2018 05 14 Hypotestest

Modell: vi har en nollhypotes (rådande paradigm/ grundantagande) som vi beteckar med H_0. Mot denna ställs en (ny) alternativ hypotes betecknad med H_A eller H_B.

Ex 1 singla slant & låt p= sanno. att få T. Vi har

$H_0: p = 1/2$

$H_1: p > 1/2$ (manipulerad / asymmetrisk slant).

Ex 2 aspirin förebygger hjärtattack?

H_0: aspirin har ej förebyggande effekt mot h-a.

H_1: —//— har föreb. —//—

Anm: bevisbördan ligger naturligt på H_1.

Terminologi

1) om vi förkastar H_0 fastän H_0 är sann, gör vi ett typ I fel

2) vi låter α beteckna sanno. för ett typ I fel, dvs. $\alpha = \mathbb{P}$(förkastar $H_0 | H_0$ sann), α kallas signifikansnivån.

3) att acceptera H_0 då H_0 är falsk kallas för ett typ II fel. Vi låt $\beta = \mathbb{P}$(accepterar $H_0 | H_0$ falsk).

4) styrkan (kraften) hos ett test ges av $1-\beta = \mathbb{P}$(förkastar $H_0 | H_0$ falsk).

Anm: H_0 & H_1 behöver ej vara komplementa. Dvs. H_0 sann $\neq H_1$ falsk (i allmänhet).

Def en test-statistika är en funk. av datamängden med olika fördelning under H_0 resp. H_1.

Ex: vi singlar slant 100 ggr & låter $\hat{p} = \frac{X}{100}$ där $X = \#T$.

Vi förkastar H_0 om $\hat{p}(x) \geq 0.7$

Varför just 0.7? Vad ska vi använda ist?

Anm: 1) här är $\hat{p}$ vår test-statistika

2) $\alpha = \mathbb{P}(\text{förkastar } H_0 \mid H_0 \text{ sann})$

$= \mathbb{P}(X \geq 70 \mid X \sim Bin(100, 1/2))$

$= \mathbb{P}(X \geq 70)$ om $X \sim Bin(100, 1/2)$

3) $1-\beta = \mathbb{P}(X \geq 70 \mid X \sim Bin(100, p))$

OBS: styrkan är en funk. av p

Def förkastningsområdet (RR = rejection region) för ett test är de värden på test-statistikan för vilka vi förkastar H_0.

Ex: vi använder $RR = [0.7, 1]$

Anm: 1) om RR ökar, förkastar vi mer & därmed ökar även signifikansnivån α. Detsamma gäller styrkan $1-\beta$

2) vi vill ha låg signifikansnivå men samtidigt ha en hög styrka.

Tal betrakta exemplet med slantsingling. Bestäm RR s.a $\alpha = 0.05$

L vi förkastar H_0 om $\hat{p}(x)$ "blir för stort", dvs.

$RR = [c, 1]$. Vi vill hitta ett c s.a. $\mathbb{P}(\hat{p}(X) \geq c \mid H_0 \text{ sann}) = 0.05$

Om H_0 är sann (dvs. $p = 1/2$) så gäller att

$$\mathbb{E}[\hat{p}(X)] = \mathbb{E}[X/100] = 1/2 \quad \& \quad Var(\hat{p}) = \frac{1}{100^2} Var(X) = \frac{\overset{1/2}{p(1-p)}}{100} = 1/400$$

CGS ger att $\hat{p}(X) \approx N(\mathbb{E}[\hat{p}], Var(\hat{p})) = N(\frac{1}{2}, \frac{1}{400})$

$\Rightarrow \frac{\hat{p}-1/2}{1/20} = 20(\hat{p}-1/2) \approx N(0,1)$.

Vi får då (under antagandet att H_0 sann)

$$0.05 = \mathbb{P}(\hat{p} \geq c) = \mathbb{P}(20(\hat{p}-1/2) \geq 20(c-1/2)) \approx \mathbb{P}(Z \geq 20(c-1/2))$$

där $Z \sim N(0,1)$.

Tabell ger då att $20(c-1/2) \approx 1.645$ s.a. $c \approx 0.582$

& därmed blir $RR = [0.582, 1]$ //

Anm: 1) $N(1/2, 1/400)$ är den (approx.) nollfördelningen för $\hat{p}(X)$ H_0 sann

2) alt. hypo. er kan vara ensidiga t.ex. $H_1: p > 1/2$, $H_1: p < 1/2$ eller också tvåsidiga $H_1: p \neq 1/2$.

Def p-värdet för ett test är den minsta signifikansnivån för vilket vårt test förkastar H_0 givet oberverade data.

Tal Aspirin-försöket

placebo: 11 034 individer, 189 h.-a.

aspirin: 11 037 –11– 104 –11–.

Låt H_0: aspirin har ingen effekt.

H_1: –11– reducerar risken för h.-a.

Testa detta på signifikansnivån 0.01 & beräkna p-värdet

Låt X = # h.-a. i placebo-gruppen, vara vår test-statistika. Under H_0 gäller att $X \sim Hg(22071, 293, \frac{11034}{22071})$

$\approx N(\mathbb{E}[X], Var(X))$.

Räkning ger att $X \approx N(146.5, 72.3)$.
Vi förkastar H_0 om X "blir för stort"
$(RR = [C, 293])$
p-värde $= \mathbb{P}(X \geq 189 \mid H_0 \text{ sann}) = \mathbb{P}\left(\frac{X - 146.5}{\sqrt{72.3}} \geq \frac{189 - 146.5}{\sqrt{72.3}}\right)$
$\approx 1 - \mathbb{P}(Z \leq 5) \approx 0.000\,000\,29.$
Om signifikansnivån hade varit $3 \cdot 10^{-7}$ hade vi förkastat H_0, så vi förkastar H_0 på signifikansnivån 0.01 //

Anm: om p-värdet $\leq \alpha$ förkastar vi H_0 på sign. nivån α.

Def en enkel hypotes H_0 är en hypotes som fullständigt preciserar en fördelning.
Ex: $H_0 : X \sim N(0, \sigma^2)$ med $\sigma = \sigma_0$

Def en hypo. som ej är enkel kallas för en *sammansatt*.
Ex: $H : X \sim N(0, \sigma^2)$ med $\sigma \neq \sigma_0$

2018 05/15 Hypotestest forts.

Proportionstest

Antag att $H_0: p = p_0$, $H_1: p \neq p_0$, med $X \sim Bin(n, p)$

1) om n är stort så blir $X \approx N(np, np(1-p))$ & vårt RR kan bestämmas ur villkoret $\left|\frac{X - np_0}{\sqrt{np_0(1-p_0)}}\right| \geq c$ ty $p = p_0$ under H_0

2) om n är litet är nollfördelningen $Bin(n, p_0)$ & $RR = [0, x_{1-\alpha/2}] \cup [x_{\alpha/2}, n]$, där $x_{\alpha/2}$ är det minsta talet s.a. $P(X \geq x_{\alpha/2} \mid H_0) \leq \alpha/2$ & $x_{1-\alpha/2}$ största talet s.a. $P(X \leq x_{1-\alpha/2} \mid H_0) \leq \alpha/2$.

Tal Saida gissar färgen på 20 kort som dras slumpmässigt med återläggning. Om X = #kort som gissar rätt $\Rightarrow X \sim Bin(20, p)$.

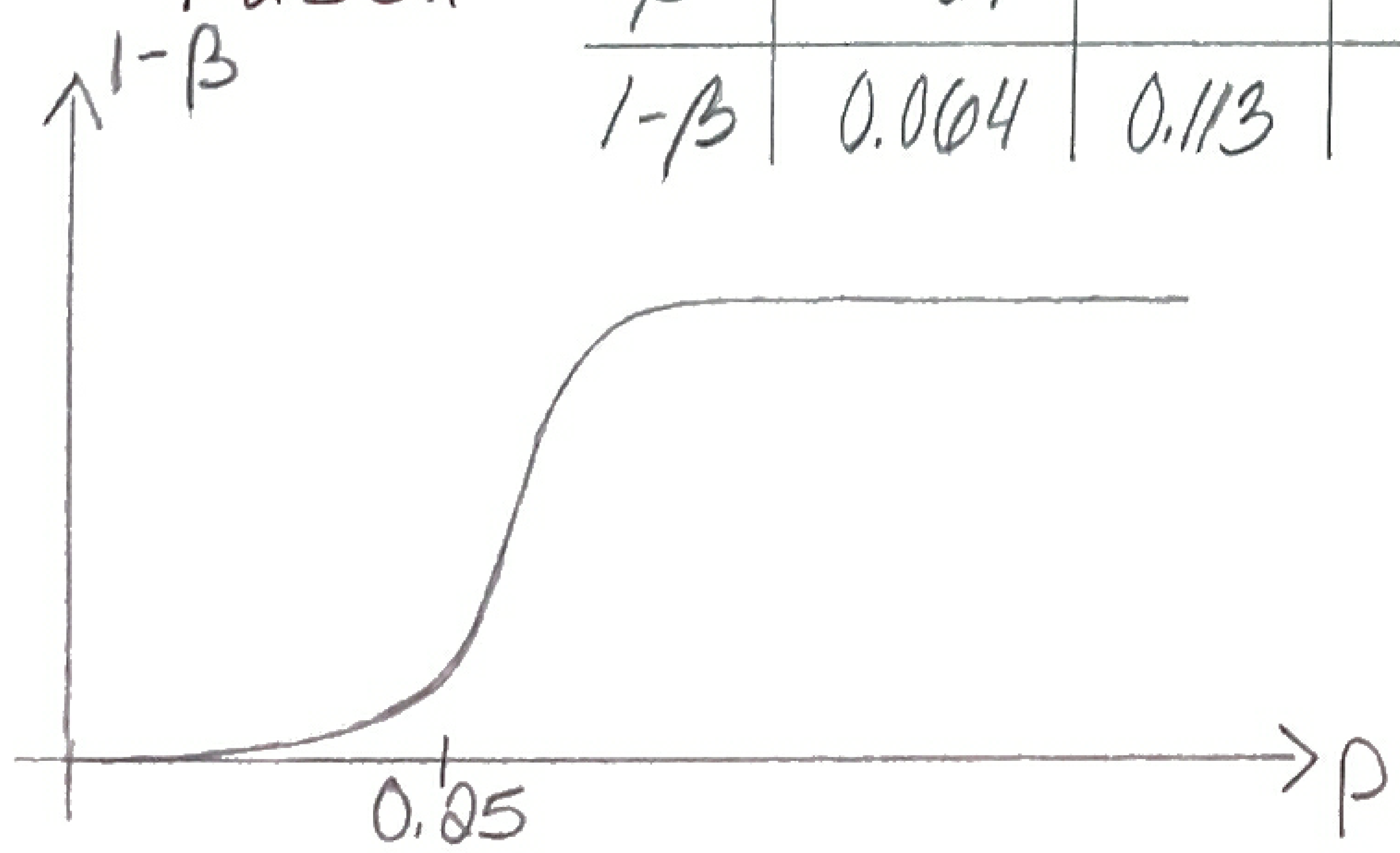
Hitta RR för $H_0: p = 1/4$, $H_1: p > 1/4$, på signif.nivån 5%, & skissa test-styrkan som funk. av p.

L Räkning ger att $P(X \geq 8 \mid H_0) \approx 0.101$, $P(X \geq 9 \mid H_0) \approx 0.041$
testat sig fram till 8 & 9
s.a. $RR = [9, 20]$.

Test-styrkan $1 - \beta = P(\text{förkastar } H_0 \mid H_0 \text{ falsk}) =$
$= P(X \geq 9)$ om $X \sim Bin(20, p)$

Tabell

p	0.27	0.3	0.35	0.4	0.45
$1-\beta$	0.064	0.113	0.238	0.404	0.586



Anm: 1) om Saida har $p=0.4$ upptäcker vi detta m.s. 0.404.

2) Om 1000 personer gör testet på $\alpha = 0.05$ ($\alpha = 0.041$) kommer antalet som av ren slump få över 9 rätt vara Bin(1000, 0.04) "fishing".

Ibland (ofta?) är det uppenbart hur test-statistikan & RR ska väljas. I allmänhet behövs en systematisk metod

Likelihood ratio test

Idé låt H_0 & H_1 vara enkla hypoteser & låt $x = (x_1, \dots, x_n)$ vara data från en diskret fördelning.

Betrakta likelihooden $f(x|H_i) = \mathbb{P}(X_1 = x_1, \dots, X_n = x_n | H_i)$ dvs. sanno. att se utfallet x om H_i är sann, $i = 0, 1$.

Om $f(x|H_0) > f(x|H_1)$ är det mer troligt att få x under H_0 än under H_1.

Vi förkastar H_0 om $\frac{f(x|H_0)}{f(x|H_1)} \le c$ (LR-test)

där c väljs så att vi får önskad sign. nivå.

Här är f tfkn/slf beroende på om vi är i det kont./diskr. fallet.

Anm: 1) LR-test ger ett lämpligt test för alla situationer (med enkla hypoteser).

2) Neyman-Pearson visade att av alla test med sign.nivå α ger LR-testet den högsta styrkan.

3) hur man ska gå till väga för att hitta c beror på situationen.

Ex: Antag att $X_1, \dots, X_n$ är iid & Poissonfördel.

Betrakta $H_0: \lambda = \lambda_0$

$H_1: \lambda = \lambda_1$.

Bilda LR $$\frac{f(X \mid H_0)}{f(X \mid H_1)} = \frac{\prod_{k=1}^{n} f(X_k \mid H_0)}{\prod_{k=1}^{n} f(X_k \mid H_1)} = \frac{\prod_{k=1}^{n} \frac{\lambda_0^{X_k}}{X_k} e^{-\lambda_0}}{\prod_{k=1}^{n} \frac{\lambda_1^{X_k}}{X_k} e^{-\lambda_1}} = e^{n(\lambda_1 - \lambda_0)} \left(\frac{\lambda_0}{\lambda_1}\right)^{X_1 + \dots + X_n}$$

Antag nu att $\lambda_1 > \lambda_0$. Då blir $\frac{f(X \mid H_0)}{f(X \mid H_1)}$ litet om $X_1 + \dots + X_n$ är stor.

Vi förkastar H_0 om $X_1 + \dots + X_n \sim Poi(n\lambda_0)$ överstiger x_α, där x_α valt s.a. $\mathbb{P}(Y \geq x_\alpha) = \alpha$ & där $Y \sim Poi(n\lambda_0)$.

Om ist $\lambda_1 < \lambda_0$ förkastar vi om $X_1 + \dots + X_n$ litet.

Anm 1) uttrycket $e^{n(\lambda_1 - \lambda_0)} \left(\frac{\lambda_0}{\lambda_1}\right)^{X_1 + \dots + X_n}$ är en test-statistika som är ekvivalent med att titta på test-sta. $X_1 + \dots + X_n$

2) $Poi(n\lambda_0)$ är nollfö'rd. för test-sta. $X_1 + \dots + X_n$.

Generaliserad LR

Låt Ω_i $i = 0, 1$ vara delmängder av alla möjliga värden på θ.

Betrakta $H_0: \theta \in \Omega_0$

$H_1: \theta \in \Omega_1$ (sammansatta hypoteser)

Man bildar sedan $\Lambda = \frac{\max_{\theta \in \Omega_0} lik(\theta)}{\max_{\theta \in \Omega} lik(\theta)}$, där $\Omega = \Omega_0 \cup \Omega$

& låter $RR = \{\Lambda \leq c\}$

Ex (s. 339) $X_1, \dots, X_n$ iid $N(\mu, \sigma^2)$ med σ^2 känd.

Vi har $H_0: \mu = \mu_0$

$H_1: \mu \neq \mu_0$.

Vi får att $\max_{\mu \in \Omega_0} lik(\mu) = lik(\mu_0) = \ldots = \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n \cdot$
$\cdot \exp\left(-\frac{1}{2\sigma^2}\sum_{k=1}^{n}(X_k-\mu_0)^2\right)$

& $\max_{\mu \in \Omega} lik(\mu) = \max_{\mu} lik(\mu) = lik(\hat{\mu})$,

där $\hat{\mu} = \bar{X}$ är MLE'n för μ

$\Rightarrow \Lambda = \ldots = \exp\left(-\frac{1}{2\sigma^2}\left(\sum_{k=1}^{n}(X_k-\mu_0)^2 - \sum_{k=1}^{n}(X_k-\bar{X})^2\right)\right)$

$\Rightarrow \sqrt{-2\log(\Lambda)} = \ldots = \frac{\sqrt{n}}{\sigma}|\bar{X}-\mu_0|$

Vi förkastar H_0 om Λ är litet $\Leftrightarrow \sqrt{-2\log(\Lambda)}$ stort $\Leftrightarrow$
$\frac{\sqrt{n}}{\sigma}|\bar{X}-\mu_0|$ stort.

Då $\frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}} \sim N(0,1)$ förkastar vi H_0 på sign. nivån α om
$|\bar{X}-\mu_0| \geq \frac{\sigma}{\sqrt{n}} z_{\alpha/2}$. Vårt RR blir därför $[\mu_0 - \frac{\sigma}{\sqrt{n}} z_{\alpha/2}, \mu_0 + \frac{\sigma}{\sqrt{n}} z_{\alpha/2}]^C$ komplement

Anm: villkoret $\bar{X} \in RR \Leftrightarrow \mu_0 \notin [\bar{X} - \frac{\sigma}{\sqrt{n}} z_{\alpha/2}, \bar{X} + \frac{\sigma}{\sqrt{n}} z_{\alpha/2}] = I_\mu$ (K.I.)
så vi förkastar H_0 på sign. nivå α om μ_0 ej tillhör vårt $(1-\alpha)\cdot 100\%$ K.I.

20180516 Räkneövning 7
Hypotestest

9.12 Låt $X_1, \dots, X_n \sim Exp(\theta)$, ober, med tfkn $f(x|\theta) = \theta e^{-\theta x}$, $x \geq 0$. Visa att RR för ett likelihood ratio-test har formen $\bar{X} e^{-\theta \bar{X}} \leq c$, där c är en konst. då $H_0 : \theta = \theta_0$ & $H_1 : \theta \neq \theta_0$.

L H_1 är en sammansatt hypotes så vi får titta på ett generaliserat LR-test.

Låt $\Lambda = \frac{L(\theta_0)}{\max_\theta L(\theta)} = \frac{f(X_1, \dots, X_n | \theta_0)}{L(\hat{\theta})}$, där $\hat{\theta}$ är MLE för θ.

$$L(\theta) = f(X_1, \dots, X_n | \theta) = \prod_{i=1}^{n} f(X_i | \theta) = \prod_{i=1}^{n} \theta e^{-\theta X_i} = \theta^n e^{-\theta \sum_{i=1}^{n} X_i} =$$
$$= \theta^n e^{-\theta n \bar{X}}$$

$$\Rightarrow \ell(\theta) = n \log \theta - \theta n \bar{X} \Rightarrow \ell'(\theta) = \frac{n}{\theta} - n\bar{X}, \quad \ell''(\theta) = -\frac{n}{\theta^2}.$$

$$\frac{n}{\hat{\theta}} - n\bar{X} = 0 \Leftrightarrow \hat{\theta} = \frac{1}{\bar{X}}, \quad \ell''(\hat{\theta}) = -n\bar{X}^2 < 0 \Rightarrow \hat{\theta} \text{ max pkt.}$$

$$\Rightarrow LR : \Lambda = \frac{\theta_0^n e^{-n\theta_0 \bar{X}}}{\left(\frac{1}{\bar{X}}\right)^n e^{-n}} = (\theta_0 \bar{X})^n e^{-n(\theta_0 \bar{X} - 1)}.$$

Vi förkastar H_0 om LR är liten, dvs. om $\Lambda \leq c$ för ngt c.

$$\Lambda \leq c \Leftrightarrow \theta_0 \bar{X} e^{-\theta_0 \bar{X} + 1} \leq c^{1/n} \Leftrightarrow \bar{X} e^{-\theta_0 \bar{X}} \leq \frac{c^{1/n}}{\theta_0 e} = \text{konst.}$$

9.21 Låt $X \sim U[0, \theta]$ vara en observ. Testa $H_0 : \theta = 1$ & $H_1 : \theta = 2$.

a) Hitta ett test med sign. nivån $\alpha = 0$. Vad är styrkan?

L α = sanno. för felaktig förkastning av nollhypotes.

$\Rightarrow$ förkasta om $X > 1$ (för $P(X > 1 | \theta = 1) = 0$)

$\Rightarrow$ styrka $= 1-\beta = \mathbb{P}(\text{förkasta } H_0 \mid H_1 \text{ sann})$
$= \mathbb{P}(X > 1 \mid \theta = 2) = 1/2$

b) Låt ett test vara sådant att H_0 förkastas om $X \in [0, \alpha]$, för ngt α, $0 < \alpha < 1$. Ange sign. nivå & styrka.

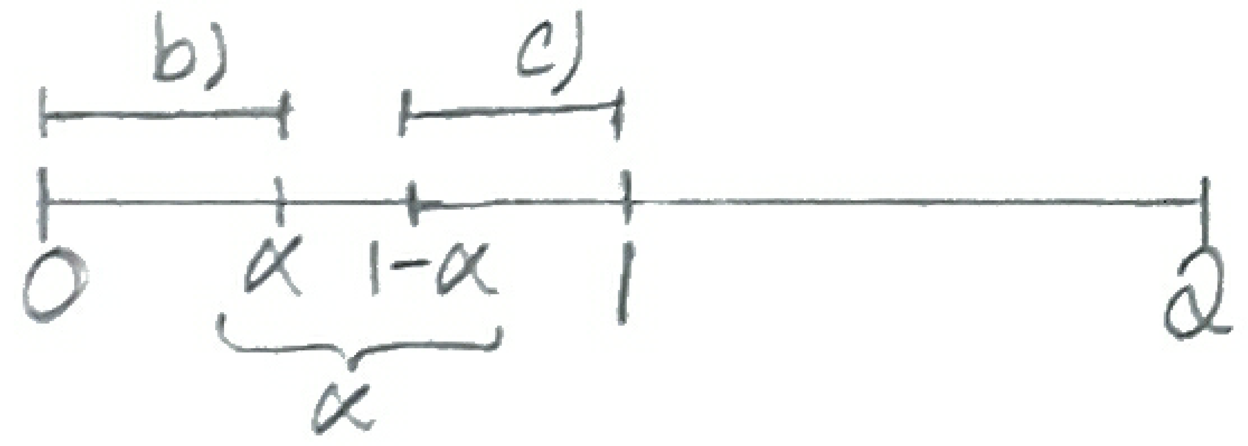
$\mathcal{L}$ sign. nivån $= \mathbb{P}(\text{förk. } H_0 \mid H_0 \text{ sann}) = \mathbb{P}(X \leq \alpha \mid \theta = 1) = \alpha$
styrkan $= \mathbb{P}(\text{förk. } H_0 \mid H_1 \text{ sann}) = \mathbb{P}(X \leq \alpha \mid \theta = 2) = \alpha/2$

c) Förkasta om $X \in [1-\alpha, 1]$. Ange sign. nivå & styrka.

$\mathcal{L}$ sign. nivån $= \mathbb{P}(1-\alpha \leq X \leq 1 \mid \theta = 1) = \alpha$
styrkan $= \mathbb{P}(1-\alpha \leq X \leq 1 \mid \theta = 2) = \alpha/2$

d) Hitta fler test med samma sign. nivå & styrka

$\mathcal{L}$

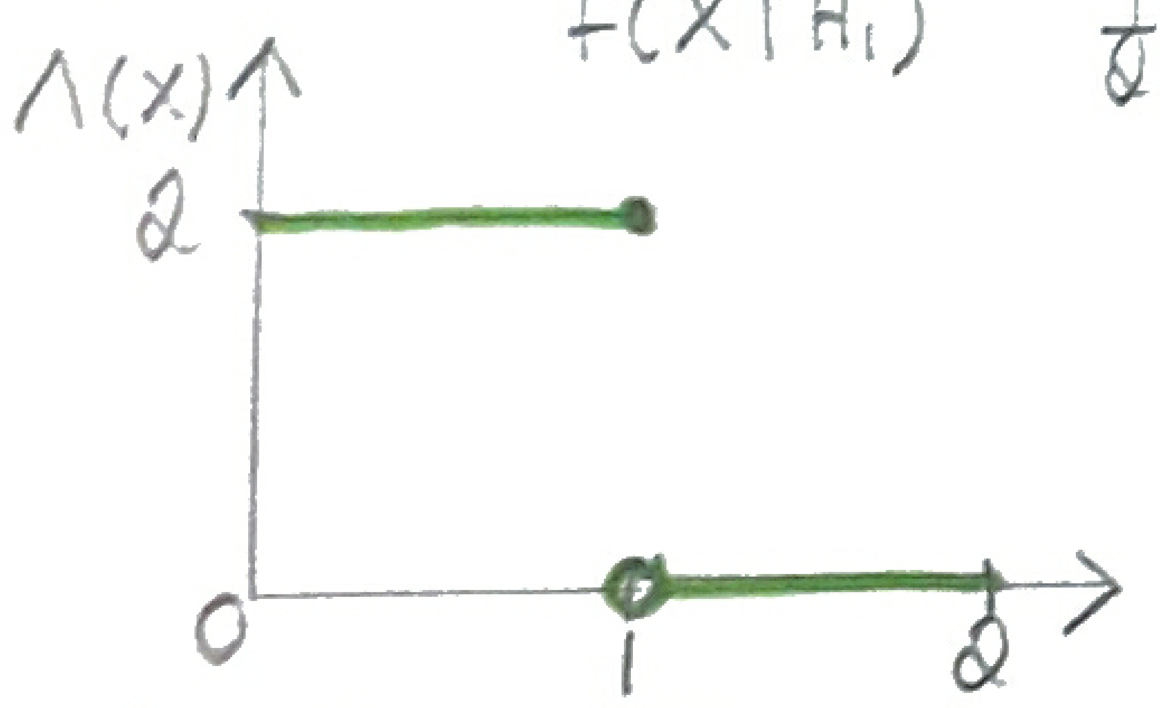


$\Rightarrow$ Alla RR med längd α som ligger helt i $[0, 1]$ uppfyller villkoren.

e) Vilken RR ges av ett LR-test?

$\mathcal{L}$ $\Lambda = \dfrac{f(X \mid H_0)}{f(X \mid H_1)} = \dfrac{\mathbb{1}(0 \leq X \leq 1)}{\frac{1}{2}\mathbb{1}(0 \leq X \leq 2)} \leq C$

$\mathbb{1}(\dots) = \begin{cases} 1 & \text{om villkoret uppfyllt} \\ 0 & \text{annars} \end{cases}$



$\Rightarrow$ $C \geq 2$ ger $X > 0$ & $C < 2$ ger $X > 1$

f) Vad händer om $H_0: \theta = 2$ & $H_1: \theta = 1$?

$\mathcal{L}$ Om förk. reg. är $X \leq \alpha$ får vi signif. $= \mathbb{P}(X \leq \alpha \mid \theta = 2) = \alpha/2$
styrka $= \mathbb{P}(X \leq \alpha \mid \theta = 1) = \alpha$

11.10 Verifiera att ett två-sticksprovs t-test av $H_0: \mu_x = \mu_y$ mot $H_1: \mu_x \neq \mu_y$ med signif. α förkastar omm K.I. för $\mu_x - \mu_y$ ej innehåller 0.

L Anta att populationerna har lika varians (se s. 428 i annat fall).

Vi har att $t = \frac{\bar{x} - \bar{y} - (\mu_x - \mu_y)}{s_p\sqrt{n^{-1}+m^{-1}}} \sim t_{m+n-2}$, där

$s_p = \left(\frac{(n-1)s_x^2+(m-1)s_y^2}{m+n-2}\right)^{1/2}$ är den poolade standardavv. skattningen

& n,m är stickprovsstorlekarna.

Vi förkastar om $|t|>c$ med $\mu_x-\mu_y=0$ (dvs. om statistikans värde är extremt givet H_0).

c ges av $\mathbb{P}(|t|>c \mid \mu_x=\mu_y)=\alpha$ (kan slås upp i tabell för t.ex. $\alpha=0.05$)

$\mathbb{P}(|t|>c \mid \mu_x=\mu_y)=\alpha \Leftrightarrow \mathbb{P}(|t|\leq c \mid \mu_x=\mu_y)=1-\alpha.$

$\Leftrightarrow \mathbb{P}\left(\left|\frac{\bar{x}-\bar{y}-(\mu_x-\mu_y)}{s_p\sqrt{n^{-1}+m^{-1}}}\right| \leq c \,\middle|\, \mu_x=\mu_y\right)=1-\alpha \Leftrightarrow \mathbb{P}\Big(+(\bar{x}-\bar{y})-cs_p\sqrt{n^{-1}+m^{-1}} <$

$< \mu_x-\mu_y < \bar{x}-\bar{y}+cs_p\sqrt{n^{-1}+m^{-1}} \,\Big|\, \mu_x=\mu_y\Big)=1-\alpha$

$\Rightarrow$ K.I för $\mu_x-\mu_y$ enl. def. av K.I.

(Ekvivalens mellan test & K.I ges generellt i sats A kap. 9)

Snarlika beräkningar ger $0 \in$ K.I. $\Leftrightarrow \left|\frac{\bar{x}-\bar{y}}{s_p\sqrt{n^{-1}+m^{-1}}}\right|>c$

11.36 Vi har att två mätserier med <u>parade</u> data, som jmfr två olika mätmetoder

x: 97.2 105.8 97.8

y: 97.2 99.5 96.2

se boken för fler värden...

Analysera data för att se om det är ngn skillnad på mätmetoderna. Vad händer om vi missar att data är parade?

L Parade data: sätt $Z_1=\underline{X}_1-\underline{Y}_1$ & $z_1=x_1-y_1$, om $Z\sim N(\mu,\sigma^2)$ så blir K.I. för μ: $\bar{z} \pm z_{\alpha/2}$.

& data/tabell ger $z_{0.025} \approx 1.96$, $\bar{Z} \approx 0.44$,
$s_Z^2 \approx 21.6$ & $s_Z \approx 4.65$

$\Rightarrow$ K.I.: $0.44 \pm \frac{4.65}{\sqrt{15}} 1.96 \approx [-1.91, 2.79]$ (15 mätpkter)

(Då $0 \in$ K.I. kan vi ej säga ngt om att metoderna skulle ge olika resultat.)

Om ej parade $\frac{\bar{x}-\bar{y}}{s_p\sqrt{n^{-1}+m^{-1}}} \sim t_{m+n-2}$, $s_p^2 = \frac{(n-1)s_x^2+(m-1)s_y^2}{n+m-2} = 8205$

$\Rightarrow s_p \approx 90.58$

$t_{15+15-2}(0.025) \approx 2.048$

K.I. $0.44 \pm 90.58\sqrt{\frac{1}{15}+\frac{1}{15}}\, 2.048 \approx [-67.3, 68.18]$