Stochastik für die Naturwissenschaften

Dr. C.J. Luchsinger

Hip hip hooray - Jetzt beginnt die eigentliche Statistik!

8. Schätzen von Parámetern und Konfidenzintervalle

Literatur Kapitel 8

* Statistik in Cartoons: Kapitel 6 und Kapitel 7
* Stahel: Kapitel 7 und 9
* Storrer: Kapitel 42 und 43

8.0 Grundbegriffe der beurteilenden Statistik

Kapitel 42 in Storrer wird *in der Vorlesung* übersprungen. Sie sollten es sicher zwei mal lesen: einmal nach dieser Doppelstunde, dann am Schluss der Vorlesung nochmals. Lesen Sie dazu auch in diesem Skript 5.3 nochmals durch. Kurz zusammengefasst (Storrer Kapitel 42):

Statistik:
aus Stichprobe Rückschlüsse auf Gesamtheit ziehen

* Küken > Durchschnittsgewicht
* neue Kartoffelsorte > besser als alte?

Alte Sorte	410	420	430	440	450	450	480
Neue Sorte	440	450	455	480	490	505	

* Münze > fair?
* Anzahl Zerfälle > Poissonverteilt? Vgl Schluss Kapitel 7. Statistische Überprüfung einer einleuchtenden Theorie anhand von empirischen Daten.

Im ersten Beispiel wird die Zufallsgrösse (zufällig ist die Auswahl des Kükens) X das Gewicht sein und $E[X]$ das durchschnittliche Gewicht aller Küken.

Wir haben normalerweise nicht nur *ein* Küken in der Stichprobe, sondern allgemein n, zum Beispiel $n = 50$. Wie ist das zu verstehen? Die Daten sind dann

$$x_1, x_2, \ldots, x_{50}. \tag{8.1}$$

Das zugrunde liegende Experiment (Auswahl eines Kükens) wird hierzu 50 mal (unabhängig) wiederholt (mit "Zurücklegen", siehe Storrer Kapitel 42). Wir haben damit eine Folge von Zufallsgrössen

$$X_1, X_2, \ldots, X_{50}, \tag{8.2}$$

welche im konkreten Fall (8.1) hervorgebracht haben. Diese Folge in (8.2) wird mit iid Zufallsgrössen modelliert - ausser wir vereinbaren etwas anderes.

Lesen Sie dazu vor allem Seite 202 in Storrer aufmerksam durch - vielleicht werden Ihnen damit Kapitel 5 und 7 eher klar.

8.1 Schätzen von Parametern (Estimation of parameters)

8.1.1 Schätzen des Erwartungswerts

Gegeben sei eine Stichprobe vom Umfang 9. Der Mittelwert der Grundgesamtheit ist leider unbekannt. Wir wollen diesen Mittelwert schätzen. Die Stichprobe sieht folgendermassen aus (bereits der Reihe nach geordnet und auf 2 Stellen nach dem Komma gerundet):

$$-0.21, 0.11, 0.39, 0.64, 1.24, 1.46, 1.51, 2.89, 3.53$$

Wie könnten wir μ schätzen?

Wir bezeichnen einen Schätzer für einen Parameter μ mit $\hat{\mu}$ (für σ^2 mit $\hat{\sigma}^2$).

Welchen Vorschlag sollten wir warum auswählen?

MathematikerInnen haben mit gutem Grund das arithmetische Mittel

$$\frac{1}{n}\sum_{i=1}^{n} X_i$$

als *den* Schätzer für den Mittelwert identifiziert.

Warum und Einwände:

1. Einer von vielen Gründen ist die Erwartungstreue: Wie in Kapitel 5 berechnet, gilt nämlich $E\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right] = \mu$. Also wird man im Mittel (wenn wir immer wieder erneut mit neuen Daten mit dem arithmetischen Mittel den Erwartungswert berechnen) den Erwartungswert genau treffen.

2. Wenn wir zudem immer mehr Daten erhalten, also das n erhöhen, wird wegen des LLN der Schätzwert immer näher an den Erwartungswert kommen (siehe Kapitel 7.1).

3. Leider hat das arithmetische Mittel einen grossen Nachteil: die starke Abhängigkeit von Ausreissern, also extremen Werten. Wenn auch nur ein Datenwert völlig falsch ist (Übermittlungs- oder Ablesefehler), dann wird das arithmetische Mittel eventuell auch total falsch. Der Median ist robuster.

Das heisst für die Praxis:

8.1.2 Schätzen der Varianz

Wir haben bereits in der beschreibenden Statistik den folgenden Schätzer für die Varianz kennengelernt:

$$\frac{1}{(n-1)}\sum_{j=1}^{n}(x_j-\overline{x})^2.$$

Warum wird vorgeschlagen, durch $(n-1)$ zu teilen? Die Antwort findet sich in (43.6), wobei die konkreten Rechnungen analog 8.1.1 - aber viel komplizierter - sind. Es geht wieder um die Erwartungstreue:

$$E\left[\frac{1}{(n-1)}\sum_{j=1}^{n}(X_j-\overline{X})^2\right]=\sigma^2.$$

8.1.3 Durchschnitt als Zufallsgrösse

Wir sind bis jetzt (vgl Teil 8.0 und 5.3) locker zwischen kleinen x und grossen X hin und her gesprungen.

$$\frac{1}{n}\sum_{i=1}^{n}x_i$$

ist eine Zahl. Aber bei anderer Auswahl würde diese Zahl wohl leicht anders aussehen. Wenn wir den Ausdruck

$$\frac{1}{n}\sum_{i=1}^{n}X_i$$

untersuchen, erfahren wir etwas über den erwarteten Wert des arithmetischen Mittels und über die Varianz des arithmetischen Mittels. Wir repetieren nicht zum ersten Mal von Kapitel 5: $E[\overline{X}]=E\left[\frac{1}{n}\sum_{i=1}^{n}X_i\right]=\mu$ und $V[\overline{X}]=V\left[\frac{1}{n}\sum_{i=1}^{n}X_i\right]=\frac{1}{n}\sigma^2$ und damit $sd[\overline{X}]=\frac{1}{\sqrt{n}}\sigma$. Des weiteren erinnern wir uns von Kapitel 6: Summe von Normal ist Normal (Summanden müssen nicht mal identisch verteilt sein) $\sum_{i=1}^{n}\mathcal{N}(\mu_i,\sigma_i^2)=\mathcal{N}(\sum_{i=1}^{n}\mu_i,\sum_{i=1}^{n}\sigma_i^2)$. Wir schauen nochmals die Z-Transformierte an: Wenn X eine $\mathcal{N}(\mu,\sigma^2)$-Verteilung hat, dann hat $Z:=\frac{X-\mu}{\sigma}$ eine $\mathcal{N}(0,1)$-Verteilung. Damit können wir interessante Schlussfolgerungen ziehen:

1. [schwache Voraussetzungen] Wenn wir eine Folge von - nicht unbedingt normalverteilten - iid Zufallsgrössen $X_1, \ldots, X_n$ haben mit Standardabweichung σ, dann hat das arithmetische Mittel (als Zufallsgrösse) eine Standardabweichung von

$$\sigma_{\overline{X}} := sd[\overline{X}] = \frac{\sigma}{\sqrt{n}}. \qquad \text{(sog. Standardfehler)}$$

Das σ ist normalerweise nicht bekannt und wir werden es mit

$$s := \sqrt{\frac{1}{(n-1)} \sum_{j=1}^{n} (x_j - \overline{x})^2}$$

schätzen. Zusammen erhalten wir als *Schätzung* des Standardfehlers folgenden Ausdruck:

$$s_{\overline{X}} := \frac{s}{\sqrt{n}} = \sqrt{\frac{\sum_{j=1}^{n} (x_j - \overline{x})^2}{n(n-1)}}.$$

Beispiel Storrer p 207 unten schauen Sie als Hausaufgabe an.

Warnung und Konsequenzen: Verwechseln Sie nicht Standardfehler ($sd[\overline{X}]$) und die Standardabweichung ($sd[X]$) selber: Im Standardfehler wird durch $\sqrt{n}$ geteilt und damit wird er immer kleiner, je grösser wir das n (=Stichprobengrösse) machen. Dies sollte ja auch so sein: je mehr Daten wir haben, desto präziser können wir schätzen. Die Präzision ist aber derart, dass man zum Beispiel vier mal mehr Daten braucht, um eine doppelte Genauigkeit (Standardfehler halbieren) zu erhalten. Warum schaut man diesen Ausdruck überhaupt an?

2. [starke Voraussetzungen]: Wenn wir sogar eine Folge $X_1, \ldots, X_n$ von iid $\mathcal{N}(\mu, \sigma^2)$-Zufallsgrössen haben, so können wir sogar viel weiter gehen: Wegen

$$\sum_{i=1}^{n} \mathcal{N}(\mu_i, \sigma_i^2) = \mathcal{N}(\sum_{i=1}^{n} \mu_i, \sum_{i=1}^{n} \sigma_i^2)$$

hat die Summe $\sum_{i=1}^{n} X_i$ eine

$$\mathcal{N}(n\mu, n\sigma^2)$$

-Verteilung. Wegen Lemma 5.3, Lemma 5.4 a) und der Z-Transformierten hat

$$\overline{X}$$

eine

$$\mathcal{N}(\mu, \frac{\sigma^2}{n})$$

-Verteilung. Dies gibt das Bild in Storrer p 208. Weil die Varianz immer kleiner wird, wird die Dichtefunktion immer enger und damit die Schätzung immer präziser. Wir rechnen hierzu noch das Zahlenbeispiel in Storrer p 209 oben.

8.2 Konfidenzintervalle (Intervall-Schätzer)

Eigentlich interessieren wir uns für einen unbekannten Parameter, den wir schätzen wollen. Konfidenzintervalle geben uns eine Vorstellung von der Präzision eines Schätzers / einer Schätzung.

Ziele von 8.2: Die StudentInnen wissen, was ein Konfidenzintervall (KI) ist - und was es *nicht* ist. Sie erkennen Formeln für KI in 4 wichtigen Fällen wieder und können diese anwenden.

8.2.1 Was ist ein Konfidenzintervall (KI) - was ist es *nicht*

Was ist es *nicht* - drei Beispiele: 1. Sozialwissenschaften: In den Nachrichten liest man oft Sätze der Art (Zahlen frei erfunden): "Aufgrund einer Befragung mit einer Stichprobe von 10'000 Personen kam man zum Schluss, dass der Anteil der Anhänger von Bundeskanzlerin Merkel mit 95 % Wahrscheinlichkeit in einem Konfidenzintervall von $[46\%, 48\%]$ liegt." Was ist hier falsch?

2. Physik:

3. Stochastik:

Definition 8.1 [Konfidenzintervalle] *Ein Konfidenzintervall KI für μ mit Konfidenzkoeffizient $(1-\alpha)$ ist eine zufällige Teilmenge von $\mathbb{R}$ mit der Eigenschaft, dass*

$$P[\mu \in KI] = 1 - \alpha$$

für alle $\mu \in \mathbb{R}$ (siehe auch Storrer p 217).

Bemerkungen zu Definition 8.1: Im Teil "was ist es *nicht*" haben wir bereits betont, dass μ nicht zufällig ist. Das Konfidenzintervall KI ist zufällig (*vor* der Realisation). Wenn danach die Realisation mit konkreten Zahlen vorliegt, sollte man Formulierungen brauchen wie: "$[46\%, 48\%]$ ist eine Realisation eines 95 % Konfidenzintervalles". Dies machen nur MathematikerInnen; Sie dürfen sagen:

"$[46\%, 48\%]$ ist ein 95 % Konfidenzintervall".

8.2.2 Konstruktion von KI's

8.2.2.1 $\mathcal{N}(\mu, \sigma^2)$: KI für μ wenn σ^2 bekannt

Daten $x_1, \ldots, x_n$, berechnen arithmetisches Mittel $\overline{x}$. Bei Luchs im Büro; vor Versuch: Wechseln zu Zufallsgrössen: betrachten $\overline{X}$. Verteilung von $\overline{X}$ ist:

Also ist

$$\frac{\overline{X} - \mu}{\sqrt{\frac{\sigma^2}{n}}}$$

wegen der "Z-Transform" eine $\mathcal{N}(0,1)$-Zufallsgrösse. Aber das ist ja fantastisch: wir kennen die kritischen Werte bei einer $\mathcal{N}(0,1)$-Verteilung:

$$P\left[-1.96 \leq \frac{\overline{X} - \mu}{\sqrt{\frac{\sigma^2}{n}}} \leq 1.96\right] = 95\%. \tag{8.3}$$

Ziel ist offensichtlich ein $(1-\alpha) = 95\%$-KI! Aber: wir können offensichtlich (8.3) (noch) nicht als KI verkaufen. Machen wir ein paar algebraische Umformungen (Resultat ist dann (8.8)):

$$P\left[-1.96\sqrt{\frac{\sigma^2}{n}} \leq \overline{X} - \mu \leq 1.96\sqrt{\frac{\sigma^2}{n}}\right] = 95\%. \tag{8.4}$$

Wir addieren μ:

$$P\left[\mu - 1.96\sqrt{\frac{\sigma^2}{n}} \leq \overline{X} \leq \mu + 1.96\sqrt{\frac{\sigma^2}{n}}\right] = 95\%, \tag{8.5}$$

formen es noch um zu

$$P\left[\overline{X} \in \left[\mu - 1.96\sqrt{\frac{\sigma^2}{n}}, \mu + 1.96\sqrt{\frac{\sigma^2}{n}}\right]\right] = 95\%. \tag{8.6}$$

Dies ist aber gleichbedeutend mit

$$P\left[\mu \in \left[\overline{X} - \frac{1.96\sigma}{\sqrt{n}}, \overline{X} + \frac{1.96\sigma}{\sqrt{n}}\right]\right] = 95\%. \tag{8.7}$$

Wir vergleichen voller Stolz und Befriedigung (8.7) mit Definition 8.1.

Also gilt (nach Versuch, mit Daten): *ein* 95 % KI für μ bei bekanntem σ^2 und Daten $x_1, \ldots, x_n$ in Modell $\mathcal{N}(\mu, \sigma^2)$ ist:

$$\left[\overline{x} - \frac{1.96\sigma}{\sqrt{n}}, \overline{x} + \frac{1.96\sigma}{\sqrt{n}}\right]. \tag{8.8}$$

99 % KI in dieser Situation ist:

Was geschieht wenn σ^2 wächst?

Was geschieht wenn n wächst?

Kleine Aufgabe zu KI I: Es wird angenommen, dass die Durchmesser der auf einer bestimmten Anlage hergestellten Stahlkugeln durch die Realisationen einer normalverteilten Zufallsgrösse mit $\sigma = 1.04$ mm beschrieben werden können. Aus einer Stichprobe vom Umfang $n = 30$ ergab sich $\bar{x} = 12.14$ mm. Bestimmen Sie für die Vertrauenswahrscheinlichkeit von 0.95 und 0.99 die Grenzen des KI für den mittleren Durchmesser dieser Kugeln.

8.2.2.2 $\mathcal{N}(\mu, \sigma^2)$: KI für μ wenn σ^2 unbekannt (realistisch), Storrer (43.8)

Daten $x_1, \ldots, x_n$, berechnen das arithmetische Mittel $\overline{x}$. Bei Luchs im Büro; vor Versuch: Wechseln auf die Ebene der Zufallsgrössen: betrachten $\overline{X}$. Verteilung von $\overline{X}$ ist:

Also hat wegen der "Z-Transform" der Ausdruck

$$\frac{\overline{X} - \mu}{\sqrt{\frac{\sigma^2}{n}}}$$

eine $\mathcal{N}(0, 1)$-Verteilung. Ooooops? σ^2 ist unbekannt! Keine Panik: Wir werden es einfach schätzen:

$$\hat{\sigma}^2 := \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2.$$

Das Tier

$$\frac{\overline{X} - \mu}{\sqrt{\frac{\hat{\sigma}^2}{n}}} = \frac{\overline{X} - \mu}{\sqrt{\frac{\frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2}{n}}} \tag{8.9}$$

hat eine t_{n-1}-Verteilung! Die t-Verteilung ist "fast" eine Normalverteilung. Das "fast" kommt daher, dass wir im Nenner von (8.9) nicht das (eben unbekannte) σ einsetzen können, sondern nur eine Schätzung hiervon. Die Schätzung von σ bedeutet mehr Unsicherheit, womit die t-Verteilungen mehr Gewicht in den Enden haben. In Storrer (43.7) haben Sie hierzu mehr Informationen. In meinem Skript ist in 6.2.6 ebenfalls die t-Verteilung aufgeführt.

Mit den gleichen Schritten wie in 8.2.2.1 folgt dann analog zu (8.8); nach Versuch, mit Daten:

$$\left[\overline{x} - \frac{CV\hat{\sigma}}{\sqrt{n}}, \overline{x} + \frac{CV\hat{\sigma}}{\sqrt{n}}\right], \tag{8.10}$$

wo

$$\hat{\sigma} := \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2}. \tag{8.11}$$

Der kritische Wert (Critical Value CV) ist jetzt nicht mehr nur von $(1-\alpha)$ abhängig (wie in 8.2.2.1), sondern auch von n (vgl. z.B. Tabelle in (51.4)). Zum Beispiel mit $n = 20$ sind die kritischen Werte für ein 95%-KI:

zu vergleichen mit den 1.96 von 8.2.2.1.

Kleine Aufgabe zu KI II: gleiche Ausgangslage wie bei Aufgabe I, aber wir kennen σ nicht und haben es mit (8.11) genau auf 1.12 mm geschätzt. Berechnen Sie das 0.95-KI nochmals in dieser Situation.

Was passiert wenn n grösser wird?

Was wenn σ^2 grösser wird?

Storrer A 43-8:

8.3. Die 4 wichtigsten KI aus der Praxis (beyond Storrer)

Orientierungsschema:

In den bisherigen Situationen (8.2.2.1 und 8.2.2.2) stellen wir fest, dass dortige 95 % Konfidenzintervalle approximativ immer nach dem gleichen Schema aufgebaut sind:

$$[\overline{x} - 2\mathrm{sd}(\overline{X}), \overline{x} + 2\mathrm{sd}(\overline{X})]. \qquad \text{(generic)}$$

Die Zahl "2" ist im Fall der Normalverteilung eigentlich präziser unser bekanntes 1.96. Den Unterschied zwischen 1.96 und 2 ignorieren Praktiker/innen gerne - wir jetzt auch. Wegen des CLT haben wir bei grossen n approximativ in der Summe ($\overline{x}$ beinhaltet eine Summe!) ebenfalls fast eine Normalverteilung (technisch: solange die X_i's eine endliche Varianz besitzen). Deshalb können wir in den noch folgenden 3 Fällen A, B und C obige generische Formel (generic) benutzen.

A (Bernoulli/Binomial): Sei x_i gleich 1 bei Erfolg und gleich 0 bei Misserfolg bei Versuch i. Dann ist $\sum_{i=1}^n x_i$ die gesamte Anzahl Erfolge und damit $\bar{x} := \frac{1}{n}\sum_{i=1}^n x_i$ der *Anteil* der Erfolge (zB bei $n = 100, \sum_{i=1}^{100} x_i = 57, \bar{x} = 0.57$). Weil wir jetzt ein Konfidenzintervall für die Erfolgs*wahrscheinlichkeit* wollen, ändern wir gleich die Bezeichnung und setzen:

$$\hat{p} := \bar{x} := \frac{1}{n}\sum_{i=1}^{n} x_i.$$

Wir wollen ein 95 % KI für die wahre Wahrscheinlichkeit p konstruieren. Dies wird auch in HHS (43.9) gemacht; hier ist es einfacher: von (generic) her haben wir - auf der Ebene von Zufallsgrössen $\hat{p} \pm 2\mathrm{sd}(\hat{p})$. Wir kennen $\mathrm{sd}(\hat{p})$ nicht. Es gilt:

$$V[\hat{p}] = V\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right] = \frac{1}{n^2}V\left[\sum_{i=1}^{n} X_i\right] = \frac{1}{n^2}\sum_{i=1}^{n} V[X_i] = \frac{1}{n^2}nV[X_1] = \frac{1}{n}p(1-p).$$

Damit gilt $\mathrm{sd}(\hat{p}) = \sqrt{\frac{p(1-p)}{n}}$. Leider ist p unbekannt; wir ersetzen es deshalb durch die Schätzung und erhalten $\mathrm{sd}(\hat{p}) \doteq \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$. Damit erhalten wir für das KI

$$\left[\hat{p} - 2\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}, \hat{p} + 2\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right].$$

Das ist Formel (4) in HHS, Seite 220. Eine (uns nicht unbekannte) Praktikerregel verlangt hier, dass $n\hat{p}(1-\hat{p}) > 9$ sein muss, sonst muss man die sogenannten Geigy-Tabellen oder die sogenannte Methode nach Wilson (das ist im Storrer) anwenden. Wir rechnen zur Kontrolle obiger Formel gleich mal das Beispiel aus 8.2.1 nach, bei der es um das KI zur Anhängerschaft von Angie ging:

Die Formel sieht für die USA (324 M Einwohner), Schweiz (8400000 Einwohner) und Nauru (10084 Einwohner) genau gleich aus!?

B (Bernoulli/Binomial; KI für *Differenz* bei Proportionen): Wir haben jetzt 2 Stichproben (zB Erfolg von 2 Operationsmethoden oder bei Abstimmungen ob Romandie anders stimmen wird als Deutschschweiz). Analog zu A haben wir m Objekte in Gruppe 1 und n Objekte in Gruppe 2. Sei x_i gleich 1 bei Erfolg und gleich 0 bei Misserfolg bei Versuch i in Gruppe 1. Dann ist $\sum_{i=1}^m x_i$ die gesamte Anzahl Erfolge und damit $\bar{x} := \frac{1}{m}\sum_{i=1}^m x_i$ der *Anteil* der Erfolge. Analog sei y_j gleich 1 bei Erfolg und gleich 0 bei Misserfolg bei Versuch j in Gruppe 2. Dann ist $\sum_{j=1}^n y_j$ die gesamte Anzahl Erfolge und damit $\bar{y} := \frac{1}{n}\sum_{j=1}^n y_j$ der *Anteil* der Erfolge. Weil wir wieder ein Konfidenzintervall für die Erfolgs*wahrscheinlichkeiten* (jetzt aber für die Differenz ebendieser Wahrscheinlichkeiten) wollen, ändern wir gleich die Bezeichnung und setzen:

$$\hat{p}_1 := \bar{x} := \frac{1}{m}\sum_{i=1}^m x_i \quad , \qquad \hat{p}_2 := \bar{y} := \frac{1}{n}\sum_{j=1}^n y_j .$$

Wir wollen ein 95 % KI für die Differenz der Wahrscheinlichkeiten $p_1 - p_2$ konstruieren: von (generic) her haben wir - auf der Ebene von Zufallsgrössen $(\hat{p}_1 - \hat{p}_2) \pm 2\mathrm{sd}(\hat{p}_1 - \hat{p}_2)$. Sie berechnen bitte analog zum Fall A $\mathrm{sd}(\hat{p}_1 - \hat{p}_2)$. Wir erhalten schlussendlich für das KI:

$$\left[(\hat{p}_1 - \hat{p}_2) - 2\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{m} + \frac{\hat{p}_2(1-\hat{p}_2)}{n}} \;,\; (\hat{p}_1 - \hat{p}_2) + 2\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{m} + \frac{\hat{p}_2(1-\hat{p}_2)}{n}}\right].$$

Praktikerrregel: $\hat{p}_i \in [0.1, 0.9]$ und m, n je grösser als 30; sonst mit der Methode nach Wilson (googeln).

C (2 Stichproben; Normalverteilung; KI für *Differenz* der Mittelwerte): Wir haben wieder 2 Stichproben (zB Ertrag Düngemittel von 2 verschiedenen Düngern). Analog zu B haben wir m Objekte in Gruppe 1 und n Objekte in Gruppe 2. Sei x_i der Wert bei Versuch i in Gruppe 1 - wir modellieren mit einer $\mathcal{N}(\mu_1, \sigma^2)$-Zufallsgrösse. Dann ist $\bar{x} := \frac{1}{m}\sum_{i=1}^{m} x_i$ der Durchschnittswert in Gruppe 1. Analog sei y_j der Wert bei Versuch j in Gruppe 2 - wir modellieren mit einer $\mathcal{N}(\mu_2, \sigma^2)$-Zufallsgrösse. **Die Varianzen der einzelnen Messungen in Gruppe 1 und 2 müssen gleich sein!** Dann ist $\bar{y} := \frac{1}{n}\sum_{j=1}^{n} y_j$ der Durchschnittswert in Gruppe 2. Wir wollen ein 95 % KI für die Differenz der Mittelwerte $\mu_1 - \mu_2$ konstruieren: von (generic) her haben wir - auf der Ebene von Zufallsgrössen $(\bar{X} - \bar{Y}) \pm 2\mathrm{sd}(\bar{X} - \bar{Y})$. Analog zu A und B wird auch hier $\mathrm{sd}(\bar{X} - \bar{Y})$ berechnet, wobei für die Schätzung von σ^2 ein hier nicht begründeter ungewöhnlicher Ausdruck verwendet wird: Wir erhalten für das KI zuerst:

$$\left[(\bar{x} - \bar{y}) \ \ \pm \ \ 2\sqrt{\left(\frac{1}{m} + \frac{1}{n}\right)\frac{\sum_{i=1}^{m}(x_i - \bar{x})^2 + \sum_{j=1}^{n}(y_j - \bar{y})^2}{m+n-2}}\right].$$

Hier wird *in Abweichung des bisherigen Vorgehens* nicht 2 benutzt, sondern der entsprechende Term aus der t_{m+n-2}-Verteilung, da man σ^2 geschätzt hat (**Fehlerquelle Prüfung**).

$$\left[(\bar{x} - \bar{y}) \ \ \pm \ \ \mathrm{CV}\sqrt{\left(\frac{1}{m} + \frac{1}{n}\right)\frac{\sum_{i=1}^{m}(x_i - \bar{x})^2 + \sum_{j=1}^{n}(y_j - \bar{y})^2}{m+n-2}}\right].$$

Alle 4 Fälle für KI müssen für die Prüfung beherrscht werden. In den Übungen werden wir aber nicht in jedem Semester zu jedem Fall ein Beispiel lösen. Schauen Sie stattdessen die 10 letzten Prüfungen im Archiv an.

8.4. KI: wie gross muss n sein?

Wichtig:

1. Lesen Sie jetzt das komplette Kapitel im Storrer II selber durch (Kapitel 42-43).
2. Lösen Sie danach mindestens 5 Aufgaben hinten im Kapitel und vergleichen Sie mit den Lösungen am Schluss des Buches. Bei Bedarf lösen Sie mehr Aufgaben.
3. Gehen Sie in die Übungsstunde. Drucken Sie das Übungsblatt dazu *vorher* aus, lesen Sie *vorher* die Aufgaben durch und machen sich erste Gedanken dazu (zum Beispiel, wie man sie lösen könnte).
4. Dann lösen Sie das Übungsblatt: zuerst immer selber probieren, falls nicht geht: Tipp von Mitstudi benutzen, falls immer noch nicht geht: Lösung von Mitstudi anschauen, 1 Stunde warten, versuchen, aus dem Kopf heraus wieder zu lösen, falls immer noch nicht geht: Lösung von Mitstudi abschreiben (und verstehen - also sollte man insbesondere keine Fehler abschreiben!).
5. Lösen Sie die entsprechenden Prüfungsaufgaben im Archiv.