

Tanel Kaart  
sügis, 2009

# Statistiline andmetöötlus

## VL.0435

### Loeng 3

- ✓ Hüpoteeside statistiline kontrollimine
- ✓ Keskmiste võrdlemine

[http://www.eau.ee/~ktanel/VL\\_0435/](http://www.eau.ee/~ktanel/VL_0435/)

## 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 hypothesis*).

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

## 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, $\beta$
Kummutame $H_0$	I liiki viga, $\alpha$	+

**Olulisuse nivoo  $\alpha$**  (*significance level*) – maksimaalne lubatav I liiki vea tõenäosus (tavaliselt  $\alpha = 0,05; 0,01; 0,001$ ), nõ 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 põhjal sisuka hüpoteesi  $H_1$  kehtimist (I liiki vea tegemise tõenäosus);
- tõenäosus saada analüüsitava struktuuriga (“nii suure erinevusega” või “nii tugeva seosega”) andmed juhuslikult –  $P(\text{valim}|H_0)$ ;

### Otsuse vastuvõtmine (1)

Võrreldakse olulisuse tõenäosust  $p$  ja olulisuse nivood  $\alpha$ :

☒ kui  $p < \alpha$ , siis on tõestatud  $H_1$ ,

☒ kui  $p \geq \alpha$ , siis jääme  $H_0$  juurde.

### Hüpoteeside kontroll

**Otsuse vastuvõtmine (2)**  
 Võrreldakse arvatud 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:

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

### 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  $\alpha$ )



#### Normaaljaotuse eeldusel t-test

Teststatistik:  $t = \frac{\bar{x} - c}{s} \sqrt{n} \underset{H_0}{\sim} t_{n-1}$   $|t| \geq t_{1-\alpha/2, n-1} \Leftrightarrow p \leq \alpha \Rightarrow H_1 : \mu \neq c$   
 $|t| < t_{1-\alpha/2, n-1} \Leftrightarrow p > \alpha \Rightarrow H_0 : \mu = c$

#### Suurte valimite ( $n > 60$ ) korral z-test

Teststatistik:  
 $Z = \frac{\bar{x} - c}{s} \sqrt{n} \underset{H_0}{\sim} N(0,1)$   $|Z| \geq z_{1-\alpha/2} \Leftrightarrow p \leq \alpha \Rightarrow H_1 : \mu \neq c$   
 $|Z| < z_{1-\alpha/2} \Leftrightarrow p > \alpha \Rightarrow H_0 : \mu = c$

### 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.

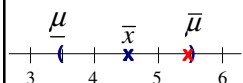
$H_0 : \mu = 5,4$   $n = 10; \bar{x} = 4,5; s \approx 1,43$

$H_1 : \mu \neq 5,4$

95%-lised usalduspiirid keskmisele nädalasele munatoodangule:

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

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



Järeldused: et  $\mu = 3,47 < 5,4 < 5,53 = \bar{\mu}$ ,

siis ei ole meil olulisuse nivoo  $\alpha = 0,05$  korral alust ümber lükata nullhüpoteesi sellest, et sassexi tõugu kanad munevad sama palju kui njuuhämpširi tõugu kanad.

NB! Näitel on illustreeriv tähendus. Praktikas ei ole sedavõrd väikese valimi ja diskreetse arvunnuse puhul toodud valemite rakendamine soovitatav.

### Populatsiooni keskväertuse 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.

$$H_0 : \mu = 5,4$$

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

$$H_1 : \mu \neq 5,4$$

või

$$H_0 : \mu \geq 5,4$$

$$H_1 : \mu < 5,4$$

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

Teststatistiku kriitiline väärtus (kahepoolne hüpotees):

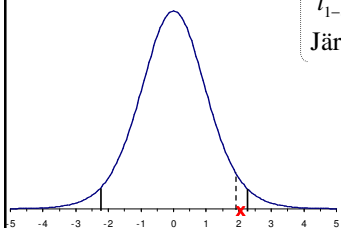
$$t_{1-\alpha/2; (n-1)} = t_{0,975; 9} = 2,26$$

$$\text{Järeldus: } |t| = 1,985 < 2,26 = t_{0,975; 9} \Rightarrow H_0 : \mu = 5,4$$

Teststatistiku kriitiline väärtus (ühepoolne hüpotees):

$$t_{1-\alpha; (n-1)} = t_{0,95; 9} = 1,83$$

$$\text{Järeldus: } t_{0,95; 9} = 1,83 < 1,985 = |t| \Rightarrow H_1 : \mu < 5,4$$



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  $H_0$ :

$$p=0,0784 \text{ (2-poolne hüp.); } p=0,0392 \text{ (1-poolne hüp.)}$$

### Kahe grupi keskmiste võrdlus, kolm t-testi

Kahe populatsiooni keskmiste võrdlus – t-test

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

Sõltuvad vaatlused (paariviisiline võrdlus)

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

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

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

Sõltumatud vaatlused

$$H_0 : \mu_1 = \mu_2$$

$$H_1 : \mu_1 \neq \mu_2$$

Kahe populatsiooni dispersioonide võrdlus – F-test

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

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

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

Võrdsed dispersioonid

$$\text{Teststatistik: } t = \frac{\bar{x}_1 - \bar{x}_2}{s} \sqrt{\frac{n_1 n_2}{n_1 + n_2}} \underset{H_0}{\sim} t_{n_1 + n_2 - 2}$$

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

Mittevõrdsed dispersioonid

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

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

### Sõltuvad vaatlused (paariviisiline võrdlus)

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

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

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

Lehm	Piim (kg/ööpäevas)	Piimatoodangu juurviljaga ( $x_1$ )	Piimatoodangu juurviljata ( $x_2$ )	Piimatoodangu muutused ( $d$ )
1	10	11	11	-1
2	8	7	7	1
3	11	10	10	1
4	10	10	10	0
5	7	6	6	1
6	8	5	5	3
7	10	6	6	4
8	9	4	4	5
9	8	6	6	2
10	10	8	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 arvatud teststatistik:  $t = \frac{\bar{d}}{s_d} \sqrt{n} \approx 3,14$   
 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 igapäev keskmiselt 8,7 päeva standardhällbega 2,3 päeva. Kontrollida hüpoteesi ettevõtte töötajate keskmiselt puudunud päevade arvu sõltuvusest ametiühingusse kuulumisest olulisuse nivool  $\alpha = 0,05$ .

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

(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:  $t = |\bar{x}_1 - \bar{x}_2| / \sqrt{s_1^2/n_1 + s_2^2/n_2} = 0,969$   
 Teststatistiku kriitiline väärtus:  $t_{1-\alpha/2, v} = t_{0,975; 95} = 1,985$   
 Järeldus:  $t = 0,969 < 1,985 = t_{0,975; 95} \Rightarrow H_0: \mu_1 = \mu_2$  ( $p = 0,335$ )

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

### Kolm *t*-testi – mis seal vahet on?

Üldine eeldus: $X \sim N(\mu; \sigma^2)$				Piim (kg/ööpäevas)	
		$p$		juurviljaga juurviljata	
		$\mu_1 > \mu_2$	$\mu_1 \neq \mu_2$	( $x_1$ )	( $x_2$ )
Andmete olemus	Lisaeeldused			10	11
				8	7
Sõltumatud vaatlused	mittevõrdne varieeruvus	–	0,026	0,053	11
					10
	võrdne varieeruvus	$\sigma_1^2 = \sigma_2^2$	0,024	0,048	7
					6
				8	5
				10	6
				9	4
Sõltuvad vaatlused	$H_0 : \mu_1 = \mu_2$	0,006	0,012	8	6
	$\Updownarrow$			10	8
	$H_0 : \mu_1 - \mu_2 = 0$				

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)!

### Permutatsioonitest

**Permutatsioonitest** (e **täpne test** [*exact test*] või randomiseerimistest [*randomization test*]) kujutab enesest teststatistiku nullhüpoteesile vastava jaotuse leidmist arvutades teststatistiku väärtused andmete kõikvõimalike ümberpaigutuste korral.

Näiteks kahe grupi keskmiste võrdlemisel, kus gruppide suurused on  $n_1$  ja  $n_2$ , arvutatakse esmalt välja andmetele vastav teststatistiku väärtus,

- seejärel moodustatakse ühine andmestik suurusega  $n_1 + n_2$ ,
- millest moodustatakse kõikvõimalikud grupid suurustega  $n_1$  ja  $n_2$  ning arvutatakse kõigil juhtudel teststatistiku väärtus;
- tulemuseks saadud teststatistiku jaotuse alusel leitakse, kui suure sagedusega (tõenäosusega) tulid teststatistiku väärtused võrdsed või suuremad originaalandmeist leitud väärtusest – saadud tõenäosus on täpne 2-poolsele hüpoteesile vastava olulisuse tõenäosuse väärtus.

Juhul, kui kõikvõimalike permutatsioonide teostamine on liiga tömahukas, valitakse neist juhuslikult üksnes teatud hulk ja arvutatakse asümptootiliselt täpne *p*-väärtus – selliseid teste tuntakse **Monte Carlo testidena**.

## Mitteparameetrilised testid

Enam levinud testid sõltuvate vaatluste korral

### Märgitest

Eeldused: uuritavad tunnused on vähemalt järjestustunnused.

Idee : võrdse keskväärtuse korral peaks paariviisiliste vaatluste vahede hulgas positiivseid ja negatiivseid (tähistatuna vastavalt “+” ja “-”, siit ka testi nimi) olema enamvähem võrdselt.

### Wilcoxon'i astakmärgitest

Eeldused: uuritavad tunnused on vähemalt järjestustunnused.

Idee : võrdse keskväärtuse korral peaks vaatluste vahede hulgas positiivseid ja negatiivseid olema enamvähem võrdselt ning, täiendusena märgitestile, peaksid mõlemad muutuma samades piirides.

## Mitteparameetrilised testid

Enam levinud testid sõltumatute vaatluste korral

### Mann-Whitney U-test, Wilcoxon'i test

Eeldused: uuritavad tunnused on vähemalt järjestatavad;  
uuritavad tunnused omandavad küllalt palju erinevaid väärtusi.

Idee: kui võrreldavate valimite keskväärtused (jaotused) on võrdsed, peaks nendest moodustatud ühine variatsioonirida olema nõ hästi segunenud, st et mõlema valimi elemendid paiknevad enamvähem vaheldumisi ega ole koondunud variatsioonirea algusesse või lõppu.

### Kolmogorov-Smirnovi test

Eeldused: uuritavad tunnused on vähemalt järjestustunnused.

Idee: test võrdleb kahe valimi jaotust (mitte üksnes keskmist taset!), leides selleks võrreldavate valimite empiiriliste jaotusfunktsioonide maksimaalse erinevuse – kui see erinevus on piisavalt suur, siis on jaotused järelikult erinevad.



### Parameetrised *versus* mitteparameetrised testid

		$P$		Piim (kg/ööpäevas)	
				juurviljaja	juurviljata
		$\mu_1 > \mu_2$	$\mu_1 \neq \mu_2$	$(x_1)$	$(x_2)$
Sõltumatud vaatlused	$t$ -test, mittevõrdne varieeruvus	0,026	0,053	10	11
	$t$ -test, võrdne varieeruvus	0,024	0,048	8	7
	Mann-Whitney U- test, Wilcoxon test	0,038	0,075	11	10
	Kolmogorov- Smirnovi test	0,055	0,164	10	10
				7	6
Sõltuvad vaatlused	$t$ -test	0,006	0,012	8	5
	Märgitest	0,020	0,039	10	6

### Mitmene võrdlus

Võrdleme näiteks 4 gruppi, lubades iga üksikvõrdluse puhul eksimist 5% tõenäosusega.

Tõenäosus, et üksikvõrdlusel viga ei tehta, on  $1 - \alpha = 0,95$ .

Tõenäosus, et kuuel üksikvõrdlusel kokku ei eksita, on  $(1 - \alpha)^6 = 0,95^6 \approx 0,735$ .

Mistõttu tõenäosus teha üks (või mitu) vale otsust(t) 4 grupi paarikaupa võrdlemisel on  $1 - 0,735 = 0,265$  (eksimise tõenäosus on üle 25%!).

Otsite näiteks põhjust, mis võiks soodustada loomadel haiguse tekkimist. Viite läbi uuringu ja fikseerite haigetel ja tervetel loomadel 100 potentsiaalselt haigestumist mõjutava tunnuse väärtused (a'la sööda toitaineterikkus, eelmise kuu keskmine temperatuur, loomi kasvata taluniku pikkus jne).

Iga potentsiaalse haigusega seotud tunnuse osas võrdlete haigeid ja terveid loomi kasutades olulisuse nivood 0,05.

Kui nüüd eeldada, et tegelikult ei mõjuta ükski valitud 100-st tunnusest haigestumist, siis sellest hoolimata võiksite antud uuringu puhul lugeda tõestatuks umbes 5 haiguse tekkimist soodustavat tegurit.

$$\begin{array}{cccc}
 & & \alpha=0,05 & \\
 & \alpha=0,05 & \alpha=0,05 & \alpha=0,05 \\
 \text{I} & \text{II} & \text{III} & \text{IV} \\
 & \alpha=0,05 & \alpha=0,05 & 
 \end{array}$$

### Mitmene võrdlus

**Bonferroni meetod:** piiramaks  $k$  üksikvõrdluse puhul ühe või enama vea tegemise tõenäosust olulisuse nivooaga  $\alpha$ , tuleb kõigil üksikvõrdlustel võtta olulisuse nivooks  $\alpha/k$ .

Näiteks 4 grupi võrdlemisel, garanteerimaks kuue võrdluse peale kokku eksimist mitte üle 5%-lise tõenäosusega, tuleb üksikvõrdlustel võtta olulisuse nivooks  $\alpha^* = \alpha/k = 0,05/6 \approx 0,0083$ .

**Bonferroni-Holmi meetod:** teostatakse kõik testid ja järjestatakse saadud olulisuse tõenäosused,  $p_1 \leq p_2 \leq \dots \leq p_k$ ; otsused nullhüpoteesi kasuks või kahjuks tehakse kasutades olulisuse nivooasid  $\alpha/k, \alpha/(k-1), \alpha/(k-2), \dots, \alpha/2, \alpha$ .

Kui teostatavate testide arv kasvab, väheneb kasutatav olulisuse nivoo kiiresti ja alternatiivse hüpoteesi tõestamine osutub sageli äärmiselt raskeks (nõuab tohutu hulga vaatluste olemasolu).

Seetõttu ei ole statistilised meetodid mitte eriti sobivad katse/eksitusemeetodil teaduse tegemiseks (proovime, kas midagi õnnestub)!

### Keskmete mitmene võrdlemine

On  $k$  gruppi, mille keskmist taset tahame võrrelda.

Sellisel juhul on sageli otstarbekas  $k(k-1)/2$  paariviisilist võrdlust ( $t$ -testi) asendada üheainsa hüpoteeside paari kontrollimisega.

Viimase võib sõnastada kujul:

$$H_0: \mu_1 = \mu_2 = \dots = \mu_k$$

$$H_1: \text{leiduvad sellised grupid } i, j, \text{ et } \mu_i \neq \mu_j$$

Eeldustel, et

- ☒ uuritav (sõltuv) tunnus on normaaljaotusega ja
- ☒ uuritava tunnuse varieeruvus võrreldavais gruppides on ühesugune, on taolise hüpoteeside paari korral rakendatavaks analüüsimeetodiks **dispersioonanalüüs**.

### Dispersioonanalüüs

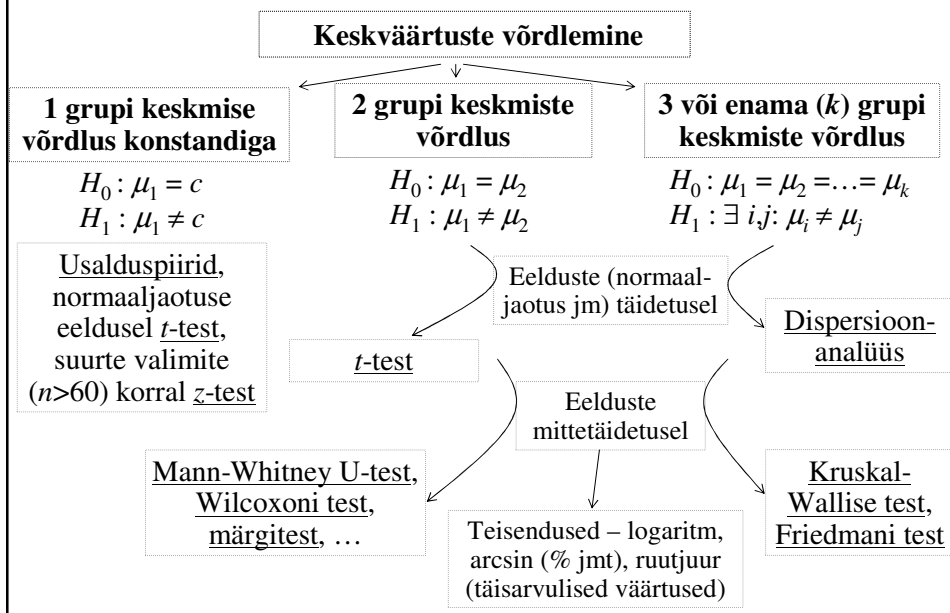
Dispersioonanalüüsil jagatakse tunnused vastavalt nende rollile kaheks:

- ⌘ tunnus, mille keskmisi võrrelda soovitakse, on **uuritav tunnus** e **funktsioontunnus** (lehma piimatoodang, forelli kasvukiirus, talle mass, sea pekipaksus, jne);
- ⌘ (diskreetne või mitteamriline) tunnus, mille väärtuste alusel võrreldavad grupid moodustatakse, on **faktortunnus** (tõug, lüpsiseade, laudatüüp jne).

Dispersioonanalüüsi tulemuste tõlgendamisel räägitaksegi enamasti faktortunnuse **mõjust** uuritavale tunnusele.

Näiteks, tõu või lüpsiseadme või laudatüübi vm mõju piimatoodangule, kasvanduse mõju forellide kasvukiirusele, omaniku mõju talle massile, genotüübi (teatud geenikombinatsioonide) mõju sigade pekipaksusele jne.

### Keskväärtuste võrdlemine



*Journal of Fish Biology* (2006) **69**, 1427–1434  
doi:10.1111/j.1095-8649.2006.01205.x, available online at <http://www.blackwell-synergy.com>

### Assessment of blood chemistry normal ranges in rainbow trout

M. MANERA\*† AND D. BRITTI‡

\*Department of Food Science, University of Teramo, Viale Crispi 212, I-64100 Teramo, Italy and †Department of Veterinary Clinical Sciences, University of Teramo, Viale Crispi 212, I-64100 Teramo, Italy

#### STATISTICAL ANALYSES

The null hypothesis, that biometrics data were a random sample from a normal distribution, was tested by the Kolmogorov–Smirnov test ( $\alpha = 0.05$ ) with SPSS 13.0.1 (SPSS Inc., Chicago, IL, U.S.A.). The null hypothesis was not rejected. Also blood chemistry values were assessed for normality with the same test in order to evaluate the normal ranges correctly. Moreover, square root and logarithmic ( $\log_{10}$ ) normalizations were performed on variables as suggested by Farver (1989) and Kolmogorov–Smirnov test reassessed. The possible impact of differences in holding time (from March to June) and, accordingly, in holding temperature (from 17 to 22° C) on analytes was assessed by means of ANOVA ( $\alpha = 0.05$ ) or Kruskal–Wallis  $H$ -test ( $\alpha = 0.05$ ), according to the distribution of the data (Gaussian or non-Gaussian), using SPSS 13.0.1.