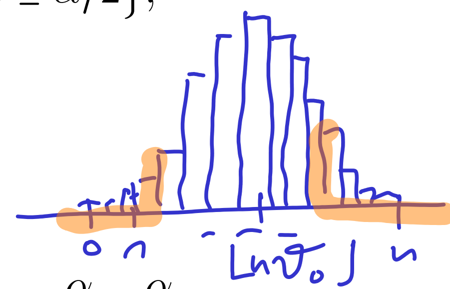
Beispiel 6.22 (Binomialtest)**.** $\Theta = [0,1]$, unter P_ϑ sei die Beobachtung $X \sim \mathrm{Bin}_{n,\vartheta}$ (oder aber Beobachtungen $X_1, \dots, X_n$ sind unter P_ϑ u.i.v. $\sim \mathrm{Ber}_\vartheta$ und wir bilden $X := X_1 + \dots + X_n$). Wähle $\alpha \in (0, 1/2)$.

1. Zweiseitiger Binomialtest: $\Theta_0 = \{\vartheta_0\}$ für ein $\vartheta \in [0,1]$, $\Theta_1 = \Theta \setminus \Theta_0$. Setze

$$c_\ell := \max\left\{x \in \{0,1,2,\dots,n\} : \mathrm{Bin}_{n,\vartheta_0}(\{0,1,\dots,x\}) \le \alpha/2\right\},$$
$$c_r := \min\left\{x \in \{0,1,2,\dots,n\} : \mathrm{Bin}_{n,\vartheta_0}(\{x,x+1,\dots,n\}) \le \alpha/2\right\},$$
$$\varphi(x) := \mathbf{1}_{\{0,1,\dots,c_\ell\}}(x) + \mathbf{1}_{\{c_r,c_r+1,\dots,n\}}(x).$$



Dann gilt

$$\begin{aligned}\mathbb{E}_{\vartheta_0}[\varphi(X)] &= P_{\vartheta_0}(X \le c_\ell) + P_{\vartheta_0}(X \ge c_r)\\ &= \mathrm{Bin}_{n,\vartheta_0}(\{0,1,\dots,c_\ell\}) + \mathrm{Bin}_{n,\vartheta_0}(\{c_r, c_r+1,\dots,n\}) \le \frac{\alpha}{2} + \frac{\alpha}{2} = \alpha\end{aligned}$$

nach Konstruktion, d.h. der Test hält Niveau α ein.
(Wegen der Diskretheit der möglichen Beobachtungen ist das tatsächliche Niveau i.A. etwas kleiner.)

Bei gegebener Beobachtung x ist der p-Wert dann $2\mathrm{Bin}_{n,\vartheta_0}(\{0,1,\dots,x\})$ falls $x < n\vartheta_0$ und $2\mathrm{Bin}_{n,\vartheta_0}(\{x, x+1,\dots,n\})$ falls $x > n\vartheta_0$.

$\Theta = [0,1]$, unter P_ϑ sei die Beobachtung $X \sim \mathrm{Bin}_{n,\vartheta}$ (oder aber Beobachtungen $X_1, \dots, X_n$ sind unter P_ϑ u.i.v. $\sim \mathrm{Ber}_\vartheta$ und wir bilden $X := X_1 + \dots + X_n$). Wähle $\alpha \in (0, 1/2)$.

2. a) Einseitiger Binomialtest (linksseitige Alternative): $\Theta_0 = [\vartheta_0, 1]$ für ein $\vartheta_0 \in (0,1]$, $\Theta_1 = [0, \vartheta_0) = \Theta \setminus \Theta_0$. Setze

$$c := \max\left\{x \in \{0,1,2,\dots,n\} : \mathrm{Bin}_{n,\vartheta_0}(\{0,1,\dots,x\}) \le \alpha\right\},$$
$$\varphi(x) := \mathbf{1}_{\{0,1,\dots,c\}}(x).$$

Nach Konstruktion ist $\mathbb{E}_{\vartheta_0}[\varphi(X)] = P_{\vartheta_0}(X \le c) \le \alpha$ und man kann (leicht) zeigen, dass für $\vartheta > \vartheta_0$ gilt $P_\vartheta(X \le c) \le P_{\vartheta_0}(X \le c)$ $(\le \alpha)$, d.h. der Test hält Niveau α ein.

Bei gegebener Beobachtung x ist der p-Wert dann $\mathrm{Bin}_{n,\vartheta_0}(\{0,1,\dots,x\})$.

sei $X_i \sim \mathrm{Ber}_\vartheta$, $W_i \sim \mathrm{Ber}_{\vartheta_0/\vartheta}$ u.a.

$\tilde{X}_i \le X_i \longrightarrow \tilde{X}_i = X_i \cdot W_i \sim \mathrm{Ber}_{\vartheta_0}$

2. b) Einseitiger Binomialtest (rechtssseitige Alternative): $\Theta_0 = [0, \vartheta_0]$ für ein $\vartheta_0 \in [0,1)$, $\Theta_1 = (\vartheta_0, 1] = \Theta \setminus \Theta_0$. Analog setze

$$C := \overset{\min}{\cancel{\max}}\left\{x \in \{0,1,2,\dots,n\} : \mathrm{Bin}_{n,\vartheta_0}(\{x, x, +1, \dots, n\}) \le \alpha\right\},$$
$$\varphi(x) := \mathbf{1}_{\{C, C+1, \dots, n\}}(x).$$

Der Test hält Niveau α ein, bei gegebener Beobachtung x ist der p-Wert dann $\mathrm{Bin}_{n,\vartheta_0}(\{x, x, +1, \dots, n\})$.

Bemerkung. Offenbar benötigt man die Verteilungsfunktion der Binomialverteilung, um die „kritischen Werte“ c_ℓ, c_r bzw. c, C für den Binomialtest bei vorgegebenem n und α zu bestimmen. Für kleine Werte von n kann man diese „von Hand“ bestimmen, für größere Werte konsultiert man entweder ein Computerprogramm oder eine entsprechende Tabelle oder man verwendet die Normalapproximation der Binomialverteilung: Mit Satz 5.1 und Korollar 5.2 ist

$$\mathrm{Bin}_{n,\vartheta_0}(\{0,1,\ldots,x\}) \approxeq \Phi\left(\frac{x-n\vartheta_0}{\sqrt{n\vartheta_0(1-\vartheta_0)}}\right)$$

wobei $\Phi = F_{\mathcal{N}_{0,1}}$ die Verteilungsfunktion der Standard-Normalverteilung ist.

Beispiel 6.23 (z-Test oder Gauß-Test)**.** $\Theta = \mathbb{R}$, unter P_ϑ seien die Beobachtungen $X_1, \dots, X_n$ u.i.v. $\sim \mathcal{N}_{\vartheta,\sigma^2}$ mit bekanntem, festem $\sigma^2 > 0$. Wähle $\alpha \in (0,1)$.

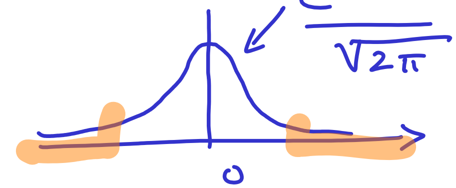
1. Zweiseitiger z-Test: $\Theta_0 = \{\vartheta_0\}$ für ein $\vartheta \in \mathbb{R}$, $\Theta_1 = \Theta \setminus \Theta_0$

$$M := \frac{1}{n}\sum_{i=1}^{n} X_i, \qquad Z := \frac{M - \vartheta_0}{\sqrt{\sigma^2/n}}$$

ist unter P_{ϑ_0} $\sim \mathcal{N}_{0,1}$

mit $q := \Phi^{-1}(1-\alpha/2)$ (das $(1-\alpha/2)$-Quantil der Standardnormalverteilung) ist

$\frac{e^{-z^2/2}}{\sqrt{2\pi}}$



$$\varphi(X_1, \dots, X_n) := \mathbf{1}_{\{|Z|>q\}}$$

ein Test von Θ_0 gegen Θ_1 zum Niveau α (denn unter P_{ϑ_0} ist $M \sim \mathcal{N}_{\vartheta_0,\sigma^2/n}$).
Der p-Wert ist dann $2\big(1 - \Phi(|Z|)\big)$, wobei Φ die Verteilungsfunktion der Standard-Normalverteilung ist.

2. Einseitiger Test: $\Theta_0 = \{\vartheta : \vartheta \le \vartheta_0\}$ für ein $\vartheta \in \mathbb{R}$, $\Theta_1 = \Theta \setminus \Theta_0 = \{\vartheta : \vartheta > \vartheta_0\}$.
Mit $q := \Phi^{-1}(1-\alpha)$ ist

$P_{\vartheta_0}(Z > q) = \alpha$

$$\varphi(X_1, \dots, X_n) := \mathbf{1}_{\{Z>q\}}$$

ein Test von Θ_0 gegen Θ_1 zum Niveau α. Als p-Wert ergibt sich $1 - \Phi(Z)$

für $\vartheta < \vartheta_0$ ist

(Je nach Anwendungssituation kann man auch $\Theta_0 = \{\vartheta : \vartheta \ge \vartheta_0\}$ betrachten, dann ist $\varphi(X_1, \dots, X_n) := \mathbf{1}_{\{Z<-q\}}$ zu wählen und der p-Wert wäre $\Phi(Z) = 1 - \Phi(-Z)$).

unter P_ϑ

$Z \sim \mathcal{N}_{\frac{\vartheta-\vartheta_0}{n},1}$

Somit $P_\vartheta(Z > q) \underset{(=)}{<} \alpha$ für $\vartheta \le \vartheta_0$ <0

Beispiel 6.24 ((ein-Stichproben- oder gepaarter) t-Test). $\Theta = \mathbb{R} \times (0, \infty) \ni \vartheta = (\mu, \sigma^2)$, unter P_ϑ seien die Beobachtungen $X_1, \dots, X_n$ u.i.v. $\sim \mathcal{N}_{\mu,\sigma^2}$ (mit unbekanntem $\mu \in \mathbb{R}$ und unbekanntem $\sigma^2 > 0$). Wähle $\alpha \in (0, 1)$.

Sei $M := \frac{1}{n} \sum_{i=1}^{n} X_i$, $S^2 := \frac{1}{n-1} \sum_{i=1}^{\overset{n}{\cancel{n-1}}} (X_i - M)^2$. (Erinnerung $E_{(\mu,\sigma^2)}[S^2] = \sigma^2$)

1. Zweiseitiger (ein-Stichproben) t-Test: $\Theta_0 = \{\vartheta = (\mu, \sigma^2) \in \Theta : \mu = \mu_0\}$ für ein $\mu_0 \in \mathbb{R}$ (man schreibt dies oft knapp als „$\Theta_0 : \mu = \mu_0$“), $\Theta_1 = \Theta \setminus \Theta_0$.

$$T := \sqrt{n}\frac{M - \mu_0}{\sqrt{S^2}} \qquad \left(= \frac{M - \mu_0}{\sqrt{S^2/n}} \right)$$

Mit $q := q_{n-1,1-\alpha/2} = (1 - \alpha/2)$-Quantil der Student-$(n - 1)$-Verteilung ist

$$\varphi(X_1, \dots, X_n) := \mathbf{1}_{\{|T|>q\}}$$

ein Test von Θ_0 gegen Θ_1 zum Niveau α (denn nach Satz 6.16 ist T für jedes $\vartheta \in \Theta_0$ unter P_ϑ Student-$(n - 1)$-verteilt).
Der p-Wert ist $2\big(1 - F_{T_{n-1}}(|T|)\big)$ mit $F_{T_{n-1}}$ der Verteilungsfunktion der Student-$(n - 1)$-Verteilung.

2. Einseitiger Test: $\Theta_0 = \{\vartheta = (\mu, \sigma^2) \in \Theta : \mu \le \mu_0\}$ für ein $\mu_0 \in \mathbb{R}$ (oft knapp geschrieben als „$\Theta_0 : \mu \le \mu_0$“), $\Theta_1 = \Theta \setminus \Theta_0$. Mit $q := q_{n-1,1-\alpha} = (1 - \alpha)$-Quantil der Student-$(n - 1)$-Verteilung ist

$$\varphi(X_1, \dots, X_n) := \mathbf{1}_{\{T>q\}}$$

ein Test von Θ_0 gegen Θ_1 zum Niveau α (und analog $\varphi(X_1, \dots, X_n) := \mathbf{1}_{\{T<-q\}}$ ein Test für $\Theta_0 = \{\mu \ge \mu_0\}$, beachte auch: $-q$ ist das α-Quantil der Student-$(n - 1)$-Verteilung).
Der p-Wert ist $1 - F_{T_{n-1}}(T)$.

(Je nach Anwendungssituation kann man auch $\Theta_0 = \{\vartheta = (\mu, \sigma^2) \in \Theta : \mu \ge \mu_0\}$ betrachten, dann ist $\varphi(X_1, \dots, X_n) := \mathbf{1}_{\{T<-q\}}$ zu wählen und der p-Wert wäre $F_{T_{n-1}}(T) = 1 - F_{T_{n-1}}(-T)$).

Anwendungsbeispiel. a) Die Wirksamkeit eines gewissen Schlafmittels soll geprüft werden. 10 Patienten erhalten das Schlafmittel, die Anzahl zusätzlicher Stunden Schlaf wird in einer Nacht beobachtet.

Wir nehmen an, die Beob. sind u.i.v. $\sim \mathcal{N}_{\mu,\sigma^2}$ und wir möchten die Nullhypothese $\mu = 0$, sagen wir, zum Niveau $\alpha = 0{,}05$ testen.

Die Daten[4]:

Patient i	1	2	3	4	5	6	7	8	9	10
zus. Schl.	0.7	-1.6	-0.2	-1.2	-0.1	3.4	3.7	0.8	0.0	2.0

Es ist $n = 10$, $\overline{x} = \frac{1}{10}\sum_{i=1}^{10} x_i = 0{,}75$, $s^2 = \frac{1}{9}\sum_{i=1}^{10}(x_i - \overline{x})^2 \approxeq 1{,}79$, $t = \frac{\overline{x}-0}{s/\sqrt{10}} \approxeq 1{,}326$

Das $0{,}975$-Quantil der Student-9-Verteilung ist $\approxeq 2{,}262$, demnach können wir die Nullhypothese nicht ablehnen. (Für ein Student-9-verteiltes T ist $P(|T| \geq 1{,}326) \approxeq 0{,}2176$, dies ist der p-Wert des Tests.)

Man kann diesen Befund folgendermaßen formulieren:

„Die Beobachtungen sind mit der Nullhypothese $\mu = 0$ (im statistischen Sinne) verträglich.“

oder

„Die beobachtete Abweichung $\overline{x} = 0{,}75$ ist nicht signifikant von 0 verschieden (t-Test, $\alpha = 0{,}05$).“

[4] Aus Student (William S. Gosset), The Probable Error of a Mean, Biometrika 6:1–25 (1908)

b) Die Wirksamkeit eines Schlafmittels soll mit der eines anderen verglichen werden. 10 Patienten erhalten Schlafmittel A, die Anzahl zusätzlicher Stunden Schlaf wird in einer Nacht beobachtet. Dann erhalten dieselben 10 Patienten Schlafmittel B, wieder wird die Anzahl zusätzlicher Stunden Schlaf in einer Nacht beobachtet.

Da dieselben Patienten untersucht werden, können (und sollten) wir die Messungen paaren: Wir interessieren uns bei jedem Patienten für die Differenz des (zusätzlichen) Schlafs bei Mittel 2 und bei Mittel 1.

Wir nehmen an, die beobachteten Differenzen sind Realisierungen von u.i.v. ZVn mit Vert. $\mathcal{N}_{\mu,\sigma^2}$ und wir möchten die Nullhypothese $\mu \leq 0$ gegen die Alternative $\mu > 0$, sagen wir, zum Niveau $\alpha = 0{,}05$ testen.
(Dies wäre beispielsweise in folgender Situation angemessen: Wir möchten darlegen, dass Mittel B wirksamer ist als Mittel A, indem wir die Nullhypothese „$\mu \leq 0$“ entkräften.)

Die Daten (wiederum aus Student, a.a.O.) :

Patient i	1	2	3	4	5	6	7	8	9	10
Mittel A	0.7	-1.6	-0.2	-1.2	-0.1	3.4	3.7	0.8	0.0	2.0
Mittel B	1.9	0.8	1.1	0.1	-0.1	4.4	5.5	1.6	4.6	3.4
Diff.	1.2	2.4	1.3	1.3	0.0	1.0	1.8	0.8	4.6	1.4

Es ist $n = 10$, $\overline{x} = \frac{1}{10}\sum_{i=1}^{10} x_i = 1{,}58$, $s^2 = \frac{1}{9}\sum_{i=1}^{10}(x_i - \overline{x})^2 \approxeq 1{,}23$, $t = \frac{\overline{x}-0}{s/\sqrt{10}} \approxeq 4{,}062$
Das $0{,}95$-Quantil der Student-9-Verteilung ist $\approxeq 1{,}833$, demnach können wir die Nullhypothese ablehnen.
(Für ein Student-9-verteiltes T ist $P(T > 4{,}062) \approxeq 0{,}0014$, dies ist der p-Wert des Tests.)

Mögliche knappe Formulierung dieses Befunds:

„Die beobachtete Differenz $\overline{x} = 1{,}58$ ist signifikant größer als 0 (einseitiger t-Test, $\alpha = 0{,}05$).“

Beispiel 6.25 (Test für die Varianz im normalen Modell)**.** In der Situation von Beispiel 6.24 sei $\Theta_0 = \{\vartheta = (\mu, \sigma^2) \in \Theta : \sigma^2 \leq v_0\}$ für ein $v_0 > 0$, $\Theta_1 = \Theta \setminus \Theta_0$. Wähle $\alpha \in (0, 1)$.

Mit $q :=$ $(1 - \alpha)$-Quantil der χ^2_{n-1}-Verteilung ist

$$\varphi(X_1, \ldots, X_n) := \mathbf{1}_{\{S^2 > qv_0/(n-1)\}}$$

ein Test von Θ_0 gegen Θ_1 zum Niveau α (vgl. Satz 6.16).

Beispiel 6.26 (zwei-Stichproben oder ungepaarter t-Test [mit Annahme gleicher Varianzen]). $\Theta = \mathbb{R}\times\mathbb{R}\times(0,\infty) \ni \vartheta = (\mu_1, \mu_2, \sigma^2)$, unter P_ϑ sind $X_1, \dots, X_m$ u.i.v. und davon unabhängig $Y_1, \dots, Y_n$ u.i.v. ($m, n \in \mathbb{N}$), $X_i \sim \mathcal{N}_{\mu_1,\sigma^2}$, $Y_j \sim \mathcal{N}_{\mu_2,\sigma^2}$. Seien

$$M_X := \frac{1}{m}\sum_{i=1}^m X_i, \quad M_Y := \frac{1}{n}\sum_{j=1}^n Y_j$$

die jeweiligen Stichprobenmittelwerte,

$$S_X^2 = \frac{1}{m-1}\sum_{i=1}^m (X_i - M_X)^2, \quad S_Y^2 = \frac{1}{n-1}\sum_{j=1}^n (Y_j - M_Y)^2,$$

die (korrigierten) Stichprobenvarianzen,

$$S^2 := \frac{(m-1)S_X^2 + (n-1)S_Y^2}{m+n-2} \quad \Big(= \frac{1}{m+n-2}\big(\sum_{i=1}^{\cancel{n}\ m} (X_i - M_X)^2 + \sum_{j=1}^n (Y_j - M_Y)^2\big)\Big),$$

(die „gepoolte Stichprobenvarianz"),

$$T = \frac{M_X - M_Y}{S\sqrt{\frac{1}{m} + \frac{1}{n}}}.$$

$\mathrm{Var}[M_X - M_Y]$
$= \mathrm{Var}[M_X] + \mathrm{Var}[M_Y] = \frac{\sigma^2}{m} + \frac{\sigma^2}{n}$

(Beachte: Stets gilt $\mathbb{E}_{(\mu_1,\mu_2,\sigma^2)}[S^2] = \sigma^2$ [S^2 ist ein erwartungstreuer Schätzer für σ] und T ist unter $\cancel{P_{(\mu_1,\mu_2,\sigma^2)}}$ $P_{(\mu_1,\mu_1,\sigma^2)}$ Student-$(n+m-2)$-verteilt, Argument analog zum Beweis von Satz 6.16).

1. Zweiseitiger ungepaarter t-Test : $\Theta_0 = \{(\mu_1, \mu_2, \sigma^2) \in \Theta : \mu_1 = \mu_2\}$ (oft knapp geschrieben als „$\Theta_0 : \mu_1 = \mu_2$"), $\Theta_1 = \Theta \setminus \Theta_0$.

Wähle $\alpha \in (0,1)$, mit $q := q_{m+n-2,1-\alpha/2} = (1-\alpha/2)$-Quantil der Student-$(m+n-2)$-Verteilung ist

$$\varphi(X_1, \dots, X_m, Y_1, \dots, Y_n) := \mathbf{1}_{\{|T|>q\}}$$

ein Test von Θ_0 gegen Θ_1 zum Niveau α.

1. Einseitiger Test : $\Theta_0 = \{(\mu_1, \mu_2, \sigma^2) \in \Theta : \mu_1 \leq \mu_2\}$ (oft knapp geschrieben als „$\Theta_0 : \mu_1 \leq \mu_2$“), $\Theta_1 = \Theta \smallsetminus \Theta_0$. Mit $q := q_{m+n-2,1-\alpha} = (1-\alpha)$-Quantil der Student-$(m+n-2)$-Verteilung ist

$$\varphi(X_1, \dots, X_m, Y_1, \dots, Y_n) := \mathbf{1}_{\{T>q\}}$$

ein Test von Θ_0 gegen Θ_1 zum Niveau α.
(Analog ist $\varphi(X_1, \dots, X_m, Y_1, \dots, Y_n) := \mathbf{1}_{\{T<-q\}}$ ein Test von $\Theta_0 = \{(\mu_1, \mu_2, \sigma^2) \in \Theta : \mu_1 \geq \mu_2\}$.)

(p-Werte werden analog zum ein-Stichproben-Fall (Bsp. 6.24) berechent, wobei $F_{T_{n-1}}$ durch $F_{T_{m+n-2}}$ ersetzt wird.)