

## PONAVLJANJE ZA IZPIT

### Primeri izpitnih nalog

1. V spodnji tabeli so podani rezultati izpitov študentov pri izbirnem predmetu v študijskem letu 2015/2016 na eni izmed slovenskih fakultet.

- a) Dopolnite tabelo s kumulativnimi frekvencami.

(3T)

	Frekvenca ( $f_k$ )	Kumulativna frekvenca
Ni opravil/a	8	8
zd(6)	13	21
db(7)	22	43
pd(8)	16	59
pd(9)	10	69
odl(10)	8	77

- b) V zgornjem primeru opredelite stvarno, krajevno in časovno populacijo. Kaj je statistična enota in kaj spremenljivka?

(6T)

Populacija

Stvarna: Študenti

Časovna: Študijsko leto 2015/2016

Krajevna: Ena izmed slovenskih fakultet

Statistična enota: Študent

Statistična spremenljivka: Rezultat izpita (ocena izpita)

Število enot: 77

c) Za mediano je kvantilni rang enak:  $R = 39$ . Katera izmed ocen torej predstavlja mediano?

(2T)

**Odgovor:** Ocena dobro (7)

2. V nekem proizvodnem podjetju so za zadnjih 12 mesecev zbrali podatke o številu proizvodov, ki jih je določen stroj izdelal z napako. Podatki so naslednji:

30    36    10    11    32    40    16    33    34    21    19    5

a) V zgornjem primeru opredelite krajevno populacijo, statistično enoto in spremenljivko.

(3T)

Krajevna populacija: Neko proizvodno podjetje

Statistična enota: Mesec

Statistična spremenljivka: Število proizvodov z napako

b) Sestavite ranžirno vrsto.

(2T)

$x_i$	5	10	11	16	19	21	30	32	33	34	36	40
$R_i$	1	2	3	4	5	6	7	8	9	10	11	12

c) Kolikšen je delež mesecev, ko je stroj izdelal več kot 31 proizvodov z napako?

(5T)

$$x = 31$$

$$x_0 = 30, R_0 = 7$$

$$x_1 = 32, R_1 = 8$$

$$R = R_0 + \frac{x - x_0}{x_1 - x_0} \cdot (R_1 - R_0) = 7 + \frac{31 - 30}{32 - 30} \cdot (8 - 7) = 7.5$$

$$P = \frac{R - 0.5}{N} = \frac{7.5 - 0.5}{12} = 0.583$$

**Delež:**  $1 - 0.583 = 0.417$

**Odgovor:** Stroj je izdelal več kot 31 proizvodov z napako v 41.7% mesecev.

d) Največ koliko proizvodov z napako je izdelal stroj v četrtini mesecev, ko je izdelal najmanj proizvodov z napako? (5T)

$$P = 0.25$$

$$R = P \cdot N + 0.5 = 0.25 \cdot 12 + 0.5 = 3.5$$

$$R_0 = 3, x_0 = 11$$

$$R_1 = 4, x_1 = 16$$

$$x_{0.8} = x_0 + \frac{R - R_0}{R_1 - R_0} \cdot (x_1 - x_0) = 11 + \frac{3.5 - 3}{4 - 3} \cdot (16 - 11) = 13.5$$

**Število:** 13

**Odgovor:** V četrtini mesecev, ko je stroj izdelal najmanj proizvodov z napako, je izdelal  
največ 13 proizvodov z napako.

### 3. Frekvenčna porazdelitev

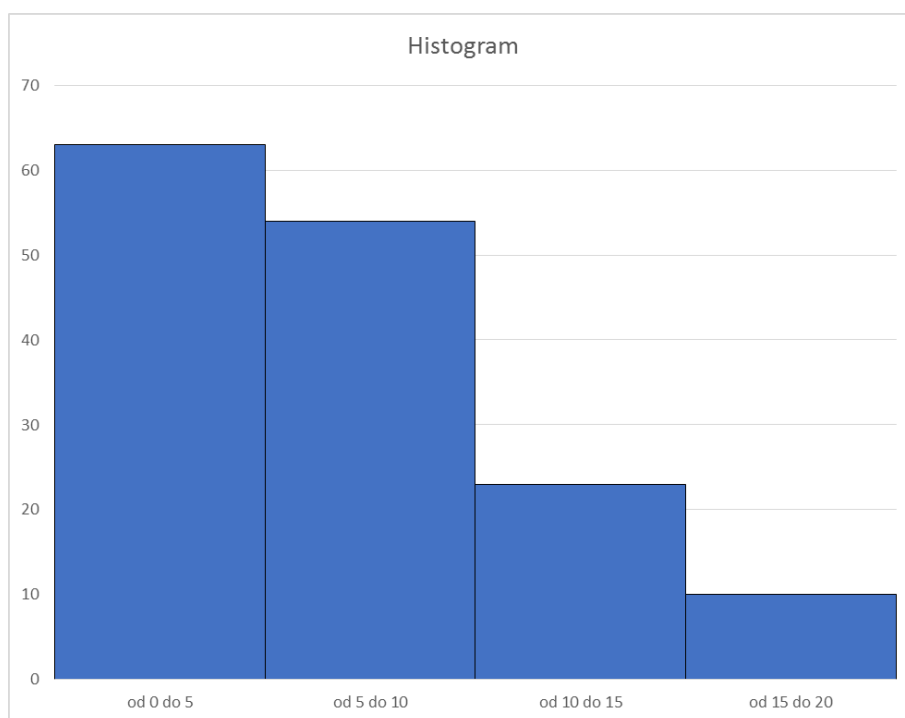
a) Dopolnite naslednjo frekvenčno tabelo (meje razredov, frekvence in kumulativne frekvence):

(6T)

	Razredi $k$	Frekvenca $f_k$	Kumulativna frekvenca $F_k$	Širina razreda $d_i$
1.	od ...0... do ...5...	63	63	5
2.	od ...5... do ...10...	54	117	5
3.	od ...10... do ...15...	23	140	5
4.	od ...15... do ...20...	10	150	5

b) Narišite histogram frekvenčne porazdelitve.

(4T)



c) Za frekvenčno porazdelitev izračunajte modus.

(4T)

MODUS:

$$\max(f_k) = 63 \quad \Rightarrow \quad k = 1$$

$$Mo = x_{k-1,max} + \frac{f_k - f_{k-1}}{2f_k - f_{k-1} - f_{k+1}} \cdot d_k = x_{0,max} + \frac{f_1 - f_0}{2f_1 - f_0 - f_2} \cdot d_1$$

$$= 0 + \frac{63 - 0}{2 \cdot 63 - 0 - 54} \cdot 5 = 0 + \frac{63}{72} \cdot 5 = 4.375$$

**Modus:** 4.375

4. Na nekem taboru so preverjali poznavanje okolja otrok glede na to, ali so gledali dokumentarni film o okolju ali ne. Podatke prikazuje spodnja tabela:

Empirične (opazovane) frekvence			
	Pozna	Ne pozna	Skupaj
Gledal film	25	6	31
Ni gledal filma	8	15	23
Skupaj	33	21	54

- a) S pomočjo hi-kvadrat testa preverite, ali sta spremenljivki povezani.

(6T)

Teoretične (pričakovane) frekvence			
	Pozna	Ne pozna	Skupaj
Gledal film	18,94	12,06	31
Ni gledal filma	14,06	8,94	23
Skupaj	33	21	54

$$\chi^2 = \sum_{i=1}^k \frac{(f_i - f_i')^2}{f_i'} = \frac{(25 - 18.94)^2}{18.94} + \frac{(6 - 12.06)^2}{12.06} + \frac{(8 - 14.06)^2}{14.06} + \frac{(15 - 8.94)^2}{8.94} = 11.69$$

**Vrednost hi-kvadrat testa:** 11.69

**Odgovor:** Vrednost hi-kvadrat testa je večja od 0, nakazuje se povezanost med  
Spremenljivkama.

b) Izračunajte ustrezen kontingenčni koeficient in opredelite intenziteto povezanosti.

(3T)

$$C = \sqrt{\frac{\chi^2}{\chi^2 + N}} = \sqrt{\frac{11.69}{11.69 + 54}} = 0.42$$

$$C_{pop} = \frac{C}{\sqrt{\frac{k-1}{k}}} = \frac{0.42}{\sqrt{\frac{1}{2}}} = \frac{0.42}{0.71} = 0.597$$

Vrednost kontingenčnega koeficienta: 0.597

Odgovor: Spremenljivki sta srednje močno povezani

če je  $0.1 < C_{pop} < 0.3$ , potem obstaja šibka povezanost med spr.;

če je  $0.3 < C_{pop} < 0.6$ , potem med spr. obstaja srednje močna povezanost;

če je  $0.6 < C_{pop} < 1.0$ , potem med spr. obstaja močna povezanost.

5. Raziskati želimo, ali je raven glukoze pri posameznikih linearno odvisna od njihove starosti. Naslednja tabela prikazuje podatke za 6 oseb:

<i>Starost</i>	<i>Raven glukoze</i>
43	99
21	65
25	79
42	75
57	87
59	81

- a) Izračunajte Pearsonov koeficient korelacije. Ali podatki nakazujejo linearno povezanost med ravno glukoze in starostjo posameznikov? Če je odgovor pritrديل, kakšno moč in kakšno smer povezanosti nakazujejo podatki?

(12T)

	<b>Starost</b>	<b>Raven glukoze</b>	$x_i - \mu_X$	$y_i - \mu_Y$	$(x_i - \mu_X) \cdot (y_i - \mu_Y)$	$(x_i - \mu_X)^2$	$(y_i - \mu_Y)^2$
1	43	99	1,83	18	33	3,36	324
2	21	65	-20,17	-16	322,67	406,69	256
3	25	79	-16,17	-2	32,33	261,36	4
4	42	75	0,83	-6	-5	0,69	36
5	57	87	15,83	6	95	250,69	36
6	59	81	17,83	0	0	318,03	0
<b>Σ</b>	<b>247</b>	<b>486</b>			<b>478</b>	<b>1240,83</b>	<b>656</b>

$$\mu_X = \frac{1}{N} \sum_{i=1}^N x_i = \frac{247}{6} = 41.17 \quad \mu_Y = \frac{1}{N} \sum_{i=1}^N y_i = \frac{486}{6} = 81$$

$$\rho_{XY} = \frac{C_{XY}}{\sigma_X \sigma_Y} = \frac{\sum_{i=1}^N (x_i - \mu_X)(y_i - \mu_Y)}{\sqrt{\sum_{i=1}^N (x_i - \mu_X)^2 \cdot \sum_{i=1}^N (y_i - \mu_Y)^2}} = \frac{478}{\sqrt{1240.83 \cdot 656}} = 0.5298$$

**Pearsonov koeficient korelacije:**  $\rho = \underline{0.5298}$

**Odgovor:** Spremenljivki sta močno pozitivno povezani.

- b) Če podatki nakazujejo linearno povezanost, izračunajte še regresijsko premico odvisnosti ravni glukoze od starosti posameznika.

(4T)

$$a = \mu_Y - \frac{C_{XY}}{\sigma_X^2} \mu_X = 81 - \frac{478}{1240.83} \cdot 41.17 = 65.14$$

$$b = \frac{C_{XY}}{\sigma_X^2} = \frac{478}{1240.83} = 0.385$$

**Enačba regresijske premice odvisnosti:** Raven Glukoze = 65.14 + 0.385 \* Starost