RAČUNANJE KVARTILOV V ELEMENTARNI STATISTIKI

JANEZ ŽEROVNIK

Fakulteta za strojništvo, Univerza v Ljubljani
Inštitut za matematiko, fiziko in mehaniko

Uvod

Definicija kvartilov se v teoriji verjetnosti nanaša na pojem porazdelitve verjetnosti, v statistiki pa se seveda hitro vprašamo, kdaj in kako lahko iz vzorca ocenimo kvartile porazdelitve. Pri vpeljavi osnovnih pojmov statistike na ravni osnovne in srednje šole abstraktnega pojma porazdelitve seveda ne moremo vpeljati, uporabljamo pa elementarne definicije za posebne primere, kar je eden od razlogov za nejasnosti, saj je končno množico vrednosti mogoče razumeti na različne načine: na primer kot diskretno porazdelitev z natanko temi danimi vrednostmi in enakimi verjetnostmi ali pa kot vzorec neke splošne in neznane porazdelitve. Naprej, iz končno mnogo podatkov lahko v nekaterih praktičnih primerih sklepamo, da gre za vzorec, ki smo ga dobili z nekaj realizacijami zvezne slučajne spremenljivke, v drugih primerih lahko verjamemo, da slučajna spremenljivka zavzame samo končno mnogo vrednosti, morda celo samo tiste, ki so že v dani množici podatkov, če omenimo samo dva primera.

Verjetno ni treba posebej utemeljevati, da morajo kvartili imeti naslednji dve lastnosti:

1. kvartili razdelijo elemente na štiri približno enake dele;

2. prvi in tretji kvartil sta mediani spodnje in zgornje polovice.

Seveda je treba tudi mediano nedvoumno definirati, prav tako je treba pojasniti pojem »polovic«, še zlasti pri lihem številu elementov. Pri definiciji kvartilov sta gotovo zaželeni vsaj še naslednji dve lastnosti:

3. vrednosti kvartilov so enake na podvojenih podatkih (če vsak element podvojimo in dobimo množico z $2n$ elementi, so vrednosti novih kvartilov enake prejšnjim);

4. definicija za končne množice podatkov in definicija za zvezne porazdelitve sta posebna primera splošne definicije (če je končna množica vzorec neke splošne porazdelitve, potem so kvartili na vzorcu dobri približki za kvartile te porazdelitve pri predpostavki, da je vzorec nastal kot realizacija neodvisnih enako porazdeljenih slučajnih spremenljivk in da kvartil res obstaja ter je enolično določen).

Kvartili so poseben primer centilov (ali percentilov) in oboji poseben primer kvantilov [2, 11]. Čeprav je prava definicija kvantilov in s tem centilov za diskretno porazdelitev (in enaka definicija za končno množico podatkov) jasna [2] in je na prvi pogled edino smiselno definirati kvartile z ustreznimi kvantili, v literaturi in praksi pri metodah za računanje kvartilov vlada precejšnja zmešnjava.[1] Statistiki uporabljajo različne metode za računanje kvartilov. Videti je, da te približne metode računanja nekateri razumejo kot definicije. Nejasnost se na žalost prenaša tudi na vpeljavo teh osnovnih pojmov v elementarni statistiki [10, 4, 3]. Zadrega je precej hujša, kot je videti na prvi pogled in kot večina misli. (Ali, kot pravi Langford [4]: »The situation is, I believe, far worse than most realize.«) Langford v članku [4] navaja sedem različnih metod in še nekaj na videz drugih, ki so ekvivalentne kateri od prvih sedem. Implementacije v splošno uporabljanih kalkulatorjih in programju (MINITAB, SAS, Mathematica, JMP, Microsoft Excel) dodajo še pet dodatnih metod. Različne metode dajejo na majhnih primerih različne rezultate in ker, kot zapisano, nekateri te metode razumejo kot definicije kvartilov, je zmešnjava popolna. Langford si predstavlja študenta, ki s svojim računalom, na fakulteti priporočenim statističnim programskim orodjem, in z računanjem »peš« dobi vsakič drugačen rezultat! V Sloveniji so bili osnovni pojmi statistike vpeljani v osnovne in srednje šole ob uvedbi devetletne osnovne šole in s tem povezani prenovi učnih načrtov [6, 5]. Na žalost v veljavnih učbenikih [1, 8, 9] najdemo različne metode za računanje

[1] Definicija (enačba 29.1 na strani 195 v [2]) kvantilov ne določa enolično, od koder deloma izhaja zmeda, o kateri govori ta članek.

kvartilov in ker manjkajo formalne definicije, je videti, kot da so te približne metode eksaktne, s čimer sta implicitno vpeljani vsaj dve problematični (da ne uporabimo besede napačni) definiciji kvartilov [1, 8, 9]. Na maturitetni komisiji smo nedavno dobili vprašanje zaskrbljene matere dvojčic, ki sta v osnovni šoli in na gimnaziji opazili različne metode z različnimi rezultati. Kaj je pravilno?

V nadaljevanju bomo definirali kvartile, najprej za splošno in potem še za diskretne porazdelitve s končno zalogo vrednosti. V primeru enakomerne diskretne porazdelitve s končno zalogo vrednosti lahko enakovredno govorimo o kvantilih končne množice, torej o nalogi, ki se obravnava v srednji in osnovni šoli. Potem bomo opisali dve metodi iz naših učbenikov in podobno metodo, ki računa prave vrednosti. Ker je zadnja metoda malenkost bolj zapletena kot prvi dve, v nadaljevanju opišemo še tri ekvivalentne metode, ki računajo prave vrednosti kvartilov in so morda primerne za vpeljavo v osnovni šoli, zagotovo pa niso preveč zahtevne za obravnavo v srednji šoli.

Definicija kvartilov

Definicija kvartilov (in kvantilov) za porazdelitve s porazdelitveno funkcijo je nesporna. q-ti kvantil je enolično določen, če ima enačba $F(x_q) = q$ natanko eno rešitev. V primeru, ko je slučajna spremenljivka zvezna in ima gostoto p, lahko enačbo zapišemo v obliki $F(x_q) = \int_{-\infty}^{x_q} p(t)dt = q$ in lahko se zgodi, da enačba nima enolične rešitve. Če je slučajna spremenljivka diskretna, potem je F stopničasta in enačba $F(x_q) = q$ praviloma ne bo enolično rešljiva. Kvantili torej v nekaterih primerih niso enolično določeni, ali drugače zapisano, obstaja več vrednosti, ki ustrezajo definiciji takega kvantila. Ker želimo obravnavo ohraniti na ravni elementarne matematike, se bomo namesto splošnih kvantilov tu omejili na kvartile in centile. Zato zapišimo, da je i-ti **centil** vsako število P_i, za katero velja $\int_{-\infty}^{P_i} p(x)dx = i/100$. (Torej centili niso nujno enolično določeni.) Kvartili so seveda 25., 50. in 75. centil, torej **prvi kvartil** $Q_1 = P_{25}$, **drugi kvartil** $Q_2 = P_{50}$, **tretji kvartil** $Q_3 = P_{75}$.

Zapisana definicija temelji na porazdelitveni funkciji $F(x) = \int_{-\infty}^{x} p(t)dt$, ki jo lahko definiramo tudi za diskretne porazdelitve. Zato lahko uporabimo isto definicijo tudi na diskretnih porazdelitvah. A ker, kot že prej omenjeno, pri tem v praksi vlada kar precej zmede, bomo tu posebej zapisali, kako lahko enakovredno definicijo centilov v primeru diskretne porazdelitve, ki ima končno mnogo vrednosti, zapišemo v preprostejšem jeziku. Dana je končna množica elementov (podatkov, meritev) $V = \{v_1, v_2, v_3, \ldots, v_n\}$. Predpostavimo, da je urejena v nepadajoče zaporedje v_k, $k = 1, 2, \ldots, n$, torej da velja $v_1 \leq v_2 \leq v_3 \leq \cdots \leq v_n$. Naj bo $i \in \{0, 1, 2, \ldots, 100\}$. i-ti **centil** P_i je vrednost, za katero velja, da je vsaj i odstotkov elementov manjših ali enakih P_i in da je vsaj $100 - i$ odstotkov elementov večjih ali enakih P_i. (Torej vrednost P_i ni nujno enaka enemu od elementov iz množice $V = \{v_1, v_2, v_3, \ldots, v_n\}$.)

Očitno velja naslednje: če $i\frac{n}{100}$ ni celo število, potem za neki k, $1 \leq k \leq n$, lahko zapišemo $\lfloor i\frac{n}{100} \rfloor = k - 1 < i\frac{n}{100} < \lceil i\frac{n}{100} \rceil = k$, tako da je $k - 1$ manj, k pa več kot i odstotkov od n. (In seveda, $n - k$ je manj, $n - k + 1$ pa več kot $100 - i$ odstotkov od n.) V tem primeru je v skladu z zgornjo definicijo $P_i = v_k$.

Če je $k = i\frac{n}{100}$ naravno število, potem je v množici prvih k elementov natanko i odstotkov elementov množice. (In seveda, $n-k$ je natanko $100-i$ odstotkov od n.) Vsako število med v_k in v_{k+1} torej ustreza definiciji P_i.

V drugem primeru je smiselna naslednja definicija: **Kanonična vrednost** centila je $\bar{P}_i = \frac{v_k + v_{k+1}}{2}$. V posebnih primerih je morda mogoče zagovarjati drugačno izbiro vrednosti med $v_k < v_{k+1}$, torej ni nujno, da za centil vedno vzamemo njegovo kanonično vrednost.

Kvartili so posebni primeri centilov, zato kot prej definiramo: $Q_1 = P_{25}$, $Q_2 = P_{50}$, $Q_3 = P_{75}$.

Pri definiciji **mediane** ni dvoma, v vseh znanih virih je mediana definirana kot kanonična vrednost petdesetega centila (ali, enakovredno, drugega kvartila): za $n = 2k + 1$ je $M = P_{50} = v_{k+1}$, za $n = 2k$ pa $M = P_{50} = \frac{v_k + v_{k+1}}{2}$.

Zgledi: Po definiciji izračunajmo kvartile za naslednje štiri množice velikosti $n = 9, 10, 11, 12$ in jih poimenujmo Zgled 1–4. Vrednosti kvartilov so zapisane krepko. V primeru, ko je kvartil na sredini med dvema elementoma, sta krepko zapisani obe vrednosti.

1. $n = 9$: 12, 15, $\boxed{\mathbf{22}}$, 23, $\boxed{\mathbf{25}}$, 44, $\boxed{\mathbf{46}}$, 51, 59

   $Q_1 = 22$, $Q_2 = M = 25$, $Q_3 = 46$.

2. $n = 10$: 12, 15, $\boxed{\mathbf{22}}$, 23, $\boxed{\mathbf{25,\ 26}}$, 44, $\boxed{\mathbf{46}}$, 51, 59

   $Q_1 = 22$, $Q_2 = M = 25{,}5$, $Q_3 = 46$.

3. $n = 11$: 12, 15, $\boxed{\mathbf{22}}$, 23, 24, $\boxed{\mathbf{25}}$, 26, 44, $\boxed{\mathbf{46}}$, 51, 59

   $Q_1 = 22$, $Q_2 = M = 25$, $Q_3 = 46$.

4. $n = 12$: 12, 15, $\boxed{\mathbf{22,\ 23}}$, 24, $\boxed{\mathbf{25,\ 26}}$, 44, $\boxed{\mathbf{46,\ 51}}$, 59, 88

   $Q_1 = 22{,}5$, $Q_2 = M = 25{,}5$, $Q_3 = 48{,}5$.

Metode v naših učbenikih in Langfordova metoda

V tem razdelku bomo opisali tri metode za računanje kvartilov, ki delujejo tako, da izračunamo mediani na polovici podatkov. Prvi dve metodi sta med najpogosteje uporabljanimi, tretjo pa je predlagal Langford [4] zato, ker prvi dve, tako kot še nekatere druge zgoraj omenjene, ne dajejo pravilnih rezultatov na majhnih množicah podatkov. Vse tri metode (tu jih bomo imenovali M1, M2 in M3) so si zelo podobne, razlikujejo se samo v koraku 2(b), a bomo zaradi nedvoumnosti vse tri postopke zapisali v celoti.

Metoda 1 (M1) je uporabljena v učbeniku [9], sodeč po zgledu na strani 172. Prej je na strani 166 [9] samo zapisano, da je prvi kvartil mediana prve polovice podatkov, tretji kvartil pa mediana druge polovice podatkov, kaj je prva in kaj druga polovica podatkov v primeru lihega števila, ni nikjer definirano. Podobno tudi vir [1] ne pove, kako obravnavati primer z liho mnogo podatki, na zgledu z 11 elementi mediana (pravilno) ni upoštevana, drugega primera z liho mnogo podatki ni.

Metoda M1

1. Izračunamo mediano.
2. Razdelimo podatke na dve podmnožici, v prvi so vsi elementi, ki so manjši od mediane, v drugi polovici so vsi elementi, ki so večji od mediane,

   natančneje:

   (a) če je $n = 2k$, potem je $M = \frac{v_k + v_{k+1}}{2}$ in $V_1 = \{v_1, v_2, \ldots, v_k\}$, $V_2 = \{v_{k+1}, v_{k+2}, \ldots, v_n\}$.
   (b) če je $n = 2k + 1$, potem je $M = v_k$ in $V_1 = \{v_1, v_2, \ldots, v_{k-1}\}$, $V_2 = \{v_{k+1}, v_{k+2}, \ldots, v_n\}$, torej mediane ne štejemo k nobeni od polovic.
3. Prvi kvartil je mediana spodnje, tretji kvartil pa mediana zgornje polovice podatkov.

Metoda 2 (M2) je uporabljena v učbeniku [8], kjer je zapisano, da mediano moramo šteti k obema polovicama podatkov. Enako je v priporočilu [6], kjer je v opombi celo navedeno, da je opis kvartilov iz didaktičnih razlogov nekoliko poenostavljen.

Metoda M2

1. Izračunamo mediano.
2. Razdelimo podatke na dve podmnožici, v prvi so vsi elementi, ki so manjši od mediane, v drugi polovici so vsi elementi, ki so večji od mediane,

   natančneje:

   (a) če je $n = 2k$, potem je $M = \frac{v_k + v_{k+1}}{2}$ in $V_1 = \{v_1, v_2, \ldots, v_k\}$, $V_2 = \{v_{k+1}, v_{k+2}, \ldots, v_n\}$.
   (b) če je $n = 2k + 1$, potem je $M = v_k$ in $V_1 = \{v_1, v_2, \ldots, v_k\}$, $V_2 = \{v_k, v_{k+1}, v_{k+2}, \ldots, v_n\}$, torej mediano dodamo k obema polovicama.
3. Prvi kvartil je mediana spodnje, tretji kvartil pa mediana zgornje polovice podatkov.

Metodo M3, ki je podobna zgornjima in da pravilen rezultat v vseh primerih, je predlagal Langford [4].

Metoda M3 [Langford]

1. Izračunamo mediano.
2. Razdelimo podatke na dve podmnožici, v prvi so vsi elementi, ki so manjši od mediane, v drugi polovici so vsi elementi, ki so večji od mediane,

   natančneje:

   (a) če je $n = 2k$, potem je $M = \frac{v_k + v_{k+1}}{2}$ in $V_1 = \{v_1, v_2, \ldots, v_k\}$, $V_2 = \{v_{k+1}, v_{k+2}, \ldots, v_n\}$.

   (b) če je $n = 2k+1$, potem je $M = v_k$ in ločimo dva primera:

   i. če je k liho število, potem je $V_1 = \{v_1, v_2, \ldots, v_k\}$, $V_2 = \{v_k, v_{k+1}, v_{k+2}, \ldots, v_n\}$, torej mediano dodamo k obema polovicama.

   ii. če je k sodo število, potem je $V_1 = \{v_1, v_2, \ldots, v_{k-1}\}$, $V_2 = \{v_{k+1}, v_{k+2}, \ldots, v_n\}$, torej mediane ne štejemo k nobeni od polovic.
3. Prvi kvartil je mediana spodnje, tretji kvartil pa mediana zgornje polovice podatkov.

V tabeli so prvi in tretji kvartili izračunani po definiciji in po metodah M1, M2 in M3, izračunani so za zglede iz prejšnjega razdelka (z 9, 10, 11 in 12 elementi). Krepko so označene vrednosti, ki se ne ujemajo z definicijo.

	Q_1				Q_3			
	Def	M1	M2	M3	Def	M1	M2	M3
Zgled 1	22	**18,5**	22	22	46	**48,5**	46	46
Zgled 2	22	22	22	22	46	46	46	46
Zgled 3	22	22	**22,5**	22	46	46	**45**	46
Zgled 4	22,5	22,5	22,5	22,5	48,5	48,5	48,5	48,5

Vidimo, da metodi M1 in M2 ne izračunata pravilnih vrednosti v vseh primerih. Ni težko videti, da za primere, ko je $n = 4r + 1$, metoda M1 ne deluje pravilno, za $n = 4r + 3$ pa se od definicije razlikuje rezultat, dobljen po metodi M2. Metoda M3 v vseh primerih pravilno izračuna kvartile, česar ni težko formalno dokazati.

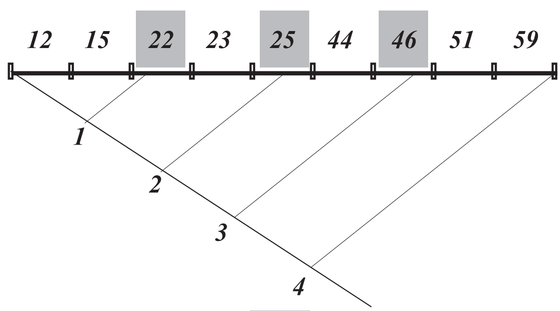
Možna obravnava v šoli

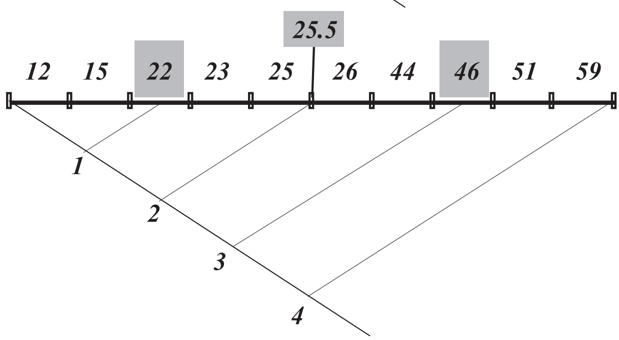
Langford [4] ob predlogu metode pove, da je ni preizkusil v praksi. Čeprav M3 ni pretirano zapletena, v tem razdelku vpeljemo kvartile še na tri druge načine, ki so verjetno primerni za obravnavo v srednji šoli, morda celo v osnovni šoli. Poudarimo, da so v vseh primerih rezultati enaki in ustrezajo pravi definiciji, zato lahko kvartile brez škode tudi »definiramo« s katerim koli od spodnjih postopkov.

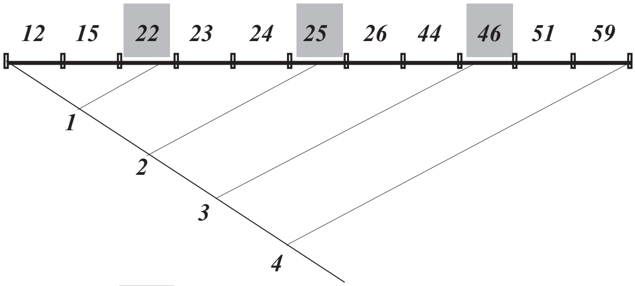
Vpeljava kvartilov s pomočjo elementarne geometrije

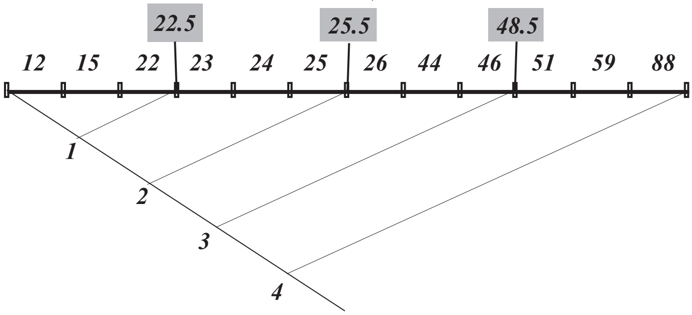
Množico podatkov uredimo in si predstavljajmo, da vsak element pokrije po en interval dolžine ena, tako da dobimo daljico dolžine n. Kvartile in tudi centile dobimo (in definiramo) tako, da daljico dolžine n razdelimo v primernem razmerju. Če je delilna točka na intervalu, potem je vrednost centila (kvartila, mediane) vrednost elementa na tem intervalu, če pa delilna točka pade natanko na mejo med dva intervala, potem za (kanonično) vrednost centila vzamemo aritmetično sredino.

Na zgledih velikosti $n = 9$, 10, 11 in 12 metoda da naslednje rezultate.

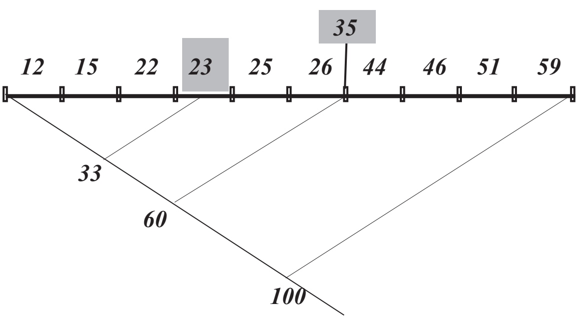








Na ta način brez težav vpeljemo in določamo tudi centile, na primer 33. centil in 60. centil (kanonično vrednost) na zgledu 2 dobimo takole:



Vpeljava kvartilov s pomočjo osnovnega izreka o deljenju

Zapišemo $n = 4r + o$ in kvartile *definiramo* takole:

$$\boxed{\begin{array}{llll}
\text{če je } o = 1, \text{ potem je } Q_1 = v_{r+1}, & Q_2 = v_{2r+1} \text{ in} & Q_3 = v_{3r+1}. \\
\text{če je } o = 2, \text{ potem je } Q_1 = v_{r+1}, & Q_2 = \frac{v_{2r+1}+v_{2r+2}}{2} \text{ in} & Q_3 = v_{3r+2}. \\
\text{če je } o = 3, \text{ potem je } Q_1 = v_{r+1}, & Q_2 = v_{2r+2} \text{ in} & Q_3 = v_{3r+3}. \\
\text{če je } o = 0, \text{ potem je } Q_1 = \frac{v_r+v_{r+1}}{2}, & Q_2 = \frac{v_{2r}+v_{2r+1}}{2} \text{ in} & Q_3 = \frac{v_{3r}+v_{3r+1}}{2}.
\end{array}}$$

Utemeljitev, ki je hkrati tudi formalen dokaz pravilnosti, je preprosta, le obravnavati je treba vsakega od primerov posebej.

- Če je $o = 1$, potem je $Q_1 = v_{r+1}$, ker veljajo neenakosti:

$$\frac{r}{4r+1} < \frac{1}{4} < \frac{r+1}{4r+1}, \qquad \frac{3r}{4r+1} < \frac{3}{4} < \frac{3r+1}{4r+1}.$$

  Podobno pokažemo, da je $Q_3 = v_{3r+1}$, na primer z uporabo očitne simetrije.

- Če je $o = 2$, potem je $Q_1 = v_{r+1}$:

$$\frac{r}{4r+2} < \frac{1}{4} < \frac{r+1}{4r+2}, \qquad \frac{3r+1}{4r+2} < \frac{3}{4} < \frac{3r+2}{4r+2}.$$

  Podobno vidimo, da je $Q_3 = v_{3r+2}$.

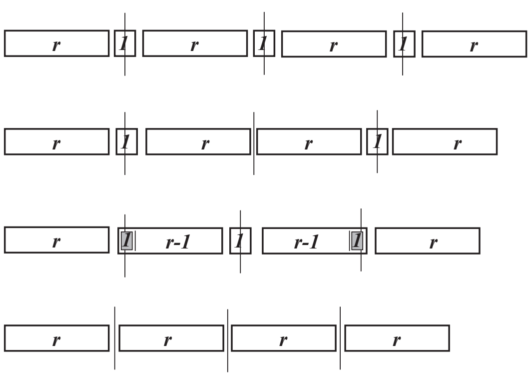
- Če je $o = 3$, potem je $Q_1 = v_{r+1}$:

$$\frac{r}{4r+3} < \frac{1}{4} < \frac{r+1}{4r+3}, \qquad \frac{3r+2}{4r+3} < \frac{3}{4} < \frac{3r+3}{4r+3}.$$

  Podobno vidimo, da je $Q_3 = v_{3r+3}$.

- Če je $o = 0$, potem lahko množico razdelimo na štiri enako velike četrtine s po r elementi, torej je $Q_1 = \frac{v_r + v_{r+1}}{2}$ in $Q_3 = \frac{v_{3r} + v_{3r+1}}{2}$.

Morda lahko namesto za dijake ne ravno zanimive formalne izpeljave za intuitivno razumevanje zadošča spodnja slika:



Podvojitev vzorca

V primeru, ko je število elementov liho, je mediana enaka vrednosti elementa v_{k+1} in če hočemo kvartile računati kot mediane polovic, se pojavi vprašanje, kaj narediti s tem srednjim elementom. Ideja, ki se ponuja in jo je mogoče tudi formalno utemeljiti, je podvojitev: če vsakega od elementov podvojimo, tako da dobimo $2n$ elementov, je delitev na enako veliki podmnožici dobro definirana. Podvojeni srednji element tako prispeva po en element v vsako od polovic. Kvartil Q_1 lahko potem izračunamo (in definiramo) kot mediano spodnje polovice, kvartil Q_3 pa kot mediano zgornje polovice podvojene osnovne množice. Ni težko preveriti, da tako dobimo prave vrednosti kvartilov.

Zaključek

V Sloveniji so bili osnovni pojmi statistike vpeljani v učne načrte osnovne šole in srednjih šol pred slabimi dvajsetimi leti ob uvedbi devetletke. Na žalost v veljavnih učbenikih najdemo različne metode za računanje kvartilov in ob izostanku formalne definicije ali opozorila, da gre za približne metode, je mogoče razumeti, da so te približne metode eksaktne, s čimer sta implicitno vpeljani vsaj dve problematični definiciji kvartilov [1, 8, 9]. Če so razlike v uporabni statistiki zaradi narave različnih aplikacij morda do neke mere razumljive, je nejasnost osnovnih pojmov v elementarni statistiki najmanj neprijetna, če ne celo nesprejemljiva. Če učenci v osnovni in srednji šoli srečajo dve nasprotujoči si definiciji preprostih osnovnih pojmov, upravičeno lahko podvomijo v konsistentnost predavane snovi in z malo posploševanja razširijo ugotovitev na nekonsistentnost celega predmeta, v tem primeru matematike. Zato je v tem in podobnih primerih nujna uskladitev med učnimi gradivi po vertikali in horizontali. Enako pomembna je seveda korektna vpeljava osnovnih pojmov in če gre za prezahtevne pojme, morajo biti razlogi za obravnavo zelo močni, sicer je obravnavo pametneje opustiti ali preložiti na kasnejše obdobje. Spomnimo se samo razvpitega primera vpeljave teorije množic pred desetletji.

Na srečo je tu obravnavani primer precej preprost. Vpeljavo osnovnih pojmov statistike, vključno s kvartili, je vsaj v srednji šoli škoda okrniti, saj za to ni nobene potrebe. Malo bolj vprašljiva je seveda smiselnost obravnave v devetletki, vsekakor je treba razloge za in proti dobro pretehtati [7]. Kvartili so uporabni na primer pri predstavitvi podatkov, kjer razpršenost podatkov lepo prikažemo s tako imenovano »škatlo z brki«.

Torej, k sedaj uporabljanim metodam za računanje kvartilov je nujno treba pripomniti, da gre za približne metode, ki za velike množice podatkov v statistiki delujejo dovolj dobro. Ali, in seveda precej boljše, kvartile vpeljati korektno, na primer z enim od tu nakazanih elementarnih pristopov.

Zahvala

Anonimnemu recenzentu se zahvaljujem za konstruktivne pripombe in predlagane popravke, ki so veliko prispevali k jasnosti in matematični korektnosti besedila. Zahvaljujem se tudi kolegom, članom maturitetne komisije za matematiko za splošno maturo, ki so mi pred pisanjem prispevka opisali svoje izkušnje pri obravnavi teh pojmov in uporabo statističnih programov v šolski praksi, in nenazadnje uredniku, ki je v zadnji različici pred tiskom opozoril še na nekaj napak v besedilu.

LITERATURA

[1] M. Bon Klajnšček, B. Dvoržak in D. Felda, *Matematika 1, učbenik za gimnazije*, DZS, 2009.

[2] R. Jamnik, *Verjetnostni račun*, Mladinska knjiga, Ljubljana, 1971.

[3] A. H. Joarder in M. Firozzaman, *Quartiles for Discrete Data*, Teaching Statistics **3**, 86–89.

[4] E. Langford, *Quartiles in Elementary Statistics*, Journal of Statistics Education **14** (2006) 16 strani, dostopno na: `ww2.amstat.org/publications/jse/v14n3/langford.html`, ogled: 22. 5. 2017.

[5] Z. Magajna in A. Žakelj, *Obdelava podatkov pri pouku matematike 6–9*, Matematika v šoli **7** (1999) 249–252.

[6] Z. Magajna in A. Žakelj, *Obdelava podatkov pri pouku matematike 6–9*, Zavod Republike Slovenije za šolstvo, Ljubljana, 2000.

[7] Z. Magajna in A. Žakelj, *Ali sodi obdelava podatkov k pouku matematike?*, Obzornik mat. fiz. **46** (1999) 113–119.

[8] A. Mohorčič in drugi, *Vega 1, i-učbenik za matematiko v gimnazijah*, Zavod Republike Slovenije za šolstvo, Ljubljana, 2013.

[9] G. Pavlič, D. Kavka, M. Rugelj in J. Šparovec, *Linea nova, matematika za gimnazije*, Modrijan, Ljubljana, 2011.

[10] Ask Dr. Math, dostopno na: `mathforum.org/library/drmath/view/60969.html`, ogled: 24. 12. 2016.

[11] *Statistični terminološki slovar*, Statistično društvo Slovenije, SAZU, 2001.