



Johan Koskinen, Statistiska institutionen, Stockholms universitet

Finansiell statistik, vt-05

F13 hypotesprövning

Kort om konfidensintervall

I de allra flesta fall använder man normalfördelningen

Alltså: ett $(1 - \alpha) \times 100\%$ -igt konfidensintervall för θ ges av

$$\hat{\theta} \pm z_{\alpha/2} SD(\hat{\theta})$$

detta gäller visserligen endast då

$$\hat{\theta} \in N(\theta, SD(\hat{\theta})^2)$$

men när n är stor gäller CGS och

$$\hat{\theta} \in \text{approximativt } N(\theta, SD(\hat{\theta})^2)$$



Johan Koskinen, Department of Statistics

2005-05-01

2

Kort om konfidensintervall

För vissa fördelningar kan man dock lösa ut B som i

fallet med normalfördelning på liknande sätt $I_{\theta} = (\hat{\theta} - B_1, \hat{\theta} + B_2)$

ober. likaförd. s.v. $X_1, X_2, \dots, X_n, X_i \in N(\mu, \sigma^2)$

och vi skattar μ med stickprovsmedelvärdet samt

$$\text{Var}(\hat{\theta}) = \frac{\sigma^2}{n} \quad \text{med} \quad \frac{s^2}{n} \quad \text{kan vi utnyttja} \quad \boxed{\frac{\bar{X} - \mu}{(s/\sqrt{n})} \in t(n-1)}$$

Alltså: ett $(1 - \alpha) \times 100\%$ -igt konfidensintervall för θ ges av

$$\hat{\theta} \pm t_{\alpha/2} \sqrt{\frac{s^2}{n}}$$



Johan Koskinen, Department of Statistics

2005-05-01

3

Kort om konfidensintervall

Om ober. likaförd. s.v. $X_1, X_2, \dots, X_n, X_i \in N(\mu, \sigma^2)$

kan vi utnyttja att

$$\begin{aligned} \sum_{i=1}^n \frac{(X_i - \mu)^2}{\sigma^2} &\in \chi^2(n) \\ \frac{\bar{X} - \mu}{(\sigma/\sqrt{n})} &\in N(0,1) \\ \downarrow \\ \left[\frac{\bar{X} - \mu}{(\sigma/\sqrt{n})} \right]^2 &\in \chi^2(1) \\ \uparrow \\ \sum_{i=1}^n \frac{(X_i - \mu)^2}{\sigma^2} &= \frac{(\bar{X} - \mu)^2}{\sigma^2/n} + \frac{(n-1)s^2}{\sigma^2} \end{aligned}$$

$$\frac{(n-1)s^2}{\sigma^2} \in \chi^2(n-1)$$

för att visa att

Ett $(1 - \alpha) \times 100\%$ -igt konfidensintervall för σ^2 ges av

$$\frac{(n-1)s^2}{\chi_{n-1, \alpha/2}^2} < \sigma^2 < \frac{(n-1)s^2}{\chi_{n-1, 1-\alpha/2}^2}$$

2005-05-01

4

Hypotesprövning fortsättning

Får igen: verkligheten är oftast inte så svart-vit



Andelen får i Derbyshire som är svarta är 1/10



Data



Om/när 1 gäller hur troligt är det att observera det vi observerat (i data)?



Johan Koskinen, Department of Statistics

2005-05-01

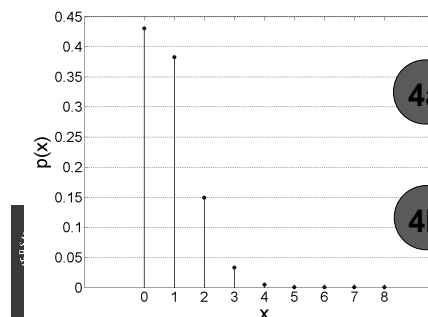
5

Hypotesprövning



Om/när 1 gäller hur troligt är det att observera det vi observerat (i data)?

Om 1 gäller: antalet får X som är svarta vi observerar i åtta "försök"



$X \in \text{Binomial}(8, 1/10)$

Om det vi har observerat inte är särskilt troligt när 1 gäller: förkasta 1

Om det vi har observerat är troligt när 1 gäller: förkasta ej 1



Hypotesprövning - hypotes

Hypotes: antagande angående populationsparameter

Ex: Antag att längden för vuxna störfiskar i Donau är $N(\theta, 0, 19^2)$

Hypotes: $\theta = 1,5$

Ex: För andelen θ svarta får i Derbyshire Hypotes: $\theta = 1/10$



Johan Koskinen, Department of Statistics

2005-05-01

7

Hypotesprövning - hypotes

Vi kallar den hypotes vi vill testa

nollhypotes, H_0

vilken skall testas mot

mot-/alternativhypotesen, $H_1 = \text{inte } H_0$

Ex: Antag att längden för vuxna störfiskar i Donau är $N(\theta, 0, 19^2)$

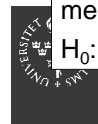
$H_0: \theta = 1,5$

$H_1: \theta \neq 1,5$

men även

$H_0: \theta \leq 1,5$

$H_1: \theta > 1,5$



Johan Koskinen, Department of Statistics

2005-05-01

8

Hypotesprövning - beslutsregel

En regel för när vi skall förkasta H_0 när vi samlat in data kallas för:

beslutsregel.

Ex: När vi drar 20 får förkastar vi

$H_0: \theta = 1/10$

när antal svarta är större än 6

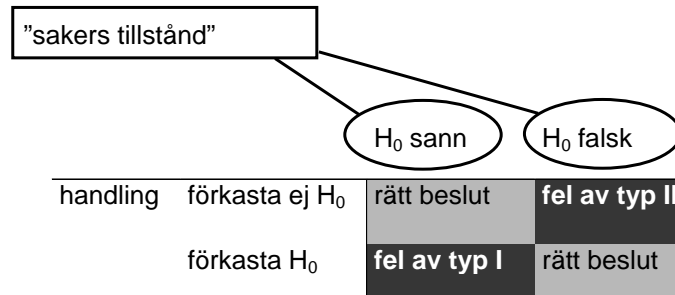


Johan Koskinen, Department of Statistics

2005-05-01

9

Hypotesprövning - feltyper



Johan Koskinen, Department of Statistics

2005-05-01

10

Hypotesprövning - feltyper

Får: med beslutsregeln

förkasta $H_0: \theta = 1/10$

när fler än 6 av tjugo (observerade) får är svarta

		H ₀ sann	H ₀ falsk
handling	förkasta ej H_0	rätt beslut	fel av typ II
	förkasta H_0	fel av typ I	rätt beslut

H_0 sann. θ är verkligen lika med $1/10$

vi drar 20 får och 8 är svarta!

enligt beslutsregeln förkastar vi H_0
(kom ihåg vi vet ej vad som är sant)

vi gör fel av typ I

$\theta = 2/10$

H_0 falsk

vi drar 20 får och 4 är svarta!

enligt beslutsregeln förkastar vi ej H_0
(kom ihåg vi vet ej vad som är sant)

vi gör fel av typ II



Johan Koskinen, Department of Statistics



2005-05-01

11

Hypotesprövning - signifikansnivå

Låt oss undersöka händelsen

vi förkastar H_0

		H ₀ sann	H ₀ falsk
handling	förkasta ej H_0	rätt beslut	fel av typ II
	förkasta H_0	fel av typ I	rätt beslut

Vill ha beslutsregel så att

vi ofta förkastar H_0 när H_0 är falsk

vi sällan förkastar H_0 när H_0 är sann

Princip: hellre fria än fälla



Johan Koskinen, Department of Statistics

2005-05-01

12

Hypotesprövning - signifikansnivå

Princip: hellre fria än fälla

Kan vi konstruera beslutsregel så att vi aldrig gör fel av typ I?

SVAR: nja...

Börja med:

ges av vår beslutsregel

$$\alpha = P(\text{förkasta } H_0 \mid H_0 \text{ sann})$$

		H_0 sann	H_0 falsk
handling	förkasta ej H_0	rätt beslut	fel av typ II
	förkasta H_0	fel av typ I	rätt beslut

2005-05-01

13

Hypotesprövning - signifikansnivå

Vad innebär det att göra

$$\alpha = P(\text{förfästa } H_0 \mid H_0 \text{ sann})$$

liten?

Ex: För s.v. $X \in N(\mu, \sigma^2)$ (tänk t.ex. på längden på störrar)

och hypotesen

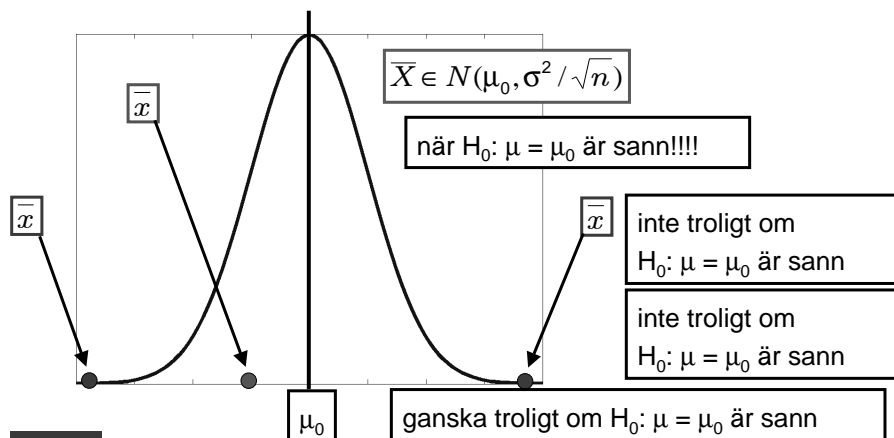
$$H_0: \mu = \mu_0$$

		H_0 sann	H_0 falsk
handling	förkasta ej H_0	rätt beslut	fel av typ
	förkasta H_0	fel av typ I	rätt beslut



Hypotesprövning - signifikansnivå

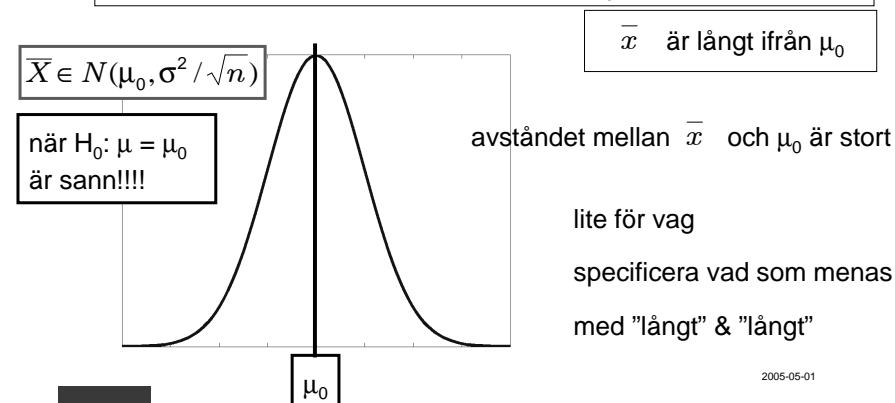
För s.v. $X \in N(\mu, \sigma^2)$, $H_0: \mu = \mu_0$, medför att



Hypotesprövning - signifikansnivå

M.a.o. om s.v. $X \in N(\mu, \sigma^2)$, $H_0: \mu = \mu_0$

verkar en beslutsregel av typen: förkasta H_0 när



2005-05-01

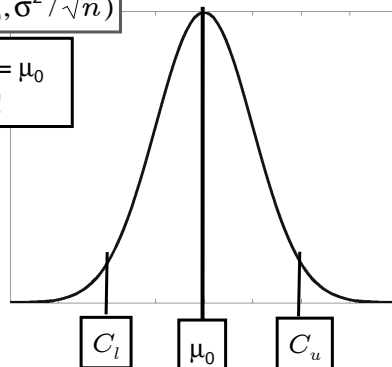
16

Hypotesprövning - signifikansnivå

Vi bestämmer kritiska värden/förkastningsregion

$$\bar{X} \in N(\mu_0, \sigma^2 / \sqrt{n})$$

när $H_0: \mu = \mu_0$
är sann!!!!



om

$$\bar{x} > C_u$$

eller

$$\bar{x} < C_l$$

\bar{x} är långt ifrån μ_0



Johan Koskinen, Department of Statistics

2005-05-01

17

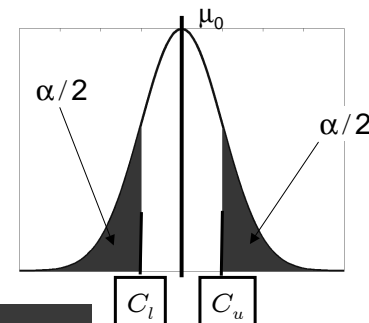
Hypotesprövning - signifikansnivå

Ofta i statistik: något är "stort" om slh att få något större är liten

Vi börjar med att bestämma signifikansnivån vi vill ha

$$\alpha = P(\text{förfasta } H_0 \mid H_0 \text{ sann})$$

För s.v. $X \in N(\mu, \sigma^2)$, $H_0: \mu = \mu_0$, och föreslagna förkastningsregel



$$\alpha = P(\text{förfasta } H_0 \mid H_0 \text{ sann})$$

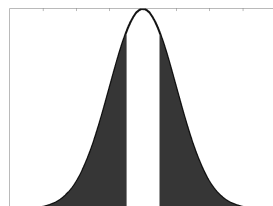
$$= P(\bar{X} < C_l \mid H_0 \text{ sann})$$

$$+ P(\bar{X} > C_u \mid H_0 \text{ sann})$$

2005-05-01

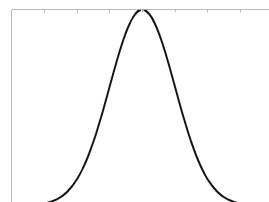
18

Hypotesprövning - signifikansnivå



α för stort

vi förkastar H_0 för ofta



α för litet

vi förkastar H_0 för
sällan?

		H_0 sann	H_0 falsk
handling	förfasta ej H_0	rätt beslut	fel av typ II
	förfasta H_0	fel av typ I	rätt beslut

alltså om ej "rätt" så "fel typ II"

2005-05-01

19

Hypotesprövning - signifikansnivå

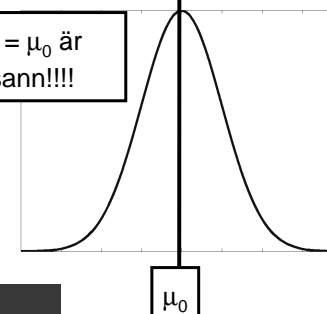
Vi kan alltså välja signifikansnivån α

så att sannolikheten att förfasta H_0 är $\alpha \dots$ givet att H_0 stämmer

Antag H_1 är sann, alltså inte H_0 !

$$\bar{X} \notin N(\mu_0, \sigma^2 / \sqrt{n})$$

$H_0: \mu = \mu_0$ är
inte sann!!!!



		H_0 sann	H_0 falsk
handling	förfasta ej H_0	rätt beslut	fel av typ II
	förfasta H_0	fel av typ I	rätt beslut

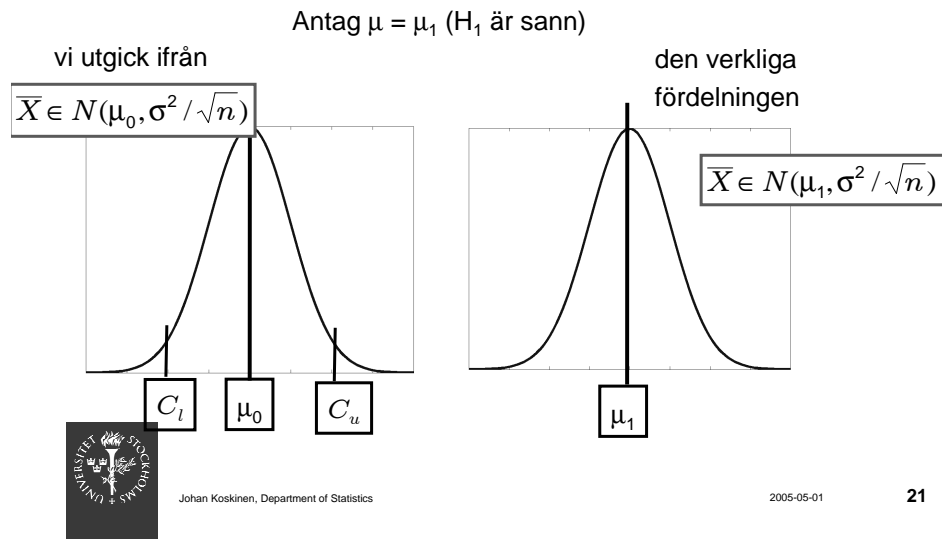


Antag $\mu = \mu_1$ (H_1 är sann)

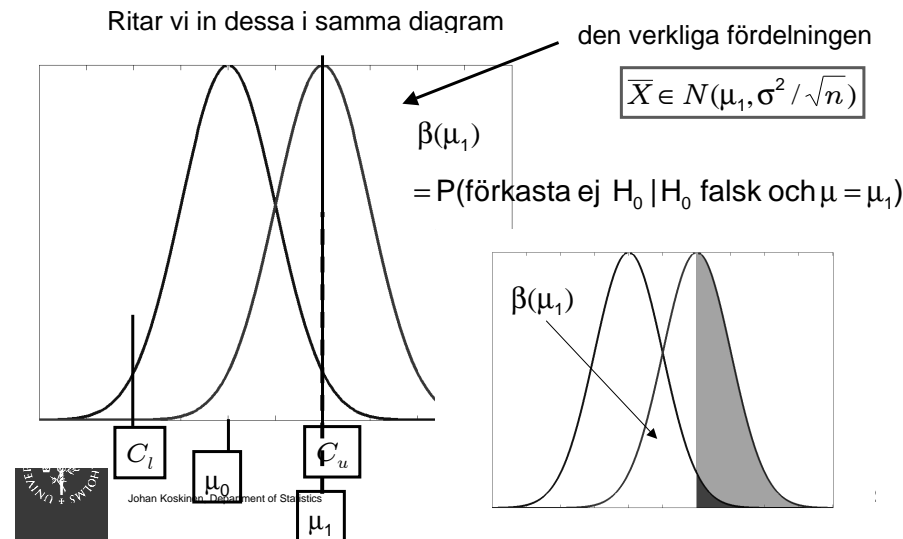
2005-05-01

20

Hypotesprövning - signifikansnivå



Hypotesprövning - signifikansnivå



Hypotesprövning - signifikansnivå

Slutsats: sätter vi för liten signifikansnivå α riskerar

$$P(\text{förfasta ej } H_0 \mid H_0 \text{ falsk}) = P(\text{typ II fel})$$

bli för stor...

Vanliga signifikansnivåer

$\alpha = 0,1$ vi riskerar att förfasta riktig H_0 i 10 fall av hundra

$\alpha = 0,05$ vi riskerar att förfasta riktig H_0 i 5 fall av hundra

$\alpha = 0,01$ vi riskerar att förfasta riktig H_0 i 1 fall av hundra

	H_0 sann	H_0 falsk
handling	förfasta ej H_0	rätt beslut
	förfasta H_0	fel av typ I
		rätt beslut

2005-05-01

23

Hypotesprövning - form

Hypotes

teststatistika/funktion

beslutsregel

insamling av "bevis"/data

beslut



Johan Koskinen, Department of Statistics

2005-05-01

24

Hypotesprövning - normal

Ex: För s.v. X_1, X_2, \dots, X_n , fördelade som $X \in N(\mu, \sigma^2)$

testa $H_0: \mu = \mu_0$

mot $H_1: \mu \neq \mu_0$ på signifikansnivån α

Teststatistikan

$$\frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}} \in N(0,1)$$

då $H_0: \mu = \mu_0$
är sann.

Vi förkastar $H_0: \mu = \mu_0$ på signifikansnivån α då det observerade värdet på teststatistikan är större än $z_{\alpha/2}$ eller mindre än $-z_{\alpha/2}$

