Sveučilište u Zagrebu
PMF-Matematički odjel
Poslijediplomski stručni studij aktuarske matematike

ISPIT

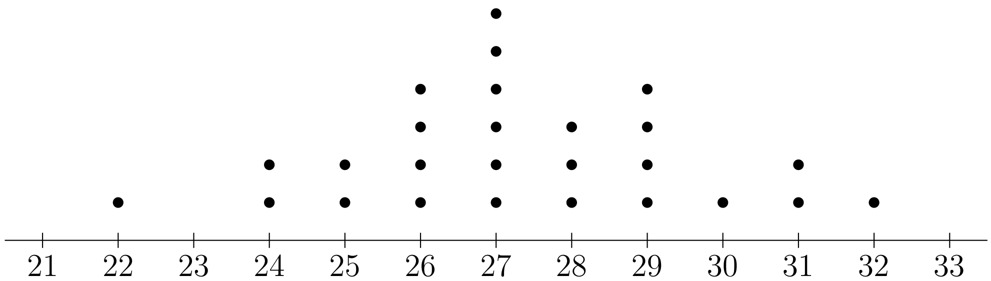
VJEROJATNOST I MATEMATIČKA STATISTIKA

17. 2. 2003.

Vrijeme trajanja ispita: 120 minuta
Ukupan broj bodova: 100
Broj zadataka: 11
Naznačeno je koliko bodova donosi svaki točan odgovor.
Odgovore i rješenja zadataka pišite na ovjerenim papirima. Na svakom korištenom papiru naznačite na koji se zadatak odnosi i čitko se potpišite.
Dozvoljeno je korištenje džepnog kalkulatora, statističkih tablica i *Formulae and Tables for Actuarial Examinations* (Institute of Actuaries).

---

1. U dijagramu točaka



prikazani su podaci dobiveni mjerenjem vremena izvođenja određene radne operacije (u sekundama). Za taj skup podataka:

(a) Izračunajte aritmetičku sredinu i varijancu. (2 boda)

(b) Izračunajte medijan i interkvartil. (3 boda)

(c) Skicirajte dijagram pravokutnika (*box and whisker*). (1 bod)

(ukupno 6 bodova)

2. Pretpostavimo da se štete po policama autoodgovornosti (u portfelju takvih polica nekog osiguravajućeg društva) pojavljuju na sljedeći način. U nekom se vremenskom trenutku može dogoditi najviše jedna šteta, a njihov intenzitet pojavljivanja je jedna šteta na dan. Nadalje, brojevi šteta iz dva vremenska razdoblja koji se ne preklapaju, su nezavisna. Neka X označava slučajan broj šteta u sljedeća tri dana po danom portfelju polica od autoodgovornosti.

(a) Postavite odgovarajući model za razdiobu od X. (2 boda)

(b) Izračunajte vjerojatnost da će u sljedeća tri dana biti barem tri štete u danom portfelju polica osiguranja. (2 boda)

(c) Koliki je očekivani (srednji) broj šteta u sljedećih tri dana u danom portfelju polica osiguranja. (1 bod)

(ukupno 5 bodova)

3. Osigurana svota polica osiguranja određenog tipa modelira se kao razdioba s matematičkim očekivanjem jednakim 9000 kn i standardnom devijacijom jednakom 3000 kn. Na slučajan način odabran je uzorak od 100 takvih polica osiguranja. Izračunajte aproksimativnu vjerojatnost da je ukupan zbroj osiguranih svota u tom uzorku veći od 945000 kn. (3 boda)

4. X je neprekidna slučajna varijabla s uniformnom razdiobom na intervalu $[0, 1]$.

(a) Izračunajte funkciju izvodnicu momenata varijable $Y = -\log X$. (3 boda)

(b) Koju razdiobu ima Y? (2 boda)

(ukupno 5 bodova)

5. Neka je su X i Y dvije Poissonove slučajne varijable, $X \sim P(\mu)$, $Y \sim P(\nu)$. Pomoću funkcija izvodnica vjerojatnosti pokažite da njihov zbroj, $Z = X+Y$, ima također Poissonovu razdiobu. S kojim parametrom? (5 bodova)

6. Broj odlazaka aktuara s posla nakon redovnog radnog vremena tijekom radnog tjedna modelira se kao binomna slučajna varijabla X s parametrima (n, θ) gdje je $n = 5$, a $\theta = \frac{4}{5}$. Za uvjetnu razdiobu ukupnog vremena Y koje je aktuar proveo na poslu tijekom tjedna (u satima) ako je taj tjedan morao na poslu ostati dulje x dana, vrijedi:

$$\mathbb{E}[Y|X = x] = 4(x + 10), \quad \mathbb{V}\text{ar}[Y|X = x] = x.$$

(a) Koliko iznosi matematičko očekivanje $\mathbb{E}[Y]$ ukupnog vremena koje aktuar provodi na poslu tijekom tjedna? (2 boda)

(b) Izračunajte bezuvjetnu varijancu od Y (tj. varijancu $\mathbb{V}\text{ar}[Y]$ marginalne razdiobe od Y). (3 boda)

(ukupno 5 bodova)

7. Standardna devijacija skupa podataka iznosa prošlogodišnjih šteta po policama osiguranja određenog tipa iznosi 78 kn. Iznos štete po polici osiguranja istoga tipa za štete koje se pojavljuju u ovoj godini modeliraramo kao normalnu razdiobu sa standardnom devijacijom jednakoj prošlogodišnjim štetama, a parametar očekivanja želimo procijeniti iz na slučajan način odabranoga uzorka ovogodišnjih šteta do na točnost od ± 10 kn. Koliko veliki uzorak (najmanje) trebamo uzeti da bi parametar očekivanja procijenili sa zadanom točnošću uz pouzdanost od 95%? (3 boda)

8. U svrhu usporedbe iznosa premija osiguranja kućanstava koje naplaćuju dva osiguravajuća društva A i B, na slučajan način i nezavisno jedan od drugoga, odabrana su dva uzorka od po pet polica tog tipa iz svakog od navedenih društva. Opaženi iznosi premija su:

društvo A:	175	155	162	186	148
društvo B:	152	141	129	120	115

Pretpostavljamo da su iznosi premija normalno distribuirani s istim varijancama: $N(\mu_A, \sigma^2)$ za društvo A i $N(\mu_B, \sigma^2)$ za društvo B.

(a) Izračunajte uzoračke sredine i varijance svakog od navedenih uzoraka posebno, te, na osnovi oba uzorka zajedno, procijenite zajedničku varijancu σ^2. (3 boda)

(b) Konstruirajte i izračunajte opaženi 95% pouzdani interval za razliku parametara očekivanja $\mu_A - \mu_B$. Izrazite ga u formi granica. (3 boda)

(c) Koliko iznosi p-vrijednost za jednostrani test

$$H_0 :\ \mu_A = \mu_B, \quad H_1 :\ \mu_A > \mu_B?$$

Zaključite je li opažena uzoračka sredina iznosa premija osiguranja kućanstava društva A značajno veća od odgovarajuće uzoračke sredine za društvo B, uz razinu značajnosti od 5%. (4 boda)

(ukupno 10 bodova)

9. Jednostavan model za kretanje cijena dionica je sljedeći. Na kraju svakog proteklog jediničnog vremenskog razdoblja, neovisno o ostalim vremenskim razdobljima, cijena dionice poraste za jednu novčanu jedinicu s vjerojatnosti $\frac{1}{4}-\theta$, ostaje nepromijenjena s vjerojatnosti $\frac{5}{8}+2\theta$, ili padne za jednu novčanu jedinicu s vjerojatnosti $\frac{1}{8}-\theta$.

(a) Odredite skup dopuštenih vrijednosti za parametar θ. (1 bod)

(b) Promatra se jedna dionica tijekom 80 uzastopnih jediničnih vremenskih razdoblja. Opaženo je sljedeće:

promjena u cijeni za:	1	0	-1
broj vremenskih razdoblja:	24	35	21

Na osnovi tih podataka procijenite nepoznati parametar θ metodom najveće vjerodostojnosti. (4 boda)

(c) Sprovedite χ^2-test prilagodbe pretpostavljenog modela za kretanje cijena dionica podacima iz (b) dijela zadatka (na osnovi dobivene p-vrijednosti donesite zaključak). (6 bodova)

(ukupno 11 bodova)

10. Na slučajan način i nezavisno jedan od drugih, uzeta su tri slučajna uzorka polica osiguranja istog tipa triju osiguravajućih društava. Zabilježene su vrijednosti osiguranih svota po tim policama. Zbirni rezultati su dani u tablici:

društvo i	veličina uzorka n_i	uzoračka sredina $\bar{y}_i$	uzoračka varijanca s_i^2
1	15	115.13	196.41
2	10	95.00	341.33
3	12	135.42	609.36

Pretpostavlja se sljedeći model za dobivene podatke:

$$Y_{ij} = \mu + \tau_i + \varepsilon_{ij}, \quad j = 1, 2, \ldots, n_i, \quad i = 1, 2, 3,$$

gdje je Y_{ij} osigurana svota po j-toj polici u uzorku iz društva i, za parametre τ_1, τ_2 i τ_3 vrijedi $\sum_{i=1}^{3} n_i\tau_i = 0$, a ε_{ij} su slučajne greške za koje vrijedi da su međusobno nezavisne i normalno $N(0, \sigma^2)$-distribuirane.

(a) Procijenite parametre μ, τ_1, τ_2, τ_3 i σ^2. (3 boda)

(b) Je li procjenitelj metodom maksimalne vjerodostojnosti za σ^2 nepristran? Objasnite. (5 bodova)

(c) Ispišite ANOVA tablicu i sprovedite test nulhipoteze da nema razlike u distribuciji iznosa osiguranih svota između tri navedena osiguravajuća društva. Je li razlika između njihovih srednjih vrijednosti značajna uz razinu značajnosti od 5%? (12 bodova)

(ukupno 20 bodova)

11. Za zadanih 12 vrijednosti x_1, x_2,..., x_{12} (nezavisne) varijable poticaja X izmjerene su pripadne vrijednosti y_1, y_2,..., y_{12} (zavisne) varijable odziva Y. Na taj način dobiven je uzorak (x_i, y_i), $i = 1, 2, \ldots, 12$ za koji vrijedi

$$\sum_{i=1}^{12} x_i = 516.4, \qquad \sum_{i=1}^{12} x_i^2 = 22741.34$$
$$\sum_{i=1}^{12} y_i = 14821, \qquad \sum_{i=1}^{12} y_i^2 = 18695125$$

$$\sum_{i=1}^{12} x_i y_i = 650264.8.$$

(a) Uz pretpostavku da je model za vezu varijabli poticaja X i odziva Y jednostavni linearni regresijski model, procijenite pravac regresije. (5 bodova)

(b) Konstruirajte i procijenite 95%-pouzdan interval za koeficijent smjera regresijskog pravca. (4 boda)

(c) Testirajte nulhipotezu da je koeficijent smjera jednak nuli u odnosu na alternativu da to nije tako, uz razinu značajnosti od 5%. (2 boda)

(d) Konstruirajte i procijenite 95%-pouzdan interval za očekivanu (srednju) vrijednost varijable Y ako je $X = 50$. (2 boda)

(e) Opišite linearni regresijski model i navedite koje ste sve pretpostavke na njega koristili u zadacima $(a - d)$. (4 boda)

(ukupno 17 bodova)

Sveučilište u Zagrebu
PMF-Matematički odjel
Poslijediplomski stručni studij aktuarske matematike

ISPIT

VJEROJATNOST I MATEMATIČKA STATISTIKA

17. 2. 2003.

RJEŠENJA ZADATAKA

1. Iz prve frekvencijske tablice

a_j	f_j	$f_j a_j$	$f_j a_j^2$
22	1	22	484
23	0	0	0
24	2	48	1142
25	2	50	1250
26	4	104	2704
27	6	162	4374
28	3	84	2352
29	4	116	3364
30	1	30	900
31	2	62	1920
32	1	32	1024
Σ	26	710	19526

d_j	f_j	$f_j d_j$	$f_j d_j^2$
-5	1	-5	25
-4	0	0	0
-3	2	-6	18
-2	2	-4	8
-1	4	-4	4
0	6	0	0
1	3	3	3
2	4	8	16
3	1	3	9
4	2	8	32
5	1	5	25
Σ	26	8	140

dobijamo:

(a) Aritmetička sredina i varijanca:

$$\begin{aligned} \bar{x} &= \frac{\sum_j f_j a_j}{n} = && (1) \\ &= \frac{710}{26} = 27.31 && (\tfrac{1}{2}) \\ s^2 &= \frac{1}{n-1}\sum_j f_j a_j^2 - \frac{n}{n-1}\bar{x}^2 = && (1) \\ &= \frac{19526}{25} - \frac{710^2}{25\cdot 26} = \frac{1788}{325} = 5.5015 && (\tfrac{1}{2}) \end{aligned}$$

Napomena: $\bar{x}$ i s^2 se mogu izračunati jednostavno na sljedeći način. Uočimo mod $a_0 = 27$. Svaki od podataka prikažimo u obliku $a_j = a_0 + d_j$. Ako sa $\bar{d}$ označimo aritmetičku sredinu, a sa s_d^2 varijancu izvedenih podataka d_j (vidjeti frekvencijsku tablicu gore desno), tada se odmah vidi da vrijede relacije

$$\bar{x} = a_0 + \bar{d}, \quad s^2 = s_d^2.$$

Iz pripadne frekvencijske tablice i računa koji je u njoj sproveden vidi se da je pomoćne vrijednosti $\bar{d}$ i s_d^2, a time i $\bar{x}$ i s^2, lakše računati na taj način nego direktnim pristupom.

(b) Medijan:

$$\begin{aligned} m &= x_{(\frac{n+1}{2})} = x_{(\frac{27}{2})} = x_{(13+\frac{1}{2})} = \frac{1}{2}(x_{(13)} + x_{(14)}) = && (1) \\ &= \frac{1}{2}(27+27) = 27 && (\tfrac{1}{2}) \end{aligned}$$

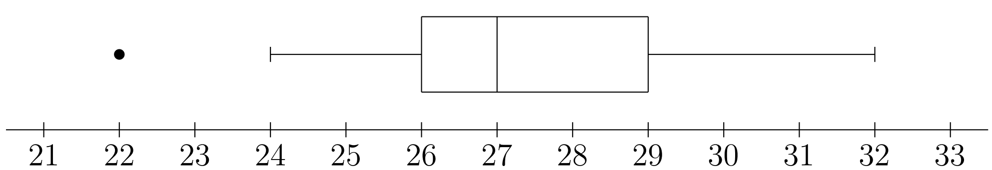
Donji i gornji kvartil:

$$\begin{aligned} q_L &= x_{(\frac{n+1}{4})} = x_{(\frac{27}{4})} = & q_U &= x_{(\frac{3(n+1)}{4})} = x_{(\frac{81}{4})} = & \\ &= \tfrac{1}{4}x_{(6)} + \tfrac{3}{4}x_{(7)} = & &= \tfrac{3}{4}x_{(20)} + \tfrac{1}{4}x_{(21)} = & (\tfrac{1}{2}+\tfrac{1}{2}) \\ &= \tfrac{1}{4}\cdot 26 + \tfrac{3}{4}\cdot 26 = 26 & &= \tfrac{3}{4}\cdot 29 + \tfrac{1}{4}\cdot 29 = 29 & (\tfrac{1}{2}+\tfrac{1}{2}) \end{aligned}$$

Interkvartil:

$$\begin{aligned} IQR &= q_U - q_L = & (1) \\ &= 29 - 26 = 3 & (\tfrac{1}{2}) \end{aligned}$$

(c) Dijagram pravokutnika (s “protezanjem” na obje strane od pravokutnika za 1 IQR):



(1)

(9)

2. (a) X ima Poissonovu razdiobu (1)
s parametrom $\lambda = 3$. (1)

(b) Vjerojatnost da će biti barem tri štete:

$$\begin{aligned} & \mathbb{P}(X \geq 3) = & (\tfrac{1}{2}) \\ = \; & 1 - \mathbb{P}(X \leq 2) = 1 - e^{-3}(1 + 3 + \tfrac{3^2}{2}) = & (1) \\ = \; & 1 - \tfrac{17}{2}\cdot e^{-3} = 0.578. & (\tfrac{1}{2}) \end{aligned}$$

(c) Srednji broj šteta je: $\mathbb{E}[X] = \lambda = 3$ (1)

(5)

3. Neka je S ukupan (slučajan) zbroj osiguranih svota u portfelju. Tada je $\mathbb{E}[S] = 100 \cdot 9000 = 900000$ i $\sigma[S] = 10 \cdot 3000 = 30000$.

$$
\begin{aligned}
& \mathbb{P}(S > 945000) = & (1)\\
= \quad & \mathbb{P}\left(\tfrac{S-\mathbb{E}[S]}{\sigma(S)} > \tfrac{945000-900000}{30000}\right) = \mathbb{P}(Z > 1.5) = & (1)\\
\overset{\text{CGT}}{\approx} \quad & 1 - \Phi(1.5) = & (1)\\
= \quad & (\text{tablice}) = 1 - 0.9332 = 0.0668. & (1)
\end{aligned}
$$

(4)

4. (a) Funkcija izvodnica momenata od Y:

$$
\begin{aligned}
& M_Y(t) = \mathbb{E}[e^{tY}] = & (1)\\
= \quad & \mathbb{E}[e^{-t\log X}] = \mathbb{E}[X^{-t}] = \int_0^1 x^{-t}\,dx = & (1)\\
= \quad & \tfrac{1}{1-t} x^{1-t} \Big|_0^1 = \tfrac{1}{1-t} \ \text{ za } \ t < 1. & (1)
\end{aligned}
$$

(b) Budući da je $M_Y(t)$ iz (a) f.i.m. eksponencijalne razdiobe s parametrom $\lambda = 1$, $Y \sim Exp(1)$. (2)

(5)

Napomena: Zadatak se može riješiti i tako da se prvo izračuna gustoća od Y kao gustoća funkcije slučajne varijable (2 boda), a zatim se izračuna f.i.m. ili se napiše pozivanjem na tablične formule (3 boda).

5. F.i.v. od X i Y: $G_X(t) = e^{\mu(t-1)}$, $G_Y(t) = e^{\nu(t-1)}$ (1)

$$
\begin{aligned}
& G_Z(t) = (\text{nezavisnost}) = G_X(t) \cdot G_Y(t) = & (1)\\
= \quad & e^{\mu(t-1)} \cdot e^{\nu(t-1)} = e^{(\mu+\nu)(t-1)}. & (1)
\end{aligned}
$$

Dobivena f.i.v. je također f.i.v. Poissonove razdiobe.
Dakle, Z ima Poissonovu razdiobu (1)
s parametrom $\mu + \nu$. (1)

(5)

6. $\mathbb{E}[X] = 5 \cdot \frac{4}{5} = 4$, $\mathbb{V}\mathrm{ar}[X] = 5 \cdot \frac{4}{5} \cdot \frac{1}{5} = \frac{4}{5}$ (1)

(a) Očekivanje od Y:

$$\begin{aligned} \mathbb{E}[Y] &= \mathbb{E}[\mathbb{E}[Y|X]] = && (1) \\ &= \mathbb{E}[4(X+10)] = 4(\mathbb{E}[X]+10) = 4 \cdot (4+10) = 56. && (1) \end{aligned}$$

(b) Varijanca od Y:

$$\begin{aligned} \mathbb{V}\mathrm{ar}[Y] &= \mathbb{E}[\mathbb{V}\mathrm{ar}[Y|X]] + \mathbb{V}\mathrm{ar}[\mathbb{E}[Y|X]] = && (1) \\ &= \mathbb{E}[X] + \mathbb{V}\mathrm{ar}[4(X+10)] = \mathbb{E}[X] + 16\mathbb{V}\mathrm{ar}[X] = && (1) \\ &= 4 + 16 \cdot \tfrac{4}{5} = \tfrac{84}{5} = 16.8. && (1) \end{aligned}$$

(6)

7. 95% pouzdan interval za μ: $\overline{X} \pm 1.96\frac{\sigma}{\sqrt{n}} \approx \overline{X} \pm 1.96 \cdot \frac{78}{\sqrt{n}}$. (1)

$$\begin{aligned} & 1.96 \cdot \tfrac{78}{\sqrt{n}} \leq 10 \quad \Leftrightarrow && (1) \\ \Leftrightarrow \quad & (7.8 \cdot 1.96)^2 \leq n \quad \Leftrightarrow \quad n \geq 233.7 \quad \Rightarrow \quad n \geq 234. && (1) \end{aligned}$$

(3)

8. (a) Procjena zajedničke varijance:

$$\begin{aligned} s_p^2 &= \tfrac{(n_A-1)s_A^2+(n_B-1)s_B^2}{n_A+n_B-2} = && (1) \\ &= \tfrac{4 \cdot 234.7 + 4 \cdot 230.3}{5+5-2} = 232.5. && (1) \end{aligned}$$

(b) 95% pouzdani interval za $\mu_A - \mu_B$: $\overline{X}_A - \overline{X}_B \pm t_{0.025}(8) S_p \sqrt{\frac{2}{5}}$. (1)
Iz tablice za t-razdiobu: $t_{0.025}(8) = 2.306$. (1)
Dakle, opaženi 95% pouzdani interval je:
$165.2 - 131.4 \pm 2.306 \cdot 15.25 \cdot \sqrt{\frac{2}{5}} = 33.8 \pm 22.2$. (1)

(c) Testna statistika: $T = (\overline{X}_A - \overline{X}_B)/(S_p\sqrt{\frac{2}{5}}) \overset{H_0}{\sim} t(8)$. (1)
Opažena vrijednost: $t = (165.2 - 131.4)/(15.25 \cdot \sqrt{\frac{2}{5}}) = 3.5$. (1)
p-vrijednost: $\mathbb{P}(T > 3.5 | H_0) < 0.005$. (1+1)
Budući da je p-vrijednost manja od 0.01, srednja vrijednost od A je značajno veća od srednje vrijednosti od B uz razinu značajnosti od 1%. (1)

(10)

9. X=cijena dionice na kraju jediničnog razdoblja$=\begin{pmatrix} -1 & 0 & 1 \\ \frac{1}{8}-\theta & \frac{5}{8}+2\theta & \frac{1}{4}-\theta \end{pmatrix}$.

(a) Skup dopuštenih vrijednosti za parametar θ:

$$\begin{array}{lll} & 0 < \frac{1}{8}-\theta < 1 \ \ \& \ \ 0 < \frac{5}{8}+2\theta < 1 \ \ \& \ \ 0 < \frac{1}{4}-\theta < 1 & \\ \Leftrightarrow & -\frac{14}{16} < \theta < \frac{2}{16} \ \ \& \ \ -\frac{5}{16} < \theta < \frac{3}{16} \ \ \& \ \ -\frac{12}{16} < \theta < \frac{4}{16} & \\ \Leftrightarrow & -\frac{5}{16} < \theta < \frac{1}{8} \ \ \Leftrightarrow \ \ \theta \in \langle -0.3125, 0.125 \rangle & (1) \end{array}$$

(b) Vjerodostojnost od θ:

$$L(\theta) = (\frac{1}{8}-\theta)^{21} \cdot (\frac{5}{8}+2\theta)^{35} \cdot (\frac{1}{4}-\theta)^{24}. \qquad (1)$$

Im$X = \{-1, 0, 1\}$ ne ovisi o θ pa MLE tražimo kao rješenje stacionarne jednadžbe $\ell'(\theta) = 0$ log-vjerodostojnosti $(\frac{1}{2})$

$$\begin{array}{lll} & \ell(\theta) = 21\log(\frac{1}{8}-\theta) + 35\log(\frac{5}{8}+2\theta) + 24\log(\frac{1}{4}-\theta). & (\frac{1}{2}) \\ & \ell'(\theta) = -\frac{21\cdot 8}{1-8\theta} + \frac{35\cdot 8\cdot 2}{5+16\theta} - \frac{24\cdot 4}{1-4\theta} = 0 & \\ \Leftrightarrow & 5120\theta^2 - 468\theta - 95 = 0 & (1) \\ \Leftrightarrow & \theta_1 = 0.1894, \quad \theta_2 = -0.09798. & (1) \end{array}$$

Budući da je jedino θ_2 u skupu iz (a), MLE je $\hat{\theta} = -0.09798$. (1)

(c) Tablica:

a_i	p_i	f_i	e_i	$(f_i - e_i)^2/e_i$
-1	0.22298	21	17.84	0.5597
0	0.42905	35	34.32	0.0135
1	0.34798	24	27.84	0.5297
Σ	0.99991	80	80.00	1.1029

(1+1)

Broj stupnjeva slobode: $3 - 1 - 1 = 1$. (1)
Testna statistika: $H \overset{H_0}{\sim} \chi^2(1)$, opažena vrijednost: $h = 1.1$. (1)
p-vrijednost: $\mathbb{P}(H > 1.1 | H_0) > 0.3$. (1)
Nemamo dokaza u korist odbacivanja H_0, dakle, model je vrlo dobro prilagođen. (1)

(12)

10. (*a*) $n = n_1 + n_2 + n_3 = 37,\ k = 3$
Procjene parametara:

$$\begin{aligned} \hat{\mu} &= \bar{y} = \tfrac{n_1\bar{y}_1 + n_2\bar{y}_2 + n_3\bar{y}_3}{n} = & (1)\\ &= \tfrac{15\cdot 115.13 + 10\cdot 95.00 + 12\cdot 135.42}{37} = 116.27. & (1) \end{aligned}$$

$$\begin{aligned} \hat{\tau}_i &= \bar{y}_i - \bar{y}, \quad i = 1,2,3 & & & (1)\\ \hat{\tau}_1 &= 115.13 - 116.27 &=& -1.14 &\\ \hat{\tau}_2 &= 95.00 - 116.27 &=& -21.27 &\\ \hat{\tau}_3 &= 135.42 - 116.27 &=& 19.15 & (1) \end{aligned}$$

$$\begin{aligned} \mathrm{SSE} &= (n_1 - 1)s_1^2 + (n_2 - 1)s_2^2 + (n_3 - 1)s_3^2 = & (1)\\ &= 14 \cdot 196.41 + 9 \cdot 341.33 + 11 \cdot 609.36 = 12524.67 & (1)\\ \hat{\sigma}^2 &= \tfrac{\mathrm{SSE}}{n-k} = & (1)\\ &= \tfrac{1254.67}{37-3} = 368.37. & (1) \end{aligned}$$

(*b*) Za statistiku SSE vrijedi $\mathrm{SSE}/\sigma^2 \sim \chi^2(n-k)$. (1)
Dakle,

$$\mathbb{E}[\tilde{\sigma}^2] = \mathbb{E}[\frac{\mathrm{SSE}}{n}] = \frac{\sigma^2}{n}\mathbb{E}[\frac{\mathrm{SSE}}{\sigma^2}] = \frac{n-k}{n}\sigma^2. \qquad (1)$$

Ako bi $\tilde{\sigma}^2$ bila nepristrani procjenitelj za σ^2, trebalo bi vrijediti $\mathbb{E}[\tilde{\sigma}^2] = \sigma^2$ što nije slučaj. (1)

(*c*) ANOVA-tablica:

izvor varijabilnosti	stupnjevi slobode	sume kvadrata	srednji kvadrati	testna statist.
društva	2	8944.29	4472.15	12.14
greške	34	2524.67	68.7	—
ukupno	36	1470.52	—	—

(1)

$$\begin{aligned} \mathrm{SST} &= n_1(\bar{y}_1 - \bar{y})^2 + n_2(\bar{y}_2 - \bar{y})^2 + n_3(\bar{y}_3 - \bar{y})^2 = & (1)\\ &= 15 \cdot (-1.14)^2 + 10 \cdot (-21.27)^2 + 12 \cdot 19.15^2 = 8944.293 & (1)\\ \mathrm{MST} &= \tfrac{\mathrm{SST}}{k-1} = \tfrac{8944.29}{3-1} = 442.1465 & (1+1)\\ \mathrm{MSE} &= \hat{\sigma}^2 = 3.37 & (1)\\ f &= \tfrac{\mathrm{MST}}{\mathrm{MSE}} = \tfrac{4472.15}{68.37} = 12.14. & (1+1) \end{aligned}$$

Za test nulhipoteze $H_0 : \tau_1 = \tau_2 = \tau_3 = 0$,
testna statistika je: $F = \frac{\mathrm{MST}}{\mathrm{MSE}} \overset{H_0}{\sim} F(2, 34)$. (1)
p-vrijednost: $\mathbb{P}(F > 12.14 | H_0) < 0.01$. (1)
Populacijske srednje vrijednosti se značajno razlikuju uz razinu značajnosti od 5%. (1)

(22)

11. (*a*) Opažene vrijednosti pomoćnih statistika:

$$\begin{aligned} \bar{x} &= \tfrac{516.4}{12} = 43.3 & \bar{y} &= \tfrac{1481}{12} = 1235.083 & (\tfrac{1}{2} + \tfrac{1}{2}) \\ S_{xx} &= \textstyle\sum_{i=1}^{12} x_i^2 - 12 \cdot \bar{x}^2 = & S_{yy} &= \textstyle\sum_{i=1}^{12} y_i^2 - 12 \cdot \bar{y}^2 = & (\tfrac{1}{2}) \\ &= 522.37 & &= 389965 & (\tfrac{1}{2}) \\ S_{xy} &= \textstyle\sum_{i=1}^{12} x_i y_i - 12 \cdot \bar{x}\bar{y} = & & & (\tfrac{1}{2}) \\ &= 12517.35 & & & (\tfrac{1}{2}) \end{aligned}$$

Parametri pravca:

$$\begin{aligned} \hat{\alpha} &= \tfrac{S_{xy}}{S_{xx}} = & \hat{\beta} &= \bar{y} - \hat{\alpha}\bar{x} = & (1 + 1) \\ &= 23.962612 & &= 203.9719 & (\tfrac{1}{2} + \tfrac{1}{2}) \end{aligned}$$

Regresijski pravac: $y = 23.963x + 203.97$ (1)

(*b*) 95% pouzdani interval za α: $\hat{\alpha} \pm t_{0.025}(10)\hat{\sigma}\sqrt{\frac{1}{S_{xx}}}$. (1)
Iz tablice za t-razdiobu: $t_{0.025}(10) = 2.228$. (1)

$$\begin{aligned} \mathrm{SSE} &= S_{yy} - \hat{\alpha}^2 S_{xx} = 90016.61 & (\tfrac{1}{2} + \tfrac{1}{2}) \\ \hat{\sigma}^2 &= \tfrac{\mathrm{SSE}}{n-2} = \tfrac{90016.61}{10} = 9001.661 \quad \Rightarrow \quad \hat{\sigma} = 94.877 & (\tfrac{1}{2} + \tfrac{1}{2}) \end{aligned}$$

Dakle, opaženi 95% pouzdani interval je:
$23.963 \pm 2.228 \cdot 94.88 \cdot \sqrt{\frac{1}{522.37}} = 23.963 \pm 9.249$. (1)

(*c*) Budući da se nula ne nalazi u 95% poudanom intervalu za α, odbacujemo nulhipotezu $H_0 : \alpha = 0$ u odnosu na dvostranu alternativu uz razinu značajnosti od 5%. (2)

(*d*) Procjena za $\mathbb{E}[Y|50]$: $\hat{\mathbb{E}}[Y|50] = 23.963 \cdot 50 + 203.97 = 1402.10$.
95% pouzdani interval: $\hat{\mathbb{E}}[Y|50] \pm t_{0.025}(10)\hat{\sigma}\sqrt{\frac{1}{12} + \frac{(50-\bar{x})^2}{S_{xx}}}$. (1)
Opaženi 95% pouzdani interval: 1402.10 ± 88.77. (1)

(e) Model je

$$Y_i = \alpha x_i + \beta + \varepsilon_i, \quad i = 1, 2, \ldots, n,$$

gdje su α, β parametri modela, a ε_i, $i = 1, 2, \ldots, n$, su slučajne varijable (slučajne greške) (1)
za koje smo pretpostavili da su:

(i) centrirane ($\mathbb{E}[\varepsilon_i] = 0$ za sve i) $(\frac{1}{2})$

(ii) jednake varijance ($\mathbb{V}\mathrm{ar}[\varepsilon_i] = \sigma^2$ za sve i) $(\frac{1}{2})$

(iii) nekorelirane ($\mathrm{cov}[\varepsilon_i, \varepsilon_j] = 0$ za sve $i \neq j$) $(\frac{1}{2})$

(iv) normalno distribuirane. $(\frac{1}{2})$

(19)