Tõenäosusteooria ja statistika II

Loeng 14

Liithüpoteeside kontrollimine

ehk mis saab siis, kui nullhüpoteese on mitu

Rakendused:

Kuidas tõestada olulist erinevust?

Kuidas tõestada nullhüpoteesi?

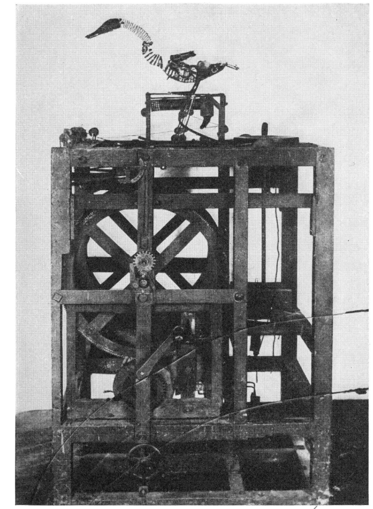
Märt Möls

Liithüpoteesid (*Composite hypothesis*)

Vanaisa masin juhuslike numbrite mehhaniseeritud tootmiseks

Vanaisa sõnul tekitas kas jaotusega N(3; 1) või N(7; 1) juhuslikke arve – ei mäletanud enam temagi, kumba Gaussi jaotust ta lõpuks masina ehitamisel kasutas.

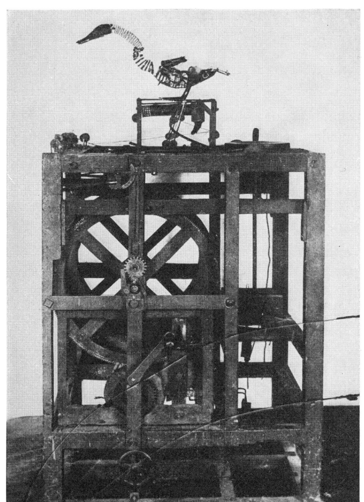
Kas masin on ikka veel töökorras ja töötab korrektselt?



Liithüpoteesid (*Composite hypothesis*)

H_0: E(X)=3 või E(X)=7

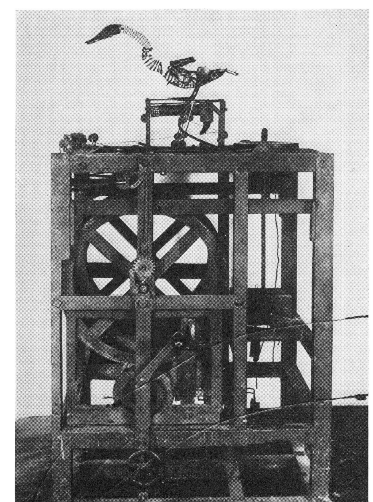
Kas masin on ikka veel töökorras ja töötab korrektselt?



Liithüpoteesid (*Composite hypothesis*)

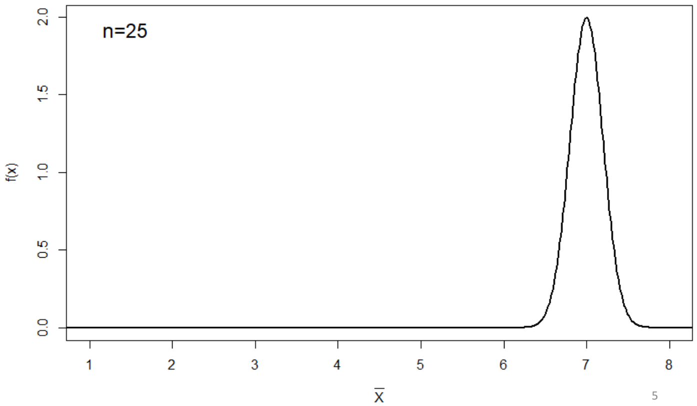
H_0: E(X)=3 või E(X)=7

Teststatistik: valimi keskmine



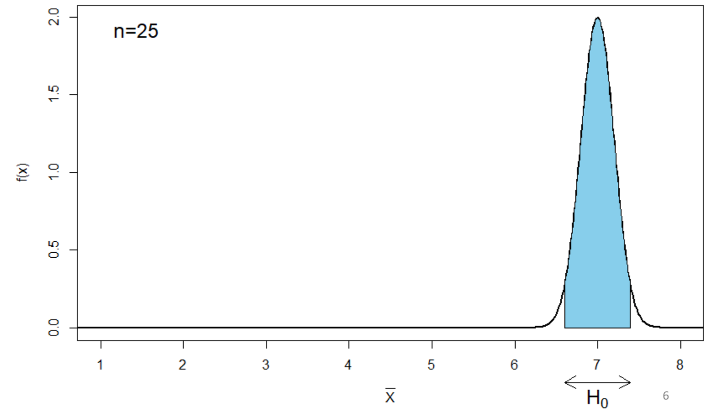
Liithüpoteesid (*Composite hypothesis*)

H_0: E(X)=3 või E(X)=7



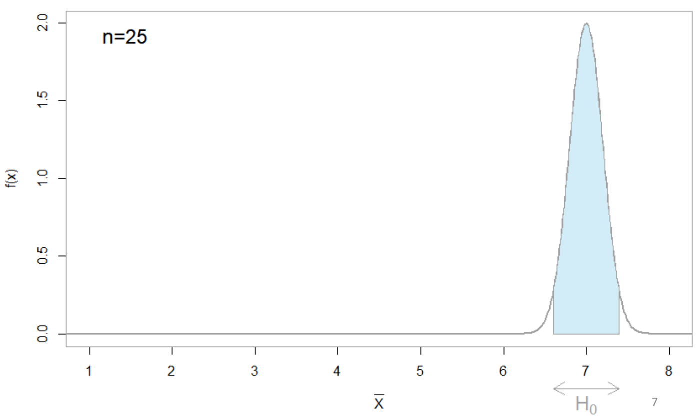
Liithüpoteesid (*Composite hypothesis*)

H_0: E(X)=3 või E(X)=7



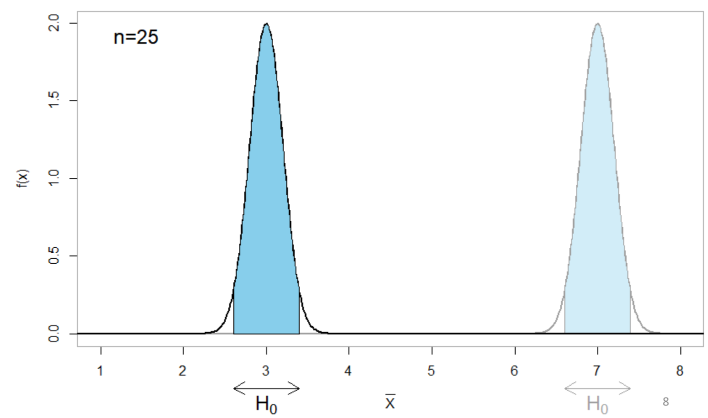
Liithüpoteesid (*Composite hypothesis*)

H_0: E(X)=3 või E(X)=7



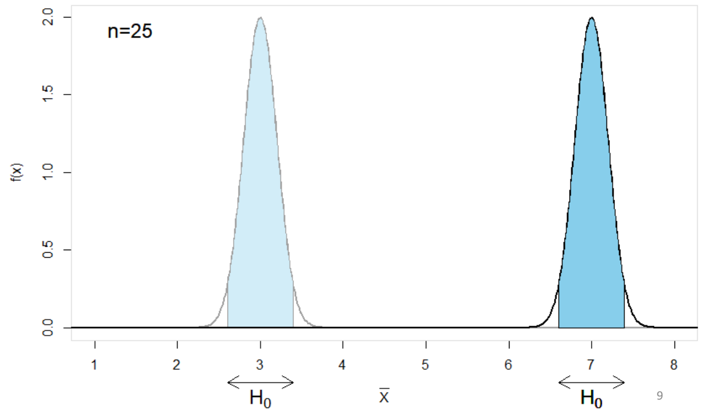
Liithüpoteesid (*Composite hypothesis*)

H_0: E(X)=3 või E(X)=7



Liithüpoteesid (*Composite hypothesis*)

H_0: E(X)=3 või E(X)=7



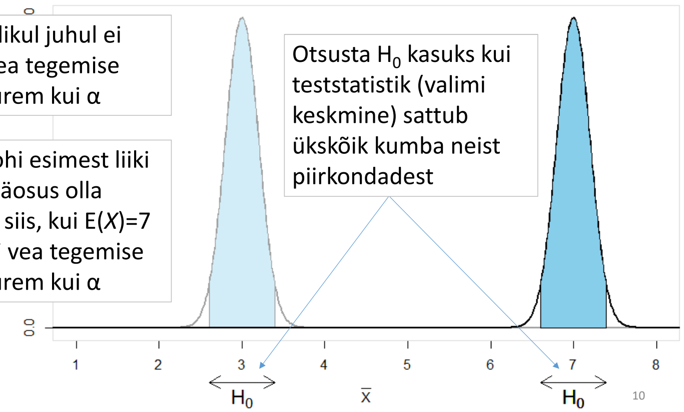
Liithüpoteesid (*Composite hypothesis*)

H_0: E(X)=3 või E(X)=7

Ka halvimal võimalikul juhul ei tohi esimest liiki vea tegemise tõenäosus olla suurem kui α

Kui E(*X*)=3 siis ei tohi esimest liiki vea tegemise tõenäosus olla suurem kui α ja ka siis, kui E(*X*)=7 ei tohi esimest liiki vea tegemise tõenäosus olla suurem kui α

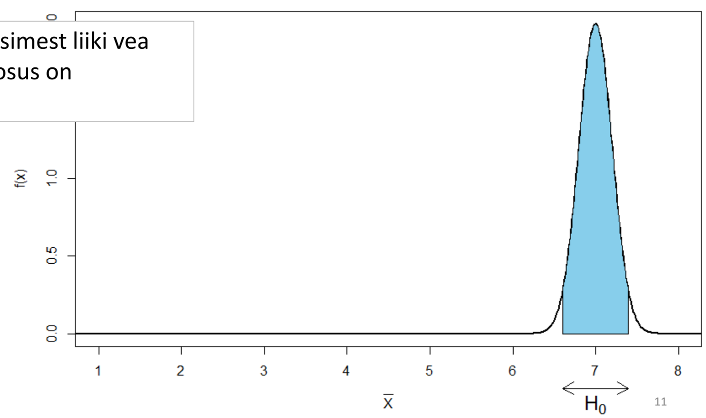
Otsusta H_0 kasuks kui teststatistik (valimi keskmine) sattub ükskõik kumba neist piirkondadest



Liithüpoteesid (*Composite hypothesis*)

H_0: E(X)=3 või E(X)=7

Kui E(X)=7, siis esimest liiki vea tegemise tõenäosus on 0,05.



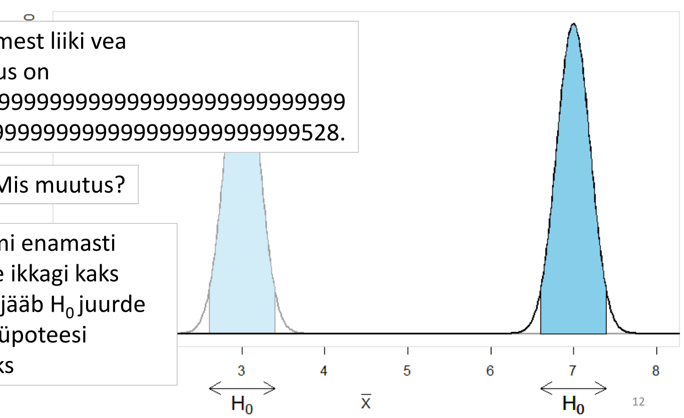
Liithüpoteesid (*Composite hypothesis*)

H_0: E(X)=3 või E(X)=7

Kui E(X)=7, siis esimest liiki vea tegemise tõenäosus on 0,04999999999999999999999999999999999999999999999999999999999999999999999999528.

Miks? Mis muutus?

Praktikas probleemi enamasti ignoreeritakse: tee ikkagi kaks testi, kui üks neist jääb H_0 juurde siis otsusta ka liithüpoteesi testimisel H_0 kasuks

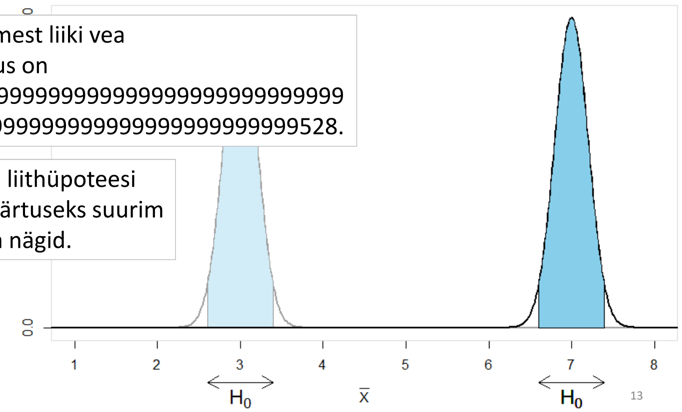


Liithüpoteesid (*Composite hypothesis*)

H_0: E(X)=3 või E(X)=7

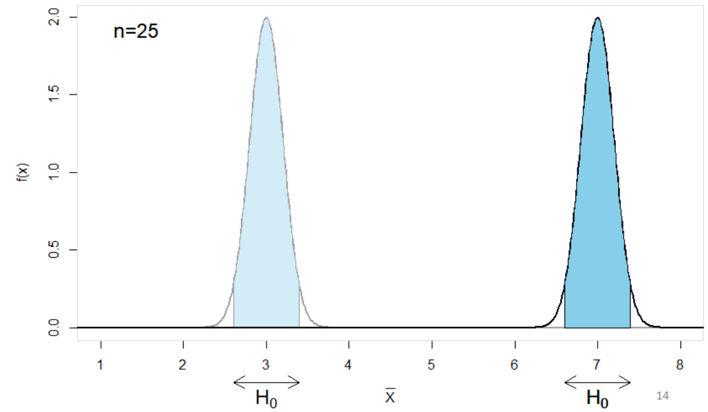
Kui E(X)=7, siis esimest liiki vea tegemise tõenäosus on 0,0499999999999999999999999999999999999999999999999999999999999999999999999999528.

Alternatiivselt: vali liithüpoteesi kontrollimisel p-väärtuseks suurim p-väärtustest mida nägid.



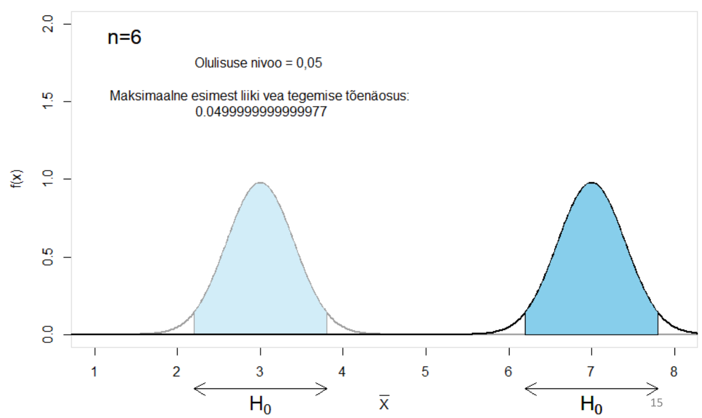
Liithüpoteesid (*Composite hypothesis*)

H_0: E(X)=3 või E(X)=7



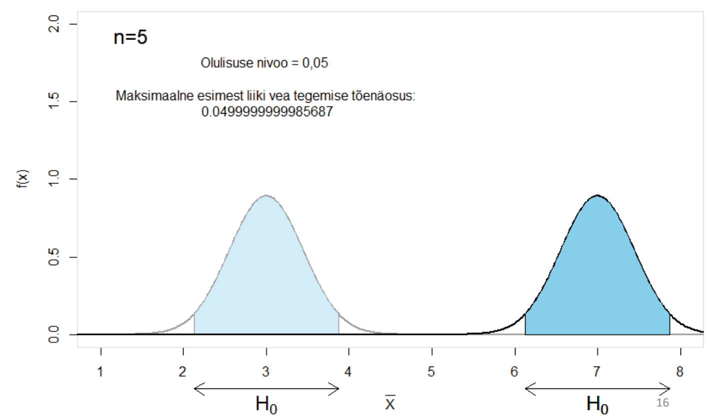
Liithüpoteesid (*Composite hypothesis*)

H_0: E(X)=3 või E(X)=7



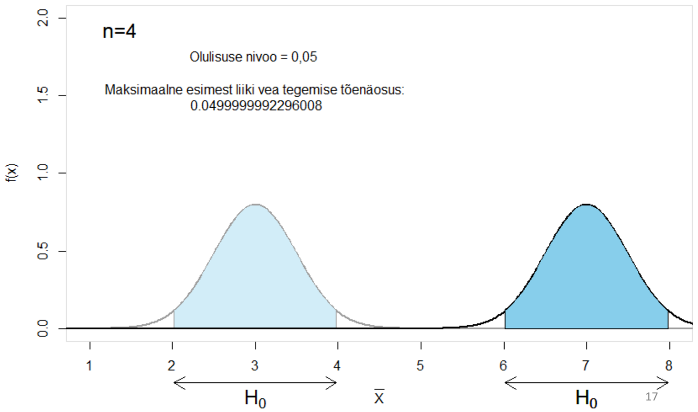
Liithüpoteesid (*Composite hypothesis*)

H_0: E(X)=3 või E(X)=7



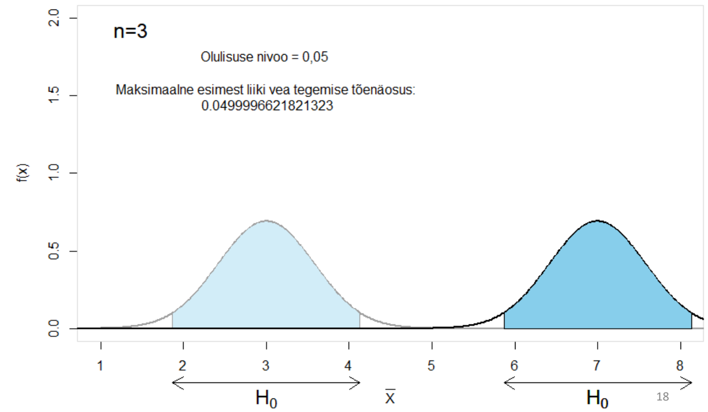
Liithüpoteesid (*Composite hypothesis*)

H_0: E(X)=3 või E(X)=7



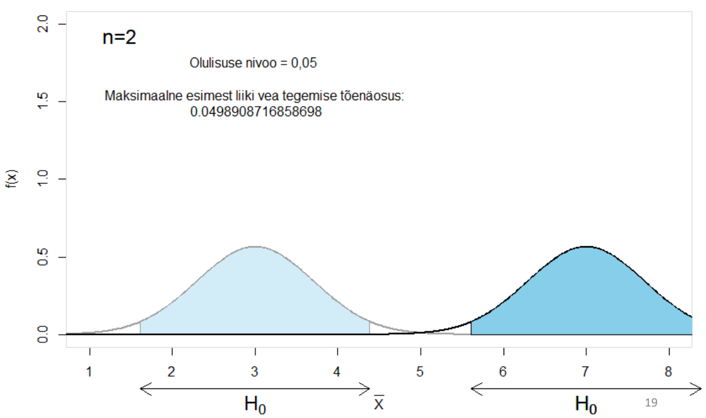
Liithüpoteesid (*Composite hypothesis*)

H_0: E(X)=3 või E(X)=7



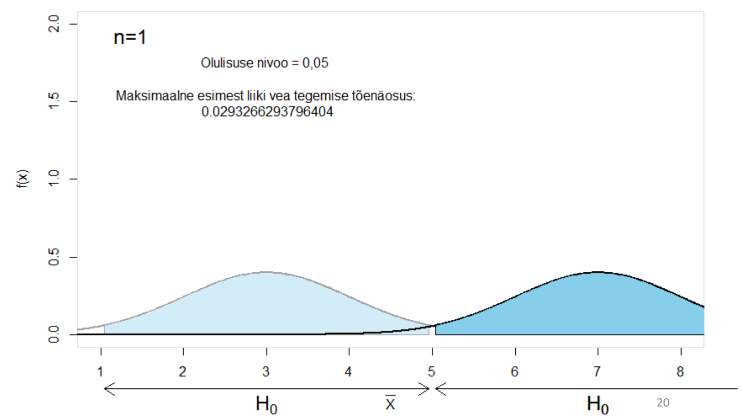
Liithüpoteesid (*Composite hypothesis*)

H_0: E(X)=3 või E(X)=7



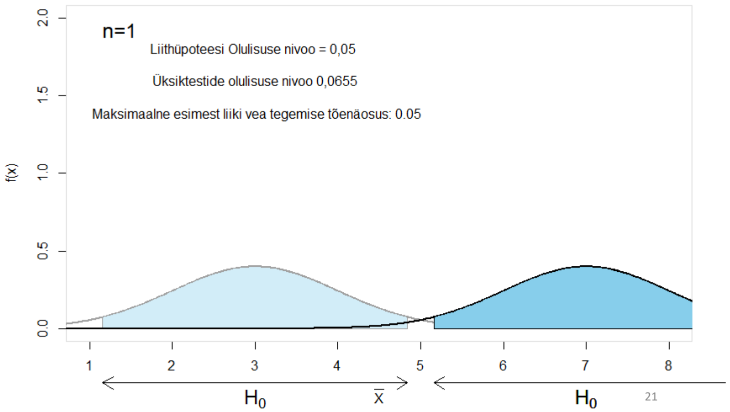
Liithüpoteesid (*Composite hypothesis*)

H_0: E(X)=3 või E(X)=7



Liithüpoteesid (*Composite hypothesis*)

H_0: E(X)=3 või E(X)=7



Liithüpoteesid (*Composite hypothesis*)

Liithüpotees: H_0: $\theta \epsilon \Theta_0$

Soovime saavutada:

- Mistahes $\theta \epsilon \Theta_0$ korral pole I liiki vea tegemise tõenäosus suurem kui α (olulisuse nivoo)

Lahendus 1:

- Kui mistahes $\theta \epsilon \Theta_0$ korral jääme H_0 juurde siis jääme H_0 juurde ka liithüpoteesi testimisel.
- Alternatiivne sõnastus: Raporteeri liithüpoteesi p-väärtusena suurimat p-väärtust mida näed testides kõiki lihtsaid nullhüpoteese kujul H_0: $\theta=\theta_0$ kus $\theta_0 \epsilon \Theta_0$.

Probleem antud lahendusega: saadud test võib teatud situatsioonides osutuda liiga konservatiivseks.

Liithüpoteesid (*Composite hypothesis*)

Liithüpotees: H_0: $\theta \epsilon \Theta_0$

Kuidas võidelda testi liigse konservatiivsuse vastu?

Sageli raske: Kui peaksin testima lihtsalt nullhüpoteesi

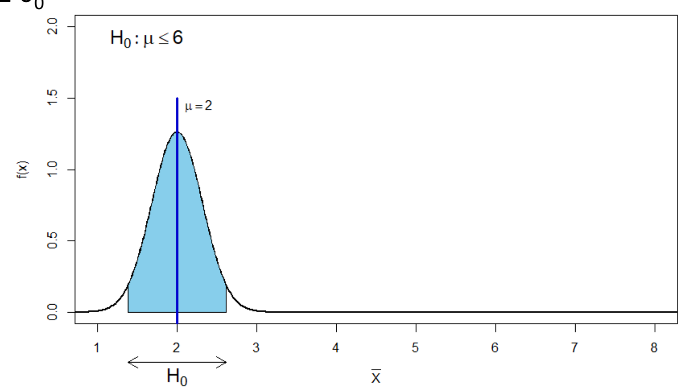
H_0: E(X)=3 või E(X)=7

ilma, et teaksin juhusliku suuruse X dispersiooni, siis oleks väga keeruline korrigeerida üksiktestide olulisuse nivoosid selliselt, et liithüpoteesi kontrollimisel esimest liiki vea tegemise tõenäosus oleks nullhüpoteesi kehtides täpselt 0,05.

Aga teatavatel erijuhtudel on sellist korrigeerimist siiski võimalik teha.

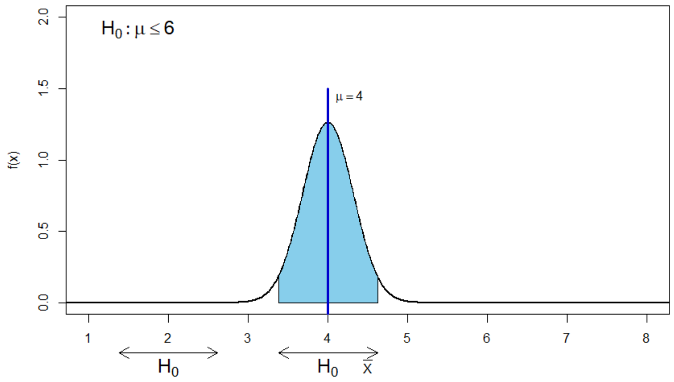
Liithüpoteesid (*Composite hypothesis*)

Liithüpotees: H_0: $\theta \leq \theta_0$



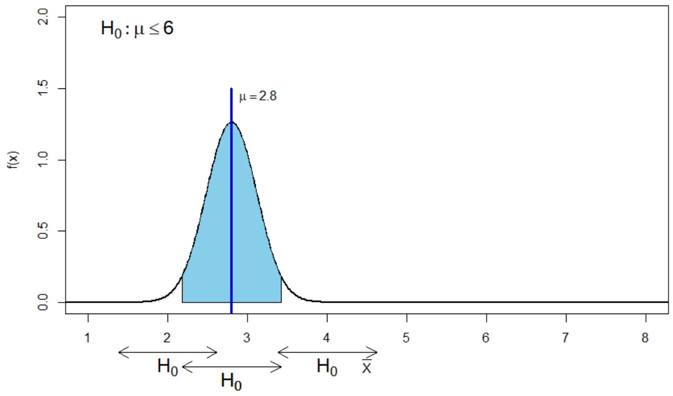
Liithüpoteesid (*Composite hypothesis*)

Liithüpotees: H_0: $\theta \leq \theta_0$



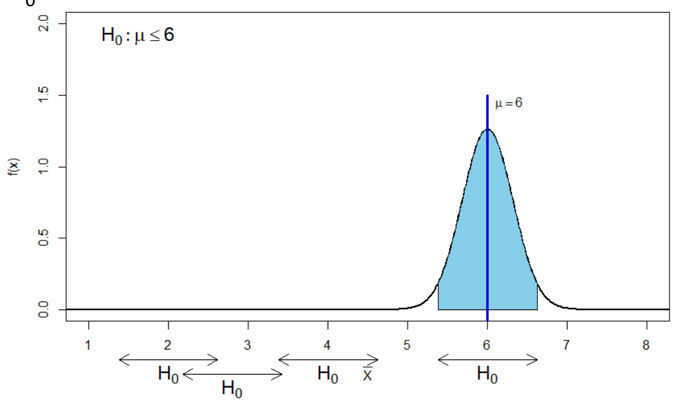
Liithüpoteesid (*Composite hypothesis*)

Liithüpotees: H_0: $\theta \leq \theta_0$



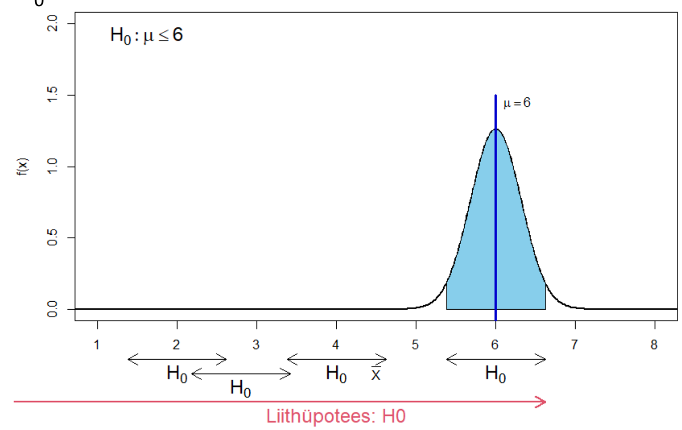
Liithüpoteesid (*Composite hypothesis*)

Liithüpotees: H_0: $\theta \leq \theta_0$



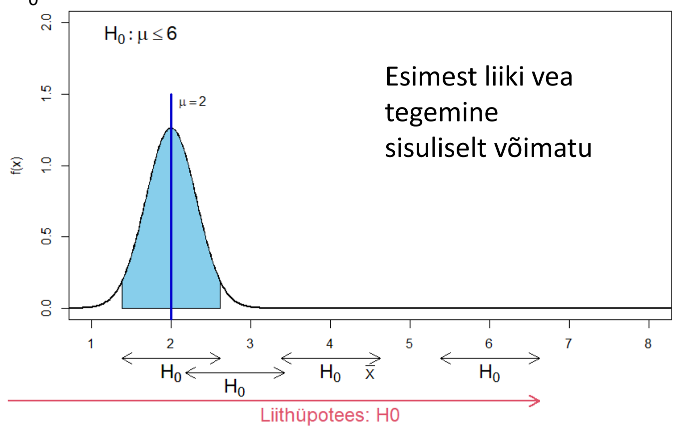
Liithüpoteesid (*Composite hypothesis*)

Liithüpotees: H_0: $\theta \leq \theta_0$



Liithüpoteesid (*Composite hypothesis*)

Liithüpotees: H_0: $\theta \leq \theta_0$

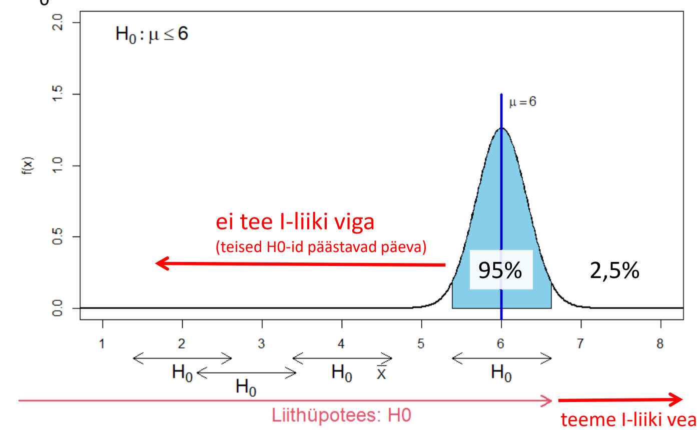


Liithüpoteesid (*Composite hypothesis*)

Liithüpotees: H_0: $\theta \leq \theta_0$

Esimest liiki vea
tegemine
veidi võimalik...

Kui tegelikult E(X)=6;
saame I-liiki viga teha
ainult ühes suunas
eksides...

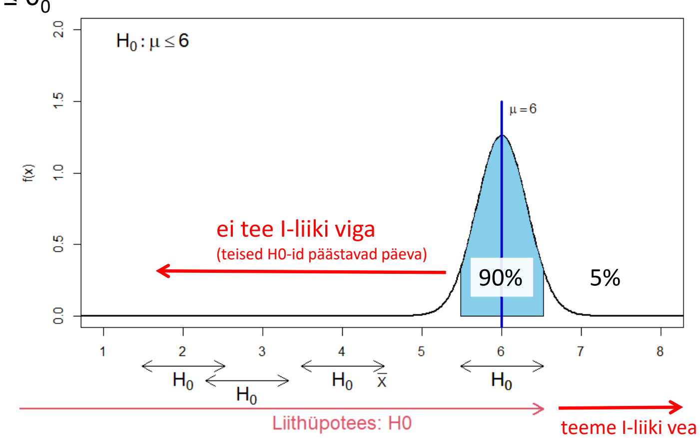


Liithüpoteesid (*Composite hypothesis*)

Liithüpotees: H_0: $\theta \leq \theta_0$

Esimest liiki vea
tegemine
veidi võimalik...

Kui tegelikult E(X)=6;
saame I-liiki viga teha
ainult ühes suunas
eksides...



Liithüpoteesid (*Composite hypothesis*)

Ühepoolne hüpotees (*One-sided test*): H_0: $\mu \leq \mu_0$; H_1: $\mu > \mu_0$

Testime olulist erinevust nullhüpoteesist (*Minimal Effect Test*): H_0: $|\mu - \mu_0| \leq d$; H_1: $|\mu - \mu_0| > d$

Testime kas nullhüpotees kehtib – kas võimalik erinevus nullhüpoteesist on ebaoluline (*Eqvivalence test*): H_0: $|\mu - \mu_0| > d$; H_1: $|\mu - \mu_0| \leq d$

(*Inferiority test*): H_0: $\mu - \mu_0 > d$; H_1: $\mu - \mu_0 \leq d$

Olulise erinevuse testimine (*Minimal Effect Test*)

Tavapärased statistilised testid:

$$H_0 : \mathrm{E}(X) = 6{,}4$$

$$H_0 : \mathrm{E}(X) = \mathrm{E}(Y)$$

$$H_0 : \mathrm{D}(X) = 12$$

Otsus: kummutame H_0-i, sest tehtud mõõtmiste keskväärtus pole 6,4 (on näiteks 6,4128...)

Kas see ikkagi veenab meid selles, et kontrollitav teooria ei kehti?

Mõõtmised:

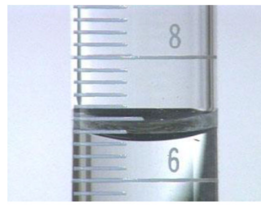


Figure 6 – a 10 mL graduated cylinder

Kui palju vedelikku on selles katseklaasis?

Olulise erinevuse testimine (*Minimal Effect Test*)

Lahendus:

Hüpoteesi $H_0 : \mathrm{E}(X) = \mu_0$

asemel testi hüpoteese

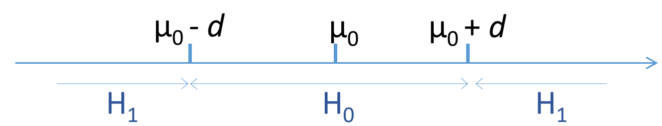
$$H_0 : |\mathrm{E}(X) - \mu_0| \le d$$

$$H_1 : |\mathrm{E}(X) - \mu_0| > d$$

Keskväärtuste erinevus on väiksem mingist suurusest *d (keskväärtused pole oluliselt erinevad)*

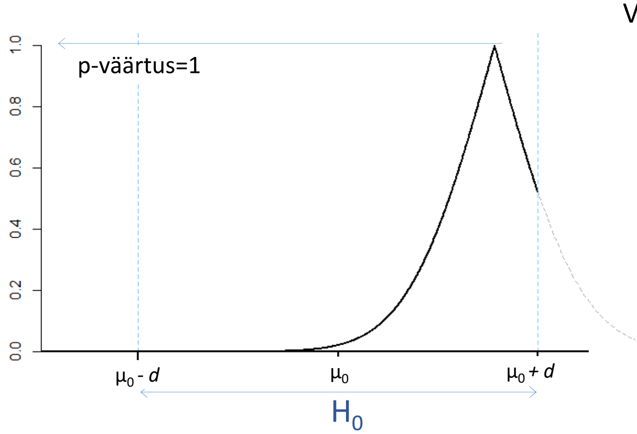
Keskväärtuste erinevus on suurem kui *d* (keskväärtused on ka meditsiinilises/bioloogilises/... mõttes oluliselt erinevad)

Kus *d* tähistab minimaalset olulist erinevus.



Liithüpoteesi näide. Variant 1: suurim p-väärtus

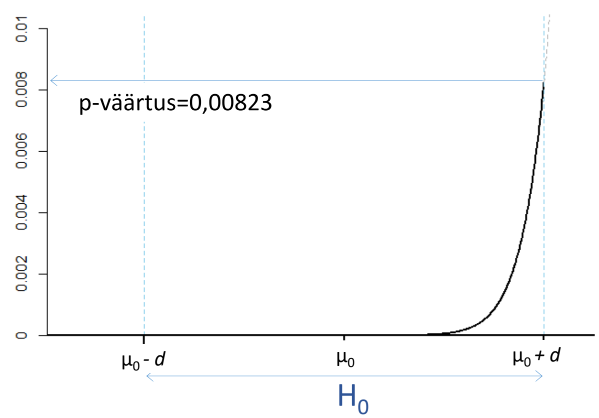
Valim 1



H_0: $|\mu - \mu_0| \le d$
H_1: $|\mu - \mu_0| > d$

Liithüpoteesi näide. Variant 1: suurim p-väärtus

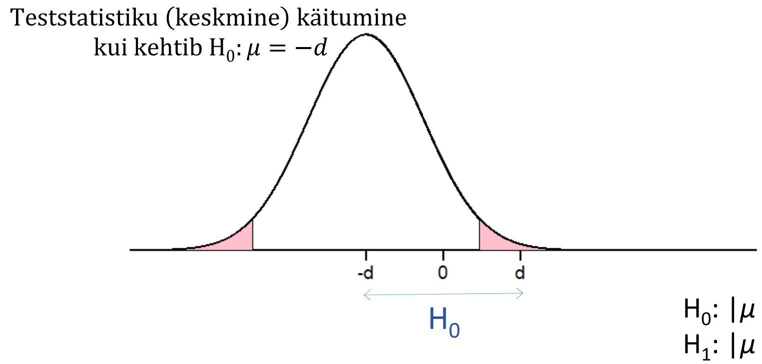
Valim 2



Ehk võtame vastu alternatiivse hüpoteesi vaid siis, kui kõigi parameetrite $\theta \in \Theta_0$ korral saame kummutada nullhüpoteesi.

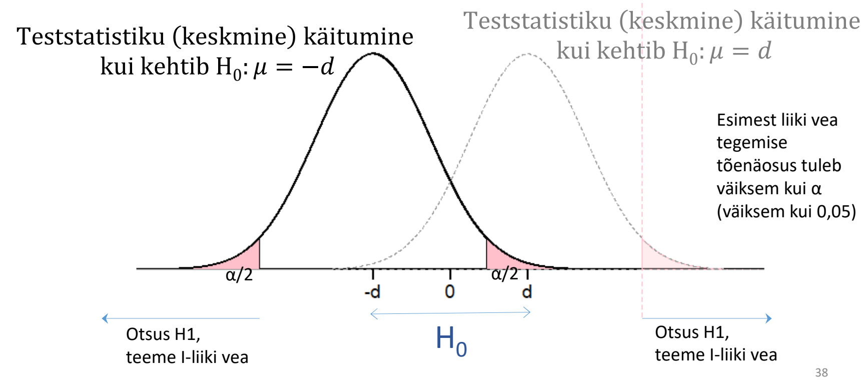
H_0: $|\mu - \mu_0| \le d$
H_1: $|\mu - \mu_0| > d$

Miks mainitud lähenemine on konservatiivne?



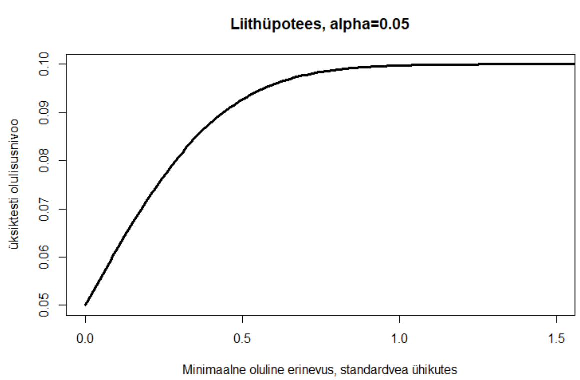
H_0: $|\mu - 0| \leq d$
H_1: $|\mu - 0| > d$

Miks mainitud lähenemine on konservatiivne?



Esimest liiki vea tegemise tõenäosus tuleb väiksem kui α (väiksem kui 0,05)

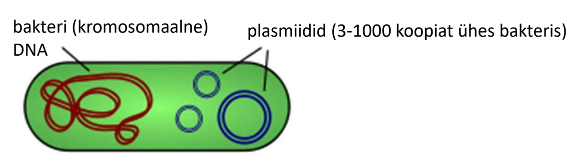
Miks mainitud lähenemine on konservatiivne?

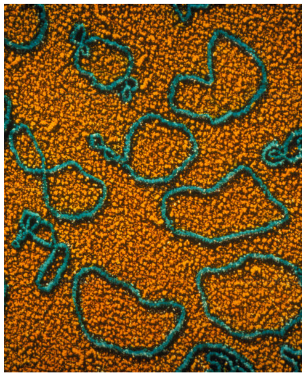


Järeldus: Kui piirkond kus kehtib H_0 on lai (võrreldes teststatistiku hajuvusega), siis võid teha regiooni otstes ühepoolsed testid. Kui mõlemad testid kummutavad nullhüpoteesi siis loe tõestatuks alternatiivne hüpotees.

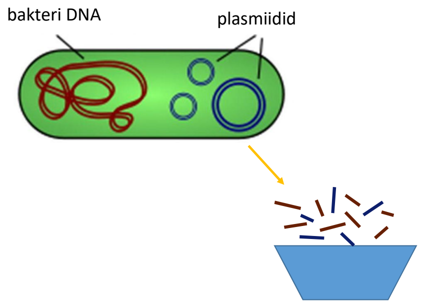
H_0: $|\mu - \mu_0| \leq d$
H_1: $|\mu - \mu_0| > d$

Olulise erinevuse testimine - kasutusnäiteid





Kas mingi DNA-lõik paikneb plasmiidis või kromosomaalses DNA-s?




```
  ACTGGGAT
GGACTGGG
       GATCAATA
    TGGGATCA
```

Uurimisalune DNA mille asukohta soovime määrata

......AACTGGGATAA.....

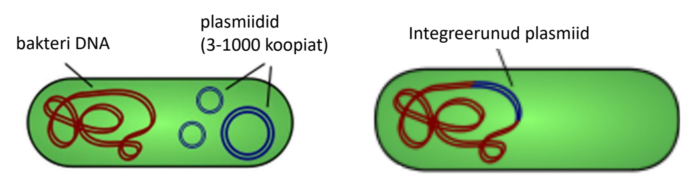
Võtame sealt välja ühe
k-tähelise jupi (*k*-meeri)

Uurimisalune DNA mille asukohta soovime määrata

......A**ACTGGGATA**A.....

Võtame sealt välja ühe
k-tähelise jupi (*k*-meeri)

… ja vaatame, kui sageli seda tähekombinatsiooni lugemites ette tuleb



Vaba plasmiidi DNA esineb mitmes korduses (ühes bakteris mitu korda):

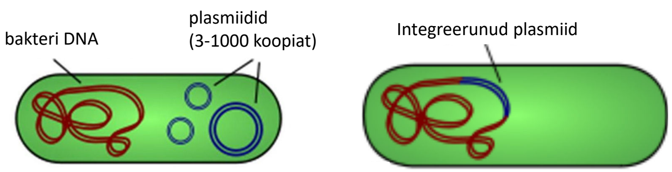
.......ACTGGGATA....
.......ACTGGGATA....
.......ACTGGGATA....

```
 AGAGAACTGGGATA
        TGGGATAAAAATAC
    GAACTGGGATAAAA
     AACTGGGATAAAAA
CAGAGAACTGGGATA
  GAGAACTGGGATAAA
```

Integeerunud plasmiidi DNA esineb üks kord bakteri kohta:

.......ACTGGGATA....

```
AACTGGGATAAAAA
 ACTGGGATAAAAAT
```



Vaba plasmiidi DNA esineb mitmes korduses (ühes bakteris mitu korda):

.......ACTGGGATA....
.......ACTGGGATA....
.......ACTGGGATA....

↓

```
AGAGAACTGGGATA
       TGGGATAAAAATAC
   GAACTGGGATAAAA
    AACTGGGATAAAAA
CAGAGAACTGGGATA
  GAGAACTGGGATAAA
```

5 ×

Integeerunud plasmiidi DNA esineb üks kord bakteri kohta:

.......ACTGGGATA....

↓

```
AACTGGGATAAAAA
 ACTGGGATAAAAAT
```

2 ×

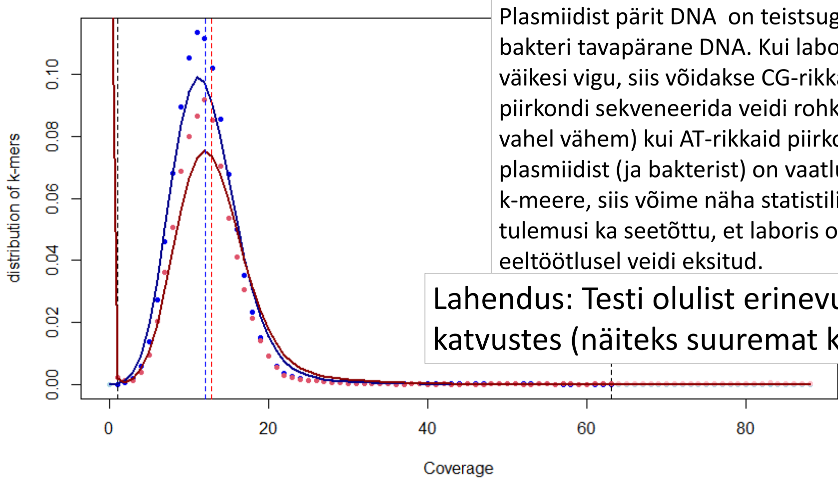
k-meerid bakteri (kromosomaalsest) DNAst

k-meer	mitu korda nähti
...	...
AACAAAACA	1
TATATTATC	2
TAGGAGAAA	2
TATCTATAA	3
AAACAAAGA	3
...	...

k-meerid oletuslikust plasmiidist

k-meer	mitu korda nähti
...	...
CGCCCGGCA	2
GGCGGGAGG	2
TTCCTCCTC	3
CGGGGCGAA	3
CCCGGCCCG	3
...	...

Võrdlen keskväärtuseid?
Näiteks *t*-testi abil?



Probleem:
Plasmiidist pärit DNA on teistsugune kui bakteri tavapärane DNA. Kui labori töös on väikesi vigu, siis võidakse CG-rikkaid DNA-piirkondi sekveneerida veidi rohkem (või vahel vähem) kui AT-rikkaid piirkondi. Kui plasmiidist (ja bakterist) on vaatluse all palju k-meere, siis võime näha statistiliselt olulisi tulemusi ka seetõttu, et laboris on proovi eeltöötlusel veidi eksitud.

Lahendus: Testi olulist erinevust katvustes (näiteks suuremat kui 5%).

Näide 2 - sotsiaalteadused

RMSEA

RMSEA is one of the only fit indices for which the asymptotic sampling distribution is known, so we can make confidence intervals and conduct hypothesis tests about its population value.

- Test of Exact Fit
  - H_0: RMSEA $= 0$ in the population
  - Equivalent to the significance test on the chi-square statistic
- Test of Close Fit (MacCallum et al., 1996)
  - H_0: Null hypothesis: $\mathrm{RMSEA} < \mathrm{RMSEA}_{good}$ in the population.
  - RMSEA_{good} is some acceptable value of RMSEA (in lavaan: 0.05)
- Test of Not-Close Fit (MacCallum et al., 1996)
  - H_0: Null hypothesis: $\mathrm{RMSEA} > \mathrm{RMSEA}_{bad}$ in the population.
  - RMSEA_{bad} is some unacceptable value of RMSEA (e.g., 0.08)

H_0: Väljapakutud mudel on absoluutselt õige

H_0: Väljapakutud mudel on põhimõtteliselt õige

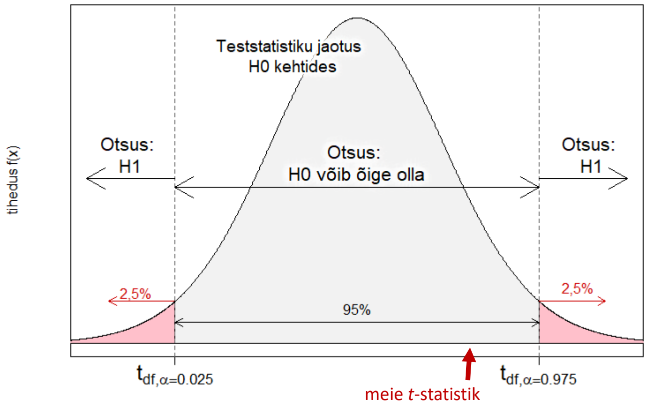
H_0: Väljapakutud mudel on halb

Statistikas ei saa (tavaliselt) nullhüpoteesi tõestada...

Kui testime näiteks hüpoteese

H_0: EX = EY
H_1: E$X \neq$ EY

Siis jääme nullhüpoteesi juurde siis, kui teststatistiku väärtus on ootuspärane (selline, millist võiks kohata nullhüpoteesi kehtides)



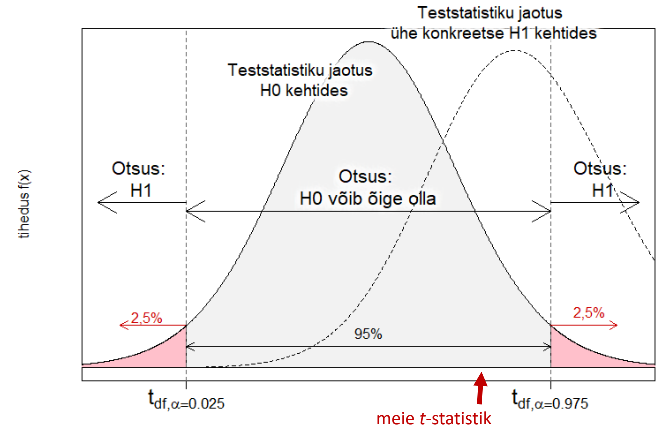
Statistikas ei saa (tavaliselt) nullhüpoteesi tõestada...

Kui testime näiteks hüpoteese

H_0: EX = EY
H_1: E$X \neq$ EY

Siis jääme nullhüpoteesi juurde siis, kui teststatistiku väärtus on ootuspärane (selline, millist võiks kohata nullhüpoteesi kehtides)

Samasugust teststatistiku väärtust võime kohata aga ka alternatiivse hüpoteesi kehtides....



Vahel aga on vaja nullhüpoteesi tõestada...

- Kas koopiaravim on sama hea kui originaalravim?
- Soovime tõestada, et mingis valdkonnas pole meeste ja naiste võimete vahel erinevust...
- Kas mõõteriista mõõtmistäpsus on selline nagu tootja poolt väidetud?
- ....

Ekvivalentsuse testimine.

H_0: $|E(X)-E(Y)| > d$
H_1: $|E(X)-E(Y)| \leq d$

Tee kaks ühepoolset testi, mõlemad olulisuse nivool α:

Test 1:
H_0: $E(X)-E(Y) > d$
H_1: $E(X)-E(Y) \leq d$

Test 2
H_0: $E(X)-E(Y) < -d$
H_1: $E(X)-E(Y) \geq -d$

Kui mõlema testi korral kummutad nullhüpoteesi (tõestad alternatiivse hüpoteesi), siis loe tõestatuks, et keskväärtused on võrdsed (keskväärtuste võimalik erinevus ei oma praktikas tähtsust).
TOST – **T**wo **O**ne-**S**ided t-**T**ests

Näiteid kasutamisest.

Oletus: ovaluatsiooni ajal muutuvad naised näost punasemaks et noormeestele just sel ajal kõige enam meeldida.

Väikseim oluline erinevus nahavärvis: väikseim värvierinevus mida mehed veel eristada suudavad

Järeldus: Tõestati, et naised ei lähe ovaluatsiooni ajal näost punaseks (võimalik nahavärvi muutus on piisavalt väike, nii et mehed võimalikku erinevust ei suuda märgata).