**Institut für Stochastik**
**Universität Karlsruhe**
**Priv.-Doz. Dr. D. Kadelka**

WS 2008/2009
30. April 2009

# Nachklausur zur Vorlesung

# Statistik für Biologen

## Musterlösungen

**Aufgabe 1** (15 Punkte)

Für die Länge $X$ und die Breite $Y$ von Samen einer Pflanze wurden die folgenden $n = 12$ Werte in $cm$ gemessen:

| $j$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $x_i$ | 4.9 | 5.2 | 5.4 | 5.5 | 5.8 | 6.1 | 6.5 | 7.1 | 7.4 | 7.5 | 8.6 | 8.7 |
| $y_j$ | 4.3 | 6.0 | 4.1 | 4.7 | 5.0 | 5.4 | 4.6 | 5.3 | 4.5 | 5.1 | 7.7 | 6.3 |
| $R_j$ | | | | | | | | | | | | |

Daraus berechnet man:

$$\bar{x}_{12} = 6.56, \quad \bar{y}_{12} = 5.25, \quad s_x^2 = 1.681, \quad s_y^2 = 1.026, \quad \sum_{j=1}^{12} x_j \cdot y_j = 422.31.$$

a) Fertigen Sie eine Stamm– und Blatt–Darstellung der $y_j$ mit der Einheit $1cm$ an. Der Stamm bestehe aus jeweils 1 Ziffer.

**Lösung:** Stamm- und Blatt-Darstellung ($n = 10$, Einheit $= 1$ $cm$):

| Stamm | Blätter |
|---|---|
| 4 | 3 1 7 6 5 |
| 5 | 0 4 3 1 |
| 6 | 0 3 |
| 7 | 7 |

b) Bestimmen Sie den Pearson–Korrelationskoeffizienten von $(x_1, y_1), \ldots, (x_{12}, y_{12})$ sowie die empirische Regressionsgerade von $y$ auf $x$.

**Lösung:** Direkt aus den Daten und den Hinweisen ergibt sich gemäß den Abschnitten 2.1, 2.8 und 3.2 des Skriptums unter Ausnützung der Beziehung

$$\sum_{j=1}^{n}(x_j - \bar{x}) \cdot (y_j - \bar{y}) = \sum_{j=1}^{n} x_j \cdot y_j - n \cdot \bar{x} \cdot \bar{y}$$

der empirische Pearson-Korrelationskoeffizient zu $r_{xy} = 0.6322$.

Nach Abschnitt 3.2 des Skriptums ist $y = a^* + b^* \cdot x$ mit $b^* = r_{xy} \cdot \dfrac{s_y}{s_x}$ und $a^* = \bar{y} - b^* \cdot \bar{x}$ die gesuchte Regressionsgerade, also

$$\begin{aligned} b^* &= 0.494 \\ a^* &= 2.010 \end{aligned}$$

und die Regressionsgerade $y = 2.010 + 0.494 \cdot x$.

c) Tragen Sie die Ränge $R_j$ der $y_j$–Werte in die obige Tabelle ein.

**Lösung:** Es ergibt sich die folgende Tabelle:

| $j$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $x_i$ | 4.9 | 5.2 | 5.4 | 5.5 | 5.8 | 6.1 | 6.5 | 7.1 | 7.4 | 7.5 | 8.6 | 8.7 |
| $y_j$ | 4.3 | 6.0 | 4.1 | 4.7 | 5.0 | 5.4 | 4.6 | 5.3 | 4.5 | 5.1 | 7.7 | 6.3 |
| $R_j$ | 2 | 10 | 1 | 5 | 6 | 9 | 4 | 8 | 3 | 7 | 12 | 11 |

d) Bestimmen Sie den Spearman–Rang–Korrelationskoeffizienten der $(x_1, y_1), \ldots, (x_{12}, y_{12})$.

**Lösung:** Für den Spearman-Rang-Korrelationskoeffizienten

$$\rho_{xy} = 1 - \frac{6}{n \cdot (n^2 - 1)} \cdot \sum_{j=1}^{n} (j - R_j)^2 = 1 - \frac{6}{12 \cdot (12^2 - 1)} \cdot \sum_{j=1}^{12} (j - R_j)^2$$

benützen wir die Ränge $R_j$ der $y_j$ aus c), die zum $j$-kleinsten $x_j$-Wert gehören. Damit

$$\rho_{xy} = 1 - \frac{6}{1716} \cdot ((1-2)^2 + (2-10)^2 + (3-1)^2 + (4-5)^2 + (5-6)^2 + (6-9)^2 + (7-4)^2$$
$$+ (8-8)^2 + (9-3)^2 + (10-7)^2) + (11-12)^2 + (12-11)^2) = 1 - \frac{1}{286} \cdot 136 = 0.5245.$$

**Aufgabe 2** (15 Punkte)

Ein Verfahren zur Bestimmung des Blutzuckerwertes liefere einen falschen Wert mit der Wahrscheinlichkeit $p = 0.015$. In einem Labor werden an einem Tag $n = 200$ Bestimmungen durchgeführt, wobei davon ausgegangen werden kann, dass sich die einzelnen Bestimmungen nicht gegenseitig beeinflussen.

a) Welche Verteilung hat $X$, die zufällige Anzahl der fehlerhaften Werte?

**Lösung:** Es liegt hier ein Treffer-Niete-Experiment vor mit $n = 200$ Versuchen und der Treffer-Wahrscheinlichkeit $p = 0.015$. (Ein Treffer liegt hier vor, wenn das Verfahren einen falschen Wert liefert.) Gemäß 5.2.1 hat also $X$, die zufällige Anzahl der Treffer, die Binomialverteilung

$$Bin(n, p) = Bin(200, 0.015) \; .$$

b) Beschreiben Sie das Ereignis $\{X \geq 3\}$.

**Lösung:** $\{X \geq 3\}$ ist das Ereignis, dass das Verfahren an einem Tag mindestens 3 falsche Werte liefert.

c) Bestimmen Sie die Wahrscheinlichkeit, dass mindestens zwei der Werte falsch sind.

**Lösung:** Gesucht ist die Wahrscheinlichkeit

$$\begin{aligned} \mathbb{P}(X \geq 2) &= 1 - \mathbb{P}(X \leq 1) = 1 - \mathbb{P}(X = 0) - \mathbb{P}(X = 1) \\ &= 1 - \binom{200}{0} \cdot 0.015^0 \cdot 0.985^{200} - \binom{200}{1} \cdot 0.015^1 \cdot 0.985^{199} \\ &= 1 - 1 \cdot 1 \cdot 0.985^{200} - 200 \cdot 0.015 \cdot 0.98^{199} = 0.8975\ . \end{aligned}$$

d) Bestimmen Sie den Erwartungswert und die Varianz von $X$.

**Lösung:** Allgemein gilt nach den Tabellen in 6.1.1 und 6.2.1, dass $Bin(n, p)$ den Erwartungswert $n \cdot p$ und die Varianz $n \cdot p \cdot (1 - p)$ hat. Der Erwartungswert von $X$ ist also $200 \cdot 0.015 = 3$ und die Varianz $200 \cdot 0.015 \cdot 0.985 = 2.955$.

e) Die Zufallsvariable $X$ besitzt näherungsweise eine Poisson-Verteilung $Po(\lambda)$ mit einem Parameter $\lambda > 0$. Wie groß ist $\lambda$ zu wählen?

**Lösung:** Nach den Überlegungen in 5.2.3 tritt die Poisson-Verteilung $Po(\lambda)$ mit $\lambda = n \cdot p$ bei großem $n$ und kleinem $p$ als Approximation von $Bin(n, p)$ auf. Diese Voraussetzungen sind hier erfüllt mit

$$\lambda = n \cdot p = 3\ .$$

**Aufgabe 3** (15 Punkte)

Der zufällige Wirkstoffgehalt $X$ (gemessen in $mg$) einer Tablette sei produktionsbedingt eine $N(300, 64)$-verteilte Zufallsvariable.

a) Mit welcher Wahrscheinlichkeit liegt der Wirkstoffgehalt $X$ einer Tablette oberhalb von 280 $mg$ und mit welcher Wahrscheinlichkeit höchstens 320 $mg$?

**Lösung:** Nach Voraussetzung gilt $X \sim N(300, 64) = N(300, 8^2)$. Gesucht ist zuerst $P(X > 280)$. Wegen Abschnitt 5.1.1 gilt

$$\begin{aligned} P(X > 280) = 1 - P(X \leq 280) = 1 - \Phi_{300,8^2}(280) &= 1 - \Phi\left(\frac{280 - 300}{8}\right) \\ &= 1 - \Phi(-2.5) = \Phi(2.5) = 0.9938\ . \end{aligned}$$

Für die gesuchte zweite Wahrscheinlichkeit ergibt sich ebenso

$$P(X \leq 320) = \Phi_{300,8^2}(320) = \left(\frac{320 - 300}{8}\right) = \Phi(2.5) = 0.9938\ .$$

b) Wie groß ist die Wahrscheinlichkeit, dass der Wirkstoffgehalt $X$ höchsten um 12 $mg$ vom Sollgehalt 300 $mg$ abweicht?

**Lösung:** Gesucht ist $P(|X - 300| \leq 12)$. Hier lässt sich Beispiel 5.6 aus dem Skriptum anwenden ($k \cdot \sigma$-Bereiche der Normalverteilung). Mit $k = 1.5$ und $\sigma = 8$ ist $k \cdot \sigma = 1.5 \cdot 8 = 12$, also

$$P(|X - 300| \leq 12) = P(|X - 300| \leq k \cdot \sigma) = 2 \cdot \Phi(k) - 1 = 2 \cdot \Phi(1.5) - 1$$

und damit die gesuchte Wahrscheinlichkeit

$$P(|X - 300| > 12) = 2 \cdot 0.9332 - 1 = 0.8664 \; .$$

c) Ein Tablettenröhrchen enthält 20 Tabletten. Wie groß ist die Wahrscheinlichkeit, dass alle 20 Tabletten mehr als 280 $mg$ Wirkstoffgehalt haben? (Setzen Sie voraus, dass die Wirkstoffgehalte der einzelnen Tabletten unabhängig voneinander sind.)
**Hinweis:** Verwenden Sie das Resultat aus a).

**Lösung:** Sei $Y$ die zufällige Anzahl der Tabletten, die einen Wirkstoffgehalt von mehr als 280 $mg$ haben. Nach Voraussetzung liegt hier ein Treffer-Niete Experiment mit $n = 20$ Versuchen und einer Trefferwahrscheinlichkeit $p = 0.9938$ vor. Die Zufallsvariable $Y$ hat also die Verteilung $Bin(n, p) = Bin(20, 0.9938)$ und die gesuchte Wahrscheinlichkeit ergibt sich zu

$$P(Y = 20) = \binom{20}{20} \cdot p^{20} \cdot (1 - p)^0 = 1 \cdot 0.9938^{20} \cdot 1 = 0.8830 \; .$$

**Aufgabe 4** (15 Punkte)

Man vermutet, dass bei Personen, die zu Herzinfarkt neigen, der Eisengehalt des Serums im allgemeinen höher ist als bei gesunden Personen. Aus umfangreichen Untersuchungen sei bekannt, dass der Median des Eisengehalts bei gesunden Männern 1.15 $\mu g/l$ (Mikrogramm pro Milliliter) beträgt.
Bei $n = 16$ Männern, die schon einen Infarkt überlebt haben und nach wie vor infarktgefährdet sind, misst man folgende Werte:

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 1.58, | 1.39, | 1.39, | 1.25, | 1.47, | 1.47, | 1.68, | 1.17, |
| 1.18, | 1.19, | 1.03, | 1.38, | 1.33, | 0.89, | 1.36, | 1.64. |

a) Testen Sie auf dem 0.01-Niveau die Hypothese, dass der (wahre) Median des Eisengehalts bei infarktgefährdeten Männern gleich 1.15 ($\mu g/l$) ist.

**Lösung:** Das Testproblem lautet: Teste $H_0\colon M = 1.15$ gegen $H_1\colon M \neq 1.15$, wobei $M$ den unbekannten Median der zugrundeliegenden Verteilung bezeichnet. Wir verwenden den Median-Vorzeichentest mit der Prüfgröße

$$N_+ = \text{ Anzahl der Stichprobenwerte, die größer als 1.15 sind.}$$

Aus den Daten folgt $N_+ = 14$.

Der Median-Vorzeichentest lehnt zum Niveau $\alpha$ die Hypothese $H_0$ genau dann ab, wenn $N_+ \leq k$ oder $N_+ \geq n - k$ mit $k$ aus Tabelle A.5 gilt. Hier ist $\alpha = 0.01$ und $n = 16$ und damit $k = 2$.

Wegen

$$N_+ = 14 \geq 16 - 2 = 14$$

wird $H_0$ zum Niveau $\alpha = 0.01$ abgelehnt.

b) Nehmen Sie nun an, dass bei infarktgefährdeten Männern der Eisengehalt des Serums eine Normalverteilung mit unbekannten Parametern $\mu$ und $\sigma^2$ besitzt. Testen Sie auf dem Niveau $\alpha = 0.05$ die Hypothese $H_0: \ \mu = 1.15 \ (\mu g/l)$.

**Hinweis:** Für die obigen 16 Werte gilt: $\bar{x}_{16} = 1.3375, \quad s_x = 0.2141.$

**Lösung:** Das Testproblem lautet jetzt: Teste $H_0\colon \mu = \mu_0 := 1.15$ gegen $H_1\colon \mu \neq 1.15$. Wir verwenden den zweiseitigen Einstichproben-$t$-Test mit der Prüfgröße $T = \dfrac{\sqrt{n} \cdot (\overline{x} - \mu_0)}{s_x}$. Aus den Daten folgt gemäß Hinweis

$$n = 16, \quad \overline{x} = 1.3375, \quad s_x \approx 0.2141,$$

und damit

$$T = \frac{\sqrt{16} \cdot (1.3375 - 1.15)}{0.2141} \approx 3.503 \ .$$

Weiter gilt für $\alpha = 0.05$ mit $1 - \alpha/2 = 0.995$ nach Tabelle A.2

$$t_{n-1,1-\alpha/2} = t_{15,0.995} \approx 2.95 \ .$$

$H_0$ wird zum Niveau $\alpha$ genau dann verworfen, wenn $|T| \geq t_{n-1,1-\frac{\alpha}{2}}$. Wegen

$$|T| = 3.503 > 2.95 = t_{15,0.995}$$

wird also die Hypothese $H_0$ zum Niveau $\alpha = 0.01$ verworfen.

c) Bestimmen Sie unter den Voraussetzungen von b) einen 0.99-Vertrauensbereich für $\mu$.

**Lösung:** Gemäß Satz 10.18 ist ein Konfidenzintervall für $\mu$ zur Konfidenzwahrscheinlichkeit $1 - \alpha = 0.01$

$$\left[\bar{x} - \frac{s_x}{\sqrt{n}} \cdot t_{n-1,1-\alpha/2}, \bar{x} + \frac{s_x}{\sqrt{n}} \cdot t_{n-1,1-\alpha/2}\right]$$

Einsetzen der Werte aus b) ergibt

$$\left[1.3375 - \frac{0.2141}{4} \cdot 2.95, 1.3375 + \frac{0.2141}{4} \cdot 2.95\right] = [1.1796, 1.4954] \ .$$

**Aufgabe 5** (15 Punkte)

Ein Imker hat zu Beginn einer Tracht 18 etwa gleichstarke Völker; 8 davon gehören zur Rasse A, die restlichen 10 zur Rasse B. Er möchte herausfinden, ob sich die beiden Rassen bei den gegebenen Verhältnissen hinsichtlich ihres Honigertrags unterscheiden. Bei der Schleuderung am Ende der Tracht erhält er von den Völkern der Rasse A die Mengen $x_1, \ldots, x_8$ an Honig, von den Völkern der Rasse B die Mengen $y_1, \ldots, y_{10}$ (gemessen jeweils in $kg$):

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Rasse A: | 13.2, | 14.2, | 16.1, | 18.0, | 18.5, | 18.9, | 19.3, | 20.4, | | |
| Rasse B: | 8.7, | 8.8, | 9.1, | 9.7, | 11.2, | 14.4, | 14.6, | 14.7, | 16.4, | 24.4 |

Man erhält daraus:

$$\bar{x}_8 = 17.325,\ s_x^2 = 6.565, \quad \bar{y}_{10} = 13.20,\ s_y^2 = 23.778.$$

a) Nehmen Sie an, dass die $x_i$ aus einer $N(\mu, \sigma^2)$-verteilten Grundgesamtheit und die $y_j$ aus einer $N(\nu, \sigma^2)$-verteilten Grundgesamtheit stammen. Testen Sie zum Niveau $\alpha = 0.05$ die Hypothese $H_0 : \mu = \nu$.

**Lösung:** Es handelt sich hier um ein Zwei-Stichproben-Problem. Die erste Stichprobe sind die $n = 8$ Werte $x_1, \ldots, x_8$ (Rasse A), die zweite Stichprobe die $m = 10$ Werte $y_1, \ldots, y_{10}$ (Rasse B). Zu testen ist

$$H_0 :\ \mu = \nu \quad \text{gegen} \quad H_1 :\ \mu \neq \nu.$$

Für diese Situation ist der Zwei-Stichproben-$t$-Test geeignet. Er hat die Prüfgröße

$$T = \frac{\sqrt{\frac{m \cdot n}{m+n}} \cdot (\bar{x} - \bar{y})}{\sqrt{\frac{1}{m+n-2} \cdot ((n-1) \cdot s_x^2 + (m-1) \cdot s_y^2)}} \approx \frac{2.108 \cdot 4.125}{\sqrt{\frac{1}{16} \cdot (7 \cdot 6.565 + 9 \cdot 23.778)}} \approx 2.157.$$

Der kritische Wert ist nach Tabelle A.2

$$t_{m+n-2,1-\alpha/2} = t_{16,0.975} \approx 2.12.$$

Wegen $|T| > t_{m+n-2,1-\alpha/2}$ wird die Hypothese abgelehnt.

b) Verzichten Sie jetzt auf die Normalverteilungsannahme, und nehmen Sie statt dessen nur noch an, dass die beiden Verteilungen stetig sind. Testen Sie jetzt auf dem Niveau $\alpha = 0.05$ die Hypothese, dass die beiden Verteilungen übereinstimmen.

**Lösung:** Wenn keine Normalverteilungsannahme vorliegt, sollte der $t$-Test nicht verwendet werden. Für dieses Problem ist der Mann-Whitney-$U$-Test geeignet. Die Hypothese $H_0$ lautet: Die Verteilungsfunktion $F$ der ersten Stichprobe und die Verteilungsfunktion $G$ der zweiten Stichprobe sind gleich, also

$$H_0 :\ F = G.$$

Zunächst werden zu den gegebenen 18 Daten die Ränge bestimmt. Wir ergänzen dazu die obige Tabelle durch die Ränge der einzelnen Werte.

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Rasse A: | 13.2, | 14.2, | 16.1, | 18.0, | 18.5, | 18.9, | 19.3, | 20.4, | | |
| | 6 | 7 | 11 | 13 | 14 | 15 | 16 | 17 | | |
| Rasse B: | 8.7, | 8.8, | 9.1, | 9.7, | 11.2, | 14.4, | 14.6, | 14.7, | 16.4, | 24.4 |
| | 1 | 2 | 3 | 4 | 5 | 8 | 9 | 10 | 12 | 18 |

Nun müssen die Ränge der $x_i$ aufaddiert werden:

$$W = 6 + 7 + 11 + 13 + 14 + 15 + 16 + 17 = 99.$$

Die Prüfgröße des Mann-Whitney-$U$-Tests ist

$$U = W - \frac{n(n+1)}{2} = 99 - \frac{8 \cdot 9}{2} = 63.$$

Der Mann-Whithney-$U$-Test lehnt die Hypothese ab, wenn eine der beiden Ungleichungen

$$U \geq U_{m,n,1-\alpha/2} \text{ oder } U \leq m \cdot n - U_{m,n,1-\alpha/2}$$

erfüllt ist. Aus Tabelle A.6 kann man ablesen:

$$U_{m,n,1-\alpha/2} = U_{8,10,0.975} = U_{10,8,0.975} = 63.$$

Wegen $U = 63 \geq U_{m,n,1-\alpha/2} = 63$ lehnt der Mann-Whitney-$U$-Test die Hypothese ab.

**Aufgabe 6** (15 Punkte)

Es sollen zwei Gastritis-Medikamente $A$ und $B$ verglichen werden. Die Medikamente wurden jeweils 500 Patienten verabreicht. Die Behandlung mit Medikament $A$ führte bei 319, mit Medikament $B$ bei 292 Patienten zum Erfolg.

a) Tragen Sie die obigen Werte an geeigneter Stelle in die Kontingenztafel ein und vervollständigen Sie die Tafel (insgesamt 9 Werte).

| Erfolg / Medikament | ja | nein | $\Sigma$ |
|---|---|---|---|
| $A$ | | | |
| $B$ | | | |
| $\Sigma$ | | | |

b) Testen Sie zum Niveau $\alpha = 0.05$ die Hypothese $H_0$, dass die Erfolgswahrscheinlichkeiten $p_1$ von Medikament $A$ und $p_2$ von Medikament $B$ gleich sind.

c) Bestimmen Sie einen 0.95-Vertrauensbereich für $p_1 - p_2$.

**Lösung:**

a) Die Kontingenztafel ergibt sich direkt aus der Aufgabenstellung:

| Medikament \ Erfolg | ja | nein | $\Sigma$ |
|---|---|---|---|
| $A$ | 319 | 181 | 500 |
| $B$ | 292 | 208 | 500 |
| $\Sigma$ | 611 | 389 | 1000 |

b) In der Gruppe der mit Medikament $A$ behandelten Patienten liegen $n_1 = 500$ Daten vor, von denen $a_1 = 319$ zum Erfolg gehören. In der Gruppe der mit Medikament $B$ behandelten Patienten liegen ebenfalls $n_2 = 500$ Daten vor, von denen $a_2 = 292$ zum Erfolg gehören. Schätzwerte für die unbekannten Erfolgswahrscheinlichkeiten sind

$$\hat{p}_1 = \frac{a_1}{n_1} = \frac{319}{500} = 0.638 \,, \qquad \hat{p}_2 = \frac{a_2}{n_2} = \frac{292}{500} = 0.584 \,.$$

Weiter ist

$$\hat{p} = \frac{a_1 + a_2}{n_1 + n_2} = \frac{319 + 292}{500 + 500} = \frac{611}{1000} = 0.611 \,.$$

Es soll die Hypothese

$$H_0: \ p_1 = p_2$$

getestet werden zum Niveau $\alpha = 0.05$ gegen die Alternative

$$H_1: \ p_1 \neq p_2 \,.$$

Es ist hier der Zwei–Stichproben–Test zum Vergleich von zwei Wahrscheinlichkeiten angebracht (S. 158–159 im Skript). Da $n_1 + n_2 = 1000 \geq 60$ ist, kann die Prüfgröße

$$T = \frac{n_1 \cdot n_2}{n_1 + n_2} \cdot \frac{(\hat{p}_1 - \hat{p}_2)^2}{\hat{p} \cdot (1 - \hat{p})} = \frac{500 \cdot 500}{1000} \cdot \frac{(0.638 - 0.584)^2}{0.611 \cdot (1 - 0.611)} \approx 3.067$$

angewendet werden. Der kritische Wert ist $\chi^2_{1,1-\alpha} = \chi^2_{1,0.95} \approx 3.84$ aus Tabelle A.3. Wegen $T < \chi^2_{1,1-\alpha}$ kann die Hypothese zum Niveau $\alpha = 0.05$ **nicht** abgelehnt werden.

c) Nach Abschnitt 10.10 ist mit dem $(1 - \alpha/2)$–Quantil der $N(0,1)$–Verteilung $c_{1-\alpha/2} = c_{0.975} = 1.96$ eine (approximative) untere Konfidenzgrenze für $p_1 - p_2$:

$$\begin{aligned}
& \hat{p}_1 - \hat{p}_2 - \frac{n_1 + n_2}{2 \cdot n_1 \cdot n_2} - c_{1-\alpha/2} \cdot \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}} \\
= \;& 0.638 - 0.584 - \frac{500 + 500}{2 \cdot 500 \cdot 500} - 1.96 \cdot \sqrt{\frac{0.638 \cdot 0.362}{500} + \frac{0.584 \cdot 0.416}{500}} \\
\approx \;& 0.054 - 0.062 = -0.008
\end{aligned}$$

und eine (approximative) obere Konfidenzgrenze für $p_1 - p_2$:

$$\begin{aligned}
& \hat{p}_1 - \hat{p}_2 + \frac{n_1 + n_2}{2 \cdot n_1 \cdot n_2} + c_{1-\alpha/2} \cdot \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}} \\
= \;& 0.638 - 0.584 + \frac{500 + 500}{2 \cdot 500 \cdot 500} + 1.96 \cdot \sqrt{\frac{0.638 \cdot 0.362}{500} + \frac{0.584 \cdot 0.416}{500}} \\
\approx \;& 0.054 + 0.062 = 0.116
\end{aligned}$$

Daher ist $[-0.008, 0.116]$ ein 0.95–Vertrauensbereich für $p_1 - p_2$.