LATVIJAS SPORTA PEDAGOĢIJAS AKADĒMIJA

Juris Dravnieks

MS EXCEL pievienojumprogramma

STATISTIKA 3.0

Mācību līdzeklis - rokasgrāmata
LSPA studentiem, maģistrantiem, doktorantiem

Papildināts

RĪGA - 2011

Mācību līdzeklis „MS EXCEL pievienojumprogramma STATISTIKA 2.1" paredzēts Latvijas sporta pedagoģijas akadēmijas studentiem, maģistrantiem un doktorantiem kā rokasgrāmata zinātniskās informācijas apstrādei. Šeit apskatīti populārāko matemātiskās statistikas metožu algoritmi, metožu izmantošanas noteikumi, aprakstīta datu analīze, izmantojot autora izstrādāto programmu STATISTIKA (pievienojumprogramma veidota, izmantojot Visual Basic vor MS Excel).

SATURS

IEVADS

Topošie bakalauri, maģistri un doktori mēdz apstrādāt savus pētījumu datus, izmantojot LSPA datorklases pakalpojumus. Matemātiskās statistikas metožu izmantošana ir neatņemama promocijas darbu izstrādes procedūra. Datoru bāzes programmatūra – operētājsistēma **Windows** un **MS Office** ļauj tikai daļēji un ar zināmām neērtībām atrisināt šos uzdevumus. Tabulu procesora MS Excel programmu paketē ir komponents **Analysis Toolpak** ar statistikas uzdevumu risināšanai paredzētām palīgprogrammām, kuras studenti detalizēti iepazīst tikai 2. studiju gada kursā "Informācijas un komunikāciju tehnoģijas sportā". Vajadzība pēc statistikas metodēm uzdevumos, kas saistīti ar bakalaura darba izstrādes jautājumiem studiju kursa "Pētniecības metodoloģija" ietvaros, parādās jau 1. studiju gadā. Līdz ar to kļūst pamanāmi arī vairāki faktori, kas studijas traucē:

- MS EXCEL standartprogrammas neaptver visus sporta zinātnes specifikai raksturīgos statistiskās analīzes variantus (nav ļoti vajadzīgo neparametrisko kritēriju un rangu korelācijas analīzes);
- standartprogrammās lietotāja dialogs ir angļu valodā, piesātināts ar daudziem jauniem terminiem un raksturojumiem, kurus lielākajā daļā gadījumu neizmanto;
- visas palīgprogrammas izskaitļotos analīzes rezultātus piedāvā neformatētā veidā (8 un vairāk decimālzīmes aiz komata);
- tas viss kopumā palēnina šo programmu apgūšanu un palielina uzdevumu risināšanas darbietilpību.

Mūsu izstrādātajai MS EXCEL tabulu procesora Visual Basic aplikācijai - programmai STATISTIKA šie trūkumi nepiemīt. Tā aptver visbiežāk sastopamos statistiskās analīzes variantus, visos gadījumos lietotāja interfeiss ir līdzīgs, dialogs – latviešu valodā, tiek aprēķināti tikai nepieciešamie raksturojumi un rezultāti attēloti formatētā veidā – atbilstoši sākuma datu precizitātei (bez liekajām decimālzīmēm aiz komata). Katra palīgprogramma paziņo secinājumu par statistiskā raksturojuma (vidējo aritmētisko starpības, vidējā pieauguma vai korelācijas) ticamību. Programma ieviesta studiju procesā, un iemācīties ar to strādāt var ļoti ātri.

Metodiskā līdzekļa pirmajā nodaļā ļoti konspektīvi aprakstīti visbiežāk izmantojamo matemātiskās statistikas metožu algoritmi un izmantošanas noteikumi. Otrajā nodaļā dots programmas STATISTIKA apraksts un izmantošanas instrukcija.

1. MATEMĀTISKĀS STATISTIKAS METODES (ALGORITMI)

1.1. Kad izmanto matemātiskās statistikas metodes?

Sporta zinātnē un praksē visbiežāk sastopami sekojoši statistiskās analīzes varianti.

1. Lai raksturotu sportistu grupu kopumā, analizē sacensību vai testēšanas rezultātus (paraugkopu) un aprēķina vidējo rādītāju (parasti vidējo aritmētisko, kas raksturo dotās grupas sagatavotības līmeni), variēšanas rādītājus - standartnovirzi un variācijas koeficientu (raksturo rezultātu blīvumu) un vidējā rādītāja standartkļūdu (raksturo vidējā rādītāja neprecizitāti, vispārinot to uz visiem šādas kategorijas sportistiem - ģenerālkopu).

2. Pētot treniņa efektivitāti, aprēķina grupas vidējo rezultātu pieaugumu aizvadītajā laika periodā un novērtē tā ticamību, izmantojot noteiktu iepriekš pieņemtu ticamības līmeni. Šai procedūrā lieto statistikas metodes, kuras sauc par atšķirību kritērijiem. Līdzīgi rīkojas, ja vēlas noskaidrot, vai kopumā ņemot viena sportistu grupa sagatavota labāk nekā otra - novērtē grupu vidējo rezultātu starpības ticamību.

3. Sportā sastopamies ar parādību, ko sauc par trenētības pārnešanu - trenējoties vienā fiziskajā vingrinājumā bieži novērojam rezultātu pieaugumu arī kādā citā vingrinājumā, kuram līdzīga kustību struktūra un fizioloģiskais mehānisms. Lai novērtētu, cik cieša ir šī sakarība, t.i., cik lielā mērā sasniegumi vienā vingrinājumā ietekmē sasniegumus citā vingrinājumā, izmanto korelācijas analīzi. Korelācijas koeficients ir atkarības ciešuma mērs.

1.2. Aprakstošā statistika

Aprakstošās statistiskās uzdevums - raksturot pētāmās pazīmes vidējo vērtību, variēšanu un to, cik reprezentatīvi ir šie raksturojumi - cik labi paraugkopa reprezentē ģenerālkopu. Aprēķina četrus skaitļus, kuri koncentrētā veidā satur vajadzīgo informāciju. Tos sauc par statistiskajiem rādītājiem. Ir trīs statistisko rādītāju grupas: vidējie rādītāji, variēšanas jeb izkliedes rādītāji un reprezentācijas jeb standartkļūdas. Pētījumos, kuru rezultāti ir tieši vai netieši mērīšanas rezultāti, no vidējiem rādītājiem parasti izmanto vidējo aritmētisko, no variēšanas rādītājiem - standartnovirzi (vidējo kvadrātisko novirzi) un variācijas koeficientu. Reprezentācijas kļūdu (standartkļūdu), kura raksturo neprecizitāti, kas rodas, vispārinot paraugkopas raksturojumus uz ģenerālkopu, aprēķina vidējam aritmētiskajam.

1.2.1. Vidējais aritmētiskais

Vidējais aritmētiskais ir visu mērīšanas rezultātu (varianšu) summa dalīta ar to skaitu:

$$\bar{x} = \frac{x_1 + x_2 + x_3 + ... + x_{n-1} + x_n}{n} \qquad (1.1),$$

Ģenerālkopai šis rādītājs nav izskaitļojams, jo nevar iegūt un apstrādāt visas šīs kopas variantes. Paraugkopas vidējam aritmētiskajam atbilstošo lielumu ģenerālkopā sauc par vidējo ģenerālo jeb pazīmes vidējo vērtību un apzīmē ar grieķu burtu μ (mī).

Mēdz teikt, ka vidējais ģenerālais ir vidējā aritmētiskā matemātiskā cerība. Vidējais aritmētiskais labi raksturo pazīmes vidējo vērtību, t.i., sportistu grupas vidējo sacensību vai kontroles vingrinājuma rezultātu, funkcionālās sagatavotības vidējo līmeni u.c. Vidējo aritmētisko aprēķina pēc formulas:

$$\bar{x} = \frac{\sum x_i}{n} \qquad (1.2),$$

kur x_i – variante jeb atsevišķs mērījuma rezultāts; n - novērojumu skaits. Vidējais aritmētiskais ir nosaukts skaitlis - tam ir tā pati mērvienība, kas atsevišķai variantei. Aprēķināto vidējo noapaļo līdz tādai precizitātei, ar kādu dotas paraugkopas variantes. Neierobežoti palielinot pazīmes (gadījumlieluma) savstarpēji neatkarīgo novērojumu skaitu, iegūto rezultātu vidējā vērtība tuvojas noteiktam konstantam lielumam - $\bar{x}$ matemātiskai cerībai μ.

1.2.2. Standartnovirze un variācijas koeficients

Vidējie rādītāji nav universāli, jo pazīmes ar vienādiem vidējiem var atšķirties pēc variēšanas lieluma un rakstura. Variēšanas sinonīmi no statistikas viedokļa ir jēdzieni: rezultātu izkliede, blīvums, vienveidība, mainīgums. Tie visi raksturo vienu parādību - pazīmes variēšanu jeb individuālo rezultātu atšķirības. Mazākajai variācijai raksturīgas mazākas individuālo rezultātu savstarpējās atšķirības. Arī sportista meistarību raksturo ne tikai atkārtotos mēģinājumos sasniegtais vidējais vai augstākais rezultāts, bet arī rezultātu stabilitāte atkārtotos mēģinājumos. Tādēļ kopā ar vidējo rādītāju pazīmes raksturošanai izmanto arī variēšanas rādītājus. Pazīmes variēšanu raksturo, izpētot variantšu izkliedi ap vidējo aritmētisko. Par pamatu šādam variēšanas vērtējumam izmanto variantšu centrālās novirzes $x_i - \bar{x}$, aprēķinot vidējo kvadrātisko novirzi jeb standartnovirzi.

Standartnovirzi aprēķina pēc darba formulas:

$$s = \sqrt{\frac{\sum x_i^2 - \frac{\left(\sum x_i\right)^2}{n}}{n-1}} \qquad (1.3)$$

Palielinoties pazīmes variēšanai pieaug standartnovirzes vērtība, savukārt mazāka standartnovirze atbilst blīvākiem, vienveidīgākiem rezultātiem. Standartnovirze ir nosaukts skaitlis, tai ir tā pati mērvienība, kas variantēm. Aprēķināto standartnovirzes vērtību noapaļo līdz precizitātei, ar kādu dotas variantes. Standartnovirzei ir divi trūkumi, kuru dēļ to nevar vienmēr izmantot. Standartnovirzes vērtība atkarīga no vidējā aritmētiskā. Tādēļ, ja vēlas salīdzināt rezultātu variēšanu divām sportistu grupām, standartnovirzi var izmantot šim nolūkam tikai tad, ja abu grupu vidējie (aritmētiskie) rezultāti ir vienādi. Ja vēlas salīdzināt divu dažādu pazīmju variēšanu, traucē arī mērvienība - standartnovirzes, kas, izteiktas dažādās mērvienībās, nav salīdzināmas. Šo iemeslu dēļ variēšanas vērtēšanai parasti izmanto standartnovirzes relatīvo vērtību - variācijas koeficientu. To veido standartnovirzes attiecība pret vidējo aritmētisko procentos:

$$s_{\%} = \frac{s}{\bar{x}} \cdot 100\% \qquad (1.4)$$

Variācijas koeficients ir nenosaukts skaitlis un līdz ar to universāls izkliedes rādītājs. Aprēķināto variācijas koeficientu noapaļo līdz vienai decimālzīmei aiz komata. Svarīgi ir atcerēties variācijas koeficienta robežvērtību - 10%. Ja $s\% \leq 10\%$ novērojumu rezultāti ir vienveidīgi, pretējā gadījumā tos par vienveidīgiem uzskatīt nevar un jānoskaidro lielās variācijas cēloņi. Tie var būt rupjas mērīšanas kļūdas vai arī dotajai grupai netipiska objekta klātbūtne novērojumos. Variācijas koeficienta izmantošana ir ierobežota, t.i., tas nav derīgs, ja mērot izmantota intervālu skala, piemēram, mērot leņķus vai temperatūru. Arī pazīmes diskrētās variēšanas gadījumā (novērojumu rezultāti – veseli skaitļi) variācijas koeficients nav informatīvs.

1.2.3. Vidējā aritmētiskā standartkļūda

Vidējā aritmētiskā standartkļūda raksturo neprecizitāti, kas rodas, vispārinot paraugkopas vidējo aritmētisko uz ģenerālkopu.

$$s_{\bar{x}} = \frac{s}{\sqrt{n}} \qquad (1.5)$$

Secinājumos, tekstā un tabulās parasti standartkļūdu raksta kopā ar vidējo aritmētisko, atdalot ar plus-mīnus zīmi: $\bar{x} \pm s_{\bar{x}}$.

1.3. Par normālo sadalījumu

Katrai pazīmei raksturīga noteikta variēšanas likumsakarība. Aprakstot šo sakarību, kas ir kopīga kādai pazīmju grupai, ar atbilstošu vienādojumu formulē teorētisko sadalījumu. Tas ir dotās grupas pazīmju variēšanas vispārīgs likums jeb matemātiskais modelis. Zināmi vairāki teorētiskie sadalījumi. Bioloģiskās pazīmes bieži variē atbilstoši normālā sadalījuma likumam, kuru pirmo reizi 1733.g. formulējis A. Muaurs, pēc tam neatkarīgi viens no otra - Laplass un K. Gauss. **Ņemot par pamatu normālā sadalījuma likumu, ir izstrādātas daudzas statistiskās analīzes metodes. To izmantošana konkrētā gadījumā korekta tikai tad, ja nav šaubu, ka pētāmā pazīme variē atbilstoši normālajam likumam.** Maza novērojumu skaita gadījumā ($n < 30$) visbiežāk šīs atbilstības nebūs, tādēļ tā jāpārbauda ar speciālām, šim nolūkam paredzētām metodēm [2]. Atsevišķos gadījumos par datu atbilstību normālajam likumam var spriest arī pēc praktiskās pieredzes un teorētiskiem apsvērumiem. Piemēram, dati, kurus veido veseli skaitļi, kas variē nelielā intervālā (piemēram, skolas atzīmes 5 vai 10 ballu sistēmā), neatkarīgi no novērojumu skaita nekad neatbilst normālajam sadalījumam. Ja novērojumu skaits ir liels ($n = 50 \div 100$) un pazīme variē nepārtraukti (mērskaitļi ar decimāldaļu), parasti par pietiekoši ticamu datu atbilstību normālajam sadalījumam nav šaubu. Atsevišķos gadījumos, kad rezultāti ir veseli skaitļi, kas variē relatīvi plašā diapazonā (piemēram, no 30 līdz 150) un novērojumu skaits sniedzas vairākos simtos, variācija tuvojas normālajam sadalījumam.

Bioloģiskajos un sporta pētījumos bieži sastopamies ar neliela apjoma eksperimentālām grupām. Ja pazīme variē atbilstoši normālajam likumam, šādas grupas rezultātu kopā biežāk sastopami skaitļi, kas atrodas tuvāk vidējai vērtībai. Ņemot vērā šo parādību, angļu statistiķis Viljams Gossets izstrādāja normālā sadalījuma variantu mazām kopām, ko publicējot nosauca sava pseidonīma vārdā par Stjūdenta sadalījumu. Tā nozīmīgākais raksturojums ir Stjūdenta sadalījuma normētā novirze - $t_{\alpha;\nu}$, kuras vērtība atkarīga no brīvības pakāpju skaita ν un pieņemtā būtiskuma līmeņa α. Stjūdenta sadalījuma normēto novirzi ar nosaukumu „Stjūdenta kritērija teorētiskā vērtība” izmanto vairākās matemātiskās statistikas metodēs. Šī rādītāja vērtību var nolasīt Stjūdenta tabulā (1.pielikums).

Datu atbilstību normālajam sadalījumam var pārbaudīt ar apakšprogrammu „Aprakstošā statistika” (asimetrijas un ekscesa metode).

1.4. Atšķirību novērtēšana

1.4.1. Nulles hipotēze un tās pārbaude

Vai starp divām sportistu grupām sagatavotības ziņā ir būtiskas atšķirības? Vai dotās sportistu grupas sagatavotības līmenis aizvadītajā laika periodā ir būtiski mainījies? Šos jautājumus risina, pārbaudot nulles hipotēzi - pieņēmumu, ka divu ģenerālkopu rādītāju starpība ir nulle, t.i., bezgalīgi palielinot salīdzināmo paraugkopu apjomus, iegūst vienu un to pašu ģenerālkopu. Pārbaudes rezultātā nulles hipotēzi pieņem vai noraida. Lēmumu pieņem nevis kā absolūtu patiesību, bet gan ar vajadzīgo ticamības līmeni (sporta pētījumos - $P = 0{,}95$) vai pieļaujamās kļūdas varbūtību – būtiskuma līmeni ($\alpha = 0{,}05$). Tātad pieļaujam, ka 5% gadījumu iespējama kļūda.

Pārbaudi veic, izmantojot kādu parametrisku vai neparametrisku metodi. Šīs metodes parasti sauc par atšķirību kritērijiem. Ja novērtē nulles hipotēzi par divu kopu parametriem - statistiskiem rādītājiem, tad metode ir parametrisks kritērijs, kas pamatojas uz normālā sadalījuma likumu. Šādu metožu izmantošanas iespējas ir ierobežotas. Metodes, ar kurām novērtē atšķirības kopumā, sauc par neparametriskajiem kritērijiem. To izmantošanai nav ierobežojumu.

Vispārējos vilcienos iespējami divi nulles hipotēzes pārbaudes iznākumi.

- Vidējo rādītāju starpība nav statistiski ticama. Tas nozīmē, ka tai ir gadījuma jeb nejaušības raksturs, tās cēlonis ir pazīmes variēšana, un praksē jāpieņem, ka šīs starpības vērtība ir nulle. Tas ir gadījums, kad divu paraugkopu vidējo rādītāju starpība vairāk vai mazāk atšķiras no nulles arī tad, ja kopas pieder vienai ģenerālkopai - tātad ir līdzīgas. Citiem vārdiem, salīdzināmie dati iegūti divās vienādi sagatavotās sportistu grupās vai arī vienīgās pētītās grupas stāvoklis (sagatavotība) aizvadītajā laika periodā nav mainījies.
- Vidējo rādītāju starpība ir statistiski ticama. Tas ir gadījums, kad salīdzināmās paraugkopas ir veidotas no dažādām ģenerālkopām. Citiem vārdiem - salīdzināmie dati iegūti divās dažāda sagatavotības līmeņa sportistu grupās vai arī vienīgās pētītās grupas stāvoklis (sagatavotība) aizvadītajā laika periodā ir mainījies, t.i., pārgājis jaunā kvalitātē.

Lai izšķirtos par to, ar kuru no šiem gadījumiem sastopamies risināmajā uzdevumā, starpību standartizē - ar skaitļošanas operāciju palīdzību atbrīvojas no mērvienības. Iegūst nenosauktu skaitli, ko statistikas valodā sauc par atšķirību kritērija empīrisko vērtību. Pieņemtajam ticamības līmenim un novērojumu skaitam atbilstošo lielāko, pieļaujamo starpību (kas pielīdzināma nullei), t.i., atšķirību kritērija teorētisko vērtību nolasa speciālā tabulā. Šāda tabula sastādīta katrai metodei un atrodama attiecīgās statistikas mācību grāmatas vai rokasgrāmatas pielikumā. Salīdzina kritērija empīrisko vērtību ar teorētisko vērtību. Ja kritērija empīriskā vērtība lielāka par teorētisko vērtību, vidējo rādītāju starpību uzskata par ticamu, t.i., nulles hipotēzi noraida. Pretējā gadījumā nulles hipotēze ir pareiza.

Ja empīriskie dati atbilst normālam sadalījumam, uzdevuma risināšanai var izmantot kā parametriskas, tā neparametriskas metodes. Pretējā gadījumā drīkst izmantot tikai neparametriskas metodes. Skaitļošanas operāciju darbietilpības ziņā neparametriskie kritēriji ir vienkāršāki. Turpretī ar parametrisku metodi iegūts slēdziens ir precīzāks, jo šo kritēriju izšķiršanas spēja jeb jūtība ir lielāka.

No turpmāk apskatītajām parametriskas metodes ir Stjūdenta t-kritērijs un lineārā pāru korelācijas analīze. Neparametriskās metodes ir Vilkoksona kritērijs, Van der Vardena kritērijs un Spirmena rangu korelācijas analīze.

1.4.2. Stjūdenta t-kritērijs saistītām paraugkopām

Par saistītām sauc paraugkopas, kas iegūtas, pētot vienu grupu atkārtoti pēc noteikta laika intervāla. Piemēram, piecas saistītas paraugkopas ir vienas un tās pašas studentu grupas sasniegumi kādā testā iestājeksāmenā un 1., 2., 3., 4. studiju gadā. Sportistu grupas vidējo sacensību vai testa rezultāta izmaiņu ticamību noteiktā laika periodā (no mēģinājuma uz mēģinājumu, dažādos treniņa periodos utt.) novērtē, izmantojot Stjūdenta kritēriju saistītām kopām. Primitīviem aprēķiniem (ar parasto kalkulatoru) izmanto darba tabulu, kuras (1. tabula) 1. kolonnā ieraksta beigu rezultātu $x_{2.i}$, 2. kolonnā - sākuma rezultātu $x_{1.i}$, 3. kolonā - rezultātu starpību - d_i , 4. kolonnā starpības kvadrātu. Aprēķina summas Σd_i ; Σd_i^2.

1. tabula

Palīg lielumu aprēķināšana Stjūdenta *t*-kritērijam

Gala rezultāts $x_{2.i}$	Sākuma rezultāts $x_{1.i}$	Starpība $d_i = x_{2.i} - x_{1.i}$	d_i^2
$x_{2.1}$	$x_{1.1}$	d_1	d_1^2
$x_{2.2}$	$x_{1.2}$	d_2	d_2^2
$x_{2.3}$	$x_{1.3}$	d_3	d_3^2
$x_{2.n}$	$x_{1.n}$	d_n	d_n^2
		Σd_i	Σd_i^2

Aprēķina vidējo pieaugumu (starpību):

$$\bar{d} = \frac{\sum d_i}{n} \qquad (1.6)$$

kur n - rezultātu pāru skaits.

Aprēķina vidējās starpības standartkļūdu:

$$s_{\bar{d}} = \sqrt{\frac{\sum d_i^2 - \frac{\left(\sum d_i\right)^2}{n}}{n(n-1)}} \qquad (1.7)$$

Aprēķina Stjūdenta kritērija empīrisko vērtību:

$$t = \frac{\left|\bar{d}\right|}{s_{\bar{d}}} \qquad (1.8)$$

Stjūdenta tabulā (1. pielikums) pēc $\alpha = 0,05$ un $\nu = n - 1$ nolasa Stjūdenta sadalījuma normēto novirzi $t_{\alpha;\nu}$. Ja $t \geq t_{\alpha;\nu}$, tad vidējais pieaugums ir statistiski ticams ($\alpha < 0,05$). Ja $t < t_{\alpha;\nu}$, tad vidējai rezultātu starpībai ir gadījuma raksturs, t.i., tā nav statistiski ticama ($\alpha > 0,05$).

1.4.3. Vilkoksona kritērijs

Vilkoksona kritēriju izmanto saistītu paraugkopu atšķirību novērtēšanai. Aprēķiniem veido darba tabulu (2. tabula), kuras 1. kolonnā ieraksta beigu rezultātu $x_{2.i}$, 2. kolonnā - sākuma rezultātu $x_{1.i}$, 3. kolonnā - rezultātu starpību - d_i , 4. kolonnā katrai starpībai pēc absolūtās vērtības piešķir rangu g_i , kuram pieraksta attiecīgās starpības zīmi. Atsevišķi aprēķina pozitīvo rangu summu $T_{(+)} = \Sigma g_{(+)}$ un negatīvo rangu summu $T_{(-)} = \Sigma g_{(+)}$. Pēc absolūtās vērtības mazāko no šīm summām izmanto par kritērija empīrisko vērtību.

2. tabula

Palīg lielumu aprēķināšana Vilkoksona kritērijam

Gala rezultāts $x_{2.i}$	Sākuma rezultāts $x_{1.i}$	Starpība $d_i = x_{2.i} - x_{1.i}$	$\lvert d_i\rvert$ rangs g_i
$x_{2.1}$	$x_{1.1}$	d_1	g_1
$x_{2.2}$	$x_{1.2}$	d_2	g_2
$x_{2.3}$	$x_{1.3}$	d_3	g_3
$x_{2.n}$	$x_{1.n}$	d_n	g_n
			Σg_i

Speciālā tabulā (2. pielikums) pēc rezultātu pāru skaita n (ja rezultātu starpība ir nulle, pāri neskaita) un $\alpha = 0,05$ nolasa $T_{\alpha;n}$ - kritērija robežvērtību. Ja $T \leq T_{\alpha;n}$ rezultātu izmaiņas ir būtiskas, pretējā gadījumā ($T > T_{\alpha;n}$) atšķirībām ir gadījuma raksturs.

1.4.4. Stjudenta t-kritērijs neatkarīgām paraugkopām

Neatkarīgas kopas veido variantes, kas iegūtas, pētot divas vai vairākas objektu grupas (piemēram, 3[A]. un 3[B]. klases zēnu sasniegumi bumbiņas mešanā).

Vispirms iegūst aprakstošo statistiku, aprēķinot abu salīdzināmo paraugkopu vidējos aritmētiskos, standartnovirzes u.c. Pēc tam aprēķina Stjūdenta kritērija empīrisko vērtību:

$$t = \frac{\left|\bar{x}_1 - \bar{x}_2\right|}{s_{\bar{x}_1 - \bar{x}_2}} \qquad (1.9),$$

kur $\bar{x}_1$ un $\bar{x}_2$ - attiecīgi pirmās un otrās paraugkopas vidējie aritmētiskie; $s_{\bar{x}_1 - \bar{x}_2}$ - vidējo aritmētisko starpības standartkļūda. Šo lielumu aprēķina pēc formulas:

$$s_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{s_1^2(n_1 - 1) + s_2^2(n_2 - 1)}{n_1 + n_2 - 2} \cdot \frac{n_1 + n_2}{n_1 \cdot n_2}} \qquad (1.10)$$

Stjūdenta tabulā (1.pielikums) pēc $\alpha = 0{,}05$ un $\nu = n_1 + n_2 - 2$ nolasa kritērija teorētisko vērtību $t_{\alpha;\nu}$. Ja $t \geq t_{\alpha;\nu}$, tad atšķirības ir statistiski ticamas, turpretī pie $t < t_{\alpha;\nu}$ rezultātu atšķirības nevar uzskatīt par pierādītām, jāpieņem, ka tām ir gadījuma raksturs.

1.4.5. Van der Vardena kritērijs

Van der Vardena kritēriju izmanto neatkarīgu paraugkopu salīdzināšanai. Tā ir viena no precīzākajām, bet arī darbietilpīgākā no neparametriskajām metodēm.

Darba tabulā abu paraugkopu variantes uzraksta vienā ranžētā rindā, bet katru kopu savā kolonnā - $x_{1.i}$ vai $x_{2.i}$ (3. tabula). 3. kolonnā ieraksta varianšu rangus g_i (vienādiem rezultātiem dod vienādu rangu, t.i., ieņemto vietu vidējo aritmētisko). 4. kolonnā **mazākās** kopas variantēm aprēķina $\frac{g_i}{n_1 + n_2 + 1}$, kur n_1 un n_2 ir attiecīgi pirmās un otrās paraugkopas apjoms.

11. pielikumā katram dalījumam nolasa $\Psi\left(\frac{g_i}{n_1 + n_2 + 1}\right)$ vērtību un ieraksta 5. kolonnā.

Ja dalījums mazāks par 0,5, tad nolasa pēc $1 - \frac{g_i}{n_1 + n_2 + 1}$ un nolasīto vērtību raksta ar mīnus zīmi. Aprēķina summu 5. stabiņā - Van der Vardena kritērija empīrisko vērtību:

$$X = \sum \Psi\left(\frac{g_i}{n_1 + n_2 + 1}\right) \qquad (1.11)$$

Pēc n_1+n_2, un $|n_1 - n_2| = 1$ un $\alpha = 0{,}05$ 3. pielikuma tabulā nolasa Van der Vardena kritērija teorētisko vērtību X_α . Nulles hipotēzi pieņem, ja $X \leq X_\alpha$. Ja $X > X_\alpha$, nulles hipotēze jānoraida, t.i., atšķirības ir ticamas.

3. tabula

Darba tabula Van der Vardena kritērija aprēķināšanai

$x_{1.i}$	$x_{2.i}$	g_i	$\frac{g_i}{n_1+n_2+1}$	$\Psi(\frac{g_i}{n_1+n_2+1})$
$x_{1.1}$		1		
$x_{1.2}$		2		
$x_{1.3}$		3		
$x_{1.4}$		4		
$x_{1.5}$		5		
$x_{1.6}$		6		
	$x_{2.1}$	7	0,46875	-0,08
	$x_{2.2}$	8	0,46875	-0,08
$x_{1.7}$		9		
	$x_{2.3}$	10	0.625	0,32
$x_{1.8}$		11		
	$x_{2.4}$	12	0.75	0,67
	$x_{2.5}$	13	0.8125	0,89
	$x_{2.6}$	14	0.875	1,15
	$x_{2.7}$	15	0.935	1,535
				$X = 4{,}405$

1.5. Korelācijas analīze

1.5.1. Lineārās pāru korelācijas koeficients

Korelācijas analīzes uzdevums - kvantitatīvi novērtēt atkarību starp divām vai vairākām pazīmēm. Korelācijas analīzes izmantošana ir korekta tikai tad, ja tiek vērtēts loģiski pamatotas pazīmju savstarpējās atkarības ciešums (programmas „Pīrsona korelācijas koeficients" vai „Spirmena rangu korelācijas koeficients"). Korelācijas analīze ir formāla matemātiska metode, kuru nedrīkst izmantot atkarību meklēšanai (tādēļ jābūt piesardzīgiem, interpretējot ar apakšprogrammu „Korelācijas matrica" iegūtos datus, jo programma nešķiro loģiskās un nejaušās, šķietamās sakarības) .

Pāru korelācijas koeficients kvantitatīvi raksturo pazīmju savstarpējās ietekmes stiprumu un virzienu. Pāru korelācija ir atkarība starp divām pazīmēm, kuru pēta neņemot vērā varbūtēju citu pazīmju ietekmi uz tām. Pāru korelācija var būt lineāra vai nelineāra. Ja vienādām faktorālās pazīmes vērtību izmaiņām atbilst aptuveni vienādas rezultatīvās pazīmes vērtību izmaiņas, tad korelācija ir lineāra, t.i., tā tuvojas lineārai funkcijai. Nelineārā atkarība tuvojas kādai līklīnijas funkcijai. Pāru korelācijas koeficientu sauc arī par Pīrsona korelācijas koeficientu, un tas raksturo lineārās atkarības virzienu un ciešumu. To var izmantot atkarības novērtēšanai, ja pierādīta novērojumu rezultātu atbilstība normālajam sadalījumam. Paraugkopas korelācijas koeficientu apzīmē ar latīņu burtu *r*, bet ģenerālo korelācijas koeficientu ar grieķu alfabēta burtu ρ (ro). Paraugkopas korelācijas koeficientu aprēķina pēc formulas:

$$r = \frac{\sum x_i y_i - \frac{\sum x_i \cdot \sum y_i}{n}}{\sqrt{\sum x_i^2 - \frac{\left(\sum x_i\right)^2}{n}} \cdot \sqrt{\sum y_i^2 - \frac{\left(\sum y_i\right)^2}{n}}} \qquad (1.12)$$

kur x_i - faktorālās pazīmes variante; y_i - rezultatīvās pazīmes variante; n - variančšu pāru skaits jeb paraugkopas apjoms. Aprēķināto korelācijas koeficientu noapaļo līdz 3 zīmēm aiz komata. Tas ir nenosaukts skaitlis un atrodas robežās no -1 līdz +1. Jo vairāk koeficienta modulis tuvojas 1, jo ciešāka korelācija, jo spēcīgāk viena pazīme ietekmē otru.

Lai pārliecinātos, vai aprēķināto Pīrsona korelācijas koeficientu var izmantot atkarības vērtēšanai, koeficienta moduli salīdzina ar kritisko vērtību $r_{\alpha;n}$, kuru nolasa tabulā (4. pielikums) atbilstoši paraugkopas apjomam n un rezultātu būtiskuma līmenim $\alpha = 0{,}05$. Ja koeficienta modulis | r | lielāks vai vienāds ar kritisko vērtību, korelācija ir ticama un var izdarīt secinājumus par tās virzienu un ciešumu. Ja koeficients nav ticams ($r < r_{\alpha;n}$), tad pie dotā novērojumu skaita atkarību nav izdevies pierādīt, un pazīmes jāuzskata par neatkarīgām. Jāņem vērā, ka pēc korelācijas analīzes rezultātiem daudzmaz nopietnu slēdzienu var izdarīt, ja novērojumu skaits nav mazāks par 50. Tikai ticamas korelācijas gadījumā ($r \geq r_{\alpha;n}$) var vērtēt tās virzienu un ciešumu. Ja $r > 0$, tad korelācija ir pozitīva, t.i., pieaugot faktorālās pazīmes X vērtībām attiecīgi palielinās rezultatīvās pazīmes Y vērtības. Negatīvas korelācijas gadījumā ($r < 0$), palielinoties vienas pazīmes vērtībām, otras pazīmes vērtības samazinās. Raksturojot pazīmju savstarpējās ietekmes spēku, korelāciju vērtē kā: vāju, ja $0{,}2 < |r| < 0{,}49$ vidēju, ja $0{,}5 < |r| < 0{,}69$ ciešu, ja $0{,}7 < |r| < 0{,}99$.

Lai veiktu aprēķinus, novērojuma rezultātus ierakstām darba tabulā (4. tabula), aprēķinām x_i^2, y_i^2, $x_i.y_i$, bet pēc tam ierakstām summu katrā tabulas stabiņā: $\sum x_i$, $\sum y_i$, $\sum x_i^2$, $\sum y_i^2$, $\sum x_i.y_i$.

4. tabula

Darba tabula korelācijas koeficienta aprēķināšanai

Faktorālā pazīme x_i	Rezultatīvā pazīme y_i	x_i^2	y_i^2	$x_i.y_i$
x_1 x_2 x_3 ... x_{n-1} x_n	y_1 y_2 y_3 ... y_{n-1} y_n	x_1^2 x_2^2 x_3^2 ... x_{n-1}^2 x_n^2	y_1^2 y_2^2 y_3^2 ... y_{n-1}^2 y_n^2	$x_1.y_1$ $x_2.y_3$ $x_3.y_3$ ... $x_{n-1}.y_{n-1}$ $x_n.y_n$
$\sum x_i$	$\sum y_i$	$\sum x_i^2$	$\sum y_i^2$	$\sum x_i.y_i$

Tabulā (4. pielikums) atrodam $r_{\alpha;n}$. Korelācija ir ticama, ja $|r| > r_{\alpha;n}$.

1.5.2. Spirmena rangu korelācijas koeficients

Šo rādītāju izmanto divu pazīmju savstarpējās atkarības novērtēšanai, ja nav pierādīta novērojumu rezultātu atbilstība normālajam sadalījumam, kā arī tad, ja viena vai abas pazīmes novērtētas, izmantojot rangu skalu. Sastāda darba tabulu (5. tabula), kuras pirmajās divās ailēs ieraksta attiecīgi pa pāriem saistītās faktorālās un rezultatīvās pazīmes variantes - x_i un y_i , 4. stabiņā ieraksta šo varianšu rangus $g_{x.i}$ un $g_{y.i}$. Aprēķina pa pāriem saistīto rangu starpības - $d_i = g_{x.i} - g_{y.i}$ un ieraksta tās 5. stabiņā, bet 6. stabiņā - šo starpību kvadrātus d_i^2. Aprēķina starpību kvadrātu summu $\sum d_i^2$.

5. tabula

Darba tabula rangu korelācijas koeficienta aprēķināšanai

Faktorālās pazīmes variante x_i	Rezultatīvās pazīmes variante y_i	x rangs $g_{x.i}$	y rangs $g_{y.i}$	Rangu starpība $d_i = g_{x.i} - g_{y.i}$	d_i^2
x_1	y_1	$g_{x.1}$	$g_{y.1}$	d_1	d_1^2
x_2	y_2	$g_{x.2}$	$g_{y.2}$	d_2	d_2^2
x_3	y_3	$g_{x.3}$	$g_{y.3}$	d_3	d_3^2
...	...	...	...	...	...
x_{n-1}	y_{n-1}	$g_{x.n-1}$	$g_{y.n-1}$	d_{n-1}	d_{n-1}^2
x_n	y_n	$g_{x.n}$	$g_{y.n}$	d_n	d_n^2
					$\sum d_i^2$

Spirmena rangu korelācijas koeficientu aprēķina pēc formulas:

$$r_s = 1 - \frac{6\sum d_i^2}{n(n^2-1)} \tag{1.13}$$

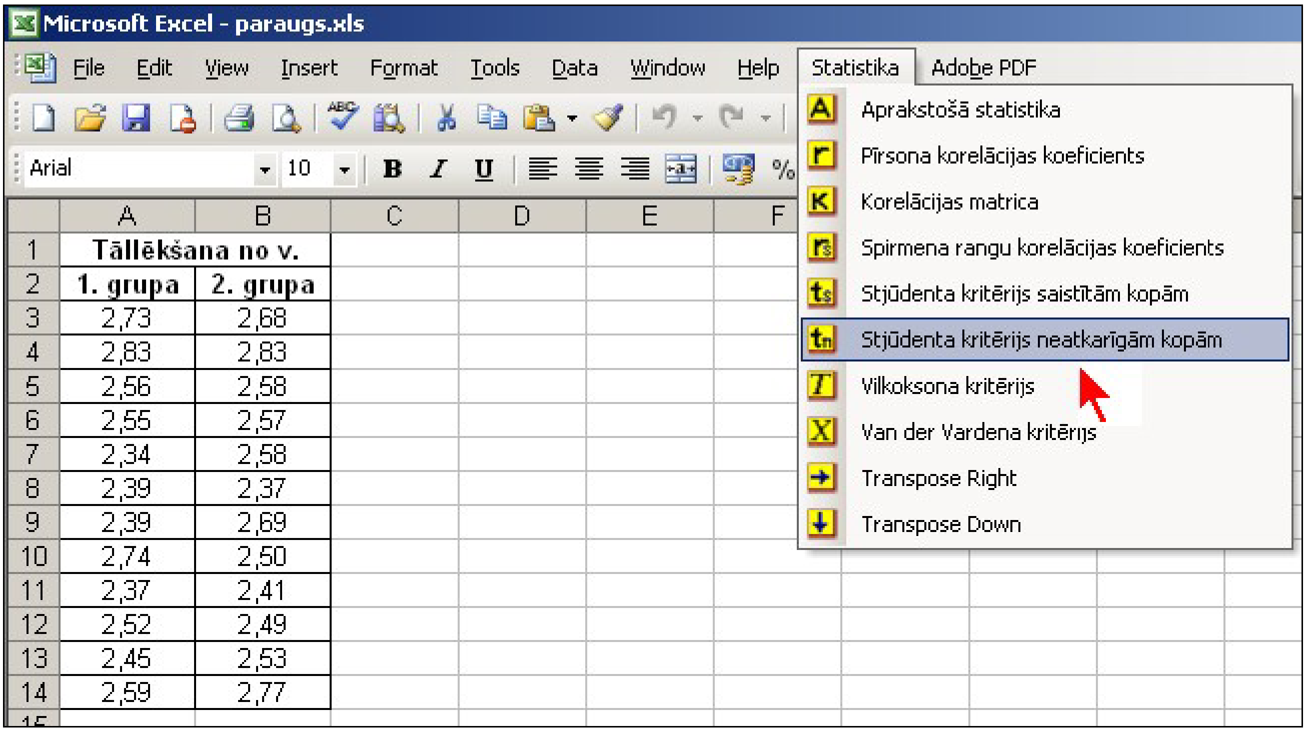
kur n - varianšu pāru skaits; d_i - varianšu rangu pāru starpība. Korelācija ir ticama, ja $|r| > r_\alpha$. Kritisko vērtību atrod tabulā. Spirmena rangu korelācijas koeficients, tāpat kā Pīrsona korelācijas koeficients, raksturo atkarības virzienu un ciešumu.

2. DATORPROGRAMMA "STATISTIKA"

2.1. Darbs ar programmu STATISTIKA

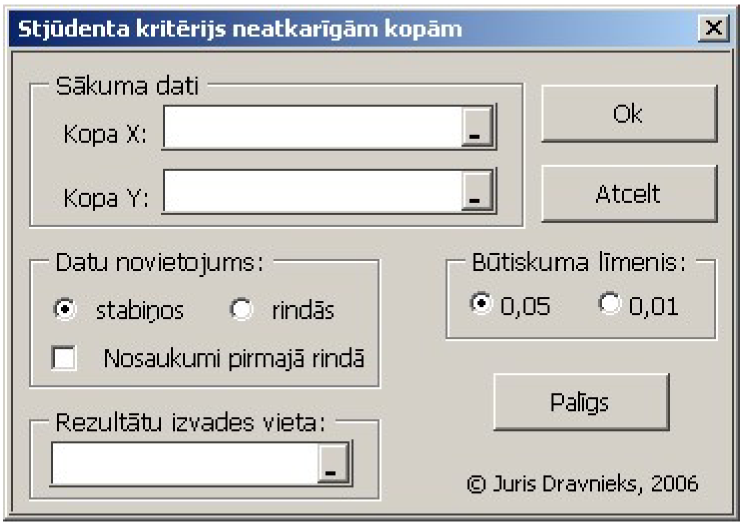
2.1.1. Sākuma datu paziņošana un analīzes rezultātu saņemšana (Stjūdenta kritērijs)

Pēc MS EXCEL palaišanas tabulas atsevišķās kolonās ievada sākuma datus. Lai attēlojot aprēķinu rezultātus būtu saprotams, kas ir kas, salīdzināmo datu stabiņu (vai rindu) nosaukumus vēlams veidot atšķirīgus. Kad dati ierakstīti tabulā, novieto peles rādītāju uz izvēlnes "Statistika" un ar klikšķi atver to (1. att.)



1. att. Operācijas izvēle

Ar peles rādītāju izvēlas apakšprogrammas nosaukumu un noklikšķinot palaiž to. Atveras sākuma datu paziņošanas logs (2. att.).



2. att. Sākuma datu paziņošanas dialogs

Lodziņos "Kopa X" un "Kopa Y" jāpaziņo tabulā ievadīto datu masīvu adreses. Atšķirību kritērijos "Kopa X" vienmēr ir pirmās grupas dati vai arī sākuma dati, bet "Kopa Y" ir otrās grupas rezultāti vai beigu dati. Korelācijas analīzes gadījumā "Kopa X" ir faktorālās pazīmes dati, bet "Kopa Y" – rezultatīvās pazīmes dati. Peles rādītāju novieto uz pogas lodziņa labajā pusē un ar klikšķi atver palīg lodziņu (3. att.).

3. att. Palīg lodziņš datu paziņošanai

Tabula kļūst pārskatāma. Novieto peles rādītāju uz datu stabiņa augšējās šūnas (vēlams ar datu nosaukumu) un, turot nospiestu peles kreiso pogu velk rādītāju pa stabiņu uz leju, tā iezīmējot šīs kopas datus. Lodziņā automātiski parādās iezīmētā masīva robežu adreses. Ar klikšķi uz pogas lodziņa labajā pusē to aizver, līdz ar to notiek atgriešanās datu paziņošanas dialoga logā (2. att.). Tādā pat veidā iezīmē otras kopas datus.

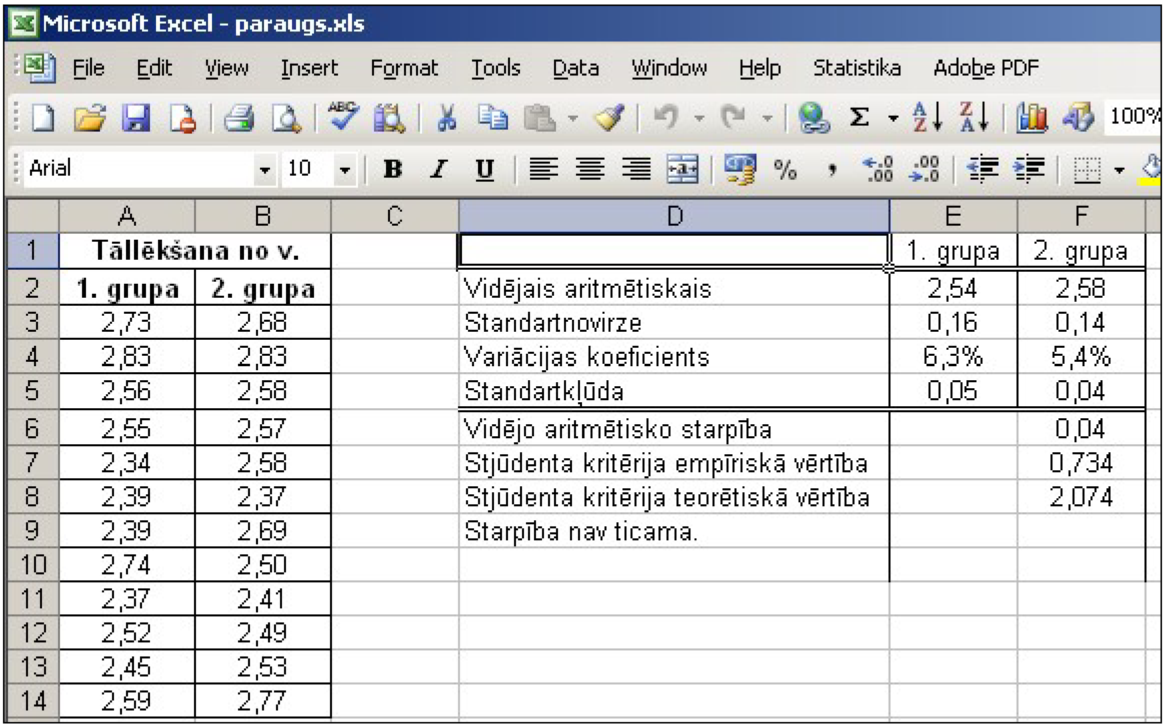
Lodziņos "Dati grupēti" atbilstoši to novietojumam, ieklikšķinām punktu pozīcijā "stabiņos" vai "rindās".

Ja datu masīvu sākuma šūnā bija ierakstīts to nosaukums, ieklikšķinām ķeksīti lodziņā "Nosaukumi pirmajā rindā" ("Nosaukumi pirmajā kolonnā"). Ja nosaukuma nav, lodziņu atstājam tukšu, tad Excel's, izvadot analīzes rezultātus, automātiski pievienos nosaukumus "Kopa X" un "Kopa Y".

Ieklikšķinām lodziņā "Rezultātu izvades zona", lai tur parādās kursors, un pēc tam ieklikšķinām tabulas šūnā, no kuras uz leju un pa labi vēlamies redzēt izskaitļotos analīzes rezultātus. Šūnas adrese parādās lodziņā „Rezultātu izvades vieta. Dati ir paziņoti (4. att.).

4. att. Dati paziņoti

Ar klikšķi uz pogas <OK> liekam datoram rēķināt. Ja viss izdarīts pareizi, saņemam rezultātus, ieskaitot statistiskās ticamības raksturojumu (5. att.).

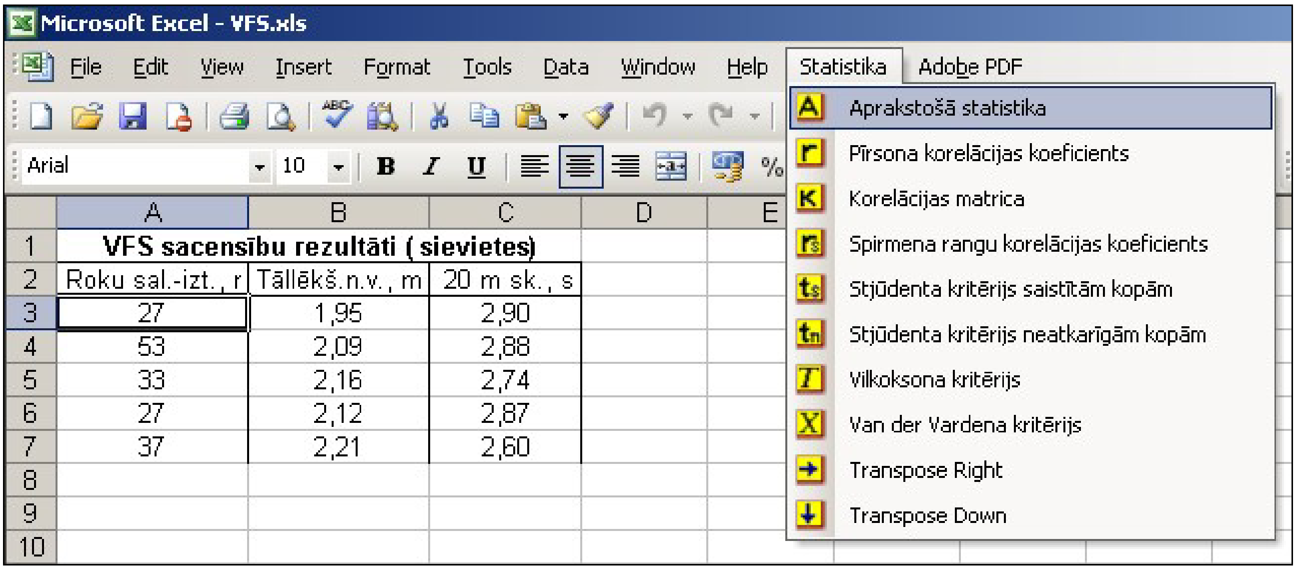


5. att. Analīzes rezultātu attēlojums

Programmās, kurās dati jāpaziņo divu atsevišķu stabiņu veidā, dialoga logi ir līdzīgi (izņemot metodes nosaukumu).

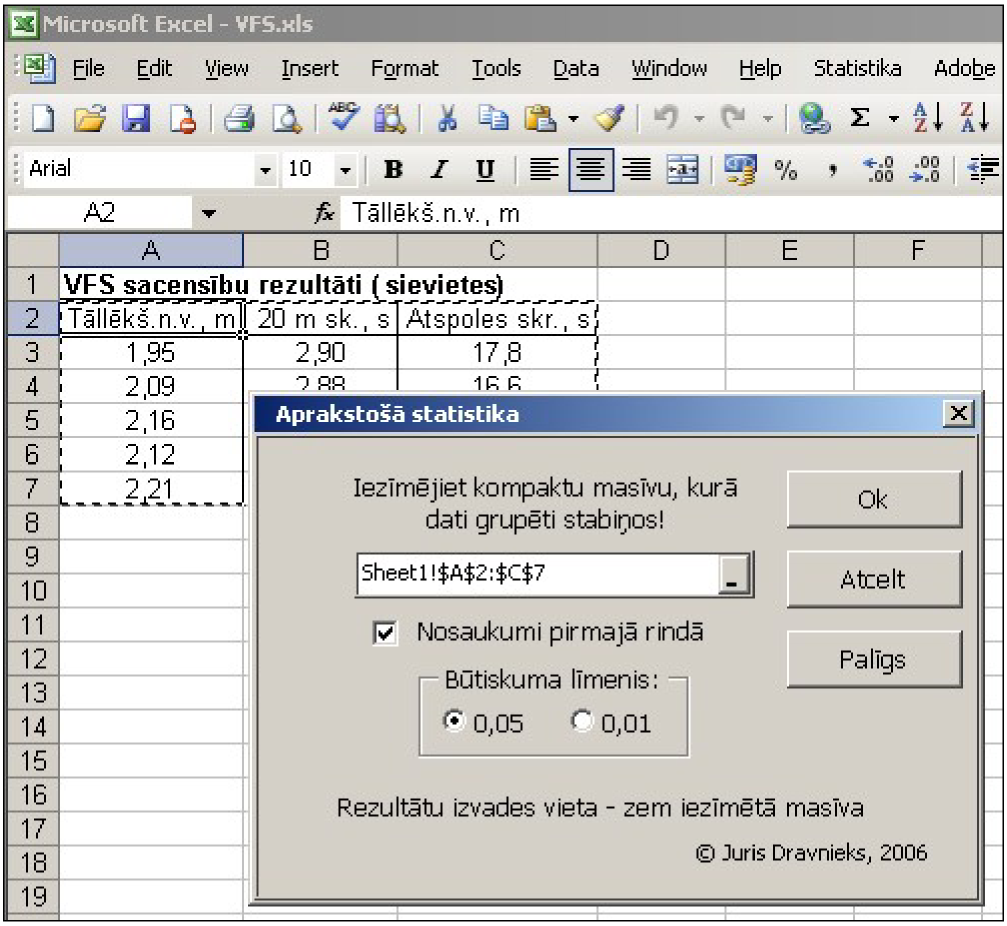
2.1.2. Sākuma datu paziņošana un analīzes rezultātu saņemšana (Aprakstošā statistika)

Ar klikšķi uz izvēlnes nosaukuma atver sarakstu <Statistika> (6. att.).



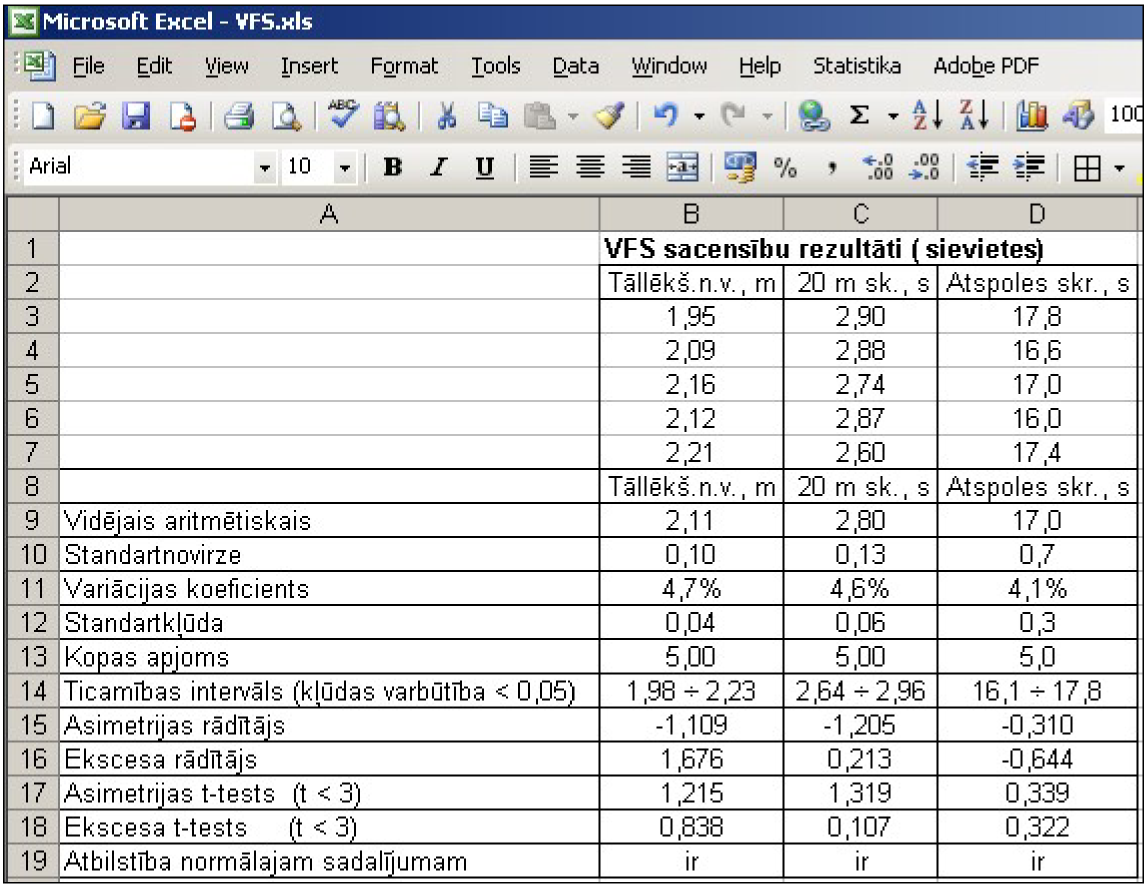
6. att. Izvēlne <Statistika>

Ar peles rādītāju izvēlas apakšprogrammas „Aprakstošā statistika” nosaukumu un ar klikšķi palaiž programmu. Atveras sākuma datu paziņošanas logs (7. att.).



7. att. Sākuma datu paziņošana

Novieto peles rādītāju uz datu stabiņa augšējās šūnas un, turot nospiestu peles kreiso pogu velk rādītāju pa diagonāli uz leju, iezīmējot visu masīvu. Lodziņā automātiski parādās iezīmētā masīva robežu adreses. Ieklikšķina lodziņā karodziņu „Nosaukumi pirmajā rindā”. Pēc klikšķa uz pogas <OK>, zem iezīmētā masīva parādās izskaitļotie rezultāti, bet papildstabiņā kreisajā pusē izskaitļoto raksturojumu nosaukumi.

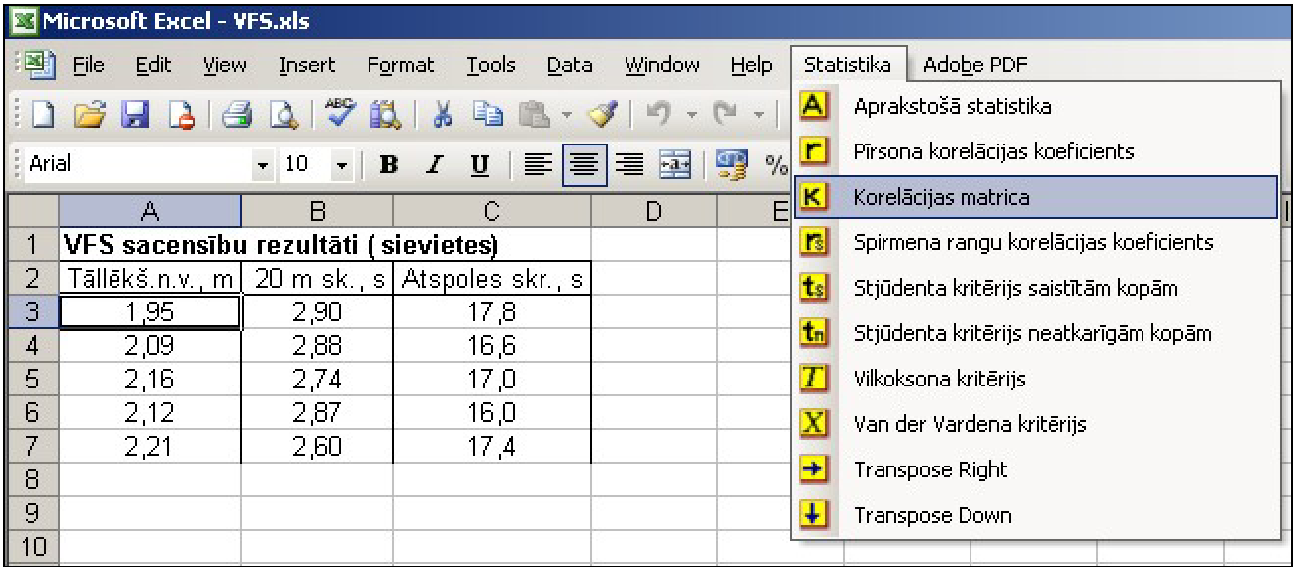


	A	B	C	D
1		VFS sacensību rezultāti (sievietes)		
2		Tāllēkš.n.v., m	20 m sk., s	Atspoles skr., s
3		1,95	2,90	17,8
4		2,09	2,88	16,6
5		2,16	2,74	17,0
6		2,12	2,87	16,0
7		2,21	2,60	17,4
8		Tāllēkš.n.v., m	20 m sk., s	Atspoles skr., s
9	Vidējais aritmētiskais	2,11	2,80	17,0
10	Standartnovirze	0,10	0,13	0,7
11	Variācijas koeficients	4,7%	4,6%	4,1%
12	Standartkļūda	0,04	0,06	0,3
13	Kopas apjoms	5,00	5,00	5,0
14	Ticamības intervāls (kļūdas varbūtība < 0,05)	1,98 ÷ 2,23	2,64 ÷ 2,96	16,1 ÷ 17,8
15	Asimetrijas rādītājs	-1,109	-1,205	-0,310
16	Ekscesa rādītājs	1,676	0,213	-0,644
17	Asimetrijas t-tests (t < 3)	1,215	1,319	0,339
18	Ekscesa t-tests (t < 3)	0,838	0,107	0,322
19	Atbilstība normālajam sadalījumam	ir	ir	ir

8. att. Izskaitļotie aprakstošās statistikas dati

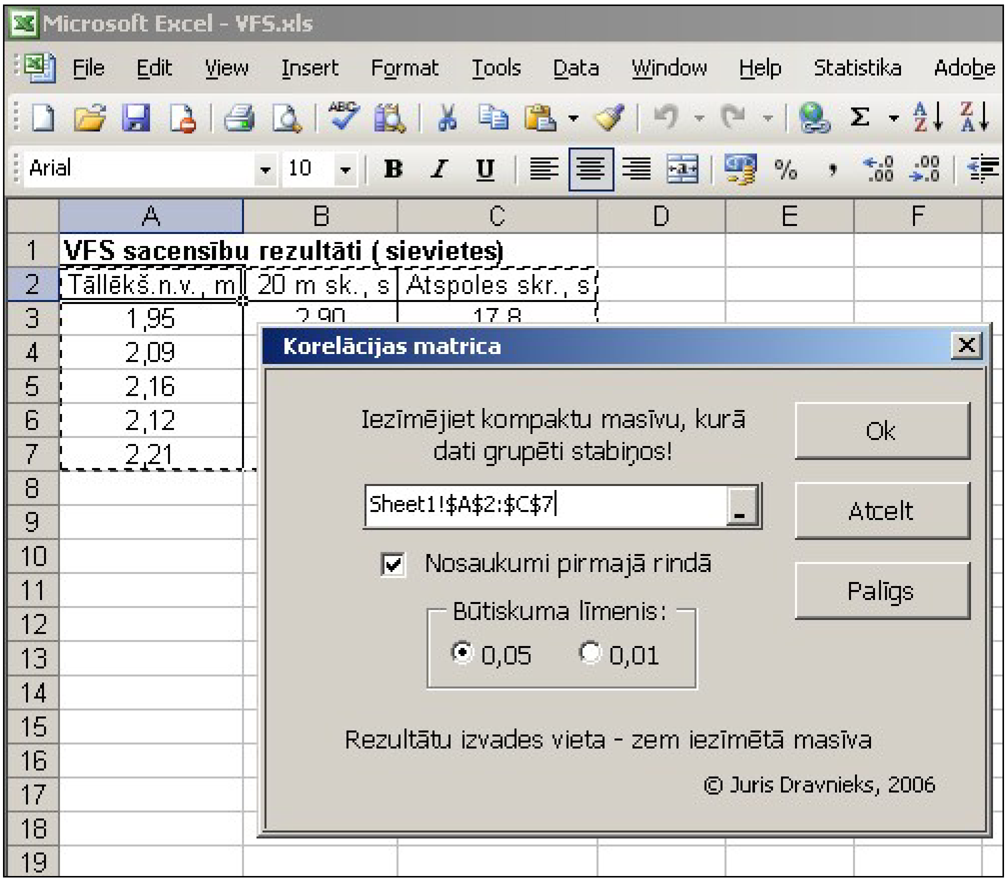
2.1.3. Sākuma datu paziņošana un analīzes rezultātu saņemšana (Korelācijas matrica)

Programmās, kurās sākuma dati jāpaziņo viena masīva veidā dialoga logi ir līdzīgi (izņemot metodes nosaukumu). Ar klikšķi uz izvēlnes nosaukuma atver sarakstu <Statistika> (9. att.).



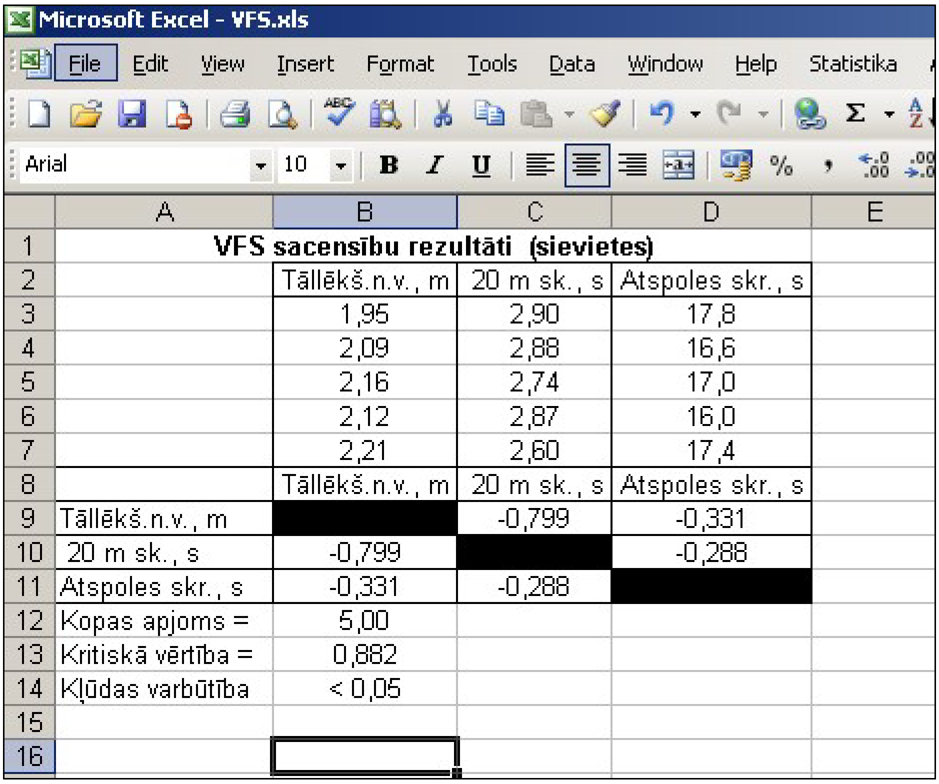
9. att. Programmas "Korelācijas matrica izvēle"

Ar peles rādītāju izvēlas apakšprogrammas „Korelācijas matrica" nosaukumu un ar klikšķi palaiž programmu. Atveras sākuma datu paziņošanas logs (10. att.).



10. att. Sākuma datu masīva paziņošana

Novieto peles rādītāju uz datu stabiņa augšējās šūnas un, turot nospiestu peles kreiso pogu, velk rādītāju pa diagonāli uz leju, iezīmējot visu masīvu. Lodziņā automātiski parādās iezīmētā masīva robežu adreses. Ieklikšķina lodziņā karodziņu „Nosaukumi pirmajā rindā”. Pēc klikšķa uz pogas <Ok>, zem iezīmētā masīva attēlojas matricas stabiņu nosaukumi un izskaitļotie rezultāti, bet papildstabiņā kreisajā pusē matricas rindu nosaukumi. Pīrsona korelācijas koeficienti izskaitļoti katram lielumam ar katru. Nederīgo koeficientu (lielumam pašam ar sevi) šunas aizkrāsotas (11. att.). Programma paziņo arī paraugkopas apjomu, korelācijas koeficienta kritisko vērtību un izvēlēto būtiskuma līmeni (kļūdas varbūtība).



	A	B	C	D	E
1	VFS sacensību rezultāti (sievietes)				
2		Tālēkš.n.v., m	20 m sk., s	Atspoles skr., s	
3		1,95	2,90	17,8	
4		2,09	2,88	16,6	
5		2,16	2,74	17,0	
6		2,12	2,87	16,0	
7		2,21	2,60	17,4	
8		Tālēkš.n.v., m	20 m sk., s	Atspoles skr., s	
9	Tālēkš.n.v., m		-0,799	-0,331	
10	20 m sk., s	-0,799		-0,288	
11	Atspoles skr., s	-0,331	-0,288		
12	Kopas apjoms =	5,00			
13	Kritiskā vērtība =	0,882			
14	Kļūdas varbūtība	< 0,05			
15					
16					

11. att. Izskaitļotā korelācijas matrica

2.1.4. Palīdzības informācija

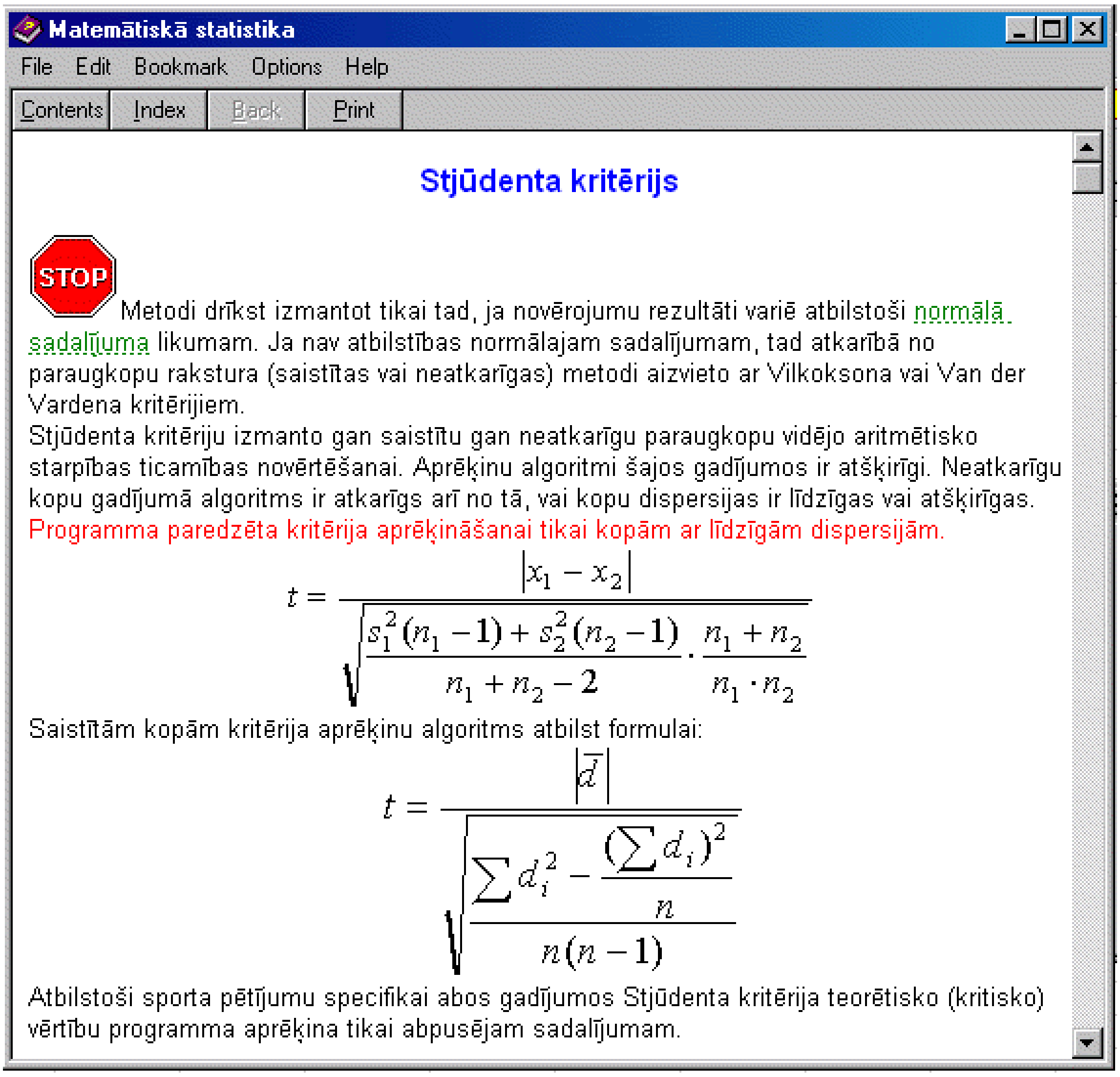
Statistikas metožu teorētiskie pamati aprakstīti šīs brošūras 1. nodaļā. Galvenie metožu izmantošanas noteikumi aprakstīti arī programmas palīdzības failā. Ja neesam pārliecināti par analīzes metodes izvēles pareizību, varam izmantot Windows Help sistēmu. “Palīga” failu var apskatīt, ja novieto peles rādītāju uz datu paziņošanas loga pogas “Palīgs” izpilda klikšķi. Sākuma lapā, klikšķinot uz metodes nosaukuma, saņemam informāciju par šo metodi (12. att.)

2.1.5. Paziņojumi par kļūdām

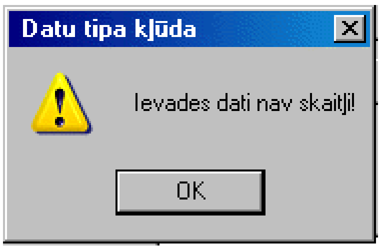
Ja, paziņojot datus, esam bijuši neprecīzi, dators parāda paziņojumu par kļūdu. Skaidrojošo tekstu datu masīva sākumā dators var pieņemt par tekstu, ja neesam ielikuši ķeksīti lodziņā “Nosaukumi pirmajā rindā” (“Nosaukumi pirmajā kolonnā”) un, sākot rēķinus rodas kļūda. Dators par to paziņo un pārtrauc programmas izpildi (13. att.).

Ja, paziņojot datus, esam iezīmējuši tukšas tabulas šūnas, dators par to paziņo un pārtrauc programmas izpildi (14. att.).

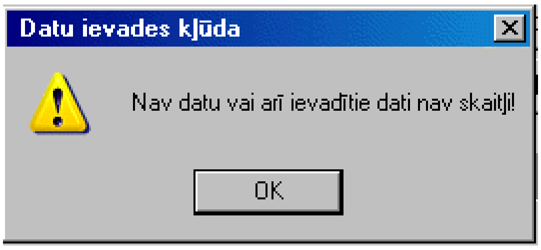
Analizējot saistītas kopas, novērojumu skaits abās kopās ir vienāds (Stjūdenta kritērijs saistītām kopām, Vilkoksona kritērijs, korelācijas analīze). Ja, paziņojot datus, šis noteikums nav ievērots, dators par to paziņo un pārtrauc programmas izpildi (15.att.).



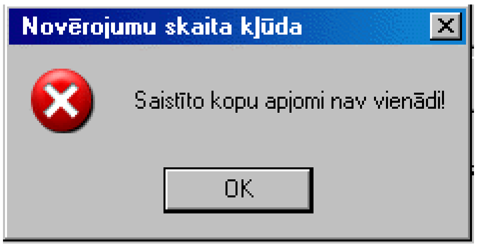
12. att. “Palīgs”



13. att. Kļūda - skaitļa vietā paziņots teksts vai arīo nav karodziņa lodziņā „Nosaukumi pirmajā rindā”

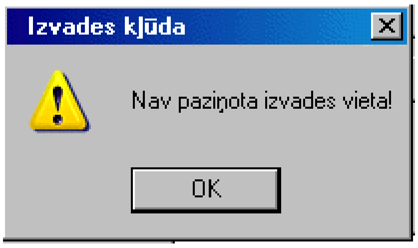


14. att. Kļūda – paziņojot datus, iezīmēta tukša šūna



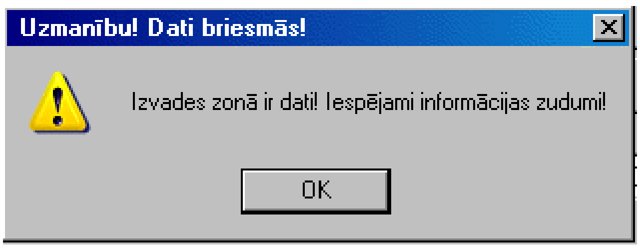
15. att. Kļūda – dažāds novērojumu skaits saistītās kopās

Ja nav paziņota šūna, no kuras sākot tabulā attēlot aprēķinātos rezultātus, dators pārtrauc programmas izpildi un paziņo par kļūdu (16. att.).



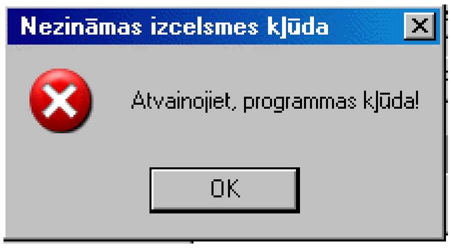
16. att. Kļūda – nav paziņots, kur izvadīt datus

Norādot rezultātu izvades zonu, pa kreisi un uz leju no norādītās šūnas tabulai jābūt tukšai. Ja kādā šūnā būs atrodama kaut tukšumzīme, kuras klātbūtni vizuāli nevar konstatēt, programmas izpilde tiek pārtraukta ar kļūdas paziņojumu (17. att.).



17. att. Kļūda – rezultātu izvades zonā ir dati

Ja nav iespējams identificēt kļūdu t.i. pieļautas vienlaicīgi vairākas no iepriekš minētajām neprecizitātēm, dators par to paziņo un pārtrauc programmas izpildi (18. att.).



18. att. Kļūda, kuru dators nespēj identificēt

Visos kļūdu paziņojumos, izņemot pēdējo (18. att.), pēc klikšķa uz pogas <Ok> notiek atgriešanās datu ievades dialogā un pēc labojumiem var pabeigt aprēķinus. Nezināmas izcelsmes kļūdas gadījumā (18. att.), klikšķis uz pogas <Ok> pārtrauc programmas izpildi - tā jāpalaiž atkārtoti un dati jāpaziņo no jauna.

PS. ZINĀMIE TRŪKUMI UN NEPILNĪBAS.

Mērfija likums saka, ka katrā programmā ir vismaz viena kļūda. Darba gaitā programma STATISTIKA tiek pilnveidota.

Šobrīd zināmi šādi trūkumi:

- Analizē izmantojamos datus nevar ņemt no dažādām lapām (***Sheet***) – tiem jāatrodas vienā un tai pašā lapā.
- Atsevišķos gadījumos sākuma datus var iezīmēt tikai izmantojot peli.
- Programma nestrādā operētājsistēmas **Windows Vista** vidē.

Uzmanību! Autors negarantē **STATISTIKAS** funkcionalitāti visos gadījumos – to var ietekmēt Jūsu datora konfigurācija (izvairieties no **Microsoft** piedāvātajām **WindowsHome** versijām) !

Veiksmīgas uzstādīšanas priekšnosacījumi:

- **datora administratora tiesības;**
- **operētājsistēma: Windows95/98/2000/XP/Vista/7 (PRO vai Enterprise);**
- **Microsoft Office 2000/XP/2003/2007/2010 (PRO) ar pilnībā instalētu Excel;**
- **atļauta visu MS Excel makrosu un VBA projektu darbība.**

Literatūra

1. Arhipova I., Bāliņa S. Statistika ar Microsoft Excel ikvienam. 1. daļa. Rīga : Datorzinību Centrs, 1999. 168. lpp.
2. Arhipova I., Bāliņa S. Statistika ar Microsoft Excel ikvienam. 2. daļa. Rīga : Datorzinību Centrs, 2000. 136. lpp.
3. Dravnieks J., Popovs E., Paeglītis A. Sporta zinātnisko pētījumu tehnoloģija. Rīga, 1997. 1. daļa. - 98. lpp.
   2. daļa. - 86. lpp.
   3. daļa. - 86. lpp.
   4., daļa. - 52 lpp.
4. Dravnieks J. Bakalaura pavārgrāmata [Tiešsaistes pakalpojums] : Labots un papildināts / LSPA. 5. labotā un papildinātā versija. Rīga : LSPA, 2011. Versija 5.0. 54 lpp. PDF formats. Pieejas veids: tīmeklis WWW.URL: http://runcis.lspa.lv/pavars.pdf
5. Dravnieks J. Matemātiskās statistikas metodes sporta zinātnē [tiešaiste] : mācību līdzeklis / LSPA. Rīga : LSPA, 2004. 75 lpp. PDF formats. Pieejas veids: tīmeklis WWW.URL: http://runcis.lspa.lv/statist2.pdf
6. Ķiņķere A. Narņicka S. Microsoft Excel 2000 no A līdz Z. 1.grāmata. Rīga : Datorzinību centrs, 2000. 136 lpp.
7. Ķiņķere A. Microsoft Excel 2000 no A līdz Z. 2.grāmata. Rīga : Datorzinību centrs, 2000. 136 lpp.

3. PIELIKUMI

3.1. Stjūdenta kritērija teorētiskās vērtības

$t_{\alpha;\nu}$

ν	α			ν	α		
	0,1	0,05	0,01		0,1	0,05	0,01
1	6,314	12,706	63,657	17	1,740	2,110	2,898
2	2,920	4,303	9,925	18	1,734	2,101	2,878
3	2,353	3,182	5,841	19	1,729	2,093	2,861
4	2,132	2,776	4,604	20	1,725	2,086	2,845
5	2,015	2,571	4,032	21	1,721	2,080	2,831
6	1,943	2,447	3,707	22	1,717	2,074	2,819
7	1,895	2,365	3,499	23	1,714	2,069	2,807
8	1,860	2,306	3,355	24	1,711	2,064	2,797
9	1,833	2,262	3,250	25	1,708	2,060	2,787
9	1,812	2,228	3,169	26	1,706	2,056	2,779
11	1,796	2,201	3,106	27	1,703	2,052	2,771
12	1,782	2,179	3,055	28	1,701	2,048	2,763
13	1,771	2,160	3,012	29	1,699	2,045	2,756
14	1,761	2,145	2,977	30	1,697	2,042	2,750
15	1,753	2,131	2,947	>30	1,645	1,960	2,576
16	1,746	2,120	2,921				

Nulles hipotēzi noraida, t.i., vērtējamā starpība ir ticama, ja $t \geq t_{\alpha;\nu}$

3.2. Vilkoksona kritērija teorētiskās vērtības

$T_{\alpha;n}$ $(\alpha = 0,05)$

n	$T_{0,05;n}$	n	$T_{0,05;n}$	n	$T_{0,05;n}$
6	1	13	17	20	52
7	2	14	21	21	59
8	4	15	25	22	66
9	6	16	30	23	73
10	8	17	35	24	81
11	11	18	40	25	89
12	14	19	46		

Nulles hipotēzi noraida, ja $T \leq T_{0,05;n}$

3.3.

Van der Vardena kritērija teorētiskās vērtības

X_α ($\alpha = 0,05$)

n	$n_1 - n_2$			n	$n_1 - n_2$		
	0 ÷ 1	2 ÷ 3	4 ÷ 5		0 ÷ 1	2 ÷ 3	4 ÷ 5
8	2,40	2,30	-	30	4,88	4,87	4,84
9	2,38	2,20	-	31	4,97	4,95	4,91
10	2,60	2,49	2,30	32	5,07	5,06	5,03
11	2,72	2,58	2,40	33	5,15	5,13	5,10
12	2,86	2,79	2,68	34	5,25	5,24	5,21
13	2,96	2,91	2,78	35	5,33	5,31	5,28
14	3,11	3,06	3,00	36	5,42	5,41	5,38
15	3,24	3,19	3,06	37	5,50	5,48	5,45
16	3,39	3,36	3,28	38	5,59	5,58	5,55
17	3,49	3,44	3,36	39	5,67	5,65	5,62
18	3,63	3,60	3,53	40	5,75	5,74	5,72
19	3,73	3,69	3,61	41	5,83	5,81	5,79
20	3,86	3,84	3,78	42	5,91	5,90	5,88
21	3,96	3,92	3,85	43	5,99	5,97	5,95
22	4,08	4,06	4,01	44	6,04	6,06	6,04
23	4,18	4,15	4,08	45	6,14	6,12	6,10
24	4,29	4,27	4,23	46	6,21	6,21	6,19
25	4,39	4,36	4,30	47	6,29	6,27	6,25
26	4,50	4,48	4,44	48	6,36	6,35	6,34
27	4,59	4,56	4,51	49	6,43	6,42	6,39
28	4,68	4,68	4,64	50	6,50	6,51	6,48
29	4,78	4,76	4,72				

Nulles hipotēzi noraida, ja $X > X_\alpha$

3.4. Pīrsona korelācijas koeficienta kritiskās vērtības

$r_{0,05;n}$ ($\alpha = 0,05$)

n	$r_{0,05;n}$	n	$r_{0,05;n}$	n	$r_{0,05;n}$	n	$r_{0,05;n}$
4	0,950	15	0,514	26	0,388	80	0,219
5	0,875	16	0,497	27	0,381	90	0,206
6	0,811	17	0,487	28	0,374	100	0,196
7	0,754	18	0,468	29	0,367	125	0,175
8	0,707	19	0,456	30	0,361	150	0,160
9	0,666	20	0,444	35	0,332	200	0,135
10	0,632	21	0,433	40	0,310	250	0,124
11	0,602	22	0,423	45	0,292	300	0,113
12	0,576	23	0,413	50	0,277	400	0,098
13	0,533	24	0,404	60	0,253	500	0,088
14	0,532	25	0,396	70	0,234	1000	0,063

Korelācija ir ticama, ja $|r| \geq r_{0,05;n}$

3.5. Spirmena rangu korelācijas koeficienta kritiskās vērtības

$r_{S_{0,05}}$ ($\alpha = 0,05$)

n	r_{S_α}	n	r_{S_α}	n	r_{S_α}
5	0,94	12	0,58	22	0,43
6	0,85	13	0,56	24	0,41
7	0,78	14	0,54	26	0,39
8	0,72	15	0,52	28	0,38
9	0,68	16	0,50	30	0,36
10	0,64	18	0,47	35	0,33
11	0,61	20	0,45	40	0,31

Korelācija ir ticama, ja $|r_S| > r_{S_{0,05}}$

3.6. Funkcijas Ψ vērtības

$\frac{g_i}{n_1+n_2+1}$	0	1	2	3	4	5	6	7	8	9
0,50	0,00	0,00	0,01	0,01	0,01	0,01	0,02	0,02	0,02	0,02
0,51	0,03	0,03	0,03	0,03	0,04	0,04	0,04	0,04	0,05	0,05
0,52	0,05	0,05	0,06	0,06	0,06	0,06	0,07	0,07	0,07	0,07
0,53	0,08	0,08	0,08	0,08	0,09	0,09	0,09	0,09	0,10	0,10
0,54	0,10	0,10	0,11	0,11	0,11	0,11	0,12	0,12	0,12	0,12
0,55	0,13	0,13	0,13	0,13	0,14	0,14	0,14	0,14	0,15	0,15
0,56	0,15	0,15	0,16	0,16	0,16	0,16	0,17	0,17	0,17	0,17
0,57	0,18	0,18	0,18	0,18	0,19	0,19	0,19	0,19	0,20	0,20
0,58	0,20	0,20	0,21	0,21	0,21	0,21	0,22	0,22	0,22	0,23
0,59	0,23	0,23	0,23	0,24	0,24	0,24	0,24	0,25	0,25	0,25
0,60	0,25	0,26	0,26	0,26	0,26	0,27	0,27	0,27	0,27	0,28
0,61	0,28	0,28	0,28	0,29	0,29	0,29	0,30	0,30	0,30	0,30
0,62	0,31	0,31	0,31	0,31	0,32	0,32	0,32	0,32	0,33	0,33
0,63	0,33	0,33	0,34	0,34	0,34	0,35	0,35	0,35	0,35	0,36
0,64	0,36	0,36	0,36	0,37	0,37	0,37	0,37	0,38	0,38	0,38
0,65	0,39	0,39	0,39	0,39	0,40	0,40	0,40	0,40	0,41	0,41
0,66	0,41	0,42	0,42	0,42	0,42	0,43	0,43	0,43	0,43	0,44
0,67	0,44	0,44	0,45	0,45	0,45	0,45	0,46	0,46	0,46	0,46
0,68	0,47	0,47	0,47	0,48	0,48	0,48	0,48	0,49	0,49	0,49
0,69	0,50	0,50	0,50	0,50	0,51	0,51	0,51	0,52	0,52	0,52
0,70	0,52	0,53	0,53	0,53	0,54	0,54	0,54	0,54	0,55	0,55
0,71	0,55	0,56	0,56	0,56	0,57	0,57	0,57	0,57	0,58	0,58
0,72	0,58	0,59	0,59	0,59	0,59	0,60	0,60	0,60	0,61	0,61
0,73	0,61	0,62	0,62	0,62	0,63	0,63	0,63	0,63	0,64	0,64
0,74	0,64	0,65	0,65	0,65	0,66	0,66	0,66	0,67	0,67	0,67

6. pielikuma turpinājums

$\frac{g_i}{n_1+n_2+1}$	0	1	2	3	4	5	6	7	8	9
0,75	0,67	0,68	0,68	0,68	0,69	0,69	0,69	0,70	0,70	0,70
0,76	0,71	0,71	0,71	0,72	0,72	0,72	0,73	0,73	0,73	0,74
0,77	0,74	0,74	0,75	0,75	0,75	0,76	0,76	0,76	0,77	0,77
0,78	0,77	0,78	0,78	0,78	0,79	0,79	0,79	0,80	0,80	0,80
0,79	0,81	0,81	0,81	0,82	0,82	0,82	0,83	0,83	0,83	0,84
0,80	0,84	0,85	0,85	0,85	0,86	0,86	0,86	0,87	0,87	0,87
0,81	0,88	0,88	0,89	0,89	0,89	0,90	0,90	0,90	0,91	0,91
0,82	0,92	0,92	0,92	0,93	0,93	0,93	0,94	0,94	0,95	0,95
0,83	0,95	0,96	0,96	0,97	0,97	0,97	0,98	0,98	0,99	0,99
0,84	0,99	1,00	1,00	1,01	1,01	1,02	1,02	1,02	1,03	1,03
0,85	1,04	1,04	1,05	1,05	1,05	1,06	1,06	1,07	1,07	1,08
0,86	1,08	1,09	1,09	1,09	1,10	1,10	1,11	1,11	1,12	1,12
0,87	1,13	1,13	1,14	1,14	1,15	1,15	1,16	1,16	1,17	1,17
0,88	1,18	1,18	1,19	1,19	1,20	1,20	1,21	1,21	1,22	1,22
0,89	1,23	1,23	1,24	1,24	1,25	1,25	1,26	1,26	1,27	1,28
0,90	1,28	1,29	1,29	1,30	1,30	1,31	1,32	1,32	1,33	1,33
0,91	1,34	1,35	1,35	1,36	1,37	1,37	1,38	1,39	1,39	1,40
0,92	1,41	1,41	1,42	1,43	1,43	1,44	1,45	1,45	1,46	1,47
0,93	1,48	1,48	1,49	1,50	1,51	1,51	1,52	1,53	1,54	1,55
0,94	1,55	1,56	1,57	1,58	1,59	1,60	1,61	1,62	1,63	1,64
0,95	1,64	1,65	1,66	1,67	1,68	1,70	1,71	1,72	1,73	1,74
0,96	1,75	1,76	1,77	1,79	1,80	1,81	1,83	1,84	1,85	1,87
0,97	1,88	1,90	1,91	1,93	1,94	1,96	1,98	2,00	2,01	2,03
0,98	2,05	2,07	2,10	2,12	2,14	2,17	2,20	2,23	2,26	2,29
0,99	2,33	2,37	2,41	2,46	2,51	2,58	2,65	2,75	2,88	3,09