

**Регрессионный анализ.**  
**Лекция 2: план, learning outcomes**

1. Допущения об ошибках в линейной регрессионной модели. Понятие Best Linear Unbiased Estimator (BLUE)
2. Распределение оценок коэффициентов в парной линейной регрессии, параметры распределения

$$\hat{Var}(\hat{\beta}_0) = \frac{\hat{\sigma}^2 \sum_{i=1}^n x_i^2}{n \sum_{i=1}^n (x_i - \bar{x})^2}$$

$$\hat{Var}(\hat{\beta}_1) = \frac{\hat{\sigma}^2}{\sum_{i=1}^n (x_i - \bar{x})^2}$$

где

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^n e_i^2}{n - 2}$$

3. Проверка статистической значимости оценок коэффициентов в парной линейной регрессии: t-статистика как результат деления оценки коэффициента на стандартную ошибку.
4. Разложение вариации зависимой переменной в регрессионной модели:

$$\sum_{i=1}^n (y_i - \bar{y})^2 = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2 + \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

5. Коэффициент детерминации: расчет и интерпретация
6. ANOVA-таблица:

- TSS (total sum of squares)
- ESS (explained sum of squares)
- RSS (residual sum of squares)
- MSS (mean sum of squares)
- F-статистика и проверка гипотезы о том, что модель с предиктором ничем не лучше, чем модель на константу