https://www.facebook.com/profile.php?id=10000654912
https://github.com/lovgager/latex

25.04.2021

Задача. *По выборке из n независимых нормально распределённых случайных величин построить доверительный интервал уровня γ для неизвестного математического ожидания в случае известной дисперсии.*

Решение. Обозначим случайные величины в выборке через $X_1, ..., X_n$. Пусть σ_0 – известное стандартное отклонение (квадратный корень из дисперсии), $\overline{X}$ – выборочное среднее, то есть

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i\,.$$

Это точечная оценка для неизвестного математического ожидания a. Но нам нужно построить доверительный интервал. Для этого можно рассмотреть случайную величину

$$Z = \frac{\overline{X} - a}{\sigma_0}\sqrt{n}.$$

Можно убедиться в том, что она имеет стандартное нормальное распределение. Действительно, суммирование нормальных величин, добавление константы или умножение на число не портят нормальность, а из независимости случайных величин $X_1, ..., X_n$ и линейности математического ожидания и дисперсии следует, что $\mathbb{E}Z = 0,\ \mathbb{D}Z = 1$, то есть $Z \sim \mathcal{N}(0, 1)$. Таким образом, эта случайная величина задаётся в зависимости от неизвестного параметра a, но её распределение от него не зависит. Мы можем вычислить вероятность принадлежности этой величины любому заданному интервалу (b, c):

$$\mathbb{P}(b < Z < c) = \Phi(c) - \Phi(b),$$

где $\Phi(x)$ – функция стандартного нормального распределения. Возьмём симметричный интервал $(-c, c),\ \ c > 0$, и приравняем вероятность к γ:

$$\gamma = \mathbb{P}(-c < Z < c) = \Phi(c) - \Phi(-c) = 2\Phi(c) - 1.$$

Здесь мы использовали свойство о том, что $\Phi(-x) = 1 - \Phi(x)$ (объясняется оно тем, что стандартное нормальное распределение симметрично, поэтому вероятность оказаться левее $-c$ – это всё равно что вероятность попасть правее c). Из равенства $\gamma = 2\Phi(c) - 1$ следует, что

$$c = \Phi^{-1}\left(\frac{1+\gamma}{2}\right).$$

Здесь $\Phi^{-1}(y)$ – обратная функция к $\Phi(x)$ или квантиль стандартного нормального распределения, то есть значение, которое стандартная нормальная случайная величина не превышает с вероятностью y.

Вернёмся к равенству для вероятности:

$$\gamma = \mathbb{P}(-c < Z < c) \quad \Leftrightarrow \quad \gamma = \mathbb{P}\left(-c < \frac{\overline{X} - a}{\sigma_0}\sqrt{n} < c\right) \quad \Leftrightarrow$$

$$\Leftrightarrow \quad \gamma = \mathbb{P}\left(\overline{X} - \frac{\sigma_0}{\sqrt{n}}\, c < a < \overline{X} + \frac{\sigma_0}{\sqrt{n}}\, c\right).$$

Число c известно и зависит от γ. Получаем интервал, границы которого являются случайными величинами и который накрывает неизвестное число a с наперёд заданной вероятностью γ:

$$\left(\overline{X} - \frac{\sigma_0}{\sqrt{n}}\, c,\ \overline{X} + \frac{\sigma_0}{\sqrt{n}}\, c\right), \quad c = \Phi^{-1}\left(\frac{1+\gamma}{2}\right).$$