Hypotesetesting, to utvalg (Kapitel 8)

Medisinsk statistikk 2009

http://folk.ntnu.no/slyderse/medstat/medstatI_h09.html

1

To-utvalgstest (def 8.1) vs ett-utvalgstest:

- To-utvalgstest: Sammenlikne den underliggende parameter for to forskjellige grupper, hvor verdiene i begge gruppene er ukjent.
- Ett-utvalgstest: Sammenlikne den underliggende parameter i en gruppe med en kjent verdi (f.eks 0 eller et kjent befolkningsgjennomsnitt)

2

Eksempel 8.2

- Er det en sammenheng mellom bruk av p-pille og blodtrykk?
- Flere studiedesign er mulige

3

Longitudinell studie (oppfølgingsstudie) - eqn 8.1

- Identifiser en gruppe ikke-gravide kvinner i fruktbar alder, som ikke bruker p-pillen. Mål blodtrykk (*baseline)*
- Etter 1 år: Identifiser en studiegruppe som ikke har vært gravide i perioden, og som nå bruker p-pillen.
- Mål blodtrykk i studiegruppen.
- Sammenlikne verdiene ved 1 år og baseline

4

Tverrsnittsstudie (cross-sectional study) - eqn 8.2

- Identifiser en gruppe som bruker, og en gruppe som ikke bruker, p-pillen blant ikke-gravide kvinner i fruktbar alder
- Sammenlikne blodtrykk mellom de to gruppene

5

Matchede par (Eks 8.6)

- Er det forskjellig fertilitet for p-pille brukere og pessar-brukere?
- Gruppe 1 består av 20 p-pille brukere.
- For hver kvinne i gruppe 1 identifiseres en pessar-bruker med samme alder (innen 5 år), rase, paritet, sosio-økonomisk status.
- Registrere tid til graviditet.

6

Parede versus uavhengige utvalg - forskjellig metode

- To utvalg er paret hvis hver observasjon i første utvalg er relatert til en bestemt observasjon i andre utvalg (f.eks longitudinell studie eller matchede par)
- To utvalg er uavhengige hvis observasjonene i første utvalg ikke er relatert til observasjonene i andre utvalg (f.eks tverrsnittsstudie)

7

Matchede par. Eksempel fra Box, Hunter & Hunter: Statistics for Experimenters 2nd ed. (2005)

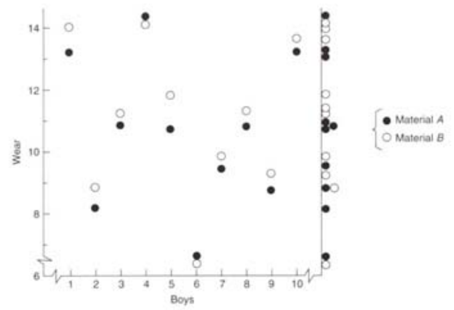


Figure 3.7. Data on two different materials A and B, used for making soles of boys' shoes.

8

Paret t-test eller konfidensintervall:

- For hvert par av observasjoner, regn ut differansen $d = x_2 - x_1$
- Forventet differanse er $\Delta = E(D)$
- H_0: $\Delta = 0$ mot H_1: $\Delta \neq 0$ (evt >0 eller <0)
- Gjennomfør en ett-utvalgs t-test eller beregn konfidensintervall for Δ basert på differansene $d_1, d_2, \ldots, d_n$

9

Repetisjon:

Hvis $X_1, X_2, ..., X_n$ er uavhengige $N(\mu, \sigma^2)$:

Da er $Z = \dfrac{\overline{X} - \mu}{\sigma / \sqrt{n}} \sim N(0,1)$.

Hvis σ er ukjent brukes $S^2 = \dfrac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2 = \dfrac{1}{n-1}\left[\left(\sum_{i=1}^{n} X_i^2\right) - n\overline{X}^2\right]$

Da er $T = \dfrac{\overline{X} - \mu}{S / \sqrt{n}} \sim t_{n-1}$

Z eller T brukes til å sette opp en hypotesetest eller konfidensintervall for μ.

Hvis n er "stor" så er T tilnærmet $N(\mu, \sigma^2)$

10

TABLE 8.1 Systolic blood-pressure levels (mm Hg) in 10 women while not using (baseline) and while using (follow-up) oral contraceptives

i	SBP level while not using OC's (x_{i1})	SBP level while using OC's (x_{i2})	d_i*
1	115	128	13
2	112	115	3
3	107	106	−1
4	119	128	9
5	115	122	7
6	138	145	7
7	126	132	6
8	105	109	4
9	104	102	−2
10	115	117	2

*$d_i = x_{i2} - x_{i1}$

11

Eksempel 8.5 (Tabell 8.1)

n=10, $\bar{d} = 4.80$, s^2=20.85=4.566^2

Tosidig test, t=3.32

Finner 0.001 < p < 0.01 vha Tabell 5 i Appendix

EXCEL: =TDIST(3,32;9;2) gir verdien p=0.00894

12

95% konfidensintervall for Δ:

$$\bar{d} \mp t_{\underbrace{n-1}_{9},\underbrace{1-\alpha/2}_{0.975}} s/\sqrt{n}$$

$$4.8 \mp 2.262 \times 4.566/\sqrt{10}$$

$$= 4.80 \mp 3.27$$

dvs 1.53 til 8.07 (mmHg)

t-test og konfidensintervall for to uavhengige utvalg

- n_1 observasjoner, antas uavh. $N(\mu_1, \sigma_1^2)$
- n_2 observasjoner, antas uavh. $N(\mu_2, \sigma_2^2)$
- H_0: $\mu_1 = \mu_2$ mot H_1: $\mu_1 \neq \mu_2$
- Ekvivalent: H_0: $\mu_1 - \mu_2 = 0$ mot H_1: $\mu_1 - \mu_2 \neq 0$
- Antar foreløpig lik varians, $\sigma_1^2 = \sigma_2^2 = \sigma^2$

Estimator for $\mu_1 - \mu_2$: $\bar{X}_1 - \bar{X}_2 \sim N\left(\mu_1 - \mu_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right)$

Altså: $\dfrac{\bar{X}_1 - \bar{X}_2 - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0,1)$

Hvis $\sigma_1 = \sigma_2 = \sigma$ så er $\dfrac{\bar{X}_1 - \bar{X}_2 - (\mu_1 - \mu_2)}{\sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim N(0,1)$

Men σ^2 er ukjent og estimeres ved "pooled estimate of the variance":

$$S^2 = \frac{1}{n_1 + n_2 - 2}\left[\sum_{i=1}^{n_1}(X_{i1} - \bar{X}_1)^2 + \sum_{i=1}^{n_2}(X_{i2} - \bar{X}_2)^2\right]$$
$$= \frac{n_1 - 1}{n_1 + n_2 - 2} S_1^2 + \frac{n_2 - 1}{n_1 + n_2 - 2} S_2^2$$

Vi bruker at

$$\frac{\bar{X}_1 - \bar{X}_2 - (\mu_1 - \mu_2)}{S\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t_{n_1 + n_2 - 2}$$

Eks 8.9

Cardiovascular Disease, Hypertension
Suppose a sample of eight 35- to 39-year-old nonpregnant, premenopausal OC users who work in a company are identified who have mean systolic blood pressure of 132.86 mm Hg and sample standard deviation of 15.34 mm Hg. A sample of twenty-one 35- to 39 year-old nonpregnant, premenopausal non-OC users are similarly identified who have mean systolic blood pressure of 127.44 mm Hg and sample standard deviation of 18.23 mm Hg. What can be said about the underlying mean difference in blood pressure between the two groups?

Eks 8.10 – lik varians

$n_1=8$, $\bar{x}_1=132.86$, $s_1=15.34$

$n_2=21$, $\bar{x}_2=127.44$, $s_2=18.23$

H_0: $\mu_1 - \mu_2 = 0$

$\bar{x}_1 - \bar{x}_2 = 5.42$, $\quad s^2 = \frac{7}{27}15.34^2 + \frac{20}{27}18.23^2 = 307.18 = 17.53^2$

$$t = \frac{5.42 - (-0)}{17.53\sqrt{\frac{1}{8} + \frac{1}{21}}} = 0.74$$

Frihetsgrader: 8+21-2=27, forhast H_0 på 5% nivå hvis |0.74|>2.052

P-verdi f.eks EXCEL TFORDELING(0,74;27;2)=0.47

Eks 8.11 Lik varians

95% konfidensintervall

$\Pr(-t \le T \le t) = 1-\alpha$, hvor $t = t_{\underbrace{n_1+n_2-2}_{27},\underbrace{1-\alpha/2}_{0.975}} = 2.052$

$$-t \le \frac{\bar{X}_1 - \bar{X}_2 - (\mu_1 - \mu_2)}{S\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \le t$$

løser mhp $\mu_1 - \mu_2$: $-9.52 \le \mu_1 - \mu_2 \le 20.36$

19

to uavhengige utvalg, ulik varians

- n_1 observasjoner, antas uavh. $N(\mu_1, \sigma_1^2)$
- n_2 observasjoner, antas uavh. $N(\mu_2, \sigma_2^2)$
- H_0: $\mu_1 = \mu_2$ mot H_1: $\mu_1 \ne \mu_2$
- Ulik varians, $\sigma_1^2 \ne \sigma_2^2$

20

To utvalg, $\sigma_1 \ne \sigma_2$: Vi bruker ”Satterthwaite’s metode”:

$$\frac{\bar{X}_1 - \bar{X}_2 - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}} \sim t_{d'} \quad \text{tilnærmet,}$$

hvor antall frihetsgrader d beregnes ut fra n_1 ,S_1, n_2 ,S_2.

$$d' = \frac{(S_1^2/n_1 + S_2^2/n_2)^2}{(S_1^2/n_1)^2/(n_1-1) + (S_2^2/n_2)^2/(n_2-1)}$$

21

Eks 8.1 (utvidet) – Ulik varians

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}} = 0.81$$

$d' = 15.04 \quad (d'' = 15)$

$p - verdi = 0.43$

22

to uavhengige utvalg, test for ulik varians

- n_1 observasjoner, antas uavh. $N(\mu_1, \sigma_1^2)$
- n_2 observasjoner, antas uavh. $N(\mu_2, \sigma_2^2)$
- H_0: $\sigma_1^2 = \sigma_2^2$ mot H_1: $\sigma_1^2 \ne \sigma_2^2$
- Ekvivalent: H_0: $\sigma_1^2/\sigma_2^2 = 1$ mot H_1: $\sigma_1^2/\sigma_2^2 \ne 1$
- Forkast H_0 hvis S_1^2/S_2^2 avviker “mye” fra 1
- Under H_0: $S_1^2/S_2^2 \sim F_{n1-1,\, n2-1}$ (Fisherfordelt med n_1-1 og n_2-1 frihetsgrader)
- SPSS bruker “Levene’s test” i steden for “Fisher’s test”

23

Eksempel 8.16

- $F = S_2^2/S_1^2 = 18.23^2/15.34^2 = 1.41$
- Forkast H_0: $\sigma_2^2/\sigma_1^2 = 1$ på nivå α=0.05 hvis
  - $F > F_{20,7,0.975} = 4.47$ (FINV(0,025;20;7) i EXCEL)
  - eller
  - $F < F_{20,7,0.0.025} = 0.33$ (FINV(0,975;20;7) i EXCEL)
- Alternativt: p-verdi = 2 * 0.335 = 0.67 (FDIST(1,41;20;7))
- Konklusjon: Vi forkaster ikke H_0

24

Equation 8.14

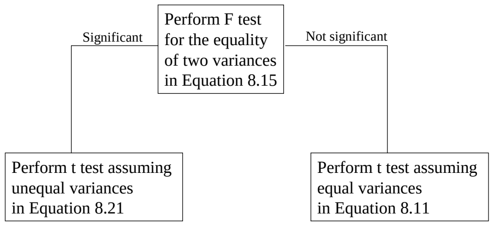
Nedre p-persentil i en F-fordeling med d_1 og d_2 frihetsgrader er den inverse av den øvre p-persentilen i en F-fordeling med d_1 og d_2 frihetsgrader:

$$F_{d_1,d_2,p} = 1 / F_{d_2,d_1,1-p}$$

(Nyttig hvis tabellen bare inneholder øvre persentiler)

25

Rosner, Figure 8.10
Strategy for testing the equality of means in two independent, normally distributed samples



26

MEN:

Navidi: “Statistics for Engineers and Scientists”, 2006, page 343-344:

Don’t Assume the Population Variances are Equal Just Because the Sample Variances are Close

27

” … the expression assuming equal variances requires that the population variances be equal, or nearly so. In situations where the *sample* variances are nearly equal, it is tempting to assume that the population variances are nearly equal as well. However, when the sample sizes are small, the sample variances are not necessarily good approximations to the population variances. Thus it is possible that the sample variances be close even when the population variances are fairly far apart. In general, population variances should be assumed equal only when there is knowledge about the processes that produced the data that justifies this assumption. “

28

“ … the expression not assuming equal variances produces good results in almost all cases, whether the population variances are equal or not. (Exceptions can occur when the sample sizes are very different.) Therefore, when in doubt, use the expression not assuming equal variances.”

29

Altså: t-test eller konfidensintervall for differanse mellom forventningsverdene i to uavhengige, normalfordelte utvalg

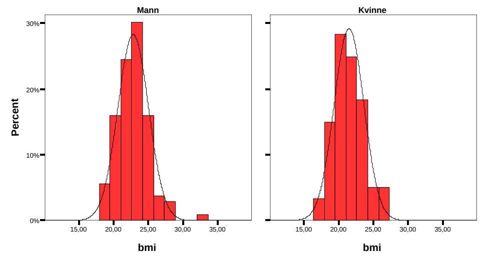
		Virkelig	
		lik varians	ulik varians
Du antar	lik varians (eqn 8.11)	korrekt	gir feil svar
	ulik varians (eqn 8.21)	tilnærmet samme svar som ovenfor	korrekt

Altså: Velg t-test for ulik varians, eller en ikke-parametrisk metode, hvis du er i tvil!

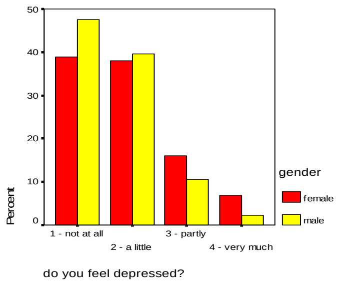
30

Hvis data ikke er normalfordelt:

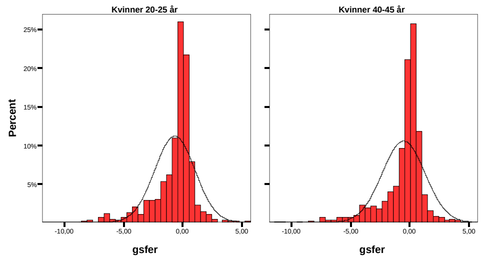
- t-tester fungerer brukbart ved begrenset variasjon i data
- t-tester er ubrukelige hvis mange sterkt avvikende verdier (“outliers”).
- Ikke-parametriske metoder er brukbare uansett.
- F-testen for sammenlikning av varians er lite robust mot avvik fra normalfordelingen.



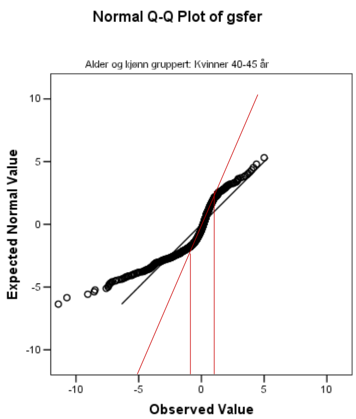
Nær normalfordeling - t-test er bra

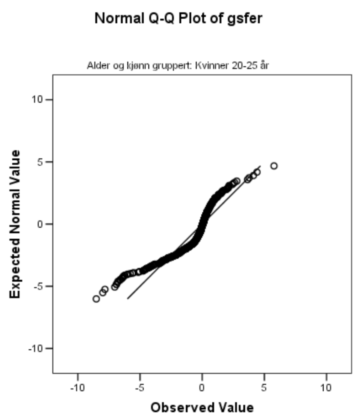


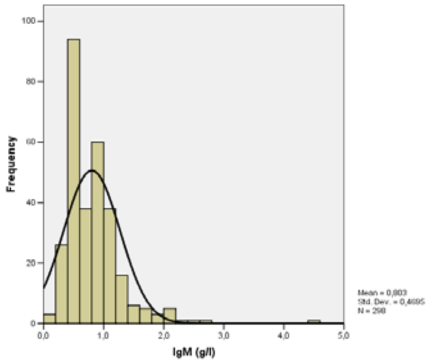
Begrenset variasjon i data.
T-test er brukbar - eller bruk ikke-parametriske metoder



T-test er ubrukelig - benytt ikke-parametriske metoder







Konsentrasjon av serum IgM (g/l) hos
298 friske barn, 6 mnd - 6 år gamle (Altman, 1991)

37