Vorlesung Statistik II

Florian Wickelmaier

EBERHARD KARLS
UNIVERSITÄT
TÜBINGEN

Sommersemester 2026

Binomialtest

Annahmen

- $X_1, \ldots, X_n$ unabhängig und identisch Bernoulli-verteilt mit Parameter p

Hypothesen

1. $H_0: \ p = p_0 \qquad H_1: \ p \neq p_0$
2. $H_0: \ p \geq p_0 \qquad H_1: \ p < p_0$
3. $H_0: \ p \leq p_0 \qquad H_1: \ p > p_0$

Teststatistik

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

Exakter Binomialtest

Verteilung der Teststatistik unter H_0

- Binomialverteilt mit Parametern n und p_0 ($B(n, p_0)$-verteilt)

Ablehnungsbereich

1. $T < b_{\alpha/2}$ oder $T > b_{1-\alpha/2}$
2. $T < b_\alpha$
3. $T > b_{1-\alpha}$

p-Quantil b_p der $B(n, p_0)$-Verteilung definiert durch

$$P(T \leq b_p - 1) < p \leq P(T \leq b_p)$$

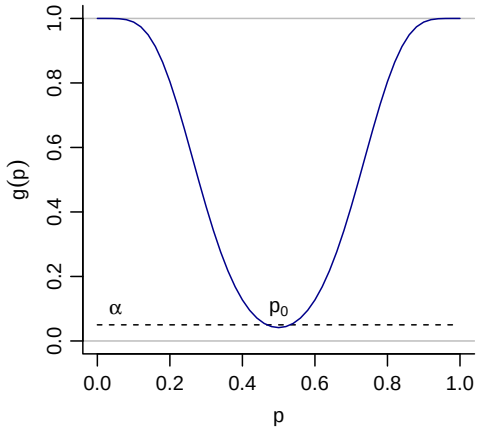
Bemerkung

- Da die Teststatistik nur endlich viele Werte annehmen kann, weicht die tatsächlich verwendete Irrtumswahrscheinlichkeit von der nominellen ab (man sagt "der Test ist konservativ")

Gütefunktion

Für $\alpha = 0.05$ und $n = 20$ erhält man $b_{0.025} = 6$ und $b_{0.975} = 14$ und daraus mit der $B(20, p)$-verteilten Zufallsvariable T für die Nullhypothese $H_0 : \; p = 0.5$

$$g(p) = P\left(T < b_{\alpha/2} \mid p\right) + P\left(T > b_{1-\alpha/2} \mid p\right)$$



Approximativer Binomialtest

Annahmen

- $X_1, \ldots, X_n$ unabhängig und identisch Bernoulli-verteilt mit Parameter p und n hinreichend groß ($n \geq 30$)

Hypothesen

1. $H_0: \; p = p_0 \qquad H_1: \; p \neq p_0$
2. $H_0: \; p \geq p_0 \qquad H_1: \; p < p_0$
3. $H_0: \; p \leq p_0 \qquad H_1: \; p > p_0$

Teststatistik

$$Z = \frac{\bar{X} - p_0}{\sqrt{p_0 \, (1 - p_0)}} \sqrt{n}$$

Approximativer Binomialtest

Verteilung der Teststatistik unter H_0

- Approximativ $N(0,1)$-verteilt

Ablehnungsbereich

1. $|Z| > z_{1-\alpha/2}$
2. $Z < z_\alpha = -z_{1-\alpha}$
3. $Z > z_{1-\alpha}$

mit dem p-Quantil z_p der $N(0,1)$-Verteilung

Bemerkung

- Da die Teststatistik auch in diesem Fall keine stetige Zufallsvariable ist, ist auch dieser Test konservativ

Poweranalyse mittels Computersimulation

Per Computersimulation wird der Anteil der signifikanten Testergebnisse unter Wiederholung des Zufallsexperiments bestimmt

Dafür müssen Stichprobengröße n, minimal relevanter Effekt und Irrtumswahrscheinlichkeit α vorher festgelegt werden

Pseudocode

```
1  Stichprobengröße beliebig setzen
2  replicate
   {
3      Stichprobenwerte aus daten-erzeugendem Modell
           mit minimal relevantem Effekt ziehen
4      Test der Nullhypothese durchführen
   }
5  Anteil der signifikanten Ergebnisse bestimmen
```

Bei der Fallzahlplanung wird die Stichprobengröße solange erhöht, bis dieser Anteil die geplante Power (z. B. 0.8 oder 0.95) erreicht

Exakter Binomialtest in R

2000-fach wiederholte Stichprobenziehung vom Umfang $n = 20$ aus einer $B(20, p = 0.8)$-verteilten Population mit exaktem Binomialtest der Nullhypothese H_0: $p = 0.5$ bei $\alpha = 0.05$

```
n <- 20                                  # Stichprobengroesse
pval <- replicate(2000, {                # Wiederholungen
  x <- rbinom(1, size = n, prob = 0.8)   # Modell mit Effekt
  binom.test(x, n = n, p = 0.5)$p.value  # p-Wert des Tests gegen H0
})
mean(pval < 0.05)                        # simulierte Teststaerke
```

Für $n = 20$ und $p = 0.8$ wird eine Teststärke (Power) gegen die Nullhypothese von ca. 80% erreicht.

Exakter Binomialtest in R

Mit $\alpha = 0.05$, $n = 20$ und der Nullhypothese $H_0: \ p = 0.5$ erhält man für die Teststatistik $T = 14$ bei Eingabe von

```
binom.test(14, n = 20, p = 0.5, conf.level = 0.95)
```

die Ausgabe

```
    Exact binomial test

data:  14 and 20
number of successes = 14, number of trials = 20, p-value = 0.1153
alternative hypothesis: true probability of success is not equal to 0.5
95 percent confidence interval:
  0.4572108 0.8810684
sample estimates:
probability of success
                   0.7
```

Approximativer Binomialtest in R

Für $\alpha = 0.05$, $n = 30$ und der Nullhypothese $H_0 \colon \; p = 0.5$ erhält man für die Teststatistik $T = 21$ bei Eingabe von

```
prop.test(21, n = 30, p = 0.5, conf.level = 0.95, correct = FALSE)
```

die Ausgabe

```
    1-sample proportions test without continuity correction

data:  21 out of 30, null probability 0.5
X-squared = 4.8, df = 1, p-value = 0.02846
alternative hypothesis: true p is not equal to 0.5
95 percent confidence interval:
  0.5212421 0.8333525
sample estimates:
  p
0.7
```

Approximativer Binomialtest in R

Mit denselben Werten, aber mit Stetigkeitskorrektur (siehe Statistik I), erhält man bei Eingabe von

```
prop.test(21, n = 30, p = 0.5, conf.level = 0.95, correct = TRUE)
```

die Ausgabe

```
    1-sample proportions test with continuity correction

data:  21 out of 30, null probability 0.5
X-squared = 4.0333, df = 1, p-value = 0.04461
alternative hypothesis: true p is not equal to 0.5
95 percent confidence interval:
  0.5044209 0.8458720
sample estimates:
  p
0.7
```

Gauß-Test für den Erwartungswert (Ein-Stichproben-z-Test)

Annahmen

- $X_1, \ldots, X_n$ unabhängig und identisch $N(\mu, \sigma^2)$-verteilt mit bekannter Varianz σ^2

Hypothesen

1. $H_0\colon\ \mu = \mu_0 \qquad H_1\colon\ \mu \neq \mu_0$
2. $H_0\colon\ \mu \geq \mu_0 \qquad H_1\colon\ \mu < \mu_0$
3. $H_0\colon\ \mu \leq \mu_0 \qquad H_1\colon\ \mu > \mu_0$

Teststatistik

$$Z = \frac{\bar{X} - \mu_0}{\sigma} \sqrt{n}$$

Gauß-Test für den Erwartungswert

Verteilung der Teststatistik unter $\mu = \mu_0$

- $N(0,1)$-Verteilung

Ablehnungsbereich

1. $|Z| > z_{1-\alpha/2}$
2. $Z < z_\alpha = -z_{1-\alpha}$
3. $Z > z_{1-\alpha}$

mit dem p-Quantil z_p der $N(0,1)$-Verteilung

Bemerkung

- Die Teststatistik ist identisch mit der Pivot-Variable in der Ableitung des entsprechenden Konfidenzintervalls, woraus dessen Identität mit dem Beibehaltungsbereich folgt

Gütefunktion, einseitig

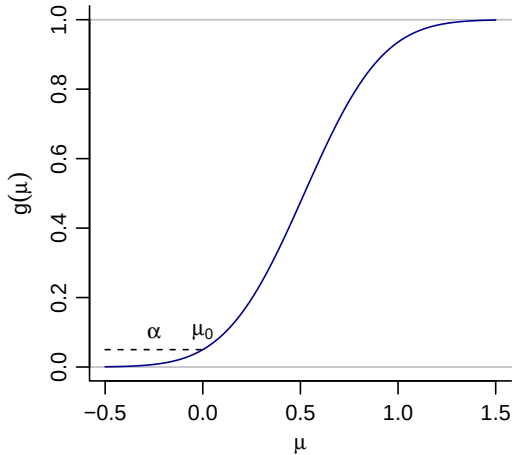
Für das einseitige Testproblem $H_0\colon\ \mu \leq \mu_0$ gilt mit der Definition $g(\mu) = P(H_0 \text{ ablehnen} \mid \mu)$ dann

$$
\begin{aligned}
g(\mu) &= P\left(\frac{\bar{X} - \mu_0}{\sigma}\sqrt{n} > z_{1-\alpha} \mid \mu\right) \\
&= P\left(\frac{\bar{X} - \mu_0 + \mu - \mu}{\sigma}\sqrt{n} > z_{1-\alpha} \mid \mu\right) \\
&= P\left(\frac{\bar{X} - \mu}{\sigma}\sqrt{n} + \frac{\mu - \mu_0}{\sigma}\sqrt{n} > z_{1-\alpha} \mid \mu\right) \\
&= P\left(\frac{\bar{X} - \mu}{\sigma}\sqrt{n} > z_{1-\alpha} - \frac{\mu - \mu_0}{\sigma}\sqrt{n} \mid \mu\right)
\end{aligned}
$$

Gütefunktion, einseitig

Mit der kumulativen Standardnormalverteilung Φ ergibt sich dann

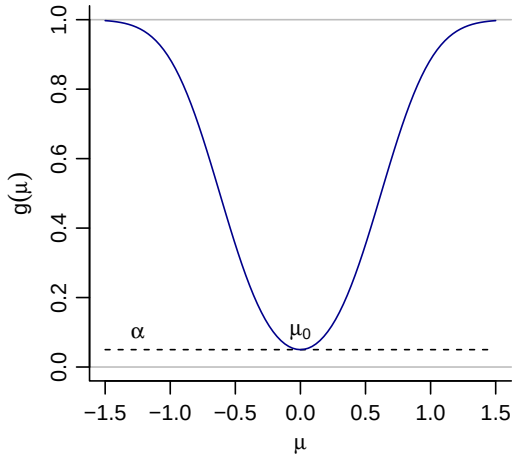
$$g(\mu) = 1 - \Phi\left(z_{1-\alpha} - \frac{\mu - \mu_0}{\sigma}\sqrt{n}\right)$$



Gütefunktion, zweiseitig

Für das zweiseitige Testproblem
$H_0\colon\ \mu = \mu_0$ berechnet man

$$g(\mu) = \Phi\left(-z_{1-\alpha/2} + \frac{\mu - \mu_0}{\sigma}\sqrt{n}\right) + \Phi\left(-z_{1-\alpha/2} - \frac{\mu - \mu_0}{\sigma}\sqrt{n}\right)$$



Approximativer Gauß-Test für den Erwartungswert

Annahmen

- $X_1, \ldots, X_n$ unabhängig und identisch verteilt (beliebige Verteilung) bei $n \geq 30$, Varianz σ^2 bekannt

Hypothesen

1. $H_0\colon\ \mu = \mu_0 \qquad H_1\colon\ \mu \neq \mu_0$
2. $H_0\colon\ \mu \geq \mu_0 \qquad H_1\colon\ \mu < \mu_0$
3. $H_0\colon\ \mu \leq \mu_0 \qquad H_1\colon\ \mu > \mu_0$

Teststatistik

$$Z = \frac{\bar{X} - \mu_0}{\sigma} \sqrt{n}$$

Approximativer Gauß-Test für den Erwartungswert

Verteilung der Teststatistik unter $\mu = \mu_0$

- Approximativ $N(0,1)$-verteilt

Ablehnungsbereich

1. $|Z| > z_{1-\alpha/2}$
2. $Z < z_\alpha = -z_{1-\alpha}$
3. $Z > z_{1-\alpha}$

mit dem p-Quantil z_p der $N(0,1)$-Verteilung

Student-Test für den Erwartungswert (Ein-Stichproben-t-Test)

Annahmen

- $X_1, \ldots, X_n$ unabhängig und identisch $N(\mu, \sigma^2)$-verteilt mit unbekannter Varianz σ^2

Hypothesen

1. $H_0\colon\ \mu = \mu_0 \qquad H_1\colon\ \mu \neq \mu_0$
2. $H_0\colon\ \mu \geq \mu_0 \qquad H_1\colon\ \mu < \mu_0$
3. $H_0\colon\ \mu \leq \mu_0 \qquad H_1\colon\ \mu > \mu_0$

Teststatistik

$$T = \frac{\bar{X} - \mu_0}{S} \sqrt{n}$$

Student-Test für den Erwartungswert

Verteilung der Teststatistik unter $\mu = \mu_0$

- $t(n-1)$-Verteilung

Ablehnungsbereich

1. $|T| > t_{1-\alpha/2}(n-1)$
2. $T < t_\alpha(n-1) = -t_{1-\alpha}(n-1)$
3. $T > t_{1-\alpha}(n-1)$

mit dem p-Quantil $t_p(n-1)$ der $t(n-1)$-Verteilung

Bemerkung

- Die Teststatistik ist identisch mit der Pivot-Variable in der Ableitung des entsprechenden Konfidenzintervalls, woraus dessen Identität mit dem Beibehaltungsbereich des Tests folgt

Beispiel

Wir betrachten wieder das oben beschriebene Experiment zur Tonhöhenwahrnehmung, in dem der Punkt der subjektiven Gleichheit zu einem Sinuston von 1000 Hz mit Hilfe der Herstellungsmethode ermittelt wird

Die Standardabweichung sei nicht bekannt

Unter der Annahme einer unterliegenden Normalverteilung prüfen wir die zweiseitige Fragestellung H_0: $\mu = 1000\,\text{Hz}$ und H_1: $\mu \neq 1000\,\text{Hz}$ bei einer Irrtumswahrscheinlichkeit von $\alpha = 0.05$

Zunächst soll der Stichprobenumfang n bestimmt werden, um mit ausreichender Power eine minimal relevante Abweichung von μ_0 um 3 Hz bei einer Standardabweichung von $\sigma = 8\,\text{Hz}$ zu entdecken

Beispiel: Teststärke und Versuchsplanung

2000-fach wiederholte Stichprobenziehung vom Umfang $n = 30$ aus einer $N(\mu = 997, \sigma^2 = 8^2)$-verteilten Population mit t-Test der Nullhypothese $H_0\colon\ \mu = 1000$

```
n <- 30; sigma <- 8; alpha <- 0.05

pval <- replicate(2000, {
  x <- rnorm(n, mean = 1000 - 3, sd = sigma)  # Modell mit Effekt
  t.test(x, mu = 1000)$p.value                # Test gegen H0
})
mean(pval < alpha)                            # Power
```

Mit $n = 30$ wird für eine Abweichung von μ_0 um $1000\,\text{Hz} - 997\,\text{Hz} = 3\,\text{Hz}$ nur eine Teststärke von ca. 50% erreicht

Auf welchen Umfang muss n erhöht werden, um eine Teststärke von mindestens 80% zu erreichen?

Beispiel

Es sei eine Stichprobe vom Umfang $n = 30$ erhoben worden

Bei der Schätzung der Standardabweichung über die Stichprobenvarianz ergebe sich der Wert $s = 7.5\,\text{Hz}$ und für den Mittelwert der Einstellungen resultiere $\bar{x} = 996\,\text{Hz}$

Als Teststatistik berechnet man

$$t = \frac{\bar{x} - \mu_0}{s}\sqrt{n} = \frac{996 - 1000}{7.5}\sqrt{30} = -2.92$$

Testentscheidung

- Die Nullhypothese H_0 wird mit einer Irrtumswahrscheinlichkeit $\alpha = 0.05$ abgelehnt, wegen
$$|t| > t_{0.975}(29) = 2.05$$
- Wegen der geringen Power ist die Schätzung des Effekts (Abweichung von μ_0) jedoch verzerrt

t-Test in R

Wir nehmen an, dass im oben beschriebenen Experiment zur Tonhöhenwahrnehmung für eine andere Versuchsperson bei $n = 10$ Reizdarbietungen folgende Einstellungen resultieren

$$998, 1003, 1004, 993, 990, 997, 1001, 992, 1000, 995$$

Mit $\alpha = 0.05$ soll unter der Annahme unabhängig und identisch normalverteilter Stichprobenwerte die Nullhypothese $\mu = 1000\,\text{Hz}$ getestet werden

Die Funktion `t.test()` wird dann mit folgender Eingabe aufgerufen

```
x <- c(998, 1003, 1004, 993, 990, 997, 1001, 992, 1000, 995)
t.test(x, mu = 1000, alternative = "two.sided", conf.level = 0.95)
```

t-Test in R

R liefert folgende Ausgabe

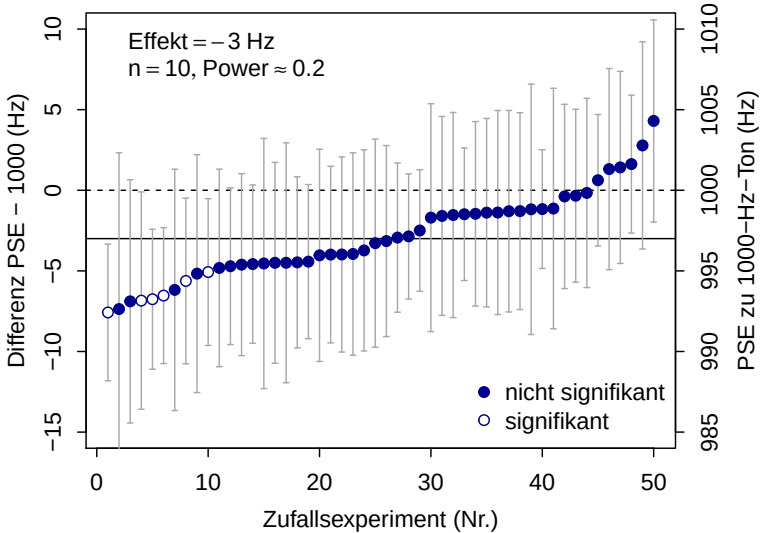
```
        One Sample t-test

data:  x
t = -1.7929, df = 9, p-value = 0.1066
alternative hypothesis: true mean is not equal to 1000
95 percent confidence interval:
  993.8934 1000.7066
sample estimates:
mean of x
    997.3
```

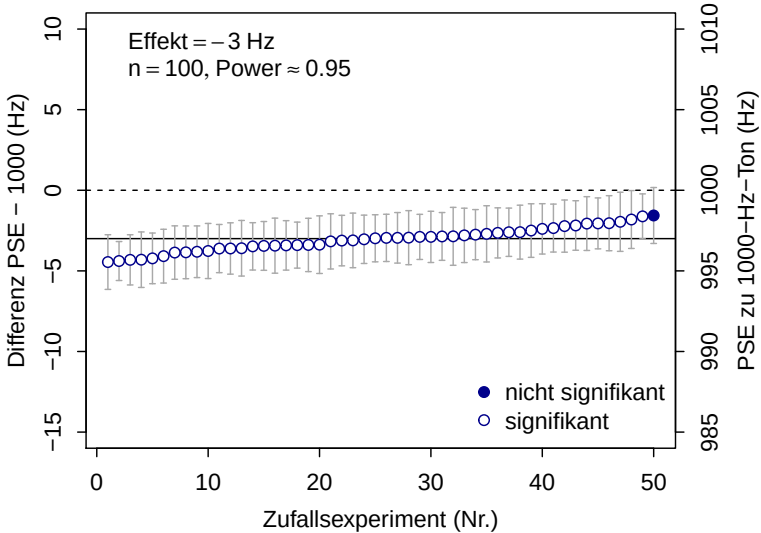
Die Nullhypothese wird nicht abgelehnt, da der resultierende p-Wert *nicht kleiner als* die vorgegebene Irrtumswahrscheinlichkeit $\alpha = 0.05$ ist

Wegen der geringen Power ist das Ergebnis wenig aussagekräftig

Beispiel: Niedrige Power



Beispiel: Hohe Power



Approximativer t-Test für den Erwartungswert

Annahmen

- $X_1, \ldots, X_n$ unabhängig und identisch verteilt (beliebige Verteilung) bei $n \geq 30$, Varianz σ^2 unbekannt

Hypothesen

1. $H_0\colon\ \mu = \mu_0 \qquad H_1\colon\ \mu \neq \mu_0$
2. $H_0\colon\ \mu \geq \mu_0 \qquad H_1\colon\ \mu < \mu_0$
3. $H_0\colon\ \mu \leq \mu_0 \qquad H_1\colon\ \mu > \mu_0$

Teststatistik

$$T = \frac{\bar{X} - \mu_0}{S}\sqrt{n}$$

Approximativer t-Test für den Erwartungswert

Verteilung der Teststatistik unter $\mu = \mu_0$

- Approximativ $t(n-1)$-verteilt

Ablehnungsbereich

1. $|T| > t_{1-\alpha/2}(n-1)$
2. $T < t_\alpha(n-1) = -t_{1-\alpha}(n-1)$
3. $T > t_{1-\alpha}(n-1)$

mit dem p-Quantil $t_p(n-1)$ der $t(n-1)$-Verteilung

Test der Varianz einer Normalverteilung

Annahmen

- $X_1, \ldots, X_n$ unabhängig und identisch $N(\mu, \sigma^2)$-verteilt mit unbekannter Varianz σ^2

Hypothesen

1. $H_0\colon\ \sigma^2 = \sigma_0^2 \qquad H_1\colon\ \sigma^2 \neq \sigma_0^2$
2. $H_0\colon\ \sigma^2 \geq \sigma_0^2 \qquad H_1\colon\ \sigma^2 < \sigma_0^2$
3. $H_0\colon\ \sigma^2 \leq \sigma_0^2 \qquad H_1\colon\ \sigma^2 > \sigma_0^2$

Teststatistik

$$T = \frac{(n-1)\,S^2}{\sigma_0^2}$$

Test der Varianz einer Normalverteilung

Verteilung der Teststatistik unter $\sigma^2 = \sigma_0^2$

- $\chi^2(n-1)$-Verteilung

Ablehnungsbereich

1. $T < \chi^2_{\alpha/2}(n-1)$ oder $T > \chi^2_{1-\alpha/2}(n-1)$
2. $T < \chi^2_{\alpha}(n-1)$
3. $T > \chi^2_{1-\alpha}(n-1)$

mit dem p-Quantil $\chi^2_p(n-1)$ der $\chi^2(n-1)$-Verteilung

Bemerkung

- Die Teststatistik ist wiederum identisch mit der Pivot-Variable in der Ableitung des entsprechenden Konfidenzintervalls, woraus dessen Identität mit dem Beibehaltungsbereich des Tests folgt

Vergleich von Erwartungswerten (z-Test bei zwei unabhängigen Stichproben)

Annahmen

- Unabhängige Stichproben $X_1, \ldots, X_n \sim N(\mu_x, \sigma_x^2)$ und $Y_1, \ldots, Y_m \sim N(\mu_y, \sigma_y^2)$, Varianzen σ_x^2, σ_y^2 bekannt

Folgerungen

- Dann ist

$$\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i \sim N(\mu_x, \sigma_x^2/n)$$

$$\bar{Y} = \frac{1}{m} \sum_{j=1}^{m} Y_j \sim N(\mu_y, \sigma_y^2/m)$$

mit $\bar{X}$ und $\bar{Y}$ stochastisch unabhängig

z-Test bei zwei unabhängigen Stichproben

Folgerungen

- Darüber hinaus gilt

$$E(\bar{X} - \bar{Y}) = E(\bar{X}) - E(\bar{Y}) = \mu_x - \mu_y$$

$$Var(\bar{X} - \bar{Y}) = Var(\bar{X}) + Var(\bar{Y}) = \frac{\sigma_x^2}{n} + \frac{\sigma_y^2}{m}$$

Hypothesen

1. $H_0\colon\ \mu_x - \mu_y = \delta_0 \qquad H_1\colon\ \mu_x - \mu_y \neq \delta_0$
2. $H_0\colon\ \mu_x - \mu_y \geq \delta_0 \qquad H_1\colon\ \mu_x - \mu_y < \delta_0$
3. $H_0\colon\ \mu_x - \mu_y \leq \delta_0 \qquad H_1\colon\ \mu_x - \mu_y > \delta_0$

z-Test bei zwei unabhängigen Stichproben

Teststatistik

$$Z = \frac{\bar{X} - \bar{Y} - E(\bar{X} - \bar{Y})}{\sqrt{Var(\bar{X} - \bar{Y})}} = \frac{\bar{X} - \bar{Y} - \delta_0}{\sqrt{\frac{\sigma_x^2}{n} + \frac{\sigma_y^2}{m}}}$$

Verteilung der Teststatistik unter $\mu_x - \mu_y = \delta_0$

- $N(0,1)$-Verteilung

Ablehnungsbereich

1. $|Z| > z_{1-\alpha/2}$
2. $Z < z_\alpha = -z_{1-\alpha}$
3. $Z > z_{1-\alpha}$

mit dem p-Quantil z_p der $N(0,1)$-Verteilung

Vergleich von Erwartungswerten (t-Test bei zwei unabhängigen Stichproben)

Annahmen

- Unabhängige Stichproben $X_1, \ldots, X_n \sim N(\mu_x, \sigma_x^2)$ und $Y_1, \ldots, Y_m \sim N(\mu_y, \sigma_y^2)$, Varianzen σ_x^2, σ_y^2 unbekannt, aber identisch (d. h. $\sigma_x^2 = \sigma_y^2$)

Folgerungen

- Wegen $\bar{X}$ und $\bar{Y}$ stochastisch unabhängig und $\sigma_x^2 = \sigma_y^2 = \sigma^2$ hat man

$$\begin{aligned} Var(\bar{X} - \bar{Y}) &= Var(\bar{X}) + Var(\bar{Y}) \\ &= \frac{\sigma^2}{n} + \frac{\sigma^2}{m} \\ &= \sigma^2 \left(\frac{1}{n} + \frac{1}{m} \right) \end{aligned}$$

t-Test bei zwei unabhängigen Stichproben

Folgerungen

- Die beste Schätzung von σ^2 erhält man, wenn man beide Stichproben in die Schätzung mit einbezieht (pooled estimate)

$$\hat{\sigma}^2 = \frac{(n-1)\,S_x^2 + (m-1)\,S_y^2}{(n-1)+(m-1)}$$

Hypothesen

1. $H_0\colon\ \mu_x - \mu_y = \delta_0 \qquad H_1\colon\ \mu_x - \mu_y \neq \delta_0$
2. $H_0\colon\ \mu_x - \mu_y \geq \delta_0 \qquad H_1\colon\ \mu_x - \mu_y < \delta_0$
3. $H_0\colon\ \mu_x - \mu_y \leq \delta_0 \qquad H_1\colon\ \mu_x - \mu_y > \delta_0$

t-Test bei zwei unabhängigen Stichproben

Teststatistik

$$T = \frac{\bar{X} - \bar{Y} - E(\bar{X} - \bar{Y})}{\sqrt{\hat{\sigma}^2 \left(\frac{1}{n} + \frac{1}{m}\right)}} = \frac{\bar{X} - \bar{Y} - \delta_0}{\sqrt{\frac{(n-1)\,S_x^2 + (m-1)\,S_y^2}{n+m-2} \cdot \left(\frac{1}{n} + \frac{1}{m}\right)}}$$

Verteilung der Teststatistik unter $\mu_x - \mu_y = \delta_0$

- $t(n+m-2)$-Verteilung

Ablehnungsbereich

1. $|T| > t_{1-\alpha/2}(n+m-2)$
2. $T < t_\alpha(n+m-2) = -t_{1-\alpha}(n+m-2)$
3. $T > t_{1-\alpha}(n+m-2)$

mit dem p-Quantil $t_p(n+m-2)$ der $t(n+m-2)$-Verteilung