Лабораторная 1

Вариант 1 (Задание 1)

Вариант 3 (Задание 2)

Ибрахим Ахмад M33391

Задание 1

$X_1 \dots X_n \sim Bern(p)$, нужно найти минимальное n, при котором $P(|EX_1 - \overline{X}| \geq \epsilon) \leq \delta$, где $\epsilon = 0.1$ и $\delta = 0.05$.

Вспомним условия ЦПТ:

$\{Y_n\}$ — i.i.d., $EY_i = \mu < \infty$, $DY_i = \sigma^2 < \infty$; $Z_n = \frac{S_n - n\mu}{\sigma\sqrt{n}}$, $Z \sim N(0,1)$. Тогда $Z_n \xrightarrow{d} Z$

Попытаемся преобразовать $P(|EX_1 - \overline{X}| \leq x)$ так, чтобы можно было выразить это в терминах ЦПТ. Положим $Y_i = X_i$ в условии теоремы.

$$P(|\underbrace{EX_1}_{\mu} - \underbrace{\overline{X}}_{S_n/n}| \leq x) = P(|n\mu - S_n| \leq nx) = P\left(\left|\frac{n\mu - S_n}{\sigma\sqrt{n}}\right| \leq \frac{x\sqrt{n}}{\sigma}\right)$$

$$\begin{aligned} &= F_{Z_n}\left(\frac{x\sqrt{n}}{\sigma}\right) - F_{Z_n}\left(-\frac{x\sqrt{n}}{\sigma}\right) \\ &\approx \Phi\left(\frac{x\sqrt{n}}{\sigma}\right) - \Phi\left(-\frac{x\sqrt{n}}{\sigma}\right) \\ &= 2\Phi\left(\frac{x\sqrt{n}}{\sigma}\right) - 1 \end{aligned}$$

Отлично, теперь выразим через это приближение исходное условие:

$$P(|EX_1 - \overline{X}| \geq \epsilon) \leq \delta$$

$$1 - P(|EX_1 - \overline{X}| \leq \epsilon) \leq \delta$$

$$P(|EX_1 - \overline{X}| \leq \epsilon) \geq 1 - \delta$$

$$2\Phi\left(\frac{\epsilon\sqrt{n}}{\sigma}\right) - 1 \geq 1 - \delta$$

$$\Phi\left(\frac{\epsilon\sqrt{n}}{\sigma}\right) \geq 1 - \delta/2$$

По строгому возрастанию Φ, можем применить Φ^{-1} к неравенству:

$$\frac{\epsilon\sqrt{n}}{\sigma} \geq \Phi^{-1}(1 - \delta/2)$$

$$\sqrt{n} \geq \frac{\sigma}{\epsilon}\Phi^{-1}(1 - \delta/2)$$

$$n \geq \left(\frac{\sigma}{\epsilon}\Phi^{-1}(1 - \delta/2)\right)^2$$

Т.к. мы пытаемся найти наименьшее натуральное n, то можно явно выразить его из неравенства выше:

$$n = \left\lceil \left(\frac{\sigma}{\epsilon}\Phi^{-1}(1 - \delta/2)\right)^2 \right\rceil$$

Вспомним, что $\mu = p$, $\sigma^2 = p - p^2$ и подставим в формулу $p = 0.4$, $\epsilon = 0.1$ и $\delta = 0.05$:

```
>>> from scipy.stats import norm
>>> p = 0.4
>>> eps = 0.1
>>> delta = 0.05
>>> mu = p
>>> sigma = math.sqrt(p - p*p)
>>> n = ((sigma/eps)*norm.ppf(1 - delta/2))**2
>>> print(n)
92.195011696659
```

$$n = \lceil 92.195011696659 \rceil = 93$$

Ответ: 93.

Анализ выборок приведён в `task-1.ipynb`

Задание 2

Решение описано в `task-2.ipynb`