Lab 1, task 2.3

Талгат Сапаров

22 марта 2020 г.

Предположим, что каждый день вероятность поломки равна фиксированному числу p. Тогда выборка описывается геометрическим распределением с параметром p, $\mathbb{P}(X = n) = (1-p)^{n-1}p$, $\mathbb{E}X = \dfrac{1}{p}$, $\mathbb{V}X = \dfrac{1-p}{p^2}$. В таком случае задачу можно поставить следующим образом:

$$\begin{gathered} X \sim \text{Geom}(p) \\ H_0 : \mathbb{V}X = 9 \\ H_1 : \mathbb{V}X > 9 \end{gathered}$$

Параметр p однозначно получается из значения дисперсии, так как $f(p) = \dfrac{1-p}{p^2}$ - монотонно убывающая функция на отрезке $[0,1]$. Поэтому дисперсия однозначно задает выборочное распределение. При $\mathbb{V}X = 9 \quad p = p_0 = \dfrac{-1+\sqrt{37}}{18} \approx 0.28$. Учитывая вышенаписанное, переформулируем задачу:

$$\begin{gathered} X \sim \text{Geom}(p) \\ H_0 : p = p_0 \\ H_1 : p < p_0 \end{gathered}$$

В качестве критерия выберем критерий меток $Z(X) = \dfrac{S(p_0)}{\sqrt{I(p_0)}} \sim N(0,1)$.

$$\log L(X^n, p) = \sum_{i=1}^{n} \log p(1-p)^{X_i - 1} = n \log p + \sum_{i=1}^{n} (X_i - 1)\log(1-p)$$

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

$$I(p) = -\mathbb{E}\left[-\frac{n}{p^2} - \frac{\sum\limits_{i=1}^{n} X_i}{(1-p)^2} + \frac{n}{(1-p)^2}\right] = n\left(\frac{1}{p^2} + \frac{1}{p(1-p)^2} - \frac{1}{(1-p)^2}\right) =$$

$$n\left(\frac{(1-p)^2 + p - p^2}{p^2(1-p)^2}\right) = n\frac{1-p}{p^2(1-p)^2} = \frac{n}{p^2(1-p)}$$

Для нашей реализации выборки $x = (3, 22, 13, 6, 18, 5, 6, 10, 7, 15)$, $n = 10, \sum\limits_{i=1}^{n} x_i = 105$. Тогда

$$S(p_0) \approx -96.23, I(p_0) \approx 177.15, Z(X) \approx -7.23$$

Гипотезу отклоняем, так как $Z(X) \approx -7.23 < z_{0.05} \approx -1.645$ (учитываем то, что альтернатива односторонняя).
Доверительный интервал построим, используя критерий Вальда $[p_{MLE} - z_{1-\alpha/2}\sqrt{I^{-1}(p_{MLE})}, p_{MLE} + z_{1-\alpha/2}\sqrt{I^{-1}(p_{MLE})}]$. Оценку максимального правдоподобия найдем из уравнения $S(p) = 0$:

$$\frac{n}{p} - \frac{\sum\limits_{i=1}^{n} X_i}{1-p} + \frac{n}{1-p} = 0$$

$$n - np = p\sum_{i=1}^{n} X_i - np$$

$$p_{MLE} = \frac{n}{\sum\limits_{i=1}^{n} X_i} \approx 0.095$$

$$I(p_{MLE}) = \frac{n}{p_{MLE}^2(1 - p_{MLE})} \approx 1218.55$$

$$z_{0.975} \approx 1.96$$

Доверительный интервал для p при $\alpha = 0.05$ — $[0.039, 0.151]$. Так как дисперсия монотонно непрерывно связана с p, то доверительный интервал для дисперсии на том же уровне будет образом интервала для p под действием связывающей их функции $f(p) = \dfrac{1-p}{p^2}$. Итого, доверительный интервал для дисперсии на требуемом уровне — $[37.2, 631.]$.