STATISTIKA ZA PSIHOLOGE 2019/20

Vaja 10 - 13 - Delne rešitve

Vaja 10

3. Izvedli smo študentsko anketo, kjer smo študente spraševali tudi o številu ur na teden, ki jih namenijo športu. Podatki so dani v tabeli.

Št. ur športa na teden	0	1	2	3	4	5	6	7	9	10	12	15	20
Ženske	2	1	10	11	7	5	2	3	1	4	1	1	1
Moški	0	1	1	4	3	2	1	3	1	5	2	1	0

Pri stopnji značilnosti $\alpha = 0,05$ preverite domnevo ali študentke namenijo več ur športu kot študentje.

Rešitev: Označimo z X ženske in z Y moške. Oznaǐcmo še število žensk $m_Z = 49$ in število moških $m_M = 24$, torej $n = 73$. Najprej računamo

$$\mu_X = \frac{2\cdot 0 + 1\cdot 1 + 10\cdot 2 + 11\cdot 3 + 7\cdot 4 + 5\cdot 5 + 2\cdot 6 + 3\cdot 7 + 1\cdot 9 + 4\cdot 10 + 1\cdot 12 + 1\cdot 15 + 1\cdot 20}{49},$$

kar je enako $\mu_X = 3,77$ in

$$\mu_Y = \frac{0\cdot 0 + 1\cdot 1 + 1\cdot 2 + 4\cdot 3 + 3\cdot 4 + 2\cdot 5 + 1\cdot 6 + 3\cdot 7 + 1\cdot 9 + 5\cdot 10 + 2\cdot 12 + 1\cdot 15 + 0\cdot 20}{24},$$

kar je $\mu_Y = 1,84$. Nato

$$\mu = \frac{m_Z}{m_Z + m_M}\mu_X + \frac{m_M}{m_Z + m_M}\mu_Y.$$

in računamo

$$\mu = \frac{49}{24+49}\cdot 3,77 + \frac{24}{24+49}\cdot 1,84 = 3,14.$$

Izračunamo standardno deviacijo **vseh podatkov**

$$\begin{aligned}\sigma &= \sqrt{\frac{2\cdot 0^2 + 1\cdot 1^2 + 10\cdot 2^2 + \ldots + 1\cdot 20^2 + 0\cdot 0^2 + \ldots + 0\cdot 20^2}{73} - (3,14)^2}\\ &= \sqrt{\frac{3262}{73} - (3,14)^2}\\ &= 5,90.\end{aligned}$$

Sedaj lahko izračunamo točkovni biserialni korelacijski koeficient

$$r_{PB} = \frac{\mu_X - \mu_Y}{\sigma}\cdot\frac{\sqrt{m_Z\cdot m_M}}{m_Z + m_M},$$

kar je v našem primeru

$$r_{PB} = \frac{3,77 - 1,84}{5,90}\cdot\frac{\sqrt{49\cdot 24}}{49+24} = 0,15.$$

Zapišemo ničelno in alternativno hipotezo:
H_0*:* $\bar{X} = \bar{Y}$.
ali H_0*: Spremenljivki Št. ur športa na teden in Spol v populaciji nista korelirani.*
H_1*:* $\bar{X} > \bar{Y}$.
Glede na našo alternativno domnevo izvedemo enostranski T-test v desno z $n-2$ *stopinjami prostosti. To pomeni, da bomo ničelno domnevo zavrnili natanko tedaj, ko bo veljalo*

$$T > t_{1-\alpha}(n-2).$$

Izračunamo vrednost testne statistike, ki je enaka

$$T = \frac{r_{PB}}{\sqrt{1-r_{PB}^2}} \cdot \sqrt{n-2},$$

torej

$$T = \frac{0,15}{\sqrt{1-0,15^2}} \cdot \sqrt{73-2} = 1,28.$$

Primerjamo vrednost testne statistike z vrednostjo iz tabele

$$t_{1-0,05}(73-2) = t_{0,95}(71) = 1,67.$$

Ker NE velja $1,28 > 1,67$*, ničelne domneve ne moremo zavrniti. Torej naši podatki ne potrjujejo, da bi študentke več časa posvečale športu kot študentje.*

4. [1] Želeli ste oceniti razliko v povprečni širini dlak med psi in mačkami. Izmerili ste širino dlake pri 10 psih in 10 mačkah ter dobili naslednje podatke.

Širina (mm)	Povprečje	Standardni odklon	Število (n)
Psi (1)	103.2	25.0	10
Mačke (2)	68.2	13.9	10

Ali je povprečna širina dlake pri psih in mačkah razlikuje?

Rešitev: Ker imamo podani standardni deviaciji obeh skupin, torej psov in mačk, bomo računali Cohenov koeficient. Z X *označimo pse in z* Y *mačke. Formula za Cohenov koeficient*

$$d = \frac{\mu_X - \mu_Y}{\sigma_W}.$$

Rabimo σ_W

$$\sigma_W = \sqrt{\frac{m_P}{m_P+m_M}\sigma_X^2 + \frac{m_M}{m_P+m_M}\sigma_Y^2},$$

kjer m_P *označuje število psov in* m_M *število mačk vključenih v raziskavo. V našem primeru je* $m_P = m_M = 10$. *Standardni deviaciji imamo že podani* $\sigma_X^2 = 25$ *in* $\sigma_Y = 13,9$. *Torej*

$$\sigma_W = \sqrt{\frac{10}{10+10}25^2 + \frac{10}{10+10}13,9^2} = 20,22.$$

[1] L. Lusa: Naloge iz biostatistike

Sedaj lahko izračunamo Cohenov koeficient

$$d = \frac{103,2 - 68,2}{20,22} = 1,73.$$

Zapišemo ničelno in alternativno hipotezo:
H_0: $\bar{X} = \bar{Y}$.
H_1: $\bar{X} \neq \bar{Y}$.
Glede na našo alternativno domnevo izvedemo dvostranski T-test z $n-2$ stopinjami prostosti. To pomeni, da bomo ničelno domnevo zavrnili natanko tedaj, ko bo veljalo

$$|T| > t_{1-\frac{\alpha}{2}}(n-2).$$

Izračunamo vrednost testne statistike, ki je enaka

$$T = d \cdot \frac{\sqrt{(n-2)\cdot m_P \cdot m_M}}{n} = 1,73 \cdot \frac{\sqrt{(20-2)\cdot 10 \cdot 10}}{20} = 3,67.$$

Sedaj vrednost testne statistike primerjamo z vrednostjo iz tabele

$$t_{1-0,025}(18) = t_{0,975}(18) = 2,10.$$

Ker velja $|3,67| > 2,10$, lahko zavrnemo ničelno domnevo. Torej naši podatki kažejo, da se povprečna širina dlak pri psih in mačkah statistično značilno razlikuje.

Preverimo še, če se povprečna širina dlak pri psih in mačkah statistično zelo značilno razlikuje (domnevamo, da je odgovor da).
Naša α je sedaj $\alpha = 0,01$. Ničelno domnevo bomo zavrnili natanko tedaj, ko bo veljalo

$$|T| > t_{1-\frac{\alpha}{2}}(n-2) = t_{0,995}(18) = 2,88.$$

Ker $|3,67| > 2,88$, lahko tudi v tem primeru zavrnemo ničelno domnevo.

Vaja 11

3. V populaciji 50 letnikov smo želeli preveriti ali je raven holesterola v krvi povezana s statusom (S-samski, P-poročen, L-ločen). Z enostavnim slučajnim vzorčenjem smo dobili naslednje podatke

(S, 5,0), (P, 4,7), (P, 5,1), (L, 4,7), (P, 5,2), (P, 4,8), (L, 5,8), (S, 5,2)
(P, 4,5), (L, 5,0), (S, 5,3), (P, 4,9), (S, 5,2), (P, 5,3), (S, 4,8), (L, 4,7)

Naredite statistično analizo in pri stopnji značilnosti $\alpha = 0,05$ preverite ali sta Raven holesterola in Status v populaciji 50-letnikov povezana.

Rešitev: Označimo S=1.skupina, P=2. skupina, L=3.skupina. Velja $n_1 = 5$, $n_2 = 7$, $n_3 = 4$ in $n = 16$. Računamo

$$\mu_1 = \frac{5,0 + 5,2 + 5,3 + 5,2 + 4,8}{5} = 5,10$$

$$\mu_2 = \frac{4,7+5,1+5,2+4,8+4,5+4,9+5,3}{7} = 4,93$$

$$\mu_3 = \frac{4,7+5,8+5,0+4,7}{4} = 5,05$$

Naprej računamo nepojasnjeno varianco, zato potrebujemo

$$\sigma_1^2 = \frac{5,0^2+5,2^2+5,3^2+5,2^2+4,8^2}{5} - 5,10^2 = 0,03,$$

$$\sigma_2^2 = \frac{4,7^2+5,1^2+5,2^2+4,8^2+4,5^2+4,9^2+5,3^2}{7} - 4,93^2 = 0,057,$$

$$\sigma_3^2 = \frac{4,7^2+5,8^2+5,0^2+4,7^2}{4} - 5,05^2 = 0,20.$$

In nepojasnjena varianca je

$$\sigma_W^2 = \frac{n_1}{n}\sigma_1^2 + \frac{n_2}{n}\sigma_2^2 + \frac{n_3}{n}\sigma_3^2,$$

torej

$$\sigma_W^2 = \frac{5}{16}\cdot 0,03 + \frac{7}{16}\cdot 0,057 + \frac{4}{16}\cdot 0,20 = 0,084.$$

Za izračun pojasnjene variance potrebujemo še skupno aritmetično sredino

$$\mu = \frac{n_1}{n}\mu_1 + \frac{n_2}{n}\mu_2 + \frac{n_3}{n}\mu_3,$$

torej

$$\mu = \frac{5}{16}5,10 + \frac{7}{16}4,93 + \frac{4}{16}5,05 = 5,01.$$

Izračunajmo pojasnjeno varianco

$$\sigma_B^2 = \frac{n_1}{n}(\mu_1-\mu)^2 + \frac{n_2}{n}(\mu_2-\mu)^2 + \ldots + \frac{n_k}{n}(\mu_k-\mu)^2$$

kar je v našem primeru

$$\sigma_B^2 = \frac{5}{16}(5,10-5,01)^2 + \frac{7}{16}(4,93-5,01)^2 + \frac{4}{16}(5,05-5,01)^2 = 0,0057.$$

Zapišemo
H_0: *Med intervalsko (Raven holesterola) in imensko spremenljivko (Status) ni povezave. oz. Povprečja so v populaciji na vseh treh skupinah enaka.*
H_1: *Med intervalsko (Raven holesterola) in imensko spremenljivko (Status) je povezava. oz. Vsaj dve povprečji skupin sta v populaciji različni.*

Izvedemo F-test, kjer je testna statistika enaka

$$F = \frac{n-k}{k-1}\cdot\frac{\sigma_B^2}{\sigma_W^2}.$$

Porazdelitev testne statistike ob veljavni ničelni domnevi je Fisherjeva (Snedecorjeva) porazdelitev s $(k-1, n-k)$ *prostostnimi stopinjami. Ničelno domnevo zavreno, če velja*

$$F > F_{1-\alpha}(k-1, n-k).$$

Izračunajmo torej vrednost testne statistike

$$F = \frac{n-k}{k-1} \cdot \frac{\sigma_B^2}{\sigma_W^2} = \frac{16-3}{3-1} \cdot \frac{0,0057}{0,084} = 0,44.$$

Sedaj vrednost testne statistike primerjamo z vrednostjo iz tabele

$$F_{1-\alpha}(k-1, n-k) = F_{0,95}(2, 13) = 3,89.$$

Ker NE velja $0,44 > 3,89$*, ne moremo zavrniti ničelne domneve. Torej naši podatki ne kažejo statistično značilne povezanosti med spremenljivkama Raven holesterola in Statusom v populaciji 50-letnikov.*

4. Z raziskavo bi radi preverili ali sta poslušanje določene slovenske radijske postaje in starost povezana. Označimo štiri najpogosteje poslušane radijske postaje v Sloveniji z A, B, C in D. Naredimo enostavni slučajni vzorec in 21 polnoletnih ljudi povprašamo o njihovi starosti in najljubši slovenski radijski postaji. Dobimo naslednje rezultate

Postaja	Starosti
A	20, 32, 45, 33, 29, 44
B	26, 29, 41, 49, 35, 71
C	61, 30, 36, 42
D	51, 47, 58, 39, 77

Pri stopnji značilnosti $\alpha = 0,01$ preverite ali sta v Sloveniji Najljubša radijska postaja in Starost povezani spremenljivki.

Rešitev: Označimo radijske postaje A=1.skupina, B=2.skupina, C=3.skupina in D=4.skupina. Potem velja $n_1 = 6$, $n_2 = 6$, $n_3 = 4$ *in* $n_4 = 5$. *Računamo*

$$\mu_1 = \frac{20+32+45+33+29+44}{6} = 33,83$$

$$\mu_2 = \frac{26+29+41+49+35+71}{6} = 41,83$$

$$\mu_3 = \frac{61+30+36+42}{4} = 42,25$$

$$\mu_4 = \frac{51+47+58+39+77}{5} = 54,40$$

Izračunajmo še skupno aritmetično sredino

$$\mu = \frac{n_1}{n}\mu_1 + \frac{n_2}{n}\mu_2 + \frac{n_3}{n}\mu_3 + \frac{n_4}{n}\mu_4,$$

torej

$$\mu = \frac{6}{21}33,83 + \frac{6}{21}41,83 + \frac{4}{21}42,25 + \frac{5}{21}54,40 = 42,62.$$

Naprej računamo nepojasnjeno varianco, zato potrebujemo

$$\sigma_1^2 = \frac{20^2 + 32^2 + 45^2 + 33^2 + 29^2 + 44^2}{6} - 33,83^2 = 74,70,$$

$$\sigma_2^2 = \frac{26^2 + 29^2 + 41^2 + 49^2 + 35^2 + 71^2}{6} - 41,83^2 = 227,75,$$

$$\sigma_3^2 = \frac{61^2 + 30^2 + 36^2 + 42^2}{4} - 42,25^2 = 135,19.$$

$$\sigma_4^2 = \frac{51^2 + 47^2 + 58^2 + 39^2 + 77^2}{5} - 54,40^2 = 165,44.$$

In nepojasnjena varianca je

$$\sigma_W^2 = \frac{n_1}{n}\sigma_1^2 + \frac{n_2}{n}\sigma_2^2 + \frac{n_3}{n}\sigma_3^2 + \frac{n_4}{n}\sigma_4^2,$$

torej

$$\sigma_W^2 = \frac{6}{21} \cdot 74,70 + \frac{6}{21} \cdot 227,75 + \frac{4}{21} \cdot 135,19 + \frac{5}{21} \cdot 165,44 = 151,56.$$

Izračunajmo še pojasnjeno varianco

$$\sigma_B^2 = \frac{n_1}{n}(\mu_1 - \mu)^2 + \frac{n_2}{n}(\mu_2 - \mu)^2 + \ldots + \frac{n_k}{n}(\mu_k - \mu)^2$$

kar je v našem primeru

$$\sigma_B^2 = \frac{6}{21}(33,83-42,62)^2+\frac{6}{21}(41,83-42,62)^2+\frac{4}{21}(42,25-42,62)^2+\frac{5}{21}(54,40-42,62)^2 = 55,32.$$

Zapišemo
H_0: *Med intervalsko (Starost) in imensko spremenljivko (Najljubša slovenska radijska postaja) ni povezave. oz. Povprečja so v populaciji na vseh treh skupinah enaka.*
H_1: *Med intervalsko (Starost) in imensko spremenljivko (Najljubša slovenska radijska postaja) je povezava. oz. Vsaj dve povprečji skupin sta v populaciji različni.*

Izvedemo F-test, kjer je testna statistika enaka

$$F = \frac{n-k}{k-1} \cdot \frac{\sigma_B^2}{\sigma_W^2}.$$

Porazdelitev testne statistike ob veljavni ničelni domnevi je Fisherjeva (Snedecorjeva) porazdelitev s $(k-1, n-k)$ *prostostnimi stopinjami. Računamo*

$$F = \frac{(21-4)}{(4-1)} \cdot \frac{55,32}{151,56} = 2,07.$$

Ničelno domnevo zavreno, če velja

$$F > F_{1-\alpha}(k-1, n-k).$$

Vrednost testne statistike primerjamo z vrednostjo iz tabele

$$F_{1-\alpha}(k-1, n-k) = F_{0{,}99}(3, 17) = 5{,}29.$$

Ker NE velja $2{,}07 > 5{,}29$, *ne moremo zavrniti ničelne domneve. Torej naši podatki ne kažejo statistično značilne povezanosti med spremenljivkama Starost in Najljubšo slovensko radijsko postajo.*

Vaja 12

2. V obdobju stroge karantene smo med študenti UP Famnit naredili anketo o tem kako se počutijo in kako pogosto spremljajo dnevna poročila. Vprašanji sta se glasili: Kako se počutite med trenutno karanteno:

   (a) Slabo.
   (b) Niti slabo, niti dobro.
   (c) Dobro.

   Kako pogosto spremljate dnevna poročila:

   (a) Nikoli.
   (b) Občasno.
   (c) Redno.

   Pridobili smo podatke, ki so zbrani v naslednji kontingenčni tabeli:

Spremlj. poročil\Počutje	Slabo	Niti slabo, niti dobro	Dobro
Nikoli	1	3	7
Občasno	5	6	7
Redno	9	6	3

   (a) Izračunajte Spearmanov korelacijski koeficient.
   Rešitev: Ko je podana kontingenčna tabela, se formule za računanje rahlo razlikujejo. Formula za povprečni rang ostaja enaka. Če z $R^{(x)}(a)$ *označimo vezani rang vrednosti* a *glede na spremenljivko* X *in z* $R^{(y)}(b)$ *vezani rang vrednosti* b *glede na spremenljivko* Y, *potem velja*

$$K_{R^{(X)},R^{(Y)}} = \frac{1}{n}\sum_{i=1}^{k}\sum_{j=1}^{l} f_{ij}(R^{(x)}(a_i) - \bar{R})(R^{(y)}(b_j) - \bar{R}),$$

$$s_{R^{(x)}} = \sqrt{\frac{1}{n}\sum_{i=1}^{k} f_{i\cdot}(R^{(x)}(a_i) - \bar{R})^2}$$

$$s_{R^{(y)}} = \sqrt{\frac{1}{n}\sum_{j=1}^{l} f_{\cdot j}(R^{(y)}(b_j) - \bar{R})^2}.$$

   Poglejmo najprej kako velik je naš vzorec

Spremlj. poročil \ Počutje	*Slabo*	*Niti slabo, niti dobro*	*Dobro*	*Skupaj*
Nikoli	*1*	*3*	*7*	*11*
Občasno	*5*	*6*	*7*	*18*
Redno	*9*	*6*	*3*	*18*
Skupaj	*15*	*15*	*17*	*47*

Vidimo, da je $n = 47$. *Torej bo povpreav cni rang enak*

$$\bar{R} = \frac{n+1}{2} = 48/2 = 24.$$

Rangi posameznih odgovorov in njihovi odmiki od povprečja so enaki

a_i	$R^{(x)}(a_i)$	$R^{(x)}(a_i) - \bar{R}$
Slabo	*(1+15)/2=8*	*-16*
NSND	*(16+30)/2=23*	*-1*
Dobro	*(31+47)/2=39*	*15*

b_i	$R^{(y)}(b_j)$	$R^{(y)}(b_j) - \bar{R}$
Nikoli	*(1+11)/2=6*	*-18*
Občasno	*(12+29)/2=20,5*	*-3,5*
Redno	*(30+47)/2=38,5*	*14,5*

Sledi

$$\begin{aligned}
K_{R^{(X)},R^{(Y)}} &= \frac{1}{47}[1 \cdot (-16) \cdot (-18) + 3 \cdot (-1) \cdot (-18) + 7 \cdot 15 \cdot (-18) \\
&\quad +5 \cdot (-16) \cdot (-3,5) + 6 \cdot (-1) \cdot (-3,5) + 7 \cdot 15 \cdot (-3,5) \\
&\quad +9 \cdot (-16) \cdot 14,5 + 6 \cdot (-1) \cdot 14,5 + 3 \cdot 15 \cdot 14,5] \\
&= -66,75
\end{aligned}$$

Z izračunanima tabelama si bomo pomagali tudi pri standardnih odklonih

$$s_{R^{(x)}} = \sqrt{\frac{1}{47}\left(15 \cdot (-16)^2 + 15 \cdot (-1)^2 + 17 \cdot 15^2\right)} = 12,78$$

$$s_{R^{(y)}} = \sqrt{\frac{1}{47}\left(11 \cdot (-18)^2 + 18 \cdot (-3,5)^2 + 18 \cdot 14,5^2\right)} = 12,69$$

Sedaj lahko izračunamo Spearmanov korelacijski koeficient

$$\rho = \rho_{X,Y} = \frac{K_{R^{(X)},R^{(Y)}}}{s_{R^{(X)}} s_{R^{(Y)}}} = \frac{-66,75}{12,78 \cdot 12,69} = -0,41.$$

(b) Pri stopnji značilnosti $\alpha = 0,05$ preverite ali sta spremenljivki negativno korelirani.

Rešitev: Zapišemo ničelno in alternativno hipotezo
H_0: *X (Počutje) in Y (Spremljanje poročil) nista povezani spremenljivki.*
H_1: *X (Počutje) in Y (Spremljanje poročil) sta negativno povezani spremenljivki.*

Izračunamo vrednost testne statistike

$$T = \frac{\rho}{\sqrt{1-\rho^2}}\sqrt{n-2}.$$

torej

$$T = \frac{-0,41}{\sqrt{1-(-0,41)^2}}\sqrt{47-2} = -3,02.$$

Glede na našo alternativno domnevo uporabimo T-test v levo. Ničelno domnevo zavrnemo, če velja

$$T < -t_{1-\alpha}(n-2).$$

Iz tabele razberemo

$$t_{1-\alpha}(n-2) = t_{0,95}(45) = 1,68.$$

Ker velja $-3,02 < -1,68$, *lahko zavrnemo ničelno domnevo. Torej naši podatki kažejo, da sta spremenljvki Počutje in Spremljanje poročil v populaciji študentov UP Famnit statistično značilno negativno korelirani. Z drugimi besedami študentje UP Famnit, ki so bolj pogosto spremljali dnevna poročila, so se v karanteni slabše počutili.*

Vaja 13

2. Med 70-letniki smo izvedli anketo, da bi ugotovili ali obstaja povezava med počutjem in letnim časom. V vsakem letnem času smo poklicali nekaj starostnikov in jih povpraševali po počutju. Možni odgovori so bili Zelo slabo (ZS), Slabo (S), Nevtralno (N), Dobro (D), Odlično (O). Dobili smo naslednje rezultate.

Počutje\Letni čas	Pomlad	Poletje	Jesen	Zima
ZS	0	5	8	10
S	1	4	7	8
N	9	0	3	5
D	11	5	1	2
O	13	8	0	2

Pri stopnji značilnosti $\alpha = 0,05$ preverite ali v populciji 70-letnikov obstaja povezava med počutjem in letnim časom.

Rešitev:vZapišimo ničelno in alternativno domnevo:
H_0: *Med spremenljivkama Počutje in Letni čas ni povezave.*
H_1: *Med spremenljivkama Počutje in Letni čas je povezava.*

Za izračun testne statistike potrebujemo pojasnjeno varianco, kar pomeni, da moramo poračunati vse range in povprečne range po skupinah. Range urejenostne spremenljivke Počutje izračunamo kar v tabeli

Počutje | Letni čas	*Pomlad*	*Poletje*	*Jesen*	*Zima*	*Skupaj*	*Rang*
ZS	*0*	*5*	*8*	*10*	*23*	$\frac{1+23}{2} = 12$
S	*1*	*4*	*7*	*8*	*20*	$\frac{24+43}{2} = 33,5$
N	*9*	*0*	*3*	*5*	*17*	$\frac{44+60}{2} = 52$
D	*11*	*5*	*1*	*2*	*19*	$\frac{61+79}{2} = 70$
O	*13*	*8*	*0*	*2*	*23*	$\frac{80+102}{2} = 91$
Skupaj	*34*	*22*	*19*	*27*	*102*	

Iz tabele vidimo, da je $n = 102$. *Izračunamo povprečne range po skupinah imenske spremenljivke (v našem primeru so to Pomlad, Poletje, Jesen, Zima)*

$$\bar{R}_1 = \frac{1}{34}(0 \cdot 12 + 1 \cdot 33,5 + 9 \cdot 52 + 11 \cdot 70 + 13 \cdot 91) = 72,19$$

$$\bar{R}_2 = \frac{1}{22}(5 \cdot 12 + 4 \cdot 33,5 + 0 \cdot 52 + 5 \cdot 70 + 8 \cdot 91) = 57,82$$

$$\bar{R}_3 = \frac{1}{19}(8 \cdot 12 + 7 \cdot 33,5 + 3 \cdot 52 + 1 \cdot 70 + 0 \cdot 91) = 29,29$$

$$\bar{R}_4 = \frac{1}{27}(10 \cdot 12 + 8 \cdot 33,5 + 5 \cdot 52 + 2 \cdot 70 + 2 \cdot 91) = 35,93$$

Povprečni rang je

$$\bar{R} = \frac{n+1}{2} = \frac{102+1}{2} = 51,5.$$

Varianca ranga med skupinami oziroma pojasnjena varianca (skupine imenske spremenljivke, kar so v našem primeru pomlad, poletje, jesen, zima) pa je enaka

$$s_B^2 = \frac{1}{102}\left(34 \cdot (72,19 - 51,5)^2 + 22 \cdot (57,82 - 51,5)^2 + 19 \cdot (29,29 - 51,5)^2 + 27 \cdot (35,93 - 51,5)^2\right)$$

kar je enako

$$s_B^2 = 307,3646.$$

Minimalni pogoj za uporabo testa hi kvadrat, da je v vsaki skupini vsaj 5 enot, je izpolnjen. Izračunamo vrednost testne statistike

$$K = \frac{12}{n+1} s_B^2 = \frac{12}{102+1} \cdot 307,3646 = 35,81.$$

To vrednost primerjamo z vrednostjo iz tabele za χ^2 *porazdelitev*

$$\chi^2_{1-\alpha}(k-1) = \chi^2_{0,95}(3) = 7,815.$$

Ker velja $35,81 > 7,815$, *lahko zavrnemo ničelno domnevo. Naši podatki kažejo, da sta počutje in letni čas v populaciji 70-letnikov statistično značilno povezani spremenljivki.*