

Bonusquiz I – Woche 1

Es werden zwei unterschiedliche Würfel gleichzeitig geworfen. Die Würfel sind fair, d.h., die Augenzahlen 1 bis 6 eines Würfels treten mit gleicher Wahrscheinlichkeit auf.

Wählen Sie alle Aussagen aus, die korrekt sind (Runden Sie auf drei Nachkommastellen):

Wählen Sie eine oder mehrere Antworten:

- ☐ a. Die Wahrscheinlichkeit, dass die Augensumme 8 ist, ist 0.139 .
- ☐ b. Die Wahrscheinlichkeit, dass die Augensumme kleiner gleich 6 ist, ist 0.583 .
- ☐ c. Die Wahrscheinlichkeit eine gerade Augensumme zu wuerfeln ist 0.569 .
- ☐ d. Die Wahrscheinlichkeit, dass beide Wuerfel die gleiche Augenzahl zeigen ist 0.222 .

An einer Supermarktkasse stellen sich 5 Personen (A,B,C,D,E) in einer völlig zufälligen Reihenfolge an. Nehmen Sie an, dass alle Reihenfolgen gleich wahrscheinlich sind.

Wählen Sie alle Aussagen aus, die korrekt sind, wenn keine weiteren Annahmen gemacht werden (Runden Sie auf drei Nachkommastellen):

Wählen Sie eine oder mehrere Antworten:

- ☐ a. Die Wahrscheinlichkeit, dass die Reihenfolge an der Supermarktschlange A, E, D, B, C ist, ist 0.076 .
- ☐ b. Die Wahrscheinlichkeit, dass Person C an Position 1 und Person A an Position 5 steht, ist 0.054 .
- ☐ c. Die Wahrscheinlichkeit, dass die erste Person an der Supermarktschlange nicht Person A ist, ist 0.8 .
- ☐ d. Die Wahrscheinlichkeit, dass die Reihenfolge an der Supermarktschlange nicht B, E, A, C, D ist, ist 0.938 .

Verwenden Sie Venn-Diagramme um folgende Aussagen bzgl. der beiden Ereignisse A und B zu prüfen.

Wählen Sie alle Aussagen aus, die korrekt sind, wenn keine weiteren Annahmen gemacht werden:

Wählen Sie eine oder mehrere Antworten:

- ☐ a. $P(A \setminus B) = P(A - B)$
- ☐ b. $P(A) - P(B^c) = P(A \cap B^c)$
- ☐ c. $P(A) + P(B) = P(A \cap B)$
- ☐ d. $P(A) - P(A \cap B) = P(A \cup B^c)$

Bonusquiz I – Woche 2

Im Folgenden finden Sie verschiedene Aussagen zum Thema Wahrscheinlichkeit, odds und log-odds. Wählen Sie alle Aussagen aus, die korrekt sind (Runden Sie auf drei Nachkommastellen):

Wählen Sie eine oder mehrere Antworten:

- ☐ a. Angenommen die log-odds fuer ein Ereignis sind 1.45 , dann sind die odds fuer das gleiche Ereignis 7.048 .
- ☐ b. Angenommen die odds erniedrigen sich fuer ein gewisses Ereignis, dann erhoehen sich die log-odds fuer dasselbe Ereignis.
- ☐ c. Folgender Zusammenhang gilt ohne weitere Annahmen fuer alle Ereignisse A : $P(A^c) = 1 - P(A)$
- ☐ d. Die log-odds mit einem fairen Wuerfel eine Zahl in der Menge 4,6,1,3,2 zu wuerfeln, sind 1.609 .
- ☐ e. Keine der oben genannten Aussagen ist richtig.

Ein fairer, sechsseitiger Wuerfel wird einmal geworfen. Das Ereignis A tritt ein, wenn eine Zahl in der Menge $\{5,6,4,2\}$ gewuerfelt wird. Das Ereignis B tritt ein, wenn eine Zahl in der Menge $\{4,6\}$ gewuerfelt wird. Wählen Sie alle Aussagen aus, die korrekt sind (Runden Sie auf drei Nachkommastellen):

Wählen Sie eine oder mehrere Antworten:

- ☐ a. Die Wahrscheinlichkeit, dass A nicht eintritt gegeben B ist eingetreten ist 0.245 .
- ☐ b. Die Wahrscheinlichkeit, dass B nicht eintritt gegeben A ist eingetreten ist 0.565 .
- ☐ c. Die Wahrscheinlichkeit, dass A nicht eintritt gegeben B ist nicht eingetreten ist 0.775 .
- ☐ d. Die Wahrscheinlichkeit, dass B nicht eintritt gegeben A ist nicht eingetreten ist 0.198 .
- ☐ e. Keine der oben genannten Aussagen ist richtig.

Ein neuer Spam-Filter wurde trainiert. Eine eingehende mail kann spam oder korrekt sein. Der Spam-Filter kann dann Alarm schlagen (also die mail als spam deklarieren) oder nicht. Die Wahrscheinlichkeit, dass eine zufaellig ausgewaehlte, eingehende email Spam ist, ist 0.08. Eine Spam-mail wird mit einer Wahrscheinlichkeit von 0.77 als solche erkannt. Wenn eine korrekte email eingeht, wird sie mit einer Wahrscheinlichkeit von 0.17 faelschlicherweise als Spam markiert. Wählen Sie alle Aussagen aus, die korrekt sind (Runden Sie auf drei Nachkommastellen):

Wählen Sie eine oder mehrere Antworten:

- ☐ a. Eine eingehende mail wird zufaellig ausgewaehlt. Die Wahrscheinlichkeit, dass diese mail als spam-mail deklariert wird, ist 0.562 .
- ☐ b. Eine eingehende mail wird zufaellig ausgwaeht. Die Wahrscheinlichkeit, dass diese mail spam ist und auch als spam deklariert wird, ist 0.062 .
- ☐ c. Die Wahrscheinlichkeit, dass eine mail eine echte spam-mail ist, wenn unser Spam-Filter Alarm geschlagen hat, ist 0.517 .
- ☐ d. Die Wahrscheinlichkeit, dass eine mail eine korrekte mail ist, wenn sie als Spam-Mail erkannt wurde, ist 0.716 .
- ☐ e. Keine der oben genannten Aussagen ist richtig.

Bonusquiz I – Woche 3

Bei einem Casinospiel wird zufällig eine Zahl aus der Menge 1,2,3,4,5,6 gezogen. Die gezogene Zahl entspricht gerade dem Gewinn X in SFr, der einem dann ausgezahlt wird. Die Teilnahme an dem Spiel kostet 2.5 SFr pro Runde. Folgende Tabelle zeigt die Wahrscheinlichkeiten, mit denen jede einzelne Zahl gezogen wird:

Zahl	1	2	3	4	5	6
Wa.	0.25	0.27	0.21	0.10	0.10	0.07

Wählen Sie alle Aussagen aus, die korrekt sind (Runden Sie auf drei Nachkommastellen):

Wählen Sie eine oder mehrere Antworten:

- ☐ a. Die Wahrscheinlichkeit, dass die gezogene Zahl grösser gleich 2 ist, ist 0.25 .
- ☐ b. Die Wahrscheinlichkeit, dass die gezogene Zahl verschieden von 3 ist, ist 0.9 .
- ☐ c. Die Varianz des Gewinns nach einer Runde ist 2.374 (Einheit: SFr²).
- ☐ d. Angenommen, jeder mögliche Gewinn in dem Spiel vergrössert sich um plus 3 SFr. Die Zufallsvariable, die den neuen Gewinn beschreibt ist also nicht mehr X sondern $X+3$. Dann gilt: $\text{Var}(X+3)=\text{Var}(X)+3$.
- ☐ e. Keine der oben genannten Aussagen ist richtig.

Bei klinischen Studien wird in der Phase 2 oft die Wirksamkeit von Medikamenten abgeschätzt. Dazu wird meist die Binomialverteilung verwendet. Nehmen Sie an, dass $n = 12$ Patienten das neue Medikament erhalten haben. Anschliessend wird beobachtet, wie viele Patienten gesund werden. Wir nehmen an, dass das Medikament bei allen Patienten mit der Erfolgswahrscheinlichkeit $p = 0.52$ wirkt und dass die Patienten voneinander unabhängig sind. Dann kann die Anzahl gesunder Patienten X mit einer Binomialverteilung approximiert werden: $X \sim \text{Bin}(n, p)$.

Wählen Sie alle Aussagen aus, die korrekt sind (Runden Sie auf drei Nachkommastellen):

Wählen Sie eine oder mehrere Antworten:

- ☐ a. Angenommen, das Ereignis A beschreibt die Situation, in der mindestens 7 Patienten gesund werden. Dann gilt $P(A^c) = P(X \leq 7)$, wobei $X \sim \text{Bin}(n=12, p=0.52)$.
- ☐ b. Der Erwartungswert für die Anzahl gesunder Patienten ist 4.519 .
- ☐ c. Der Varianz für die Anzahl gesunder Patienten ist 1.763 .
- ☐ d. Wir suchen die kleinste Zahl x , sodass $P(X \leq x) \geq 0.9$. Die gesuchte Zahl ist 8.
- ☐ e. Keine der oben genannten Aussagen ist richtig.

Im Folgenden kommen unterschiedliche Aussagen. Wählen Sie alle Aussagen aus, die korrekt sind (Runden Sie auf drei Nachkommastellen):

Wählen Sie eine oder mehrere Antworten:

- ☐ a. Für Zufallsvariablen verwenden wir die Notation: Grossbuchstaben für die Funktion (z.B. X) und Kleinbuchstaben für einen konkreten Wert, den die Zufallsvariable annehmen kann (z.B. x). Das Ereignis ' Die Zufallsvariable nimmt den Wert 3 an ' wird dann beschrieben mit ' $X = 3$ '.
- ☐ b. X ist eine beliebige diskrete Zufallsvariable, die n verschiedene Werte $x_1, \dots, x_n$ annehmen kann. Die Summe $\sum_{i=1} P(X = x_i)$ kann man ohne weitere Angaben nicht berechnen.
- ☐ c. Wir sind bei einem Abendessen mit Freunden insgesamt 13 Personen. Jeder stösst mit jedem einmal an. Man hört 67 mal die Gläser klingen.
- ☐ d. Wir betrachten eine Gruppe aus 2 Männern und 2 Frauen. Es wird zufällig ein Team aus 3 Personen gebildet. Sei X die Zufallsvariable, die die Anzahl Frauen in dem Team angibt. Folgendes Modell beschreibt die Situation nicht perfekt aber in sehr guter Näherung : $X \sim \text{Bin}(n = 3, \pi = 0.5)$
- ☐ e. Keine der oben genannten Aussagen ist richtig.

Bonusquiz I – Woche 4

Bei einem Multiple-Choice Quiz gibt es $n = 59$ TRUE/FALSE-Fragen. Student Meier hat sich nicht vorbereitet und ist jetzt so nervös, dass er zwar alle Fragen bearbeitet aber jedes mal (ohne die Aufgabe auch nur zu lesen) mit Wahrscheinlichkeit $p = 0.5$ den Wert TRUE ankreuzt (unabhängig von den anderen Aufgaben). Die Anzahl Aufgaben, die Student Meier auf diese Weise richtig beantwortet (X) kann mit einer Binomialverteilung approximiert werden: $X \sim \text{Bin}(n, p)$.

Wählen Sie alle Aussagen aus, die korrekt sind (Runden Sie auf drei Nachkommastellen):

Wählen Sie eine oder mehrere Antworten:

- ☐ a. Die Wahrscheinlichkeit, dass genau 7 Fragen richtig angekreuzt wurden ist 0.513 .
- ☐ b. Der Erwartungswert fuer die Anzahl richtig angekreuzter Fragen ist 29.5 .
- ☐ c. Die Varianz fuer die Anzahl richtig angekreuzter Fragen ist 10.693 .
- ☐ d. Der Lehrer will natuerlich, dass jemand nicht besteht, der nur geraten hat. Als Mindestpunktzahl fuer die Note 4 sucht er die kleinste Zahl x , sodass $P(X \leq x) \geq 0.9$. Die gesuchte Zahl ist 34 . (Hier verwenden Sie am besten die Funktion pbinom() in R.)

Der TCS interessiert sich fuer die Anzahl Autopannen pro Tag im Kanton Zuerich (Zufallsvariable X). Erfahrungsgemaess haben pro Tag im Mittel $m=24.1$ Autos im Kanton Zuerich eine Panne. Als Approximation nehmen wir an, dass $X \sim \text{Pois}(m)$.

Wählen Sie alle Aussagen aus, die korrekt sind (Runden Sie auf drei Nachkommastellen):

Wählen Sie eine oder mehrere Antworten:

- ☐ a. Die Wahrscheinlichkeit, dass genau 28 Autos an einem Tag eine Panne haben ist 0.145 .
- ☐ b. Angenommen, das Ereignis A beschreibt die Situation, in der weniger als 24 Autos pro Tag eine Panne haben. Dann gilt $P(A^c) = P(X > 24)$, wobei $X \sim \text{Pois}(m = 24.1)$.
- ☐ c. Der Erwartungswert fuer die Anzahl Autopannen pro Tag ist 24.1 .
- ☐ d. Die Standardabweichung fuer die Anzahl Autopannen pro Tag ist 7.127 .

- ☐ a. Bei einer Telefonzentrale treffen pro Stunde im Mittel 20 Anrufe ein. Die Verteilung der Anzahl Anrufe pro Stunde lässt sich gut mit einer Binomialverteilung beschreiben.
- ☐ b. Angenommen $X \sim \text{Pois}(10)$ und $Y \sim \text{Pois}(30)$. Zudem sind X und Y unabhängig. Dann gilt $E(\frac{1}{2}(X + Y)) = 20$.
- ☐ c. Die Binomialverteilung (BV) und die Hypergeometrische Verteilung (HV) beschreiben eine sehr ähnliche Situation: Die HV beschreibt eine 'kleine' Urne, d.h., durch das Ziehen einer Kugel ändert sich die Wahrscheinlichkeit für die nächste markierte Kugel merklich. Die BV beschreibt eine 'unendlich grosse' Urne, d.h., die Wahrscheinlichkeit, eine markierte Kugel zu ziehen (Gewinnwahrscheinlichkeit) ist bei jedem Zug gleich. In dieser Aufgabe wollen wir diesen Unterschied etwas genauer untersuchen. In einer Urne sind 30 Bälle. Davon sind 25 Bälle mit einem roten Kreuz markiert ('Gewinne'). Nun werden 7 Bälle zufällig gezogen. Wir interessieren uns für die Wahrscheinlichkeit genau 7 markierte Bälle zu ziehen. Berechnen Sie diese Wahrscheinlichkeit zunächst mit dem Hypergeometrischen Modell (p_H). Anschließend berechnen Sie die Wahrscheinlichkeit mit dem Binomial Modell (p_B ; als Gewinnwahrscheinlichkeit nehmen Sie Anzahl markierter Bälle durch Gesamtanzahl Bälle in der Urne). Die Differenz $p_B - p_H$ ist -0.248 .
- ☐ d. Die Poisson-Verteilung ist eine sehr gute Näherung für die Binomialverteilung, falls n gross und p klein ist (dann sind Gewinne 'seltene' Ereignisse). In dieser Aufgabe wollen wir diesen Zusammenhang untersuchen. Wir nehmen nun an, dass Gewinne selten sind. Konkret: $X \sim \text{Bin}(26, 0.048)$. Berechnen Sie $P(X = 1) = p_B$. Nun approximieren wir die Binomialverteilung mit der Poissonverteilung (Parameter $\lambda = E[X] = n \cdot p \approx 1.248$), also $Y \sim \text{Pois}(\lambda = 1.248)$. Damit berechnen Sie $P(Y = 1) = p_P$. Beurteilen Sie: Die Differenz $p_B - p_P$ ist -0.556 .

Bonusquiz I – Woche 5

Im Folgenden kommen unterschiedliche Aussagen. Wählen Sie alle Aussagen aus, die korrekt sind (Runden Sie auf drei Nachkommastellen):

Wählen Sie eine oder mehrere Antworten:

- ☐ a. Bei einem Binomialtest kann die Nullhypothese auf dem 1 % Signifikanzniveau verworfen werden. Dann kann die Nullhypothese auf dem 10 % Signifikanzniveau eventuell (aber nicht sicher) auch verworfen werden.
- ☐ b. Angenommen wir verwenden einen einseitigen Binomialtest um die Wirksamkeit eines neuen Medikaments zu prüfen. Wir haben die Nullhypothese ($H_0 : \pi = 0.33$) und die Alternative ($H_A : \pi > 0.33$) formuliert. Insbesondere interessieren wir uns für die Macht und den Fehler 2. Art für eine konkrete Alternativhypothese ($\pi = 0.45$). Um den Test unseren Bedürfnissen genau anzupassen, überlegen wir, welche Folgen es hat, wenn man das Signifikanzniveau ändert. Wenn der Fehler 1. Art zunimmt, dann nimmt die Macht (bei konstanter Stichprobengröße) zu.
- ☐ c. Wir wollen mit einem einseitigen Binomialtest prüfen, ob eine Münze zu häufig Zahl zeigt. Dazu werfen wir die Münze 120 mal und zählen, wie oft Zahl oben gelandet ist. Die Zufallsvariable X beschreibt die Anzahl Würfe, bei denen Zahl oben gelandet ist (bei total 120 Würfungen) und folgt der Verteilung $X \sim \text{Bin}(n = 120, \pi)$. Die Nullhypothese ist $H_0 : \pi = 0.5$. Die einseitige Alternative ist $H_A : \pi > 0.5$. Das Signifikanzniveau ist 5%. Um die Grenze des Verwerfungsbereichs zu finden muss man die kleinste Zahl k finden, für die gilt $P(X = k) \leq 0.05$.
- ☐ d. Angenommen der Verwerfungsbereich von unserem einseitigen Binomialtest ($\alpha = 0.05$) geht von 20 bis 36. Der beobachtete Wert der Teststatistik (d.h. die beobachtete Anzahl 'Gewinne') ist 5. Dann kann die Nullhypothese (auf dem genannten Signifikanzniveau) verworfen werden.

Im Folgenden kommen unterschiedliche Aussagen. Wählen Sie alle Aussagen aus, die korrekt sind (Runden Sie auf drei Nachkommastellen):

Wählen Sie eine oder mehrere Antworten:

- ☐ a. Von einem neuen Medikament wird erwartet, dass es deutlich besser als das herkömmliche Medikament wirkt. Um diese Vermutung zu unterstützen wird ein einseitiger Binomialtest gemacht. 28 kranke Patienten erhalten das neue Medikament. Die Zufallsvariable X beschreibt die Anzahl gesunder Patienten am Ende der Studie unter allen Patienten, die das neue Medikament erhalten haben. Wir nehmen an, dass $X \sim \text{Bin}(n = 28, \pi)$. Die Nullhypothese ist $H_0 : \pi = 0.2$. Die Alternativhypothese ist $H_A : \pi > 0.2$. Das Signifikanzniveau ist 10 %. Der Verwerfungsbereich des Tests geht dann von 8 bis 28 (beide Grenzen eingeschlossen).
- ☐ b. Von einem neuen Medikament wird erwartet, dass es deutlich besser als das herkömmliche Medikament wirkt. Um diese Vermutung zu unterstützen wird ein einseitiger Binomialtest gemacht. 38 kranke Patienten erhalten das neue Medikament. Die Zufallsvariable X beschreibt die Anzahl gesunder Patienten am Ende der Studie unter allen Patienten, die das neue Medikament erhalten haben. Wir nehmen an, dass $X \sim \text{Bin}(n = 38, \pi)$. Die Nullhypothese ist $H_0 : \pi = 0.4$. Die Alternativhypothese ist $H_A : \pi > 0.4$. Das Signifikanzniveau ist 5 %. Der Verwerfungsbereich des Tests geht dann von 21 bis 38 (beide Grenzen eingeschlossen).
- ☐ c. In den Medien wird behauptet, dass ein neuer Lebensmittelfarbstoff als Nebenwirkung leichte Kopfschmerzen verursachen kann. Der Hersteller behauptet daraufhin, dass diese Nebenwirkungen mit einer Wahrscheinlichkeit von höchstens 35 % auftreten. Eine Vereinigung zum Konsumentenschutz vermutet, dass die Wahrscheinlichkeit für diese Nebenwirkung grösser sein könnte und macht einen einseitigen Binomialtest. 21 Probanden bekommen ein Getränk, das mit dem neuen Lebensmittelfarbstoff gefärbt wurde. Nach einer Stunde werden sie gefragt, ob sie Kopfschmerzen haben. Die Zufallsvariable X beschreibt die Anzahl Probanden mit Kopfschmerzen nach einer Stunde. Wir nehmen an, dass $X \sim \text{Bin}(n = 21, \pi)$. Die Nullhypothese ist $H_0 : \pi = 0.35$. Die Alternativhypothese ist $H_A : \pi > 0.35$. Das Signifikanzniveau ist 7 %. Der Verwerfungsbereich des Tests geht dann von 11 bis 21 (beide Grenzen eingeschlossen).
- ☐ d. Eine Marketingfirma will ein neues provokatives Plakat veröffentlichen. Allerdings gibt es innerhalb der Firma Bedenken, dass ein grosser Anteil der Bevölkerung das Plakat als 'zu provokativ' einstufen könnte. Akzeptabel wäre maximal ein Anteil von 50 % in der Gesamtbevölkerung, die das Plakat als 'zu provokativ' einstufen. In einer kleinen Pilotstudie soll nun mit einem einseitigen Binomialtest geprüft werden, ob der Anteil der provozierten Personen in der Bevölkerung kleiner als 50 % ist. Dazu werden zufällig 11 Personen ausgewählt und über das geplante Plakat befragt. Die Zufallsvariable X beschreibt die Anzahl befragter Personen, die das Plakat als 'zu provokativ' einstufen. Wir nehmen an, dass $X \sim \text{Bin}(n = 11, \pi)$. Die Nullhypothese ist $H_0 : \pi = 0.5$. Die Alternativhypothese ist $H_A : \pi < 0.5$. Das Signifikanzniveau ist 5 %. Der Verwerfungsbereich des Tests geht dann von 0 bis 3 (beide Grenzen eingeschlossen).

Im Folgenden kommen unterschiedliche Aussagen. Wählen Sie alle Aussagen aus, die korrekt sind (Runden Sie auf drei Nachkommastellen):

Wählen Sie eine oder mehrere Antworten:

- ☐ a. In einem einseitigen Binomialtest ($H_A : p > p_0$) haben wir den Verwerfungsbereich bestimmt: $\{19, \dots, 29\}$. Angenommen die Gewinnwahrscheinlichkeit ist in Wahrheit 0.1. Die Wahrscheinlichkeit unter diesen Umständen die Nullhypothese zu verwerfen ist 0.
- ☐ b. In einem einseitigen Binomialtest ($H_A : p < p_0$) haben wir den Verwerfungsbereich bestimmt: $\{0, \dots, 19\}$ bei einer Stichprobengröße von 25. Angenommen die Gewinnwahrscheinlichkeit ist in Wahrheit 0.45. Die Wahrscheinlichkeit unter diesen Umständen die Nullhypothese zu verwerfen ist 0.45.
- ☐ c. Eine Phase-2 klinische Studie wurde mit einem einseitigen Binomialtest ($H_A : p > p_0$) sorgfältig geplant. Folgende Parameter wurden fixiert: Die Wahrscheinlichkeit unter der Nullhypothese p_0 und das Signifikanzniveau α . Zudem wurde eine konkrete Alternative formuliert (p_A), die mit einer vorgegebenen Macht $1 - \beta$ detektiert werden soll. Dazu wurde die nötige Stichprobengröße n bestimmt. Kurz bevor die Studie beginnt wird allerdings bei der Uebermittlung der Parameter ein Fehler entdeckt: Das Signifikanzniveau ist grösser als ursprünglich geplant. Beurteilen Sie: Es ist sichergestellt, dass die geforderte Macht auch mit den neuen Parametern eingehalten wird.
- ☐ d. Der Verwerfungsbereich in einem einseitigen Binomialtest ($H_A : p < p_0$) ist: $\{0, \dots, 21\}$. Die Stichprobengröße ist 39. Angenommen die Erfolgswahrscheinlichkeit ist tatsächlich 0.5. Die Wahrscheinlichkeit, dass die Nullhypothese dann nicht verworfen wird ist 0.261.

Bonusquiz I – Woche 6

Im Folgenden kommen unterschiedliche Aussagen. Wählen Sie alle Aussagen aus, die korrekt sind (Runden Sie auf drei Nachkommastellen):

Wählen Sie eine oder mehrere Antworten:

- ☐ a. Angenommen der p-Wert bei einem zweiseitigen Binomialtest ist 0.086 . Wenn man den Hypothesentest mit der gleichen Nullhypothese und mit den gleichen Daten nochmals auswerten wuerde, koennte man die Nullhypothese auf dem Signifikanzniveau 0.1 verwerfen.
- ☐ b. Angenommen die Gewinnwahrscheinlichkeit unter der Nullhypothese p_0 befindet sich im (zweiseitigen) 97 %-Vertrauensintervall fuer die Gewinnwahrscheinlichkeit. Es ist dann plausibel, dass der p-Wert fuer einen zweiseitigen Binomialtest mit der Nullhypothese $H_0 : p = p_0$ den Wert 0.065 hat.
- ☐ c. Angenommen das zweiseitige 95%-Vertrauensintervall fuer die Erfolgswahrscheinlichkeit eines Binomialtests geht von 0.5 bis 0.839 . Wenn wir mit den gleichen Daten einen zweiseitigen Binomialtest mit der Nullhypothese $H_0 : p = 0.35$ auf dem 5% Signifikanzniveau durchfuehren wuerden, koennten wir die Nullhypothese verwerfen.
- ☐ d. Die Wahrscheinlichkeit, dass ein 95 %-Vertrauensintervall den wahren Parameter enthaelt ist 99 %.

Im Folgenden kommen unterschiedliche Aussagen. Wählen Sie alle Aussagen aus, die korrekt sind (Runden Sie auf drei Nachkommastellen):

Wählen Sie eine oder mehrere Antworten:

- ☐ a. Eine gewisse genetische Mutation tritt bei einer Insektenart mit einer Wahrscheinlichkeit von 0.35 auf. Nun wird vermutet, dass sich diese Wahrscheinlichkeit in der Nachfolge-Generation ändern könnte, wenn die Insekten einer radioaktiven Strahlung ausgesetzt werden. Es ist allerdings nicht klar, ob die Mutationsrate durch die Strahlung grösser oder kleiner wird. Um diese Frage zu beantworten, wird ein zweiseitiger Binomialtest ($H_0 : p = p_0$ wobei $p_0 = 0.35$) auf dem 7 % Signifikanzniveau mit 26 Insekten durchgeführt. Der Verwerfungsbereich geht von 4 bis 15 .
- ☐ b. Aus der Vergangenheit weiss man, dass die Wahlbeteiligung bei Initiativen zum Thema 'Soziales' in etwa 0.5 ist. Nun kommt es wieder zu einer Initiative im gleichen Themenkreis. Allerdings wurde die Initiative dieses mal fast nur im Internet und nicht auf herkömmlichen Plakaten beworben. Es soll nun geprüft werden, ob sich die Wahlbeteiligung nennenswert geändert hat. Zu diesem Zweck, wird ein zweiseitiger Binomialtest ($H_0 : p = p_0$ wobei $p_0 = 0.5$) auf dem 10 % Signifikanzniveau mit 36 zufällig ausgewählten Personen (über Wahlteilnahme befragt) durchgeführt. Der Verwerfungsbereich geht von 12 bis 24 .
- ☐ c. Eine grosse Zeitschrift hat einen Lektorendienst engagiert, der alle Artikel vor Veröffentlichung auf grammatikalische Fehler prüfen soll. Der Dienst garantiert, dass die korrigierten Texte keine grösseren Fehler mehr enthalten. Allerdings könnte es sein, dass es noch stilistische Mängel gibt. Zudem wird garantiert, dass die Wahrscheinlichkeit für stilistische Mängel nach der Korrektur pro Dokument nicht mehr als 50 % ist. Die Leitung der Zeitschrift vermuten aber, dass es mehr sein könnte. Also machen sie einen einseitigen Binomialtest. Dazu lassen sie 16 Artikel mit bekannten stilistischen Mängeln korrigieren. Die Zufallsvariable X beschreibt die Anzahl Artikel mit stilistischen Mängeln nach der Korrektur. Wir nehmen an, dass $X \sim \text{Bin}(n = 16, \pi)$. Die Nullhypothese ist $H_0 : \pi = 0.5$. Die Alternativhypothese ist $H_A : \pi > 0.5$. Das Signifikanzniveau ist 10 %. Der Verwerfungsbereich des Tests geht dann von 12 bis 16 (beide Grenzen eingeschlossen).
- ☐ d. Ein neuer Wirkstoff sieht vielversprechend aus und der Finanzchef der Pharma-Firma will prüfen, ob es sich lohnt, das Medikament weiter zu untersuchen. Der Standard-Wirkstoff wirkt mit einer Wahrscheinlichkeit von ca. 15 %. Für uns wäre der neue Wirkstoff nur dann interessant, wenn er eine grössere Wirkungswahrscheinlichkeit als der Standardwirkstoff hat. Um das zu untersuchen machen wir einen einseitigen Binomialtest ($H_A : p > p_0$) mit $n=15$ Probanden und beobachten $x=15$ Probanden, bei denen der Wirkstoff die gewünschte Wirkung zeigt. Dann ist der p-Wert 0.766 .

Im Folgenden kommen unterschiedliche Aussagen. Wählen Sie alle Aussagen aus, die korrekt sind (Runden Sie auf drei Nachkommastellen):

Wählen Sie eine oder mehrere Antworten:

- ☐ a. Angenommen, der Verwerfungsbereich bei einem zweiseitigen Binomialtest besteht aus $\{0, \dots, 16\}$ und $\{28, \dots, 37\}$. Die Macht fuer die konkrete Alternative $p_A = 0.3$ ist dann 0.181 .
- ☐ b. Angenommen $X \sim \text{Bin}(n = 39, p = 0.7)$. Dann gilt $P(X \leq 2) + P(X \geq 2) > 1$.
- ☐ c. Angenommen, der Verwerfungsbereich bei einem zweiseitigen Binomialtest auf dem Signifikanzniveau 10 % besteht aus $\{0, \dots, 7\}$ und $\{16, \dots, 18\}$. Falls man das Signifikanzniveau groesser macht wird der Verwerfungsbereich gleich viele oder mehr (aber nicht weniger) Elemente enthalten.
- ☐ d. Angenommen die beobachtete Teststatistik ist nicht im Verwerfungsbereich (mit Signifikanzniveau α). Dann haben wir (mit Fehlerwa. hoechstens α) nachgewiesen, dass die Nullhypothese nicht stimmt .

Bonusquiz I – Woche 7

Im Folgenden kommen unterschiedliche Aufgaben zu Kennzahlen. Verwenden Sie dabei immer die Formeln aus der Vorlesung (für manche Kennzahlen, z.B. das Quantil, gibt es verschiedene Varianten). Wählen Sie alle Aussagen aus, die korrekt sind (Runden Sie auf drei Nachkommastellen):

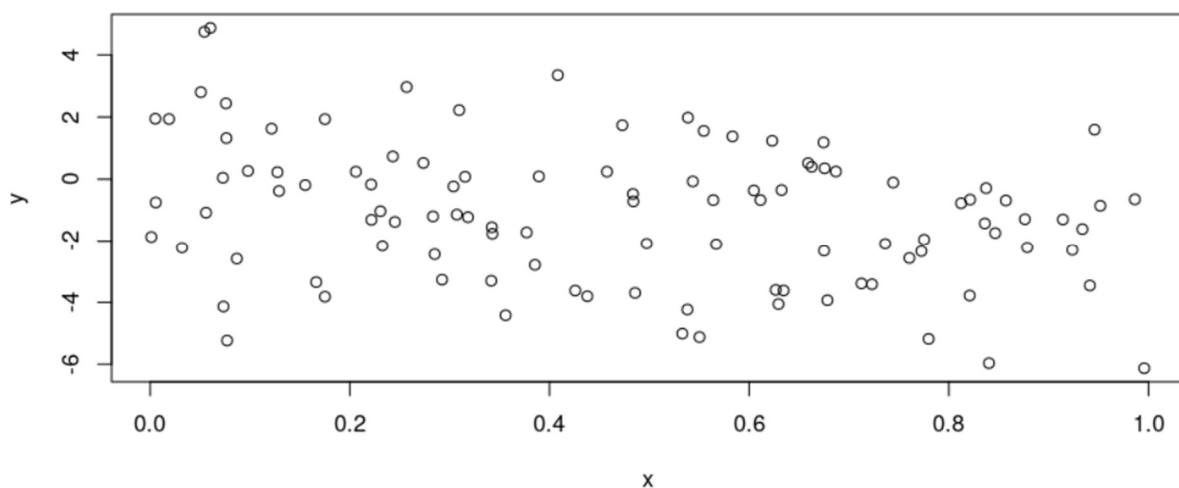
Wählen Sie eine oder mehrere Antworten:

- ☐ a. Berechnen Sie den Mittelwert folgender 6 Zahlen: 0.832, 0.002, 0.45, 0.638, 0.967, 0.065. Verwenden Sie dazu die Formeln aus der Vorlesung. Der Mittelwert ist 0.492 .
- ☐ b. Berechnen Sie den Median folgender 8 Zahlen: 0.64, 0.303, 0.521, 0.46, 0.704, 0.659, 0.863, 0.92. Verwenden Sie dazu die Formeln aus der Vorlesung. Der Median ist 0.6495 .
- ☐ c. Berechnen Sie die empirische Standardabweichung folgender 10 Zahlen: 0.191, 0.016, 0.55, 0.13, 0.777, 0.12, 0.739, 0.59, 0.685, 0.651. Verwenden Sie dazu die Formeln aus der Vorlesung. Die empirische Standardabweichung ist 0.277 .
- ☐ d. Berechnen Sie das 10 %-Quantil folgender 6 Zahlen: 0.008, 0.837, 0.926, 0.194, 0.722, 0.284. Verwenden Sie dazu die Formeln aus der Vorlesung. Das 10 %-Quantil ist 0.008 .

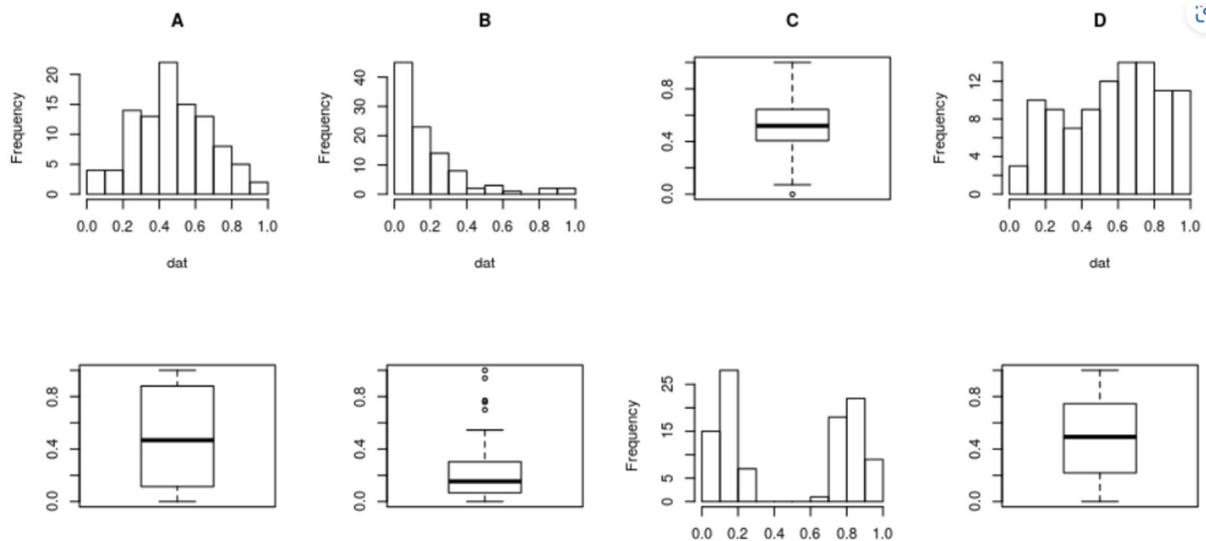
Im Folgenden kommen unterschiedliche Aufgaben zu Kennzahlen. Verwenden Sie dabei immer die Formeln aus der Vorlesung (für manche Kennzahlen, z.B. das Quantil, gibt es verschiedene Varianten). Wählen Sie alle Aussagen aus, die korrekt sind (Runden Sie auf drei Nachkommastellen):

Wählen Sie eine oder mehrere Antworten:

- ☐ a. Der Median entspricht immer einem Wert, der in der Stichprobe auch vorkommt.
- ☐ b. Angenommen, wir haben folgende Werte aufgenommen: 0.741, 0.384, 0.263, 0.026, 0.892, 0.237, 0.716, 0.656. Beim grössten Wert in der Stichprobe $x_{max} = 0.892$ ist jedoch ein Tippfehler aufgetreten und es wird fälschlicherweise $x_{max} = 0.892 = 8.92$ eingetragen. Durch diesen Tippfehler wird sich der Mittelwert der Stichprobe ändern.
- ☐ c. Angenommen, wir haben zwei kontinuierliche Zufallsvariablen X und Y . Falls es keinen Zusammenhang zwischen X und Y gibt, dann ist die Korrelation zwischen X und Y mit Sicherheit nicht null.
- ☐ d. Betrachten Sie das Streudiagramm. Der Absolutbetrag der empirischen Korrelation zwischen x und y gemäss diesem Streudiagramm ist grösser als 0.7.



Im folgenden Bild sehen Sie acht Bilder in zwei Reihen und vier Spalten (A, B, C und D). Die Frage ist immer, ob es plausibel ist, dass die beiden Bilder (Histogramm oder Boxplot) in einer Spalte von den gleichen Daten erzeugt wurden.



Wählen Sie eine oder mehrere Antworten:

- ☐ a. Es ist plausibel, dass die beiden Grafiken in Spalte A mit den gleichen Daten erzeugt wurden.
- ☐ b. Es ist plausibel, dass die beiden Grafiken in Spalte B mit den gleichen Daten erzeugt wurden.
- ☐ c. Es ist plausibel, dass die beiden Grafiken in Spalte C mit den gleichen Daten erzeugt wurden.
- ☐ d. Es ist plausibel, dass die beiden Grafiken in Spalte D mit den gleichen Daten erzeugt wurden.

Bonusquiz I – Woche 8

Im Folgenden kommen verschiedene Fragen zu kontinuierlichen Verteilungsfunktionen. Wählen Sie alle Aussagen aus, die korrekt sind (Runden Sie auf drei Nachkommastellen):

Wählen Sie eine oder mehrere Antworten:

- ☐ a. Die Wartezeit bei einem Produktionsprozess ist uniformverteilt zwischen $a = 3.7$ Minuten und $b = 9.7$ Minuten. Die Wahrscheinlichkeit, dass man höchstens 9.165 Minuten warten muss ist 0.732 .
- ☐ b. Angenommen an einem SBB-Schalter lässt sich die Bearbeitungszeit pro Kunde in etwa mit einer Exponentialverteilung mit Parameter $\lambda = 1.9$ (Einheit ' 1 Minute ') modellieren. Die Varianz fuer die Wartezeit ist 0.14.
- ☐ c. Die Zufallsvariable X folgt einer Normalverteilung: $X \sim N(51, 3.6)$. Dann gilt $E(X)=3.6$.
- ☐ d. Die Zufallsvariable X folgt einer Normalverteilung: $X \sim N(3, 3.6)$. Dann gilt $\sigma_X = 1.897$.

Im Folgenden kommen verschiedene Fragen zur Normalverteilung und zum Standardisieren. Wählen Sie alle Aussagen aus, die korrekt sind (Runden Sie auf drei Nachkommastellen):

Wählen Sie eine oder mehrere Antworten:

- ☐ a. Die Zufallsvariable Z folgt einer Standardnormalverteilung. Dann gilt $P(Z \leq 1.29) = 0.902$. (Verwenden Sie die Tabelle am Ende vom Skript S.112.)
- ☐ b. Die Zufallsvariable Z folgt einer Standardnormalverteilung. Angenommen es gilt $P(Z \leq c) = 0.9984$. Dann gilt $c \approx 3.54$. (Verwenden Sie die Tabelle am Ende vom Skript S.112.)
- ☐ c. Angenommen es gilt $X \sim N(m = 82, s^2 = 0.04)$. Dann gilt $P(X \leq 81.508) = 0.0069$.
- ☐ d. Angenommen es gilt $X \sim N(m = -23, s^2 = 1)$. Nehmen wir weiter an, es gilt $P(X \leq c) = 0.9953$. Dann gilt $c \approx -20.4$.

Im Folgenden kommen verschiedene Fragen zu Funktionen von Zufallsvariablen. Wählen Sie alle Aussagen aus, die korrekt sind (Runden Sie auf drei Nachkommastellen):

Wählen Sie eine oder mehrere Antworten:

- ☐ a. Die Zufallsvariable X hat Erwartungswert $E(X) = -0.5$ und Varianz $\text{Var}(X) = 2.7$. Nun wird die Zufallsvariable $Y = 1.5 - 4.9 \cdot X$ berechnet. Dann ist $E(Y) = 3.95$.
- ☐ b. Die Zufallsvariable X hat Erwartungswert $E(X) = 9.1$ und Varianz $\text{Var}(X) = 0.6$. Nun wird die Zufallsvariable $Y = 3.4 - 0 \cdot X$ berechnet. Dann ist $\text{Var}(Y) = 0$.
- ☐ c. Die Zufallsvariable X hat Erwartungswert $E(X) = -2.9$ und Varianz $\text{Var}(X) = 0.7$. Nun wird die Zufallsvariable $Y = 0.1 - 2.6 \cdot X$ berechnet. Dann ist $\sigma_Y = 2.175$.
- ☐ d. Die Zufallsvariable X hat 45 %-Quantil $q_X = -3.3$. Nun wird die Zufallsvariable $Y = 2.1 + 4.3 \cdot X$ berechnet. Dann ist das 45 %-Quantil von Y gleich $q_Y = -6.93$.

Im Folgenden kommen verschiedene Fragen zu Funktionen von Zufallsvariablen. Wählen Sie alle Aussagen aus, die korrekt sind (Runden Sie auf drei Nachkommastellen):

Wählen Sie eine oder mehrere Antworten:

- ☐ a. Die Exponentialverteilung ist gut geeignet um die Lebensdauer von radioaktiven Atomen zu modellieren.
- ☐ b. Die Zufallsvariable X_1 hat Erwartungswert $E(X_1) = -6.5$ und Varianz $Var(X_1) = 2$. Die Zufallsvariable X_2 hat Erwartungswert $E(X_2) = 2.5$ und Varianz $Var(X_2) = 3$. X_1 und X_2 sind unabhängig. Nun wird die Zufallsvariable $Y = 1.7 + 3.9 \cdot X_1 - 3.8 \cdot X_2$ berechnet. Dann ist $Var(Y) = 73.74$.
- ☐ c. Ein Passagierflugzeug hat 325 Sitze. Sicherheitsvorschriften geben vor, dass die Maschine nur dann starten darf, falls das Gesamtgewicht aller Passagiere höchstens 22130 kg ist. Angenommen, die Maschine ist voll besetzt und das Gewicht X_i , ($i = 1, \dots, 325$) der Passagiere ist unabhängig voneinander mit Erwartungswert $E(X_i) = 67$ kg und Standardabweichung $\sigma_{X_i} = 11$ kg. Die Wahrscheinlichkeit, dass das Flugzeug wegen zu grossem Passagiergewicht nicht starten darf ist etwa 0.266 . (Verwenden Sie den zentralen Grenzwertsatz)
- ☐ d. Spiele im Kasino bringen im Durchschnitt dem Kasino einen Gewinn. Betrachten wir einen Spielautomaten in einem Kasino einen Abend lang. Angenommen es spielen 117 Personen an diesem Abend an diesem Automaten. Wir nehmen an, der Gewinn aller Spieler ist unabhängig und gleichverteilt. Pro Spieler bringt dieser Automat dem Kasino einen erwarteten Gewinn $E(X_i) = 0.49$ SFr mit einer Standardabweichung $\sigma_{X_i} = 0.37$ SFr. Die Wahrscheinlichkeit, dass dieser Spielautomat an diesem Abend einen Totalgewinn von weniger als 50 SFr fuer das Kasino abwirft, ist 0.034 .(Verwenden Sie den zentralen Grenzwertsatz)

Bonusquiz I – Woche 9

Im Folgenden kommen verschiedene Fragen zum Wurzel-n Gesetz und zum t-Test. Wählen Sie alle Aussagen aus, die korrekt sind (Runden Sie auf drei Nachkommastellen):

Wählen Sie eine oder mehrere Antworten:

- ☐ a. Angenommen, wir haben das Körpergewicht von 16 zufällig ausgewählten Personen in der Vorlesung bestimmt. Der Hersteller der Waage gibt an, dass jede Einzelmessung eine Standardabweichung von 0.2 kg hat. Die Standardabweichung des arithmetischen Mittels der 16 Personen ist dann 3.2 .
- ☐ b. Angenommen, eine Einzelmessung X_i folgt der Verteilung $N(\mu, \sigma_X^2)$. Das arithmetische Mittel $\bar{X}_n = \frac{1}{n} \sum i$ aus n unabhängigen Einzelmessungen folgt der Verteilung $N(\mu, \left(\frac{\sigma_X}{n}\right)^2)$. Zwischen der Standardabweichung einer Einzelmessung σ_X und der Standardabweichung des arithmetischen Mittels $\sigma_{\bar{X}_n}$ gibt es folgenden Zusammenhang: $\sigma_{\bar{X}_n} = \frac{\sigma_X}{\sqrt{n}}$.
- ☐ c. Angenommen, das arithmetische Mittel ist folgendermassen verteilt: $N(\mu, \left(\frac{\sigma_X}{n}\right)^2)$. Die Verteilung von $\frac{\bar{X}-\mu}{\sigma_{\bar{X}}}$ ist dann $N(\mu, 1)$.
- ☐ d. Bei einem (zweiseitigen) t-Test mit $n = 10$ Beobachtungen ist der Wert der Teststatistik 0.883 . Der p-Wert ist dann etwa 0.35 .

Im Folgenden kommen verschiedene Fragen zum zweiseitigen Ein-Stichproben t-Test. Wählen Sie alle Aussagen aus, die korrekt sind (Runden Sie auf drei Nachkommastellen):

Wählen Sie eine oder mehrere Antworten:

- ☐ a. Für 9 Patienten wurde der Blutdruck vor und nach Verabreichung eines neuen blutdrucksenkenden Medikaments gemessen. Die Differenzen der Blutdruckwerte (vorher minus nachher) sind: 1.25, -17.53, 15.44, 6.72, -0.86, -0.14, 5.33, 7.55, 6.1. Wir wollen nun mit einem t-Test prüfen, ob die Differenz der Blutdruckwerte signifikant von null verschieden sein könnte (zweiseitiger t-Test). Beurteilen Sie: Der beobachtete Wert der Teststatistik ist 0.881.
- ☐ b. Angenommen das Signifikanzniveau in obigem Test ist 0.05. Der Verwerfungsbereich des zweiseitigen t-Tests ist dann $[-\infty; -2.306] \cup [2.306; \infty]$.
- ☐ c. Angenommen der Verwerfungsbereich bei einem zweiseitigen Ein-Stichproben t-Test auf dem 5%-Signifikanzniveau ist $[-\infty; -2.14] \cup [2.14; \infty]$ und der beobachtete Wert der Teststatistik ist $t = -0.769$. Die Nullhypothese kann dann auf dem 5%-Signifikanzniveau verworfen werden.
- ☐ d. Wir konstruieren nun mit obigen Daten ein 99 %-Vertrauensintervall für den Erwartungswert der Blutdruckdifferenz. Das 99 %-Vertrauensintervall ist: $[-6.478; 10.692]$.

Im Folgenden kommen verschiedene Fragen zum zweiseitigen Ein-Stichproben z-Test (wir nehmen also an, dass die Standardabweichungen der Einzelbeobachtungen bekannt ist). Wählen Sie alle Aussagen aus, die korrekt sind (Runden Sie auf drei Nachkommastellen):

Wählen Sie eine oder mehrere Antworten:

- ☐ a. Für 9 Patienten wurde der Blutdruck vor und nach Verabreichung eines neuen blutdrucksenkenden Medikaments gemessen. Die Differenzen der Blutdruckwerte (vorher minus nachher) sind: 4.03, 0.92, 4.72, 4.73, 8.92, 1.65, 3.59, 27.6, -19.66. Angenommen, wir kennen von einer früheren sehr grossen Studie die wahre Standardabweichung der Blutdruckdifferenzen: $\sigma_X = 12$. Wir wollen nun mit einem z-Test prüfen, ob die Differenz der Blutdruckwerte signifikant von null verschieden sein könnte (zweiseitiger z-Test). Beurteilen Sie: Der beobachtete Wert der Teststatistik ist 1.014.
- ☐ b. Angenommen das Signifikanzniveau in obigem Test ist 0.09. Der Verwerfungsbereich des zweiseitigen z-Tests ist dann $[-\infty; -1.695] \cup [1.695; \infty]$.
- ☐ c. Angenommen der Verwerfungsbereich bei einem zweiseitigen Ein-Stichproben z-Test auf dem 5%-Signifikanzniveau ist $[-\infty; -2.498] \cup [2.498; \infty]$ und der beobachtete Wert der Teststatistik ist $z=2.551$. Die Nullhypothese kann dann auf dem 5%-Signifikanzniveau verworfen werden.
- ☐ d. Wir konstruieren nun mit obigen Daten ein 95 %-Vertrauensintervall für den Erwartungswert der Blutdruckdifferenz. Dazu nehmen wir weiterhin an, dass $\sigma_X = 12$ die wahre Standardabweichung der Daten ist. Das 95 %-Vertrauensintervall ist: $[-3.881; 10.736]$.

Bonusquiz I – Woche 10

Im Folgenden finden Sie mehrere Beispiele fuer Vergleiche von 2 Stichproben. Beurteilen Sie, ob es sich um gepaarte oder um ungepaarte Stichproben handelt.

Wählen Sie eine oder mehrere Antworten:

- ☐ a. Die folgenden Daten sind aus einer Studie von Charles Darwin ueber die Fremd- und Selbstbefruchtung. 15 Paare von Setzlingen mit demselben Alter, je einer durch Selbst- und einer durch Fremdbefruchtung produziert, wurden gezuechtet. Beide Teile je eines Paares hatten nahezu gleiche Bedingungen. Das Ziel bestand darin zu sehen, ob die fremdbefruchteten Pflanzen mehr Lebenskraft besitzen als die selbstbefruchteten (d.h., ob sie groesser werden). Es wurden die Hoehen jeder Pflanze nach einer fixen Zeitspanne gemessen. Es handelt sich um gepaarte Stichproben.
- ☐ b. In einem Experiment wurde untersucht, ob Maeuse zwei Formen von Eisen (Fe^{2+} und Fe^{3+}) unterschiedlich gut aufnehmen. Dazu wurden 36 Maeuse zufaellig in zwei Gruppen zu je 18 unterteilt und die eine Gruppe mit Fe^{2+} und die andere mit Fe^{3+} gefuettert. Da das Eisen radioaktiv markiert war, konnte sowohl die Anfangskonzentration wie auch die Konzentration einige Zeit spaeter gemessen werden. Daraus wurde fuer jede Maus der Anteil des aufgenommenen Eisens berechnet. Es handelt sich um ungepaarte Stichproben.
- ☐ c. 20 zufaellig ausgewaehlte Patienten mit Fieber werden zufaellig in zwei Gruppen aufgeteilt. Die eine Gruppe (12 Personen) wird mit einem neuen Medikament behandelt. Die andere Gruppe (8 Personen) wird mit dem herkoemmlichen Medikament behandelt. Es wird bei jeder Person der Rueckgang des Fieber nach einer Stunde gemessen. Wir wollen nun pruefen, ob das neue Medikament das Fieber innerhalb einer Stunde signifikant staerker senken kann und verwenden einen Zwei-Stichproben t-Test. Es handelt sich um ungepaarte Stichproben.
- ☐ d. Es wurden neue Augentropfen entwickelt, die den Augeninnendruck senken sollen. Um die Wirkung des neuen Medikaments zu pruefen, werden 10 Personen zufaellig ausgewaehlt. Jede Person bekommt in das linke Auge die neuen Augentropfen. In das rechte Auge bekommt sie herkoemmliche Augentropfen. Gemessen wird nun fuer jedes Auge und jedem Patienten die Senkung des Augeninnendrucks nach fuenf Minuten. Wir wollen nun mit einem Zwei-Stichproben t-Test pruefen, ob der Augeninnendruck mehr sinkt, wenn das neue Medikament verwendet wird. Es handelt sich um gepaarte Stichproben.

Im Folgenden kommen verschiedene Fragen zum ungepaarten, zweiseitigen Zwei-Stichproben t-Test. Wählen Sie alle Aussagen aus, die korrekt sind (Runden Sie auf drei Nachkommastellen):

Wählen Sie eine oder mehrere Antworten:

- ☐ a. Wir wollen prüfen, ob die normalverteilten Zufallsvariablen X und Y den gleichen Erwartungswert haben (wir nehmen zudem an, dass X und Y gleiche Varianz haben). Dazu haben wir $n_1 = 5$ Beobachtungen von X gemacht: 20.94, 8.09, -15.54, -4.68, 32.76. Zudem haben wir $n_2 = 7$ Beobachtungen von Y : 11.29, 11.3, 7.49, 5.16, 3.55, 3.62, -20.35. Beurteilen Sie: S_{pool}^2 wird geschätzt als 280.722.
- ☐ b. Der Wert der Teststatistik beim zweiseitigen Zwei-Stichproben t-Tests mit obigen Daten ist 0.594.
- ☐ c. Angenommen das Signifikanzniveau im zweiseitigen Zwei-Stichproben t-Test mit obigen Daten ist 0.01. Der Verwerfungsbereich des Tests ist dann $[-\infty; -3.827] \cup [3.827; \infty]$.
- ☐ d. Angenommen bei einem zweiseitigen Zwei-Stichproben t-Test mit $n_1 = 14$ und $n_2 = 11$ Beobachtungen ist der beobachtete Wert der Teststatistik $t = 1.319$. Der p-Wert ist dann 0.2.

Wählen Sie alle Aussagen aus, die korrekt sind (Runden Sie auf drei Nachkommastellen):

Wählen Sie eine oder mehrere Antworten:

- ☐ a. Ein Verwerfungsbereich ist ein Bereich, der unplausible Werte der Teststatistik enthaelt, wenn die Nullhypothese stimmt .
- ☐ b. Ein Vertrauensintervall ist ein Bereich, der unplausible Parameterwerte enthaelt .
- ☐ c. Ein Forscher versucht ein Gewicht moeglichst genau zu bestimmen. Dazu misst er das Gewicht mit seiner Waage 5 mal und berechnet ein 95% Vertrauensintervall fuer das Gewicht (die Waage hat einen zufaelligen Messfehler). Um ganz sicher zu gehen, bittet er noch seinen Kollegen, mit dem gleichen Gewicht und der gleichen Waage nochmals 5 Messungen zu machen und aus seinen 5 Messungen auch ein 95% Vertrauensintervall zu berechnen. Am Schluss vergleichen die beiden Forscher ihre beiden Vertrauensintervalle und stellen fest, dass es zwar einen grossen Ueberlapp gibt, aber dass sie nicht identisch sind. Beurteilen Sie: Es ist zu erwarten, dass die beiden Vertrauensintervalle nicht genau identisch sind, denn die beiden Forscher werden auf Grund der Zufallsschwankung der Waage nicht genau die gleichen Messungen gemacht haben .
- ☐ d. Wir machen einen Einstichproben t-Test mit der Nullhypothese $H_0 : \mu = 0$ und der Alternative $H_A : \mu \neq 0$. Die gemessenen Werte sind $x_1 = 0.3, x_2 = 0.2, x_3 = -0.1, x_4 = 0.6, x_5 = 0.8$. Der beobachtete Wert der Teststatistik ist $t = 2.30$. Daraus ergibt sich der p-Wert $p = 0.08$. Angenommen, wir haben die Messung x_2 falsch abgetippt und stellen fest, dass in Wirklichkeit gilt $x_2 = 0.7$. Wir rechnen erneut den beobachteten Wert der Teststatistik aus und erhalten nun $t = 2.82$. Beurteilen Sie: Der p-Wert bleibt gleich.

Bonusquiz I – Woche 11

Es wurde eine einfache lineare Regression mit 11 Datenpunkten geschätzt ($y = \beta_0 + \beta_1 x$). Der R-Output wurde in folgende Tabelle übertragen:

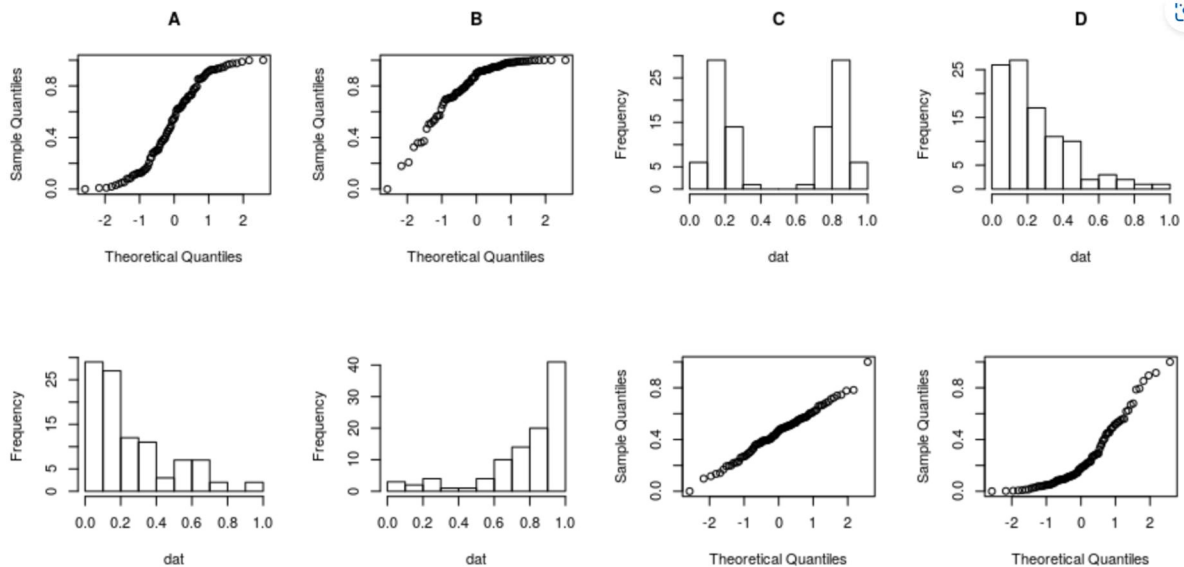
	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	1.752	1.023	1.712	0.121
x	1.829	0.534	3.426	?

Beurteilen Sie folgende Aussagen (Runden Sie auf drei Nachkommastellen):

Wählen Sie eine oder mehrere Antworten:

- ☐ a. Gemäss der Tabelle ist der p-Wert des geschätzten Achsenabschnitts gleich 1.023 .
- ☐ b. Die Obergrenze des 95%-Vertrauensintervalls (approximativ) für den Achsenabschnitt ist 3.688.
- ☐ c. An der Stelle des Fragezeichens sollte in der Tabelle der Wert 0.1 stehen.
- ☐ d. Die Nullhypothese $H_0 : \beta_0 = 0$ kann auf dem 5%-Signifikanzniveau verworfen werden.
- ☐ e. Die t-values in der Tabelle beziehen sich auf die t-Verteilung mit 28 Freiheitsgraden.
- ☐ f. Falls man x um eine Einheit erhöht, sagt unser Modell voraus, dass sich y um den Wert 1.829 erhöht.
- ☐ g. Angenommen Null ist im 95%-Vertrauensintervall für β_0 nicht enthalten. Dann kann die Nullhypothese $H_0 : \beta_0 = 0$ auf dem 5%-Signifikanzniveau verworfen werden.
- ☐ h. Angenommen der Schätzwert für β_1 wäre 4.132 und der Schätzwert von β_0 ist wie in der Tabelle angegeben (1.752). Für $x=1.596$ sagt dieses Modell dann ein erwartetes y von $y=6.518$ voraus.
- ☐ i. Angenommen der Schätzwert für β_1 wäre 4.307 und der Schätzwert von β_0 ist wie in der Tabelle angegeben (1.752). Wenn unser Modell dann ein erwartetes $y=2.492$ voraussagt, dann wurde als Input $x=0.607$ verwendet.
- ☐ j. Angenommen in der Zeile x in obiger Tabelle stehen nur Werte für Estimate (Wert= -0.984) und Std.Error (Wert= 0.355). Der p-Wert für den zweiseitigen Test der Nullhypothese $H_0 : \beta_1 = 0$ ist dann etwa 0.022. (Der angegebene p-Wert ist entweder korrekt oder mindestens fuenfmal so gross wie der wahre p-Wert oder hoechstens nur ein fuenftel des wahren p-Wertes.)

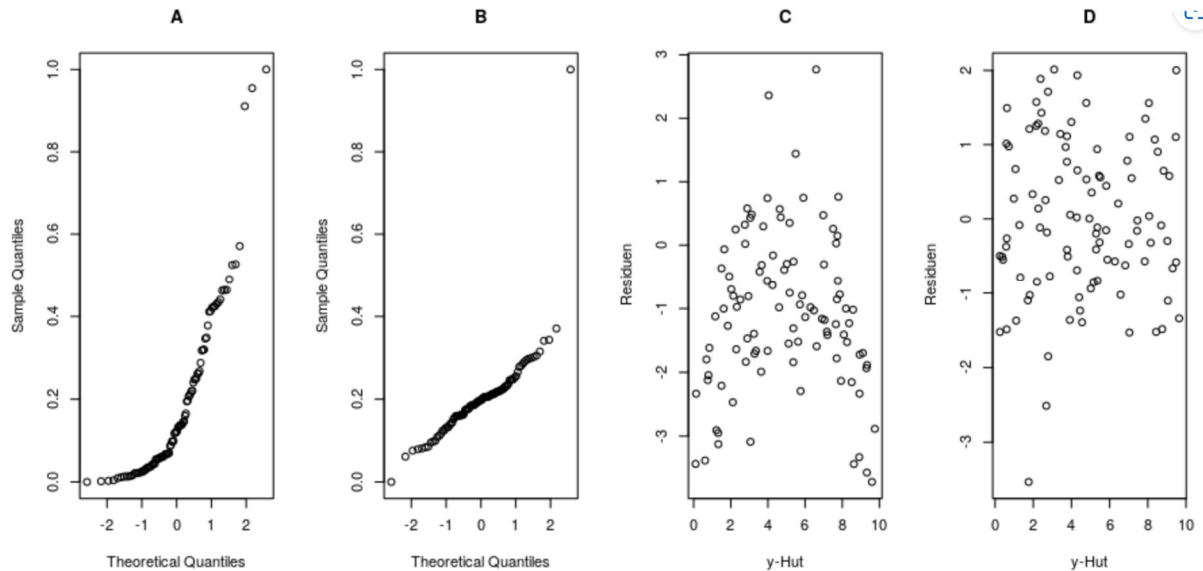
Im folgenden Bild sehen Sie acht Bilder in zwei Reihen und vier Spalten (A, B, C und D). Die Frage ist immer, ob es plausibel ist, dass die beiden Bilder (QQ-Plot oder Histogramm) in einer Spalte von den gleichen Daten erzeugt wurden.



Wählen Sie eine oder mehrere Antworten:

- ☐ a. Es ist plausibel, dass die beiden Grafiken in Spalte A mit den gleichen Daten erzeugt wurden.
- ☐ b. Es ist plausibel, dass die beiden Grafiken in Spalte B mit den gleichen Daten erzeugt wurden.
- ☐ c. Es ist plausibel, dass die beiden Grafiken in Spalte C mit den gleichen Daten erzeugt wurden.
- ☐ d. Es ist plausibel, dass die beiden Grafiken in Spalte D mit den gleichen Daten erzeugt wurden.

In der folgenden Abbildung sehen Sie vier Bilder. Die beiden Bilder links ('A' und 'B') sind QQ-Plots der Residuen von zwei verschiedenen Regressionsanalysen. Die beiden Bilder rechts ('C' und 'D') sind Tukey-Anscombe Plots zu zwei verschiedenen Regressionsanalysen. Die zugrunde liegenden Daten sind fuer alle vier Bilder unterschiedlich. Beantworten Sie zu jedem Plot eine Frage zur Residuenanalyse (die Aussagen sind entweder richtig oder relativ offensichtlich falsch):



Wählen Sie eine oder mehrere Antworten:

- ☐ a. Betrachten Sie den QQ-Plot in Abbildung 'A'. Die Residuen stammen von einer Rechtsschiefen Verteilung .
- ☐ b. Betrachten Sie den QQ-Plot in Abbildung 'B'. Die Residuen stammen von einer Rechtsschiefen Verteilung .
- ☐ c. Betrachten Sie den TA-Plot in Abbildung 'C'. Der Plot zeigt , dass die Geradengleichung des Modells nicht angebracht ist.
- ☐ d. Betrachten Sie den TA-Plot in Abbildung 'D'. Der Plot zeigt , dass die Geradengleichung des Modells nicht angebracht ist.