
Anvendt Statistik

Lektion 4

Hypotesetest generelt
Test for middelværdi
Test for andele

Hypoteser og Test

■ Hypotese

- I statistik er en hypotesen en påstand om en populationsparameter. Typisk en påstand om at parameteren tager en bestemt værdi.

■ Signifikanstest

- En signifikanstest opsummerer, på baggrund af data, beviser for/imod hypotesen.

■ Eksempel

- En virksomhed har 50/50 kønsfordeling blandt dem der kan forfremmes. 9 af 10 forfremmelser går til mænd. Virksomhedens påstand: Der er ingen kønsdiskriminering. Det skal testes!

Signifikanstest - De fem dele

- Et (signifikans)test består af følgende fem dele:
 - 1) Antagelser
 - 2) Hypoteser
 - 3) Test-størrelse
 - 4) P -værdi
 - 5) Konklusion
-

1) Antagelser

- En test baserer sig typisk på **et antal antagelser** for at være gyldig. Antagelser omhandler typisk:
 - **Type af data**
 - Kvalitative eller kvantitative data
 - **Randomisering**
 - På hvilken måde er stikprøven udtaget
 - **Populations-fordelingen**
 - Antages det er populationen er normalfordet?
 - **Stikprøvestørrelsen**
 - Er der antagelser om at stikprøven skal have en hvis størrelse?

2) Hypoteser

■ Nul-hypotesen

- Nul hypotesen er en påstand om at en parameter tager en **bestemt værdi**. Betegnes H_0 .

■ Virksomhedens påstand (Nul-hypotesen):

- $H_0: \pi = 1/2$ (π er andelen af forfremmede der er mænd)

■ Alternativ-hypotesen

- Alternativ-hypotesen er en påstand om at parameteren ligge i **et alternativt interval** (i forhold til nul-hypotesen). Betegnes H_a .

■ Kvindernes påstand (Alternativ-hypotesen):

- $H_a: \pi > 1/2$

Udgangspunktet er H_0

- **Grundlæggende princip:**

- H_0 antages at være sand indtil beviserne imod (baseret på data) er så stærke, at vi må afvise H_0 – og derfor acceptere H_a .

- Man kan *ikke* bevise H_0 – kun afvise.
- Det er derfor et stærkere udsagn at afvise H_0 end ikke at kunne afvise.
- **Retssags-analogi:** Uskyldig (H_0 : uskyldig) indtil det modsatte (H_a : skyldig) er bevist (ud over almindelig tvivl).

3) Test-størrelse

■ Teststørrelse

- En teststørrelse er et mål for hvor langt data ligger fra påstanden i nul-hypotesen.

- Jo mere teststørrelsen peger på alternativhypotesen, jo mere **kritisk** er teststørrelsen (for nul-hypotesen).

■ Eksempel

- Teststørrelse: Stikprøveandelen $\hat{\pi}$.
- Fra virksomhedens synspunkt (H_0) er det kritisk hvis $\hat{\pi}$ er højere end 0.5, da det støtter kvindernes påstand (H_a).
- **Spørgsmål:** Hvor meget højere end 0.5 skal $\hat{\pi}$ før beviserne mod virksomheden er "stærke nok"?

4) P -værdi

- P -værdien er et mål for *hvor kritisk* en given teststørrelse er. P -værdien betegnes P .

- **P -værdi**

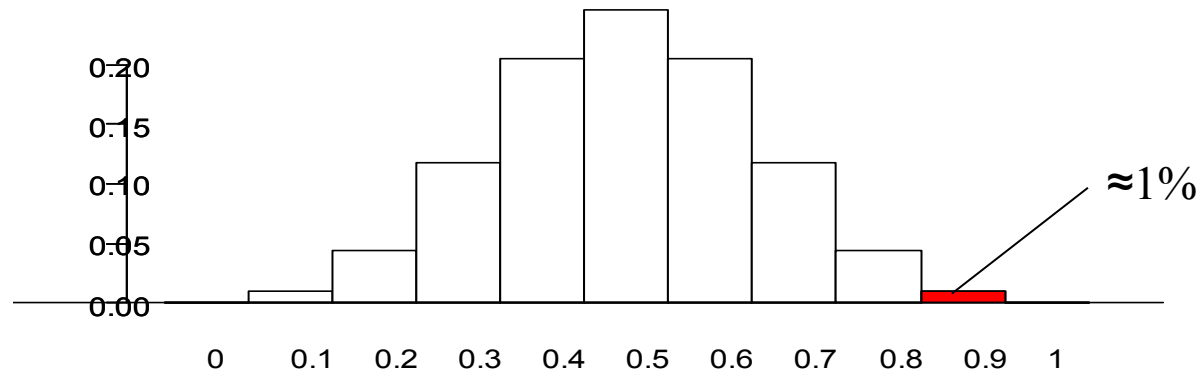
- En P -værdi er **sandsynligheden** for at "den næste" test-størrelse er **mindst lige så kritisk** som den observerede, under antagelse af, at **H_0 er sand**.

- **Eksempel**

- Hvis P -værdien er 0.01 og nul-hypotesen er sand, så er der kun 1% sandsynlighed for at observere en mere kritisk test-størrelse "næste gang".

P-værdi : Eksempel

- Hvis H_0 er sand, dvs. $\pi = 0.5$, så ser fordelingen af andel mænd blandt 10 kandidater således ud:



- **Teststørrelse** Vi har observeret en andel på $\hat{\pi} = 9/10 = 0.9$.
- Dvs. en andel på 0.9 eller 1.0 er lige så eller mere kritisk.
- Det røde område er P -værdien – sandsynligheden for en mindst lige så kritisk observation "næste gang", hvis H_0 er sand.

5) Konklusion

- Jo mindre P -værdien er, jo mindre tror vi på, at nulhypotesen er sand og jo mere hælder vi til alternativhypotesen.
- Typisk afvises H_0 hvis $P \leq 0.05$. De 0.05 kaldes **signifikans-niveauet** og betegnes α .
- Hovedet-under-armen-konklusion
 - $P \leq \alpha$ så afvises H_0 og H_a accepteres
 - $P > \alpha$ så kan H_0 ikke afvises
 - α er typisk 0.05 eller 0.1

Test for middelværdi

- Vi vil nu se på hvordan, man kan opstille og teste hypoteser for populationens middelværdi μ .
 - Vi skal derfor igennem 'de fem trin':
 - **Antagelser**
 - Tilfældig stikprøve
 - Normalfordelt population
 - Testen er relativt robust over for afvigelser fra normal-antagelsen, hvis stikprøven ikke er for lille og populationen ikke er alt for "unormal" – fx meget skæv.
-

Test for middelværdi: Hypoteser

- **Nul-hypotesen** har formen

- $H_0: \mu = \mu_0$ fx $H_0: \mu = 0$

- Dvs. nul hypotesen er at den (ukendte) middelværdien μ tager værdien μ_0 .

- **Alternativ-hypotesen**

- $H_a: \mu \neq \mu_0$ fx $H_0: \mu \neq 0$

- Dvs. alternativ-hypotesen er at μ er forskellig fra μ_0 – enten større eller mindre.
- Denne test er **to-sidet**, da afvigelser i begge retninger falder ind under alternativ-hypotesen.

Test for middelværdi: Teststørrelse

- **Husk:** Stikprøvegennemsnittet \bar{y} følger en normalfordelingen, hvis populationen er normal eller stikprøven er tilstrækkelig stor (jf. CLT).
- Værdier af \bar{y} langt fra μ_0 er kritiske for H_0 .
- Vi måler afstanden mellem \bar{y} og μ_0 i antal standardfejl.
- Da σ er ukendt bruger vi den estimerede standardfejl.

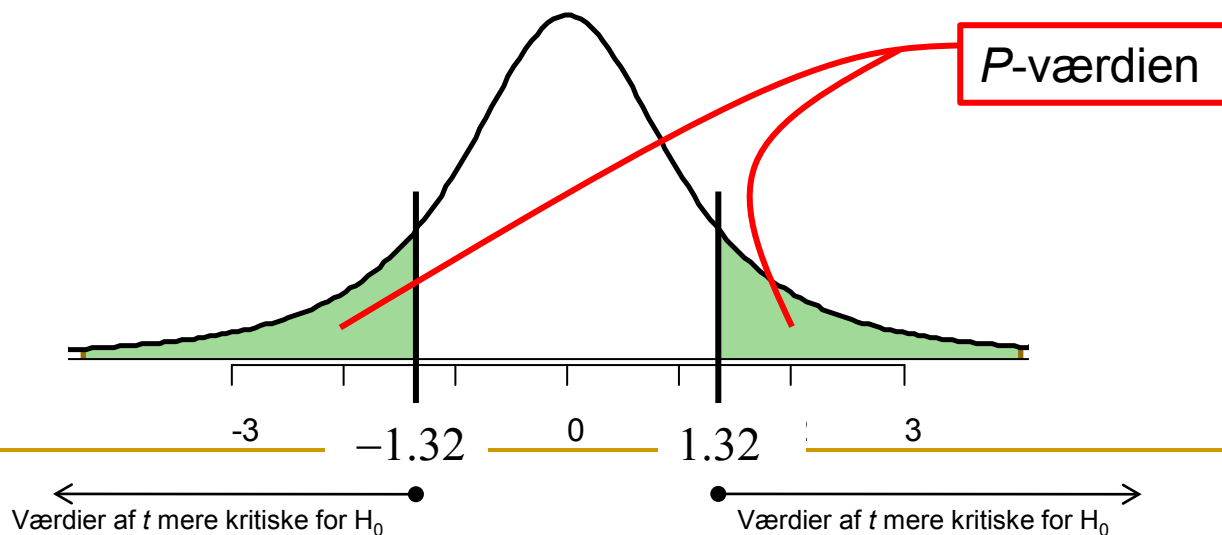
■ Teststørrelse:

$$t = \frac{\bar{y} - \mu_0}{se}$$

hvor $se = s/\sqrt{n}$.

Test for middelværdi: P -værdi

- P -værdien er sandsynligheden for at observere en mindst lige så kritisk teststørrelse, som den der er observeret, hvis H_0 -hypotesen ($\mu = \mu_0$) er sand.
- Hvis H_0 -hypotesen ($\mu = \mu_0$) er sand, så følger teststørrelsen en t -fordeling med $n-1$ frihedsgrader.
- **Eksempel:** Antag at t -værdien er 1.32:



Test for middelværdi: Konklusion

- Jo mindre P -værdi, jo mindre tror vi på nul-hypotesen.

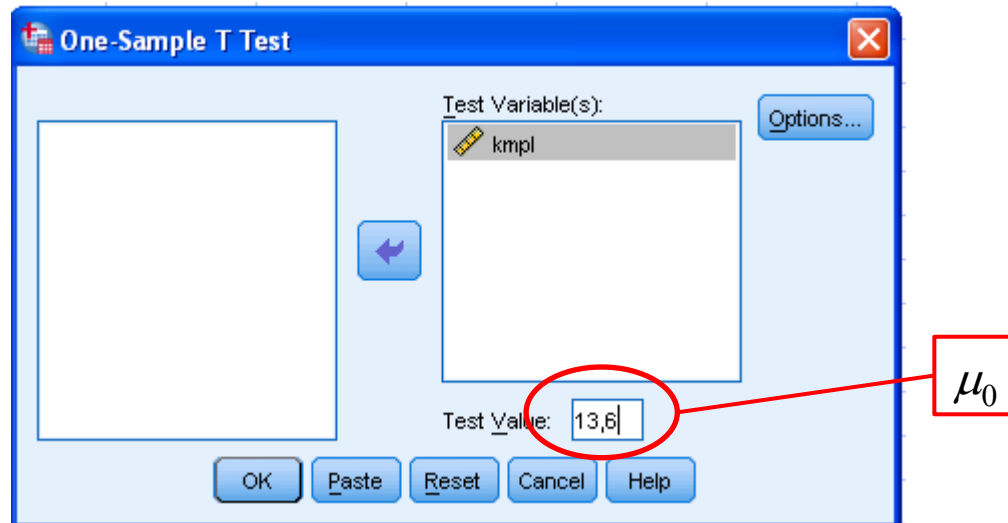


Eksempel: Hvor langt på literen?

- En bil-producent påstår at en specifik model kører 13.2 km/l. Holder det i praksis?
- $n = 132$ ejere af den givne model har indrapporteret deres forbrug.
- Gennemsnitsforbruget er $\bar{y} = 12.97$ og standardafvigelsen er $s = 1.02$.
- Test hypotesen $H_0: \mu = 13.2$ vs. $H_a: \mu \neq 13.2$

Test i SPSS

- Analyze → Compare Means → One-Sample T Test



One-Sample Test

Test Value = 13.2

	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
kmp1	-2,612	131	,010	-.23104	-.4061	-.0560

P-værdien

To-sidet test og Konfidensintervaller

- Hvis vi udregner et 95% konfidensinterval for middelværdien får vi

$$\bar{y} \pm t_{0.025} \frac{s}{\sqrt{n}} \quad 12.97 \pm 1.96 \frac{1.02}{\sqrt{132}} = 12.97 \pm 0.17 \quad (12.80; 13.14)$$

- Bemærk at konf. intervallet *ikke* indeholder 13.2!
- **Intuition:** Da et 95% konf. interval *ikke* indeholder μ_0 , så bør vi afvise H_0 .
- **Præcist:** Hvis vi afviser H_0 , når μ_0 *ikke* er i et 95% konf. interval, så svarer det præcist til at afvise H_0 , når $P < 0.05$.

En-sidet test

- **Nul-hypotesen** har formen

- $H_0: \mu = \mu_0$

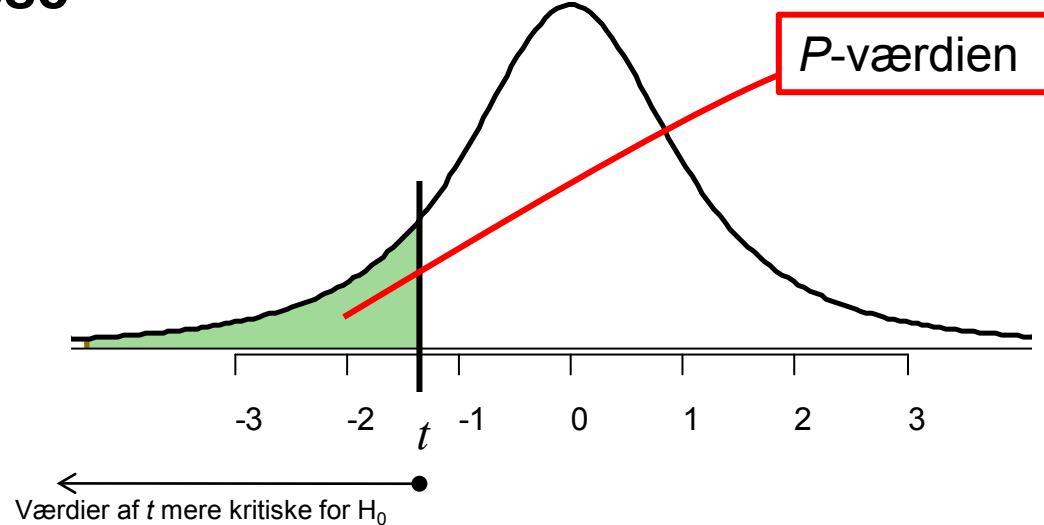
- **Alternativ-hypotese**

- $H_a: \mu < \mu_0$

- **Teststørrelse**

- $$t = \frac{\bar{y} - \mu_0}{se}$$

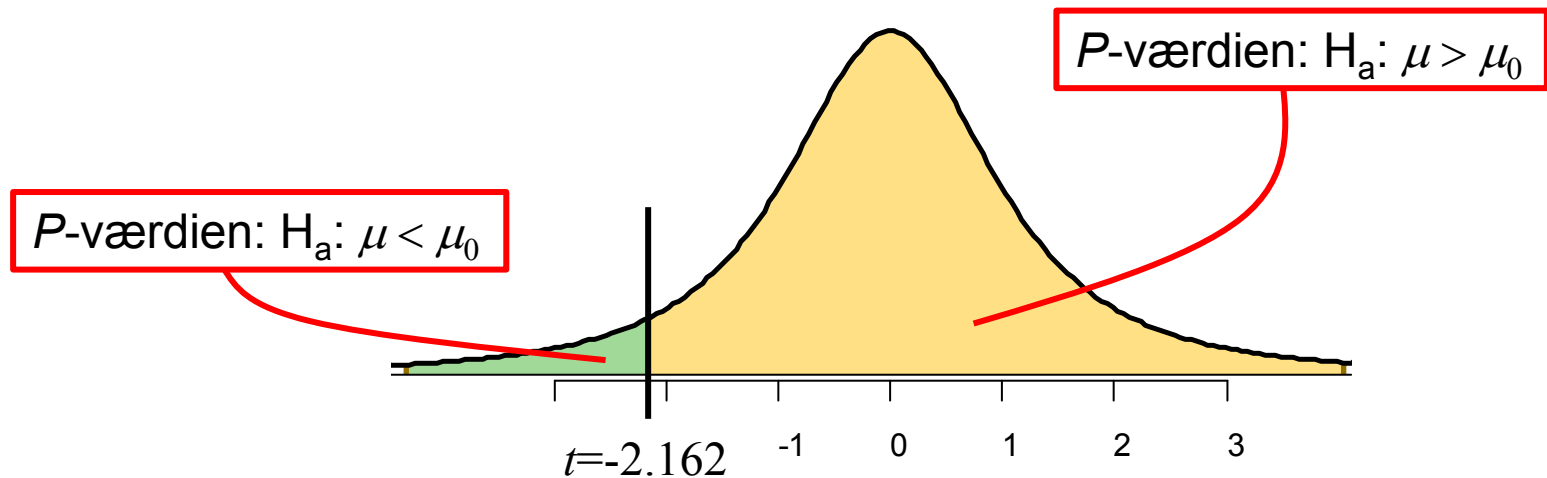
- **P-værdi**



- Hvad med bilproducenten?

Eksempel

- Bilproducenten igen
- **Nul-hypotesen** har formen
 - $H_0: \mu = \mu_0$
- **Alternativ-hypotesen**
 - $H_a: \mu < \mu_0$ Bilerne kører *kortere* pr. liter end påstået
 - $H_a: \mu > \mu_0$ Bilerne kører *længere* pr. liter end påstået



Test af middelværdi: Opsummering

- **Antagelser:**

- Normal population og/eller stor stikprøve

- **Nul-hypotese**

- $H_0: \mu = \mu_0$

- **Alternativ-hypoteser**

- $H_a: \mu \neq \mu_0$ eller $H_a: \mu < \mu_0$ eller $H_a: \mu > \mu_0$

- **Teststørrelse**

- $t = \frac{\bar{y} - \mu_0}{se}$, hvor $se = s/\sqrt{n}$

- **P-værdi...**

Test af andele: Antagelser

- I det følgende vil vi opstille og teste hypoteser for populations-andelen π .
 - **Antagelser:**
 - Tilfældig stikprøve
 - Stikprøven er så stor, at stikprøve-andelen kan antages approksimativt at følge en normalfordeling.
-

Test af andele: Hypoteser

- **Nul-hypotesen** har formen

- $H_0: \pi = \pi_0$ fx $H_0: \pi = 0.5$

- Dvs. hypotesen er at den (ukendte) andel π tager den bestemte værdi π_0 .

- **Alternativ-hypoteser**

- $H_a: \pi \neq \pi_0$ fx: $H_a: \pi \neq 0.5$ (to-sidet test)

- $H_a: \pi > \pi_0$ eller $H_a: \pi < \pi_0$ (en-sidet test)

Test af andele: Teststørrelse

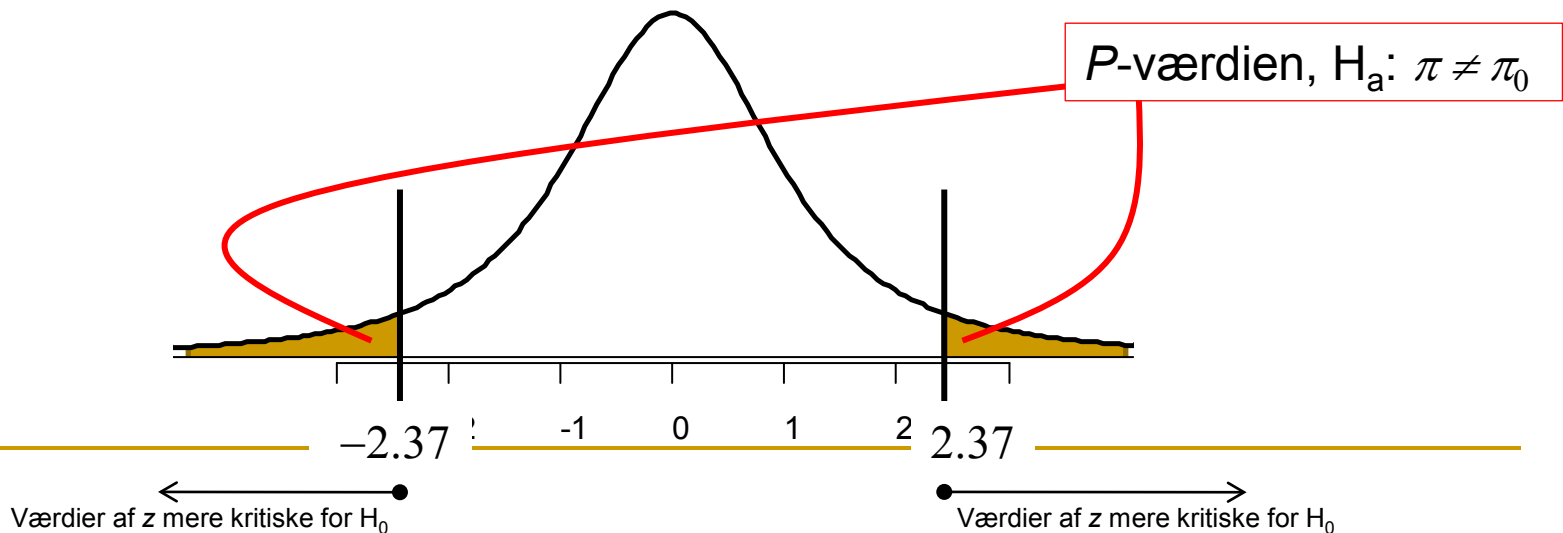
- Stikprøveandelen p følger en normalfordeling.
- Hvis H_0 er sand, dvs. $\hat{\pi} = \pi_0$, så er har $\hat{\pi}$
 - middelværdi π_0
 - standardfejl $se = \sqrt{\pi_0(1 - \pi_0)/n}$
- Værdier af $\hat{\pi}$ langt fra π_0 er kritiske for H_0 .
- Vi måler afstanden i antal standardfejl

- **Teststørrelse:**

$$z = \frac{\hat{\pi} - \pi_0}{se}$$

Test af andele: P -værdi

- P -værdien er sandsynligheden for at observere en mindst lige så kritisk teststørrelse, som den er observeret, hvis H_0 -hypotesen ($\pi = \pi_0$) er sand.
- Hvis H_0 -hypotesen ($\pi = \pi_0$) er sand, så følger teststørrelsen en standard-normalfordeling med.
- **Eksempel:** Antag $H_a: \pi \neq \pi_0$ og $z = 2.37$:



Eksempel: Hvem kører bilen?

- Blandt de $n = 132$ test-chauffører var 41% kvinder.
 - Lad π betegne andelen af kvinder blandt de testene.
 - Test hypotesen $H_0: \pi = 0.5$ vs. $H_a: \pi \neq 0.5$
-

Test i SPSS

- SPSS bruger ikke normal-approksimationen, men fastholder at antallet af kvinder følger en binomialfordeling.
- Analyze → Nonparametric Test → One-Sample...

The image shows two overlapping dialog boxes from SPSS. The main dialog is 'One-Sample Nonparametric Tests' with the 'Settings' tab selected. Under 'Choose Tests', the 'Binomial' test is selected. The 'Options...' button is circled in red. The 'Binomial Options' sub-dialog is open, showing 'Hypothesized proportion:' set to '0,5', which is also circled in red. A red box labeled π_0 points to this value. The 'Compare observed binary probabilities' checkbox is checked.

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The categories defined by koen = 0.00 and 1.00 occur with probabilities 0.5 and 0.5.	One-Sample Binomial Test	.045	Reject the null hypothesis.

P-værdien

Beslutningsskema

- Vælg signifikansniveau α , typisk 0.1 eller 0.05.
- Udregn den relevante P -værdi.
- Vi kan derefter konkludere efter følgende skema

	Konklusion	
P -værdi	H_0	H_a
$P \leq \alpha$	Afvise	Acceptere
$P > \alpha$	Ikke afvise	Ikke acceptere

Man begår fejl...

- Ind imellem vil data lede os til at drage en forkert konklusion.

- **Type I og Type II Fejl**

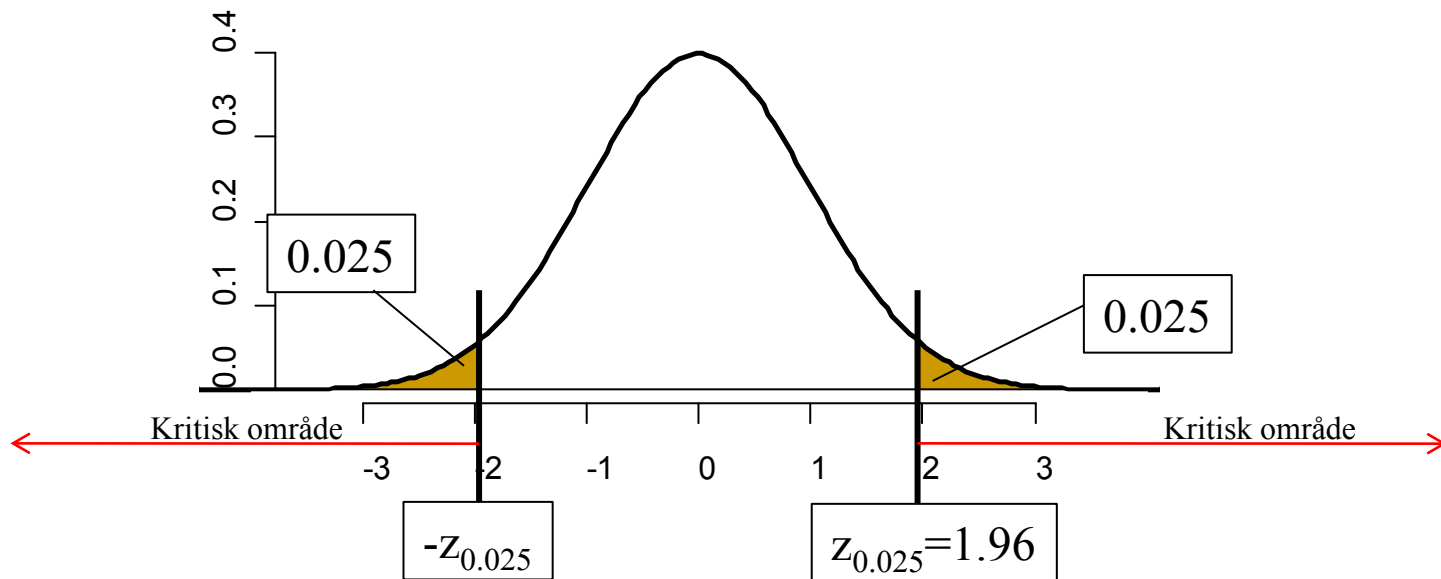
- Hvis vi afviser H_0 selvom der er sand, så har vi begået en Type I fejl.
- Hvis vi *ikke* afviser H_0 selvom der er falsk, så har vi begået en Type II fejl.

- Oversigt

	Beslutning	
	Afvis H_0	Ikke afvis H_0
H_0 sand	Type I fejl	Korrekt beslutning
H_0 falsk	Korrekt beslutning	Type II fejl

Kritisk område

- Det kritiske område er de værdier af teststørrelsen, der fører til en afvisning af nul-hypotesen.
- **Eksempel:** Test af andele
- Antag et signifikans-niveau $\alpha = 0.05$.



Signifikansniveau og Type I fejl

- Hvis H_0 hypotesen er sand, så er sandsynligheden for at "lande" i det kritiske område α .
- Dvs. sandsynligheden for at begå en Type I fejl er α signifikansniveauet.
- Når $P(\text{Type I fejl})$ går ned, går $P(\text{Type II fejl})$ op.
 - Antag $\alpha = 0.00001$
 - Hvis H_0 **sand**: Næsten umuligt at begå en Type I fejl
 - Hvis H_0 **falsk**: Næsten umuligt at undgå en Type II fejl