BI1363 HT 2020

t-test för två oberoende stickprov

Oktober 2020

Adam Flöhr, BT, SLU

t-test för två oberoende stickprov

Motsvarar *Biometri*, kap 7

I korthet

t-test för två stickprov testar om två grupper har samma populationsmedelvärde

Vid **matchade** stickprov beräknar man differensen mellan två observationer och tillämpar ett t-test på differensserien

Vid **oberoende** stickprov baseras testet på skillnaden i skattade medelvärden

Bakgrunden till hypotestest

Slumpmässiga variationer *döljer* underliggande fenomen

Statistiska metoder gör det möjligt att dra slutsatser trots slumpmässig variation

Den vanligaste metoden är ett *hypotestest*

Vid ett hypotestest jämförs en hypotes med observerad data

Ett resultat är *signifikant* om observationerna är osannolika givet hypotesen

Hypotestestens steg

Formell gång för att ställa observation mot hypotes

1. Hypoteser
2. Testfunktion
3. Testfördelning
4. p-värde (beräkning eller uppskattning)
5. Slutsats

HTTPS

Hypotestest för två stickprov: matchat och oberoende

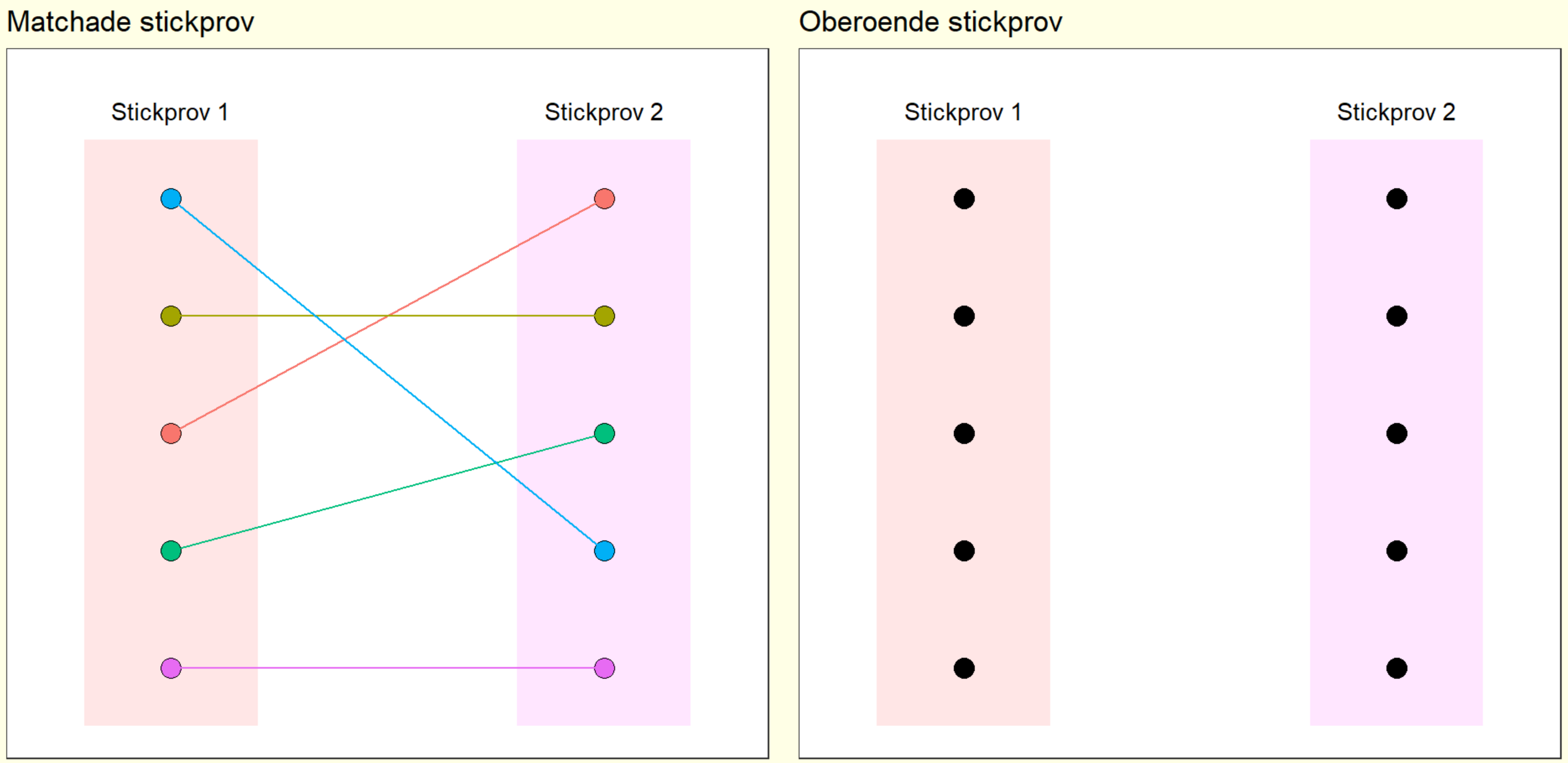
Har en utfallsvariabel som är kontinuerlig och kan antas normalfördelad

Vill jämföra två grupper och drar ett stickprov ur respektive grupp

Två designval

Matchade stickprov - varje observation i grupp 1 är parad med *en* observation i grupp 2

Oberoende stickprov - observationerna är inte kopplade till varandra



Matchat eller oberoende?

1. Jämför två jordars effekt på biomassa i sallad. 10 replikat av varje, slumpmässigt utplacerade på ett bord i växthuset
2. Planterar två ärtgenotyper på tio fält runtom i Skåne. Varje fält delat i två med en genotyp på respektive del
3. Jämför två foders effekt på tillväxt hos gris. Tio slumpmässigt utvalda grisar för respektive foder
4. Två havresorter testas genom årlig avkastning över en tioårsperiod
5. Likt tre, men nu tio syskonpar där den ena får ett foder och den andra ett annat
6. Tio individer var av två arter samlas in från vilt tillstånd. Plantorna placeras i växthus och testas för reaktion under värmerelaterad stress
7. Tio igelkottars blodtryck mäts före och efter intag av betablockerare

2, 4, 5, 7 är matchade. Övriga oberoende

Två matchade stickprov

Återför till ett prov

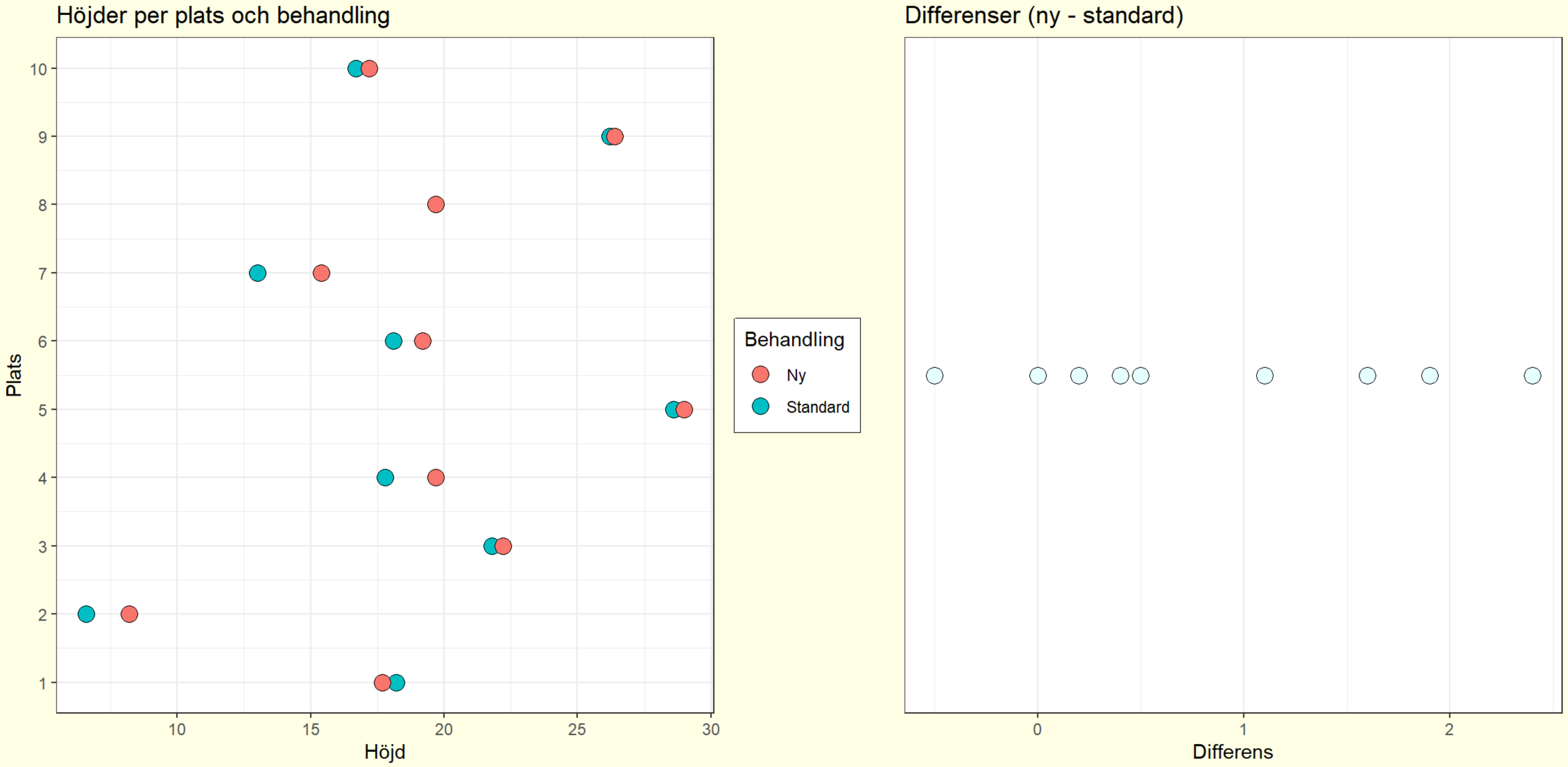
Vid två matchade stickprov kan man ta differensen och genomföra ett t-test för ett stickprov

För att testa om en ny behandling ger samma planthöjd planterar man på tio platser två plantor - en som får ny behandling, en som får en standardbehandling. För varje plats räknar man skillnaden i planthöjd mellan den nya metoden och standardmetoden. Resultat: ett medelvärde på 0.808 och en standardavvikelse på 0.945.

a. Avgör med ett test om den nya behandlingen ger högre plantor än standardmetoden.

Plats	Standard	Ny	Differens
1	18.2	17.7	-0.5
2	6.6	8.2	1.6
3	21.8	22.2	0.4
4	17.8	19.7	1.9
5	28.6	29.0	0.4
6	18.1	19.2	1.1
7	13.0	15.4	2.4
8	19.7	19.7	0.0
9	26.2	26.4	0.2
10	16.7	17.2	0.5

Matchade stickprov, illustration



Om vi har matchade stickprov kan vi beräkna differensen för varje par

Differensserien kan testas med ett t-test för ett stickprov

Om differensen är skild från noll finns det en skillnad mellan behandlingarna

Matchade stickprov, exempel

Hypoteser

Nollhypotes och alternativhypotes

$H_0 : \mu = 0$ (d.v.s ingen skillnad)

$H_1 : \mu > 0$ (ensidig mothypotes)

Testfunktion

Vanligt t-test, så samma testfunktion som innan

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$

Notera att n är antalet matchade par

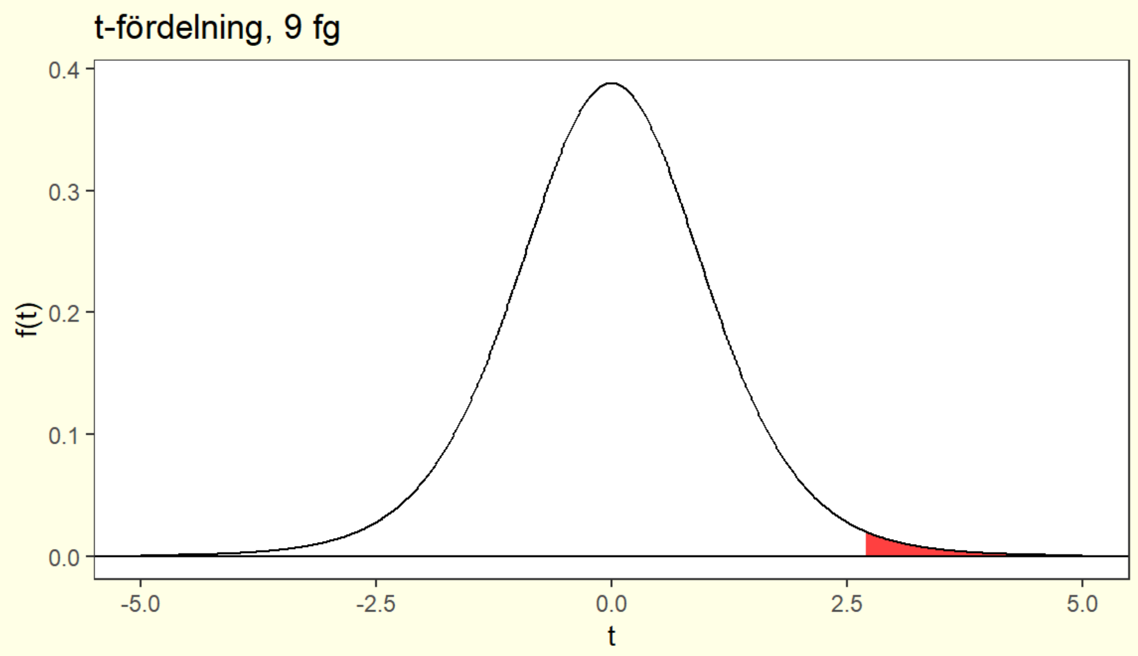
Sätter in värden i testfunktionen

$$t = \frac{0.808 - 0}{0.945/\sqrt{10}} = 2.704$$

Testfördelning

Om nollhypotesen stämmer ska t komma från en t-fördelning med $n-1$ frihetsgrader

p-värde



p-värdet beräknas som ytan bortom t i en t-fördelning med $n-1=9$ frihetsgrader

Eftersom alternativhypotesen är ensidig medtas enbart den ena svansen

Man kan beräkna p-värdet i valfritt datorprogram och få $p = 0.0171$

Eller titta i tabell 5, rad 9. Eftersom t ligger mellan 2.262 och 2.821 måste det motsvara ett ensidigt p-värde mellan 2.5 och 1 procent

	F(t)								
d.f.	0.75	0.90	0.95	0.975	0.99	0.995	0.9975	0.999	0.9995
9	0.703	1.383	1.833	2.262	2.821	3.250	3.690	4.297	4.781

Svar

p-värdet är mindre än 5 procent men större än 1 procent

Vi *förkastar* på 5-procentsnivån

Konfidensintervall

Konfidensintervall för matchade stickprov hanteras som ett stickprov på differenser

$$\bar{x} \pm t_{(1-0.05/2,n-1)} \frac{s}{\sqrt{n}}$$

$$= 0.808 \pm 2.262 \frac{0.945}{\sqrt{10}} = 0.808 \pm 0.676$$

eller $[0.132, 1.484]$

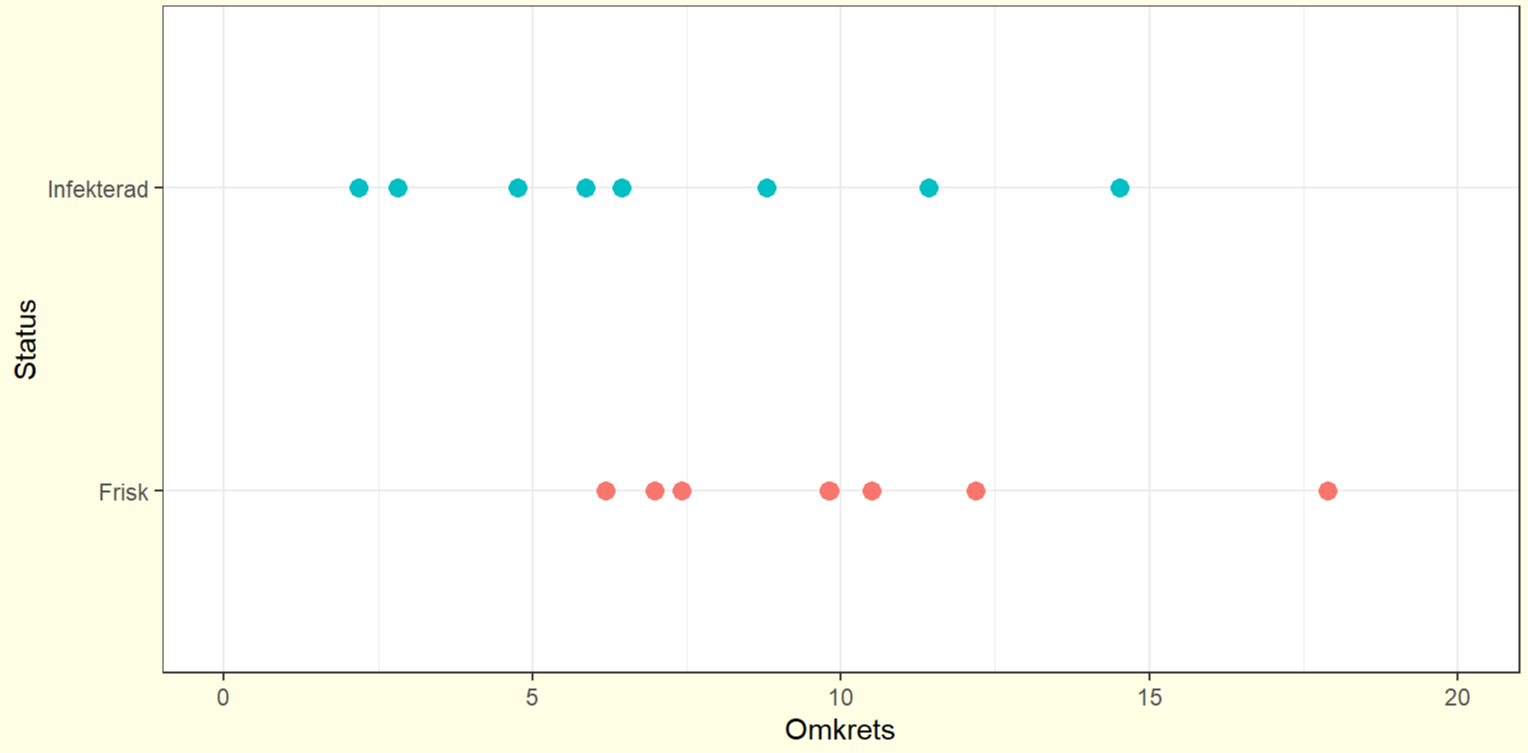
Två oberoende stickprov

Fortsatt i situationen med en kontinuerlig variabel som kan antas vara normalfördelad

Vill jämföra två grupper genom att testa om deras populationsmedelvärden är skilda

Vid två oberoende stickprov kan vi inte längre återföra det till ett stickprov

Men vi har samma princip och samma steg (https)



Två fördelningar med populationsmedelvärden μ_1 och μ_2

Vad är sannolikheten för den observerade skillnaden mellan grupperna om populationsmedelvärdena är lika?

t-test för två oberoende stickprov, schema

Hypoteser

$H_0 : \mu_1 = \mu_2$

$H_1 : \mu_1 \neq \mu_2$

Testfunktion

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{s^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

där $s^2 = \frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}$

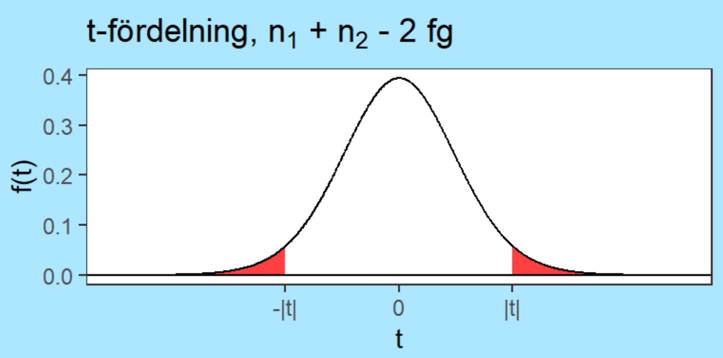
Testfördelning

Under nollhypotesen följer t en t-fördelning med $n_1 + n_2 - 2$ frihetsgrader

P-värde

P-värdet ges av arean bortom $|t|$ i testfördelningen

Vid handräkning uppskattas p-värdet genom att ställa $|t|$ mot ett tabellvärde



Svar

P-värdet ställs mot en förbestämd *signifikansnivå* (ofta 5 procent)

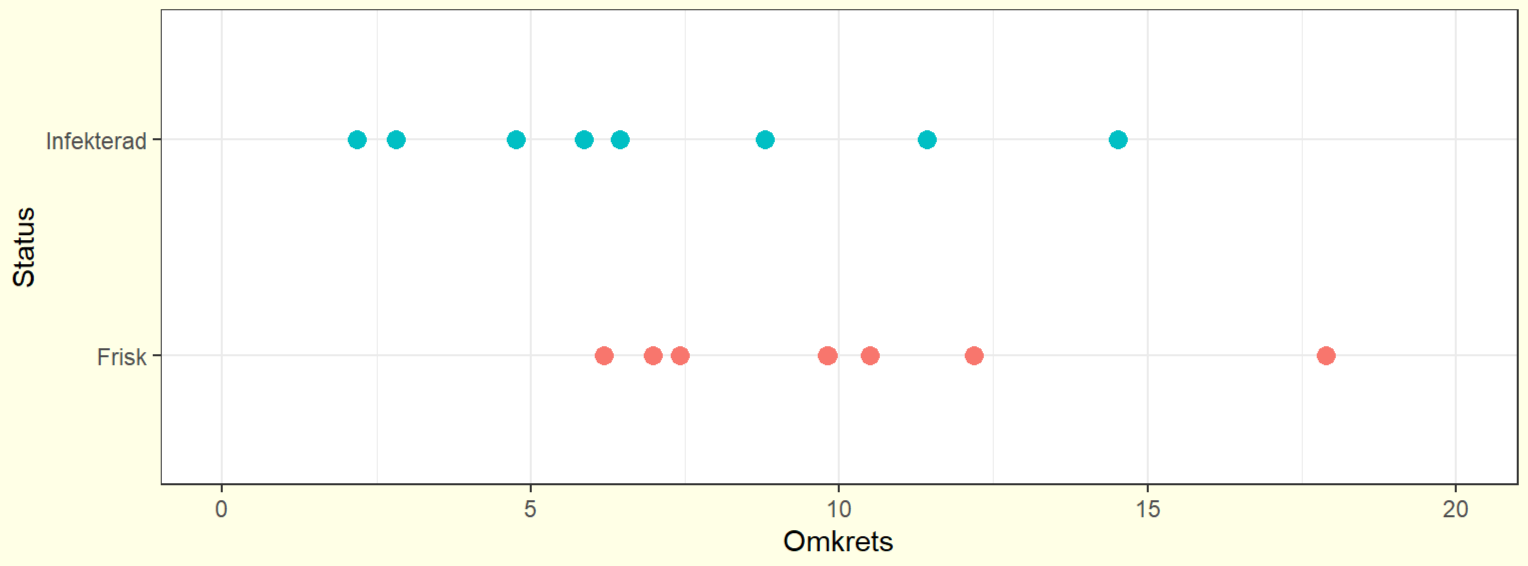
Vid ett lågt p-värde förkastas nollhypotesen

Vid ett högt p-värde förkastas ej nollhypotesen

Oberoende stickprov, exempel

Vi vill undersöka om en sjukdom hos björk påverkar trädens omkrets

Väljer slumpmässigt ut åtta friska träd och åtta sjuka träd



Skattningar från stickproven ges av

	n	Medel	s^2	s
Infekterat	8	7.1	18.1	4.25
Frisk	8	10.1	13.9	3.73

Testa för skillnad i omkrets. Använd signifikansnivån 1 procent.

t-test för oberoende stickprov

Hypoteser

Nollhypotes och alternativhypotes

$H_0 : \mu_1 = \mu_2$ d.v.s ingen skillnad

$H_1 : \mu_1 \neq \mu_2$

Testfunktion

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{s^2(\frac{1}{n_1} + \frac{1}{n_2})}}$$

där $s^2 = \frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}$

Sätter in värden i testfunktionen

$$s^2 = \frac{7 \cdot 18.1 + 7 \cdot 13.9}{8 + 8 - 2} = 16.0$$

Väldigt vanligt fel att räkna med standardavvikelser här

s^2 ska ligga mellan s_1^2 och s_2^2

$$t = \frac{7.1 - 10.1}{\sqrt{16(\frac{1}{8} + \frac{1}{8})}} = \frac{-3}{\sqrt{4}} = -1.5$$

Testfördelning

Om nollhypotesen stämmer kommer t från en t-fördelning med $n_1 + n_2 - 2 = 8 + 8 - 2 = 14$ frihetsgrader

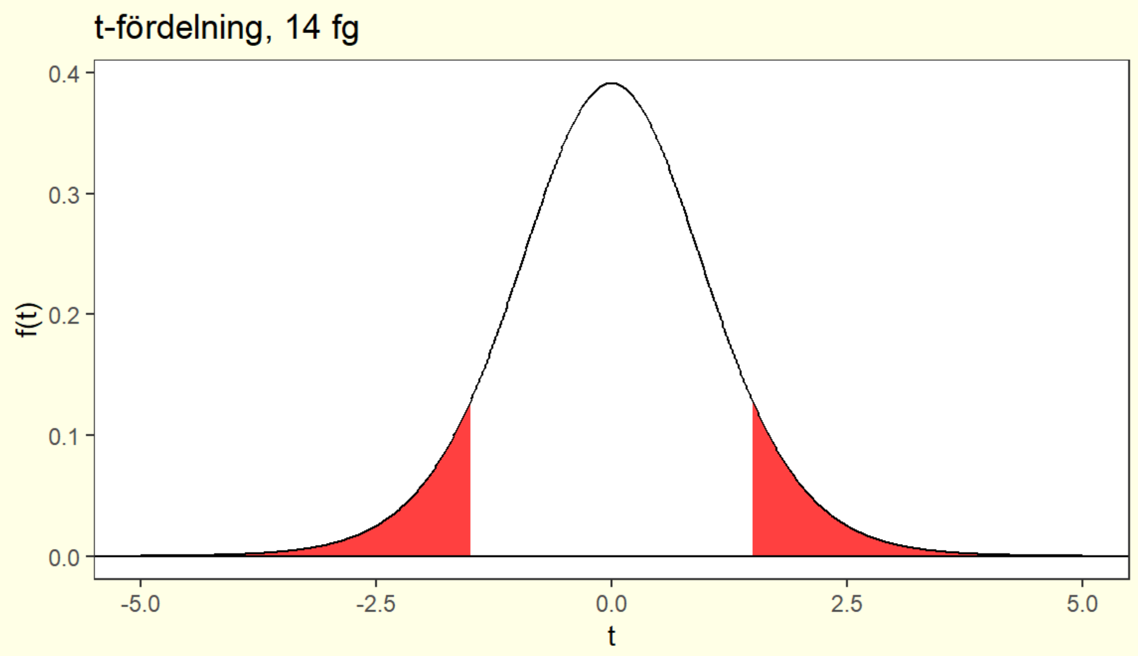
p-värde

p-värdet beräknas som ytan bortom $\boldsymbol{t}$ i en t-fördelning med $\mathbf{14}$ frihetsgrader

Eftersom alternativhypotesen är tvåsidig medtas bägge svansarna

Man kan beräkna p-värdet i valfritt datorprogram och få $\boldsymbol{p} = \mathbf{0.1558}$

Eller titta i tabell 5, rad 14. Eftersom $|\boldsymbol{t}| = \mathbf{1.5}$ är mindre än $\boldsymbol{t}_{(0.995,14)} = \mathbf{2.977}$ måste det motsvara ett tvåsidigt p-värde större än $\mathbf{1}$ procent



	F(t)								
d.f.	0.75	0.90	0.95	0.975	0.99	0.995	0.9975	0.999	0.9995
14	0.692	1.345	1.761	2.145	2.624	2.977	3.326	3.787	4.140

Svar

p-värdet är större än $\mathbf{1}$ procent

Vi förkastar ej på $\mathbf{1}$-procentsnivån

Vi kan ej påvisa någon säkerställd skillnad i diameter mellan friska och infekterade träd

Konfidensintervall

Vid två oberoende stickprov ges ett 95-procentigt konfidensintervall av

$$(\bar{x}_1 - \bar{x}_2) \pm t_{(1-0.05/2, n_1+n_2-2)} \sqrt{s^2 (\frac{1}{n_1} + \frac{1}{n_2})}$$

Från exempel

$$(10.1 - 7.1) \pm 2.145 \sqrt{16 (\frac{1}{8} + \frac{1}{8})}$$

$$= 3 \pm 2.145 \sqrt{4} = 3 \pm 4.29$$

eller $[-1.29, 7.29]$

Avslutande punkter

Antaganden

t-testerna för ett stickprov eller matchade stickprov bygger på

- normalfördelning
- oberoende observationer inom stickprovet

t-testet för oberoende stickprov bygger dessutom på

- oberoende observationer mellan stickprov
- lika varians

Brister i antaganden

t-testet sägs vara *robust* och används ofta i fall där normalfördelningsantagandet inte klart gäller t.ex. för heltalsdata

Skillnader i varianser kan testas och t-test för två stickprov kan justeras så att antagandet inte krävs

Slut