Асимптотические доверительные интервалы на основе ЦПТ

Предпосылки

- Выборка достаточно велика ($n \to \infty$)
- Наблюдения независимы
- Нет выбросов

Основа: Центральная предельная теорема (ЦПТ)

Д.И. для мат. ожидания выборок из любого распределения

$X_1 \dots, X_n \sim iid$ и $Y_1, \dots, Y_n \sim iid$ с мат. ожиданиями $\mu_y,\ \mu_x$

- для одного μ

$$\bar{X} - Z_{1-\frac{\alpha}{2}} \frac{\hat{\sigma}}{\sqrt{n}} < \mu < \bar{X} + Z_{1-\frac{\alpha}{2}} \frac{\hat{\sigma}}{\sqrt{n}}$$

- для разности мат. ожиданий $\mu_y - \mu_x$ и независимых наблюдений (X_i, Y_i)

$$\bar{Y} - \bar{X} - Z_{1-\frac{\alpha}{2}} \sqrt{\frac{\hat{\sigma}_x^2}{n_x} + \frac{\hat{\sigma}_y^2}{n_y}} < \mu_y - \mu_x < \bar{Y} - \bar{X} + Z_{1-\frac{\alpha}{2}} \sqrt{\frac{\hat{\sigma}_x^2}{n_x} + \frac{\hat{\sigma}_y^2}{n_y}}$$

Д.И. для теоретической доли распределения Бернулли

$X, \dots, X_n \sim iid\ Ber(p_x)$ и $Y_1, \dots, Y_n \sim iid\ Ber(p_y)$ с параметрами успеха $p_x,\ p_y$

- для одной теоретической доли p

$$\hat{p} - Z_{1-\frac{\alpha}{2}} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} < p < \hat{p} + Z_{1-\frac{\alpha}{2}} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

- для разности долей $p_y - p_x$

$$\hat{p}_y - \hat{p}_x - Z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}_y(1-\hat{p}_y)}{n_y} + \frac{\hat{p}_x(1-\hat{p}_x)}{n_x}} < p_y - p_x < \hat{p}_y - \hat{p}_x + Z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}_y(1-\hat{p}_y)}{n_y} + \frac{\hat{p}_x(1-\hat{p}_x)}{n_x}}$$

Д.И. для функции от некоторого параметра для $\boldsymbol{X_1 \dots, X_n \sim F(\theta)}$ (дельта-метод)

- для $g(\theta)$

$$g(\hat{\theta}_{ML}) - Z_{1-\frac{\alpha}{2}} \sqrt{\frac{(g'(\hat{\theta}_{ML}))^2}{var(\hat{\theta}_{ML})}} < g(\theta) < g(\hat{\theta}_{ML}) + Z_{1-\frac{\alpha}{2}} \sqrt{\frac{(g'(\hat{\theta}_{ML}))^2}{var(\hat{\theta}_{ML})}}$$

Доверительные интервалы для нормальных выборок

Предпосылки

- Наблюдения независимы
- Требования на размер выборки нет
- Выборка из нормального распределения

$$X_1, \dots, X_n \sim N(\mu_x, \sigma_x^2) \text{ и } Y_1, \dots, Y_n \sim N(\mu_y, \sigma_y^2)$$

Основа:

- свойства нормального распределения
- распределения хи-квадрат, Стьюдента, Фишера
- теорема Фишера

Для мат. ожидания μ

- **если известна $\boldsymbol{\sigma^2}$**

$$\bar{X} - Z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} < \mu < \bar{X} + Z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$$

- **если неизвестна $\boldsymbol{\sigma^2}$**

$$\bar{X} - t_{1-\frac{\alpha}{2};n-1} \frac{\hat{\sigma}}{\sqrt{n}} < \mu < \bar{X} + t_{1-\frac{\alpha}{2};n-1} \frac{\hat{\sigma}}{\sqrt{n}}$$

Для разности мат. ожиданий $\boldsymbol{\mu_y - \mu_x}$ и независимых наблюдений: $(\boldsymbol{X_i}, \boldsymbol{Y_i})$

- **если известны $\boldsymbol{\sigma_x^2, \sigma_y^2}$**

$$\bar{Y} - \bar{X} - Z_{1-\frac{\alpha}{2}} \sqrt{\frac{\sigma_x^2}{n_x} + \frac{\sigma_y^2}{n_y}} < \mu_y - \mu_x < \bar{Y} - \bar{X} + Z_{1-\frac{\alpha}{2}} \sqrt{\frac{\sigma_x^2}{n_x} + \frac{\sigma_y^2}{n_y}}$$

- **если неизвестны $\boldsymbol{\sigma_x^2, \sigma_y^2}$, но равны**

$$\bar{Y} - \bar{X} - t_{1-\frac{\alpha}{2};n_x+n_y-2} \hat{\sigma}_o \sqrt{\frac{1}{n_x} + \frac{1}{n_y}} < \mu_y - \mu_x < \bar{Y} - \bar{X} + t_{1-\frac{\alpha}{2};n_x+n_y-2} \hat{\sigma}_o \sqrt{\frac{1}{n_x} + \frac{1}{n_y}}$$

где $\sigma_0^2 = \dfrac{\hat{\sigma}_x^2(n_x - 1) + \hat{\sigma}_y^2(n_y - 1)}{n_x + n_y - 2}$

- **если неизвестны $\boldsymbol{\sigma_x^2, \sigma_y^2}$ и не равны**

$$\bar{Y} - \bar{X} - t_{1-\frac{\alpha}{2};d} \sqrt{\frac{\hat{\sigma}_x^2}{n_x} + \frac{\hat{\sigma}_y^2}{n_y}} < \mu_y - \mu_x < \bar{Y} - \bar{X} + t_{1-\frac{\alpha}{2};d} \sqrt{\frac{\hat{\sigma}_x^2}{n_x} + \frac{\hat{\sigma}_y^2}{n_y}}$$

где

$$d = \frac{\left(\dfrac{\sigma_x^2}{n_x} + \dfrac{\sigma_y^2}{n_y}\right)^2}{\dfrac{\sigma_x^4}{n_x^2(n_x - 1)} + \dfrac{\sigma_y^4}{n_y^2(n_y - 1)}}$$

Для дисперсии σ^2

- если известно μ

$$\frac{\sum_{i=1}^{n}(x_i-\mu)^2}{\mathcal{X}^2_{1-\frac{\alpha}{2};n}} \le \sigma^2 \le \frac{\sum_{i=1}^{n}(x_i-\mu)^2}{\mathcal{X}^2_{\frac{\alpha}{2};n}}$$

- если неизвестно μ

$$\frac{(n-1)\hat{\sigma}^2}{\mathcal{X}^2_{1-\frac{\alpha}{2};n-1}} \le \sigma^2 \le \frac{(n-1)\hat{\sigma}^2}{\mathcal{X}^2_{\frac{\alpha}{2};n-1}}$$

Для отношения дисперсия $\boldsymbol{\sigma_x^2/\sigma_y^2}$ и независимых наблюдений: $(\boldsymbol{X_i},\boldsymbol{Y_i})$

$$\frac{1}{F_{\frac{\alpha}{2};n_y-1,n_x-1}}\frac{\hat{\sigma}_x^2}{\hat{\sigma}_y^2} \le \frac{\sigma_x^2}{\sigma_y^2} \le \frac{\hat{\sigma}_x^2}{\hat{\sigma}_y^2}F_{\frac{\alpha}{2};n_y-1,n_x-1}$$

Для разности мат. ожиданий $\boldsymbol{\mu_y-\mu_x}$ и зависимых наблюдений: $(\boldsymbol{X_i},\boldsymbol{Y_i})$

$$\bar{Y}-\bar{X}-t_{1-\frac{\alpha}{2};n-1}\frac{\hat{\sigma}_\Delta}{\sqrt{n}} < \mu_y-\mu_x < \bar{Y}-\bar{X}+t_{1-\frac{\alpha}{2};n-1}\frac{\hat{\sigma}_\Delta}{\sqrt{n}}$$

где $\sigma_\Delta^2 = \frac{1}{n-1}\sum_{i=1}^{n}[(Y_i-X_i)-(\bar{Y}-\bar{X})]^2$