger und der konkreten Fragestellung angemessen ist.

Aufgabe 2.9: Ausprägungsliste mit Laborwerten

Lösung: ***E*** (vollständig, aber nicht disjunkt)

Antwort A ist natürlich unsinnig, denn selbstverständlich dürfen qualitative Merkmale numerisch codiert werden (wenn die Ausprägungsliste vollständig und disjunkt ist). – Es ist klar, dass die Liste vollständig ist; wegen der Angabe 5 (sonstige path. Werte) kann alles erfasst werden. Sie ist aber nicht disjunkt, weil 2 Ausprägungen nicht unbedingt unterscheidbar sind. So kann z. B. die Codierung 5 bedeuten: „Blutwerte und Gerinnungswerte pathologisch" oder „andere Werte pathologisch". Dieses Problem kann man umgehen, indem man statt der Zahlen 1–5 die 2er Potenzen 1, 2, 4, 8 und 16 zur Codierung verwendet.

Aufgabe 2.10: Ausprägungsliste mit klassierten Daten

Lösung: ***D*** (vollständig, aber nicht disjunkt)

Die Aussage A trifft so allgemein nicht zu. Ob diese unpräzise Erfassung brauchbar ist oder nicht, hängt von der speziellen Fragestellung ab. – Die Liste ist vollständig, denn mit den Codierungen 0 und 6 können auch die extremsten Körpergrößen erfasst werden. Allerdings ist sie nicht disjunkt – bei den Körpergrößen 150 cm, 160 cm etc. ist eine eindeutige Zuordnung zu einer Codierung nicht möglich. Die Codierungen schließen sich gegenseitig ***nicht*** aus. Bei allen Klasseneinteilungen muss deshalb darauf geachtet werden, dass Messwerte, die auf eine Klassengrenze fallen, eindeutig zuordenbar sind.

3 Univariate Datenbeschreibung

Aufgabe 3.1: Balkendiagramm

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

Aufgabe 3.2: Relative Häufigkeiten – Darstellung

Im Rahmen einer klinischen Studie wird die Wirksamkeit einer therapeutischen Maßnahme an 22 Patienten untersucht. Bei $n_1 = 14$ Patienten ist die Therapie erfolgreich. Welche Darstellung der entsprechenden relativen Häufigkeit ist am sinnvollsten?

A. $h_1 = 64\%$
B. $h_1 = 0{,}63636$
C. $h_1 = 14/22$
D. h_1 liegt über 50 %.
E. h_1 beträgt zwischen 60 % und 70 %.

Aufgabe 3.3: Relative Häufigkeiten – Blutdruck

In einer Stichprobe von $n = 200$ Personen werden die in der Tabelle aufgelisteten Blutdruckwerte (in mmHg) 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

Aufgabe 3.4: Klasseneinteilung – allgemein

Welche Aussage bzgl. klassierter Daten ist ***falsch***?

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 Historgramm sind charakteristische Eigenschaften der Merkmalsvertei- lung (Lage, Streuung, Verteilungsform) erkennbar.

Aufgabe 3.5: Klasseneinteilung – Berechnung des Mittelwerts

Die Daten zur Körpergröße von männlichen 50 Studenten sei in folgende Klassen eingeteilt:

Klasse	Grenzen	absolute Häufigkeiten
1	$\leq 160\,\text{cm}$	$n_1 = 2$
2	$(160\,\text{cm}, 170\,\text{cm}]$	$n_2 = 18$
3	$(180\,\text{cm}, 190\,\text{cm}]$	$n_3 = 20$
4	$> 190\,\text{cm}$	$n_4 = 10$

Welche Aussage bzgl. des Mittelwerts trifft zu?

A. Man kann aus diesen Angaben einen Mittelwert ermitteln und erhält dabei genau den selben Wert, der sich aus den nicht-klassierten Originalmesswerten ergibt.
B. Man kann aus diesen Angaben einen Mittelwert berechnen, der jedoch mit einer Ungenauigkeit behaftet ist.
C. Die Berechnung eines Mittelwerts ist wegen der offenen Restklassen nicht möglich.
D. Die Berechnung eines Mittelwerts ist möglich, indem man als Klassenmitten für die Restklassen 155 cm bzw. 195 cm ansetzt.
E. Die Berechnung eines Mittelwerts ist nicht möglich, da die Klassen unterschiedlich breit sind.

Aufgabe 3.6: Empirische Verteilungsfunktion – allgemein

Welche Aussage bzgl. der empirischen Verteilungsfunktion $F(x)$ ist ***falsch***?

A. $F(x)$ ist für alle $x \in \left]-\infty,+\infty\right[$ definiert.
B. Zwischen dem kleinsten und dem größten Stichprobenwert wächst $F(x)$ monoton von 0 bis 1.
C. Ein Funktionswert $F(x)$ gibt den relativen Anteil der Beobachtungen an, die kleiner oder gleich x sind.
D. Ein Funktionswert $F(x)$ gibt den relativen Anteil der Beobachtungen an, die größer oder gleich x sind.
E. Es gilt: $F(x) = 1$ für alle $x \geq x_{\max}$.

Aufgabe 3.7: Empirische Verteilungsfunktion bei klassierten Daten

Aus klassierten Daten wird die empirische Verteilungsfunktion bestimmt. Für ein bestimmtes x_1 gelte: $F(x_1) = 0{,}25$. Worum handelt es sich bei diesem x_1?

A. Die relative Häufigkeit von x_1 beträgt 25%.
B. x_1 ist das empirisch ermittelte, untere Quartil.
C. x_1 ist das empirisch ermittelte, obere Quartil.
D. Dem x_1 kann keine dieser Eigenschaften zugewiesen werden.
E. Bei klassierten Daten kann die Verteilungsfunktion nicht bestimmt werden.

Aufgabe 3.8: Vergleich Mittelwert und Median

Welche der folgenden Aussagen ist ***falsch***?

A. Der Mittelwert wird wesentlich stärker von Ausreißern beeinflusst als der Median.
B. Die Berechnung des Mittelwerts setzt ein quantitatives Merkmal voraus.
C. Der Mittelwert und der Median sind Lagemaße.
D. Bei schiefen Verteilungen weichen der Mittelwert und der Median voneinander ab.
E. Wenn die Berechnung des Medians erlaubt ist, kann auch der Mittelwert berechnet werden.

Aufgabe 3.9: Bestimmung des empirischen Medians

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:

Anzahl Punkte	≤ 4	5	6	7	8	9	10
Häufigkeit	12	9	14	18	10	4	2

Der empirische Median ist:

A. 6
B. 6,5
C. 7
D. 14
E. nicht bestimmbar

Aufgabe 3.10: Eigenschaften des Medians

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

A. alle Werte außerhalb des Intervalls $\bar{x} \pm 2s$ 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

Aufgabe 3.11: Lage- und Streuungsmaße

Im folgenden sind insgesamt 8 Lage- und Streuungsmaße aufgelistet:

a.	Varianz	d.	Modus
b.	Spannweite	e.	Standardabweichung
c.	Variationskoeffizient	f.	Minimum
d.	Quartilsabstand	g.	Maximum

Gefragt ist nach der Anzahl der Streuungsmaße.

A. 3
B. 4
C. 5
D. 6
E. 7

Aufgabe 3.12: Maßzahlen für die ASA-Risikogruppe

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	d.	Varianz
b.	Median	e.	Standardabweichung

c. Modus f. Spannweite

A. alle angegebenen Maßzahlen können berechnet werden
B. nur a, d und e
C. nur c
D. nur b, c und f
E. nur a und b

Aufgabe 3.13: Maßzahlen für die Aufenthaltsdauer

Bei jedem Patienten, der mit einer bestimmten Diagnose in eine Klinik eingeliefert wird, wird die Aufenthaltsdauer (in Tagen) ermittelt. Welche Maßzahlen lassen sich bei diesem Merkmal berechnen?

a. Mittelwert d. Standardabweichung
b. Median e. Variationskoeffizient
c. Modus f. Spannweite

A. alle angegebenen Maßzahlen können berechnet werden
B. nur a und d
C. nur a, d und e
D. nur c und f
E. nur a, b und d

Aufgabe 3.14: Maßzahlen und Stichprobenumfang

Aus einer Stichprobe vom Umfang $n = 10$ ermittelt man für ein quantitatives Merkmal den Mittelwert, den Median, die Spannweite sowie die Varianz. Danach wählt man aus derselben Grundgesamtheit weitere 10 Beobachtungseinheiten und berechnet die angegebenen Maßzahlen aus der größeren Stichprobe des Umfangs $n = 20$. Welche Maßzahlen können dabei in keinem Fall kleiner werden?

1. Mittelwert
2. Median
3. Spannweite
4. Summe der Abweichungsquadrate vom Mittelwert (Zähler der Varianz)
5. Varianz

A. keine der Maßzahlen 1–5 kann kleiner werden
B. nur 3 und 4 können nicht kleiner werden
C. nur 3, 4 und 5 können nicht kleiner werden
D. alle Maßzahlen können kleiner werden
E. diese Frage ist abhängig von dem Skalenniveau des Merkmals

Aufgabe 3.1: Balkendiagramm

Lösung: ***E*** (alle diskreten Merkmale)

Die Antworten A und B sind falsch: Voraussetzung für ein Balkendiagramm ist ein diskretes Merkmal (bei stetigen Merkmalen mit klassierten Daten verwendet man statt dessen ein Histogramm). Das Merkmal, das durch ein Balkendiagramm dargestellt wird, kann qualitativ (nominalskaliert oder ordinalskaliert) oder auch quantitativ-diskret sein. Die Aussagen der Antworten C und D sind daher zu eingeschränkt und somit falsch. – Bei quantitativen Merkmalen sind sowohl die Reihenfolge der Balken als auch deren Abstand von Bedeutung; bei ordinalen Merkmalen ist nur die Reihenfolge interessant.

Aufgabe 3.2: Relative Häufigkeiten – Darstellung

Lösung: ***C*** (14/22)

Die richtige Lösung anzugeben mag für manchen Leser schwierig sein, denn eigentlich ist keine einzige der Antworten A–E gänzlich falsch. Gefragt ist jedoch nicht nach einer ***richtigen*** Häufigkeitsangabe, sondern nach einer ***sinnvollen***. Bei 22 Beobachtungseinheiten würde man mit einer Prozentangabe oder einer Häufigkeit mit 5 Dezimalstellen eine Genauigkeit vortäuschen, die nicht vorhanden ist. Deshalb sind die Antworten A und B nicht sinnvoll. Die Antworten D und E sind zwar richtig, aber zu unpräzise. Die Angabe C ist dagegen präzise – und verheimlicht dennoch nicht, dass die Berechnung der Häufigkeit auf einer relativ kleinen Anzahl von Beobachtungseinheiten basiert.

Aufgabe 3.3: Relative Häufigkeiten – Blutgruppe

Lösung: ***E*** (39/200)

Das Wort „oder“ wird bei statistischen Berechnungen im nicht-ausschließlichen Sinn gebraucht. Die 39 Patienten mit Hypertonie teilen sich auf in 23 Patienten (systolischer Blutdruck > 150 mmHg, diastolischer normal), einen Patienten (diastolischer Blutdruck > 100 mmHg, systolischer normal) und 15 Patienten (diastolischer Blutdruck > 100 mgHg ***und*** systolischer > 150 mmHg). Man muss freilich aufpassen, dass man die 15 Patienten, die beide Kriterien bzgl. Hypertonie erfüllen, nicht doppelt zählt – damit erhielte man das falsche Ergebnis in Antwort D.

Aufgabe 3.4: Klasseneinteilung – allgemein

Lösung: ***D*** (Aussage ist ***falsch***)

Vorsicht: Hier ist nach einer falschen Aussage gefragt – und dies ist Aussage D. Es ist zwar rechentechnisch günstig und übersichtlich, wenn die Klassen gleich breit sind – dies ist aber keine unabdingbare Voraussetzung bei einer Klasseneinteilung. Manchmal (etwa bei Ausreißern oder bei Randwerten, die nicht genau erfasst sind) werden auch breite oder halboffene Randklassen verwendet. – Die Aussage A ist richtig (aus qualitativen Daten kann man keine Klassen bilden), ebenso B (die optimale Klassenanzahl ist ungefähr $\sqrt{n}$) und C (wegen der übersichtlichen Darstellung wird die Klasseneinteilung ja vorgenommen). Auch die Aussage E ist korrekt: einen ersten Eindruck bzgl. der Häufigkeitsverteilung gewinnt man über das Histogramm. Siehe dazu auch Seite 38 ff. im Buch.

Aufgabe 3.5: **Klasseneinteilung – Berechnung des Mittelwerts**

Lösung: ***C*** (Berechnung des Mittelwerts nicht möglich)

Die Berechnung eines Mittelwerts (ebenso der Varianz) aus klassierten Daten ist nur möglich, wenn keine offenen Restklassen vorliegen. Deshalb ist C richtig, A und B sind falsch. Es ist auch nicht sinnvoll, irgendwelche Klassenmitten anzunehmen (daher ist D falsch). – Nur bei abgeschlossenen Klassen ist eine Abschätzung des Mittelwerts aus den Klassenmitten möglich (auch wenn diese unterschiedlich breit sind, E ist falsch). Dieser Wert wäre allerdings etwas ungenau (d. h. wenn keine offenen Restklassen vorlägen, wäre B die richtige Antwort).

Aufgabe 3.6: **Empirische Verteilungsfunktion – allgemein**

Lösung: ***D*** (Aussage ist ***falsch***)

Die Aussagen A, B und E folgen sofort aus der Definition nach Formel (3.7) (Seite 36, diskrete Merkmale) bzw. nach Formel (3.9) (Seite 41, stetige Merkmale). Nach der Definition der empirischen Verteilungsfunktion ist auch Antwort C richtig, Aussage D muss dann falsch sein.

Aufgabe 3.7: **Empirische Verteilungsfunktion bei klassierten Daten**

Lösung: ***B*** (x_1 ist unteres Quartil)

Offensichtlicher Unsinn ist die Antwort E. Selbstverständlich kann man bei klassierten Daten die Verteilungsfunktionen $F(x)$ bestimmen, und zwar nach Formel (3.9) auf Seite 41 im Buch. Diese Funktion gibt keine relativen Häufigkeiten an (deshalb ist Antwort A falsch), sondern kumulative (also aufaddierte) Häufigkeiten. Wenn also gilt: $F(x_1) = 0{,}25$, dann heißt dies, dass 25% der Stichprobenwerte kleiner oder gleich x_1 sind – demnach ist x_1 das untere Quartil (Antwort B). Für das obere Quartil (Antwort C) gilt: $F(x_3) = 0{,}75$. „Empirisch ermittelt" bedeutet hier: das Quartil wird aufgrund von Stichprobenwerten bestimmt.

Aufgabe 3.8: **Vergleich Mittelwert und Median**

Lösung: ***E*** (Aussage ist ***falsch***)

Der Mittelwert wird von Ausreißern stark beeinflusst, während Ausreißer bei der Berechnung des Medians kaum eine Rolle spielen (siehe Bsp. 3.10, Seite 48); Antwort A ist also richtig. Generell können bei quantitativen Merkmalen der Mittelwert und der Median als Lagemaße berechnet werden (Antworten B und C). Wenn diese beiden Maße voneinander abweichen, ist die Stichprobenverteilung schief (Antwort D). – Bei ordinalskalierten Merkmalen kann der Median berechnet werden, der Mittelwert dagegen nicht. Deshalb ist Antwort E falsch. Achtung: vielfach wird bei ordinalskalierten Daten fälschlicherweise der Mittelwert angegeben – z. B. bei Schulnoten. Dann werden von den Daten Informationen ausgewertet, die diese gar nicht enthalten. Siehe auch Bsp. 3.9 auf Seite 48.

Aufgabe 3.9: **Bestimmung des empirischen Medians**

Lösung: ***A*** (Median ist 6)

Der Stichprobenumfang ist $n = 69$, also ungerade. Mit Formel (3.15) ergibt sich für den Median der Stichprobenwert mit dem Rang 35; dies ist 6. Entscheidend ist also nur der mittlere Wert. Der Mittelwert könnte aus diesen Angaben nicht berechnet werden, da dafür ***alle*** Stichprobenwerte erforderlich sind. Aus der Häufigkeitstabelle ist jedoch nur ersichtlich, dass 12 Studenten durchgefallen sind, es ist aber nicht zu entnehmen, wie viele Studenten 0, 1, 2, 3 oder 4 Punkte erreicht haben.

Aufgabe 3.10: **Eigenschaften des Medians**

Lösung: ***B*** (zum größten Wert kann man eine Zahl addieren)

Wenn man zum größten Wert eine positive Zahl addiert, bleibt dies der größte Wert. Die Rangzahlen und der Median ändern sich dadurch nicht (B ist richtig). Wenn man jedoch einen oder mehrere Werte aus der Stichprobe entfernt, ändert sich deren Umfang und damit eventuell auch der Median (Antworten A und E sind deshalb falsch). Wenn man alle Werte mit der gleichen Zahl multipliziert oder zu allen Werten eine Zahl addiert, ändert sich der Median in der gleichen Weise (obgleich dessen Rang unverändert bleibt). Deshalb sind auch die Antworten C und D falsch.

Aufgabe 3.11: **Lage- und Streuungsmaße**

Lösung: ***C*** (5 Streuungsmaße)

Streuungsmaße sind: die Varianz, die Spannweite, der Variationskoeffizient, der Quartilsabstand und die Standardabweichung – sie beschreiben die Variabilität einer Stichprobe. Der Modus, das Maximum und das Minimum sind dagegen Lagemaße.

Aufgabe 3.12: Maßzahlen für die ASA-Risikogruppe

Lösung: ***D*** (Median, Modus und Spannweite)

Es handelt sich bei der ASA-Risikogruppe um ein ordinal skaliertes Merkmal (siehe Aufgabe 2.2, Seite 1). Bei dieser Merkmalsart können als Lagemaße nur der Median und der Modus bestimmt werden (b und c), nicht jedoch der Mittelwert (siehe auch Aufgabe 3.8). Als Streuungsmaß eignet sich nur die Spannweite (f); die Varianz und die Standardabweichung setzen metrische Merkmale voraus.

Aufgabe 3.13: Maßzahlen für die Aufenthaltsdauer

Lösung: ***A*** (alle Maßzahlen können berechnet werden)

Es handelt sich bei der Aufenthaltsdauer um ein quantitativ-diskretes Merkmal mit höchstem Niveau (Verhältnisskala). Deshalb können theoretisch alle Maßzahlen berechnet werden. Ob diese Angaben immer sinnvoll sind, ist eine andere Frage. Oft werden bei metrischen Merkmalen nur der Mittelwert und die Standardabweichung angegeben; die anderen Maßzahlen nur bei besonderen Verteilungsformen oder Fragestellungen. Unbeliebt ist vor allem die Spannweite, die nur die beiden extremsten Werte berücksichtigt und daher in der Regel ein verzerrtes Bild der Variabilität wiedergibt.

Aufgabe 3.14: Maßzahlen und Stichprobenumfang

Lösung: ***B*** (Spannweite und Summe der Abweichungsquadrate)

Es ist einleuchtend, dass eine größere Stichprobe bessere Schätzungen ermöglicht als eine kleinere. Das bedeutet: die Maßzahlen Mittelwert, Median und Varianz liegen erwartungsgemäß „näher" am entsprechenden Parameter der Grundgesamtheit. „Näher" bedeutet dabei freilich nicht, dass z. B. der Mittelwert generell größer (oder generell kleiner) wird. Anders verhält es sich bei der Spannweite: wenn die neuen 10 Beobachtungseinheiten ein größeres Maximum (oder ein kleineres Minimum) haben als die ersten 10, vergrößert sich die Spannweite – ansonsten bleibt sie gleich. Da Abweichungsquadrate niemals negativ sind, kann sich auch durch eine Erhöhung des Stichprobenumfangs deren Summe nicht verkleinern – höchstens die Varianz, bei der durch den Nenner $(n-1)$ dividiert wird.

4. Bivariate Datenbeschreibung

Aufgabe 4.1: Punktwolke

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

A. ob der Zusammenhang 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

Aufgabe 4.2: Kovarianz

Bei welchen Skalenniveaus ist die Berechnung der Kovarianz erlaubt?

A. Bei 2 Merkmalen mit beliebigem Skalenniveau
B. Beide Merkmale müssen verhältnisskaliert sein.
C. Es genügt, wenn beide Merkmale ordinalskaliert sind.
D. Es genügt, wenn mindestens 1 Merkmal metrisch skaliert ist.
E. Beide Merkmale müssen metrisch skaliert sein.

Aufgabe 4.3: Wertebereich des Korrelationskoeffizienten

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

A. $\left]-\infty,+\infty\right[$
B. $\left[0,1\right]$
C. $\left[0,+\infty\right[$
D. $\left]-1,+1\right[$
E. $\left[-1,+1\right]$

Aufgabe 4.4: Interpretation eines Korrelationskoeffizienten

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.

Aufgabe 4.5:* Wertebereich des Bestimmtheitsmaßes

Welches Intervall umfasst alle Zahlenwerte, die das Bestimmtheitsmaß annehmen kann?

A. $]-\infty,+\infty[$
B. $[0,1]$
C. $[-1,+1]$
D. $]-1,+1[$
E. $[0,+\infty[$

Aufgabe 4.6: Regressionsgerade

Eine Regressionsgerade habe die Form: $y = -0{,}8 + 0{,}3 \cdot x$. 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.

Aufgabe 4.7: Regressionskoeffizient

Der Regressionskoeffizient ist

A. die Steigung der Regressionsgeraden
B. der Schwerpunkt der Punktwolke
C. der Schnittpunkt der Regressionsgeraden mit der y-Achse
D. immer eine Zahl zwischen 0 und 1
E. ein Punkt auf der Regressionsgeraden

Aufgabe 4.8: Korrelationskoeffizient und Regressionsgerade

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.

Aufgabe 4.9: Vergleich zweier Messverfahren

Eine neue Messmethode Y wird mit einer Referenzmethode X verglichen. Womit lässt sich die Güte der neuen Methode beurteilen?

A. durch den Regressionskoeffizienten
B. durch den Korrelationskoeffizienten
C. durch das Bestimmtheitsmaß
D. durch die Regressionsgerade ***und*** den Korrelationskoeffizienten
E. durch den Schwerpunkt der Punktwolke

Aufgabe 4.10: Prognostizieren mit der Regressionsgeraden

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 $y = 180 - 4 \cdot x$ ermittelt (wobei x die Behandlungstage und y die Blutdruckwerte in mmHg sind). Welche der folgenden Aussagen lassen sich aus diesen Informationen schlussfolgern?

1. Der Blutdruck sinkt während der Therapie um durchschnittlich 0,89 mmHg pro Tag.
2. Der Blutdruck sinkt während der Therapie um durchschnittlich 4 mmHg pro Tag.
3. Der Schätzwert für den letzten Tag der Therapie beträgt 120 mmHg.
4. Zu Beginn der Therapie hatte der Patient einen Blutdruck von etwa 180 mmHg.
5. 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

Aufgabe 4.11: Geeignetes Zusammenhangsmaß

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 Kovarianz
B. der Korrelationskoeffizient nach Pearson
C. der Korrelationskoeffizient nach Spearman
D. das Produkt der beiden Standardabweichungen
E. keine der Angaben aus A–D

Aufgabe 4.12: Wertebereiche

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

Aufgabe 4.13: Zusammenhang zwischen Körpergröße und Gewicht

Beurteilen Sie die folgende Aussagen:

1. Der Korrelationskoeffizient, der den Zusammenhang zwischen Körpergröße und Körpergewicht von männlichen Erwachsenen quantifiziert, ist positiv, 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	——
D.	falsch	richtig	——
E.	falsch	falsch	

Aufgabe 4.14: Abhängiges und unabhängiges Merkmal

Der Zusammenhang zwischen Körpergröße und Gewicht bei erwachsenen Frauen im Alter von 20 bis 40 Jahren soll durch eine Regressionsgleichung beschrieben werden. Welches Merkmal sollte sinnvollerweise als das unabhängige x-Merkmal und welches als das abhängige y-Merkmal aufgefasst werden?

A. Dieser Zusammenhang lässt sich nicht durch eine Regressionsgleichung beschreiben.
B. Es ist vollkommen gleichgültig, welches der beiden Merkmale als abhängig bzw. unabhängig angesehen wird.
C. Das Gewicht sollte als das unabhängige x-Merkmal gewählt werden.
D. Die Größe sollte als das unabhängige x-Merkmal gewählt werden.
E. Man berechnet 2mal den Korrelationskoeffizienten (einmal mit der Größe und einmal mit dem Gewicht als unabhängigem x-Merkmal). Der größere Koeffizient liefert die Entscheidung.

Aufgabe 4.1: **Punktwolke**

Lösung: ***E*** (kausaler Zusammenhang nicht erkennbar)

Die Darstellung 2er Merkmale mittels einer Punktwolke liefert sehr wertvolle Informationen. Man sieht, ob die Punkte um eine Gerade liegen (ob der Zusammenhang linear ist, Antwort A). Ausreißer sind auf einen Blick erkennbar (Antwort B). Wenn die Punkte nahe an der Geraden liegen, ist der Zusammenhang stark, je weiter die Punkte streuen, desto schwächer ist er (Antwort C). Wenn die Geradensteigung positiv ist, ist der Zusammenhang gleichsinnig, bei negativer Steigung gegensinnig (Antwort D). Siehe dazu auch Seiten 72ff. im Buch. Ob und wie die Merkmale kausal zusammenhängen, ist allerdings anhand der Punktwolke nicht nachvollziehbar. Man kann eine wunderschöne Punktwolke erstellen, indem man das Storchenaufkommen und die Geburtenhäufigkeit in Deutschland während der letzten 100 Jahre jeweils als einen Punkt in ein Koordinatenssystem einzeichnet. Der Punktwolke ist es egal, ob der Zusammenhang sachlich begründet und kausal ist, oder ob es sich um eine Nonsens-Korrelation handelt.

Aufgabe 4.2: **Kovarianz**

Lösung: ***E*** (beide Merkmale metrisch)

Die Berechnung der Kovarianz basiert auf den Mittelwerten $\bar{x}$ und $\bar{y}$ (siehe Formel (4.4), Seite 74). Dazu müssen beide Merkmale metrisch sein; deshalb sind Antworten A, C und D falsch. Es genügt, wenn sie intervallskaliert sind; in der Antwort B (verhältnisskaliert) wird zuviel verlangt.

Aufgabe 4.3: **Wertebereich des Korrelationskoeffizienten**

Lösung: ***E*** $[-1,+1]$

Es ist ja gerade die herausragende Eigenschaft des Korrelationskoeffizienten, dass er normiert ist und deshalb nur Werte zwischen –1 und +1 annehmen kann (und damit sehr gut interpretierbar ist). Die Grenzen sind eingeschlossen. Wenn der Zusammenhang funktional ist und exakt durch eine Geradengleichung beschrieben werden kann (was zwar praktisch kaum vorkommt, aber theoretisch möglich ist), beträgt der Korrelationskoeffizient $r = 1$ oder $r = -1$. Bei stochastischen Zusammenhängen hat er einen Betrag, der kleiner als 1 ist.

Aufgabe 4.4: **Interpretation eines Korrelationskoeffizienten**

Lösung: ***C*** $r = -0{,}95$

Barer Unsinn ist die Aussage E: dieses Maß ist sehr gut geeignet, um einen linearen Zusammenhang zu beschreiben. Auch die Aussage D erweist sich sofort als falsch: r kann nur

Werte annehmen, deren Betrag höchstens 1 ist. Nun ist bekannt: je näher der Betrag von r bei 1 liegt, desto stärker ist der Zusammenhang. Deshalb ist ein Zusammenhang mit $r = -0{,}95$ stärker als einer mit $r = 0{,}8$ (bzgl. der Stärke spielt das Vorzeichen keine Rolle). Die Antwort A ($r = 0$) bezieht sich auf 2 Merkmale, bei denen überhaupt kein linearer Zusammenhang erkennbar ist.

Aufgabe 4.5:* Wertebereich des Bestimmtheitsmaßes

Lösung: ***B*** $[0{,}1]$

Dazu muss man wissen, dass das Bestimmtheitsmaß durch r^2 quantifiziert wird (Formel (4.12) auf Seite 88). Da sich r zwischen –1 und +1 erstreckt, hat r^2 einen Wertebereich zwischen 0 und 1 (inklusive der Grenzen).

Aufgabe 4.6: Regressionsgerade

Lösung: ***B*** (Zusammenhang ist gleichsinnig)

Wenn die Steigung der Regressionsgerade wie hier mit 0,3 positiv ist, ist der Zusammenhang gleichsinnig (Antwort A falsch, B richtig) – weitergehende Aussagen sind nicht jedoch möglich. Man weiß also nur, dass der Korrelationskoeffizient positiv ist; die Gleichung der Regressionsgeraden enthält keine Aussagen über dessen Betrag.

Aufgabe 4.7: Regressionskoeffizient

Lösung: ***A*** (Steigung der Regressionsgeraden)

Dies ist eine reine Definitionssache (siehe Seite 84). Der Wertebereich des Regressionskoeffizienten ist nicht eingeschränkt, mit anderen Worten: die Steigung der Regressionsgeraden kann beliebig steil oder schwach sein. Sie beschreibt die Art des Zusammenhangs, der Korrelationskoeffizient dessen Stärke.

Aufgabe 4.8: Korrelationskoeffizient und Regressionsgerade

Lösung: ***A*** (Steigung ist positiv)

Der Korrelationskoeffizient und die Steigung der Regressionsgeraden haben dasselbe Vorzeichen (siehe auch Aufgabe 4.6). Wenn man also weiß, dass $r = 0{,}2$, kann man auf Antwort A schließen – spezifischere Aussagen bzgl. der Geradensteigung oder des y-Achsenabschnitts sind jedoch nicht möglich.

Aufgabe 4.9: Vergleich zweier Messverfahren

Lösung: ***D*** (Korrelationskoeffizient und Regressionsgerade)

Von einem neuen Messverfahren erwartet man, dass es dieselben Werte misst wie eine bekannte Referenzmethode (abgesehen von geringen, zufällig bedingten Abweichungen). Dann müsste der Korrelationskoeffizient $r \approx 1$ betragen (er sollte nur wenig kleiner sein). Das allein reicht jedoch nicht aus, um die Güte eines neuen Verfahrens zu bestimmen. Auch bei einem systematischen Fehler könnte sich ein Korrelationskoeffizient nahe bei 1 ergeben. Es ist deshalb wichtig, auch die Regressionsgerade zu ermitteln; für sie müsste gelten: $y \approx x$. Als Beispiel sei folgendes genannt: wenn man die Körpergröße mehrerer Personen mit einem normalen Maßband x und einem um exakt 10 cm verkürzten Maßband y misst, ergibt sich als Regressionsgerade: $y \approx x + 10$ und als Korrelationskoeffizient $r \approx 1$. Am Korrelationskoeffizient ist die systematische Abweichung nicht erkennbar.

Aufgabe 4.10: Prognostizieren mit der Regressionsgeraden

Lösung: ***C*** (Aussagen 2, 3 und 4 sind herleitbar)

Die Regressionsgerade beschreibt die Art des Zusammenhangs, in diesem Fall: der Blutdruck sinkt um 4 mmHg pro Tag (Aussage 2 richtig). Der Korrelationskoeffizient $r = -0{,}89$ beschreibt die Stärke des Zusammenhangs, nicht dessen Art (deshalb ist Aussage 1 falsch). Man erhält über die Gleichung der Regressionsgeraden für $x = 15$:
$y = (180 - 4 \cdot 15)\,\text{mmHg} = 120\,\text{mmHg}$ (Aussage 3). Zu Beginn der Therapie ergibt sich mit $x = 0$: $y = 180\,\text{mmHg}$ (Aussage 4). Es ist aber nicht erlaubt, über den Beobachtungsbereich hinaus zu extrapolieren. Die 5. Aussage ist deshalb nicht mehr herleitbar. Eine solche Vorgehensweise würde total unsinnige Werte liefern (man setze mal spaßeshalber in die Regressionsgerade ein: $x = 50$).

Aufgabe 4.11: Geeignetes Zusammenhangsmaß

Lösung: ***C*** (Korrelationskoeffizient nach Spearman)

Wenn man nicht genau weiß, ob ein Zusammenhang zwischen 2 metrischen Merkmalen linear ist, sollte man weder die Kovarianz noch den Korrelationskoeffizienten nach Pearson (der ja auf der Kovarinz basiert) berechnen. Diese beiden Maßzahlen eignen sich nur zur Darstellung eines linearen Zusammenhangs. Die Antworten A und B sind somit falsch, ebenso die Antwort D (das Produkt der Standardabweichungen ist generell kein Maß für die Stärke eines Zusammenhangs). Einen Ausweg liefert der Korrelationskoeffizient nach Spearman, der schwächere Voraussetzungen hat (der Zusammenhang muss nur monoton, nicht unbedingt linear sein).

Aufgabe 4.12: Wertebereiche

Lösung: ***D*** (4 Maßzahlen sind nicht negativ)

Diese sind: Spannweite, Varianz, Standardabweichung und Bestimmtheitsmaß. Die ersten 3 sind nur dann 0, wenn alle Stichprobenwerte übereinstimmen und ansonsten positiv. Das ergibt sich aus der Definition dieser Kenngrößen. Das Bestimmtheitsmaß ist das Quadrat r^2 und kann deshalb niemals negativ sein. Dagegen können – wenn die Stichprobe negative Werte enthält – das Minimum, das Maximum, der Modus und der Median auch negativ sein. Der Korrelationskoeffizient und die Kovarianz sind negativ, falls der Zusammenhang gegensinnig ist.

Aufgabe 4.13: Zusammenhang zwischen Körpergröße und Gewicht

Lösung: ***B*** (beide Aussagen richtig, Verknüpfung falsch)

Der Zusammenhang ist gleichsinnig, demnach ist der Korrelationskoeffizient positiv (große Leute wiegen viel, kleine eher weniger). Die Werte der Körpergröße und des Gewicht sind immer positiv. Also sind die Aussagen 1 und 2 richtig; allerdings nicht deren Verknüpfung. Die Aussage 2 enthält keinerlei Information bzgl. des Korrelationskoeffizienten; dieser kann positiv oder negativ sein (siehe auch Beispiel auf Seite 73 im Buch, Zusammenhang zwischen Isofluran und artieriellem Blutdruck).

Aufgabe 4.14: Abhängiges und unabhängiges Merkmal

Lösung: ***D*** (die Größe sollte das unabhängige Merkmal sein)

Antwort A ist Unsinn: natürlich lässt sich der Zusammenhang durch eine Regressionsgleichung (dies ist ein allgemeinerer Begriff als Regressionsgerade) beschreiben. Dabei ist es keineswegs gleichgültig, welches das x- und welches das y-Merkmal ist (Antwort B ist also auch falsch). Aus dem x-Merkmal kann das y-Merkmal anhand der Regressionsgleichung prognostiziert werden; daher ist im Einzelfall aufgrund fachlicher Überlegungen diese Frage zu beantworten. Im vorliegenden Fall ist die Größe quasi konstant vorgegeben; sie beeinflusst in gewissem Maße auch das Gewicht. Umgekehrt ist dies nicht der Fall: Frauen können ihr Gewicht durchaus beeinflussen; dadurch ändert sich jedoch nicht deren Größe. Durch eine Hungerkur nimmt man ab, wird deshalb aber nicht kleiner. Während einer Schwangerschaft nimmt man zu, ohne dabei zu wachsen. Aus diesen Überlegungen folgt, dass Antwort C falsch und D richtig ist. Schließlich sei noch auf die sinnlose Aussage von E hingewiesen: für die Berechnung des Korrelationskoeffizienten ist es belanglos, welches der beiden Merkmale abhängig bzw. unabhängig ist.

Teil II: Wahrscheinlichkeitsrechnung

5 Grundlagen der Wahrscheinlichkeitsrechnung

Aufgabe 5.1: Wertebereich einer Wahrscheinlichkeit

Geben Sie das kleinste Intervall an, das den gesamten Wertebereich einer Wahrscheinlichkeit enthält.

A. $(-\infty, +\infty)$
B. $[-1, +1]$
C. $[0, \infty)$
D. $[0,1]$
E. $(0,1)$

Aufgabe 5.2: Unabhängige Ereignisse

Die folgenden Sätze beinhalten jeweils 2 Ereignisse. Bei welchen Aussagen sind die Ereignisse ***un***abhängig voneinander?

1. Der systolische Blutdruck bei Patient M. betrug 180 mmHg vor einer blutdrucksenkenden Therapie und 145 mmHg danach.
2. Das 1. Kind einer Familie ist weiblich, das 2. ebenfalls (keine eineiigen Zwillinge).
3. Das Geschlecht eines Kindes ist männlich, das Geschlecht der Mutter weiblich.
4. Ein männlicher Patient erkrankt an Hämophilie.
5. Eine Person hat Blutgruppe A und Rhesusfaktor negativ.
6. Ein Vater ist 195 cm groß, dessen Sohn nur 182 cm.
7. Eine 20jährige Frau erkrankt an einem Mammakarzinom.

A. bei allen Aussagen
B. bei keiner Aussage
C. nur bei 2 und 5
D. nur bei 2, 3 und 5
E. nur bei 2, 5 und 6

Aufgabe 5.3: Wahrscheinlichkeit beim Kinderkriegen

Ein Vater von 4 Jungen plant mit seiner Partnerin ein weiteres Kind und wünscht sich sehnsüchtig ein Mädchen. Wie groß ist die Wahrscheinlichkeit p, dass sein Wunsch in Erfüllung geht? Wir nehmen an, dass mit der Wahrscheinlichkeit 1/2 ein neugeborenes Kind männlich ist, und dass keine Mehrlinge geboren werden.

A. $p = 31/32$
B. $p = 15/16$
C. $p = 1/2$
D. $p = 1/32$
E. Die Wahrscheinlichkeit kann aus den vorliegenden Angaben nicht ermittelt werden.

Aufgabe 5.4: Komplementäre Ereignisse

Bei wie vielen Ereignispaaren handelt es sich um komplementäre Ereignisse?

1. Rhesusfaktor positiv – Rhesusfaktor negativ
2. Blutgruppe A – Blutgruppe B
3. Geschlecht männlich – Geschlecht weiblich
4. schwanger – Geschlecht männlich
5. Würfeln: Augenzahl 6 – Augenzahl ≤ 5
6. herzkrank – an Diabetes erkrankt

A. 6
B. 5
C. 3
D. 2
E. 0

Aufgabe 5.5: Additionssatz

A und B seien 2 beliebige Ereignisse. Welche Aussage gilt generell?

A. $P(A \cup B) = P(A) + P(B) - 2P(A \cap B)$
B. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
C. $P(A \cup B) = P(A) + P(B)$
D. $P(A \cup B) = P(A) + P(B) - P(A) \cdot P(B)$
E. $P(A \cup B) > P(A)$

Aufgabe 5.6: Wahrscheinlichkeiten bei diagnostischen Verfahren

In einer gynäkologischen Klinik wird routinemäßig jede Frau auf das Vorhandenseins eines Mammakarzinoms untersucht. Dabei werden die Mammographie und die Palpation angewandt; die Testeregebnisse dieser beiden Methoden seien unabhängig voneinander. Die Wahrscheinlichkeit bei einer erkrankten Frau, ein Karzinom mit Mammographie zu entdecken, betrage 90%; bei der Palpation liegt diese Wahrscheinlichkeit nur bei 60%. Wie groß ist dann die Wahrscheinlichkeit, dass ein vorhandenes Karzinom bei einer Frau ***nicht*** entdeckt wird?

A. 1%
B. 4%
C. 10%
D. 40%
E. 45%

Aufgabe 5.7: Krankheiten und Wahrscheinlichkeit

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%

Aufgabe 5.8: 10 Merkmale eines Patienten

Bei einem bestimmten Patienten werden die Werte von 10 medizinisch relevanten Merkmalen erhoben. Die Wahrscheinlichkeit, dass ein Messwert innerhalb des Normbereichs liegt, beträgt für jedes dieser Merkmale 95%. Der Einfachheit halber nehmen wir an, dass die Ereignisse unabhängig voneinander sind. Wie groß ist dann die Wahrscheinlichkeit, dass ein Wert außerhalb des Normbereichs liegt?

A. $1-0{,}95^{10} \approx 0{,}40$
B. $0{,}05$
C. $0{,}95^9 \cdot 0{,}05 \approx 0{,}03$
D. $10 \cdot 0{,}95^9 \cdot 0{,}05 \approx 0{,}32$
E. $10 \cdot 0{,}05 = 0{,}50$

Aufgabe 5.9: Multiplikationssatz

Wann gilt: $P(A \mid B) = P(A)$?

A. Diese Gleichung ist generell richtig.
B. Diese Gleichung gilt nur dann, wenn A und B unabhängige Ereignisse sind.

C. Diese Gleichung gilt nur dann, wenn A und B disjunkte Ereignisse sind.
D. Diese Gleichung gilt nur dann, wenn A und B komplementäre Ereignisse sind.
E. Diese Gleichung gilt generell nie.

Aufgabe 5.10: Semmelweis – Berechnen einer Wahrscheinlichkeit

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%

Aufgabe 5.11: Karzinom am Versuchstier

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

Aufgabe 5.12: Erwartungswert beim Würfeln

Mit einem roten und einem blauen Würfel wird gleichzeitig gewürfelt. Die Augenzahl X des roten Würfels wird verdoppelt; davon wird die Augenzahl Y des blauen Würfels subtrahiert. Was ist der Erwartungswert der so berechneten Zufallsvariable?

A. 7
B. 10,5
C. 2
D. 3,5
E. Dieser Wert ist ohne zusätzliche Informationen nicht zu berechnen.

Aufgabe 5.13: Transformation einer Zufallsvariablen

Eine stetige Zufallsvariable X habe den Erwartungswert μ und die Varianz σ^2. Alle Werte von X werden nun transformiert nach $X \to aX + b$ (a und b sind konstante Zahlen). Wie ändern sich dadurch der Erwartungswert und die Varianz?

A. $E(aX+b) = a\mu + b$, $\mathrm{Var}(aX+b) = a^2\sigma^2$

B. $E(aX+b) = a\mu + b$, $\mathrm{Var}(aX+b) = a\sigma^2 + b$

C. $E(aX+b) = a\mu + b$, $\mathrm{Var}(aX+b) = a\sigma + b$

D. $E(aX+b) = a\mu$, $\mathrm{Var}(aX+b) = a\sigma$

E. $E(aX+b) = \mu$, $\mathrm{Var}(aX+b) = \sigma^2$

Aufgabe 5.14:* Empirisches Ermitteln einer Wahrscheinlichkeit

In der medizinischen Forschung wird eine Wahrscheinlichkeit in der Regel empirisch ermittelt; d. h. eine hinreichend große Stichprobe wird bezüglich eines Merkmals untersucht. Der Wert der relativen Häufigkeit einer Ausprägung wird dann als Näherungswert für die entsprechende Wahrscheinlichkeit zugrunde gelegt. Wonach ist dieses Vorgehen gerechtfertigt?

A. Dieses Vorgehen hat zwar eine lange Tradition, ist aber in keiner Weise gerechtfertigt.

B. Nach den Axiomen von Kolmogoroff.

C. Nach dem Gesetz der großen Zahl.

D. Nach der Definition der Wahrscheinlichkeit von Laplace.

E. Nach der Tschebyscheff'schen Ungleichung.

Aufgabe 5.1: **Wertebereich einer Wahrscheinlichkeit**

Lösung: ***D*** (alle Werte zwischen 0 und 1)

Eine Wahrscheinlichkeit kann – wie eine relative Häufigkeit – alle Werte zwischen 0 und 1 (inklusive dieser Grenzen) annehmen, aber keine Werte außerhalb dieses Bereichs. Da nach dem ***kleinsten*** Intervall gefragt ist, das diesen Bereich enthält, ist Antwort D korrekt.

Aufgabe 5.2: **Unabhängige Ereignisse**

Lösung: ***D*** (nur 2, 3 und 5)

Ob zwei Ereignisse abhängig oder unabhängig voneinander sind, ergibt sich aus sachlogischen Überlegungen. Es bedarf keiner langen Ausführungen, dass das Geschlecht eines Kindes nicht das Geschlecht des nachfolgenden Geschwisters beeinflusst (2) und ebenso, dass das Geschlecht eines Babys unabhängig ist von der Tatsache, dass seine Mutter weiblich ist (3). Ein Medizinstudent sollte wissen, dass die Wahrscheinlichkeit für Blutgruppe A ungefähr 42% beträgt, unabhängig davon, ob der Rhesusfaktor positiv oder negativ ist (5). – Alle anderen Ereignispaare sind jedoch abhängig voneinander. Dies ergibt sich aufgrund einfacher Überlegungen und minimaler medizinischer Fachkenntnisse.

Aufgabe 5.3: **Wahrscheinlichkeit beim Kinderkriegen**

Lösung: ***C*** ($p = 1/2$)

Die Wahrscheinlichkeit, dass ein neugeborenes Kind weiblich ist, beträgt bei ***jeder*** Geburt $p = 1/2$. Also ist auch das 5. Kind mit 50%-iger Wahrscheinlichkeit ein Mädchen, ***unabhängig*** davon, welches Geschlecht die älteren 4 Geschwister haben. Eigentlich ganz einfach – dennoch löst die Antwort immer wieder Erstaunen aus, so nach dem Motto: bei fast allen Familien mit 5 Kindern ist auch ein Mädchen dabei (die Wahrscheinlichkeit dafür beträgt in der Tat nur 31/32, Antwort A). Aber: hier ist nicht gefragt nach der Wahrscheinlichkeit, dass mindestens ein Kind weiblich ist, sondern nach der Wahrscheinlichkeit, dass das 5. Kind weiblich ist.

Aufgabe 5.4: **Komplementäre Ereignisse**

Lösung: ***C*** (Anzahl 3)

Komplementäre Ereignisse sind disjunkt und ergänzen sich zum Ereignisraum. Disjunkt sind alle Ereignispaare 1–5 (d. h. die beiden Ereignisse schließen sich gegenseitig aus). Bei den Paaren 1, 3 und 5 ist leicht nachvollziehbar, dass eines der beiden Ereignisse eintreten muss und nur eines eintreten kann. Dies trifft jedoch nicht zu für die Paare 2 und 4 (es gibt andere Blutgruppen als A und B und auch Personen, die weder männlich noch schwanger sind). –

Das 6. Paar (Diabetes- und herzkrank) ist leider nicht disjunkt und damit auch nicht komplementär.

Aufgabe 5.5: Additionssatz

Lösung: ***B***

Der Additionssatz in seiner allgemeinen Form ist die Gleichung unter B. Er beschreibt die Wahrscheinlichkeit, dass das Ereignis *A*, das Ereignis *B* oder auch beide Ereignisse eintreten (siehe Seite 106 im Buch). Die Gleichungen A, C und D sind Spezialformen. Die Gleichung unter A quantifiziert die Wahrscheinlichkeit, dass ***entweder A oder B*** eintritt (d. h. nur ein Ereignis, aber nicht beide zusammen). Wenn *A* und *B* disjunkt sind, ist die Antwort C richtig; falls *A* und *B* unabhängig sind, gilt D (siehe Seite 108). Auch die Ungleichung unter E ist nicht immer korrekt: falls *B* in *A* enthalten ist (z. B. *A* weiblich und *B* schwanger), gilt $P(A \cup B) = P(A)$.

Aufgabe 5.6: Wahrscheinlichkeiten bei diagnostischen Verfahren

Lösung: ***B***

Die Wahrscheinlichkeit, ***dass*** ein Karzinom entdeckt wird, berechnet sich nach dem Additionssatz für unabhängige Ereignisse (Seite 108) als: $0{,}9 + 0{,}6 - 0{,}9 \cdot 0{,}6 = 0{,}96$. Für die Wahrscheinlichkeit des komplementären Ereignisses (das Karzinom wird ***nicht*** entdeckt) berechnet man nach Formel (5.2) sofort: $P = 1 - 0{,}96 = 0{,}04$. – Man kann auch anders argumentieren: die Wahrscheinlichkeiten, dass die Mammographie bzw. die Palpation das Karzinom nicht entdecken, betragen jeweils 0,1 bzw. 0,4. Mit dem Multiplikationssatz für unabhängige Ereignisse (Formel 5.11, Seite 108) ergibt sich $P = 0{,}1 \cdot 0{,}4 = 0{,}04$. – Der Vollständigkeit halber seien angegeben: die Wahrscheinlichkeit, dass sowohl die Mammographie als auch die Palpation ein Karzinom entdecken, beträgt $0{,}9 \cdot 0{,}6 = 0{,}54$. Die Wahrscheinlichkeit, dass nur die Mammographie, nicht aber die Palpation hilfreich ist, beträgt dann: $0{,}9 - 0{,}54 = 0{,}36$. Die Wahrscheinlichkeit, dass nur die Palpation, aber nicht die Mammographie zum Erfolg führt, ist $0{,}6 - 0{,}54 = 0{,}06$. Die Summe dieser Wahrscheinlichkeiten $0{,}04 + 0{,}54 + 0{,}36 + 0{,}06$ ergibt 1 – logisch, denn genau eine der 4 Möglichkeiten muss in jedem Fall eintreten.

Aufgabe 5.7: Krankheiten und Wahrscheinlichkeit

Lösung: ***D*** (0,40)

Mit dem Additionssatz ergibt sich für die Wahrscheinlichkeit, dass ein Patient an einer Krankheit leidet (Diabetes oder Herz oder beide): $0{,}2 + 0{,}3 - 0{,}05 = 0{,}45$. Hier ist jedoch nach der Wahrscheinlichkeit gefragt, dass ein Patient ***entweder*** herzkrank oder diabeteskrank ist. Die Möglichkeit, dass beide Krankheiten vorliegen, ist also ausgeschlossen.

Deshalb muss die Wahrscheinlichkeit der Schnittmenge subtrahiert werden – so ergibt sich: $0{,}45 - 0{,}05 = 0{,}40$.

Aufgabe 5.8: **10 Merkmale eines Patienten**

Lösung: ***A*** ($\approx 0,40$)

Auch hier ermittelt man am besten zuerst die Wahrscheinlichkeit für das komplementäre Ereignis: die Wahrscheinlichkeit, dass alle 10 Parameter innerhalb des Normbereichs liegen, beträgt (bei Unabhängigkeit) $0,95^{10}$. Also ist die Wahrscheinlichkeit dafür, dass ein Parameter außerhalb des Normbereichs liegt, nach Formel (5.2): $1-0,95^{10} \approx 0,40$. Die Ausdrucksweise: „... ***ein*** Parameter liegt außerhalb..." bedeutet, dass ***mindestens*** einer außerhalb liegt; d. h. dies kann für einen, 2, 3 oder mehr Parameter gelten. – Diese Aufgabe mag nicht sehr praxisbezogen sein. Wenn von einem Patienten 10 Werte erfasst werden, ist kaum davon auszugehen, dass diese unabhängig voneinander sind. Die Aufgabe zeigt dennoch, dass man als behandelnder Arzt vorsichtig sein sollte etwa bei der Beurteilung von Laborwerten. Die Wahrscheinlichkeit, dass einer von 10 Werten „aus der Reihe fällt", ist recht hoch. Ein einzelner ausgefallener Wert sollte deshalb nicht überbewertet werden.

Aufgabe 5.9: **Multiplikationssatz**

Lösung: ***B*** (gilt nur bei unabhängigen Ereignissen)

Man mache sich die Bedeutung dieser Aussage klar: wenn A und B unabhängig sind, beeinflusst das Ereignis B in keiner Weise die Wahrscheinlichkeit für das Eintreten von A. Unabhängig sind beispielsweise Blutgruppe und Geschlecht. So beträgt die Wahrscheinlichkeit für Blutgruppe A 42%; egal, ob sie sich auf Frauen, auf Männer oder auf die Gesamtbevölkerung bezieht. Siehe auch Seite 108 im Buch.

Aufgabe 5.10: **Semmelweis – Berechnen einer Wahrscheinlichkeit**

Lösung: ***E*** (etwa 19%)

Antwort A ist falsch – man kann die gefragte Wahrscheinlichkeit sehr wohl berechnen. Offensichtlicher Unsinn ist auch Antwort B: eine Wahrscheinlichkeit kann nie über 100% liegen. Unter C ist die bedingte Wahrscheinlichkeit angegeben, dass eine erkrankte Frau stirbt (die Letalität) mit $P(T \mid K) = 0,80$. Unter D findet man die Wahrscheinlichkeit, dass eine Frau erkrankt (die Inzidenz) mit $P(K) = 0,24$. Nach dem Multiplikationssatz (5.9) berechnet man für die Wahrscheinlichkeit, zu erkranken und zu sterben (die Mortalität) $P(T \cap K) = P(T \mid K) \cdot P(K) = 0,80 \cdot 0,24 \approx 0,19$.

Aufgabe 5.11: Karzinom am Versuchstier

Lösung: ***E*** (0,86)

Man wendet ganz einfach den Additionssatz für unabhängige Ereignisse an (Formel (5.10), Seite 108) und erhält: $0{,}8+0{,}3-0{,}8\cdot 0{,}3=1{,}1-0{,}24=0{,}86$ (sollte auch ohne Taschenrechner zu bewältigen sein).

Aufgabe 5.12: Erwartungswert beim Würfeln

Lösung: ***D*** (3,5)

Der Erwartungswert bei jeder der Zufallsvariablen X und Y ist 3,5; gesucht ist der Erwartungswert der Variablen $2X-Y$. Nach Gleichungen (5.23) und (5.24) gilt:
$E(2X-Y)=2\cdot E(X)-E(Y)=7-3{,}5=3{,}5$. Siehe auch Beispiel 5.11 auf Seite 116.

Aufgabe 5.13: Transformation einer Zufallsvariablen

Lösung: ***A***

Der Erwartungswert ändert sich analog zur Zufallsvariablen (also zu aμ+b). Wegen der quadratischen Dimension der Varianz wird diese mit dem Faktor a^2 multipliziert; die Konstante b beeinflusst die Varianz nicht. Siehe Formeln (5.23) und (5.29), Seite 116 ff. Noch eine kleine Anmerkung: bei der Formel (5.30) auf Seite 118 muss es heißen: $\mathrm{Var}(b)=0$. Bitte verbessern Sie das in Ihrem Buch. Pardon für diesen Tippfehler.

Aufgabe 5.14:* Empirisches Ermitteln einer Wahrscheinlichkeit

Lösung: ***C*** (Gesetz der großen Zahl)

Diese Aufgabe wird gestellt, weil oft irgendwelche Namen genannt werden, ohne zu wissen, was eigentlich dahinter steckt. Das Gesetz der großen Zahl rechtfertigt, dass eine Wahrscheinlichkeit über eine relative Häufigkeit und ein Erwartungswert durch einen Stichproben-Mittelwert geschätzt wird. Dies ist in einer empirischen Wissenschaft wie der Medizin ein gängiges Verfahren. Die Axiome von Kolmogoroff und die Definition von Laplace definieren Wahrscheinlichkeiten; die Ungleichung von Tschebyscheff ermöglicht Abschätzungen, wie weit die Messwerte vom Erwartungswert entfernt liegen (siehe Seite 120).

6 Spezielle Wahrscheinlichkeiten in der Medizin

Aufgabe 6.1: Epidemiologische Maßzahlen

Sei K das Ereignis, während eines Beobachtungszeitraums an einer bestimmten Krankheit zu erkranken und T das Ereignis, daran zu sterben. Ordnen Sie die Wahrscheinlichkeiten den entsprechenden Begriffen zu:

1. $P(K)$	a. krankheitsspezifische Mortalität
2. $P(T \cap K)$	b. Letalität
3. $P(T \mid K)$	c. Inzidenz
	d. Prävalenz

A. 1a, 2b, 3c
B. 1c, 2b, 3a
C. 1c, 2a, 3b
D. 1d, 2a, 3b
E. 1d, 2a, 3c

Aufgabe 6.2: Berechnen einer Mortalität

In einer Intensivstation erkranken 20% der Patienten an einer Lungenentzündung, die Letalität bei dieser Krankheit betrage 70%. Wie hoch ist dann die krankheitsspezifische Mortalität?

A. 90%
B. 45%
C. 3,5%
D. 14%
E. Die Mortalität kann aus den Angaben nicht berechnet werden.

Aufgabe 6.3: Risiken bei Mammakoarzinom

Vonden 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

Aufgabe 6.4: Diagnostischer Test in der Notfallmedizin

Was ist bei einem diagnostischen Test, der in der Notfallmedizin bei lebensbedrohlichen Erkankungen 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

Aufgabe 6.5: Diagnostischer Test – falsche Ergebnisse

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. Mit der Formel von Bayes lässt sich der positive Vorhersagewert herleiten.
D. Die Spezifität beträgt mindestens 95%.
E. Keine dieser Aussagen ist herleitbar.

Aufgabe 6.6: Diagnostischer Test – positiver Vorhersagewert

Die Prävalenz einer Krankheit sei 0,0001. Ein diagnostischer Test habe eine Spezifität von 0,995 und eine Sensititvität 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.

Aufgabe 6.7:* Diagnostischer Test – Ändern des Schwellenwerts

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.

Aufgabe 6.8: Diagnostischer Test – Interpretation des Ergebnisses

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.

Aufgabe 6.9: Zusammenhang zwischen Sensitivität und Spezifität

Bei manchen Tests lassen sich die Sensitivität und die Spezifität beeinflussen. Welche Argumente sprechen für eine hohe Sensitivität?

1. es handelt sich um eine Krankheit mit gravierenden Folgen für den Patienten
2. es gibt eine erfolgversprechende Therapie
3. diese Therapie hat unter Umständen schwere Nebenwirkungen
4. die Therapie ist sehr teuer
5. die Therapie stellt die Patienten vor enorme psychische Belastungen
6. falsch-positive Befunde können relativ einfach geklärt werden

A. alle Argumente sprechen für eine hohe Sensitivität
B. nur die Argumente, die die Therapie betreffen (2, 3, 4 und 5)
C. nur die Argumente, die die Belastungen der Patienten betreffen (1 und 5)
D. nur die Argumente 1, 2 und 6
E. keines dieser Argumente

Aufgabe 6.10:* Verteilungsfunktion bei Sterbetafeln

Stellen Sie sich eine Sterbetafel vor mit den Sterbeziffern der im Jahr 1900 geborenen Personen. Sei X die Variable für die gelebten Jahre. Ein bestimmter Funktionswert $F(x)$ der Verteilungsfunktion beschreibt dann die Wahrscheinlichkeit für einen im Jahr 1900 lebendgeborenen Menschen,

A. höchstens das Alter x zu erreichen
B. mindestens das Alter x zu erreichen
C. zwischen dem x. und dem $(x+1)$. Geburtstag zu sterben
D. den x. Geburtstag zu erleben

E. am x. Geburtstag bereits tot zu sein

Aufgabe 6.1: Epidemiologische Maßzahlen

Lösung: ***C***

Dazu muss man nur die Definitionen der Begriffe kennen (siehe Buch Seite 125f.)

Aufgabe 6.2: Berechnen einer Mortalität

Lösung: ***D*** (14 %)

Nach Formel (6.1) berechnet sich die Mortalität als das Produkt aus Inzidenz und Letalität – dies ergibt $0{,}2 \cdot 0{,}7 = 0{,}14$.

Aufgabe 6.3: Risiken bei Mammakarzinom

Lösung: ***E*** (zuschreibbares Risiko 0,25)

Das Erkrankungsrisiko ohne den familiären Risikofaktor beträgt $P(K \mid \overline{R}) = 0{,}1$, bei familiärer Belastung ist es 3,5mal so hoch, also $P(K \mid R) = 0{,}35$. Dann beträgt nach (6.3) das ***zuschreibbare*** Risiko $P(K \mid R) - P(K \mid \overline{R}) = 0{,}35 - 0{,}1 = 0{,}25$. D. h. der Anteil 0,25 ist der familiären Belastung zuzuschreiben, der Anteil 0,1 geht auf andere Ursachen zurück. Das ***relative*** Risiko ist nach (6.2) gleich 3,5.

Aufgabe 6.4: Diagnostischer Test in der Notfallmedizin

Lösung: ***B*** (hohe Sensitivität)

Dies ist ein extremes Beispiel, das aber die Bedeutung einer hohen Sensitivität verdeutlicht. Wenn es um eine lebensbedrohliche Krankheit (und dazu mit einer guten Therapie) geht, sollte ein Test fast alle kranken Personen erkennen. Der damit verbundene Nachteil ist, dass dann auch einige nicht-erkrankte Personen fälschlicherweise ein positives Testergebnis erhalten und vielleicht unnötigerweise therapiert werden. Dies ist aber zu rechtfertigen, wenn dadurch anderen Menschen möglicherweise das Leben gerettet wird.

Aufgabe 6.5: Diagnostischer Test – falsche Ergebnisse

Lösung: ***B*** (die Wahrscheinlichkeit für falsch negativ ist 5%)

Sensitivität ist die Wahrscheinlichkeit, dass der Test ***richtig positiv*** erkennt. Das dazu komplementäre Ereignis ist ***falsch negativ.*** Die Wahrscheinlichkeit berechnet sich dann als $1 - 0{,}95 = 0{,}05$. Analog ergänzen sich Spezifität (richtig negativ) und die Wahrscheinlichkeit für falsch positiv.

Aufgabe 6.6: Diagnostischer Test – positiver Vorhersagewert

Lösung: ***A*** (0,02)

Es geht bei dieser Aufgabe keineswegs darum, mit der Formel von Bayes nach (6.14) den Vorhersagewert zu berechnen. Ein Kandidat sollte wissen, dass bei einer geringen Prävalenz der positive Vorhersagewert sehr niedrig ist und nur vorsichtig interpretiert werden darf. Siehe Beispiel 6.5 auf Seite 135.

Aufgabe 6.7:* Diagnostischer Test – Ändern des Schwellenwerts

Lösung: ***C***

Je kleiner der Schwellenwert ist, um so mehr positive Testergebnisse wird man erhalten. Das bedeutet einerseits, dass mehr kranke Personen richtig positiv eingestuft werden (dadurch steigt die Sensitivität), aber andererseits, dass mehr gesunde Personen falsch-positiv erkannt werden. Ein höherer Anteil falsch-positiver Ergebnisse geht einher mit einer niedrigeren Spezifität (siehe auch Aufgabe 6.5).

Aufgabe 6.8: Diagnostischer Test – Interpretation des Ergebnisses

Lösung: ***E***

Es ist klar, dass die Antworten A und B falsch sind. Ein positives Testergebnis allein kann niemals eindeutig belegen, dass die Krankheit vorliegt (und schon gar nicht, dass sie ausgeschlossen werden kann). Die meisten Mediziner entscheiden intuitiv nach Aussage C, indem sich die Sensitivität mit dem positiven Vorhersagewert gleichsetzen. Das ist aber grundlegend falsch. Falsch ist auch der Inhalt von Antwort D. Die Tatsache, dass der Patient drogenabhängig ist, bedeutet eine höhere Prävalenz und damit auch einen höheren positiven Vorhersagewert. Richtig ist die vorsichtige Interpretation des Testergebnisses (unter Berücksichtigung aller bekannten Faktoren) gemäß Antwort E.

Aufgabe 6.9: Zusammenhang zwischen Sensitivität und Spezifität

Lösung: ***D*** (1, 2 und 6)

Hohe Sensitivität bedeutet: möglichst viele Kranke werden richtig positiv erkannt (dafür sprechen eindeutig die Argumente 1 und 2), falsch-negative Ergebnisse werden weitgehend vermieden. Nach Argument 6 ist ein falsch-positiver Befund weit weniger tragisch als ein falsch-negativer – auch dies spricht für eine hohe Sensitivität. – Dagegen sprechen die Argumente 3-5 eher dafür, dass unnötige Therapiemaßnahmen vermieden werden (also für wenig falsch-positive Ergebnisse und damit für eine hohe ***Spezifität***).

Aufgabe 6.10:* **Verteilungsfunktion bei Sterbetafeln**

Lösung: ***A*** (höchstens das Alter x)

$F(x)$ ist – wie jede Verteilungsfunktion – monoton wachsend und beschreibt Summenhäufigkeiten (bzw. aufaddierte Wahrscheinlichkeiten). Höchstens das Alter x erreichen heißt: man stirbt vor dem x. Geburtstag. Deshalb ist $F(0) = 0$ (niemand, der lebend geboren wird, stirbt vor seinem 0. Geburtstag), steigt dann langsam an und erreicht den Wert 1 etwa bei $x = 120$. Es ist ganz sicher (Wahrscheinlichkeit 1), dass man höchstens 120 Jahre alt wird.

7 Einige theoretische Verteilungen

Aufgabe 7.1: Binomialverteilung – Therapie

Eine Therapie ist mit der Wahrscheinlichkeit $p = 0{,}9$ erfolgreich. Dann ist die Wahrscheinlichkeit, dass von 10 Behandlungen mindestens 8 erfolgreich verlaufen, gleich

A. $\sum_{k=8}^{10} \binom{10}{k} \cdot 0{,}9^k \cdot 0{,}1^{10-k} \approx 0{,}93$

B. $\binom{10}{8} \cdot 0{,}9^8 \cdot 0{,}1^2 \approx 0{,}19$

C. $1 - \binom{10}{8} \cdot 0{,}9^8 \cdot 0{,}1^2 \approx 0{,}81$

D. $8 \cdot 0{,}9 \cdot 0{,}1 = 0{,}72$

E. keine der Antworten A-D ist korrekt

Aufgabe 7.2: Binomialkoeffizient

Welcher der folgenden Ausdrücke quantifiziert die Anzahl der Möglichkeiten, aus 49 Lottokugeln 6 verschiedene auszuwählen?

A. 49^6

B. $49 \cdot 48 \cdot 47 \cdot 46 \cdot 45 \cdot 44$

C. $\frac{49!}{43!}$

D. $\binom{49}{6} = \frac{49!}{43! \cdot 6!}$

E. $\frac{49^6}{6!}$

Aufgabe 7.3: Binomialverteilung – Erbleiden

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 ***nicht*** 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.

Aufgabe 7.4:* Poissonverteilung – Kinder mit Down–Syndrom

Die Anzahl der am Klinikum Mannheim geborenen Kinder beträgt etwa $n = 2000$ pro Jahr. Die Wahrscheinlichkeit, dass ein Neugeborenes mit einem Down-Syndrom zur Welt kommt, betrage $p = 1/1000$. Wir nehmen an, dass die Ereignisse unabhängig voneinander sind. Die Anzahl der Kinder mit Down-Syndrom sei X. Welche Aussage trifft ***nicht*** zu?

A. X kann durch eine Poissonverteilung approximiert werden.
B. Der Erwartungswert von X beträgt 2.
C. Die Standardabweichung von X beträgt 2.
D. Die Varianz von X beträgt 2.
E. X ist eine diskrete Zufallsvariable, die theoretisch alle Werte zwischen 0 und 2000 annehmen kann.

Aufgabe 7.5:* Poissonverteilung – Geburtstag

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

A. $\frac{30}{365} \cdot e^{-30/365} \approx 0{,}0757$
B. $1 - e^{-30/365} \approx 0{,}0789$
C. $\frac{30}{365} \approx 0{,}0822$
D. $1 - \frac{364}{365} \approx 0{,}0027$
E. Keine der Antworten A-D ist korrekt.

Aufgabe 7.6 Verteilung von Wochentagen

Wir nehmen an, dass der Wochentag keinen Einfluss auf das Eintreten einer Spontangeburt hat. Dann folgt die Verteilung der Geburtszeiten auf die Wochentage

A. einer Poissonverteilung
B. einer diskreten Gleichverteilung
C. einer Binomialverteilung
D. einer Normalverteilung
E. einer Polynomialverteilung

Aufgabe 7.7 Spezielle Normalverteilungen

Welche der folgenden Merkmale bzw. Zufallsvariablen können als normalverteilt angesehen werden?

1. die Körpergröße erwachsener Frauen
2. die Körpergröße der gesamten erwachsenen Bevölkerung
3. das Körpergewicht erwachsener Frauen
4. Messfehler
5. Lebensdauern
6. die Mittelwerte, die aus zahlreichen Stichproben des Umfangs $n = 30$ aus einer bestimmten Grundgesamtheit berechnet werden
7. eine binomialverteilte Zufallsvariable $X : B(100;0,2)$

A. nur 1, 4, 6 und 7
B. nur 6 und 7
C. nur 1-5
D. keines
E. alle

Aufgabe 7.8 Normalverteilung – allgemeine Eigenschaften

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. Je kleiner die Varianz, desto flacher verläuft die Glockenkurve.
D. Der Erwartungswert ist immer gleich 0.
E. Der Erwartungswert kann keine negativen Werte annehmen.

Aufgabe 7.9 Normalverteilung – Dichtefunktion

Welche Aussage ist ***falsch***?

A. Die spezielle Form der Glockenkurve ist unabhängig von der Varianz σ^2.
B. Die Dichtefunktion wird durch eine Glockenkurve graphisch dargestellt.
C. Die Dichtefunktion ist symmetrisch bzgl. des Erwartungswerts μ.
D. Das Integral unter der gesamten Kurve hat den Wert 1.
E. Die Dichtefunktion hat für alle x-Werte zwischen $-\infty$ und $+\infty$ einen Funktionswert, der größer als 0 ist.

Aufgabe 7.10 Standardnormalverteilung

Sei $X : N(\mu, \sigma^2)$ und Z standardnormalverteilt. Welche Aussagen sind richtig?

1. $Z : N(0,1)$
2. $Z : N(1,0)$
3. Durch $(X - \mu)/\sigma^2$ wird X in Z transformiert.
4. Eine Rücktransformation von Z in X ist nicht immer möglich.

A. Nur 1 und 3 sind richtig.
B. Nur 2 und 3 sind richtig.
C. Nur 1, 3 und 4 ist richtig.
D. Nur 2, 3 und 4 ist richtig.
E. Nur 1 ist richtig.

Aufgabe 7.11: Normalverteilung – Transformationen von *X*

Sei $X : N(\mu, \sigma^2)$ normalverteilt. Welche der folgenden Aussagen ist ***falsch***?

A. $Y = X - a$ ist normalverteilt mit $Y : N(\mu - a, \sigma^2)$
B. $Y = (X - \mu)/\sigma$ ist standardnormalverteilt.
C. $Y = X/\sigma$ ist normalverteilt mit $Y : N(\mu/\sigma, 1)$
D. $Y = X^2$ ist normalverteilt mit $Y : N(\mu^2, \sigma^4)$
E. $Y = aX + b$ ist normalverteilt mit $Y : N(a\mu + b, a^2\sigma^2)$

Aufgabe 7.12: Normalverteilung – Referenzbereiche

Die Körpergröße männlicher Studenten sei normalverteilt mit einem Erwartungswert von $\mu = 180$ cm und einer Standardabweichung von $\sigma = 6$ 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 %

Aufgabe 7.13:* Symmetrische Verteilungen

Welche der folgenden Verteilungen sind symmetrisch?

1. Standardnormalverteilung
2. Normalverteilung allgemein
3. Binomialverteilung $B(n, p)$ mit $p = 0{,}5$
4. Poissonverteilung $P(\lambda)$
5. diskrete Gleichverteilung
6. Verteilung zur Beschreibung von Lebensdauern von Personen
7. t-Verteilung
8. Chi^2-Verteilung

A. nur 1 ist symmetrisch
B. nur 1 und 2 sind symmetrisch
C. nur 1, 2, 3, 5 und 7 sind symmetrisch
D. alle außer 4 und 6 sind symmetrisch
E. alle außer 6 sind symmetrisch

Aufgabe 7.14: Verteilung von Mittelwerten

Der systolische Blutdruck bei gesunden Männern zwischen 20 und 30 Jahren ist symmetrisch verteilt mit $\mu = 120$ mmHg und $\sigma = 10$ mmHg. Wie sind dann die Mittelwerte verteilt, die aus den Blutdruckwerten von 25 zufällig ausgewählten Studenten berechnet werden?

A. genauso verteilt wie die Blutdruckwerte der Studenten
B. normalverteilt mit $\mu = 120$ mmHg und unbekannter Varianz
C. normalverteilt mit $\mu = 120$ mmHg und $\sigma = 0{,}4$ mmHg
D. normalverteilt mit $\mu = 120$ mmHg und $\sigma = 2$ mmHg
E. Über die Verteilung der Mittelwerte kann nichts ausgesagt werden.

Aufgabe 7.1: Binomialverteilung – Therapie

Lösung: ***A***

„Mindestens 8 sind erfolgreich" heißt, dass 8, 9 oder 10 Behandlungen erfolgreich sind. Dies lässt sich ausrechnen über die Wahrscheinlichkeitsfunktion der Binomialverteilung, Formel (7.6) auf Seite 145. Mit $n = 10$, $p = 0{,}9$ und $q = 0{,}1$ erhält man: $P(X = 8) = 0{,}19$, $P(X = 9) = 0{,}39$ und $P(X = 10) = 0{,}35$ – die Summe ergibt 93%. Die Angabe in Antwort B quantifiziert die Wahrscheinlichkeit, dass ***genau*** 8 Therapien erfolgreich sind.

Aufgabe 7.2: Binomialkoeffizient

Lösung: ***D***

Der Binomialkoeffizient quantifiziert die Anzahl der Möglichkeiten, aus einer 49-elementigen Menge 6 Objekte auszuwählen. Genau danach ist gefragt. Dieser Ausdruck entspricht ungefähr 14 Millionen – dementsprechend gering sind die Aussichten auf einen 6er im Lotto. – Es ist wichtig, sich folgendes klarzumachen: die Kugeln werden nacheinander gezogen und ***nicht*** wieder zurückgelegt; die Reihenfolge der gezogenen Kugeln spielt keine Rolle. Unter A ist die Anzahl der Möglichkeiten angegeben, 6 Elemente aus 49 zu ziehen, wenn jede Kugel nach dem Zug wieder zurückgelegt wird und die Reihenfolge der Ziehungen wichtig ist. Der Ausdruck in B (der gleichbedeutend zu dem in C ist) gibt an, wie viele Möglichkeiten es gibt, 6 aus 49 zu ziehen, wenn die Kugeln nicht zurückgelegt werden und die Reihenfolge der Ziehungen eine Rolle spielt. Da die Reihenfolge bei den Lottokugeln unerheblich ist, wird nochmal durch 6! dividiert und man erhält den Binomialkoeffizienten unter D.

Aufgabe 7.3: Binomialverteilung – Erbleiden

Lösung: ***E*** (trifft nicht zu)

Jedes Kind erhält mit der Wahrscheinlichkeit $p = 1/2$ heterozygote Erbanlagen. Die Anzahl *X* lässt sich dann durch eine Binomialverteilung beschreiben; nach (7.3) und (7.4) berechnet man den Erwartungswert $n/2$ und die Varianz $n/4$. Für $p = 1/2$ ist die Verteilung symmetrisch (siehe Abbildung 7.2 Seite 150). Also sind die Antworten A-D korrekt. Falsch ist E. Die Wahrscheinlichkeit, dass ein Kind homozygot erkrankt (oder auch homozygot gesund) ist, beträgt 1/4; der Erwartungswert ist dann $n/4$.

Aufgabe 7.4:* Poissonverteilung – Kinder mit Down-Syndrom

Lösung: ***C*** (Standardabweichung ist ***nicht*** 2)

Es ist hoffentlich einsichtig, dass man die Anzahl der Kinder mit Down-Syndrom durch eine Binomialverteilung beschreiben kann. Weil hier *n* sehr groß und *p* sehr klein ist, lässt sich die

Binomialverteilung durch eine Poissonverteilung approximieren (bei der Erwartungswert und Varianz übereinstimmen). Für dieses Beispiel haben sie den Wert: $\lambda = np = 2000/1000 = 2$. Also sind A, B und D richtig, C ist falsch (die Standardabweichung beträgt $\sqrt{2}$). Richtig ist auch E: X kann theoretisch alle Werte zwischen 0 und 2000 annehmen, auch wenn damit zu rechnen ist, dass X in der Praxis nicht größer als 6 wird (siehe auch Beispiel 7.6 auf Seite 151).

Aufgabe 7.5:* Poissonverteilung – Geburtstag

Lösung: ***B***

Ein Student hat Geburtstag, bedeutet: mindestens einer. Also löst man diese Aufgabe am besten dadurch, dass man zunächst die Wahrscheinlichkeit für das komplementäre Ereignis (***keiner*** hat Geburtstag) berechnet. Diese ergibt sich mit Formel (7.14) und $k = 0$ als $e^{-30/365}$ (auch hier lässt sich die Binomialverteilung durch eine Poissonverteilung mit $\lambda = 30/365$ approximieren). Die Antwort A gibt die Wahrscheinlichkeit an, dass ***genau*** ein Student an diesem Tag Geburtstag hat.

Aufgabe 7.6: Verteilung von Wochentagen

Lösung: ***B***

Es gibt hier 7 Ausprägungen (nämlich die Wochentage Montag bis Sonntag) mit jeweils der Wahrscheinlichkeit 1/7. Dies ist eine diskrete Gleichverteilung.

Aufgabe 7.7: Spezielle Normalverteilungen

Lösung: ***A*** (Körpergröße, Messfehler, Mittelwerte und $X : B(100;0{,}2)$)

Die Körpergröße erwachsener Frauen ist ein Merkmal, das – im Gegensatz zum Körpergewicht – kaum beeinflussbar ist. Beim Körpergewicht gibt es insbesondere im oberen Bereich Ausreißer; dieses Merkmal ist (ebenso wie Lebensdauern) rechtsschief verteilt. Wenn man die Größe ***aller*** Erwachsenen (heterogene Population) betrachtet, wird man eine 2-gipfelige Verteilung erhalten mit einem „weiblichen" und einem „männlichen" Gipfel – also keine Normalverteilung. Messfehler sind ein klassisches Beispiel für eine Normalverteilung – das wusste schon Carl Friedrich Gauss. Nach dem zentralen Grenzwertsatz sind auch Mittelwerte normalverteilt, ebenso Binomialverteilungen, falls $n \cdot p \cdot (1-p) > 9$ (in unserem Beispiel ist $n \cdot p \cdot (1-p) = 16$).

Aufgabe 7.8: **Normalverteilung – allgemeine Eigenschaften**

Lösung: ***B*** (große Varianz, flache Kurve)

Innerhalb des 3σ-Bereichs liegen zwar 99,7% aller Werte, aber eben nicht alle (Antwort A ist falsch). Dass bei einer großen Varianz die Glockenkurve flach und bei einer kleinen Varianz die Kurve hoch-gestreckt ist, erkennt man in Abbildung 7.4 auf Seite 160 (also C falsch). Der Erwartungswert kann beliebige Werte – auch negative – annehmen, deshalb sind D und E falsch.

Aufgabe 7.9: **Normalverteilung – Dichtefunktion**

Lösung: ***A*** (Aussage ist falsch)

Die spezielle Form der Glockenkurve ist keineswegs ***un***abhängig von der Varianz; je größer die Varianz, desto flacher die Kurve (siehe Aufgabe 7.8). Alle anderen Aussagen sind richtig; siehe Seite 156 ff.

Aufgabe 7.10: **Standardnormalverteilung**

Lösung: ***E*** (nur 1 ist richtig)

Die Standardnormalverteilung hat den Erwartungswert 0 und die Varianz 1 (Aussage 1), nicht umgekehrt (Aussage 2 ist deshalb falsch). Die Transformation ist $(X-\mu)/\sigma$; es wird also nicht durch σ^2 dividiert (deshalb ist 3 falsch). Jede transformierte Variable kann über $X=\mu+\sigma Z$ zurücktransformiert werden; Aussage 4 ist also auch falsch.

Aufgabe 7.11: **Normalverteilung – Transformationen von *X***

Lösung: ***D*** (X^2 ist nicht normalverteilt)

Alle linearen Transformationen (A, B, C und E) ändern nichts an der Tatsache der Normalverteilung. Dabei ändern sich lediglich der Erwartungswert und die Varianz; diese ergeben sich nach den Formeln (5.23) und (5.29). Bei nichtlinearen Transformationen wie z. B. X^2 gehen die Eigenschaften der Normalverteilung verloren.

Aufgabe 7.12: **Normalverteilung – Referenzbereiche**

Lösung: ***C*** (etwa 18,5 %)

Der σ-Bereich ist in diesem Fall das Intervall [174,186], der 2σ-Bereich ist [168,192]. Innerhalb des σ-Bereichs liegen 68% aller Messwerte, außerhalb 32%; d. h. größer als die obere Grenze (also 186 cm) sind 16%. Außerhalb des 2σ-Bereichs liegen etwa 5% aller Werte; d. h. kleiner als die untere Grenze (168 cm) sind 2,5%. 16% + 2,5% ergibt 18,5 % (siehe auch Tabelle 7.1 auf Seite 161).

Aufgabe 7.13:* Symmetrische Verteilungen

Lösung: ***C*** (1, 2, 3, 5 und 7 sind symmetrisch)

Die Normalverteilung ist immer symmetrisch – nicht nur die Standardnormalverteilung (siehe Abb. 7.4, Seite 160). Die Binomialverteilung ist nur für $p = 0{,}5$ symmetrisch, die Poissonverteilung ist generell rechtsschief (siehe Abb. 7.1 – 7.3, Seite 150). Auch die Gleichverteilung (bei der jede Wahrscheinlichkeit gleich sind) ist symmetrisch. Lebensdauer-Verteilungen sind dagegen rechtsschief (z. B. die Lognormal-Verteilung oder die Exponential-Verteilung). Die *t*-Verteilung ähnelt der Normalverteilung und ist wie diese symmetrisch im Gegensatz zur Chi^2-Verteilung, die rechtsschief ist (siehe Abb. 7.8, Seite 179).

Aufgabe 7.14: Verteilung von Mittelwerten

Lösung: ***D*** (μ=120 mmHg, σ=6 mmHg)

Nach dem zentralen Grenzwertsatz sind Mittelwerte normalverteilt mit dem Erwartungswert μ und der Varianz s^2/n (siehe Seite 166). Demnach ist der Erwartungswert $m = 1200 \text{ mmHg}$, die Standardabweichung $\sigma/\sqrt{n}$, also $10 \text{ mmHg} / 5 = 2 \text{ mmHg}$.

Teil III: Induktive Statistik

8 Schätzverfahren

Aufgabe 8.1: Eigenschaften eines Schätzverfahrens

Welche Aussage ist ***falsch***?

A. Der zu schätzende Parameter der Grundgesamtheit ist unbekannt.
B. Der zu schätzende Parameter der Grundgesamtheit ist eine konstante Größe.
C. Die Schätzfunktion ist eine Zufallsvariable, deren Realisation abhängig ist von der Stichprobe.
D. Das Konfidenzintervall ist ein Bereich, das den zu schätzenden Parameter mit Sicherheit enthält.
E. Um eine erwartungstreue Schätzung zu ermöglichen, muss eine repräsentative Stichprobe vorliegen.

Aufgabe 8.2: Konfidenzintervall für den Erwartungswert

Welche der folgenden Aussagen bzgl. eines Konfidenzintervalls für μ ist richtig?

A. Je größer der Stichprobenumfang n ist, um so größer ist das Konfidenzintervall.
B. Je größer n ist, um so kleiner ist das Konfidenzintervall.
C. Die Breite des Konfidenzintervalls ist unabhängig von n.
D. Jedes Konfidenzintervall, das aus einer repräsentativen Stichprobe ermittelt wird, enthält μ.
E. Der Stichprobenmittelwert ist unerheblich für die Bestimmung der Intervallgrenzen.

Aufgabe 8.3: Standardfehler des Mittelwerts

Wie muss der Stichprobenumfang n geändert werden, um den Standardfehler des Mittelwerts zu halbieren?

A. n muss ebenfalls halbiert werden.
B. n muss verdoppelt werden.
C. n muss vervierfacht werden.
D. Dazu ist ein Stichprobenumfang von $\sqrt{2}n$ erforderlich.
E. n muss nicht gerändert werden, da der Standardfehler des Mittelwerts unabhängig von n ist.

Aufgabe 8.4: Ermitteln eines Konfidenzintervalls mit der *t*-Verteilung

Mit Hilfe der *t*-Verteilung soll ein 2-seitiges Konfidenzintervall für den Erwartungswert μ gebildet werden. Welcher Parameter wird dafür ***nicht*** benötigt?

A. der Stichproben-Mittelwert $\bar{x}$
B. die empirische Standardabweichung s
C. der Stichprobenumfang n
D. das Quantil $t_{n-1;1-\alpha/2}$ der *t*-Verteilung
E. die Standardabweichung σ der Grundgesamtheit

Aufgabe 8.5: Breite eines Konfidenzintervalls

Wovon ist die Breite eines Konfidenzintervalls für μ ***nicht*** abhängig?

A. vom Stichprobenumfang n
B. vom Stichproben-Mittelwert $\bar{x}$
C. von der Irrtumswahrscheinlichkeit α
D. von der Variabilität der Messwerte
E. davon, ob das Intervall 1-seitig oder 2-seitig ist

Aufgabe 8.6: Konfidenzintervall für die Körpergröße

Wir nehmen an, dass die Körpergröße X erwachsener Frauen normalverteilt ist mit der Standardabweichung $\sigma = 5\,\text{cm}$. Aus einer Stichprobe von 25 Frauen ergibt sich ein Mittelwert $\bar{x} = 168\,\text{cm}$. Damit lässt sich als 2-seitiges Konfidenzintervall zur Irrtumswahrscheinlichkeit $\alpha = 0{,}05$ ermitteln:

A. $[163\,\text{cm}\,; 173\,\text{cm}]$
B. $[158\,\text{cm}\,; 178\,\text{cm}]$
C. $[164\,\text{cm}\,; \infty[$
D. $[166\,\text{cm}\,; 170\,\text{cm}]$
E. Dieses Intervall kann nicht bestimmt werden, da die empirische Standardabweichung nicht angegeben ist.

Aufgabe 8.7:* Schätzeigenschaften von Mittelwert und Median

Der Erwartungswert einer Grundgesamtheit soll geschätzt werden. Welche Aussage trifft zu?

A. Bei schief-verteilten Merkmalen ist die Schätzung über den empirischen Median nicht erwartungstreu.

B. Bei symmetrisch verteilten Merkmalen ist die Schätzung durch den empirischen Median zwar erwartungstreu, aber nicht konsistent.
C. Bei normalverteilten Merkmalen ergeben der Mittelwert und der empirische Median einer Stichprobe stets denselben Schätzwert.
D. Der empirische Median liefert stets einen Schätzwert, dessen Abstand zum Erwartungswert größer ist als der Abstand zwischen Mittelwert und Erwartungswert.
E. Bei symmetrischen Merkmalen ist die Schätzung durch den empirischen Median ebenso wie die Schätzung durch den Mittelwert erschöpfend.

Aufgabe 8.8:* Schätzen von Varianz und Standardabweichung

Welche Aussage ist ***falsch***?

A. Bei der Berechnung der empirischen Varianz wird durch $n-1$ dividiert, damit die Schätzung erwartungstreu ist.
B. Die Schätzung der Standardabweichung durch die Wurzel aus der empirischen Varianz ist nicht erwartungstreu.
C. Bei der Berechnung der empirischen Varianz wird durch $n-1$ dividiert, damit die Schätzung konsistent ist.
D. Die Schätzung der Standardabweichung durch die Wurzel aus der empirischen Varianz ist konsistent.
E. Je größer der Stichprobenumfang n ist, desto genauer ist die Schätzung der Varianz.

Aufgabe 8.9: Interpretation eines Konfidenzintervalls

In Beispiel 8.4 (Seite 196) erhält man für den Anteil weiblicher Medizinstudenten aus einer Stichprobe einen Schätzwert von $\hat{p} = 0{,}435$ und als 95%-Konfidenzintervall $[0{,}311; 0{,}559]$. Was besagt dieses Intervall bezüglich des Anteils weiblicher Medizinstudenten in der Grundgesamtheit?

A. Dieser Anteil liegt mit Sicherheit zwischen den Werten 0,311 und 0,559.
B. Der Anteil liegt mit 95%-iger Wahrscheinlichkeit zwischen 0,311 und 0,559.
C. Der „weibliche" Anteil ist mit einer Wahrscheinlichkeit von 2,5% geringer als 0,311.
D. Dieser Anteil ist mit einer Wahrscheinlichkeit von 2,5% größer als 0,559.
E. Es ist letzten Endes unbekannt, ob der zu schätzende Anteil innerhalb des Konfidenzintervalls liegt. Man weiß nur, dass das angewandte Verfahren – sofern seine Voraussetzungen erfüllt sind – mit 95%-iger Wahrscheinlichkeit ein Konfidenzintervall erzeugt, das den Anteil der Grundgesamtheit enthält.

Aufgabe 8.10:* Zensierte Daten

Bei Überlebensstudien treten oft zensierte Beobachtungen auf. Diese können mit einem Verfahren nach Kaplan und Meier ausgewertet werden. Welche Aussage ist richtig?

A. Die Gründe, die eine Zensur notwendig machen, sollten in keinem Zusammenhang zu den Endereignissen stehen.
B. Zensierte Beobachtungen haben keinerlei Einfluss auf das Ergebnis der Studie.
C. Bei der Kaplan-Meier-Methode werden zensierte Daten vollständig eliminiert.
D. Zensierte Beobachtungen sind vollkommen unproblematisch, solange der Stichprobenumfang genügend groß bleibt.
E. Bei den Schätzungen nach der Kaplan-Meier-Methode werden zu jedem Beobachtungszeitpunkt gleich viele Beobachtungseinheiten berücksichtigt.

Aufgabe 8.1: **Eigenschaften eines Schätzverfahrens**

Lösung: ***D*** (diese Aussage ist falsch!)

Man erhält (bei einer Irrtumswahrscheinlichkeit von 5%) nur in 95% aller Fälle ein Konfidenzintervall, das den zu schätzenden Parameter enthält. Alle anderen Aussagen sind korrekt: der Parameter ist eine unbekannte, konstante Größe (A und B); die Schätzfunktion liefert mit jeder Stichprobe einen anderen Schätzwert (C). Logischerweise muss die Stichprobe repräsentativ sein (E); ansonsten begeht man einen systematischen Fehler und die Schätzung ist nicht erwartungstreu.

Aufgabe 8.2: **Konfidenzintervall für den Erwartungswert**

Lösung: ***B*** (je größer n, desto kleiner das Intervall)

Es ist klar: ein großes n ermöglicht eine präzisere Schätzung. Dies bedeutet: ein kleines Konfidenzintervall. Dies Abhängigkeit kommt auch in den Formeln (8.8) oder (8.9) zum Ausdruck. Daher ist B richtig, A und C sind falsch. Dass die Aussage D nicht stimmt, wurde oben bei Aufgabe 8.1 erläutert. Auch E ist Unsinn: basierend auf dem Mittelwert werden die Intervallgrenzen berechnet.

Aufgabe 8.3: **Standardfehler des Mittelwerts**

Lösung: ***C*** (n muss vervierfacht werden)

Der Standardfehler des Mittelwerts wird durch $s/\sqrt{n}$ quantifiziert (siehe Seite 195 oben). Weil n als Wurzel im Nenner steht, bewirkt ein 4facher Umfang eine Halbierung des Intervalls.

Aufgabe 8.4: **Ermitteln eines Konfidenzintervalls mit der *t*–Verteilung**

Lösung: ***E*** (σ ist nicht notwendig)

Ein Blick auf die Formel zur Bestimmung des Konfidenzintervalls (8.10) auf Seite 194 zeigt, dass alle Parameter, die unter A–D genannt werden, zur Bestimmung des Konfidenzintervalls notwendig sind. Den Parameter σ der Grundgesamtheit (der ja meist unbekannt ist), geht in diese Berechnung nicht ein. Er wird durch die empirische Standardabweichung s geschätzt. Das ist der Vorteil der t-Verteilung !

Aufgabe 8.5: **Breite eines Konfidenzintervalls**

Lösung: ***B*** (Mittelwert)

Der Mittelwert ist die Mitte des Konfidenzintervalls; er beeinflusst jedoch nicht dessen Breite (siehe auch Formel (8.16)). Alle anderen Größen schon: je größer n und je größer α, desto kleiner ist das Konfidenzintervall (A und B). Die Variabilität der Messwerte wirkt sich aus auf σ bzw. s ; je größer die Variabilität, desto größer das Intervall (Antwort D). 1-seitige Intervalle haben sind unendlich breit; 2-seitige nicht (E).

Aufgabe 8.6: Konfidenzintervall für die Körpergröße

Lösung: ***D*** $[166\,\text{cm}\,;170\,\text{cm}]$

Hier ist ausnahmsweise die Standardabweichung σ der Grundgesamtheit bekannt; daher lässt sich das Konfidenzintervall nach (8.8) bestimmen. Der Ausdruck $1{,}96 \cdot \sigma / \sqrt{n}$ ergibt ungefähr 2; demnach ergeben sich die Grenzen 166 cm und 170 cm. Bei der Angabe unter C handelt es sich um ein 1-seitiges Intervall.

Aufgabe 8.7:* Schätzeigenschaften von Mittelwert und Median

Lösung: ***A*** (Median ist bei schiefen Verteilungen nicht erwartungstreu)

Bei schiefen Verteilungen stimmen Mittelwert und Median ***nicht*** überein; deshalb wäre die Schätzung über den Median verzerrt. Nur bei symmetrischen Verteilungen ist diese Schätzung erwartungstreu und auch konsistent (B ist also falsch). Offensichtlich unsinnig ist Antwort C; es wäre ein großer Zufall, wenn Mittelwert und Median einer Stichprobe übereinstimmten. Man kann auch nicht sagen, dass der geschätzte Median stets weiter weg vom Erwartungswert liegt als der Mittelwert (D ist also auch falsch). Er ist aber im Durchschnitt ein schlechterer Schätzwert als der Mittelwert, da er weniger effizient ist, und er ist nicht erschöpfend, weil nur ein oder zwei Messwerte in dessen Berechnung einfließen (E ist also falsch).

Aufgabe 8.8:* Schätzen von Varianz und Standardabweichung

Lösung: ***C*** (diese Aussage ist ***falsch***)

A und B sind richtig: die empirische Varianz ist ein erwartungstreuer Schätzer für die Varianz der Grundgesamtheit; die Standardabweichung dagegen nicht (siehe Ausführungen Seite 186). Allerdings wird nur wegen der Erwartungstreue, nicht wegen der Konsistenz durch n–1 dividiert (konsistent wäre die Schätzung auch bei der Division durch n).

Aufgabe 8.9: Interpretation eines Konfidenzintervalls

Lösung: ***E***

Hierzu sei auf die Anmerkung auf Seite 192 im Buch verwiesen. Man kann niemals sagen, dass ein Parameter „mit Sicherheit“ oder einer bestimmten Wahrscheinlichkeit innerhalb oder außerhalb des Konfidenzintervalls liegt. Dem Parameter haftet nämlich nichts zufälliges an; zufällig sind dagegen die Stichprobe und auch das aus ihr berechnete Konfidenzintervall. Richtig ist deshalb nur die Aussage E.

Aufgabe 8.10:* **Zensierte Daten**

Lösung: ***A***

Zensierte Daten sind problematisch, weil Information verloren geht und weil sie die Ergebnisse einer Studie verzerren können (B und D sind also falsch). Wenn sie dennoch auftreten, hat man darauf zu achten, dass die Gründe in keinem Zusammenhang mit den kritischen Endereignissen stehen (also A). Die Kaplan-Meier-Methode wertet soviel Information wie möglich aus (also ist C falsch). Leider reduziert sich der Stichprobenumfang mit wachsenden Zeitpunkten (E falsch).

9 Statistische Tests

Aufgabe 9.1: Formulieren der Hypothesen

Ein Forscher hofft, dass ein von ihm entwickeltes Medikament zur Blutdrucksenkung besser ist als ein herkömmliches Standardmedikament und will dies durch einen Test absichern. Wie soll er seine Vermutung formulieren?

A. als Nullhypothese
B. als Alternativhypothese
C. dies ist gleichgültig
D. dies hängt von den Folgen einer Fehlentscheidung ab
E. dies hängt von ethisch-moralischen Überlegungen ab

Aufgabe 9.2: Grundlagen

Welche Aussage ist ***falsch***?

A. Bei jedem Test schließen sich Annahme- und kritischer Bereich aus.
B. Die Größe des α-Fehlers beeinflusst die Größe des ß-Fehlers.
C. Ob 1- oder 2-seitig getestet wird, muss vor der Testdurchführung aufgrund sachlogischer Überlegungen entschieden werden.
D. Die Messwerte der Beobachtungseinheiten innerhalb einer Stichprobe müssen unabhängig voneinander sein.
E. Der Stichprobenumfang hat keinerlei Einfluss auf das Testergebnis.

Aufgabe 9.3: Testentscheidungen

Welche Aussage ist richtig?

A. Liegt nach der Durchführung eines Tests die Testgröße nicht im Annahmebereich, wird die Nullhypothese abgelehnt.
B. Die Größe des Fehlers 1. Art ist zufällig.
C. Wenn die Nullhypothese nicht abgelehnt wird, wird stets ein Fehler 2. Art gemacht.
D. Der Ablehnungsbereich ist immer ein zusammenhängendes Intervall.
E. Wenn die Alternativhypothese richtig ist, beträgt die Wahrscheinlichkeit, aufgrund des Testergebnisses falsch zu entscheiden, höchstens α.

Aufgabe 9.4: a -Fehler

Beim Test einer Nullhypothese H_0 gegen eine Alternativhypothese H_1 bedeutet eine Wahrscheinlichkeit $\alpha = 0,05$ für den Fehler 1. Art: die Wahrscheinlichkeit ist höchstens 0,05 dafür, dass man

A. H_1 annimmt, wenn H_1 richtig ist
B. H_0 beibehält, wenn H_0 richtig ist
C. H_0 nicht ablehnt, wenn H_1 richtig ist
D. H_0 ablehnt, obwohl H_0 richtig ist
E. H_1 fälschlicherweise ablehnt

Aufgabe 9.5: Tests und Merkmale

Welcher Test setzt qualitative Merkmale voraus?

A. der *t*-Test für 2 verbundene Stichproben
B. der Wilcoxon-Rangsummentest
C. der *F*-Test
D. der Chi^2-Homogenitätstest
E. keiner dieser Tests

Aufgabe 9.6: ß-Fehler – Allgemeines

Welche Aussage bzgl. des ß-Fehlers ist richtig?

A. Der ß-Fehler wird vor der Testdurchführung festgelegt und beträgt üblicherweise 5%.
B. Der ß-Fehler ist immer größer als der α-Fehler.
C. Der ß-Fehler kann durch den Stichprobenumfang beeinflusst werden.
D. Je größer die Power (Güte) eines Tests ist, um so größer ist auch der ß-Fehler.
E. Der α-Fehler und der ß-Fehler sind unabhängig voneinander.

Aufgabe 9.7: ß-Fehler beim *t*-Test

Gegeben seien 2 Grundgesamtheiten mit den Erwartungswerten μ_1 und μ_2 und der selben Varianz σ^2. Daraus werden 2 Stichproben gezogen und deren Mittelwerte mit dem *t*-Test für unverbundene Stichproben überprüft. Dabei wird die Wahrscheinlichkeit für den ß-Fehler größer, wenn alle Größen gleichbleiben, aber

A. der Stichprobenumfang größer wird
B. die Irrtumswahrscheinlichkeit α größer wird
C. der Betrag der Differenz $|\mu_1 - \mu_2|$ größer wird
D. die Varianz σ^2 größer wird
E. die Varianz σ^2 kleiner wird

Aufgabe 9.8: Nicht-signifikantes Testergebnis

Ein Forscher hat ein schmerzstillendes Präparat entwickelt und überprüft dessen Wirkung über einen statistischen Test. Er erhält ein nicht-signifikantes Testergebnis (mit $\alpha=0,05$). Wie ist dies zu interpretieren?

A. Damit ist bewiesen, dass sich das neue Präparat von einem Placebo grundlegend unterscheidet.
B. Damit ist bewiesen, dass das neu entwickelte Medikament wirkungslos ist.
C. Mit 95%-iger Wahrscheinlichkeit gibt es keinen Unterschied zwischen neuem Präparat und Placebo.
D. Das Testergebnis besagt, dass weitergehende Forschungen auf diesem Gebiet sinnlos sind.
E. Aufgrund des Ergebnisses lässt sich ein Unterschied zwischen neuem Präparat und Placebo nicht nachweisen. Ein ß-Fehler ist dabei jedoch nicht ausgeschlossen. Über dessen mögliche Ursachen muss nachgedacht werden.

Aufgabe 9.9: Voraussetzungen

Welche Tests setzen normalverteilte Daten voraus?

A. *t*-Tests
B. Wilcoxon-Rangsummentests
C. Chi^2-Tests
D. alle diese Tests
E. keiner dieser Tests

Aufgabe 9.10: Auswahl eines Tests bei normalverteilten Daten

Gegeben seien 2 unverbundene Stichproben mit sehr günstigen Voraussetzungen: die Daten entstammen aus 2 normalverteilten Grundgesamtheiten mit gleich großen Varianzen. Es soll überprüft werden, ob man Gleichheit der Erwartungswerte annehmen kann. Welchen Test sollte man bevorzugen?

A. den *t*-Test für 2 unverbundene Stichproben
B. den Welch-Test
C. den *U*-Test nach Mann, Whitney und Wilcoxon
D. Man wendet alle 3 Tests an und entscheidet sich dann für einen, der ein signifikantes Ergebnis liefert.
E. Es ist vollkommen gleichgültig, welchen Test man anwendet, weil die Voraussetzungen für jeden Test erfüllt sind.

Aufgabe 9.11: Auswahl eines Tests bei schief-verteilten Daten

Man vergleicht das mittlere Körpergewicht einer Patientengruppe, die ein Jahr lang eine bestimmte Diät zu sich genommen hat, mit dem mittleren Körpergewicht einer vergleichbaren Gruppe, die sich mit Normalkost ernährt hat. Man weiß, dass die Gewichte schief-verteilt sind und der Stichprobenumfang pro Gruppe nicht größer als 10 ist. Welcher Test eignet sich am ehesten?

A. der *t*-Test für verbundene Stichproben
B. der *t*-Test für unverbundene Stichproben
C. der *U*-Test nach Mann, Whitney und Wilcoxon
D. der Welch-Test
E. der Vorzeichentest

Aufgabe 9.12: *t*-Test für 2 unverbundene Stichproben

Zum *t*-Test für zwei unverbundene Stichproben werden 2 Stichproben der Umfänge n_1 und n_2 herangezogen. Welche Aussage ist falsch?

A. Dieser Test setzt gleiche Varianzen der Grundgesamtheiten voraus.
B. Die Umfänge müssen gleich groß sein.
C. Dieser Test setzt normalverteilte Grundgesamtheiten voraus.
D. Die Anzahl der Freiheitsgrade beträgt $f = n_1 + n_2 - 2$.
E. Man legt beim Testen die Nullhypothese $\mu_1 = \mu_2$ zugrunde.

Aufgabe 9.13: Chi^2-Test

Bei einem Chi^2-Test ergibt sich für den Wert der Prüfgröße χ^2=0. Was besagt dieses Ergebnis?

A. Dieses Ergebnis ist unmöglich, da die Prüfgröße nur positive Werte annehmen kann.
B. Aufgrund des Testergebnisses behält man die Nullhypothese bei; ein ß-Fehler ist bei dieser Entscheidung jedoch nicht auszuschließen.
C. χ^2=0 belegt eindeutig, dass die Nullhypothese richtig ist.
D. χ^2=0 belegt eindeutig, dass die Alternativhypothese richtig ist.
E. Ob man die Null- oder die Alternativhypothese annimmt, ist abhängig von der Größe des α-Fehlers.

Aufgabe 9.14: Vierfeldertest

Welche Aussage trifft ***nicht*** zu?

A. Dem Vierfelder-Test liegt die Chi^2-Verteilung zugrunde.
B. Die Anzahl der Freiheitsgrade beträgt immer 1.
C. Falls die Prüfgröße einen Wert größer als 3,84 annimmt, wird die Nullhypothese zugunsten der Alternativhypothese abgelehnt.
D. Mit diesem Test lässt sich die Unabhängigkeit 2er Alternativmerkmale überprüfen.
E. Die Prüfgröße kann generell Werte zwischen $-\infty$ und $+\infty$ annehmen.

Aufgabe 9.15: Vorzeichentest

Welche Aussage ist falsch?

A. Der Vorzeichentest ist bei 2 verbundenen Stichproben mit einem quantitativ-stetigen Merkmal anwendbar.
B. Das zugrunde liegende Modell ist die Binomialverteilung mit $p = 0{,}5$.
C. Der Test setzt 2 verbundene Stichproben mit gleichen Varianzen voraus.
D. Vor der Durchführung dieses Tests muss die Größe des α-Fehlers festgelegt werden.
E. Er hat eine geringere Power als der t-Test, wenn dessen Voraussetzungen erfüllt sind.

Aufgabe 9.16: Mehrfaches Testen

Gegeben seien 3 unverbundene Stichproben, die paarweise mit dem t-Test (jeweils $\alpha = 5\%$) getestet werden. Insgesamt werden also 3 Tests durchgeführt. Wie groß ist bei diesem Verfahren insgesamt der Fehler 1. Art?

A. $\alpha / 3$
B. 3α
C. $1 - 3\alpha$
D. α^3
E. $1-(1-\alpha)^3$

Aufgabe 9.1: **Formulieren der Hypothesen**

Lösung: ***B*** (Alternativhypothese)

Die Nullhypothese besagt immer, dass 2 Parameter (z. B. Erwartungswerte) gleich sind. Der Test geht davon aus, dass die Nullhypothese richtig ist; die Alternativhypothese (also Unterschiede) wird erst dann angenommen, wenn die Prüfgröße schwer mit der Nullhypothese zu vereinbaren ist.

Aufgabe 9.2: **Grundlagen**

Lösung: ***E*** (Aussage ist falsch)

Aussagen A–D sind richtig. Annahme- und kritischer Bereich schließen sich aus (siehe Tabelle Seite 210 und Abbildung Seite 211). Je größer der α-Fehler, desto kleiner der ß-Fehler (und umgekehrt; siehe Seite 206). Die Frage, ob 1- oder 2-seitig getestet wird, hat weniger mit Statistik als mit sachlogischen Überlegungen zu tun. Es ist auch leicht nachvollziehbar, dass sich die Beobachtungseinheiten innerhalb einer Stichprobe nicht beeinflussen dürfen, da sonst Ergebnisse verzerrt würden. Der Stichprobenumfang (Antwort E) hat sehr wohl Einfluss auf das Ergebnis: ein kleiner Umfang führt eher zur Beibehaltung der Nullhypothese, während ein extrem großer Umfang zur Annahme der Alternativhypothese tendiert.

Aufgabe 9.3: **Testentscheidungen**

Lösung: ***A***

Wenn die Prüfgröße nicht im Annahmebereich liegt, muss sie im kritischen Bereich liegen – dann nimmt man die Alternativhypothese an und lehnt die Nullhypothese ab. Zu B: die Größe des 1. Fehlers (α-Fehler) wird vor der Testdurchführung (üblicherweise mit α=5%) festgelegt, ist also nicht zufällig. Zu C: Wenn die Nullhypothese beibehalten wird, kann dies auch dadurch begründet sein, dass deren Aussage tatsächlich richtig ist – dann macht man keinen Fehler. Zu D: Der Ablehnungsbereich ist nicht immer zusammenhängend (z. B. bei 2-seitigem *t*-Test, Abb. 9.1). Zu E: Wenn die Alternativhypothese in Wirklichkeit richtig ist, kann man keinen α-Fehler machen (höchstens einen ß-Fehler).

Aufgabe 9.4: a **-Fehler**

Lösung: ***D***

Dass die Aussage D richtig ist, wird an der Tabelle 9.1 auf Seite 206 klar. Zu A und B: diese Entscheidungen wären richtig; es sind ***keine*** Fehler. Unter C wird der ß-Fehler beschrieben.

Zu E: diese Formulierung ist schlecht. Die Alternativhypothese kann man annehmen, aber nicht ablehnen (der Test geht ja von der Nullhypothese aus).

Aufgabe 9.5: Tests und Merkmale

Lösung: ***D*** (Chi^2-Homogenitätstest)

Für qualitative Merkmale eignen sich generell Chi^2-Tests. t-Tests und Rangsummentests setzen stetige (also quantitative) Daten voraus; ebenso der F-Test, der die Gleichheit 2er Varianzen untersucht.

Aufgabe 9.6: ß-Fehler – Allgemeines

Lösung: ***C*** (abhängig vom Stichprobenumfang)

Der ß-Fehler kann nicht festgelegt werden; seine Größe ist von vielen Faktoren abhängig (A und B sind falsch). Die Güte eines Tests lässt sich ausdrücken als 1–ß (das ist die Wahrscheinlichkeit, bei Gültigkeit der Alternativhypothese ein signifikantes Testergebnis zu erhalten). Damit ist D falsch. Auch E ist falsch: je kleiner α, um so größer wird ß (und umgekehrt). Durch einen großen Stichprobenumfang, der üblicherweise vor der Testdurchführung festgelegt wird, lässt sich ß jedoch klein halten (Antwort C).

Aufgabe 9.7: ß-Fehler beim *t*-Test

Lösung: ***D*** (größere Varianz)

Durch A–C wird der ß-Fehler verkleinert (siehe Aufgabe 9.6 und Ausführungen Seite 206f.). Es ist logisch, dass sich bei kleinerer Varianz ein Unterschied einfacher nachweisen lässt, was ebenfalls zu einer Verkleinerung des ß-Fehlers führt (E). Umgekehrt bewirkt eine Vergrößerung der Varibilität der Messwerte, dass sich ein Unterschied schwerer nachweisen lässt – also eine Vergrößerung des ß-Fehlers.

Aufgabe 9.8: Nicht-signifikantes Testergebnis

Lösung: ***E*** (vorsichtige Interpretation)

Mit einem Testergebnis lässt sich ***nichts*** beweisen – Antworten A und B sind offenkundig Unsinn. Ebenso C: ein eventueller Unterschied ist nicht vom Zufall abhängig. Zu voreilig wäre auch die Interpretation nach C. Anzuraten sind in jedem Fall vorsichtige Schlussfolgerungen, wie sie in Antwort E formuliert werden.

Aufgabe 9.9: Voraussetzungen

Lösung: ***A*** (t-Tests)

t-Tests setzen generell normalverteilte Daten voraus, auch wenn diese Voraussetzungen bei praktischen Anwendungen abgeschwächt werden können. Rangsummentests werden bei quantitativen Merkmalen verwendet, wenn diese Voraussetzungen nicht erfüllt sind. Chi^2-Tests setzten qualitative Merkmale voraus.

Aufgabe 9.10: Auswahl eines Tests bei normalverteilten Daten

Lösung: ***A*** (t-Test für 2 unverbundene Stichproben)

Man sollte generell alle Informationen so weit wie möglich ausnutzen. Wenn also bekannt ist, dass die Daten normalverteilt sind und die Varianzen gleich sind, sollte man den t-Test benutzen (der diese Eigenschaften voraussetzt). Die anderen beiden Tests haben schwächere Voraussetzungen (Welch-Test: keine gleich großen Varianzen, *U*-Test: keine Normalverteilung). Obgleich deren Voraussetzungen auch erfüllt sind, wäre es leichtsinnig, sie anzuwenden – sie haben nämlich eine geringere Power, und es ist schwieriger, einen Unterschied nachzuweisen. E ist also falsch, ebenso wie D: man sollte sich immer vor der Testdurchführung überlegen, welcher Test geeignet ist und nicht versuchen, im nachhinein zu manipulieren.

Aufgabe 9.11: Auswahl eines Tests bei schief-verteilten Daten

Lösung: ***C*** (*U*-Test)

Zunächst fallen der *t*-Test für verbundene Stichproben und der Vorzeichentest weg, da hier 2 unverbundene Stichproben vorliegen. Da die Voraussetzungen des *t*-Tests (auch des Welch-Tests) bzgl. Normalverteilung (u. a. wegen der kleinen Stichprobenumfänge) nicht einmal annähernd erfüllt sind, kann man diese nicht anwenden. Bleibt also nur noch der *U*-Test (der auch schief verteilte Daten zulässt).

Aufgabe 9.12: *t*-Test für 2 unverbundene Stichproben

Lösung: ***B*** (Umfänge müssen nicht gleich groß sein)

Dieser Test hat formal sehr strenge Voraussetzungen (siehe Abschnitt 9.2.3 im Buch), aber die Stichprobenumfänge müssen nicht gleich groß sein. Sie sollten aber wegen der Teststärke (Power) nicht allzu unterschiedlich sein.

Aufgabe 9.13: Chi²-Test

Lösung: ***B*** (Nullhypothese beibehalten)

Eine Prüfgröße mit dem Wert 0 kann sich durchaus ergeben, wenn nämlich $ad = bc$ (siehe Vierfeldertafel Seite 236). A ist also falsch. Mit einem Testergebnis – auch mit einem extremen – lässt sich nichts eindeutig belegen; demnach sind C und (erst recht) D falsch. Ebenso E: das Ergebnis liegt im Annahmebereich, deshalb muss man immer die Nullhypothese beibehalten. Eine vorsichtige Interpretation nach B ist angebracht.

Aufgabe 9.14: Vierfeldertest

Lösung: ***E*** (Prüfgröße ist nicht negativ)

Falsch ist E. Die Prüfgröße χ^2 beim Vierfeldertest ist – wie bei allen Chi²-Tests – größer oder gleich 0. Alle anderen Aussagen sind richtig (siehe Abschnitt 9.5.1).

Aufgabe 9.15: Vorzeichentest

Lösung: ***C*** (gleiche Varianzen werden ***nicht*** vorausgesetzt)

Der Vorzeichentest wird normalerweise angewandt zum Vergleich eines stetigen Merkmals bei 2 verbundenen Stichproben (Antwort A). Weiter beinhaltet er keine Voraussetzungen – auch nicht gleiche Varianzen. Das Testverfahren basiert auf der Binomialverteilung; unter der Nullhypothese gilt: $P(X > Y) = P(X < Y) = 0{,}5$ (Antwort B). Der α-Fehler muss immer vor der Durchführung festgelegt werden (Antwort D); dies ist keine Besonderheit des Vorzeichentests. Auch Antwort E ist richtig; der Vorzeichentest ist zwar universeller anwendbar, führt aber seltener zu einem signifikanten Ergebnis als der entsprechende *t*-Test.

Aufgabe 9.16: Mehrfaches Testen

Lösung: ***E*** ($1-(1-\alpha)^3 \approx 14\%$)

Wenn die Nullhypothese richtig ist, beträgt die Wahrscheinlichkeit, die Nullhypothese aufgrund des Testergebnisses beizubehalten $(1-\alpha) = 95\%$. Bei 3 Tests, die unabhängig voneinander durchgeführt werden, ist diese Wahrscheinlichkeit $(1-\alpha)^3 \approx 86\%$. Die Wahrscheinlichkeit, einmal fälschlicherweise eine Entscheidung zugunsten der Alternativhypothese zu treffen, beträgt dann $1-(1-\alpha)^3 \approx 14\%$. Dieser Wert ist fast 3mal so groß wie α – d. h. bei mehrfachem Testen steigt der Fehler 1. Art an.

Teil IV: Versuchsplanung

10 Grundlagen der Versuchsplanung

Aufgabe 10.1: Zufällige Stichproben

Von 20.000 Anästhesien sollen in den folgenden Monaten etwa 2.000 als Stichprobe gezogen und unter verschiedenen Gesichtspunkten ausgewertet werden. Die unter 1–5 genannten Verfahren liefern ungefähr die benötigte Anzahl. Welche der folgenden Mengen stellen zufällige Stichproben dar?

1. alle Anästhesien, für die bestimmte Oberärzte verantwortlich sind
2. alle Anästhesien, die an Patienten durchgeführt werden, deren Nachnamen mit einem der Buchstaben A–C beginnen
3. alle Anästhesien von Patienten, deren Geburtsdatum im Pass der 1., 2. oder 3. eines Monats ist
4. alle Anästhesien, die von der allgemeinchirurgischen Klinik durchgeführt wird
5. alle Anästhesien von Patienten im Alter zwischen 20 und 29 Jahren
6. alle Anästhesien, die montags durchgeführt werden

A. Alle Stichproben sind zufällig.
B. Nur 1–4 sind zufällig.
C. Nur 2–4 und 6 sind zufällig.
D. Nur 2 und 3 sind zufällig.
E. Keine Stichprobe ist zufällig.

Aufgabe 10.2: Repräsentative Stichprobe

Im Rahmen einer Doktorarbeit soll untersucht werden, wie die Heilung bei 120 Patienten verlaufen ist, die im vergangenen Jahr an einem Finger operiert worden sind. Zur Datenerhebung schickt man jedem dieser Patienten einen Fragebogen mit der Bitte, diesen auszufüllen und zurückzusenden. Man erhält die ausgefüllten Bögen von 80 Patienten. Kann man dann davon ausgehen, dass diese 80 Patienten eine repräsentative Stichprobe der 120 Patienten darstellen?

A. Ja, da der Stichprobenumfang $n = 80$ sehr groß ist
B. Ja, da der Stichprobenumfang im Verhältnis zur Grundgesamtheit sehr groß ist
C. Ja, da die Teilnahme an der Fragebogenaktion freiwillig erfolgte
D. Eine Antwort auf diese Frage hängt von den Skalenniveaus der auszuwertenden Merkmale ab.
E. Nein

Aufgabe 10.3: Versuchsplanung

Zu welchem Zeitpunkt sollte man sich bei einer klinischen Studie überlegen, welche statistische Analysemethode(n) verwendet werden sollen?

A. vor dem Formulieren der Fragestellungen
B. vor Beginn der Datensammlung
C. unmittelbar nachdem alle Daten vorliegen (vorher ist nicht möglich)
D. der geeignete Zeitpunkt ergibt sich von selbst im Laufe der Studie
E. dieser Zeitpunkt ist irrelevant

Aufgabe 10.4: Systematischer Fehler

Was trägt – wenn mehrere Stichproben untersucht werden – ***nicht*** dazu bei, den systematischen Fehler zu vermeiden?

A. Beobachtungsgleichheit
B. Strukturgleichheit
C. die Wahl eines geeigneten statistischen Modells
D. große Stichproben
E. repräsentative Stichproben

Aufgabe 10.5: Zufälliger Fehler

Was können die Ursachen für den zufälligen Fehler sein?

1. kleine Stichproben
2. die intraindividuelle Variabilität der Beobachtungseinheiten
3. der interindividuelle Variabilität der Beobachtungseinheiten
4. nicht-repräsentative Stichproben
5. fehlerhaft durchgeführte Messungen (z. B. falsch geeichtes Messgerät)

A. alles unter 1–5
B. nur 1–3
C. nur 1–4
D. nur 2 und 3
E. nur 1

Aufgabe 10.6: Blockbildung

Zum Vergleich 2er Augensalben bildet man sinnvollerweise aus den Augenpaaren mehrerer Patienten Blöcke. Welche Aussagen treffen zu?

1. Durch das Bilden von Blöcken wird der systematische Fehler weitgehend ausgeschaltet.
2. Innerhalb eines Blocks wird randomisiert.
3. Die Randomisierung trägt zur Strukturgleichheit der beiden Behandlungsgruppen bei.
4. Die Blockbildung trägt zur Beobachtungsgleichheit bei.
5. Der Einfluss von Störgrößen wird ausgeschaltet.

A. nur 2 und 3
B. nur 1, 3 und 5
C. nur 3, 4 und 5
D. nur 2, 3 und 4
E. nur 1, 3, 4 und 5

Aufgabe 10.7: Bilden von Schichten

Bei größeren Untersuchungen fasst man Beobachtungseinheiten, die sich in wesentlichen Eigenschaften ähneln, in einer Schicht zusammen. Was ist daran vorteilhaft?

1. Dadurch wird der systematische Fehler ausgeschaltet.
2. Es entstehen übersichtliche Gruppen.
3. Der zufällige Fehler wird reduziert.
4. Der Versuchsfehler insgesamt wird reduziert.
5. Unterschiede zwischen den einzelnen Schichten sind dann klarer erkennbar.

A. 1–5
B. nur 2 und 3
C. nur 2–4
D. nur 2–5
E. nur 1 und 3

Aufgabe 10.8: Anzahl von Schichten

Bei einer klinischen Untersuchung erscheint es wünschenswert, nach 3 Merkmalen zu schichten: 1. nach Geschlecht, 2. nach Alter (5 Klassen von jeweils 10 Jahren) und 3. nach Krankheitsstatus (leicht, mittelmäßig und schwer erkrankt). Wie viele Schichten ergeben sich nach diesem Verfahren?

A. 1
B. 3
C. 30
D. 60

E. Die Anzahl der Schichten ergibt sich erst im Laufe der Untersuchung.

Aufgabe 10.9: Zuteilen zu Behandlungsgruppen

Eine streng zufällige Zuteilung auf 2 Behandlungsgruppen wird am ehesten erzielt

A. mit Hilfe eines Würfels oder einer Zufallszahl
B. indem man den Patienten die Gruppe wählen lässt
C. indem man den behandelnden Arzt die Gruppe wählen lässt
D. indem man eine zufällig anwesende Person bittet, eine Zahl zwischen 1 und 8 zu nennen und danach entscheidet, ob diese Zahl gerade oder ungerade ist
E. durch systematisches Alternieren

Aufgabe 10.10: Doppelblindstudien

Wozu trägt die Doppelblindheit bei klinischen Studien ***nicht*** bei?

A. systematische Fehler zu vermeiden
B. Beobachtungsgleichheit zu erreichen
C. Datenschutz zu gewährleisten
D. Autosuggestion auf Seiten des behandelnden Arztes auszuschalten
E. Autosuggestion auf Seiten der Patienten auszuschalten

Aufgabe 10.1: Zufällige Stichproben

Lösung: ***E*** (keine)

Zufällig bedeutet: jedes Element der Grundgesamtheit hat dieselbe Chance, in die Stichprobe zu gelangen. Damit ist eigentlich klar, dass keine dieser Stichproben zufällig ist; sie sind alle systematisch. Man darf deshalb auch nicht annehmen, dass sie repräsentativ für die Menge aller Anästhesien sind. Zu 1: Oberärzte führen schwierigere OPs durch als beispielsweise AIP-ler. Zu 2: Damit wären evtl. Mitglieder einer Familie mit dem selben Namen oder auch Patienten aus bestimmten Ländern übermäßig häufig vertreten. Zu 3: Hier ist zu berücksichtigen, dass bei einigen Leuten das Geburtsdatum nicht bekannt ist und im Pass dann oft der 1. eines Monats eingetragen ist. Zu 4: Die Allgemeinchirurgie ist ***eine*** Klinik und nicht unbedingt repräsentativ. Zu 5: Eine einzige Altersgruppe kann unmöglich repräsentativ für alle Patienten sein. Zu 6: Montags gibt es hauptsächlich planmäßige Operationen (und weniger Notfälle als am Wochenende). Fazit: Am ehesten erhält man eine zufällige (und damit repräsentative) Stichprobe, wenn man einen Zufallszahlengenerator entscheiden lässt. Falls dies nicht möglich ist, wird manchmal eine systematische Stichprobe gezogen (am ehesten nach 2 oder 3). Dabei muss man aber sehr aufpassen, dass man keinen allzu großen systematischen Fehler begeht.

Aufgabe 10.2: Repräsentative Stichprobe

Lösung: ***E*** (nein)

Zunächst: ob eine Stichprobe repräsentativ ist, hängt nicht mit ihrer Größe oder den Merkmalen zusammen. A, B und E sind demnach falsch. Dass die Teilnahme freiwillig erfolgte, ist keinerlei Hinweis auf Repräsentativität (Antwort D). Man hat in jedem Fall nach den Gründen zu fragen, aus denen immerhin 40 Patienten ihren Bogen nicht ausgefüllt haben. Das könnte auch damit zusammenhängen, dass sie Schwierigkeiten mit dem operierten Finger haben und deshalb nicht schreiben können. Wenn man diese 40 Patienten nicht berücksichtigt, würden die Ergebnisse der Studie sehr verzerrt werden (systematischer Fehler). Fazit: Fragebogenaktionen sind immer problematisch. Man sollte in jedem Fall bei Nicht-Respondern nachhaken und versuchen, die Gründe für die Nicht-Teilnahme herauszufinden.

Aufgabe 10.3: Versuchsplanung

Lösung: ***B*** (vor Datenerhebung)

Man sollte so früh wie möglich überlegen, welche Analysemethoden angewandt werden. Vor dem Formulieren der Fragestellungen wäre allerdings zu früh (Antwort A): zunächst müssen ja alle Merkmale bekannt sein, die man auswerten will – und diese ergeben sich erst aus der Fragestellung. Aus den gewählten Verfahren ergeben sich der benötigte Stichprobenumfang und damit wichtige Informationen zum weiteren Ablauf der Studie. Nachdem die Daten vorliegen (Antwort C) wäre also spät. Die Antworten D und E sind offensichtlich unsinnig.

Aufgabe 10.4: Systematischer Fehler

Lösung: ***D*** (große Stichproben)

Große Stichproben tragen dazu bei, den ***zufälligen*** Fehler zu reduzieren; repräsentative Stichproben tragen dazu bei, den systematischen zu vermeiden. Auch Beobachtungs- und Strukturgleichheit 2er Stichproben sind wichtig (sonst würde man Äpfel mit Birnen vergleichen; Antworten A und B). Ein ungeeignetes Modell (Antwort C) kann zu verzerrten Ergebnissen und damit auch zu einem systematischen Fehler führen.

Aufgabe 10.5: Zufälliger Fehler

Lösung: ***B***

Je kleiner der Stichprobenumfang ist und je mehr die Messwerte variieren, um so größer wird der zufällige Fehler (der Standardfehler des Mittelwerts ist $s/\sqrt{n}$; siehe Seite 267). Die intra- und interindividuelle Variabilität sind hier maßgebend. An dieser Stelle sei darauf hingewiesen, dass im Buch intra- und interindividuelle Variabilität verwechselt ist (Seite 266). Es muss heißen: ***interindividuelle Variabilität*** (bei mehreren Beobachtungseinheiten erhält man beim Messen einer Größe unterschiedliche Ergebnisse) und ***intraindividuellen Variabilität*** (bei einer Beobachtungseinheit ergeben sich beim Messen einer Größe unter denselben Bedingungen unterschiedliche Werte). Pardon! Zu den Punkten 4 und 5: nicht-repräsentative Stichproben und fehlerhafte Messungen sind verantwortlich für den systematischen Fehler.

Aufgabe 10.6: Blockbildung

Lösung: ***A***

Richtig ist: durch die Blockbildung wird der zufällige Fehler innerhalb eines Blocks reduziert (weil sich dessen Einheiten weitgehend gleichen). Mit dem systematischen Fehler hat dies zunächst nichts zu tun; also ist 1 falsch. Innerhalb des Blockes wird randomisiert; das heißt, der Zufall entscheidet, welches Auge mit welcher Therapie behandelt wird. Dadurch erhält man 2 Gruppen (die jeweils genau ein Auge von jedem Patienten enthalten), die sich in wesentlichen Eigenschaften gleichen, und trägt damit zur Strukturgleichheit bei. Also sind 2 und 3 richtig. Die Beobachtungsgleichheit (z. B. gleiche Messmethoden in beiden Gruppen) wird nicht durch die Blockbildung beeinflusst; ebenso wenig werden Störgrößen dadurch ausgeschaltet (diese lassen sich nie komplett ausschalten). Daher sind 4 und 5 falsch.

Aufgabe 10.7: Bilden von Schichten

Lösung: ***D***

Bei Schichten handelt es sich wie bei Blöcken um übersichtliche, weitgehend homogene Gruppen, wodurch der zufällige Fehler reduziert wird. Dadurch wird auch der gesamte Versuchsfehler, der sich aus dem zufälligen und dem systematischen Fehler zusammensetzt, reduziert. Es ist klar, dass dann Unterschiede zwischen den Schichten deutlicher sind. Also sind 2–5 richtig. Mit dem systematischen Fehler hat dies nichts zu tun; 1 ist also falsch.

Aufgabe 10.8: Anzahl von Schichten

Lösung: ***C*** (30)

Dies ist leicht auszurechnen, indem man die Anzahl der Merkmalsausprägungen bzw. Klassen multipliziert: $2(\text{Geschlecht}) \cdot 5(\text{Altersklassen}) \cdot 3(\text{Krankheitsstatus}) = 30$. Dies zeigt, dass eine Schichtung nach wichtigen Einflussfaktoren zwar wünschenswert ist, in der Praxis aber an Grenzen stößt, weil sehr viel Beobachtungseinheiten erforderlich wären.

Aufgabe 10.9: Zuteilen zu Behandlungsgruppen

Lösung: ***A*** (Würfel)

Alles andere wäre eine Zuteilung nach einem bestimmten System und damit nicht zufällig, auch wenn dies auf den ersten Blick nicht erkennbar ist. Sowohl die Patienten als auch der behandelnde Arzt können subjektiven Einflüssen bei der Wahl der Behandlungsgruppe unterliegen (B und C). Wenn man eine Person bittet, eine Zahl zwischen 1 und 8 zu wählen, wird erfahrungsgemäß die 7 häufiger als andere Zahlen genannt (D). Auch systematisches Alternieren kann nicht gewährleisten, dass die Zuteilung zufällig erfolgt (E).

Aufgabe 10.10: Doppelblindstudien

Lösung: ***C*** (Datenschutz ***nicht*** gewährleistet)

Studien werden doppelblind durchgeführt, um Autosuggestion zu vermeiden (D und E). Dies trägt zur Beobachtungsgleichheit und damit zur Vermeidung systematischer Fehler bei (B und A). Mit dem Datenschutz hat dies aber nichts zu tun.

11 Studientypen

Aufgabe 11.1: Studie bzgl. OP-Risiken

In einer Klinik sollen die Häufigkeiten für das Auftreten von Komplikationen bei Operationen und deren mögliche Ursachen untersucht werden. Als Grundlage werden alle Anästhesieprotokolle aus den vergangenen 12 Monaten herangezogen. Um welchen Studientyp handelt es sich?

A. retrospektve Studie
B. prospektive Studie
C. klinisch-kontrollierte Studie
D. Kohortenstudie
E. Experiment

Aufgabe 11.2: Fall-Kontroll-Studien

Welche Aussage bzgl. Fall-Kontroll-Studie ist richtig?

A. Es handelt sich um prospektive Studien.
B. Die Anzahl der Fälle muss gleich der Anzahl der Kontrollen sein.
C. Diese Studien dienen in der Regel zur Klärung ätiologischer Faktoren.
D. Es besteht die Notwendigkeit, Patienten mehrfach zu untersuchen.
E. Die Matched-Pairs-Technik dient der Beobachtungsgleichheit.

Aufgabe 11.3: Retrospektive Studien

Welche möglichen Nachteile haben retrospektive Studien?

1. mangelhafte Datenqualität
2. kein Einfluss auf die Wahl der Beobachtungseinheiten
3. erhöhter Zeitbedarf im Vergleich zu prospektiven Studien
4. meist wesentlich höhere Kosten als bei prospektiven Studien
5. sie sind nicht geeignet, um kausale Zusammenhänge nachzuweisen

A. 1–5 sind richtig
B. nur 1–3 ist richtig
C. nur 1 und 2 sind richtig
D. nur 1, 2 und 5 sind richtig
E. nur 1 ist richtig

Aufgabe 11.4: Matched-Pairs-Technik

Warum wird bei einigen Fall-Kontroll-Studien die Matched-Pairs-Technik angewandt?

A. Um gleich große Gruppen zu erhalten.
B. Um Strukturgleichheit zu erreichen.
C. Um Beobachtungsgleichheit zu erreichen.
D. Um Störgrößen auszuschalten.
E. Um den zufälligen Fehler zu reduzieren.

Aufgabe 11.5: Kontrollierte klinische Therapiestudien

Worauf ist das Attribut „kontrolliert" zurückzuführen?

A. darauf, dass der α-Fehler kontrollierbar ist
B. auf die Kontrollgruppe
C. darauf, dass der Umfang der Behandlungsgruppen kontrolliert werden kann
D. darauf, dass der Therapieerfolg in Abhängigkeit der Behandlungsform evaluiert werden kann
E. darauf, dass die Ausprägungen der Einflussgrößen (also die Behandlungsform) vorgegeben werden können

Aufgabe 11.6: Kohortenstudien

Welche Aussage trifft ***nicht*** zu? Bei einer Kohortenstudie

A. geht man oft von einer Gruppe von exponierten und einer Gruppe von nicht-exponierten Personen aus
B. muss man meist längere Zeit auf das Eintreten der Zielereignisse warten
C. handelt es sich um eine reine Beobachtungsstudie
D. werden die Ausprägungen der Einflussfaktoren vom behandelnden Arzt bestimmt
E. hat der Versuchsleiter Einfluss auf die Datenerfassung und die vollständige und richtige Dokumentation der Daten

Aufgabe 11.7: Klinisch kontrollierte Studien – Eigenschaften

Welche Aussage bezüglich randomisierter, klinischer Therapiestudien trifft ***nicht*** zu?

A. Doppelblindheit ist unbedingt erforderlich.
B. Struktur- und Beobachtungsgleichheit sind unbedingt erforderlich.
C. Kein Patient darf gezwungen werden, an einer solchen Studie teilzunehmen.
D. Die Randomisierung dient der Strukturgleichheit.
E. Sie kommt in ihrer wissenschaftlichen Aussagekraft einem Experiment nahe.

Aufgabe 11.8: Klinisch kontrollierte Studien – Vorgehensweise

Im Rahmen einer klinischen Therapiestudie soll ein neu entwickeltes Medikament N gegen die herkömmliche Standardtherapie S getestet werden. Welche Vorgehensweise ist korrekt?

A. Man teilt die Patienten zufällig einer der Therapien N oder S zu und behandelt die beiden Gruppen im selben Zeitraum in 2 verschiedenen Krankenhäusern.
B. Man behandelt nur eine Gruppe mit N und vergleicht dann mit einer Gruppe, die in der Vergangenheit mit S behandelt wurde.
C. Um Risiken zu vermeiden, behandelt man nur die leichter erkrankten Patienten mit N und die schwereren Fälle mit S (gleichzeitig, in der selben Einrichtung).
D. Man teilt die Patienten zufällig einer der Therapien N oder S zu. Die beiden Gruppen werden gleichzeitig und von demselben Personal behandelt und beobachtet.
E. Man lässt die Patienten entscheiden, ob sie mit N oder S behandelt werden wollen.

Aufgabe 11.9: Vergleich retrospektive und prospektive Studie

Es soll untersucht werden, ob die Rauchgewohnheiten schwangerer Frauen einen Einfluss auf das Auftreten von Frühgeburten haben. Dies ist theoretisch mit einer retro- oder einer prospektiven Studie möglich. Welche der folgenden Aussagen ist falsch?

A. Eine retrospektive Studie basiert auf einer Gruppe von Müttern mit frühgeborenen Kindern und einer anderen Gruppe, deren Babys nach normaler Schwangerschaftsdauer zur Welt kamen.
B. Wenn die Studie prospektiv durchgeführt wird, ist es sinnvoll, eine große Anzahl schwangerer Frauen (bestehend aus Nicht-Raucherinnen, schwachen, mäßigen und starken Raucherinnen) heranzuziehen.
C. Beide Studientypen sind geeignet, einen kausalen Zusammenhang zwischen Rauchgewohnheiten werdender Mütter und Frühgeburten nachzwueisen.
D. Die Daten der retrospektiven Studie können sofort analysiert werden, während man bei der prospektiven Studie mehrere Monate warten muss.
E. Bei beiden Studientypen handelt es sich um Beobachtungsstudien.

Aufgabe 11.10: Letzte Aufgabe

Eine pharmazeutische Firma hat ein neues Schmerzmittel entwickelt und lässt dies von 10 verschiedenen Kliniken unter annähernd gleichen Bedingungen gegen ein Placebo testen. Die Ergebnisse werden für jede einzelne Klinik mit dem selben statistischen Test ausgewertet, wobei jeweils $\alpha = 5\%$ festgelegt wird. Nur bei einer einzigen Klinik XY konnte ein signifikanter Unterschied nachgewiesen werden. Daraufhin veröffentlicht die Firmenleitung einen Fachartikel mit den Untersuchungsergebnissen der Klinik XY und schreibt: „Die schmerzstillende Wirkung unseres neu-entwickelten Präparats wurde in einer Untersuchung an der Klinik XY nachgewiesen ($\alpha = 5\%$)". Wie beurteilen Sie dieses Verfahren?

A. Dies ist korrekt, weil ja nur die Ergebnisse der Klinik XY publiziert werden.
B. Dies ist korrekt, weil der α-Fehler angegeben wurde.
C. Man kann dies nur beurteilen, wenn man den gesamten Artikel gelesen hat.
D. Dieses Verfahren ist abzulehnen, weil es nicht zu verantworten ist, gegen ein Placebo zu testen.
E. Diese Vorgehensweise grenzt an Betrug und ist rundweg abzulehnen.

Aufgabe 11.1: Studie bzgl. OP-Risiken

Lösung: ***A*** (retrospektiv)

Die relevanten Ereignisse geschahen in der Vergangenheit und werden im nachhinein analysiert. Dies ist das Kennzeichen einer retrospektiven Studie (siehe Seite 269). Hier wird der Nachteil dieser Studienart bzgl. der mangelhaften Datenqualität sehr deutlich: man kann nicht annehmen, dass jeder Anästhesist die aufgetretenen Komplikationen fein säuberlich und vollständig dokumentiert hat.

Aufgabe 11.2: Fall-Kontroll-Studien

Lösung: ***C*** (zur Klärung ätiologischer Faktoren)

Man geht aus von einer Gruppe erkrankter Personen (Fälle) und einer Gruppe nicht-erkrankter (Kontrollen) und versucht mögliche Ursachen für diese Krankheit herauszufinden. Es handelt sich um eine retrospektive Studie (A ist falsch); die Anzahl der Fälle muss nicht unbedingt der Anzahl der Kontrollen entsprechen; oft steht eine Kontroll-Gruppe zur Verfügung, die wesentlich größer ist als die Gruppe der Fälle (B ist falsch). Patienten müssen nicht untersucht werden, weil die relevanten Ereignisse bei Studienbeginn bereits eingetreten sind (D ist falsch). Wenn Matched-Pairs-Technik angewandt wird (dies muss nicht sein), dient dies der Strukturgleichheit (E ist falsch).

Aufgabe 11.3: Retrospektive Studien

Lösung: ***D*** (1, 2 und 5)

Weil bei dieser Studienart die Ereignisse bereits eingetreten sind, muss man auf alte Dokumente zurückgreifen oder Patienten befragen. Man kann nicht unbedingt sicher sein, dass alle Daten richtig und vollständig dokumentiert sind oder dass man bei Befragungen richtige Antworten erhält. Im nachhinein hat man auch keinen Einfluss mehr auf die Auswahl der Beobachtungseinheiten. Deshalb sind 1 und 2 richtig. Man kann zwar evtl. Zusammenhänge nachweisen, aber es wäre zu gewagt, Kausalitäten herzuleiten. Dafür eignen sich prospektive Studien weit mehr. Also ist 5 richtig. – Der Zeitbedarf und die Kosten sind allerdings bei retrospektiven Studien weit geringer als bei prospektiven (3 und 4 sind also falsch).

Aufgabe 11.4: Matched-Pairs-Technik

Lösung: ***B*** (Strukturgleichheit)

Man sucht dabei zu jedem Fall eine passende Kontrolle (siehe Seite 274). Damit erreicht man, dass die beiden Gruppen bzgl. wichtiger Einflussgrößen ähnlich sind

(Strukturgleichheit). Man erreicht damit auch, dass die beiden Gruppen gleich groß sind – dies ist aber ein Nebeneffekt und nicht der eigentliche Sinn dieser Technik (das könnte man auch einfacher erreichen). A ist also falsch. Zu C, D und E: Mit Beobachtungsgleichheit hat dies nichts zu tun; Störgrößen werden dadurch nicht ausgeschaltet. Die Strukturgleichheit trägt dazu bei, den systematischen Fehler auszuschalten, nicht den zufälligen zu reduzieren.

Aufgabe 11.5: Kontrollierte klinische Therapiestudie

Lösung: ***E*** (Einflussgrößen sind kontrollierbar)

Hierin ähnelt eine klinische Studie einem Experiment, bei dem ja auch vom Versuchsleiter die Einflussgrößen vorgegeben werden (man spricht dann von ***kontrollierten*** Bedingungen). Der wesentliche Unterschied zu einem Experiment besteht darin, dass das Wohl des Patienten im Vordergrund steht und dass deshalb – falls erforderlich – der behandelnde Arzt modifizierend eingreifen darf.

Aufgabe 11.6: Kohortenstudien

Lösung: ***D*** (ist falsch)

Typisch ist die unter A beschriebene Situation: man geht von einem Kollektiv gesunder Personen (exponierte und nicht-exponierte) aus und wartet, bis eine bestimmte Krankheit eintritt. Dies dauert oft lange (B ist richtig). Dabei werden die Personen nur beobachtet; d. h. es werden keine Ausprägungen vorgegeben (C richtig, D falsch). Vorteilhaft ist, dass es der Versuchsleiter s11.9in der Hand hat, vollständige und richtige Daten zu erhalten (E, im Gegensatz zu einer retrospektiven Studie).

Aufgabe 11.7: Klinisch-kontrollierte Studien – Eigenschaften

Lösung: ***A*** (Doppelblindheit ist ***nicht*** erforderlich)

Doppelblindheit ist zwar wünschenswert, aber nicht immer realisierbar (z. B. bei chirurgischen Eingriffen). Alles andere B–E ist richtig. Strukturgleichheit und Beobachtungsgleichheit sind immer erforderlich, wenn mehrere Gruppen verglichen werden. Siehe auch Seite 275ff.

Aufgabe 11.8: Klinisch-kontrollierte Studien – Vorgehensweise

Lösung: ***D***

Nur bei dieser Vorgehensweise ist Beobachtungs- und Strukturgleichheit möglich. Bei A (verschiedene Kliniken) wäre Beobachtungsgleichheit nicht gegeben. Auch nicht bei B, weil hier die Zeit als Störgröße die Ergebnisse evtl. verzerren könnte. Dieses Verfahren kann nur unter ganz bestimmten Bedingungen angewandt werden (z. B. wenn ein neues Medikament geprüft wird, das lebensrettend ist und für das keine Alternative existiert). Der Vorschlag unter C würde zu sehr verzerrten Ergebnissen führen, die gar nicht sinnvoll interpretierbar sind. Auch das Verfahren unter E ist nicht zufällig und deshalb abzulehnen.

Aufgabe 11.9: Vergleich retrospektive und prospektive Studie

Lösung: ***C*** (kausaler Zusammenhang kann nicht nachgewiesen werden)

Es ist allenfalls möglich, einen Zusammenhang nachzuweisen. Die Frage, ob eine Frühgeburt durch die Rauchgewohnheiten der Mütter kausal bedingt ist, kann mit einer retrospektiven Studie jedoch nicht beantwortet werden. Eine prospektive Studie kann zwar einen kausalen Zusammenhang nicht direkt nachweisen (dazu müsste man ein Experiment durchführen); sie hat jedoch eine wissenschaftlich höhere Aussagekraft als eine retrospektive Studie.

Aufgabe 11.10: Letzte Aufgabe

Lösung: ***E*** (Betrug)

Dieses Verfahren ist wissenschaftlich in allerhöchstem Maße unseriös. Nicht etwa weil mit einem Placebo verglichen wird (eine Placebostudie kann durchaus gerechtfertigt und ethisch vertretbar sein), sondern weil bei 10-maligem Testen die Wahrscheinlichkeit, unter der Nullhypothese eine falsche Entscheidung zu treffen, sehr groß ist (der Fehler 1. Art beträgt dann nicht 0,05, sondern $1-0{,}95^{10} \approx 0{,}40$). Wenn die Firma in hinreichend vielen Kliniken testen lässt, wird sie bestimmt irgendwo rein zufällig einen Unterschied bzgl. der Wirkung (fälschlicherweise) nachweisen können. In einer Publikation müssten korrekterweise die Ergebnisse aus allen Kliniken dargestellt werden; dies wäre allerdings wenig werbewirksam.