FORELESNING 16 — MET 3431

EIVIND ERIKSEN — MAR 20 2012 — STATISTIKK

PLAN:

① Hypotesetesting: Type I og II feil [T] 8.5
Testing av gjennomsnitt
Repetisjon av testing av andelen

② Oppgaver fra Kap. 8: Hypotese-testing

Husk: Studentveiledning I i dag kl 17.15 – 20.15 i B2
Oppgavesett vil bli delt ut

Type I og Type II feil

En hypotesetest kan gi feil konklusjon. Det kan skje på to måter.

Type I feil

- H_0 er egentlig sann, men du forkaster H_0
- α - signifikansnivået- er sannsynligheten for å begå en type I feil.

Type II feil

- H_0 er egentlig gal, men du forkaster ikke H_0
- β er symbolet for sannsynligheten for en Type II feil, ~~også kalt testens *styrke* (eng: power)~~

$1-\beta$ er testens styrke (power)

$=$ sannsynligheten for at H_0 er gal og at vi forkaster H_0.

Type I og Type II feil

Fire ting kan skje når du tester en hypotese:

Type I

Table 8-1	Type I and Type II Errors		
		True State of Nature	
		The null hypothesis is true	The null hypothesis is false
Decision	We decide to reject the null hypothesis	[illegible]	Correct decision
	We fail to reject the null hypothesis	Correct decision	[illegible]

Type II

Oppsummering

Vi har diskutert:

- Null- og alternativ hypoteser H_0 og H_1
- Testobservator (z ellr t)
- Signifikansnivå α
- p-verdi
- Beslutningsregel (tradisjonell og p-verdi metoden)
- Type I og II feil

Testen for en andel i populasjonen

Notasjon

- n: størrelsen på stikkprøven
- $\hat{p}$: andelen i stikkprøven
- p: andelen i populasjonen, ifølge nullhypotesen
- $q = 1 - p$

Forutsetninger for testen

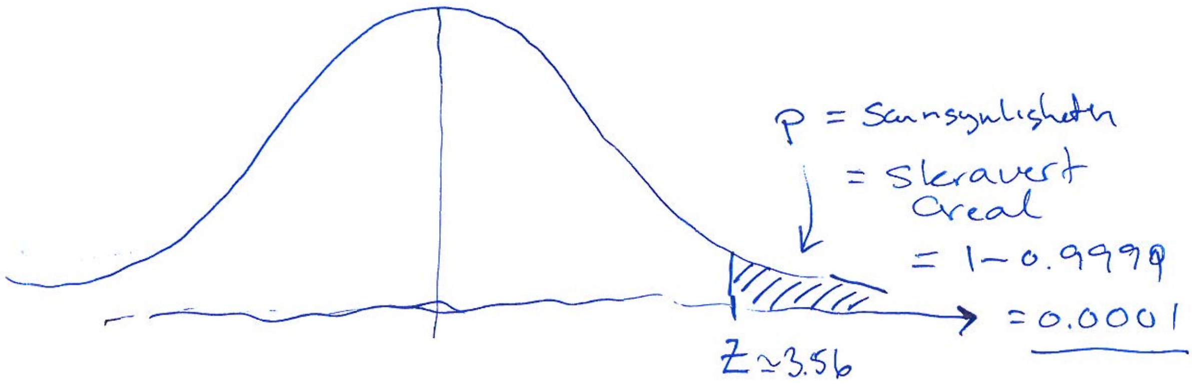
- Stikkprøven er et tilfeldig utvalg
- Betingelser for binomialfordeling holder (section 5-3)
- $np \geq 5$ og $n(1-p) \geq 5$

Testobservatoren for test om en andel

Testobservator

$$z = \frac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{n}}}$$

er tilnærmet normalfordelt når forutsetningene på forrige side holder.



Example

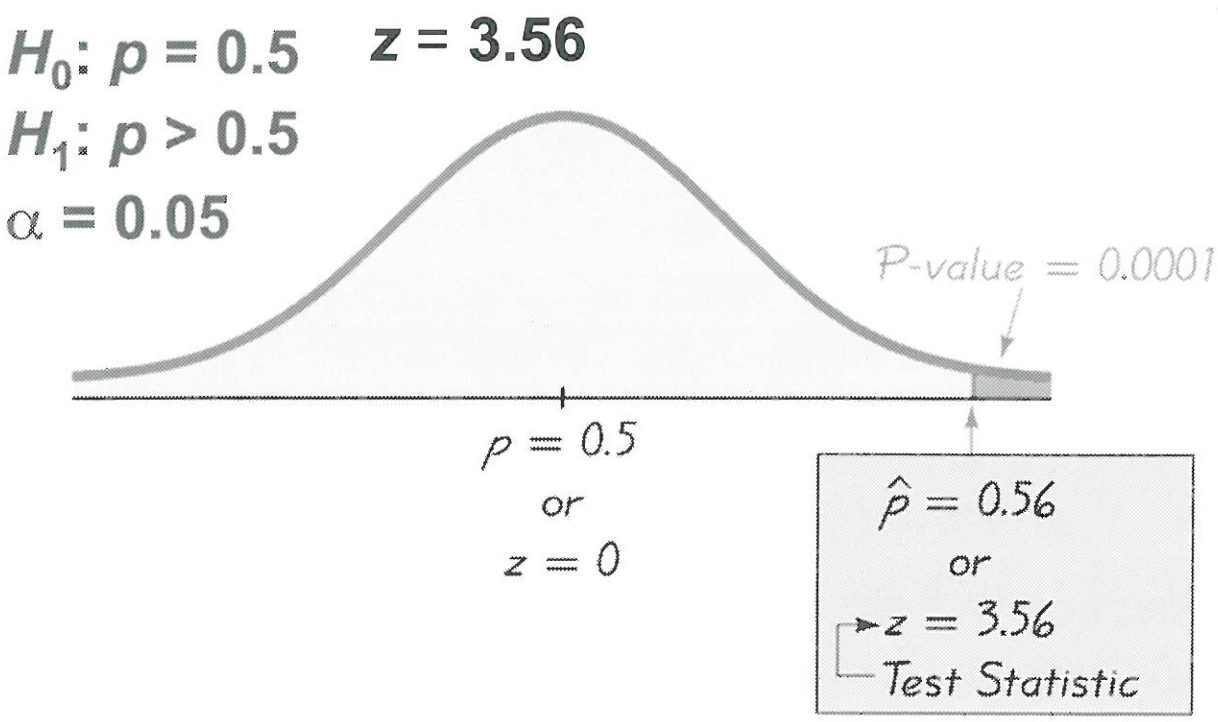
- I en spørreundersøkelse oppgir 56% av 880 studenter at de bruker lesesalen ukentlig
- Studentavisa OUTSIDE skriver i en overskrift at flere enn halvparten av BI studentene bruker lesesalen ukentlig
- Har OUTSIDE grunnlag for denne påstanden?
- $H_0 : p = 0.5$ og $H_1 : p > 0.5$. Høyresidig test

Vi bruker p-verdi metoden:

1. $np = 880 \cdot 0.5 = 440 \geq 5$ and $n(1-p) = 880 \cdot 0.5 = 440 \geq 5$
2. $\hat{p} = 0.56$ Testobservator er $z = \frac{0.56-0.5}{\sqrt{\frac{0.5(1-0.5)}{880}}} = 3.56$
3. Tabel A2: Arealet til venstre for $z = 3.56$ er 0.9999
4. p-verdien er $1 - 0.9999 = 0.0001$, mye lavere enn $\alpha = 0.05$
5. OUTSIDE har grunnlag for påstanden sin

Lesesal eksempel

p-verdien er svært liten, så nullhypotesen forkastes.



Lesesal eksempel; Tradisjonell metode

Example

Vi kan også bruke tradisjonell metode:

1. En høyrehalet test, så $\alpha = 0.05$ forkastningsområdet ligger i høyre hale
2. Den kritiske verdien er da $z = 1.645$
3. $\hat{p} = 0.56$ Testobservator er

$$z = \frac{0.56 - 0.5}{\sqrt{\frac{0.5(1-0.5)}{880}}} = 3.56$$

4. Siden testobservatoren $z = 3.56$ er større enn den kritiske verdien $z = 1.645$, så forkaster vi H_0
5. Det er tilstrekkelig grunnlag til å si at andelen er over 0.5

Example

Vi kan også bruke tradisjonell metode:

1. En høyrehalet test, så $\alpha = 0.05$ forkastningsområdet ligger i høyre hale
2. Den kritiske verdien er da $z = 1.645$
3. $\hat{p} = 0.56$ Testobservator er

$$z = \frac{0.56 - 0.5}{\sqrt{\frac{0.5(1-0.5)}{880}}} = 3.56$$

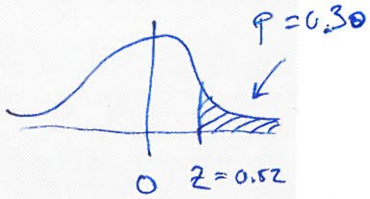
4. Siden testobservatoren $z = 3.56$ er større enn den kritiske verdien $z = 1.645$, så forkaster vi H_0
5. Det er tilstrekkelig grunnlag til å si at andelen er over 0.5

Bank2008.jmp

Example

1. En leder i Sparebank1 påstår at andelen kunder med universitetsutdannelse er mer enn 50 %
2. I stikkprøven har 48 av 91 universitetsutdannelse
3. $H_0 : p = 0.5$ versus $H_1 : p > 0.5$. Vi tester på $\alpha = 0.05$ nivået
4. $\hat{p} = 0.527$ Testobservator er

$$z = \frac{0.527 - 0.5}{\sqrt{\frac{0.5 \cdot 0.5}{91}}} = 0.52$$



5. Siden testen er ensidig, er p-verdien arealet i halen fra $z = 0.52$ og oppover Tabell A2: 0.70. Den kritiske verdien er da $1 - 0.7 = 0.3$
6. Siden p-verdien er større enn $\alpha = 0.05$, konkluderer vi med at det ikke er tilstrekkelig grunnlag til å forkaste H_0

$p = 0.30$ er ikke mindre enn $\alpha = 0.05$

$\Rightarrow$ vi forkaster ikke H_0

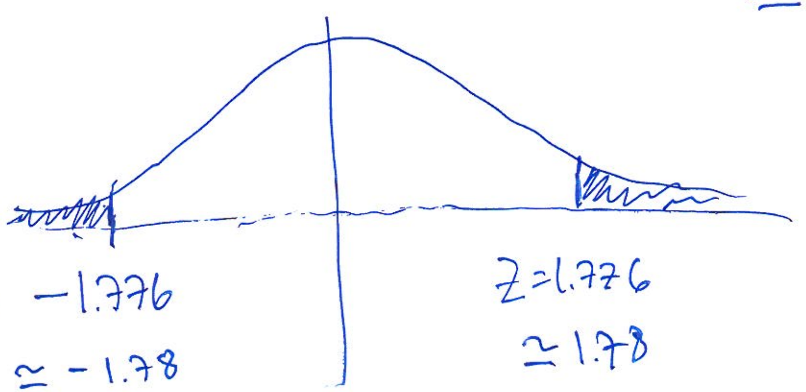
(beholder H_0)

$H_0: p = 0.265 \qquad H_1: p \neq 0.265$

$$\hat{p} = \frac{186}{784} \simeq 0.237 \Rightarrow z = \frac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{n}}} = \frac{0.237 - 0.265}{\sqrt{\frac{0.265 \cdot 0.735}{784}}}$$

1-0.265

$$\simeq -1.776$$



p = summen av arealene

$= 2 \cdot 0.0375 = \underline{0.075} = 7.5\%$

$p = 7.5\% < \alpha = 10\% \Rightarrow$ Forkaster H_0

$\boxed{p \neq 0.265}$

z-testen for én andel

Example

- Stortingsvalget 2009: FrP fikk 26.5 % i Rogaland
- 2010 meningsmåling: av 186 av 784 ville stemt FrP
- Er det en signifikant forskjell fra 2009? Signifikansnivå $\alpha = 0.1$

Svar

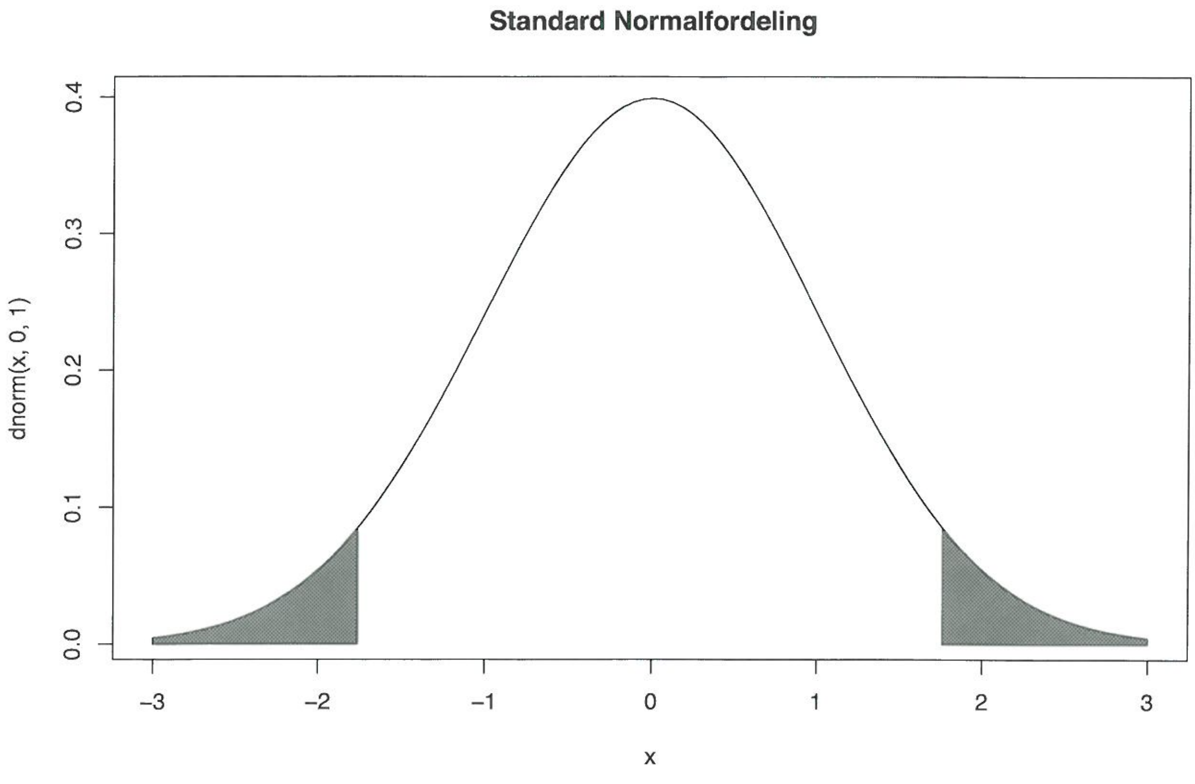
Tosidig test med hypotesene $H_0 : p = 0.265 \quad \text{vs} \quad H_1 : p \neq 0.265$.
Testobservator er

$$z = \frac{\frac{186}{784} - 0.265}{\sqrt{\frac{0.265(1-0.265)}{784}}} = -1.76$$

Tabell A2: arealet til venstre for $z = -1.76$ er 0.0392 så p-verdien er $2 \cdot 0.0392 = 0.0784$. p-verdien mindre enn 0.1: Det er grunnlag til å hevde at FrP sin oppslutning er endret

FrP eksempel

p-verdien er 0.0784, dvs. arealet i de to halene. Siden signifikansnivået var satt til $\alpha = 0.1$ så forkastes H_0.



Hypotesetest for gjennomsnittet

Forutsetninger

Vi forutsetter at:

- Stikkprøven er tilfeldig
- Dataene er normalfordelt, eller at $n > 30$

Hypotesene

- Nullhypotesen er at gjennomsnittet i populasjon μ er en spesifikk verdi $H_0: \mu = \mu_0$
- Alternativhypotesen kan være tosidig ($\neq$), eller ensidig ($<$ eller $>$) $H_1: \mu \neq \mu_0$ eller $\mu > \mu_0$ eller $\mu < \mu_0$

Testobservator

$\bar{x}$ = gjennomsnitt i utvalget

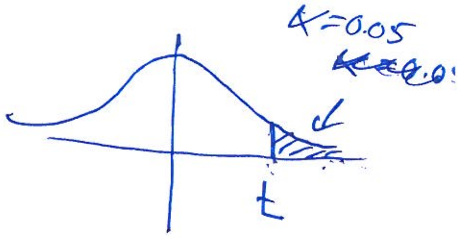
$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}}$$

μ: "riktig gjennomsnitt" = gjennomsnitt i H_0. = μ_0

er t-fordelt med $n - 1$ frihetsgrader. Kritiske verdier i Tabell A3

s = std. avvik i utvalget

Eks: $\boxed{\begin{array}{l}\alpha = 0.05 \\ n = 12\end{array}}$ i ensidige hypotese for μ



Kritisk t-verdi: $t_\alpha = 1.796$

$t > 1.796 \Rightarrow$ forkaster H_0

Tabell A3

- For hypotesetest om en **andel** er *p*-verdi metoden enklest. Tabell A2 gir *p*-verdier.
- For hypotesetest om **gjennomsnitt** er det bedre med tradisjonell metode, der testingen skjer med kritiske verdier
- Tabell A3 gir kritiske verdier og forkastningsområdet
- Dersom testobservatoren er mer ekstrem enn den kritiske verdien, så forkastes H_0

Table A-3 Finding *P*-Values from Table A-3

	Area in One Tail				
	0.005	0.01	0.025	0.05	0.10
	Area in Two Tails				
Degrees of Freedom	0.01	0.02	0.05	0.10	0.20
•					
•					
•					
11	3.106	2.718	2.201	1.796	1.363
	•				
	•				

Hypotesetest for gjennomsnittet

Example

Tenk deg følgende 3 situasjoner. Bruk tabell A3 til å avgjøre om H_0 forkastes

1. Venstresidig test. $n = 12$, $\alpha = 0.05$ og testobservatoren er $t = -2.01$
2. En høyresidig test med $n = 12$, $\alpha = 0.1$ og $t = 1.22$
3. En tosidig test med $n = 12$ og $\alpha = 0.01$ og $t = -3.45$

Svar

1. Kritisk verdi er $t = 1.796$. Forkastningsområdet er fra -1.796 og ned. Vår t-verdi ligger her. Konklusjon: Vi forkaster H_0
2. Kritisk verdi er $t = 1.363$. Dvs at forkastningsområdet er fra 1.363 og opp. Vår t-verdi ligger ikke her. Behold H_0
3. Kritisk verdi $t = 3.106$.Forkastningsområdet er utenfor ± 3.106. Forkaster H_0

Example

- På side 9 så vi at p-verdien var 0.023 for en venstresidig test av $H_0 : \mu = 7.0$ versus $H_1 : \mu < 7.0$. Vi brukte JMP til å finne det ut.
- Vi skal nå gjøre testen uten JMP, ved å bruke A3. Vi bruker $\alpha = 0.05$
- Postbanken har $\overline{x} = 6.410$ og $s = 2.753$ for 91kunder
- Testobservator er da $t = \frac{6.41-7.0}{2.753/\sqrt{91}} = -2.02$
- Tabell A3 med $91 - 90$ frihetsgrader gir da kritisk verdi 1.662
- Vi forkaster H_0: Det er grunnlag for å hevde at $\mu < 7$ i populasjonen

	t Test
Test Statistic	-2.0181
Prob > |t|	0.0466*
Prob > t	0.9767
Prob < t	0.0233*

Arealet er p-verdien. JMP gir at dette er 0.023.