Прогноз

Линейная регрессия

Novosibirsk State University
*THE REAL SCIENCE

2GIS

Как построить точечный прогноз по линейной регрессии?

$$y_i = \theta_0 + \theta_1 x_i + \varepsilon_i, i = 1, 2, \ldots, n$$

Новое наблюдение $(x_{n+1}, ?)$

$$\hat{y}_{n+1} = \hat{\theta}_0 + \hat{\theta}_1 x_{n+1}$$

Как построить интервальный прогноз? Вариант № 1

$$a = \frac{1}{n}\sum_{i=1}^{n}\left|y_i - \hat{y}_i\right| = \frac{1}{n}\sum_{i=1}^{n}\left|\varepsilon_i\right|$$

$$\hat{y} \pm a$$

$$\hat{y} \pm 2a$$

Как построить интервальный прогноз? Вариант № 2

$$s_\varepsilon = \sqrt{S_\varepsilon^2} = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(\varepsilon_i - \bar{\varepsilon})^2}$$

$$[\hat{y}_i - \lambda_\alpha S_\varepsilon ; \hat{y}_i + \lambda_\alpha S_\varepsilon]$$

S — среднеквадратическое отклонение остатков

λ — квантиль стандартного нормального распределения

$$\lambda_\alpha = \Phi_0^{-1}\left(1 - \frac{\alpha}{2}\right)$$

Как построить интервальный прогноз? Вариант № 3

$$y \pm t_{\int,\alpha} \times s_{e_{n+1}}$$

$$t_{f,\alpha} = t^{-1}(\mathrm{n-m-1})$$

$$\mathrm{s}_{e_{n+1}} = \sqrt{s_{\varepsilon}^{2}\left(1 + \frac{1}{n} + \frac{(x_{n+1} - \bar{x})^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}\right)}$$

Интервальный прогноз для множественной регрессии

$$s_{e_{N+1}}^2 = s_\varepsilon^2 \left(1 + x_{N+1}^T \left(X^T X \right)^{-1} x_{N+1} \right)$$