

## ПРИПРЕМА ЗА ИСПИТ – 1. ЗАДАТАК

### Пример 1.

На основу датих података о продаји (у хиљадама тона) и профиту (у милионима динара) једног предузећа:

- оценити једначину регресије и графички је приказати на дијаграму распршености;
- израчунати стандардну грешку регресије помоћу оцењених вредности са линије регресије;
- предвидети са 99% и 95% поузданости колики ће бити остварени профит, уколико је обим продаје 7.000 тона;
- испитати степен линеарне зависности између висине профита и обима продаје;
- израчунати са колико се процената промена у висини профита објашњава утицајем промене у обиму остварене продаје.

**Табела 1.**

Продаја (у 000 т)	Профит (у 000 000 дин.)
2	5
3	7
5	10
6	9
9	14
10	16
12	18

Решење:

- а) Да би се оценио регресиони модел на основу обрасца:

$$\hat{y}_i = a + b \cdot x_i, i = 1, 2, \dots, n$$

прво је потребно израчунати непознате коефицијенте  $a$  и  $b$  преко следећих образаца:

$$b = \frac{\sum x_i y_i - n \bar{x} \bar{y}}{\sum x_i^2 - n \bar{x}^2}$$

$$a = \bar{y} - b\bar{x}$$

Елементи неопходни за израчунавање непознатих коефицијената приказани су у наредној табели:

Табела 2.

Продаја 000 т $x_i$	Профит (у 000 000 дин). $y_i$	$x_i y_i$	$x_i^2$	$y_i^2$
2	5	10	4	25
3	7	21	9	49
5	10	50	25	100
6	9	54	36	81
9	14	126	81	196
10	16	160	100	256
12	18	216	144	324
<b>47</b>	<b>79</b>	<b>637</b>	<b>399</b>	<b>1031</b>

Пошто је у овом примеру  $n = 7$  ( $n$  представља број података) онда је:

$$\bar{x} = \frac{\sum x_i}{n} = \frac{47}{7} = 6,71 \quad ; \quad \bar{y} = \frac{\sum y_i}{n} = \frac{79}{7} = 11,29$$

$$b = \frac{\sum x_i y_i - n \bar{x} \bar{y}}{\sum x_i^2 - n \bar{x}^2} = \frac{637 - 7 \cdot 6,71 \cdot 11,29}{399 - 7 \cdot 6,71^2} = \frac{106,71}{83,83} = 1,27$$

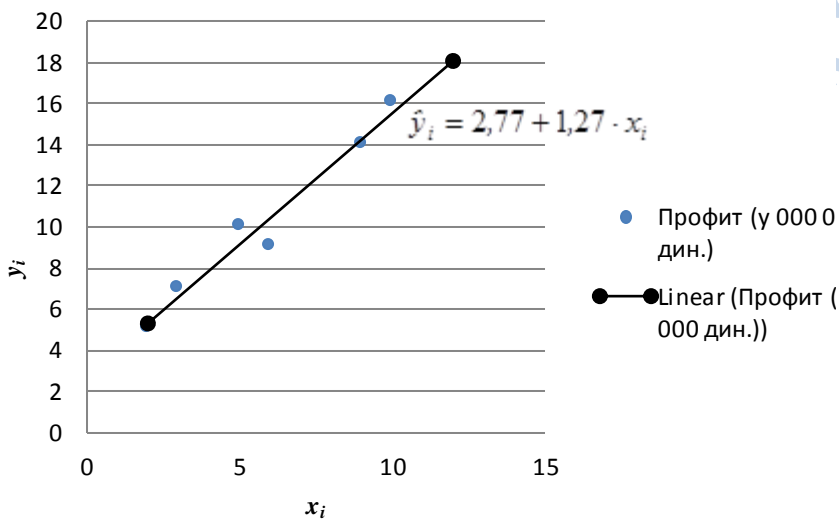
$$a = \bar{y} - b\bar{x} = 11,29 - 1,27 \cdot 6,71 = 2,77$$

Према томе, оцењена једначина регресије је:

$$\hat{y}_i = 2,77 + 1,27 \cdot x_i$$

Да би се оцењени регресиони модел графички приказао, на хоризонталној линији у координатном систему потребно је унети вредности за независну променљиву  $x_i$ , а на вертикалној линији вредности за зависну променљиву  $y_i$ . Парови података  $(x_i, y_i)$  уносе се у виду тачака у координатни систем, а линија регресије се уписује одређивањем вредности оцењеног модела за било које две тачке које су мање или једнаке најмањој вредности за  $x_i$ , односно веће или једнаке највећој вредности за  $x_i$ .

На пример, за  $x_i=2$ , следи да је  $\hat{y}_i = 2,77 + 1,27 \cdot 2 = 5,31$ , а за  $x_i=12$ , следи да је  $\hat{y}_i = 2,77 + 1,27 \cdot 12 = 18,01$ . Дакле, тачке са координатама (2;5,31) и (12;18,01) припадају линији регресије. Њиховим спајањем у координатној равни добија се график линије регресије.



Слика 1. Дијаграм распршености

Коефицијент  $a$  представља пресек са ординатном осом, односно одсечак на  $Y$  оси на дијаграму растурања. Ова оцењена вредност у пракси обично нема посебну важност, али како је то вредност зависне променљиве кад независна променљива има вредност нула, у нашем примеру се може рећи да у случају када нема промена у висини обима продаје профит износи 2.770.000 дин. Коефицијент  $b$  код праве линије представља коефицијент њеног правца, односно величину промене зависне променљиве када се независна променљива промени за јединицу. Стога се може закључити да када се обим продаје повећа за 1.000 тона, профит се у просеку повећава за 1.270.000 дин.

Закључак који треба извести је следећи: Вредност коефицијента  $a$  показује да у случају када нема промена у висини обима продаје, профит износи 2.770.000 дин., док се на основу коефицијента  $b$  може закључити да када се обим продаје повећа за 1.000 тона, профит се у просеку повећава за 1.270.000 дин.

б) Стандардна грешка се може израчунати коришћењем оцењених коефицијената  $a$  и  $b$ , односно:

$$s_e = \sqrt{\frac{\sum y_i^2 - a \cdot \sum y_i - b \cdot \sum x_i y_i}{n-2}} = \sqrt{\frac{1.031 - 2,77 \cdot 79 - 1,27 \cdot 637}{7-2}}$$

$$= \sqrt{\frac{3,18}{5}} = \sqrt{0,636} = 0,797$$

в) Да би се извршило предвиђање користи се следећи образац:

$$\hat{y}_0 - t_{\alpha/2;n-2} \cdot s_p < y_0 < \hat{y}_0 + t_{\alpha/2;n-2} \cdot s_p$$

Како би се предвидео ниво профита уколико је обим продаје 7.000 тона, како је дато у поставци задатка, односно за  $x_0 = 7$ , следи да је:

$$\hat{y}_0 = 2,77 + 1,27 \cdot x_0 = 2,77 + 1,27 \cdot 7 = 11,66$$

Дакле, на основу једначине регресије, при обиму продаје од 7.000 тона, може се очекивати да ће профит износити 11.660.000 дин.

Стандардна грешка прогнозе је:

$$s_p = s_e \sqrt{1 + \frac{1}{n} + \frac{(\bar{x} - x_0)^2}{\sum x_i^2 - n \cdot \bar{x}^2}} =$$

$$= 0,797 \cdot \sqrt{1 + \frac{1}{7} + \frac{(6,71 - 7)^2}{399 - 7 \cdot 6,71^2}} = 0,797 \cdot \sqrt{1,14} = 0,797 \cdot 1,06 = 0,85.$$

За 99%-ни интервал поузданости читавамо табличну вредност за број степени слободe  $n-2 = 7-2 = 5$ , која износи  $t_{0,995;5} = 4,03$ . Заменом свих ових вредности у формули:

$$\hat{y}_0 - t_{\alpha/2;n-2} \cdot s_p < y_0 < \hat{y}_0 + t_{\alpha/2;n-2} \cdot s_p$$

добија се да је

$$11,66 - 4,03 \cdot 0,85 < y_0 < 11,66 + 4,03 \cdot 0,85$$

тј.

$$8,23 < y_0 < 15,09.$$

Дакле, са 99% поузданости може се очекивати да ће се при обиму продаје од 7.000 тона, остварити профит између 8.230.000 и 15.090.000 дин.

За 95%-ни интервал поузданости читавамо табличну вредност за број степени слободе  $n - 2 = 7 - 2 = 5$ , која износи  $t_{0,975;5} = 2,57$ . Заменом свих ових вредности у формули:

$$\hat{y}_0 - t_{\alpha/2;n-2} \cdot s_p < y_0 < \hat{y}_0 + t_{\alpha/2;n-2} \cdot s_p$$

добива се да је

$$11,66 - 2,57 \cdot 0,85 < y_0 < 11,66 + 2,57 \cdot 0,85$$

тј.

$$9,48 < y_0 < 13,84.$$

Дакле, са 95% поузданости може се очекивати да ће се при обиму продаје од 7.000 тона, остварити профит између 9.480.000 и 13.840.000 дин.

г) Испитивање степена линеарне зависности заправо подразумева израчунавање коефицијента корелације преко следећег обрасца:

$$r = \frac{\sum x_i y_i - n \cdot \bar{x} \cdot \bar{y}}{\sqrt{(\sum x_i^2 - n \cdot \bar{x}^2) \cdot (\sum y_i^2 - n \cdot \bar{y}^2)}}$$

$$r = \frac{637 - 7 \cdot 6,71 \cdot 11,29}{\sqrt{(399 - 7 \cdot 6,71^2) \cdot (1.031 - 7 \cdot 11,29^2)}} = \frac{106,71}{\sqrt{83,83 \cdot 138,75}} = \frac{106,71}{107,85} = 0,9894$$

На основу израчунатог коефицијента корелације, може се закључити да постоји висок степен линеарне зависности између обима продаје и висине профита (вредност коефицијента корелације је блиска 1).

д) Процент зависности промене одређује се израчунавањем коефицијента детерминације преко следећег обрасца:

$$r^2 = \frac{a \cdot \sum y_i - b \cdot \sum x_i y_i - n \cdot \bar{y}^2}{\sum y_i^2 - n \cdot \bar{y}^2}$$

$$r^2 = \frac{2,77 \cdot 79 + 1,27 \cdot 637 - 7 \cdot 11,29^2}{1.031 - 7 \cdot 11,29^2} = \frac{135,57}{138,75} = 0,9771 \cdot 100 = 97,71\%$$

На основу израчунатог коефицијента детерминације, може се закључити да је 97,71% варијација у висини профита одређено променама у обиму остварене продаје.

Др Наташа Патић-Благојевић, проф.