1 Regresja liniowa

$Y_i = \beta_0+\beta_1 X_i+\xi_i$, β_0 β_1 $\xi_i \sim N\left(0, \sigma^2\right)$ $E(Y_i|X_i) = \beta_0+\beta_1 X_i$, $Var(Y_i|X_i) = \sigma^2$

2 Estymacja

$\hat{Y} = b_0+b_1 X_i$, $e_i = Y_i-\hat{Y} = Y_i-(b_0+b_1 X_i)$ $b_1 = \frac{\sum\left(X_i-\bar{X}\right)\left(Y_i-\bar{Y}\right)}{\sum\left(X_i-\bar{X}\right)^2}$, $b_0 = \bar{Y}-b_1\bar{X}$

$$s^2 = \frac{\sum\left(Y_i - \hat{Y}_i\right)^2}{n-2} = \frac{\sum e_i^2}{n-2} = \frac{SSE}{dfE} = MSE$$

3 Przedziały ufności

3.1 Dla β_1

$b_1 \sim N(\beta_1, \sigma^2(b_1))$, $\sigma^2(b_1) = \frac{\sigma^2}{\sum\left(X_i-\bar{X}\right)^2}$, $t = \frac{(b_1-\beta_1)}{s(b_1)} \sim t(n-2)$, $s^2(b_1) = \frac{s^2}{\sum\left(X_i-\bar{X}\right)^2}$,

$b_1 \pm t_c s(b_1)$, $t_c = (1-\alpha/2, n-2)$

$P(|z| \geqslant |t|)$, $z \sim t(n-2)$

3.2 Dla β_0

$b_1 \sim N(\beta_0, \sigma^2(b_0))$,

$\sigma^2(b_0) = \sigma^2\left[\frac{1}{n} + \frac{\bar{X}^2}{\sum\left(X_i-\bar{X}\right)^2}\right]$,

$t = \frac{b_0+\beta_0}{s(b_0)} \sim t(n-2)$

$s^2(b_0) = s^2\left[\frac{1}{n} + \frac{\bar{X}^2}{\sum\left(X_i-\bar{X}\right)^2}\right]$,

$b_0 \pm t_c s(b_0)$, $t_c = (1-\alpha/2, n-2)$,

$t = \frac{b_0+\beta_{00}}{s(b_0)}$, $P(|z| \geqslant |t|)$,

$$z \sim t(n-2)$$

4 Moc

$t = \frac{b_1}{s(b_1)} \sim t(n-2, \delta)$, $\delta = \frac{\beta_1}{\sigma(b_1)}$,

$b_1 \sim N(\beta_1, \sigma_2(b_1))$,

$\sigma^2(b_1) = \frac{\sigma^2}{\sum\left(X_i-\bar{X}\right)^2}$,

5 Estymacja średniej

$E(Y_h) = \mu_h = \beta_0 + \beta_1 X_h,$
$\hat{\mu}_h = b_0 + b_1 X_h,$
$\hat{\mu}_h \sim N(\mu_h, \sigma^2),$
$\sigma^2(\hat{\mu}_h) = \sigma^2 \left[\frac{1}{n} + \frac{(X_h - \bar{X})^2}{\sum (X_i - \bar{X})^2}\right],$
$s^2(\hat{\mu}_h) = s^2 \left[\frac{1}{n} + \frac{(X_h - \bar{X})^2}{\sum (X_i - \bar{X})^2}\right],$
$\frac{\hat{\mu}_h - E(Y_h)}{s(\hat{\mu}_h)} \sim t(n-2)$
$\hat{\mu}_h \pm t_c s(\hat{\mu}_h)$

6 Prognoza

$Y_h = \beta_0 + \beta_1 X_h + \xi_h,$
$s^2(pred) = s^2 \left[1 + \frac{1}{n} + \frac{(X_h - \bar{X})^2}{\sum (X_i - \bar{X})^2}\right],$
$\frac{Y_h - \hat{\mu}_h}{s(pred)} \sim t(n-2),$
$\hat{\mu}_h \pm t_c s(pred)$

7 ANOVA

$SST = \sum (Y_i - \bar{Y})^2$, $dfT = n - 1$, $MST = \frac{SST}{dfT}$
$SSM = \sum (\hat{Y}_i - \bar{Y})^2$, $dfM = 1$, $MSM = \frac{SSM}{dfM}$
$SSE = \sum (Y_i - \hat{Y}_i)^2$, $dfE = n - 2$, $MSE = \frac{SSE}{dfE}$
$F = \frac{MSM}{MSE} \sim F(dfM, dfE) = F(1, n-2)$

8 Porównywanie modeli

$$F = \frac{\frac{SSE(R) - SSE(F)}{dfE(R) - dfE(F)}}{MSE(F)}$$

$SSE(R) = \sum (Y_i - b_0)^2 = \sum \left(Y_i - \bar{Y}\right)^2 = SST$ $SSE(F) = SSE$ $dfE(R) = n - 1$ $dfE(F) = n - 2$

$$r^2 = SSM/SST = 1 - SSE/SST$$