CAPITOLUL 2. MODELUL DE REGRESIE LINIARĂ SIMPLĂ

<table>
<tr><th colspan="2">ESTIMAREA</th><th>INTERPRETARE</th></tr>
<tr><td rowspan="2">punctuală a parametrilor
$\boldsymbol{\beta_0}$ (ordonata la origine) şi
$\boldsymbol{\beta_1}$ (panta dreptei de regresie)</td><td>$\boldsymbol{b_0}$</td><td>$\boldsymbol{b_0}$: nivelul mediu estimat al variabilei dependente (Y) atunci când variabila independenta (X) ia valoarea zero</td></tr>
<tr><td>$\boldsymbol{b_1}$
$\boldsymbol{b_1 = \frac{\Delta Y}{\Delta X}}$</td><td>$\boldsymbol{b_1}$: la o creştere a variabilei independente X cu 1 unitate, variabila dependentă (Y) variază (scade sau creşte în funcţie de semn), în medie, cu $\boldsymbol{b_1}$</td></tr>
<tr><td rowspan="2">prin interval de încredere a parametrilor $\boldsymbol{\beta_0}$ şi $\boldsymbol{\beta_1}$</td><td>IC(β_0): $\left[\boldsymbol{b_0 \pm t_{\alpha/2;n-2} s_{\hat{\beta}_0}}\right]$</td><td>Cu o probabilitate de $(1-\alpha)$, se poate garanta că parametrul β_0 (sau ordonata la origine) este acoperit de intervalul $\left[\boldsymbol{b_0 \pm t_{\alpha/2;n-2} s_{\hat{\beta}_0}}\right]$
În condiţiile unui risc α, se consideră că valoarea reală a parametrului β_0 (sau ordonata la origine) se află în afara limitelor intervalului $\left[\boldsymbol{b_0 \pm t_{\alpha/2;n-2} s_{\hat{\beta}_0}}\right]$</td></tr>
<tr><td>IC(β_1): $\left[\boldsymbol{b_1 \pm t_{\alpha/2;n-2} s_{\hat{\beta}_1}}\right]$</td><td>Cu o probabilitate de $(1-\alpha)$, se poate garanta că parametrul β_1 (sau panta dreptei de regresie) este acoperit de intervalul $\left[\boldsymbol{b_1 \pm t_{\alpha/2;n-2} s_{\hat{\beta}_1}}\right]$
În condiţiile unui risc α, se consideră că valoarea reală a parametrului β_1 (sau panta dreptei de regresie) se află în afara limitelor intervalului $\left[\boldsymbol{b_1 \pm t_{\alpha/2;n-2} s_{\hat{\beta}_1}}\right]$</td></tr>
<tr><td>coeficientului de corelaţie</td><td>între Y şi X: $\boldsymbol{r}$
$\boldsymbol{-1 \le r \le 1}$
$\boldsymbol{r = \tilde{b}_1}$</td><td>$\boldsymbol{r}$: indică sensul (după semn) şi măsoară intensitatea (după valoarea în modul) legăturii dintre două variabile.
<table>
<tr><th>$\boldsymbol{|r| = 0}$</th><th>$\boldsymbol{0 \leftarrow |r|}$</th><th>$\boldsymbol{|r| \rightarrow\ 0,5\ \leftarrow |r|}$</th><th>$\boldsymbol{|r| \rightarrow 1}$</th><th>$\boldsymbol{|r| = 1}$</th></tr>
<tr><td>nu există o leg. liniară între Y şi X</td><td>leg. liniară de intensitate slabă între Y şi X</td><td>leg. liniară de intensitate moderată între Y şi X</td><td>leg. liniară de intensitate puternică între Y şi X</td><td>leg. liniară perfectă între Y şi X</td></tr>
</table>
Exemplu:
$\boldsymbol{r = 0,25}$: între Y şi X există o legătură liniară directă (după semn) şi de intensitate slabă (după valoarea în modul $|\boldsymbol{r}|$)</td></tr>
<tr><td>raportului de determinaţie $\boldsymbol{\eta^2}$</td><td>$\boldsymbol{R^2 = \frac{ESS}{TSS} = 1 - \frac{RSS}{TSS}}$</td><td>$\boldsymbol{R^2}$: măsoară cât din variaţia totală a variabilei dependente (Y) este explicată de modelul de regresie SAU de variabila independentă (X)</td></tr>
</table>

<table>
<tr>
<td></td>
<td>$\mathbf{0 \leq R^2 \leq 1}$
$\boldsymbol{R^2 = (R)^2}$
$\boldsymbol{R^2 = (r)^2}$</td>
<td>$\mathbf{1 - R^2}$: arată cât din variaţia totală a variabilei dependente (Y) este explicată de influența factorilor reziduali (aleatori) sau de factorii nespecificați (neincluși) în model
Exemplu:
$\boldsymbol{R^2 = 0,311}$ $(\boldsymbol{R^2}\% = \mathbf{31,1}\%)$
Variația variabilei dependente (Y) este explicată în proporție de 31,1% de variația variabilei independente (X). Restul de 68,9% reprezintă influența altor factori nespecificați în model (factorii reziduali).</td>
</tr>
<tr>
<td>raportului de corelaţie multiplă $\boldsymbol{\eta}$</td>
<td>$\boldsymbol{R = \sqrt{R^2}}$
$\mathbf{0 \leq R \leq 1}$</td>
<td>$\boldsymbol{R}$: măsoară intensitatea legăturii liniare dintre variabile
<table>
<tr><th>$\boldsymbol{R = 0}$</th><th>$\boldsymbol{0 \leftarrow R}$</th><th>$\boldsymbol{R \rightarrow \; 0,5 \; \leftarrow R}$</th><th>$\boldsymbol{R \rightarrow 1}$</th><th>$\boldsymbol{R = 1}$</th></tr>
<tr><td>nu există o leg. liniară între variabile</td><td>leg. liniară de intensitate slabă între variabile</td><td>leg. liniară de intensitate moderată între variabile</td><td>leg. liniară de intensitate puternică între variabile</td><td>leg. liniară perfectă între variabile între variabile</td></tr>
</table>
Exemplu:
$\boldsymbol{R = 0,558}$
Între variabila dependentă (Y) și variabila independentă (X) există o legătură liniară de intensitate moderată.</td>
</tr>
</table>

<table>
<tr><th colspan="4">TESTAREA</th></tr>
<tr><td rowspan="4">Testarea constantei sau a ordonatei la origine $\boldsymbol{\beta_0}$ şi a pantei dreptei de regresie $\boldsymbol{\beta_1}$</td><td>Ipoteze:</td><td>$H_0: \beta_0 = 0$ (parametrul β_0 nu diferă semnificativ de 0 SAU constanta modelului nu este semnificativă statistic)
$H_1: \beta_0 \neq 0$ (parametrul β_0 diferă semnificativ de 0 SAU constanta modelului este semnificativă statistic)</td><td>$H_0: \beta_1 = 0$ (coeficientul de regresie β_1 nu diferă semnificativ de 0, ceea ce înseamnă că variabila independentă X nu are o influență liniară semnificativă asupra variabilei dependente Y SAU X nu explică semnificativ variația lui Y)
$H_1: \beta_1 \neq 0$ (coeficientul de regresie β_1 diferă semnificativ de 0, ceea ce înseamnă că variabila independentă X are o influență liniară semnificativă asupra variabilei dependente Y SAU X explică semnificativ variația lui Y)</td></tr>
<tr><td>Valoarea teoretică a statisticii test Student:</td><td colspan="2">$\boldsymbol{t_{teoretic} = t_{\alpha/2;n-2}}$</td></tr>
<tr><td>Valoarea calculată a statisticii test Student:</td><td>(β_0): $\boldsymbol{t_{calc} = \frac{b_0}{s_{\hat{\beta}_0}}}$</td><td>$(\beta_j)$: $\boldsymbol{t_{calc} = \frac{b_j}{s_{\hat{\beta}_j}}}$</td></tr>
<tr><td>Decizia:</td><td colspan="2">Dacă se ţine cont de valoarea calculată a testului, regula de decizie este următoarea:
- dacă $\boldsymbol{|t_{calc}| \leq t_{\alpha/2;\, n-2}}$, nu se respinge ipoteza nulă (H_0)
- dacă $\boldsymbol{|t_{calc}| > t_{\alpha/2;\, n-2}}$, se respinge ipoteza nulă (H_0), în condițiile unui α
Dacă se ţine cont de semnificaţia testului, regula de decizie este următoarea:
- dacă $\boldsymbol{Sigt \geq \alpha}$, nu se respinge ipoteza nulă (H_0)
- dacă $\boldsymbol{Sigt < \alpha}$, se respinge H_0, în condițiile unui risc α</td></tr>
<tr><td rowspan="2">Testarea coeficientului de corelație $\boldsymbol{\rho}$</td><td>Ipoteze:</td><td colspan="2">$H_0: \rho_{y1} = 0$ (coeficientul de corelație ρ nu diferă semnificativ de 0, ceea ce înseamnă că între cele două variabile (Y şi X) nu există o legătură liniară semnificativă SAU cele două variabile (Y şi X) nu sunt corelate semnificativ)
$H_1: \rho_{y1} \neq 0$ (coeficientul de corelație ρ diferă semnificativ de 0, ceea ce înseamnă că între cele două variabile (Y şi X) există o legătură liniară semnificativă SAU cele două variabile (Y şi X) sunt corelate semnificativ)</td></tr>
<tr><td>Valoarea teoretică a statisticii test Student:</td><td colspan="2">$\boldsymbol{t_{teoretic} = t_{\alpha/2;n-2}}$</td></tr>
</table>

<table>
<tr><td rowspan="2"></td><td>Valoarea calculată a statisticii test Student:</td><td colspan="2">$$\boldsymbol{t_{calc}=\frac{r_{y1}}{\sqrt{\frac{1-r^2}{n-2}}}}$$</td></tr>
<tr><td>Decizia:</td><td colspan="2">Dacă se ţine cont de valoarea calculată a testului, regula de decizie este următoarea:
- dacă $|\boldsymbol{t_{calc}}| \leq \boldsymbol{t_{\alpha/2;\, n-2}}$, nu se respinge ipoteza nulă (H_0)
- dacă $|\boldsymbol{t_{calc}}| > \boldsymbol{t_{\alpha/2;\, n-2}}$, se respinge ipoteza nulă (H_0), în condițiile unui risc α
Dacă se ţine cont de semnificaţia testului, regula de decizie este următoarea:
- dacă $\boldsymbol{Sigt \geq \alpha}$, nu se respinge ipoteza nulă (H_0)
- dacă $\boldsymbol{Sigt < \alpha}$, se respinge H_0, în condițiile unui risc α</td></tr>
<tr><td rowspan="4">Testarea modelului de regresie și a raportului de determinație $\boldsymbol{\eta^2}$ (sau raportului de corelație $\boldsymbol{\eta}$)</td><td>Ipoteze:</td><td>$H_0: \beta_0 = 0, \beta_1 = 0$ (modelul de regresie nu explică semnificativ dependența liniară dintre cele două variabile SAU între cele două variabile nu există o legătură liniară semnificativă SAU modelul de regresie construit nu este corect specificat)
$H_1: \beta_1 \neq 0$ (modelul de regresie explică semnificativ dependența liniară dintre cele două variabile SAU între cele două variabile există o legătură liniară semnificativă)</td><td>$H_0: \eta = 0$ (raportul de determinație η^2 sau raportul de corelația η nu diferă semnificativ de 0, ceea ce înseamnă că între variabile nu există o legătură liniară semnificativă)
$H_1: \eta > 0$ (raportul de determinație η^2 sau raportul de corelația η este semnificativ mai mare decât 0, ceea ce înseamnă că între variabile există o legătură liniară semnificativă)</td></tr>
<tr><td>Valoarea teoretică a statisticii test Fisher:</td><td colspan="2">$$\boldsymbol{F_{teoretic} = F_{\alpha;k-1;n-k}}$$</td></tr>
<tr><td>Valoarea calculată a statisticii test Fisher:</td><td>$$\boldsymbol{F_{calc} = \frac{\frac{ESS}{k-1}}{\frac{RSS}{n-k}} = \frac{ESS}{RSS}\cdot\frac{n-k}{k-1}}$$</td><td>$$\boldsymbol{F_{calc} = \frac{R^2}{1-R^2}\cdot\frac{n-k}{k-1}}$$</td></tr>
<tr><td>Decizia:</td><td colspan="2">Dacă se ţine cont de valoarea calculată a testului, regula de decizie este următoarea:
- dacă $\boldsymbol{F_{calc}} \leq \boldsymbol{F_{\alpha;\, k-1;\, n-k}}$, nu se respinge ipoteza nulă (H_0)
- dacă $\boldsymbol{F_{calc}} > \boldsymbol{F_{\alpha;\, k-1;\, n-k}}$, se respinge ipoteza nulă (H_0), în condițiile unui risc α
Dacă se ţine cont de semnificaţia testului, regula de decizie este următoarea:
- dacă $\boldsymbol{SigF \geq \alpha}$, nu se respinge ipoteza nulă (H_0)
- dacă $\boldsymbol{SigF < \alpha}$, se respinge H_0, în condițiile unui risc α</td></tr>
</table>