STATISTIKA - Kvantili

Nosilec:

prof.dr.Srečko Devjak

KVANTILI

- ***Problem*** :

  Določanja položaja (mesta, pozicije) statistične enote v množici glede na vrednost spremenljivke :

  - **Rang in**
  - **Kvantilni rang**

- ***Obratni problem:***

  ***D*oločanje vrednosti** spremenljivke, da bi se razporedila na opredeljeno mesto v množici – imenujemo ga **kvantil**.

Ranžirna vrsta in rang

Ranžirna vrsta za **$y: y_1 < y_2 < y_3 < \ldots < y_N$**

Rang za y **R: 1 2 3 ... N**

- **Ranžirna vrsta** je razpored enot po vrednosti številske spremenljivke od najmanjše do največje (ali obratno).
- **Rang** je mesto ki ga enota zaseda v ranžirni vrsti.
- **Povprečni rang** je rang, ki je opredeljen enotam z enako vrednostjo y.

Tabela 4.1: Vrednost proračunskih prihodkov na prebivalca v letu 2002 za 11 slovenskih mestnih občin
(Vir: URL: http://www.fu.uni-lj.si/sib/vhod.htm)

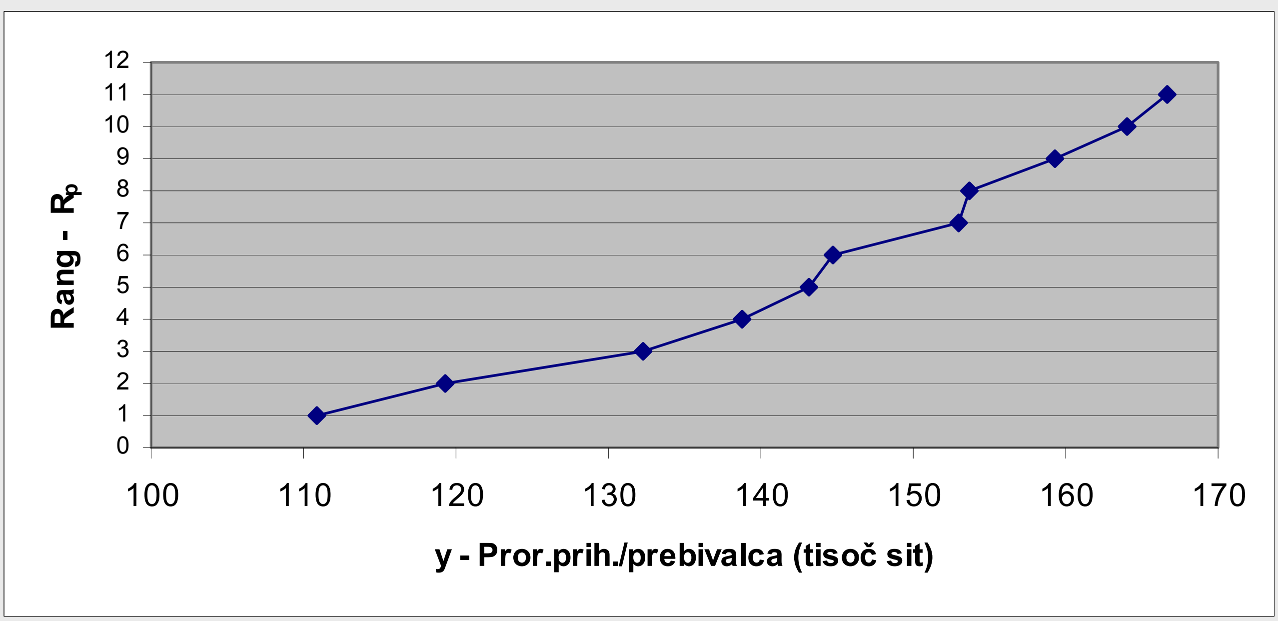
Mestna občina	Prihodki/prebivalca (tisoč sit	Rang
N. GORICA	167	1
KOPER	164	2
LJUBLJANA	159	3
N. MESTO	154	4
CELJE	153	5
S. GRADEC	145	6
MARIBOR	143	7
M. SOBOTA	139	8
PTUJ	132	9
KRANJ	119	10
VELENJE	111	11

Primer:

Tabela 4.1: Vrednost proračunskih prihodkov na prebivalca v letu 2002 za 11 slovenskih mestnih občin

(Vir: URL: http://www.fu.uni-lj.si/sib/vhod.htm)

Mestna občina	Prihodki/prebivalca (tisoč sit	Rang
VELENJE	111	1
KRANJ	119	2
PTUJ	132	3
M. SOBOTA	139	4
MARIBOR	143	5
S. GRADEC	145	6
CELJE	153	7
N. MESTO	154	8
LJUBLJANA	159	9
KOPER	164	10
N. GORICA	167	11



Kvantilni rang

Kvantilni rang P nam pove (v deležu), na katerem delu ranžirnega razmika leži določena enota ki ji pripada rang R.

$$P = \frac{R_P - 0.5}{N}$$

$$R_P = P * N + 0,5$$

P- kvantilni rang
R_P - rang
N- število enot v populaciji

***Primer*: Izračunajmo za mestno občino, ki ima v letu 2002, vrednost proračunski prihodkov na prebivalca 143 tisoč sit njen kvantilni rang P!**

Tabela 4.2: proračunski prihodki slovenskih mestnih občin na prebivalca za leto 2002 (Vir:Tabela 4.1)

Prihodki/prebivalca (tisoč sit)	167	164	159	154	153	145	143	139	132	119	111
Rang	1	2	3	4	5	6	7	8	9	10	11

$$P_{143} = \frac{R_{143} - 0.5}{11} = \frac{7 - 0.5}{11} = 0.59$$

59 % slov.mestnih občin je imelo v letu 2002 proračunske prihodke višje od 143 tisoč sit /prebivalca.

Kvantili

Kvantil (y_P) je vrednost spremenljivke y za katero ima¸100*P odstotkov enot populacije manjše ali enake vrednosti od te vrednosti.

<u>Postopek za posamične podatke:</u>

- Podatke uredimo v ranžirno vrsto,
- Za izbrani kvantilni rang P izračunamo rang R_P:

R_P=N*P +0,5

- Poiščemo v ranžirni vrsti : $R_{-1} \leq R_P \leq R_0$
- določim R_{-1} in R_0 pripadajoče y_{-1} in y_0
- Z metodo linearne interpolacije izračunam:

$$\frac{y_p - y_{-1}}{y_0 - y_{-1}} = \frac{R_p - R_{-1}}{R_0 - R_{-1}}$$

$$y_p = y_{-1} + (y_0 - y_{-1})\frac{R_p - R_{-1}}{R_0 - R_{-1}}$$

V knjigi: $y_P = y_{-1} + (y_0 - y_{-1})*(R_P - R_{-1})$ če je $R_o - R_{-1} = 1$

<u>Postopek:</u> $P \longrightarrow R_p \longrightarrow y_p$

Najpogosteje uporabljani kvantili

- **Mediana** - Me, množico razdeli na dva enaka dela,
- **Kvartili** - Q1 ,Q2, Q3, množico razdelijo na 4 enake dele,
- **Decili** -D1 , D2…….D8, D9, množico razdelijo na 10 enaki delov.

<u>Primer:</u>Katera je tista vrednost proračunskih prihodkov na prebivalca med mestnimi občinami, od katere je 80 % občin imelo višjo vrednost?

Tabela 4.3:Uporabimo podatke iz tabele 4.2:

Prihodki/prebivalca (tisoč sit)	167	164	159	154	153	145	143	139	132	119	111
Rang	1	2	3	4	5	6	7	8	9	10	11

Za dani kvantilni rang izračunamo ustrezni rang: $R_p = N*P + 0,5$

$$R_p = P * N + 0.5 = 0.8 * 11 + 0.5 = 9.3$$

Pogledamo med katera ranga spada izračunan rang:

$$R_{-1} < R_p \leq R_0$$

Opredelimo ustrezne vrednosti: $y_{-1}=132$ in $y_0=119$

Vstavimo opredeljene vrednosti v obrazec

$$y_P = y_{-1} + (y_0 - y_{-1}) * (R_P - R_{-1})$$

$$y_{0.8} = 132 + (119 - 132) * (9.3 - 9) = 128,1$$

Odgovor: 80% mestnih občin je imelo proračunske prihodke večje ali enake 128,1 tisoč sit.

Primer: Katera je tista vrednost proračunskih prihodkov na prebivalca od katere je bila v letu 2002 pri 35% slovenskih mestnih občinah ta vrednost nižja?

Tabela 4.4: Uporabimo potatke iz tabele 4.2

Prihodki/ prebivalca (tisoč sit)	111	119	132	139	143	145	153	154	159	164	167
Rang	1	2	3	4	5	6	7	8	9	10	11

$$R_{0.35} = P * N + 0.5 = 0.35 * 11 + 0.5 = 4.35$$

$$y_P = y_{-1} + (y_0 - y_{-1}) * (R_P - R_{-1})$$

$$y_{0.35} = 139 + (143 - 139) * (4.35 - 4)$$

$$y_{0.35} = 140.4$$

Odgovor: 35 % od 11 slovenskih mestnih občin je imela v letu 2002 proračunske prihodke na prebivalca nižje od 140,4 tisoč sit.

KVANTILNI RANG
za vrednost y določimo pripadajoči kvantilni rang P_y

- **Postopek pri posamičnih podatkih**

- Podatke uredimo v ranžirno vrsto,
- Poiščemo sosednji vrednosti spremenljivk , da velja za poznano vrednost y:

  $y_{-1} \leq y \leq y_0$

- Za izbranivrednosti y_{-1} in y_0 ugotovimo R_{-1} in zračunamo R:

$$\frac{y_p - y_{-1}}{y_0 - y_{-1}} = \frac{R_p - R_{-1}}{R_0 - R_{-1}}$$

$$R_y = R_{-1} + \frac{y - y_{-1}}{y_0 - y_{-1}}(R_0 - R_{-1})$$

$$P_y = \frac{R_y - 0.5}{N}$$

Postopek:

$$y_P \longrightarrow R_P \longrightarrow P_y$$

Primer:Koliko % slovenskih mestnih občin, je imelo v letu 2002 vrednost prihodkov na prebivalca nižjo od 150 tisoč sit?

Tabela 4.5: Uporabimo potatke iz tabele 4.2

Prihodki/ prebivalca (tisoč sit)	111	119	132	139	143	145	153	154	159	164	167
Rang	1	2	3	4	5	6	7	8	9	10	11

$y=150,\ y_{-1}=145\ \ y_0=153$

$R_{-1}=6$

$$R_y = R_{-1} + (R_0 - R_{-1})\frac{y - y_{-1}}{y_0 - y_{-1}} = 6 + 1.\frac{150-145}{153-150} = 6{,}63$$

$$P_y = \frac{R_y - 0.5}{N} = \frac{6{,}63 - 0{,}5}{11} = 0{,}56$$

Odgovor: 56% slovenskih mestnih občin je imelo proračunske prihodke na prebivalca nižje od 150 tisoč sit

Izračunavanje kvantilov y_P iz frekvenčnih porazdelitev

- Za dani kvantilni rang P izračunamo rang:

  $R_P = P*N+0,5$

- V kumulativni frekvenčni porazdelitvi ugotovimo kvantilni **razred** da velja:

  $$F_{-1} < R_P \leq F_0$$

- V kvantilnem razredu označimo z:

  F_{-1}, f_0, d_0, $y_{0,s}$

- na osnovi linearne interpolacije izračunamo ustrezni kvantil

Izračun kvantila

1. $\boldsymbol{R_P = PN + 0{,}5}$

2. $$y_p = y_{0,s} + d_0 * \frac{R_p - F_{-1}}{f_0}$$

Obrazec izhaja iz zveze:

$$\frac{y - y_{O,s}}{d_0} = \frac{R_y - F_{-1}}{F_0 - F_{-1}}$$

$$\frac{y - y_{O,s}}{d_0} = \frac{R_y - F_{-1}}{f_0}$$

Postopek: $P \Longrightarrow R_p \Longrightarrow y_p$

Izračun kvantilnega ranga P za izbrano vrednost kvantika y_P

- V frekvenčni poraz. določimo razred za – kvantil y_P.
- Ta razred imenujemo **kvantilni razred.**
- Za kvantilni razred ugotovimo:

$$F_{-1},\ f_0,\ d_0,\ y_{0,s}$$

- Za dani kvantil y_P izračunamo rang

Postopek:

$$y_P \longrightarrow R_P \longrightarrow P_y$$

- **Ran R_y za znani y:**

$$R_y = F_{-1} + f_0 * \frac{y - y_{0,s}}{d_0}$$

$$UPORABLJENO:$$

$$\frac{y - y_{o,s}}{d_0} = \frac{R_y - F_{-1}}{f_0}$$

- **Za izračunan rang R izračunamo kvantilni rang:**

$$P_y = \frac{R_y - 0{,}5}{N}$$

Zveza kvantilni rang in rang:

$$R_P = P * N + 0{,}5$$

Posamični podatki

$$\frac{y_p - y_{-1}}{y_0 - y_{-1}} = \frac{R_p - R_{-1}}{R_0 - R_{-1}}$$

Frekvenčna porazdelitev:

$$\frac{y - y_{O,s}}{d_0} = \frac{R_y - F_{-1}}{F_0 - F_{-1}}$$

$$\frac{y - y_{O,s}}{d_0} = \frac{R_y - F_{-1}}{f_0}$$