

REGRESIONA I KORELACIONA ANALIZA

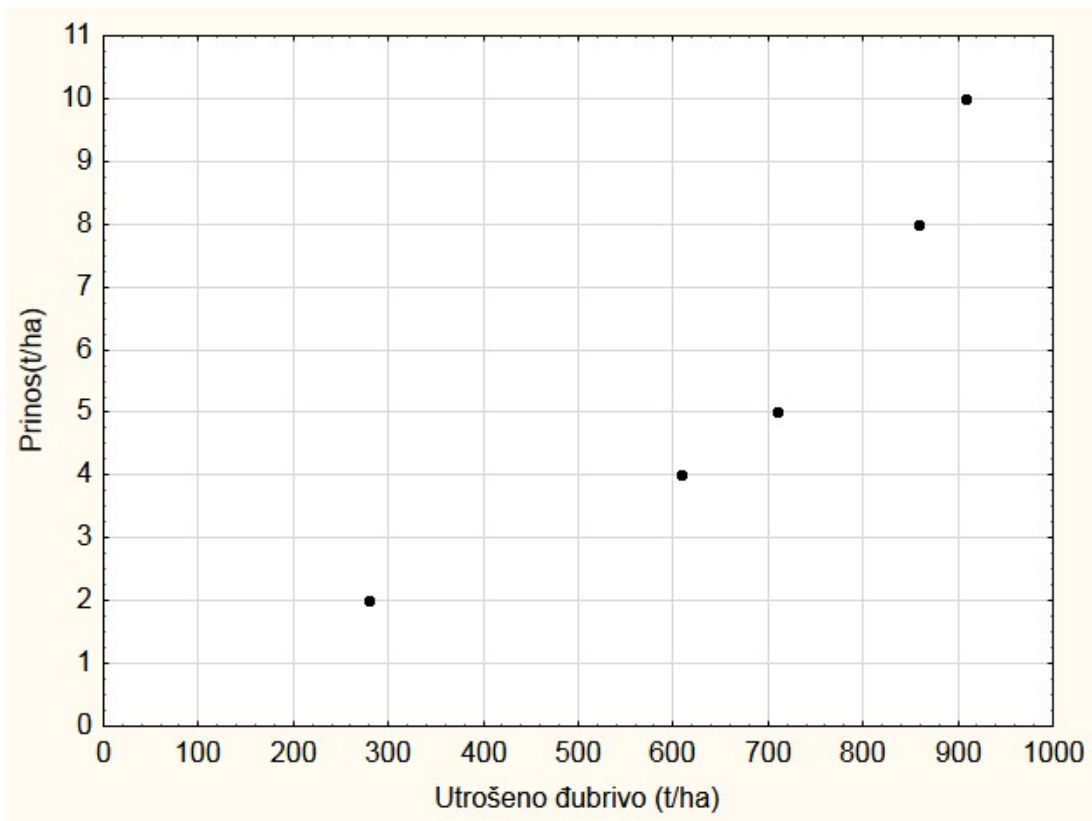
1. Na 5 parcela istog kvaliteta vrednosti prinosa suncokreta (t/ha) i utrošak mineralnog đubriva (kg) date su u tabeli:

Prinosa suncokreta (t/ha)	4	2	5	8	10
Količina đubriva (kg)	609	279	709	859	909

- Formirati dijagram rasturanja,
- Oceniti i grafički predstaviti linearni regresioni model: $\hat{Y} = a + bX$,
- Koliki prinos može da se očekuje ukoliko je utrošena količina mineralnog đubriva 750 (kg) ?
- Izračunati ocenjene vrednosti regresije i standardnu grešku regresije,
- Testirati statističku značajnost linearnog regresionog modela?
- Da li je utrošna količina mineralnog đubriva statistički značajno utiče na prinos?
- Odrediti 95% poverenja za parametar β regresije?
- Izračunati koeficijente korelacije i determinacije, objasniti njihovo značenje i testirati njihovu značajnost.

Rešenje:

a)



b) **I način**

X	Y	$X - \bar{X}$	$Y - \bar{Y}$	$(X - \bar{X})(Y - \bar{Y})$	$(X - \bar{X})^2$	$(Y - \bar{Y})^2$
609	4	-64	-1,8	115,2	4096	3,24
279	2	-394	-3,8	1497,2	155236	14,44
709	5	36	-0,8	-28,8	1296	0,64
859	8	186	2,2	409,2	34596	4,84
909	10	236	4,2	991,2	55696	17,64
3365	29	0	0	2984,0	250920	40,80

$$\bar{X} = \frac{\sum X}{n} = \frac{3365}{5} = 673 \text{ (kg)} \quad \bar{Y} = \frac{\sum Y}{n} = \frac{29}{5} = 5,8 \text{ (t/ha)}$$

$$b = \frac{\sum (X - \bar{X})(Y - \bar{Y})}{\sum (X - \bar{X})^2} = \frac{2984,0}{250920} = 0,011892 \quad a = \bar{Y} - b\bar{X} = 5,8 - 0,011892 \cdot 673$$

$$a = -2,203316$$

Ocenjeni linearni regresioni model je oblika:

$$\hat{Y} = -2,2033 + 0,01189X$$

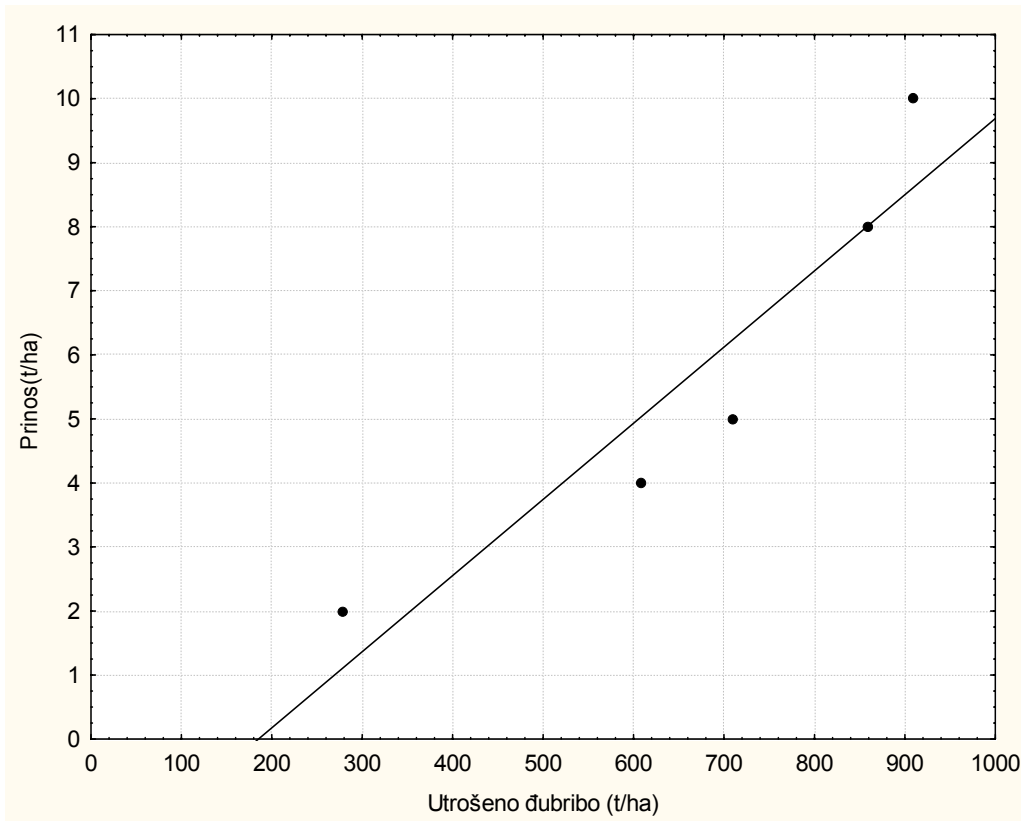
II način

X	Y	X^2	Y^2	XY
609	4	370881	16	2436
279	2	77841	4	558
709	5	502681	25	3545
859	8	737881	64	6872
909	10	826281	100	9090
3365	29	2515565	209	22501

$$b = \frac{\sum XY - \frac{(\sum X \sum Y)}{n}}{\sum X^2 - \frac{(\sum X)^2}{n}} = \frac{22501 - \frac{3365 \cdot 29}{5}}{2515565 - \frac{3365^2}{5}} = \frac{2984}{250920} = 0,011892$$

$$a = \frac{\sum Y(\sum X^2) - \sum X \sum XY}{n(\sum X^2) - (\sum X)^2} = \frac{29 \cdot 2515565 - 3365 \cdot 22501}{5 \cdot 2515565 - 3365^2} = \frac{-2764480}{1254600}$$

$$a = -2,203475$$



c) $\hat{Y}_0 = -2,2033 + 0,01189X_0 = -2,2033 + 0,01189 \cdot 750 = 6,71(\text{t/ha})$

d)

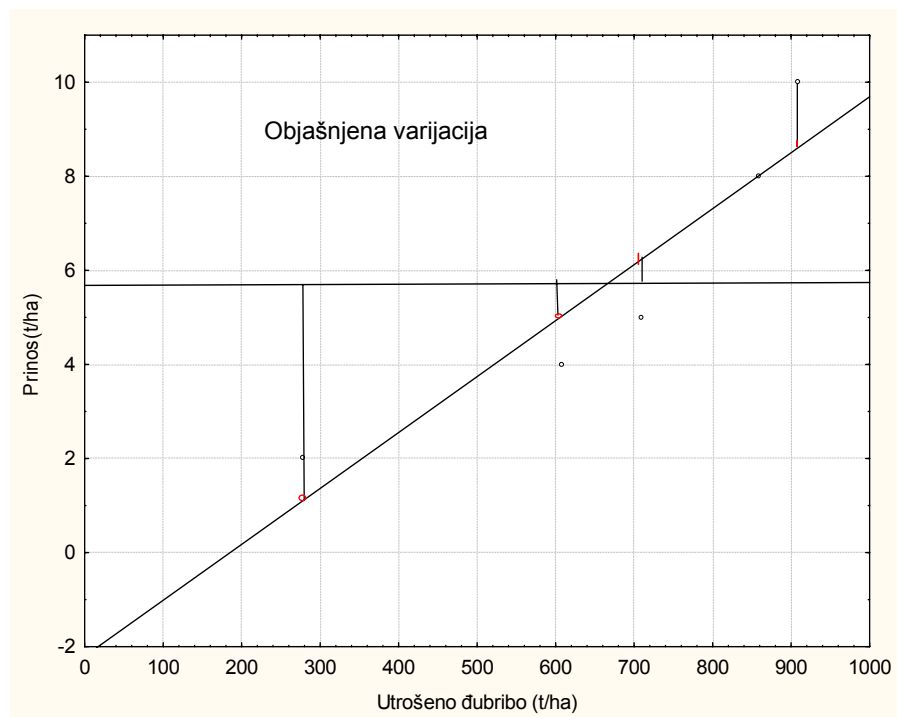
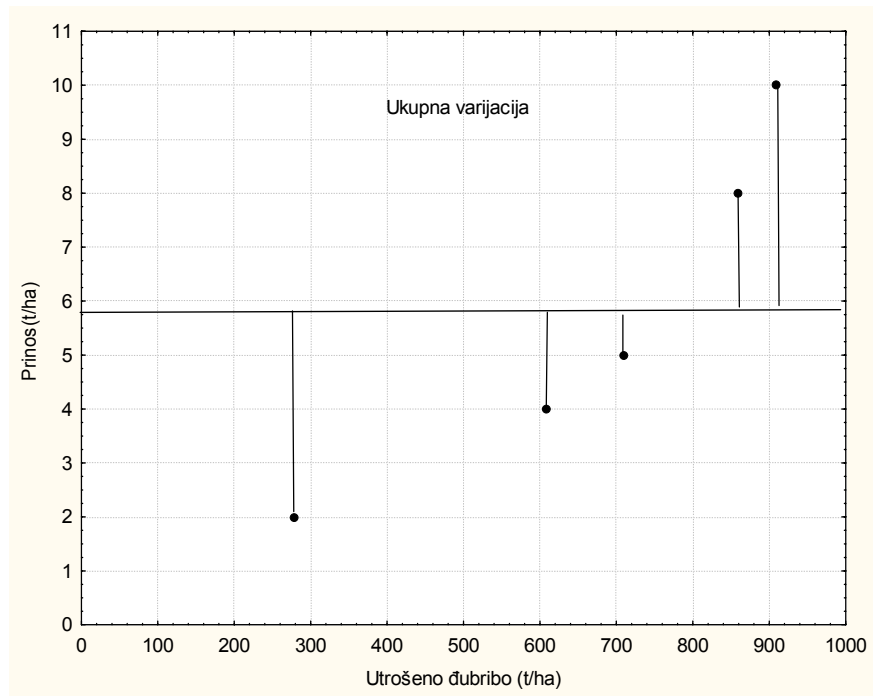
X	Y	\hat{Y}	$Y - \hat{Y}$	$(Y - \hat{Y})^2$
609	4	5,0377	-1,0370	1,0754
279	2	1,1140	0,8860	0,7850
709	5	6,2267	-1,2267	1,5048
859	8	8,0102	0,0102	0,0001
909	10	8,6047	1,3953	1,9469
3365	29	28,9933 \approx 29	0,0067	5,3122

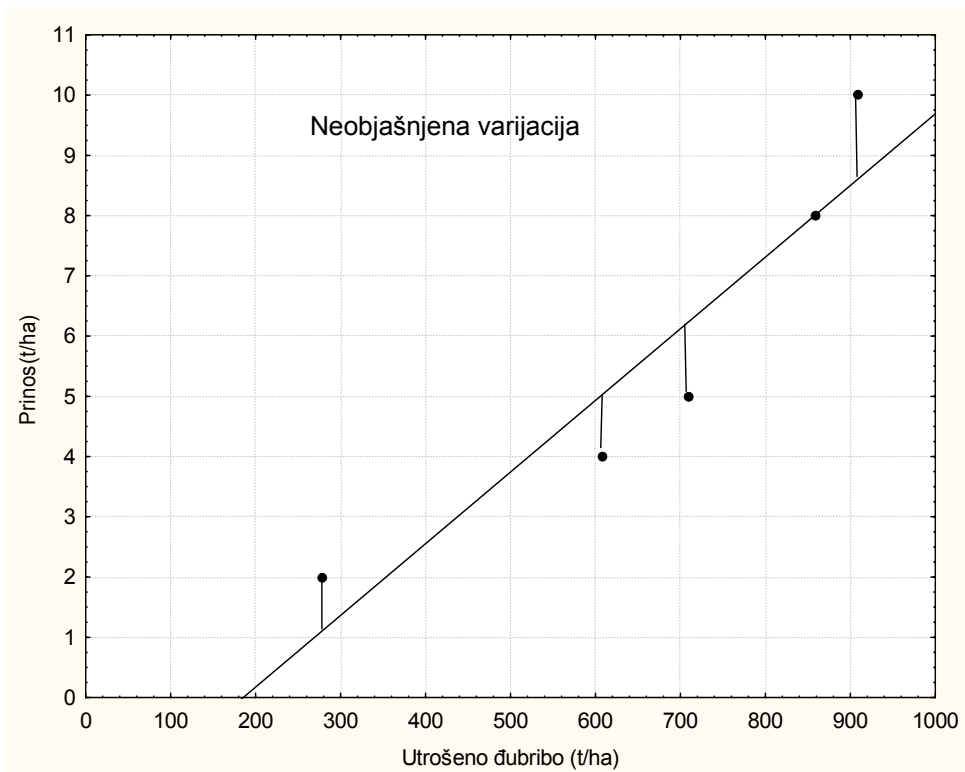
I način
$$s_e = \sqrt{\frac{\sum (Y - \hat{Y})^2}{n-2}} = \sqrt{\frac{5,3122}{3}} = \sqrt{1,77073} = 1,3307(\text{t/ha})$$

II način I

$$s_e = \sqrt{\frac{\sum Y^2 - a \sum Y - b \sum XY}{n-2}} = \sqrt{\frac{209 + 2,2033 \cdot 29 - 0,011892 \cdot 22501}{3}} = 1,3309(\text{t/ha})$$

e) Analiza varijanse





I način

$$Q = \sum (Y - \bar{Y})^2 = 40,8$$

$$Q = \sum Y^2 - \frac{(\sum Y)^2}{n} = 209 - \frac{29^2}{5} = 40,8$$

$$Q_R = \frac{[\sum (X - \bar{X})(Y - \bar{Y})]^2}{\sum (X - \bar{X})^2} = \frac{2984,0^2}{250920} = 35,4864$$

$$Q_R = \frac{\left[\sum XY - \frac{(\sum X \sum Y)}{n} \right]^2}{\sum X^2 - \frac{(\sum X)^2}{n}} = \frac{\left[22501 - \frac{3365 \cdot 29}{5} \right]^2}{2515565 - \frac{3365^2}{5}} = \frac{2984^2}{250920} = 35,4864$$

$$Q_{VR} = Q - Q_R = 40,8 - 35,4864 = 5,3136$$

II način

$$Q_{VR} = \sum (Y - \hat{Y})^2 = 5,3122$$

$$Q_R = Q - Q_{VR} = 40,8 - 5,3122 = 35,4878$$

$$F = \frac{Q_R}{\frac{Q_{VR}}{n-2}} = \frac{35,4864}{\frac{5,3136}{3}} = 20,035^* \quad F_{1,3;0,05} = 10,13 \quad F_{1,3;0,01} = 34,12$$

TABELA ANALIZE VARIJANSE

Izvori varijacije	Sume kvadrata	Sredine suma kvadrata	F	F _{0,05}	F _{0,01}
Regresija	35,4864	35,4864	20,035 *	10,13	34,12
Varijacija oko regresije	5,3136	1,7712			
Ukupno	40,8000				

f) Nulta hipoteza je da utrošena količina mineralnih đubriva statistički značajno ne utiče na prinos tj. $\beta = 0$, dok je alternativna hipoteza da je uticaj đubriva na prinos statistički značajan.

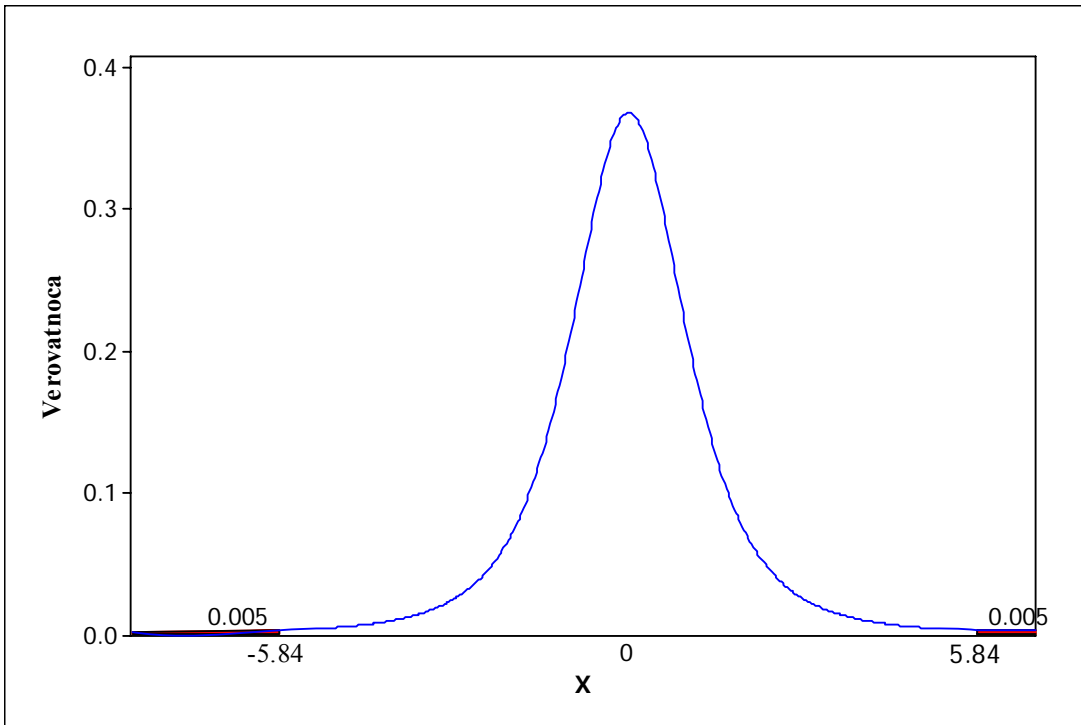
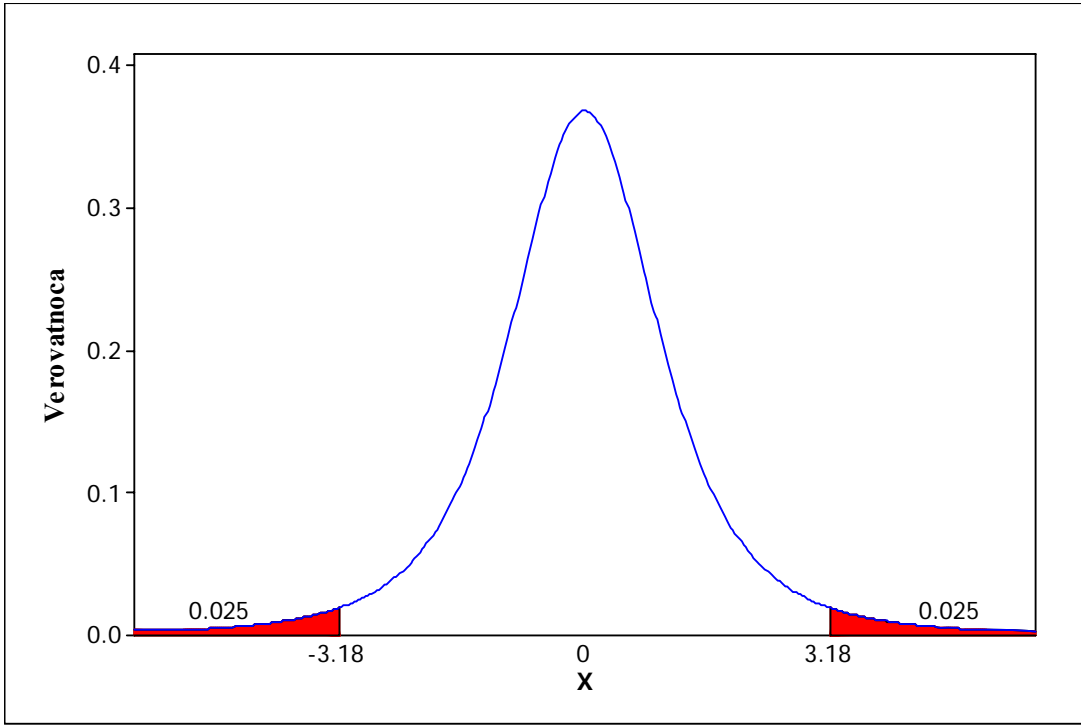
$$\begin{array}{l} H_0 : \beta = 0 \\ H_1 : \beta \neq 0 \end{array}$$

$$t = \frac{b}{s_b} \sim t_{n-2}$$

$$s_b = \sqrt{\frac{s_e^2}{\sum (X - \bar{X})^2}} = \sqrt{\frac{1,77119}{250920}} = 0,002657$$

$$t = \frac{0,01189}{0,002657} = 4,47^* \quad t_{3;0,05} = 3,182 \quad t_{3;0,01} = 5,841$$

Nulta hipoteza se odbacuje tako da može da se zaključi da utrošna količina mineralnog đubriva statistički značajno utiče na prinos.



gg)

$$b - t_{n-2;\alpha} \cdot s_b < \beta < b + t_{n-2;\alpha} \cdot s_b$$

$$0,01189 - 3,182 \cdot 0,002657 < \beta < 0,01189 + 3,182 \cdot 0,002657$$

$$0,003435 < \beta < 0,02034$$

Kako hipotetička vrednost $\beta_0 = 0 \notin (0,003435, 0,02034) \Rightarrow H_1$ tj. koeficijent pravca regresije je statistički značajan.

h)

I način

$$r = \frac{\sum (X - \bar{X})(Y - \bar{Y})}{\sqrt{\sum (X - \bar{X})^2 \sum (Y - \bar{Y})^2}} = \frac{2984}{\sqrt{250920 \cdot 40,80}} = 0,932612$$

II način

$$r = \frac{\sum XY - \frac{(\sum X \sum Y)}{n}}{\sqrt{\left[\sum X^2 - \frac{(\sum X)^2}{n} \right] \cdot \left[\sum Y^2 - \frac{(\sum Y)^2}{n} \right]}} = \frac{22501 - \frac{3365 \cdot 29}{5}}{\sqrt{\left(2515564 - \frac{3365^2}{5} \right) \cdot \left(209 - \frac{29^2}{5} \right)}} =$$

$$= \frac{2984}{\sqrt{250920 \cdot 40,8}} = 0,932612$$

I način

Koeficijent determinacije: $r^2 = 0,86978$ ($r^2 = 86,98\%$)

II način

$$r^2 = \frac{Q_R}{Q} = \frac{35,4864}{40,8} = 0,86978$$

$H_0 : \rho = 0$
$H_1 : \rho \neq 0$

$$t = \frac{r}{s_r} \sim t_{n-2} \quad s_r = \sqrt{\frac{1-r^2}{n-2}}$$

$$t = \frac{0,93261}{\sqrt{1-0,86978}} \cdot \sqrt{3} = 4,48$$

$$\begin{array}{l} H_0 : \rho^2 = 0 \\ H_1 : \rho^2 \neq 0 \end{array}$$

$$F = \frac{Q_R}{Q_{VR}} \cdot (n-2) = \frac{Q_R}{Q - Q_R} \cdot (n-2) = \frac{r^2}{1-r^2} \cdot (n-2)$$

$$F = \frac{0,86978}{1-0,86978} \cdot 3 = 20,04 \quad F_{1,3;0.05} = 10,13 \quad F_{1,3;0.01} = 34,12$$

Kako je izračunata vrednost F statistike veća od prve tablične vrednosti može da se zaključi da se nulta hipoteza odbacuje na nivou značajnosti 5% tj. F statistika je statistički značajna. Nulta hipoteza se ne odbacuje na nivou značajnosti 1%.

