Задача 2.2

Агафонов Артём

- Из курса случайных процеесов знаем, что время работы какого либо прибора до поломки описывается экспоненциальным распределением. Его дискретный аналог – геометрическое распределение. Действительно, логично предположить, что каждый день автобус может сломаться равновероятно. Тогда вероятость поломки на i - ый день:

$$\mathbb{P}\{X = i\} = (1-p)^{i-1}p,$$

где p – вероятность, что автобус сломается в каждый конкретный день.

- Формально задачу можно записать в следующем виде

$$X^n = \{3, 22, 13, 6, 18, 5, 6, 10, 7, 15\}$$
$$X \sim Geom(p)$$
$$H_0 : \ \mathbb{V}x = 9$$
$$H_1 : \ \mathbb{V}x > 9.$$

Переформулируем её эквивалентным образом

$$X^n = \{3, 22, 13, 6, 18, 5, 6, 10, 7, 15\}$$
$$X \sim Geom(p)$$
$$H_0 : \ p = \frac{\sqrt{37}-1}{18} = p_0$$
$$H_1 : \ p < \frac{\sqrt{37}-1}{18},$$

где мы воспользовались тем, что $\mathbb{V}X = \dfrac{1-p}{p^2}$.

- Будем пользоваться критерием меток (критерий Вальда плохо работает на конечных выборках, а для критерия χ^2 у нас слишком маленькая выборка). Статистика критерия

$$Z_S(X^n) = \frac{S(p_0)}{\sqrt{I(p_0)}}.$$

Вычислим $S(p)$

$$S(p) = \frac{\partial}{\partial p} \log L(p) = \frac{\partial}{\partial p} \sum_{i=0}^{n} \log p(1-p)^{X_i - 1} =$$
$$= \frac{\partial}{\partial p} \left(\sum_{i=0}^{n} \log p + \sum_{i=0}^{n} (X_i - 1) \log(1-p) \right) = \frac{n}{p} + \sum_{i=0}^{n} (X_i - 1) \frac{-1}{1-p}.$$

Вычислим $I(p)$

$$I(p) = -\mathbb{E}\left(\frac{\partial S(p)}{\partial p}\right) = -\mathbb{E}\left(-\frac{n}{p^2} + \sum_{i=0}^{n} (X_i - 1) \frac{1}{(1-p)^2}\right) =$$
$$= \frac{n}{p^2} + \frac{n}{p(1-p)^2} - \frac{n}{(1-p)^2} = n\frac{1-p}{p^2(1-p)^2} = n\frac{1}{p^2(1-p)},$$

где мы воспользовались тем, что $\mathbb{E}X = \dfrac{1}{p}$.

- Для нашей выборки $X^n = \{3, 22, 13, 6, 18, 5, 6, 10, 7, 15\}$ и $p_0 \approx 0.28$

$$S(p_0) \approx \frac{10}{0.28} - \frac{105}{0.72} + \frac{10}{0.72} = -96.23$$
$$I(p_0) \approx 10 \times \frac{1}{0.28^2 \times 0.72} = 177.15$$
$$Z_S(p_0) \approx -7.23$$

Для критерия меток известно, что $Z_S \sim \mathcal{N}(0, 1)$. 0.05-квантиль стандартного нормального распределения есть -1.64, что больше -7.23. Следовательно, гипотезу отвергаем.

- Построим доверительный интервал Уилсона, основанный на критерии меток.

$$\frac{\hat{p} + z_{1-\frac{\alpha}{2}}/2}{n + z_{1-\frac{\alpha}{2}}} \pm \frac{z_{1-\frac{\alpha}{2}}\sqrt{n}}{n + z^2_{1-\frac{\alpha}{2}}} \sqrt{\hat{p}(1-\hat{p}) + \frac{z^2_{1-\frac{\alpha}{2}}}{4n}},$$

где $\hat{p}$ – оценка максимлального правдоподобия, $z_{1-\frac{\alpha}{2}} = z_{0.975} \approx$

$\approx 1.96,\ n = 10.$

Вычислим $\hat{p}$

$$S(\hat{p}) = 0 \Leftrightarrow \frac{n}{\hat{p}} - \frac{\sum\limits_{i=0}^{n} X_i}{1-\hat{p}} + \frac{n}{1-\hat{p}} = 0$$

$$(1-\hat{p})n - \hat{p}\sum_{i=0}^{n} X_i + \hat{p}n = 0$$

$$\hat{p} = \frac{n}{\sum\limits_{i=0}^{n} X_i} \approx 0.095.$$

Если подставить эти значения, то получим, что доверительный интервал для p есть $[-0.1, 0.28]$. Отсюда можем получить интервал на дисперсию, т.к. это монотонная и непрерывная функция от p.

95%-доверительный интервал для дисперсии есть $[9.2, 110]$.