2025. április 30-i gyakorlat [1]

Kétváltozós lineáris és lineárisra visszavezethető regresszió

Az *elaszticitás* (rugalmasság) azt méri, hogy az X változó 1%-os növekedése hány százalékos növekedést/csökkenést eredményez az Y változónál. Az elaszticitás kiszámítása a becsült eredményváltozóra:

$$El(\hat{y}, x) = \frac{\partial \hat{y}}{\partial x} \cdot \frac{x}{\hat{y}}$$

Kétváltozós lineáris esetben:

$$El(\hat{y}, x) = \hat{\beta}_1 \cdot \frac{x}{\hat{y}} = \frac{\hat{\beta}_1 x}{\hat{\beta}_0 + \hat{\beta}_1 x}$$

Kétváltozós hatvány regresszió esetén:

$$El(\hat{y}, x) = \hat{\beta}_0 \hat{\beta}_1 x^{\hat{\beta}_1 - 1} \cdot \frac{x}{\hat{y}} = \frac{\hat{\beta}_0 \hat{\beta}_1 x^{\hat{\beta}_1}}{\hat{\beta}_0 x^{\hat{\beta}_1}} = \hat{\beta}_1$$

Intervallumbecslés a függvényértékekre kétváltozós lineáris regresszió esetén:

- az átlagos értékre

$$Int_{1-\alpha}(E(Y_*)) = \hat{y}_* \pm t_{1-\frac{\alpha}{2}}(n-2) s_e \sqrt{\frac{1}{n} + \frac{(x_* - \overline{x})^2}{\sum d_x^2}},$$

- az egyedi értékre

$$Int_{1-\alpha}(Y_*) = \hat{y}_* \pm t_{1-\frac{\alpha}{2}}(n-2) s_e \sqrt{1 + \frac{1}{n} + \frac{(x_* - \overline{x})^2}{\sum d_x^2}},$$

ahol $s_e = \sqrt{\frac{\sum_{i=1}^{n} e_i^2}{n-2}}$ a *korrigált reziduális szórás.*

1. Egy taxivállalat 15 véletlenszerűen kiválasztott fuvar alapján vizsgálja, hogy hogyan függ a menetidő a távolságtól (megtett km-től). A 15 fuvar esetén a távolság és a menetidő:

távolság (km)	menetidő (perc)	távolság (km)	menetidő (perc)
3	8	9	20
4	19	12	23
4	13	15	44
6	21	16	47
6	11	16	41
7	19	20	46
8	14	26	48
8	19		

$$\sum y = 393, \qquad \sum xy = 5\,433, \qquad \sum x \ln y = 545.8033,$$

$$\sum \ln y = 46.7381, \qquad \sum x = 160, \qquad \sum x^2 = 2\,328, \qquad \sum e^2_{\text{lineáris}} = 473.7289$$

$$\sum \ln x \ln y = 106.6887, \qquad \sum (\ln x)^2 = 77.2063, \qquad \sum \ln x = 32.7487.$$

[1] A feladatok Keresztély-Sugár-Szarvas: Statisztika közgazdászoknak Példatár és feladatgyűjteményből, továbbá korábbi ZH feladatokból származnak.

(a) Jellemezze a távolság és a menetidő közötti lineáris, exponenciális illetve hatvány kapcsolatot és értelmezze a paramétereket!

(b) Becsülje meg mindegyik modell alapján, hogy egy 15 km távolságú út hány percet vesz igénybe!

(c) A lineáris modell esetén adjon 95%-os konfidencia intervallumot a 15 km távolságú fuvar átlagos menetidejére!

(d) A lineáris modell esetén adjon 95%-os konfidencia intervallumot egy 15 km távolságú fuvar egyedi menetidejére!

SPSS: Graphs ⇒ Scatter/Dot

Analyze ⇒ Regression ⇒ Curve Estimation: Linear, Compound, Power

Transform ⇒ Compute variable: becslés, hiba

Analyze ⇒ Regression ⇒ Linear

2. Az eszpresszó kávék hőmérséklet-változását vizsgálták az idő függvényében. Ennek érdekében 14 eszpresszót főztek, s különböző időpontokban megmérték az egyes kávék hőmérsékletét. A kapott eredményeket az alábbi táblázat tartalmazza:

Idő (perc) X	Hőmérséklet (°C) Y
1	82
5	76
8	70
11	65
15	61
18	57
22	52
25	51
30	47
34	45
38	43
42	41
45	39
50	38

Néhány számítási eredmény:

$$\sum_{i=1}^{14} x_i = 344, \qquad \sum_{i=1}^{14} x_i^2 = 11\,658 \qquad \sum_{i=1}^{14} \ln x_i = 40 \qquad \sum_{i=1}^{14} (\ln x_i)^2 = 129$$

$$\sum_{i=1}^{14} y_i = 767, \qquad \sum_{i=1}^{14} y_i^2 = 44\,629 \qquad \sum_{i=1}^{14} \ln y_i = 55.63 \qquad \sum_{i=1}^{14} (\ln y_i)^2 = 222$$

$$\sum_{i=1}^{14} x_i y_i = 16\,048, \qquad \sum_{i=1}^{14} x_i (\ln y_i) = 1\,316.19 \qquad \sum_{i=1}^{14} (\ln x_i) y_i = 2\,003.9 \qquad \sum_{i=1}^{14} \ln x_i \ln y_i = 155.76$$

(a) Írja fel az adatokra illeszkedő exponenciális regressziós modellt!

(b) Hány fokos a frissen főzött eszpresszó?

(c) 1 óra múlva várhatóan mennyi lesz egy csésze kávé hőmérséklete?

(d) 1992-ben egy nő beperelte a McDonald's gyorséttermet, ugyanis a kapott 82°C-os kávét magára borította, ami égési sérüléseket okozott. A szakértői vizsgálat szerint a 82°C-os kávé 2-7 másodperc alatt harmadfokú égési sérülést okoz az emberi bőrön. Megállapították, hogy 68°C -ra kell lehűteni a kávét, hogy elkerülhető legyen a súlyos égési sérülés. A nő 2,7 millió dollár kártérítést kapott. Ezen nevezetes eset kapcsán sok étterem 68°C-os kávét szolgál fel. Mennyi ideig kell várnia az étteremnek mielőtt felszolgálja a kávét (a kávéscsészébe töltés után), hogy az ne legyen melegebb 68°C -nál?

3. 1912-ben Tokió kormányzója 3 000 cseresznyefát ajándékozott Washington D.C.-nek a barátság jelképeként. A Nyugat-Potomac nemzeti parkban elültetett fák virágzása minden tavasszal rendkívüli látványosság és sok turistát csal a fővárosba. Ezen fák növekedési ütemét vizsgálták, s néhány véletlenszerűen kiválasztott fa magassága alapján 11 különböző időpontban megnézték a cseresznyefák átlagmagasságát. Elültetéskor mindegyik fa 1 éves volt és 6 láb magas (1 láb = 0.3048 m).

A fa kora (év) X	Átlagos magasság (láb) Y
1	6
2	9.5
2.5	10.5
4	15
5.5	17
6	17.5
7	18.5
8	19
8.5	19.5
10	19.7
11	19.8

Néhány számítási eredmény:

$$\sum_{i=1}^{11} x_i = 66, \qquad \sum_{i=1}^{11} x_i^2 = 500 \qquad \sum_{i=1}^{11} \ln x_i = 17 \qquad \sum_{i=1}^{11} (\ln x_i)^2 = 33$$

$$\sum_{i=1}^{11} y_i = 172, \qquad \sum_{i=1}^{11} y_i^2 = 2\,920 \qquad \sum_{i=1}^{11} \ln y_i = 29.6 \qquad \sum_{i=1}^{11} (\ln y_i)^2 = 81$$

$$\sum_{i=1}^{11} x_i y_i = 1\,171.8, \qquad \sum_{i=1}^{11} x_i(\ln y_i) = 187.64 \qquad \sum_{i=1}^{11} (\ln x_i) y_i = 307.42 \qquad \sum_{i=1}^{11} \ln x_i \ln y_i = 49.61$$

(a) Írja fel az adatokra illeszkedő hatvány regressziós modellt!

(b) Várhatóan milyen magas lesz egy 20 éves cseresznyefa? Reális-e az előrejelzés? Indokolja meg az állítását!

(c) Találtak egy régi fényképet, melyen a cseresznyefák átlagosan 10 láb magasak voltak. Hány évesek lehettek akkor a fák?