



Frage: Ist das neue Medikament besser?

Man muss entscheiden zwischen

- der Nullhypothese  $H_0: \mu \in \Theta_0 := [0, \mu_{alt}]$   
(neues Medikament nicht besser)
- der Alternative  $H_1: \mu \in \Theta_1 := [\mu_{alt} + \epsilon, \infty]$

Ein Test ist eine Entscheidungsregel, ob man aufgrund der Daten  $H_0$  oder  $H_1$  glauben sollte. Dabei kann man zwei mögliche Fehler machen:

- Fehler 1. Art: Man verwirft Hypothese  $H_0$ , obwohl sie richtig ist ( $\mu \in \Theta_0$ ).
- Fehler 2. Art: Man nimmt die Hypothese  $H_0$  an, obwohl sie falsch ist ( $\mu \notin \Theta_0$ ).

10/11

Man kann den Fehler 1. Art radikal minimieren, indem man  $H_0$  immer annimmt. Dann würde er aber  $H_1$  auch immer ablehnen, insbesondere auch dann, wenn es richtig ist. Analog kann man den Fehler 2. Art radikal minimieren, indem man  $H_1$  immer annimmt also  $H_0$  stets verwirft. Beides ist unlos! Die beiden Fehler sind also komplementär zueinander, in dem Sinne, dass die Minimierung der W'keit des eine zu einer Anstieg der W'keit des andere führt. ~~Alle  $H_0$  an~~

Dabei führt man eine Rangordnung ein. Als der gravierendere Fehler (persönliche Fehler) wird das <sup>intentionale</sup> Verwerfen der Nullhypothese festgelegt. Ein intentionales Festhalten an der Nullhypothese wird als weniger gravierend eingestuft. ("vernünftiger")

↳ Best. Zuordnung von Null- und Alternativhypothese dient der Festlegung welcher der beiden Testfehler als prioritär bzw. vernünftiger anzusehen wird.

Ziel: Finde einen Test, dessen W'keit für Fehler 1. Art unter einem geg. Irrtumsniveau  $\alpha \in (0,1)$  liegt.

Durchführung eines stat. Entscheidungsproblems:

1) Formulierung des stat. Modells  $(\mathcal{X}, \mathcal{F}, (\mathbb{P}_\theta, \theta \in \Theta))$

2) Formulierung der Nullhypothese und Alternative

↳ Seien  $\Theta_0, \Theta_1 \subset \Theta$  mit  $\Theta_0 \cap \Theta_1 = \emptyset$ .

$H_0: \theta \in \Theta_0$

$H_1: \theta \in \Theta_1$

3) Wahl des Irrtumsniveaus: Wähle  $\alpha$  und fordere, dass Fehler 1. Art höchstens W'keit  $\alpha$  haben (~~ist~~ <sup>typisch ist</sup>  $\alpha \in \{5\%, 1\%, 0,5\%\}$ )

4) Wahl der Entscheidungsregel:

Wähle eine Statistik  $\phi: \mathcal{X} \rightarrow \{0,1\}$  wie folgt:

$\phi = 0 \iff$  Nehme  $H_0$  an

$\phi = 1 \iff$  Verwerfe  $H_0$  und nehme  $H_1$  an

$\phi \in (0,1) \iff$  Nehme  $H_0$  mit W'keit  $1-\phi$  an.  
Andersten verwerfe  $H_0$  und nehme  $H_1$  an.

5) Durchführung des Experimentes.

Formal:

Def <sup>P.2</sup> Sei  $(\mathcal{X}, \mathcal{F}, (\mathbb{P}_\theta, \theta \in \Theta))$  ein stat. Modell und  $\Theta_0, \Theta_1 \subset \Theta$  mit  $\Theta_0 \cap \Theta_1 = \emptyset$  definiere Nullhypothese und Alternative (wie in Punkt 2) oben).

i) Eine Statistik  $\phi: \mathcal{X} \rightarrow \{0,1\}$  heißt Test von  $\Theta_0$  gg  $\Theta_1$

ii) Ein Test  $\phi$  heißt nicht randomisiert, falls  $\phi \in \{0,1\}$   $\mathbb{P}_\theta$ -f.s. für alle  $\theta \in \Theta$ . In diesem Fall heißt

$K = \{ \omega : \phi(\omega) = 1 \}$  Ablehnungsbereich oder Kritischer Bereich

iii) Falls  $\phi \in (0,1)$  mit pos. ~~W'keit~~  $P_{\theta_0}$ -W'keit für kein ein  $\theta \in \Theta$ , so heißt  $\phi$  randomisiert.

iv) Das effektive Niveau von  $\phi$  ist def. durch  $\sup_{\theta \in \Theta} P_{\theta}^{\phi}(\text{Ablehnen}) = \sup_{\theta \in \Theta} E_{\theta}[\phi] = \sup_{\theta \in \Theta} P_{\theta}^{\phi}$

+ 4) vorige Seite

Das entspricht der größten W'keit für einen Fehler 1. Art.

v) Ein Test  $\phi$  hat (Interims-) Niveau  $\alpha \in (0,1)$ , wenn  $\sup_{\theta \in \Theta_0} E_{\theta}[\phi] \leq \alpha$ .

vi) Die Güte funktion ist definiert durch

$$G_{\phi} : \Theta \rightarrow [0, 1], \theta \mapsto E_{\theta}[\phi]$$

vii) Die Macht von  $\phi$  bei  $\theta$  ist  $G_{\phi}(\theta)$ .

Für  $\theta \in \Theta_1$  ist  $\beta_{\phi}(\theta) = 1 - G_{\phi}(\theta)$  die W'keit für einen Fehler 2. Art.

Fehlerrate

Bsp 81 Medikament. Die beobachtete Erfolgsrate ist  $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$

Wir möchten die Hypothese  $\theta \leq \theta_0$  ablehnen, falls  $\bar{x}$  deutlich über  $\theta_0$  liegt.

Prüfstatistik  $\phi = I_K$  (nicht-randomisiert)

mit Ablehnbereich

$$K = \left\{ x : \frac{1}{n} \sum_{i=1}^n x_i \geq c \right\}$$

Wähle dann  $c$  groß genug, so dass

$$\sup_{\theta \leq \theta_0} E_{\theta}[\phi] = \sup_{\theta \leq \theta_0} P_{\theta}[\bar{x} \geq c]$$

$$= P_{\theta_0}[\bar{x} \geq c] \leq \alpha$$

( $\hookrightarrow c$  kann über Quantile festgelegt werden)

Die Form  $\phi = I_K$  mit  $K = \{x : T(x) \geq c\}$  ist typisch. Man nennt die entsprechende Statistik  $T$  eine Teststatistik.

102 Alternativ zum Niveau kann man auch den P-Wert (oder Signifikanzniveau) eines Tests festlegen. Anschaulich ist der P-Wert das kleinste Niveau, bei dem man aufgrund der Daten  $X$  die Hypothese  $H_0$  noch ablehnen würde.  
 Formel: Sei  $\phi_\alpha = \mathbb{1}_{K_\alpha}$  Test zum Niveau  $\alpha$ .

P-Wert.  $W = \inf \{ \alpha \in [0,1], X \in K_\alpha \}$ .

8.3 Bem ~~8.3~~ Ziel: zerlegen Tests und Konfidenzbereiche

Sei  $\Theta_0 = \{ \theta_0 \}$  d.h.  $H_0: \theta = \theta_0$  einfache Hypothese und sei  $C(X)$  ein Konfidenzbereich zum Niveau  $1-\alpha$ , d.h.

$P_{\theta_0} [C(X) \ni \theta] \geq 1-\alpha$  für alle  $\theta$ .

Wähle  $\phi = \mathbb{1}_K$  mit  $K = \{ C(X) \not\ni \theta_0 \}$

$\Rightarrow P_{\theta_0} [\phi] = P_{\theta_0} [C(X) \not\ni \theta_0] = 1 - P_{\theta_0} [C(X) \ni \theta_0] \leq \alpha$   
 $\hookrightarrow$  also ist  $\phi$  Test zum Niveau  $\alpha$  ( $\Theta_0 = \{ \theta_0 \}$ )

Bsp 8.4: Seien  $X_1, \dots, X_n$  i.i.d.  $\mathcal{N}(\theta, \sigma^2)$ -verteilt mit  $\sigma^2$  bekannt.

Nach ~~Bsp 7.3~~ ist  $C(X) = \left[ \bar{X} - \overset{\text{Sidewerts}}{c_\alpha \sigma}, \bar{X} + \overset{\text{quantil}}{c_\alpha \sigma} \right]$ ,  $c_\alpha = \frac{1}{\sigma n} \Phi_{1-\frac{\alpha}{2}}(\mathcal{N}(0,1))$   
 ein Konfidenzbereich zum Niveau  $1-\alpha$ .

Wir möchten die Hypothese  $H_0: \theta = \theta_0$  auf dem Niveau  $\alpha$  testen. Nach ~~Bem 8.3~~ ist dann

$\phi = \mathbb{1}_K$  mit  $K = \{ \overset{X:}{\text{}} C(X) \not\ni \theta_0 \} = \{ \overset{X:}{\text{}} \theta_0 \notin \left[ \bar{X} - \overset{\text{quantil}}{c_\alpha \sigma}, \bar{X} + \overset{\text{quantil}}{c_\alpha \sigma} \right] \}$   
 $= \{ X: \left| \frac{\bar{X} - \theta_0}{\sigma} \right| > c_\alpha \} = \{ X: |\bar{X} - \theta_0| > \alpha \}$

Test zum Niveau  $\alpha$

### 8.1 Neyman-Pearson-Tests

Anforderungen an einen guten Test:

- 1)  $G_\phi(\theta) = \mathbb{P}_\theta[\phi] \geq \alpha \quad \forall \theta \in \Theta_0$   
↳ W'keit für Fehler 1. Art kleiner als  $\alpha$
- 2)  $G_\phi(\theta)$  soll möglichst groß sein für alle  $\theta \in \Theta_1$   
↳ W'keit für Fehler 2. Art wird klein.

Def 8.5 Ein Test  $\phi$  von  $\Theta_0$  gegen  $\Theta_1$  heißt UMP (uniformly most powerful) zum Niveau  $\alpha$ , falls

- i)  $\phi$  hat Niveau  $\alpha$
- ii) Für alle Tests  $\psi$  zum Niveau  $\alpha$  gilt  $G_\phi(\theta) \geq G_\psi(\theta)$  für alle  $\theta \in \Theta_1$

Das bedeutet, dass bei gleichen Schranken für Fehler 1. Art  $\phi$  den kleinsten Fehler 2. Art hat.

Def 8.6 Sei  $\Theta_0 = \{\theta_0\}$  und  $\Theta_1 = \{\theta_1\}$ ,  $L_{\theta_1}(x)$  die Likelihood.

i)  $g(x) = \frac{L_{\theta_1}(x)}{L_{\theta_0}(x)}$  heißt Likelihood-Quotient

ii) Ein Test von der Form

$$\phi = \begin{cases} 1 & \text{falls } g(x) \geq c \\ 0 & \text{falls } g(x) < c \end{cases}$$

keine Aussage über

heißt Neyman-Pearson-Test (NP-Test) oder Likelihood-Quotienten-Test (LQ-Test) zum Schwellenwert  $c$ .

Satz 8.7 (Neyman-Pearson-Lemma (1932))

Sei  $(X, \mathcal{G}, (P_\theta, \mathbb{P}_\theta))$  ein stat. Modell auch mit Likelihood-  
fkt  $L_{\theta_1, x}$  ~~und~~ eines Referenzmaßes  $\nu$  (man kann z.B. nehmen  $\nu = \frac{1}{2} P_0 + \frac{1}{2} P_1$  wählen.)

104

Sei  $\mathcal{P}$  der Likelihood-Quotient und  $\alpha \in (0, 1)$  ein vorgegebenes Niveau. Dann gilt:

- i) Es gibt einen NP-Test mit  $\mathbb{P}_{H_0}[\phi] = \alpha$   
(insbesondere hat  $\phi$  Niveau  $\alpha$ )
- ii) Jeder NP-Test  $\phi$  mit  $\mathbb{P}_{H_0}[\phi] = \alpha$  ist UMP  
zum Niveau  $\alpha$ .

Beweis: i) Sei  $\mu$  die Vflg von  $\mathcal{P}$  unter  $\mathbb{P}_{H_0}$ , d. h.  $\mu$  ist ein  $\mathbb{W}$ -Maß auf  $\mathbb{R}_+$ . Sei  $c$  ein  $(1-\alpha)$ -Quantil von  $\mu$ . Dann gilt

$$\mu((c, \infty)) = \mathbb{P}_{H_0}[\mathcal{P}(X) \geq c] \geq \alpha$$

$$\text{und } \mu((-\infty, c)) = \mathbb{P}_{H_0}[\mathcal{P}(X) < c] \\ = 1 - \underbrace{\mu((-\infty, c])}_{\geq 1-\alpha} \leq \alpha$$

( $c$  existiert, da  $\mathbb{P}_{H_0}(\mathcal{P}(X) = \infty) = 0$ )

Es folgt, dass

$$\mathbb{P}_{H_0}[\mathcal{P}(X) = c] = \mathbb{P}_{H_0}[\mathcal{P}(X) \geq c] - \mathbb{P}_{H_0}[\mathcal{P}(X) > c] \\ \geq \alpha - \mathbb{P}_{H_0}[\mathcal{P}(X) > c]$$

Fall 1:  $\mathbb{P}_{H_0}[\mathcal{P}(X) = c] = 0$ . Dann ist

$\phi = \mathbb{I}_{\{\mathcal{P}(X) > c\}}$  ein NP-Test mit

$$\mathbb{P}_{H_0}[\phi] = \mathbb{P}_{H_0}[\mathcal{P}(X) > c] = \alpha$$

Fall 2:  $\mathbb{P}_{H_0}[\mathcal{P}(X) = c] > 0$ . Setze

$$\gamma = \frac{\alpha - \mathbb{P}_{H_0}[\mathcal{P}(X) > c]}{\mathbb{P}_{H_0}[\mathcal{P}(X) = c]} \in (0, 1]$$

Dann ist

$$\phi = \begin{cases} 1 & \mathcal{P}(X) > c \\ \gamma & \mathcal{P}(X) = c \\ 0 & \mathcal{P}(X) < c \end{cases}$$

ein NP-Test zum Niveau  $\alpha$ .

$$E_{\nu_0}[\phi] = 1 \cdot P_{\nu_0}[S > c] + \gamma \cdot P_{\nu_0}[S = c] = \alpha$$

ii) Sei  $\phi$  ein NP-Fest mit  $E_{\nu_0}[\phi] = \alpha$  und Schwellenwert  $c$  und sei  $\psi$  ein Fest zum Niveau  $\alpha$

Zu zeigen:  $E_{\nu_0}[\phi] - E_{\nu_0}[\psi] = \int (\phi(x) - \psi(x)) L_{\nu_0}(x) dx \geq 0$

Es gilt falls  $\psi(x) < \phi(x) \Rightarrow \phi(x) > 0 \Rightarrow P(x) \geq c \Rightarrow L_{\nu_0}(x) \geq c L_{\nu_0}(x)$

und falls  $\psi(x) > \phi(x) \Rightarrow \psi(x) < 1 \Rightarrow P(x) \leq c \Rightarrow L_{\nu_0}(x) \leq c L_{\nu_0}(x)$

$\hookrightarrow (\phi(x) - \psi(x)) L_{\nu_0}(x) \geq (\phi(x) - \psi(x)) c L_{\nu_0}(x) \forall x$

$\Rightarrow E_{\nu_0}[\phi] - E_{\nu_0}[\psi] \geq c (E_{\nu_0}[\phi] - E_{\nu_0}[\psi]) \geq 0$

$\Rightarrow \phi$  ist U-NP zum Niveau  $\alpha$  □

Bem 8.8. Der Beweis von Satz 4.3 liefert, dass das NP-Fest

$$\phi(x) = \begin{cases} 1 & \text{falls } P(x) > c \\ \gamma & \text{--- } P(x) = c \\ 0 & \text{--- } P(x) < c \end{cases} \quad \mu \text{ (Volatilität)}$$

für  $c \in (1-\alpha)$ -Quantil von  $P$  und

$$\gamma = \begin{cases} \frac{\alpha - P_{\nu_0}[S > c]}{P_{\nu_0}[S = c]} & \text{falls } P_{\nu_0}[S = c] > 0 \\ 0 & \text{sonst} \end{cases}$$

U-NP ist.

Prop 8.5 (Eindeutigkeit)

In der Situation aus Satz 8.7 mit  $\phi$  wie in Satz 8.7 sei  $\psi$  irgendein U-NP Test für  $\{\nu_0\}$  gegen  $\{\nu_1\}$  zum Niveau  $\alpha = E_{\nu_0}[\phi]$ . Dann gilt  $\phi(x) = \psi(x)$  für a. f. o.  $x$  außer auf der Menge  $\{x: P(x) = c\}$ .

Beweis: Nach dem Beweis von Satz 8.7 ii) (Ungl. (a)) gilt

$$E_{\mathcal{H}_1}[\phi] - c E_{\mathcal{H}_0}[\phi] \leq E_{\mathcal{H}_1}[\phi] - c E_{\mathcal{H}_0}[\phi]$$

Reduzierte gilt  $E_{\mathcal{H}_0}[\phi] = c = E_{\mathcal{H}_0}[\phi]$  und

$$E_{\mathcal{H}_1}[\phi] \geq E_{\mathcal{H}_1}[\phi], \text{ da } \phi \text{ UAI-P ist.}$$

Da  $c \geq 0$  folgt somit:

$$E_{\mathcal{H}_1}[\phi] - c E_{\mathcal{H}_0}[\phi] \geq E_{\mathcal{H}_1}[\phi] - c E_{\mathcal{H}_0}[\phi]$$

Womit Gleichheit gilt. Also

$$(*) (*) \int (\psi(x) - \phi(x)) (E_{\mathcal{H}_1}(x) - c E_{\mathcal{H}_0}(x)) d\mu(x) = 0$$

Nach dem Beweis von Satz 8.7 wissen wir

$$(\psi(x) - \phi(x)) (E_{\mathcal{H}_1}(x) - c E_{\mathcal{H}_0}(x)) \leq 0 \quad \forall x$$

Wahlg (\* \*) folgt daher

$$(\psi(x) - \phi(x)) (E_{\mathcal{H}_1}(x) - c E_{\mathcal{H}_0}(x)) = 0 \quad \text{für } \psi \neq \phi \text{ o. } x$$

$$\Rightarrow \psi(x) = \phi(x) \quad \text{für } \mu\text{-f.a. } x \text{ außer auf}$$

$$\{E_{\mathcal{H}_1}(x) = c E_{\mathcal{H}_0}(x)\} = \{S(x) = c\} \text{ a.s.}$$

Multinomialer Bsp. Teste

$$H_0: \theta_1 = \theta_0 = 0,5 \quad \text{gegen } H_1: \theta_1 = \theta_1 = 0,6$$

$$S_{\mathcal{H}_1}(x) = \theta_1^{\sum_{i=1}^n x_i} (1 - \theta_1)^{n - \sum_{i=1}^n x_i} \quad \text{und somit}$$

ist der Likelihood-Quotient

$$S(x) = \frac{S_{\mathcal{H}_1}(x)}{S_{\mathcal{H}_0}(x)} = \left( \frac{\theta_1}{\theta_0} \right)^{\sum_{i=1}^n x_i} \left( \frac{1 - \theta_1}{1 - \theta_0} \right)^{n - \sum_{i=1}^n x_i}$$

$$= \left( \frac{\theta_1 \cdot (1 - \theta_0)}{\theta_0 \cdot (1 - \theta_1)} \right)^{\sum_{i=1}^n x_i} \left( \frac{1 - \theta_1}{1 - \theta_0} \right)^n$$

$> 1$ , da  $\theta_1 > \theta_0$

$$\Rightarrow Y \mapsto \left( \frac{\theta_1}{\theta_0} \cdot \frac{1 - \theta_0}{1 - \theta_1} \right)^Y \left( \frac{1 - \theta_1}{1 - \theta_0} \right)^n \quad \text{Wahrschein}$$

⇒ Kritische Bereich:

$$K = \{x : \sum_{i=1}^n X_i \geq c\} = \{x : \sum_{i=1}^n X_i \geq \tilde{c}\},$$

wobei  $\sum_{i=1}^n X_i \sim \text{Ber}(n, p_0)$  unter  $P_{p_0}$ .

Wähle nun  $\tilde{c}$ :

$$P_{p_0} \left[ \sum_{i=1}^n X_i \geq \tilde{c} \right] = \sum_{k=\tilde{c}}^n \binom{n}{k} p_0^k (1-p_0)^{n-k}$$

Für  $p_0 = \frac{1}{2}$ ,  $n = 20$  ergibt dies =  $\begin{cases} 0,0013 & \text{für } \tilde{c} = 17 \\ 0,0591 & \text{für } \tilde{c} = 16 \end{cases}$

$\tilde{c} = 16$  ist also  $(1-\alpha)$ -Quantil von  $\text{Ber}(20, \frac{1}{2})$  mit  $\alpha = 0,05$  denn

$$P_{p_0} \left[ \sum_{i=1}^{20} X_i \geq 16 \right] \geq \alpha \quad \text{und} \quad P_{p_0} \left[ \sum_{i=1}^{20} X_i > 16 \right] < \alpha$$

Wähle nun  $\gamma = \frac{\alpha - P_{p_0} \left[ \sum_{i=1}^n X_i > \tilde{c} \right]}{P_{p_0} \left[ \sum_{i=1}^n X_i = \tilde{c} \right]} = \frac{0,05 - 0,0013}{0,0591 - 0,0013} = 0,843$

↳ NP-Test 
$$\phi = \begin{cases} 1 & \text{falls } \sum_{i=1}^{20} X_i \geq 17 \\ \gamma & \text{falls } \sum_{i=1}^{20} X_i = 16 \\ 0 & \text{falls } \sum_{i=1}^{20} X_i < 16 \end{cases}$$

Bem: Im Bsp spielen die Werte  $p_0 = 0,5$  und  $p_1 = 0,6$  keine Rolle. Wichtig war nur  $p_0 < p_1$ , was konstante implizierte. Dies führt zu dem modernen Gleichheitstests, eine Gegenstand der Statistik-Vorlesung ist.

↳ Def 9.6. Sei  $\Theta \subset \mathbb{R}$  und  $T$  eine Statistik. Die Familie  $(P_{\theta})_{\theta \in \Theta}$  mit Gleichgewicht  $S_{\theta}$  hat monotonen Gleichgewicht-Verhalten in  $T$ , falls für alle  $\theta_0 < \theta_1$ ,

$$\frac{P_{\theta_1}(x)}{P_{\theta_0}(x)} = f_{\theta_1, \theta_0}(T(x)) \quad \forall x \in \mathbb{R}$$

gilt, für eine monoton wachsende Fkt  $f_{\theta_1, \theta_0}$  auf  $\mathbb{R}$ .