

Biomeetria praks 4

Illustreeritud (mittetäielik) tööjuhend

Eeltöö

1. Avage MS Excel'is oma kursuse ankeedivastuseid sisaldav andmestik,
2. lisage uus tööleht, nimetage see ümber leheküljeks 'Praks4' ja
3. kopeerige kogu 'Andmed'-lehel paiknev andmetabel lehekülje 'Praks4' ülemisse vasakusse nurka.

Ülesanne 1.

Kas teie kursuse neidude pikkused vastavad Eesti standardile (Eesti naiste keskmine pikkus on 168 cm)?

Tööjuhend

1. Sorteerige kogu andmetabel veeru 'SUGU' järgi.

1) Näiteks pannes kursori veergu 'SUGU' ja valides *Home*-sakilt käsud *Sort & Filter* -> *Sort Z to A* (sorteerimaks neide noormeestest ettepoole).

2) Või klikkides suvalisel lahtril veerus 'SUGU' hiire parempoolse klahviga ja valides avanenud rippmenüüst käsud *Sort* -> *Sort Z to A*.

	A	B	C	D	E	F	G	H	I	J	K
1	RIIK	SUGU	PIKKUS	MASS	PEA_P	JALANR	ODE_VENI	MAT_HINI	HOMMIK	PUDER	LEMMI
2	Eesti	N	170	63	57	39	1	5	võileib	jah	ei
3	Eesti	N	161								
4	Eesti	M	183								
5	Eesti	N	182	81	58	42	2	5	võ		
6	Eesti	M				43	2	3	pu		
7	Eesti	N				39	3	4	pu		
8	Eesti	N				37	1	5	võ		
9	Eesti	N				41	2	4	võ		
10	Eesti	N				40	3	4	ei		
11	Eesti	N				39	0	4	võ		
12	Soome	N				40	2	4	ei		
13	Eesti	N				40	2	5	võ		
14	Eesti	N				37	1	3	võ		
15	Eesti	N				37	2	4	he		

2. Leidke neidude arv, nende keskmine pikkus ja pikkuse standardhälve kasutades funktsioone COUNT, AVERAGE ja STDEV.

	A	B	C	P	Q	R	S	T	U	V	W
1	RIIK	SUGU	PIKKUS	OLU	SUITS	TEATER	KINDO	Neidude pikkus			
2	Eesti	N	170	0	ei	rohkem k	viimase aasta jooksul	Vaatluste arv		47	
3	Eesti	N	161	0	ei	rohkem k	viimase aasta jooksul	Keskmine		168,091	
4	Eesti	N	183	0	ei	viimase a	viimase kuu jooksul	Standardhälve		=STDEV(C2:C48)	
5	Eesti	N	171	0	ei	viimase a	viimase aasta jooksul				
6	Eesti	N	161	1	ei	viimase k	viimase kuu jooksul				
7	Eesti	N	172	1	ei	rohkem k	viimase kuu jooksul				
8	Eesti	N	183,5	0	ei	viimase k	viimase kuu jooksul				
9	Eesti	N	171	0	ei	viimase k	viimase kuu jooksul				
10	Soome	N	175	1	jah	viimase a	viimase aasta jooksul				
11	Eesti	N	168	0	ei	rohkem k	rohkem kui aasta tagasi				
12	Eesti	N	152	0,5	jah	viimase a	viimase kuu jooksul				
13	Eesti	N	164	2	jah	rohkem k	viimase aasta jooksul				
14	Eesti	N	160	2	jah	rohkem k	viimase aasta jooksul				
15	Eesti	N	163,3	0	ei	viimase a	viimase 10 päeva jooksul				
16	Eesti	N	163	0	enam ei	viimase a	viimase aasta jooksul				
17	Soome	N	162	1	ei	viimase a	viimase 10 päeva jooksul				
18	Soome	N	158	0	ei	viimase a	viimase aasta jooksul				
19	Eesti	N	170	0	ei	rohkem k	viimase aasta jooksul				
20	Eesti	N	170	0,3	enam ei	viimase a	viimase kuu jooksul				
21	Eesti	N	162	0	ei	viimase a	viimase aasta jooksul				
22	Eesti	N	169	0	jah	viimase a	viimase aasta jooksul				
23	Eesti	N	176	0	ei	viimase a	viimase kuu jooksul				
24	Eesti	N	173	0	ei	viimase a	viimase aasta jooksul				
25	Eesti	N	170	1	ei	viimase a	viimase aasta jooksul				
26	Eesti	N	167	1	ei	rohkem k	viimase kuu jooksul				
27	Eesti	N	167	0	ei	rohkem k	viimase kuu jooksul				
28		N	166	2	enam ei	viimase k	viimase 10 päeva jooksul				
29	Soome	N	169	2	ei	viimase a	viimase aasta jooksul				
30	Soome	N	180	0	ei	ei ole kun	viimase aasta jooksul				
31	Soome	N	183	0	enam ei	viimase a	viimase aasta jooksul				
32	Soome	N	175	0	ei	viimase k	viimase aasta jooksul				
33	Eesti	N	178	0	ei	viimase a	viimase kuu jooksul				
34	Eesti	N	168	0	ei	rohkem k	viimase 10 päeva jooksul				
35		N	160	0	ei	viimase a	viimase 10 päeva jooksul				
36	Eesti	N	175	0	jah	viimase a	viimase aasta jooksul				
37	Eesti	N	167	0	ei	viimase a	viimase a	Neidude pikkus			
38	Eesti	N	176	0	ei	rohkem k	rohkem k	Vaatluste arv		47	
39	Eesti	N	161	2	ei	viimase a	viimase k	Keskmine		168,0915	
40	Soome	N	163	0	ei	rohkem k	viimase 1	Standardhälve		7,065219	
41	Eesti	N	161	2	ei	viimase a	viimase k				
42	Eesti	N	161,5	0	ei	viimase a	viimase k				
43	Eesti	N	163	0	ei	viimase a	viimase k				
44	Eesti	N	172	0	ei	rohkem k	viimase aasta jooksul				
45	Eesti	N	164	0	ei	rohkem k	viimase 10 päeva jooksul				
46	Eesti	N	168	0	ei	viimase a	viimase 10 päeva jooksul				
47	Eesti	N	168	0,5	ei	viimase a	viimase kuu jooksul				
48	Eesti	N	160	0	ei	viimase k	viimase kuu jooksul				
49	Eesti	M	183	2	ei	viimase a	viimase kuu jooksul				
50	Eesti	M	174	1	ei	viimase a	viimase aasta jooksul				

Seega on 47 teie kursuse neiu keskmine pikkus 168,1 cm standardhälvega 7,1 cm

- st, et keskmiselt erineb neidude tegelik pikkus 168,1 sentimeetrist 7,1 cm võrra;
- ehk, eeldades, et pikkus jaotub normaaljaotuse järgi, jääb ligikaudu 68,3% neidude pikkus vahemikku $168,1 \pm 7,1$ cm ($\bar{x} \pm s$) ja ligikaudu 95,5% neidude pikkus vahemikku $168,1 \pm 14,2$ cm ($\bar{x} \pm 2s$).

3. Sõnastage kontrollitav hüpoteeside paar ja pange see leitud arvarakteristikute alla ka kirja.

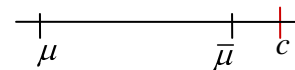
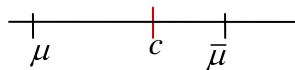
Neidude pikkus	
Vaatluste arv	47
Keskmine	168,0915
Standardhälve	7,065219
H ₀ : kursuse neidude pikkused vastavad Eesti standardile (168 cm)	
H ₁ : kursuse neidude pikkused ei vasta Eesti standardile (168 cm)	
või	
H ₀ : kursuse neidude keskmine pikkus ei erine 168 sentimeetrist	
H ₁ : kursuse neidude keskmine pikkus erineb 168 sentimeetrist	
või	
H ₀ : $\mu_T = 168$	μ_T - kursuse neidude keskmine pikkus
H ₁ : $\mu_T \neq 168$	

Meeldetuletuseks teooriast - seos hüpoteeside kontrolli ja usalduspiiride vahel

- Juhul, kui kontrollitavaks hüpoteesiks on mingi andmete alusel hinnatud suuruse võrdlemine konstandiga (mingi arvuga), tehakse otsus sageli baseeruvana uuritava suuruse usaldusintervallil:
 - kui arv, millega andmeist arvatud suurust võrreldakse, jääb usalduspiiride vahele, siis **ei ole alust väita**, et **arvutatud suurus erineb** ette antud konstandist;
 - kui aga ette antud arv jääb usaldusintervallist väljapoole, **on arvutatud suurus konstandist erinev**.
- Näiteks kui soovitakse võrrelda andmeist arvatud keskmist mingi konstandiga (et kas andmed vastavad teatud standardile), on kontrollitav hüpoteeside paar kujul:

$$H_0: \mu = c \text{ ja } H_1: \mu \neq c.$$

Kui nüüd $c \in [\underline{\mu}, \bar{\mu}]$, siis kehtib $H_0: \mu = c$; kui aga $c \notin [\underline{\mu}, \bar{\mu}]$, siis kehtib $H_1: \mu \neq c$.



4. Arvutage liidetav neidude keskmise pikkuse 95% usaldusintervalli leidmiseks (so pool usaldusintervalli laiust) ja tehke seda **kahel viisil**:

a) funktsiooni CONFIDENCE.NORM abil

(funktsioonile tuleb ette anda 3 argumenti:

olulisuse nivoo α , neidude pikkuste standardhälve ja neidude arv);

The image shows a screenshot of the Excel 'Insert Function' dialog box. The 'Or select a category' dropdown is set to 'Statistical'. In the 'Select a function' list, 'CONFIDENCE.NORM' is selected and highlighted with a red line. Below the list, the function signature is shown: **CONFIDENCE.NORM(alpha;standard_dev;size)** with a description: 'Returns the confidence interval for a population mean, using a normal distribution.' There are 'OK' and 'Cancel' buttons at the bottom right.

On the left, there is a smaller inset window titled 'Funktsioon CONFIDENCE.NORM' with a text box. A dashed line connects this text box to the 'CONFIDENCE.NORM' entry in the main dialog box.

Selle vahepealkirja võiks ise trükkida, et oleks selgem, mis funktsiooni on rakendatud. Ja kursor pange enne funktsiooni tellimist ikka sellesse lahtrisse, kuhu soovite tulemust saada!

	T	U	V	W	X	Y	Z	AA	AB	AC	AD	AE
1		Neidude pikkus										
2		Vaatluste arv	47									
3		Keskmine	168,0915									
4		Standardhälve	7,065219									
5												
6												
7		H_0 : kursuse neidude										
8		H_1 : kursuse neidude										
9												
10		või										
11												
12		H_0 : kursuse neidude										
13		H_1 : kursuse neidude										
14												
15	il	või										
16												
17	il	H_0 : $\mu_T = 168$	μ_T									
18		H_1 : $\mu_T \neq 168$										
19												
20												
21		Funktsioon CONFIDENCE.NORM										
22		=CONFIDENCE.NORM(0,05;V4;V2)										

Function Arguments

CONFIDENCE.NORM

Alpha 0,05 = 0,05

Standard_dev V4 = 7,065218843

Size V2 = 47

= 2,019876333

Returns the confidence interval for a population mean, using a normal distribution.

Size is the sample size.

Formula result = 2,019876333

[Help on this function](#)

OK Cancel

Kuna tahame leida 95% usaldusintervalli, siis olulisuse nivoo $\alpha = 0,05$.

Funktsioon CONFIDENCE.NORM

2,019876

b) funktsiooni CONFIDENCE.T abil

(funktsiooni süntaks on analoogne funktsiooni CONFIDENCE.NORM süntaksiga).

Tulemus:

Funktsioon CONFIDENCE.NORM	
	2,019876
Funktsioon CONFIDENCE.T	
	2,074426

5. Arvutage alumine ja ülemine usalduspiir mõlema tulemuse alusel.

	T	U	V	W	X	Y	Z	AA
1		Neidude pikkus						
2		Vaatluste arv	47					
3		Keskmine	168,0915					
4		Standardhälve	7,065219					
5								
6								
7		H ₀ : kursuse neidude pikkused vastavad Eesti standardile (168 cm)						
8		H ₁ : kursuse neidude pikkused ei vasta Eesti standardile (168 cm)						
9								
10		või						
11								
12		H ₀ : kursuse neidude keskmine pikkus ei erine 168 sentimeetrist						
13		H ₁ : kursuse neidude keskmine pikkus erineb 168 sentimeetrist						
14								
15	l	või						
16								
17	l	H ₀ : μ _T = 168	μ _T - kursuse neidude keskmine pikkus					
18		H ₁ : μ _T ≠ 168						
19								
20								
21		Funktsioon CONFIDENCE.NORM						
22			2,019876		Alumine usalduspiir	166,0716	=V3-V22	
23					Ülemine usalduspiir	170,1114		
24		Funktsioon CONFIDENCE.T						
25			2,074426		Alumine usalduspiir	166,0171	=V3-V25	
26					Ülemine usalduspiir	170,1659		
27								

Kumb 95%-usaldusintervallidest on laiem? Miks?

Vastus. Funktsiooniga CONFIDENCE.T arvatatu on pisut laiem. Põhjuseks see, et funktsioon

CONFIDENCE.T arvutab usalduspiirid t -jaotuse baasil valemist $\bar{x} \pm t_{1-\alpha/2, n-1} \frac{s}{\sqrt{n}}$,

funktsioon CONFIDENCE.NORM arvutab aga **asümptootilised** (ligikaudsed, kusjuures täpsus on seda suurem, mida rohkem on andmeid) ja väikeste valimite korral pisut liiga kitsad usalduspiirid standardse normaaljaotuse baasil valemist $\bar{x} \pm z_{1-\alpha/2} \frac{s}{\sqrt{n}}$.

Suurused $t_{1-\alpha/2, n-1} = t_{0,975;46} = 2,01$ ja $z_{1-\alpha/2} = z_{0,975} = 1,96$ näitavad, milline on see väärtus, millest vastava t -jaotuse või siis standardse normaaljaotuse korral on väiksemad 97,5% väärtustest (ehk, millest suuremaid väärtusi võib esineda vaid tõenäosusega 0,025), neist suurusel esimene on *Excelis* 2010 leitav näiteks funktsiooniga =T.INV(0,975;46) ja teine funktsiooniga =NORM.S.INV(0,975).

Kui aru ei saanud, tutvuge 3. loengu materjalidega.

NB! Vanemates *Exceli* versioonides ei leidu kumbagi funktsioonidest CONFIDENCE.T ja CONFIDENCE.NORM. Neist teisega, mis leiab usalduspiirid normaaljaotuse baasil, on analoogne funktsioon CONFIDENCE, t -jaotusel baseeruvate usalduspiiride leidmiseks aga vanemais *Exceli* versioonides funktsioon puudub ja kasutada tuleb protseduuri *Descriptive Statistics* valikut *Confidence Level for Mean*, mis arvutab usalduspiiride leidmiseks vajaliku liidetava t -jaotuse baasil (vt eelmise praktikumi viimast ülesannet).

6. Otsus püstitatud hüpoteesi osas – kas teie kursuse tütarlaste pikkused vastavad Eesti standardile (168 cm)? **Sõnastage lõppjärelus (koos põhjendusega) ja pange see kirja.**

Spikker. Et Eesti naiste keskmine pikkus 168 cm jääb teie kursuse neidude keskmise pikkuse 95% usaldusintervalli sisse, $168 \in (166,0; 170,2)$, siis ei ole alust lugeda tõestatult alternatiivset hüpoteesi keskmise pikkuse erinevusest 168 sentimeetrist ja tuleb jääda nullhüpoteesi H_0 juurde: kursuse neidude keskmine pikkus ei erine 168 sentimeetrist.

7. Lisa. Naiste keskmine pikkus maailmas on 154 cm. Kas on alust väita, et teie kursuse tütarlaste pikkused erinevad maailma keskmisest?

Ülesanne 2.

Kas viimase poolaasta jooksul viirushaiguseid põdenud tudengite keskmised pikkused on erinevad?

Tööjuhend

- Moodustage samale töölehele abitabel veergudest 'PIKKUS' ja 'HAIGE' ning sorteerige abitabel haigestumise järgi.

A	B	C	D	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB	AC	AD	
1	RIIK	SUGU	PIKKUS	MASS	LEMMIK	HAIGE	SPORT	SUUSEK	AUTO	OLU	SUITS	TEATER	KINO	Neidude pikkus								PIKKUS	HAIGE	
2	Eesti	N	170	63	ei	ah	jah	jah	jah	0	ei	rohkem	ku viimase aasta jooksul	Vastuste arv	47								170	ei
3	Eesti	N	163	57	jah	ah	ah	ei	ei	0	ei	rohkem	ku viimase aasta jooksul	Keskmine	168,0915								161	jah
4	Eesti	N	183	81	jah	ah	ei	jah	jah	0	ei	viimase	st viimase kuu jooksul	Standardhälve	7,065219								183	jah
5	Eesti	N	174	70	ei	ah	ah	ah	ei	0	ei	viimase	st viimase aasta jooksul										171	ei
6	Eesti	N	164	59	jah	ah	ah	jah	ei	1	ei	viimase	ku viimase kuu jooksul										161	jah
7	Eesti	N	172	80	ah	ei	ah	ei	ei	1	ei	rohkem	ku viimase kuu jooksul	H ₀ : kursuse neidude pikkused vastavad Eesti standardile (168 cm)								172	ei	
8	Eesti	N	183,5	88	jah	ah	ah	jah	jah	0	ei	viimase	ku viimase kuu jooksul	H ₁ : kursuse neidude pikkused ei vasta Eesti standardile (168 cm)								183,5	jah	
9	Eesti	N	174	57	jah	ah	ah	jah	ei	0	ei	viimase	ku viimase kuu jooksul										171	jah
10	Soome	N	174	68	ah	ei	ei	ei	jah	1	jah	viimase	st viimase aasta jooksul	või									175	ei
11	Eesti	N	165	70	ah	jah	ah	jah	jah	0	ei	rohkem	ku rohkem kui aasta tagasi										168	jah
12	Eesti	N	152	30	ah	ei	ah	ei	jah	0,5	jah	viimase	st viimase kuu jooksul	H ₀ : kursuse neidude keskmine pikkus ei erine 168 sentimeetrist								152	ei	
13	Eesti	N	164	55	ah	ei	ah	jah	ei	2	jah	rohkem	ku viimase aasta jooksul	H ₁ : kursuse neidude keskmine pikkus erineb 168 sentimeetrist								164	ei	
14	Eesti	N	160	60	ah	ei	ah	jah	ei	2	jah	rohkem	ku viimase aasta jooksul										160	ei
15	Eesti	N	163,5	67	ah	ah	ah	jah	ei	0	ei	viimase	st viimase 10 päeva jooksul	või									163,5	jah
16	Eesti	N	163	33	ah	ei	ah	jah	jah	0	enam	ei	viimase	st viimase aasta jooksul									163	ei
17	Soome	N	162	62	ah	ei	ah	jah	ei	1	ei	viimase	st viimase 10 päeva jooksul	H ₀ : μ = 168	μ - kursuse neidude keskmine pikkus								162	ei
18	Soome	N	159	65	ah	jah	ah	ei	jah	0	ei	viimase	st viimase aasta jooksul	H ₁ : μ ≠ 168									159	jah
19	Eesti	N	170	57	ah	ei	ei	jah	jah	0	ei	rohkem	ku viimase aasta jooksul										170	ei
20	Eesti	N	170	58	ah	ah	ah	ei	ei	0,3	enam	ei	viimase	st viimase kuu jooksul									170	jah
21	Eesti	N	162	37	ah	jah	ei	jah	ei	0	ei	viimase	st viimase aasta jooksul	Funktsioon CONFIDENCE.NORM								162	jah	
22	Eesti	N	165	63	ah	ei	ah	jah	jah	0	jah	viimase	st viimase aasta jooksul	2,019876	Alumine usalduspiir	166,0716 =V3-V22							169	ei
23	Eesti	N	179	67	ah	ei	ah	jah	ei	0	ei	viimase	st viimase kuu jooksul		Ülemine usalduspiir	170,1114							176	ei
24	Eesti	N	173	64	ei	jah	ah	jah	ei	0	ei	viimase	st viimase aasta jooksul	Funktsioon CONFIDENCE.T								173	jah	
25	Eesti	N	170	63	ah	ei	ah	jah	jah	1	ei	viimase	st viimase aasta jooksul	2,074426	Alumine usalduspiir	166,0171 =V3-V25							170	ei
26	Eesti	N	167	73	ah	ei	ah	jah	ei	1	ei	rohkem	ku viimase kuu jooksul		Ülemine usalduspiir	170,1659							167	ei
27	Eesti	N	167	38	ah	ah	ah	ei	ei	0	ei	rohkem	ku viimase kuu jooksul										167	jah
28	N	166	64	ah	jah	ah	ei	ei	ei	2	enam	ei	viimase	ku viimase 10 päeva jooksul									166	jah
29	Soome	N	169	65	ah	jah	ah	ei	ei	2	ei	viimase	st viimase aasta jooksul										169	jah
30	Soome	N	180	65	ah	jah	ah	ei	ei	0	ei	ei	rohkem	ku viimase aasta jooksul									180	jah
31	Soome	N	182	ah	jah	ah	ah	jah	ei	0	enam	ei	viimase	st viimase aasta jooksul									183	jah
32	Soome	N	179	60	ah	ei	ah	jah	ei	0	ei	viimase	ku viimase aasta jooksul										170	ei
33	Eesti	N	176	77	ah	ah	ah	jah	ei	0	ei	viimase	ku viimase kuu jooksul	Copy -> Paste + Sort ...								176	jah	
34	Eesti	N	168	52	ah	ah	ei	ei	ei	0	ei	rohkem	ku viimase 10 päeva jooksul										168	jah
35	N	160	62	ah	jah	ei	jah	ei	ei	0	ei	viimase	st viimase 10 päeva jooksul										160	jah
36	Eesti	N	175	65	ah	ei	ei	ei	jah	0	jah	viimase	st viimase aasta jooksul										175	ei
37	Eesti	N	167	65	ah	ei	ei	ei	ei	0	ei	viimase	st viimase aasta jooksul										167	ei
38	Eesti	N	176	74	ah	ei	ei	jah	ei	0	ei	rohkem	ku rohkem kui aasta tagasi										176	ei
39	Eesti	N	164	62	ah	ei	ah	jah	ei	2	ei	viimase	st viimase kuu jooksul										161	ei
40	Soome	N	163	52	ah	ei	ah	jah	ei	0	ei	rohkem	ku viimase 10 päeva jooksul										163	ei
41	Eesti	N	164	62	ah	ei	ah	jah	ei	2	ei	viimase	st viimase kuu jooksul										161	ei
42	Eesti	N	161,5	55	ah	jah	ah	jah	jah	0	ei	viimase	st viimase kuu jooksul										161,5	jah
43	Eesti	N	163	53	ah	ei	ah	ei	ei	0	ei	viimase	st viimase kuu jooksul										163	ei
44	Eesti	N	172	65	ah	ei	ah	ei	ei	0	ei	rohkem	ku viimase aasta jooksul										172	ei
45	Eesti	N	164	75	ah	jah	ah	jah	ei	0	ei	rohkem	ku viimase 10 päeva jooksul										164	jah
46	Eesti	N	163	57	ah	ei	ah	jah	ei	0	ei	viimase	st viimase 10 päeva jooksul										168	ei
47	Eesti	N	169	60	ah	ei	ah	jah	ei	0,5	ei	viimase	st viimase kuu jooksul										169	ei
48	Eesti	N	160	49	ah	jah	ah	jah	ei	0	ei	viimase	ku viimase kuu jooksul										160	jah
49	Eesti	M	185	80	ei	ei	ah	jah	ei	2	ei	viimase	st viimase kuu jooksul										183	ei
50	Soesti	M	174	74	ah	ei	ah	ei	jah	1	ei	viimase	st viimase aasta jooksul										174	ei
51	Eesti	M	178	80	ah	jah	ah	ei	jah	0	ei	viimase	st viimase aasta jooksul										178	jah
52	Eesti	M	189	71	ah	ei	ah	ei	jah	0	jah	viimase	st viimase kuu jooksul										189	ei
53	Eesti	M	181	61	ah	ei	ah	jah	ei	0	ei	rohkem	ku viimase 10 päeva jooksul										181	ei
54	Eesti	M	189	71	ah	ei	ah	ei	jah	0	jah	viimase	st viimase kuu jooksul										189	ei
55	Eesti	M	175	64	ah	ei	ah	jah	jah	0	enam	ei	viimase	st viimase kuu jooksul									170	ei
56	Eesti	M	178	65	ah	ei	ah	ei	jah	0	ei	viimase	st viimase kuu jooksul										178	ei
57	Eesti	M	185	80	ei	ei	ah	jah	jah	3	ei	rohkem	ku viimase aasta jooksul										183	ei
58	Eesti	M	184	83	ah	jah	ah	jah	jah	1	ei	viimase	st viimase kuu jooksul										184	jah

2. Leidke nii viirushaiguseid põdenud kui ka mitte põdenud tudengite arvud, keskmised pikkused ja pikkuste standardhälbed.

Kasutada võite nii vastavaid funktsioone kui ka Pivot Table'i abi. Kui soovite, arvutage mõlemal viisil.

PIKKUS	HAIGE	Haigus			
		Ei	Jah		
170	ei				
171	ei				
172	ei	Tudengite arv	33	24	
175	ei	Keskmine pikkus	171,0	169,7	
152	ei	Standardhälve	8,5	8,6	
164	ei				
160	ei				
163	ei				
162	ei	Column Labels			
170	ei	Values	ei	jah	Grand Total
169	ei	Count of PIKKUS	33	24	57
176	ei	Average of PIKKUS2	171,030303	169,6791667	170,4614035
170	ei	StdDev of PIKKUS3	8,520141022	8,619794165	8,511814361
167	ei				
175	ei				
175	ei				
167	ei				
176	ei				
161	ei				
163	ei				
161	ei				
163	ei				
172	ei				
168	ei				
168	ei				
183	ei				
174	ei				
189	ei				
181	ei				
189	ei				
175	ei				
178	ei				
185	ei				
161	jah				
183	jah				
161	jah				
183,5	jah				
171	jah				
168	jah				
163,3	jah				
158	jah				
170	jah				
162	jah				
173	jah				
167	jah				
166	jah				
169	jah				
180	jah				
183	jah				
178	jah				
168	jah				
160	jah				
161,5	jah				
164	jah				
160	jah				
178	jah				
184	jah				

3. Sõnastage kontrollitav hüpoteeside paar ja pange see ka kirja.

t-test	
H ₀ : Viirushaiguseid põdenud ja mitte põdenud tudengite keskmised pikkused on võrdsed.	
H ₁ : Viirushaiguseid põdenud ja mitte põdenud tudengite keskmised pikkused on erinevad.	
või	
H ₀ : $\mu_{ei} = \mu_{jah}$	μ_{ei} - viirushaiguseid mitte põdenud tudengite keskmine pikkus
H ₁ : $\mu_{ei} \neq \mu_{jah}$	μ_{jah} - viirushaiguseid põdenud tudengite keskmine pikkus

4. Millist t-testi nende keskmiste võrdlemisel kasutada?

NB! t-testi on 3 tüüpi, vt lk 11 (punkt 7 b).

- Et tegu on sõltumatute vaatlustega (võrreldavad grupid koosnevad erinevatest tudengitest), tuleb enne keskmiste võrdlemist võrrelda dispersioone, otsustamaks, millist t-testi kasutada (kas seda, mis eeldab keskmiste võrdlemisel võrdset varieeruvust, või seda, mis arvutab mõlema grupi tarvis eraldi dispersioonid).
- Varieeruvuse (dispersioonide) võrdlemiseks kasutatakse F-testi.

5. Pange kirja kontrollitav hüpoteeside paar ja viige läbi F-test, otsustamaks erinevatesse gruppidesse kuuluvate tudengite pikkuste varieeruvuse võrdumise või mittevõrdumise üle (funktsioon F.TEST).

NB! Protseduur *F-test* (Data-sakk -> Data Analysis... -> F-Test Two-Sample for Variances) testib vaid ühepoolset hüpoteesi ega ole seetõttu otseselt rakendatav, otsustamaks dispersioonide võrdumise või mittevõrdumise üle.

t-test	
H ₀ : Viirushaiguseid põdenud ja mitte põdenud tudengite keskmised pikkused on võrdsed.	
H ₁ : Viirushaiguseid põdenud ja mitte põdenud tudengite keskmised pikkused on erinevad.	
või	
H ₀ : $\mu_{ei} = \mu_{jah}$	μ_{ei} - viirushaiguseid mitte põdenud tudengite keskmine pikkus
H ₁ : $\mu_{ei} \neq \mu_{jah}$	μ_{jah} - viirushaiguseid põdenud tudengite keskmine pikkus
Tegu on sõltumatute vaatlustega. Seega tuleb enne keskmiste võrdlemist võrrelda dispersioone, otsustamaks, millist t-testi kasutada.	
F-test (võrdleme dispersioone)	
H ₀ : $\sigma_{ei}^2 = \sigma_{jah}^2$	(pikkuste varieeruvus viirushaiguseid põdenud ja mitte põdenud tudengite gruppides ei ole erinev)
H ₁ : $\sigma_{ei}^2 \neq \sigma_{jah}^2$	(pikkuste varieeruvus viirushaiguseid põdenud ja mitte põdenud tudengite gruppides on erinev)

f_x =F.TEST(AC2:AC34;AC35:AC58)

AC	AD	AE	AF	AG	AH	AI	AJ	AK	AL
PIKKUS	HAIGE				Haigus				
170	ei			Ei	Jah				
171	ei		Tudengite arv	33	24				
172	ei		Keskmine pikkus	171,0	169,7				
175	ei		Standardhälve	8,5	8,6				
152	ei								
164	ei								
160	ei								
163	ei		Column Labels						
162	ei		Values	ei	jah	Grand Total			
170	ei		Count of PIKKUS	33	24	57			
169	ei		Average of PIKKUS2	171,030303	169,6791667	170,4614035			
176	ei		StdDev of PIKKUS3	8,520141022	8,619794165	8,511814361			
170	ei								
167	ei								
175	ei		t-test						
175	ei		H ₀ : Viirushaiguseid põdenud ja mitte põdenud tudengite keskmised pikkused on võrdsed.						
167	ei		H ₁ : Viirushaiguseid põdenud ja mitte põdenud tudengite keskmised pikkused on erinevad.						
176	ei		või						
161	ei		H ₀ : $\mu_{ei} = \mu_{jah}$	μ_{ei} - viirushaiguseid mitte põdenud tudengite keskmine pikkus					
163	ei		H ₁ : $\mu_{ei} \neq \mu_{jah}$	μ_{jah} - viirushaiguseid põdenud tudengite keskmine pikkus					
161	ei								
163	ei		Tegu on sõitumatute vaatlustega. Seega tuleb enne keskmiste võrdlemist võrrelda dispersioone, otsustamaks, millist t-testi kasutada.						
172	ei								
168	ei		F-test (võrdleme dispersioone)						
168	ei		H ₀ : $\sigma_{ei}^2 = \sigma_{jah}^2$	(pikkuste varieeruvus viirushaiguseid põdenud ja mitte põdenud tudengite gruppides ei ole erinev)					
183	ei		H ₁ : $\sigma_{ei}^2 \neq \sigma_{jah}^2$	(pikkuste varieeruvus viirushaiguseid põdenud ja mitte põdenud tudengite gruppides on erinev)					
174	ei								
189	ei		funktsioon F.TEST	=F.TEST(AC2:AC34;AC35:AC58)					
181	ei								
189	ei								
175	ei								
178	ei								
185	ei								
161	jah								
183	jah								
161	jah								
183,5	jah								
171	jah								
168	jah								
163,3	jah								
158	jah								
170	jah								
162	jah								
173	jah								
167	jah								
166	jah								
169	jah								
180	jah								
183	jah								
178	jah								
168	jah								
160	jah								
161,5	jah								
164	jah								
160	jah								
178	jah								
184	jah								

Function Arguments

F.TEST

Array1 AC2:AC34 = {170;171;172;175;152;164;160;1...}

Array2 AC35:AC58 = {161;183;161;183,5;171;168;163,3;...}

= 0,935567378

Returns the result of an F-test, the two-tailed probability that the variances in Array1 and Array2 are not significantly different.

Array1 is the first array or range of data and can be numbers or names, arrays, or references that contain numbers (blanks are ignored).

Formula result = 0,935567378

[Help on this function](#)

OK Cancel

6. Sõnastage järeldus F-testist, põhjendage.

See ongi põhjendus.
Aru saite?

F-test (võrdleme dispersioone)		
H ₀ : $\sigma_{ei}^2 = \sigma_{jah}^2$	(pikkuste variatsioonid ei erine)	Otsustusreegel: ei põdenud ja mitte põdenud tudengite gruppides ei ole erinev
H ₁ : $\sigma_{ei}^2 \neq \sigma_{jah}^2$	(pikkuste variatsioonid erinevad)	otsus: ei põdenud ja mitte põdenud tudengite gruppides on erinev
funktsioon F.TEST	0,935567378	= p > 0,05 => H ₀ : pikkuste varieeruvus võrreldavates gruppides ei ole erinev
Kasutada võib t-testi, mis eeldab pikkuste võrdset varieeruvust võrreldavates gruppides.		

7. Viige läbi t-test võrdlemaks keskmisi pikkusi.

Tehke seda kahel viisil:

a) kasutades funktsiooni T.TEST:

f_x =T.TEST(AC2:AC34;AC35:AC58;2;2)

AC	AD	AE	AF	AG	AH	AI	AJ	AK	AL
PIKKUS	HAIGE			Haigus					
170	ei			Ei	Jah				
171	ei		Tudengite arv	33	24				
172	ei		Keskmine pikkus	171,0	169,7				
175	ei		Standardhälve	8,5	8,6				
152	ei								
164	ei								
160	ei								
163	ei								
162	ei		Values	Column Labels	jah	Grand Total			
170	ei		Count of PIKKUS	ei	33	24	57		
169	ei		Average of PIKKUS2	171,030303	169,6791667	170,4614035			
176	ei		StdDev of PIKKUS3	8,520141022	8,619794165	8,511814361			
170	ei								
167	ei								
175	ei		t-test						
175	ei		H ₀ : Viirushaiguseid põdenud ja mitte põdenud tudengite keskmised pikkused on võrdsed.						
167	ei		H ₁ : Viirushaiguseid põdenud ja mitte põdenud tudengite keskmised pikkused on erinevad.						
176	ei		või						
161	ei		H ₀ : $\mu_{ei} = \mu_{jah}$	μ_{ei} - viirushaiguseid mitte põdenud tudengite keskmine pikkus					
163	ei		H ₁ : $\mu_{ei} \neq \mu_{jah}$	μ_{jah} - viirushaiguseid põdenud tudengite keskmine pikkus					
161	ei								
163	ei		Tegu on sõltumatute vaatlustega. Seega tuleb enne keskmiste võrdlemist võrrelda dispersioone, otsustamaks, millist t-testi kasutada.						
172	ei								
168	ei		F-test (võrdleme dispersioone)						
168	ei		H ₀ : $\sigma_{ei}^2 = \sigma_{jah}^2$	(pikkuste varieeruvus viirushaiguseid põdenud ja mitte põdenud tudengite gruppides ei ole erinev)					
183	ei		H ₁ : $\sigma_{ei}^2 \neq \sigma_{jah}^2$	(pikkuste varieeruvus viirushaiguseid põdenud ja mitte põdenud tudengite gruppides on erinev)					
174	ei								
189	ei		funktsioon F.TEST	0,935567378	= p > 0,05 => H ₀ : pikkuste varieeruvus võrreldavais gruppides ei ole erinev				
181	ei								
189	ei		Kasutada võib t-testi, mis eeldab pikkuste võrdset varieeruvust võrreldavates gruppides.						
175	ei								
178	ei		Võrdleme keskmisi						
185	ei		funktsioon T.TEST	=T.TEST(AC2:AC34;AC35:AC58;2;2)					
161	jah								
183	jah								
161	jah								
183,5	jah								
171	jah								
168	jah								
163,3	jah								
158	jah								
170	jah								
162	jah								
173	jah								
167	jah								
166	jah								
169	jah								
180	jah								
183	jah								
178	jah								
168	jah								
160	jah								
161,5	jah								
164	jah								
160	jah								
178	jah								
184	jah								

Function Arguments

T.TEST

Array1 AC2:AC34 = {170;171;172;175;152;164;160;163;16}

Array2 AC35:AC58 = {161;183;161;183,5;171;168;163,3;...}

Tails 2

Type 2

Testime kahepoolset hüpoteesi!

Võrdse varieeruvuse eeldusel (miks võib seda eeldada?) sobiv t-testi tüüp.

Returns the probability associated with a Student's t-Test.

Type is the kind of t-test: paired = 1, two-sample equal variance (homoscedastic) = 2, two-sample unequal variance = 3.

Formula result = 0,558781961

[Help on this function](#)

OK Cancel

b) vastava statistikaprotseduuri abil (Data-sakk -> Data Analysis... -> t-Test: ...):

Sõltuvate gruppide (paariviisiline) võrdlus; funktsioonis T.TEST tüüp nr 1

Sõltumatute gruppide võrdlus võrdsete dispersioonide eeldusel; funktsioonis T.TEST tüüp nr 2

Sõltumatute gruppide võrdlus mittevõrdsete dispersioonide eeldusel; funktsioonis T.TEST tüüp nr 3

PIKKUS	HAIGE	Haigus		
		Ei	Jah	
170	ei			
171	ei			
172	ei			
175	ei			
152	ei			
164	ei			
160	ei			
163	ei			
162	ei			
170	ei			
169	ei			
176	ei			
170	ei			
167	ei			
175	ei			
175	ei			
167	ei			
176	ei			
161	ei			
163	ei			
161	ei			
163	ei			
172	ei			
168	ei			
168	ei			
183	ei			
174	ei			
189	ei			
181	ei			
189	ei			
175	ei			
178	ei			
185	ei			
161	jah			
183	jah			
161	jah			
183,5	jah			
171	jah			
168	jah			
163,3	jah			
158	jah			
170	jah			
162	jah			
173	jah			
167	jah			
166	jah			
169	jah			
180	jah			
183	jah			
178	jah			
168	jah			
160	jah			
161,5	jah			
164	jah			
160	jah			
178	jah			
184	jah			

Column Labels: ei, jah, Grand Total

Values:

	ei	jah	Grand Total
Count of PIKKUS	33	24	57
Average of PIKKUS2	171,030303	169,6791667	170,4614035
StdDev of PIKKUS3	8,520141022	8,518791165	8,511810361

t-test

H₀: Viirushaiguseid põdenud

H₁: Viirushaiguseid põdenud või

H₀: μ_e=μ_j

H₁: μ_e≠μ_j

Tegu on sõltumatute vaatl

F-te

H₀: d

H₁: d

funk

Kasu

Võrdleme keskmisi

funktsioon T.TEST

0,558781961

protseduur t-test

8. Sõnastage lõppjäreltus, põhjendage.

funktsioon T.TEST	0,558781961 = $p > 0,05 \Rightarrow H_0$: viirushaiguseid põdenud ja mitte põdenud tudengite		
	keskmised pikkused ei ole erinevad		
protseduur t-test	t-Test: Two-Sample Assuming Equal Variances		
		Variable 1	Variable 2
Mean		171,030303	169,6791667
Variance		72,59280303	74,30085145
Observations		33	24
Pooled Variance		73,30707782	
Hypothesized Mean		0	
df		55	
t Stat		0,588235718	
P(T<=t) one-tail		0,27939098	
t Critical one-tail		1,673033965	
P(T<=t) two-tail		0,558781961 = $p > 0,05 \Rightarrow H_0$: viirushaiguseid põdenud ja mitte põdenud	
t Critical two-tail		2,004044783	tudengite keskmised pikkused ei ole erinevad

Veel võimalusi lõppjäreltuse sõnastamiseks:

- „viirushaiguseid põdenud ja mitte põdenud tudengite keskmiste pikkuste vaheline erinevus ei ole statistiliselt oluline ($p > 0,05$)“, so selline teaduslikumalt sõnastatud järeltus;
- „viirustega nakatumine ei sõltu tudengi pikkusest“ (vähe teise nurga alt sõnastatud järeltus, aga ka õige).

Märkus. Kui nüüd tulemusi veelkord vaadata, siis ilmneb, et viirushaiguseid mitte põdenud tudengid on peaaegu 1,4 sentimeetrit pikemad võrreldes viirushaiguseid põdenud tudengitega. Tõsi ta on. Ainult et tulemuse üldistamiseks (et näiteks soovitada teie kursuse neidudel viirushaigustest hoidumiseks kõrgema kontsaga kingi kanda) on andmeid natuke vähe (või on pikkuse varieeruvus liiga suur, et erinevust selgelt tuvastada; või on ühe-kahe-sentimeetrine erinevus lihtsalt liiga väike).

Seda, et tuvastatud erinevus eksisteerib vaid analüüsitud valimis ega pruugi esineda üldkogumis, näitabki olulisuse tõenäosuse p väärtus 0,56. Seega väites, et kõigi viirushaiguseid mitte põdenud tudengite (ka nende, kelle kohta meil andmeid ei ole) keskmine pikkus erineks viirushaiguseid põdenud tudengite keskmisest pikkusest, eksiksime me 56%-lise tõenäosusega.

Märkus2. Omaette küsimus on muidugi püstitatud hüpoteesi sisuline mõttekus reaalse uuringu jaoks ...

9. Aga mida nendest protseduuri *t-Test: Two-Sample Assuming Equal Variances* tulemustest veel välja saab lugeda?

- Näiteks keskmise pikkuse, pikkuste dispersiooni ja vaatluste arvu võrreldavates gruppides (seejuures on mõistlik peale analüüsi teostamist kirjutada veergudele peale, mis grupi kohta seal olevad arvud käivad):

	Haigus = Ei Variable 1	Haigus = Jah Variable 2
Mean	171,030303	169,6791667
Variance	72,59280303	74,30085145
Observations	33	24

- Ülejäänud osa väljatrükist on seotud juba kahe grupi keskmiste võrdlemisega:

Pooled Variance	73,30707782	Võrdsete dispersioonide eeldusel arvatud ühine dispersioon
Hypothesized Mean	0	
df	55	<i>t</i> -statistiku empiiriline (andmete alusel arvatud) väärtus
t Stat	0,588235718	
P(T<=t) one-tail	0,27939098	1-poolsele hüpoteesile vastav <i>p</i> -väärtus
t Critical one-tail	1,673033965	1-poolsele hüpoteesile vastav <i>t</i> -statistiku kriitiline väärtus
P(T<=t) two-tail	0,558781961	2-poolsele hüpoteesile vastav <i>p</i> -väärtus
t Critical two-tail	2,004044783	2-poolsele hüpoteesile vastav <i>t</i> -statistiku kriitiline väärtus

Kahepoolne hüpotees tähendab statistikas võrdumise ja mitte võrdumise testimist:

$$H_0: \mu_{ei} = \mu_{jah},$$

$$H_1: \mu_{ei} \neq \mu_{jah}.$$

Ühepoolne hüpotees testib vaid ühepoolset erinevust – kas näiteks üks keskmine on teisest väiksem või mitte. Seejuures konstrueerib *Excel* hüpoteeside paari vastavalt keskmistele väärtustele, testides alati seda, kas suurem keskmine on ikka statistiliselt oluliselt suurem või mitte. Antud juhul on testitav ühepoolne hüpoteeside paar seega kujul

$$H_0: \mu_{ei} \leq \mu_{jah},$$

$$H_1: \mu_{ei} > \mu_{jah},$$

sest haige mitte olnute keskmine pikkus on suurem.

Olulisuse tõenäosuse *p* asemel võib otsuse vastu võtmisel lähtuda ka teststatistiku empiirilise (andmete alusel arvatud) väärtuse absoluutväärtuse $|t|$ ja teststatistiku kriitilise väärtuse $t_{critical}$ võrdlusest (vt hüpoteeside kontroll, 3. loeng).

Nimelt, kuna $|t| = 0,59 < 2,00 = t_{critical}$, siis ei ole põhjust nullhüpoteesi ümber lükata

(*t*-statistiku väärtus jääb sellesse piirkonda, kuhu ta nullhüpoteesi kehtides 95%-tõenäosusega peabki jääma).

10. Kas testides ühepoolset hüpoteesi võinuks lugeda alternatiivse hüpoteesi H_1 tõestatuks?

Kuidas kõlab lõppjärelus testitud ühepoolse hüpoteesi kohta?