Inferential Statistics

VB

Institut für Veterinär-Epidemiologie und Biometrie, FU Berlin

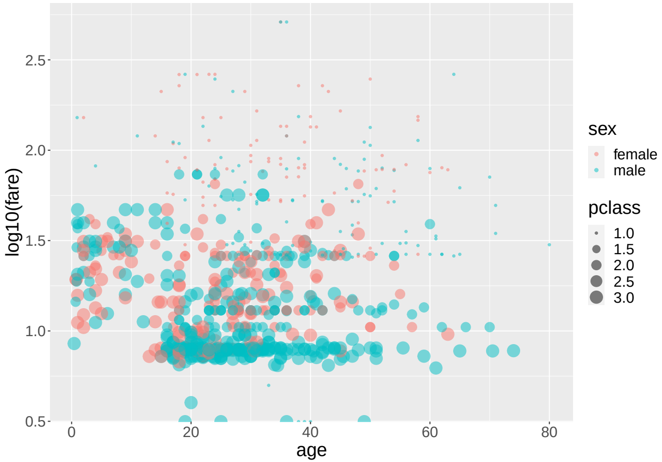
10/20/2020

Wiederholung

In der letzen Vorlesung haben wir über die über Beschreibende Statistik gersprochen.

- Charakteristische Maßzahlen
- Darstellungsforment: Tabellen, Histogram, Boxplot

Bivariate Daten



Was haben wir letztes mal gelernt?

Schliessende Statistik

- Test-Theorie
- Intervallschätzung

Schätzwerte und Tests

Basierend auf Schtichprobe, *schätz* man die Parameter der Grundgesamtheit.

Beispiel

Langjährige Beobachtungen (Grundgesamtheit): $\rho = 48$ der Neugeborenen sind Mädchen.

Erhebung aus drei Krankenhäusern (Stichprobe): 680 Geburten mit $\hat{\rho} = 51$ Mädchen.

Ist die Erhöhung *zufällig* oder nicht und aufgrund einer Ursache *signifikan*t?

Test-Theorie

Die Test-Theorie stellt eine Verbindung zwischen Stichtproben und Grundgesamtheit.

Es wird geprüft, aufgrund von Stichprobenwerten, ob gewisse *Hypothesen* über die Grundgesamtheit wahr sind oder nicht.

Es soll entschieden werden ob eine Hypothese beizubehalten oder zu verwerfen ist.

Hypothesen H_0 und H_1

Nullhypothese ist normalerweise die Behauptung, dass eine Behandlung oder Maßnahme in einem Versuch *keine Auswirkungen* hat und jegliche Unterschiede zwischen den Messwerten nur durch *Zufall* entstanden sind.

Beispiel

Man hat die Hypothese, dass die normalverteilte Grundgesamtheit einen wahren Mittelwert von $\mu = 18$ hat. Der aus einer Stichprobe ermittelte Mittelwert beträgt $\hat{\mu} = 19.5$.

- *Nullhypothese* H_0: der wahre Mittelwert μ ist gleich dem theoretischen Wert
- *alternative Hypothese* H_1: der wahre Mittelwert μ ist *nicht* gleich dem theoretischen Wert

Fehler 1. Art und 2. Art

Die Tests können nur *statistische* Aussagen über die “Wahrheit” der Hypothesen machen. Dabei können Fehler von zwei Arten vorkommen.

α-Fehler (1. Art)

Durch den Test wird die *Nullhypothese verworfen, obwohl sie in Wirklichkeit richtig* ist (*false negatives*).

β-Fehler (2. Art)

Die *Nullhypothese wird beibehalten, obwohl sie in Wirklichkeit falsch* ist (*false positives*).

α-Fehler

Die Wahrscheinlichkeit des α-Fehlers können wir für den Test selbst festlegen. Bekannt auch als *Signifikanzniveau* α, α-*Risiko*, *Irrtumswahrscheinlichkeit* α.

Beispiel

Eine Äpfellieferung wird auf Qualität geprüft und darf nicht mehr als 15% von Obst schlechter Qualität enthalten. Von jeder Palette (*Grundgesamtheit*) werden 10 Äpfel (*Stichprobe*) untersucht.

- H_0: die Untersuchte Palette ist "gut", i.e. enthält höchstens 15% schlechte Äpfel
- H_1: die Untersuchte Palette ist nicht "gut", i.e. enthält mehr als 15% schlechte Äpfel

α-Fehler(1)

Anzahl schlechtere Äpfel	i	0	1	2	3	4	5	≥ 6
WS für genau i schlechtere Äpfel	$P(i)$	0.197	0.347	0.276	0.130	0.040	0.009	0.001
WS für höchstens i schlechtere Äpfel	$\sum P(i)$	0.197	0.544	0.820	0.950	0.990	0.999	1.000

Figure 1: Wahrscheinlichkeiten berechnet nach Binomialverteilung mit $k = 10$ Ziehungen und $p = 0.15$. [Köhler et al.]

Bei Richtigkeit unserer Nullhypothese mit einer Wahrscheinlichkeit von 0.95 finden wir höchstens 3 Äpfel schlechter Qualität pro Stichprobe. Nur in 5% der Fällen würden wir in der Stichprobe mehr als 3 schlechte Äpfel vorfinden.

Wenn wir bereit sind in 5% der Fälle die Nullhypothese abzulehnen, würden wir ein α-Risiko von $\alpha = 5\%$ akzeptieren.

α-Fehler(2)

In der Stichprobe von Umfang $k = 10$ werden i schlechtere Äpfel vorgefunden.

- Ist $i \leq K$, wird H_0 angenommen.
- Ist $i > K$, wird H_0 abgelehnt.

Dabei wird $K = 3$ *kritischer Wert* genannt.

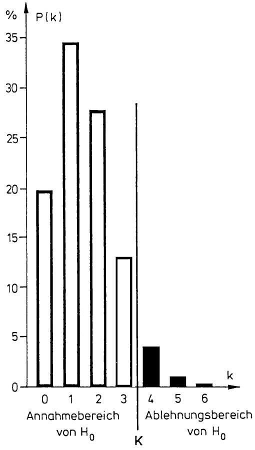


Figure 2: Wahrscheinlichkeiten berechnet nach Binomialverteilung mit $k = 10$ Ziehungen und $p = 0.15$. [Köhler et al.]

β-Fehler

Die *Nullhypothese wird beibehalten, obwohl sie in Wirklichkeit falsch* ist (*false positives*).

Beispiel

Falls die Äpfelpaletten 50% Äpfel minderer Qualität anstatt 15% hätten und unser Testverfahren mit $K = 3$ würde "schlechte" Paletten als "gute" akzeptieren, wäre das ein Fehler 2. Art.

i	0	1	2	3	4	5	6	7	≥ 8
$P(i)$	0.001	0.010	0.044	0.117	0.205	0.246	0.205	0.117	0.055
$\sum P(i)$	0.001	0.011	0.055	0.172	0.377	0.623	0.828	0.945	1.000

Figure 3: Wahrscheinlichkeiten berechnet nach Binomialverteilung mit $k = 10$ Ziehungen und $p = 0.5$. [Köhler et al.]

Mit Wahrscheinlichkeit von 0.172 würden wir die schlechten Paletten als gut akzeptieren, weil $\sum_j^3 P(j) = 0.172$ für $K = 3$.

β-Fehler(1)

Wir konnten die Größe des β-Fehlers nur berechnen, indem wir unterstellten, der wahre Anteil schlechterer Äpfel beträge $p = 50\%$.

Meist kennt man den wahren Wert von p nicht. Daher ist β unbekannt und im Falle der Beibehaltung von H_0 weiss man nicht wie groß die Wahrscheinlichkeit ist, dass die beibehaltene H_0 falsch ist.

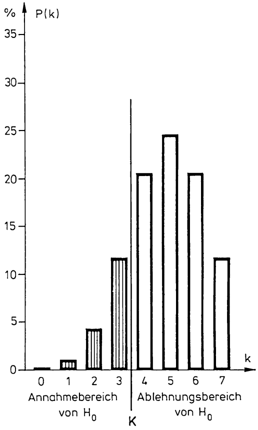


Figure 4: Wahrscheinlichkeiten berechnet nach Binomialverteilung mit $k = 10$ Ziehungen und $p = 0.5$. [Köhler et al.]

Größere Stichproben verkleinern β

Beispiel

Blutdruckmedikament soll geprüft werden. Stichprobe besteht aus 160 Personen vor (u_i) und nach (b_i) der Behandlung. Insgesamt 320 Blutdruck-Werte. Man bildet die Differenzen $d_i = b_i - u_i$ und deren Mittelwert $\bar{d}$.

Das Medikament hat eine Wirkung wenn $\bar{d}$ signifikant von null abweicht und nicht zufällig. Mit $\bar{d}$ soll geklärt werden ob der wahre Wert δ gleich oder ungleich null ist.

- $H_0(\delta = 0)$
- $H_1(\delta \neq 0)$

Größere Stichproben verkleinern β (1)

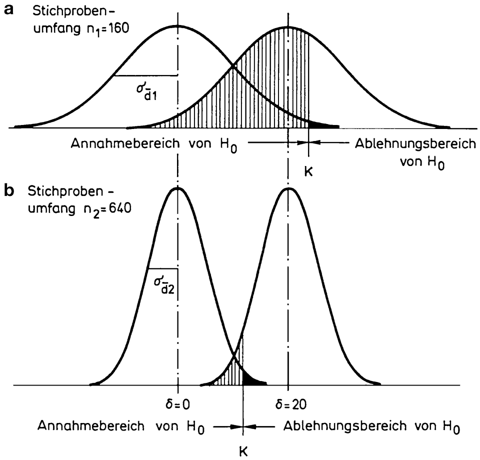


Figure 5: Mögliche Entscheidungen beim Testen.[Köhler et al.]

Um β-Fehler darstellen zu können mussten wir die Hypothese konkretesieren ($\delta = 20$).

		Wahrer Sachverhalt	
		$\delta = 0$	$\delta \neq 0$
Entscheidung des Tests	Annehmen von $H_0(\delta = 0)$	*richtige Entscheidung:* wahrer Sachverhalt stimmt mit Testergebnis überein. Wahrscheinlichkeit: $(1-\alpha)$	*falsche Entscheidung:* $H_1(\delta \neq 0)$ wäre richtig, Testergebnis führt aber zu $H_0(\delta = 0)$. (Fehler 2. Art) Wahrscheinlichkeit: β
	Annehmen von $H_1(\delta \neq 0)$	*falsche Entscheidung:* $H_0(\delta = 0)$ wäre richtig, Testergebnis führt aber zu $H_1(\delta \neq 0)$. (Fehler 1. Art) Wahrscheinlichkeit: α	*richtige Entscheidung:* wahrer Sachverhalt stimmt mit Testergebnis überein. Wahrscheinlichkeit: $(1-\beta)$

Figure 6: Mögliche Entscheidungen beim Testen.[Köhler et al.]

Schätzer

Punktschätzer

Der Parameter (z.B. μ oder σ) wird aus einer Stichprobe berechnet und als Schätzwert für die Grundgesamtheit angegeben.

Intervalschätzer

Man gibt ein ganzes Intervall als mögliche Wert des Schätzers. Man nennt solche Intervalle *Vertrauensbereiche* oder *Konfidenzintervalle*.

Konfidenzintervalle

- Das Konfidenzintervall enhält mit der *vorgegebenen* Wahrscheinlichkeit $(1-\alpha)$ den wahren Parameterwert.
- Je größer $(1-\alpha)$ desto größer das Konfidenzintervall.

Konfidenzintervalle(1)

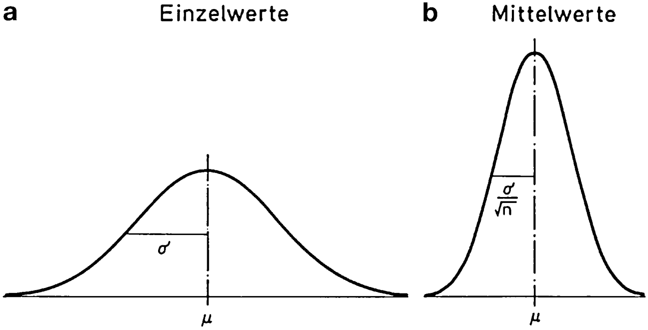


Figure 7: (a) Verteilung der Einzelwerte x der Grundgesamtheit. (b) Verteilung der Mittelwerte $\bar{x}$ von Stichproben des Umfangs n. Die Standardabweichung ist σ. [Köhler et al.]

Konfidenzintervalle für μ bei bekannter Varianz σ

Wenn die Varianz σ bekannt ist, kann man wie folgt das Intervall bestimmen, wo ungefähr 95% der $\bar{x}$-Werte liegen ($\alpha = 0.05$):

$$\mu - 2 \cdot \frac{\sigma}{\sqrt{n}} \lesssim \bar{x} \lesssim \mu + 2 \cdot \frac{\sigma}{\sqrt{n}}$$

$$\bar{x} - 2 \cdot \frac{\sigma}{\sqrt{n}} \lesssim \mu \lesssim \bar{x} + 2 \cdot \frac{\sigma}{\sqrt{n}}$$

Freihetsgrade

Anzahl der Messwerte minus Anzahl der aus den Messwerten ermittelten Parameter

Konfidenzintervalle für μ bei Normalverteilung

Fragestellung: finde das Intervall, dass den wahren Mittelwert μ mit der Sicherheitswahrscheinlichkeit $(1-\alpha)$ enthält.

Voraussetzung: Die Grundgesamtheit is normalverteilt mit *unbekannten* μ und σ.

$$s_x = \sqrt{\frac{1}{1-n}\left(\sum_{i=1}^{n} x_i^2 - \frac{(\sum_{i=1}^{n} x_i)^2}{n}\right)}$$

$$s_{\bar{x}} = \frac{s_x}{\sqrt{n}}$$

$(1-\alpha)$-Konfidenzintervall

$$[\bar{x} - t_{\mathrm{Tab}}(n-1,\alpha)\cdot s_{\bar{x}}; \bar{x} + t_{\mathrm{Tab}}(n-1,\alpha)\cdot s_{\bar{x}}]$$

$t(m,\alpha)$ ist Studentsche t-Verteilung mit m Freiheitsgraden

t-Verteilung

Mehrkmal ist t-verteilt, wenn die Varianz des Merkmals unbekannt ist und mit der Stichprobenvarianz geschätzt werden muss.

$$f_m(x) = \frac{\Gamma\left(\frac{m+1}{2}\right)}{\sqrt{m\pi}\Gamma\left(\frac{m}{2}\right)}\left(1 + \frac{x^2}{m}\right)^{-\frac{m+1}{2}}$$

für $-\infty \leq x \leq \infty$

Beispiel

```
data <- c(341, 345, 338, 339, 340, 343, 341, 343, 341, 328, 343, 347, 337, 348, 339)
t.test(data, conf.level = 0.95)$conf.int
```

```
## [1] 338.2368 343.4965
## attr(,"conf.level")
## [1] 0.95
```

Konfidenzintervalle für die Differenz von μ_x und μ_y bei Normalverteilung

Fragestellung: finde das Intervall, dass den Betrag der Differenz der wahren Mittelwerte μ_x und μ_y der Grundgesamtheiten X uns Y mit der Sicherheitswahrscheinlichkeit $(1 - \alpha)$ enthält.

Voraussetzung: Beide Grundgesamtheiten sind normalverteilt mit gelichen *unbekannten* Varianzen. Die Stichproben sind unabhängig.

$$s_D = \sqrt{\frac{(n_X - 1) \cdot s_x^2 + (n_Y - 1) \cdot s_y^2}{n_X - n_Y - 2}}$$

$$s_{\bar{D}} = S_D \sqrt{\frac{n_X + n_Y}{n_X \cdot n_Y}}$$

Konfidenzintervalle für die Differenz von μ_x und μ_y bei Normalverteilung(1)

$(1-\alpha)$-Konfidenzintervall

$$[|\bar{x}-\bar{y}| - t_{\text{Tab}}(n_X+n_Y-2,\alpha)\cdot s_{\bar{D}}; |\bar{x}-\bar{y}| + t_{\text{Tab}}(n_X+n_Y-2,\alpha)\cdot s_{\bar{D}}]$$

Beispiel

```
data1 <- c(341, 345, 338, 339, 340, 343, 341, 343, 341, 328, 343, 347, 337, 348, 339)
data2 <- data1 + 10 + runif(length(data1))*20
t.test(data2, data1, conf.level = 0.95)$conf.int
```

```
## [1] 15.43699 25.19541
## attr(,"conf.level")
## [1] 0.95
```

Statistische Tests

Intervalldaten

Ordinaldaten

Nominale Daten