

Vastaa Stack-tehtäviin (stack3.aalto.fi/course/view.php?id=30)
 viimeistään 8.12.2014 kl. 12.00.

Palauta P-tehtävät viimeistään 8.12.2014 kl. 12.

Muista kirjoittaa nimesi, opiskelijanumerosi ja harjoitusryhmäsi!

P1. Muuttujien x ja Y havaitut arvot ovat

| | | | | | | |
|-----|------|-----|-----|-----|-----|-----|
| x | 0.4 | 1.1 | 1.4 | 2.2 | 2.3 | 2.4 |
| y | -0.5 | 1.4 | 1.6 | 3.3 | 4.1 | 4.2 |

Tästä otoksesta on laskettu seuraavat tunnusluvut:

| | | |
|---------|---------|----------|
| s_x^2 | s_y^2 | s_{xy} |
| 0.64267 | 3.3950 | 1.4640 |

- Määritä regressiomallin $Y_j = \beta_0 + \beta_1 x_j + \varepsilon_j$ kertoimien β_0 ja β_1 (pienimmän neliösumman) estimaatit.
- Piirrä pistediagrammi sekä estimoitu suora.
- Laske jäännösvarianssin estimaatti s^2 ja testaa nollahypoteesia $\beta_0 = -1$ merkitsevyystasolla 0.05 (olettaen, että satunnaismuuttujat ε_j ovat riippumattomia ja $N(0, \sigma^2)$ -jakautuneita).

P2. Otoksesta (x_j, y_j) , $j = 1, 2, \dots, 16$ laskettiin seuraavat tunnusluvut:

| | | | | |
|-----------|-----------|---------|---------|----------|
| \bar{x} | \bar{y} | s_x^2 | s_y^2 | r_{xy} |
| 0.5 | 4.8 | 4.4 | 6.2 | 0.7 |

- Määritä regressiomallin $Y = \beta_0 + \beta_1 x + \varepsilon$ kertoimien β_0 ja β_1 (pienimmän neliösumman) estimaatit.
- Laske jäännösvarianssin estimaatti s^2 ja testaa nollahypoteesia $\beta_1 \leq 0.3$ merkitsevyystasolla 0.05 (olettaen, että satunnaismuuttujat ε_j ovat riippumattomia ja $N(0, \sigma^2)$ -jakautuneita).
- Mikä on mallin selitysaste?

P3. Fossiiliset hiilidioksidipäästöt (CO₂-foss) olivat vuosina 2008–2012 kotitalouksien osalta seuraavat (yksikkönä miljoona tonnia)

| | | | | |
|------|------|------|------|------|
| 2008 | 2009 | 2010 | 2011 | 2012 |
| 6.79 | 6.32 | 6.25 | 5.60 | 5.66 |

Määritä näiden lukujen perusteella 95%:n luottamusväli vuoden 2013 päästöjen odotusarvolle kun oletat että päästöjä voidaan kuvata mallilla $Y_j = \beta_0 + \beta_1 x_j + \varepsilon_j$ missä Y_j on päästöt vuonna x_j ja satunnaismuuttujat ε_j ovat riippumattomia ja $N(0, \sigma^2)$ -jakautuneita.

Voit käyttää hyväksi seuraavia tunnuslukuja:

| | | | | |
|-----------|-----------|---------|---------|----------|
| \bar{x} | \bar{y} | s_x^2 | s_y^2 | s_{xy} |
| 2010 | 6.124 | 2.5 | 0.24693 | -0.745 |

ja muista että jos vaihdat vuosiluvut luvuiksi $x_j - 2013$ ainoastaan \bar{x} muuttuu ja sinun pitää laskea luottamusväli kertoimelle β_0 .

P4. Tutkija T tutki monen eri satunnaismuuttujan välisiä riippuvuuksia ja kun hän laski 400 otoskorrelaatiokerrointa hän löysi 27 kertaa kertoimen joka poikkesi nolasta merkitsevyytasolla 5%.

Valitse ensin nolalahypoteesiksi, että kaikki satunnaismuuttujat ovat riippumattomia. Tästä voit johtaa uuden nolalahypoteesin joka koskee todennäköisyyttä, että löydat korrelaatiokertoimen joka poikkeaa nolasta merkitsevyytasolla 5% ja voit ottaa tämän nolalahypoteesin sellaiseksi, että vaihtoehto sille on yksisuuntainen. Testaa tätä nolalahypoteesia merkitsevyytasolla 5% käyttäen tutkija T:n tuloksia.

P5. Havaintoaineistosta (x_i, y_i) , $i = 1, \dots, n$ on laskettu regressiosuora $y = b_0 + b_1x$ kun x on ”selittävä” muuttuja (eli kun oletetaan, että $Y_i = \beta_0 + \beta_1x_i + \varepsilon_i$) ja (käänteisregressiosuora) $x = a_0 + a_1y$ kun y on ”selittävä” muuttuja (eli kun oletetaan, että $X_i = \alpha_0 + \alpha_1y_i + \epsilon_i$). Lisäksi on laskettu otoskorrelaatiokerroin r_{xy} ja on todettu, että nolalahypoteesi $H_0 : \beta_1 = 0$ voidaan hylätä merkitsevyytasolla 0.01. Mitä voidaan sanoa seuraavista väitteistä (olettaen, että $s_x > 0$ ja $s_y > 0$):

- (a) Suorat $y = b_0 + b_1x$ ja $x = a_0 + a_1y$ ovat samat jos ja vain jos $|r_{xy}| = 1$.
- (b) Jos testataan hypoteesia $H_0 : \alpha_1 = 0$ ne sekin voidaan hylätä merkitsevyytasolla 0.01.

Vihje: Regressiosuorat voidaan kirjoittaa muodossa $y - \bar{y} = b_1(x - \bar{x})$ ja $x - \bar{x} = a_1(y - \bar{y})$.