

## Семинар 7.

## Безусловное прогнозирование.

1. Проверьте формулу (из лекции) для среднеквадратической ошибки прогноза:

$$\mathbb{E}(\hat{y}_{n+1} - y_{n+1})^2 = \sigma_\varepsilon^2 \left( 1 + x_{n+1} (X^T X)^{-1} x_{n+1}^T \right).$$

Здесь  $\hat{y}_{n+1}$  — прогнозное значение зависимой переменной для нового наблюдения,  $y_{n+1}$  — истинное значение зависимой переменной для нового наблюдения,  $x_{n+1}$  —  $1 \times k$  вектор-строка объясняющих переменных для нового наблюдения.

2. (самостоятельно, используя формулу из задания 1) Докажите равенство для среднеквадратической ошибки прогноза в случае парной регрессии с константой вида  $y_i = \beta_1 + \beta_2 x_i + \varepsilon_i$  следующему выражению:

$$\mathbb{E}(\hat{y}_{n+1} - y_{n+1})^2 = \sigma^2 \left( 1 + \frac{1}{n} + \frac{(x_{n+1} - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2} \right).$$

3. Для модели парной регрессии  $y_i = \beta_1 + \beta_2 x_i + \varepsilon_i$ ,  $i = 1, \dots, 10$ , известно, что

$$\sum_{i=1}^{10} y_i = 8, \sum_{i=1}^{10} x_i = 40, \sum_{i=1}^{10} y_i^2 = 26, \sum_{i=1}^{10} x_i^2 = 200, \sum_{i=1}^{10} y_i x_i = 20.$$

Для некоторого наблюдения дано  $x_{11} = 10$ . Предполагая, что данное наблюдение удовлетворяет исходной модели,

- (а) вычислите наилучший линейный несмещенный прогноз величины  $y_0$ ;
  - (б) оцените стандартную ошибку прогноза.
4. Рассмотрим следующую регрессионную модель зависимости логарифма заработной платы  $\ln(W)$  от уровня образования  $Edu$ , опыта работы  $Exp$ ,  $Exp^2$  и уровня образования родителей  $Fedu$  и  $Medu$ :

$$\widehat{\ln}(W_i) = \hat{\beta}_1 + \hat{\beta}_2 Edu_i + \hat{\beta}_3 Exp_i + \hat{\beta}_4 Exp_i^2 + \hat{\beta}_5 Fedu_i + \hat{\beta}_6 Medu_i.$$

Модель регрессии была отдельно оценена по выборкам из 35 мужчин и 23 женщин, и были получены остаточные суммы квадратов  $RSS_1 = 34.4$  и  $RSS_2 = 23.4$ , соответственно. Остаточная сумма квадратов в регрессии, оцененной по объединенной выборке, равна 70.3. Протестируйте на 5% уровне значимости гипотезу об отсутствии дискриминации в оплате труда между мужчинами и женщинами.