Biomeetria

Usaldusintervall.
Hüpoteeside kontrolli põhimõisted.
Kahe populatsiooni keskväärtuste võrdlemine

Tolerantsiintervall

Milline on järgmine katsetulemus, järgmisena mõõtmiseks valitava looma turjakõrgus, järgmisena kinni püütava kala kaal, …?

Teades uuritava tunnuse jaotust populatsioonis on võimalik leida vahemik, kuhu järgmine vaatlustulemus satub küllalt suure tõenäosusega.

$$X \sim N(\mu,\sigma^2) \Rightarrow \frac{X-\mu}{\sigma} \sim N(0,1)$$

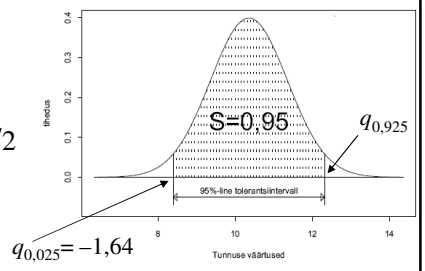
$$\mathrm{P}\left(q_{\alpha/2} > \frac{X-\mu}{\sigma}\right) = \alpha/2 \text{ ja } \mathrm{P}\left(\frac{X-\mu}{\sigma} > q_{1-\alpha/2}\right) = \alpha/2$$

$$\mathrm{P}\left(q_{\alpha/2} \le \frac{X-\mu}{\sigma} \le q_{1-\alpha/2}\right) = 1-\alpha$$

(oletame, et $\alpha = 0,05$)

$$\mathrm{P}\left(\sigma \cdot q_{0,025} \le X-\mu \le \sigma \cdot q_{0,975}\right) = 0,95$$

$$\mathrm{P}\left(\sigma \cdot q_{0,025} + \mu \le X \le \sigma \cdot q_{0,975} + \mu\right) = 0,95$$



Standardse normaaljaotuse enamkasutatavate kvantiilide (= jaotusfunktsiooni argumentide) z_α väärtused:

$\Phi(x) = \alpha$	0,005	0,025	0,05	0,5	0,95	0,975	0,995
$x = z_\alpha\ (q_\alpha)$	-2,58	-1,96	-1,64	0	1,64	1,96	2,58

Usaldusintervall

Hinnatava parameetri usaldusintervall (vahemikhinnang) kujutab enesest sellist piirkonda parameetri punkthinnangu ümber, mis katab parameetri õige väärtuse küllalt suure etteantud tõenäosusega:

$$\mathrm{P}(\underline{m} < m < \overline{m}) = 1 - \alpha$$

- $1-\alpha$ on **usaldusnivoo** (ühe lähedane, ent alati ühest väiksem);
- α, mida nimetatakse **olulisuse nivooks**, on väike positiivne arv (tavaliselt 0,01 või 0,05);
- m on õige, ent mitteteadaolev jaotusparameetri väärtus;
- $\underline{m}$ ja $\overline{m}$ on parameetri m **(1–α)-usalduspiirid** (näiteks kui α = 0,05, siis on tegu 95%-liste usalduspiiridega).

Täpsuse huvides räägitakse vahel ka alumisest ja ülemisest usalduspiirist [*lower/upper confidence limit*].

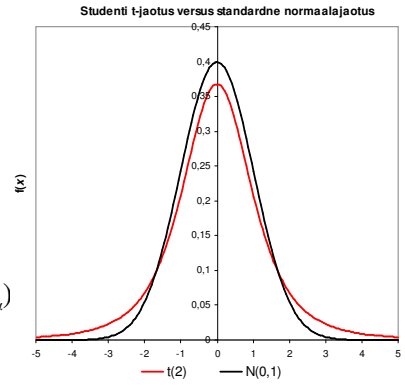
Usaldusintervall

Usaldusintervall (*confidence interval*) **keskmisele**

✓ $X \sim N(\mu, \sigma^2) \quad \Rightarrow \quad \overline{x} \sim N\left(\mu, \dfrac{\sigma^2}{n}\right)$ ehk $\dfrac{\overline{x} - \mu}{\sigma / \sqrt{n}} \sim N(0,1)$

Kui σ ei ole teada, siis $\dfrac{\overline{x} - \mu}{s / \sqrt{n}} \sim t_{n-1}$

Toodud seosed kehtivad <u>suure valimi</u> korral sõltumata uuritava tunnuse jaotusest!

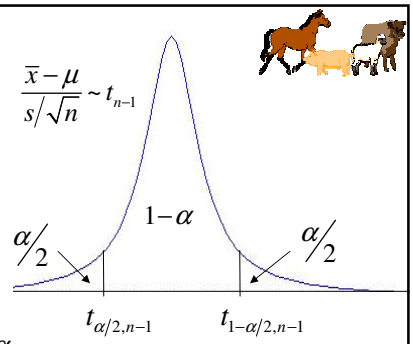


✓ α-kvantiil q_α (t-jaotuse puhul $t_{\alpha,n-1}$, standardse normaaljaotuse puhul z_α)

$$\mathrm{P}(X < q_\alpha) = \alpha$$

Sisuliselt sama, mis protsentiil;
näiteks 0,5-kvantiil on mediaan, sest P($X < med$) = 0,5.

Usaldusintervall



Usaldusintervall keskmisele

$$\mathrm{P}\left(t_{\alpha/2,n-1} \le \frac{\bar{x}-\mu}{s/\sqrt{n}} \le t_{1-\alpha/2,n-1}\right) = 1-\alpha$$

$$\mathrm{P}\left(t_{\alpha/2,n-1}\frac{s}{\sqrt{n}} \le \bar{x}-\mu \le t_{1-\alpha/2,n-1}\frac{s}{\sqrt{n}}\right) = 1-\alpha$$

$$\mathrm{P}\left(-\bar{x}+t_{\alpha/2,n-1}\frac{s}{\sqrt{n}} \le -\mu \le -\bar{x}+t_{1-\alpha/2,n-1}\frac{s}{\sqrt{n}}\right) = 1-\alpha$$

$$\mathrm{P}\left(\bar{x}-t_{\alpha/2,n-1}\frac{s}{\sqrt{n}} \ge \mu \ge \bar{x}-t_{1-\alpha/2,n-1}\frac{s}{\sqrt{n}}\right) = 1-\alpha$$

$$\mathrm{P}\left(\bar{x}-t_{1-\alpha/2,n-1}\frac{s}{\sqrt{n}} \le \mu \le \bar{x}+t_{1-\alpha/2,n-1}\frac{s}{\sqrt{n}}\right) = 1-\alpha$$

$(1-\alpha)$-usalduspiirid: $\left(\underline{\mu};\overline{\mu}\right) = \left(\bar{x}-t_{1-\alpha/2,n-1}\frac{s}{\sqrt{n}};\bar{x}+t_{1-\alpha/2,n-1}\frac{s}{\sqrt{n}}\right)$

Kui valim on suur ($n > 100$), siis võib kasutada ka normaaljaotust:

$$\left(\underline{\mu};\overline{\mu}\right) = \left(\bar{x}-z_{1-\alpha/2}\frac{s}{\sqrt{n}};\bar{x}+z_{1-\alpha/2}\frac{s}{\sqrt{n}}\right)$$

t-jaotuse *α*-kvantiilide $t_{\alpha,n-1}$ väärtused

$f = n-1$	a =0,01	a=0,025	a =0,05	a =0,95	a=0,975	a =0,99
1	-31,82	-12,71	-6,31	6,31	12,71	31,82
2	-6,97	-4,30	-2,92	2,92	4,30	6,97
3	-4,54	-3,18	-2,35	2,35	3,18	4,54
4	-3,75	-2,78	-2,13	2,13	2,78	3,75
5	-3,37	-2,57	-2,01	2,01	2,57	3,37
6	-3,14	-2,45	-1,94	1,94	2,45	3,14
7	-3,00	-2,36	-1,89	1,89	2,36	3,00
8	-2,90	-2,31	-1,86	1,86	2,31	2,90
9	-2,82	-2,26	-1,83	1,83	2,26	2,82
10	-2,76	-2,23	-1,81	1,81	2,23	2,76
12	-2,68	-2,18	-1,78	1,78	2,18	2,68
14	-2,62	-2,14	-1,76	1,76	2,14	2,62
16	-2,58	-2,12	-1,75	1,75	2,12	2,58
18	-2,55	-2,10	-1,73	1,73	2,10	2,55
20	-2,53	-2,09	-1,73	1,73	2,09	2,53
25	-2,49	-2,06	-1,71	1,71	2,06	2,49
30	-2,46	-2,04	-1,70	1,70	2,04	2,46
40	-2,42	-2,02	-1,68	1,68	2,02	2,42
60	-2,39	-2,00	-1,67	1,67	2,00	2,39
120	-2,36	-1,98	-1,66	1,66	1,98	2,36
∞	-2,33	-1,96	-1,64	1,64	1,96	2,33

Näiteks *MS Excel*is saab *t*-jaotuse 1−α/2-kvantiili $t_{1-\alpha/2,n-1}$ leidmiseks kasutada funktsiooni TINV(α;$n-1$); α-kvantiilid $t_{\alpha,n-1}$ on siis leitavad kujul 1−TINV(2α;$n-1$)

Usaldusintervall

Näide. Kanauurijat Hans Hane huvitab, mitu muna munevad keskmiselt Eestis peetavad sassexi tõugu kanad ühe nädala jooksul. Härra Hani luges ühe nädala jooksul kokku kümne kana munad: 3 5 4 6 2 6 5 6 5 3.
95%-line usaldusintervall = ?

$$\bar{x} = 4,5;\ s \approx 1,43 \qquad t_{1-\alpha/2;(n-1)} = t_{0,975;9} = 2,26$$

$$(\underline{\mu};\overline{\mu}) = \left(\bar{x} - t_{1-\alpha/2,n-1}\frac{s}{\sqrt{n}};\ \bar{x} + t_{1-\alpha/2,n-1}\frac{s}{\sqrt{n}}\right) = \left(4,5 - 2,26\frac{1,43}{\sqrt{10}};\ 4,5 + 2,26\frac{1,43}{\sqrt{10}}\right)$$

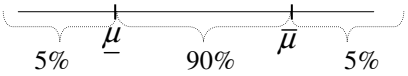
$$= (4,5 - 2,26 \times 0,45;\ 4,5 + 2,26 \times 0,45) = (3,47;\ 5,53)$$

95%-lise tõenäosusega võib väita, et keskmine nädalas munetud munade arv on kuskil vahemikus 3,47-st 5,53-ni.

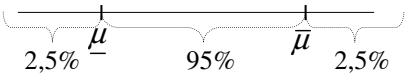
Suurendamaks hinnangu täpsust, tuleks uurida rohkem kanu, sest mida suurem on n, seda kitsamaks muutub usaldusintervall.

Usaldusintervall

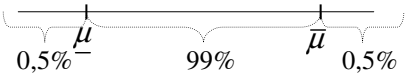
α=0,1 → 90%-usaldusintervall

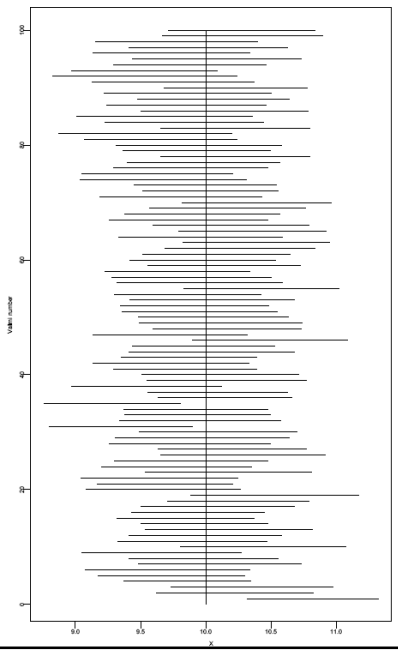


α=0,05 → 95%-usaldusintervall



α=0,01 → 99%-usaldusintervall





Hüpoteeside kontroll

Näiteid hüpoteesidest

✓ Kas jogurti toiduvärviga värvimine parandab tarbijate meelest selle maitseomadusi?
✓ Kas leidub seos lehma tiinestumise ja piimatoodangu vahel?
✓ Kas nn õnnelike sigade tailiha % on erinev tavalises sigalas kasvanud sigade vastavast näitajast?
✓ Kas Eesti ja Soome vetest püütud lõhed on geneetiliselt erinevad?

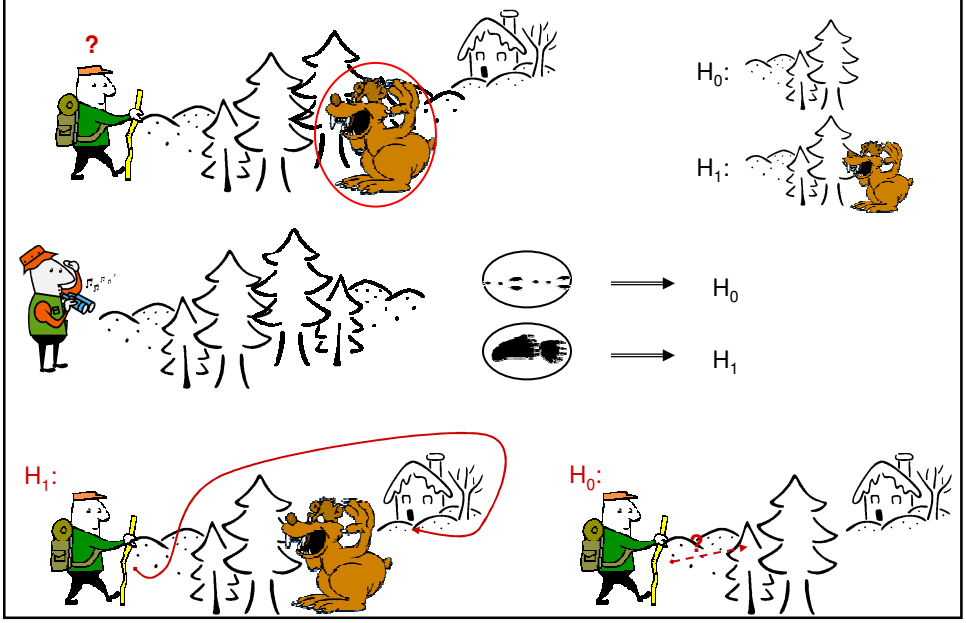
Hüpoteeside paar

H_1 - väide, mida me soovime tõestada (sisukas e alternatiivne hüpotees; *alternative hypothesis*),

H_0 - väide, et üldkogum vastab teatavale standardile (nullhüpotees; *null hyphothesis*).

Teststatistik – valimifunktsioon, mis mõõdab erinevust nullhüpoteesis väidetu ja andmetest ilmneva vahel – kui erinevus on piisavalt suur, kummutatakse nullhüpotees.

Hüpoteeside kontrolli loogika



Hüpoteeside kontroll

Vead hüpoteeside kontrollimisel

Esimest liiki viga tekib siis, kui võetakse vastu sisukas hüpotees, aga tegelikult on õige nullhüpotees.

Teist liiki viga tekib siis, kui jäädakse nullhüpoteesi juurde, kuid õige oleks sisukas hüpotees.

Tegelik olek / Otsus	Õige H_0	Õige H_1
Jääme H_0 juurde	+	II liiki viga, β
Kummutame H_0	I liiki viga, α	+

Olulisuse nivoo α (*significance level*) – maksimaalne lubatav I liiki vea tõenäosus (tavaliselt α = 0,05; 0,01; 0,001), nö valulävi.

Testi võimsus [*power*] = $1-\beta$ on tõenäosus lugeda õigeks ka tegelikult kehtiv sisukas hüpotees H_1.

Hüpoteeside kontroll

Olulisuse tõenäosus p (*probability level*, *p-value*)

– tõenäosus eksida, väites oma andmete (valimi) põhjal alternatiivse hüpoteesi H_1 kehtimist;

– tõenäosus saada uuritav (või sellest nullhüpoteesi suhtes veel ekstreemsem) valim juhuslikult, kui kehtib nullhüpotees: P(valim|H_0);

– minimaalne olulisuse nivoo alternatiivse hüpoteesi H_1 tõestamiseks, ehk I liiki vea tegemise tõenäosus antud eksperimendi korral.

Otsuse vastuvõtmine (1)

Võrreldakse olulisuse tõenäosust p ja olulisuse nivood α:

¤ **kui $p \le \alpha$, siis on tõestatud H_1,**

¤ **kui $p > \alpha$, siis jääme H_0 juurde.**

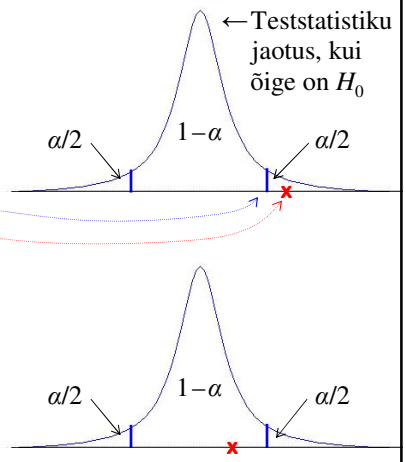
Hüpoteeside kontroll

Otsuse vastuvõtmine (2)

Võrreldakse arvutatud teststatistiku väärtust selle kriitilise väärtusega (tuginedes teoreetilistele jaotustele või simuleerimise tulemustele):

¤ kui teststatistiku absoluutväärtus on suurem tema nullhüpoteesipõhise jaotuse kriitilisest väärtusest ($1-\alpha/2$-kvantiilist), loetakse õigeks H_1,

¤ vastupidisel juhul jäädakse nullhüpoteesi H_0 juurde.



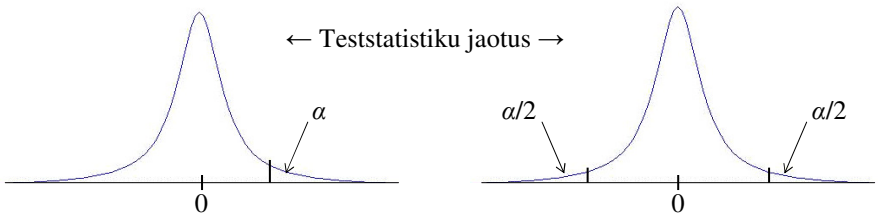
Hüpoteeside kontroll

Ühepoolne [*one-tail*] ***versus*** **kahepoolne** [*two-tail*] **hüpotees**

Näiteks:

$H_0 : \mu_1 > \mu_2$
$H_1 : \mu_1 \leq \mu_2$

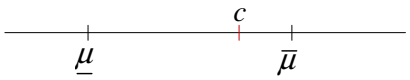
$H_0 : \mu_1 = \mu_2$
$H_1 : \mu_1 \neq \mu_2$



Seos hüpoteeside kontrolli ja usalduspiiride vahel

$H_0 : \mu = c$

võetakse vastu siis, kui c kuulub usalduspiirkonda



$H_1 : \mu \neq c$

on tõestatud siis, kui c ei kuulu usalduspiirkonda (olulisusnivool α)

Praktikas kontrollitakse sageli kas mingi kordaja või gruppide vahe erinevust nullist.
Erinevuse võib lugeda statistiliselt oluliseks ette antud olulisuse nivool (näiteks α=0,05), kui uuritavale kordajale või võrreldavate gruppide erinevusele (näiteks $\mu_1 - \mu_2$) konstrueeritud (95%-line) usaldusintervall ei sisalda nulli!

Populatsiooni keskväärtuse võrdlemine konstandiga

Usalduspiirid

$H_0 : \mu = c$

võetakse vastu siis, kui c kuulub usalduspiirkonda

$H_1 : \mu \neq c$

on tõestatud siis, kui c ei kuulu usalduspiirkonda (olulisusnivool α)

Normaaljaotuse eeldusel *t*-test

Teststatistik: $t = \frac{\bar{x} - c}{s}\sqrt{n} \underset{H_0}{\sim} t_{n-1}$

$|t| \geq t_{1-\alpha/2,n-1} \iff p \leq \alpha \Rightarrow H_1 : \mu \neq c$

$|t| < t_{1-\alpha/2,n-1} \iff p > \alpha \Rightarrow H_0 : \mu = c$

Suurte valimite (*n*>60) korral *z*-test

Teststatistik:

$Z = \frac{\bar{x} - c}{s}\sqrt{n} \underset{H_0}{\sim} N(0,1)$

$|Z| \geq z_{1-\alpha/2} \iff p \leq \alpha \Rightarrow H_1 : \mu \neq c$

$|Z| < z_{1-\alpha/2} \iff p > \alpha \Rightarrow H_0 : \mu = c$

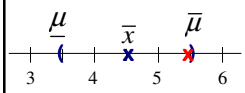
Populatsiooni keskväärtuse võrdlemine konstandiga

Näide. Kümme sassexi tõugu kana munesid nädalas vastavalt 3, 5, 4, 6, 2, 6, 5, 6, 5 ja 3 muna. Teades, et njuuhämpširi tõugu kanad munevad keskmiselt 5,4 muna nädalas, kontrollida hüpoteesi kahe tõu munatoodangute erinevusest.

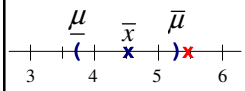
$H_0 : \mu = 5{,}4$
$H_1 : \mu \neq 5{,}4$

$\bar{x} = 4{,}5;\ s \approx 1{,}43$

95%-lised usalduspiirid keskmisele nädalasele munatoodangule:
$\underline{\mu} = 3{,}47$ ja $\overline{\mu} = 5{,}53$



Järeldused: et $\underline{\mu} = 3{,}47 < 5{,}4 < 5{,}53 = \overline{\mu}$,
siis ei ole meil olulisuse nivoo α=0,05 korral alust ümber lükata nullhüpoteesi sellest, et sassexi tõugu kanad munevad sama palju kui njuuhämpširi tõugu kanad.



Lubades 10%-list eksimist, e võttes α=0,1, saame 90%-lised usalduspiirid keskmisele nädalasele munatoodangule kujul $\underline{\mu} = 3{,}66$ ja $\overline{\mu} = 5{,}33$, mistõttu võime lugeda tõestatuks, et sassexi tõugu kanade keskmine nädalane munatoodang erineb njuuhämpširi tõugu kanade vastavast näitajast.

Populatsiooni keskväärtuse võrdlemine konstandiga

Näide. Kümme sassexi tõugu kana munesid nädalas vastavalt 3, 5, 4, 6, 2, 6, 5, 6, 5 ja 3 muna. Teades, et njuuhämpširi tõugu kanad munevad keskmiselt 5,4 muna nädalas, kontrollida hüpoteesi kahe tõu munatoodangute erinevusest.

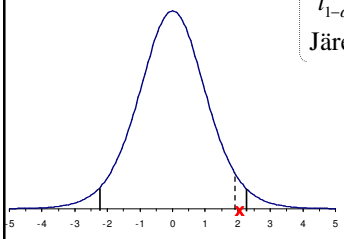
$H_0 : \mu = 5{,}4$
$H_1 : \mu \neq 5{,}4$
või
$H_0 : \mu \geq 5{,}4$
$H_1 : \mu < 5{,}4$

$n = 10;\ \bar{x} = 4{,}5;\ s \approx 1{,}43;\ \underline{\alpha = 0{,}05}$

Teststatistik: $|t| = \left|\frac{\bar{x} - 5{,}4}{s}\sqrt{n}\right| = \left|\frac{4{,}5 - 5{,}4}{1{,}43}\sqrt{10}\right| = |-1{,}985| = 1{,}985$

Teststatistiku kriitiline väärtus (kahepoolne hüpotees):
$t_{1-\alpha/2;n-1} = t_{0{,}975;9} = 2{,}26$
Järeldus: $|t| = 1{,}985 < 2{,}26 = t_{0{,}975;9} \Rightarrow H_0 : \mu = 5{,}4$

Teststatistiku kriitiline väärtus (ühepoolne hüpotees):
$t_{1-\alpha;(n-1)} = t_{0{,}95;9} = 1{,}83$
Järeldus: $t_{0{,}95;9} = 1{,}83 < 1{,}985 = |t| \Rightarrow H_1 : \mu < 5{,}4$



Näiteks *MS Excelis* funktsioon TDIST(t;n-1;2)

Arvuti abil saab leida ka täpse tõenäosuse teststatistiku väärtuse $|t|$=1,985 saamiseks eeldusel, et kehtib H_0:
p=0,0784 (2-poolne hüp.); p=0,0392 (1-poolne hüp.)

Kahe grupi keskmiste võrdlus, kolm *t*-testi – mis ja millal?

Kahe populatsiooni keskmiste võrdlus – *t*-test

Eeldus: $X \sim N(\mu;\sigma^2)$
[või suur (>60) n]

Sõltuvad vaatlused (paariviisiline võrdlus)

$H_0: \mu_1 = \mu_2$
$H_1: \mu_1 \neq \mu_2$ $\Leftrightarrow$ $H_0: \mu_1 - \mu_2 = 0$
$H_1: \mu_1 - \mu_2 \neq 0$

Teststatistik: $t = \dfrac{\bar{x}_1 - \bar{x}_2}{s_{1-2}}\sqrt{n} \underset{H_0}{\sim} t_{n-1}$

Sõltumatud vaatlused

$H_0: \mu_1 = \mu_2$
$H_1: \mu_1 \neq \mu_2$

Kahe populatsiooni dispersioonide võrdlus – *F*-test

$H_0: \sigma_1^2 = \sigma_2^2$
$H_1: \sigma_1^2 \neq \sigma_2^2$

Teststatistik:
$$F = \frac{s_1^2}{s_2^2} \underset{H_0}{\sim} F_{n_1-1,n_2-1}$$

Võrdsed dispersioonid

Teststatistik:
$$t = \frac{\bar{x}_1 - \bar{x}_2}{s}\sqrt{\frac{n_1 n_2}{n_1 + n_2}} \underset{H_0}{\sim} t_{n_1+n_2-2},$$
$$s^2 = \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}$$

Mittevõrdsed dispersioonid

Teststatistik: $t = \dfrac{|\bar{x}_1 - \bar{x}_2|}{\sqrt{s_1^2/n_1 + s_2^2/n_2}} \underset{H_0}{\sim} t_\nu,$

$$\nu = \left\{(s_1^2/n_1 + s_2^2/n_2)^2 \Big/ \left[\frac{(s_1^2/n_1)^2}{n_1+1} + \frac{(s_2^2/n_2)^2}{n_2+1}\right]\right\} - 2$$

Sõltuvad vaatlused (paariviisiline võrdlus)

$H_0: \mu_1 = \mu_2$
$H_1: \mu_1 \neq \mu_2$ $\Leftrightarrow$ $H_0: \mu_1 - \mu_2 = 0$
$H_1: \mu_1 - \mu_2 \neq 0$

Teststatistik: $t = \dfrac{\bar{x}_1 - \bar{x}_2}{s_{1-2}}\sqrt{n} \underset{H_0}{\sim} t_{n-1}$

Näide. Soovitakse uurida, kas lehmade ööpäevane piimatoodang langes pärast seda, kui neile lõpetati juurvilja söötmine (olulisuse nivool $\alpha = 0{,}05$).

Lehm	Piim (kg/ööpäevas) juurviljaga (x_1)	Piim (kg/ööpäevas) juurviljata (x_2)	Piimatoodangu muutused (d)
1	10	11	-1
2	8	7	1
3	11	10	1
4	10	10	0
5	7	6	1
6	8	5	3
7	10	6	4
8	9	4	5
9	8	6	2
10	10	8	2

Kontrollime hüpoteesi piimatoodangu languse kohta, s.t.

$H_0: \mu_d \leq 0$
$H_1: \mu_d > 0$

$n = 10;\ \bar{d} = 1{,}8;\ s_d = 1{,}81$

Andmetest arvutatud teststatistik: $t = \bar{d}/s_d \sqrt{n} \approx 3{,}14$

Teststatistiku kriitiline väärtus (ühepoolne hüpotees):
$t_{1-\alpha,\,n-1} = t_{0{,}95;9} = 1{,}83$

Järeldus: $t = 3{,}14 > 1{,}83 = t_{1-\alpha,\,n-1}$

$\Rightarrow$ $H_1: \mu_d > 0$ $(p = 0{,}006)$

Sõltumatud vaatlused

Näide. Ettevõttes võrreldi ametiühingusse kuuluvate ja sinna mittekuuluvate töötajate puudumisi aasta jooksul. Viiskümmend vaadeldud ametiühinguliiget puudusid keskmiselt 9,3 päeva, kusjuures standardhälve oli 3,1 päeva. Ametiühingusse mittekuulujad, keda oli 45, puudusid igaüks keskmiselt 8,7 päeva standardhälbega 2,3 päeva. Kontrollida hüpoteesi ettevõtte töötajate keskmiselt puudutud päevade arvu sõltuvusest ametiühingusse kuulumisest olulisuse nivool $\alpha = 0{,}05$.

$n_1 = 50;\ n_2 = 45$
$\bar{x}_1 = 9{,}1;\ \bar{x}_2 = 8{,}7$
$s_1 = 3{,}1;\ s_2 = 2{,}3$
$\alpha = 0{,}05$

$H_0: \mu_1 = \mu_2$
$H_1: \mu_1 \neq \mu_2$

(1) $H_0: \sigma_1^2 = \sigma_2^2$
$H_1: \sigma_1^2 \neq \sigma_2^2$

Teststatistik: $F = s_1^2 / s_2^2 = 1{,}817$

Teststatistiku kriitiline väärtus:
$F_{1-\alpha/2;n_1-1;n_2-1} = F_{0{,}975;\,49;\,44} = 1{,}799$

Järeldus: $F = 1{,}817 > 1{,}799 = F_{0{,}975;\,49;\,44} = F_{1-\alpha/2;n_1-1;n_2-1}$
$\Rightarrow H_1: \sigma_1^2 \neq \sigma_2^2$ $(p = 0{,}023)$

(2) Teststatistik: $t = |\bar{x}_1 - \bar{x}_2| \Big/ \sqrt{s_1^2/n_1 + s_2^2/n_2} = 0{,}969$

Teststatistiku kriitiline väärtus: $t_{1-\alpha/2,\,\nu} = t_{0{,}975;95} = 1{,}985$

Järeldus: $t = 0{,}969 < 1{,}985 = t_{0{,}975;45} \Rightarrow H_0: \mu_1 = \mu_2$ $(p = 0{,}335)$

Märkus: eeldanuks me siiski, et $\sigma_1^2 = \sigma_2^2$, jõudnuks me peale ühise dispersiooni $s^2 = 7{,}566$ ja teststatistiku $t = 0{,}708$ arvutamist samale järeldusele: $\mu_1 = \mu_2$, aga olulisustõenäosus olnuks pisut suurem ($p = 0{,}481$).

Kolm *t*-testi – mis seal vahet on?

Üldine eeldus: $X \sim N(\mu; \sigma^2)$

Andmete olemus		Lisaeeldused	p: $\mu_1 > \mu_2$	p: $\mu_1 \neq \mu_2$
Sõltumatud vaatlused	mittevõrdne varieeruvus	–	0,026	0,053
	võrdne varieeruvus	$\sigma_1^2 = \sigma_2^2$	0,024	0,048
Sõltuvad vaatlused		$H_0: \mu_1 = \mu_2 \Leftrightarrow H_0: \mu_1 - \mu_2 = 0$	0,006	0,012

Piim (kg/ööpäevas) juurviljaga (x_1)	juurviljata (x_2)
10	11
8	7
11	10
10	10
7	6
8	5
10	6
9	4
8	6
10	8

Mida enam on lihtsustavaid eelduseid e mida kitsamalt on situatsioon (andmed) enne juhuslikkuse mängu toomist piiritletud,
seda suurem on statistilise testi võimsus
(seda väiksem erinevus on vajalik populatsioonide erinevuse tõestamiseks e seda väiksem on uurija eksimistõenäosus kummutades nullhüpoteesi)!