Sheet 10

Due: To be handed in before 12.05.2023 at 12:00.

1. Exercise

We observe i.i.d. random variables $X_1, \dots, X_n$ from a Poisson distribution with an unknown intensity $\lambda \in \Theta = (0, \infty)$. We want to test

$$H_0\colon \lambda = \lambda_0 \qquad \text{versus} \qquad H_1\colon \lambda = \lambda_1 \tag{1}$$

for some $\lambda_1 > \lambda_0$. We choose $\alpha = 0.05$.

(a) Recall the density of $X \sim \text{Pois}(\lambda)$ (with respect to the counting measure).

(b) Recall the Neyman-Pearson test for simple hypotheses at level α.

(c) Find the NP-test for the testing problem in (1) at level α.

(d) Using the result that $\sum_{i=1}^n X_i \sim \text{Pois}\left(\sum_{i=1}^n \lambda_i\right)$ if $X_i \sim \text{Pois}(\lambda_i)$ and $X_1, \dots, X_n$ are independent, give the NP-test more explicitly if you know that $n = 30$, $\lambda_0 = 40$, and if $Y \sim \text{Pois}(1200)$.

x	$\mathbb{P}(Y \leq x)$
1254	0.9413
1255	0.9446
1256	0.9477
1257	0.9506
1258	0.9535
1259	0.9562
1260	0.9587

(e) The manager of a bakery would like to employ an additional person for the morning shift (7a.m. – 11a.m.). To make a case, he needs to prove that the rate at which the customers come during this period is equal to 45. Based on data collected in the past, this rate was until now assumed to be 40. During the month of April, the manager counted the number of customers. He got the following data:

$$47, 43, 50, 48, 44, 58, 43, 39, 60, 48, 51, 35, 47, 48, 45, 49, 48, 34, 49, 40, 42, 37, 40, 51, 53, 33, 45, 43, 43, 39.$$

Can the manager build their case for employing an additional person?

Solution:

(a) The density of $X \sim \text{Pois}(\lambda)$ with respect to the counting measure (or the probability mass function) is $p_\lambda(x) = \frac{1}{x!} e^{-\lambda} \lambda^x$, $x = 0, 1, 2, \dots$.

(b) Let p_0 and p_1 be the densities of the observed data X under H_0 and H_1, respectively. The NP-test is given by

$$\phi_{\text{NP}}(X) = \begin{cases} 1 & \text{if } \frac{p_1(X)}{p_0(X)} > c_\alpha, \\ q_\alpha & \text{if } \frac{p_1(X)}{p_0(X)} = c_\alpha, \\ 0 & \text{if } \frac{p_1(X)}{p_0(X)} < c_\alpha, \end{cases}$$

where $c_\alpha = (1-\alpha)$-quantile of the distribution of $\frac{p_1(X)}{p_0(X)}$ under H_0 and q_α satisfies $\mathbb{E}_{p_0}[\phi_{\text{NP}}(X)] = \alpha$, i.e.

$$\mathbb{P}_{p_0}\left(\frac{p_1(X)}{p_0(X)} > c_\alpha\right) + q_\alpha \mathbb{P}_{p_0}\left(\frac{p_1(X)}{p_0(X)} = c_\alpha\right) = \alpha.$$

(c) In this case, the observed "X" is the whole random sample $(X_1, \dots, X_n)$.
Under H_0: $p_0(x_1, \dots, x_n) = \prod_{i=1}^n \frac{1}{x_i!} e^{-\lambda_0} \lambda_0^{x_i}$
Under H_1: $p_1(x_1, \dots, x_n) = \prod_{i=1}^n \frac{1}{x_i!} e^{-\lambda_1} \lambda_1^{x_i}$

Then,

$$\frac{p_1(x_1,\dots,x_n)}{p_0(x_1,\dots,x_n)} = \frac{e^{-n\lambda_1}\lambda_1^{\sum_{i=1}^n x_i}}{e^{-n\lambda_0}\lambda_0^{\sum_{i=1}^n x_i}} = e^{-n(\lambda_1-\lambda_0)}\left(\frac{\lambda_1}{\lambda_0}\right)^{\sum_{i=1}^n x_i} = g\left(\sum_{i=1}^n x_i\right),$$

where $g(t) = e^{-n(\lambda_1-\lambda_0)}(\lambda_1/\lambda_0)^t$, which is continuous and strictly increasing on $(0,\infty)$ since $\lambda_1/\lambda_0 > 1$. Thus, the NP-test can also be given as

$$\phi_{\mathrm{NP}}(X) = \begin{cases} 1 & \text{if } \sum_{i=1}^n X_i > t_\alpha, \\ q_\alpha & \text{if } \sum_{i=1}^n X_i = t_\alpha, \\ 0 & \text{if } \sum_{i=1}^n X_i < t_\alpha, \end{cases}$$

where $t_\alpha = (1-\alpha)$-quantile of the distribution of $\sum_{i=1}^n X_i$ under H_0 and q_α is such that

$$\mathbb{P}_{p_0}\left(\sum_{i=1}^n X_i > t_\alpha\right) + q_\alpha \mathbb{P}_{p_0}\left(\sum_{i=1}^n X_i = t_\alpha\right) = \alpha$$

with $\alpha = 0.05$.

(d) We have $t_\alpha = 1257$ and

$$q_\alpha = \frac{-(1-\alpha) + \mathbb{P}(Y \le 1257)}{\mathbb{P}(Y = 1257)} = \frac{-0.95 + 0.9506}{0.9506 - 0.9477} = 0.207,$$

using $\mathbb{P}(Y = x) = \mathbb{P}(Y \le x) - \mathbb{P}(Y \le x-1)$.

(e) We want to test

$$H_0\colon \lambda = 40 \qquad \text{versus} \qquad H_1\colon \lambda = 45.$$

We have $\sum_{i=1}^n X_i = 1352 > 1257$, so the decision is: reject H_0 in favor of H_1. The manager can indeed build a case.

2. Exercise

Let X be a random variable such that $X(\Omega) = \mathcal{X} = \{1,2,3,4,5,6,7\}$. Under some null and alternative hypotheses H_0 and H_1, we know that the probability mass function f of X is f_0 and f_1, respectively, and

x	1	2	3	4	5	6	7
$f_0(x)$	0.01	0.01	0.01	0.01	0.01	0.01	0.94
$f_1(x)$	0.06	0.05	0.04	0.03	0.02	0.01	0.79

We want to test

$$H_0\colon f = f_0 \qquad \text{versus} \qquad H_1\colon f = f_1.$$

We choose $\alpha = 0.04$.

(a) What is the NP-test of level α for this testing problem?

(b) Compute the power of the NP-test found in (a).

Solution:

(a) It is

$$\phi_{\mathrm{NP}}(X) = \begin{cases} 1 & \text{if } \frac{f_1(X)}{f_0(X)} > c_\alpha, \\ q_\alpha & \text{if } \frac{f_1(X)}{f_0(X)} = c_\alpha, \\ 0 & \text{if } \frac{f_1(X)}{f_0(X)} < c_\alpha, \end{cases}$$

where $c_\alpha = (1-\alpha)$-quantile of the distribution of $\frac{f_1(X)}{f_0(X)}$ and q_α satisfies

$$\mathbb{E}_{f_0}[\phi_{\mathrm{NP}}(X)] = \mathbb{P}_{f_0}\left(\frac{f_1(X)}{f_0(X)} > c_\alpha\right) + q_\alpha \mathbb{P}_{f_0}\left(\frac{f_1(X)}{f_0(X)} = c_\alpha\right) = \alpha.$$

Now, the random variable $\frac{f_1(X)}{f_0(X)}$ takes on the following values:

$$\frac{f_1(X)}{f_0(X)} = \begin{cases} 6 & \text{if } X = 1, \\ 5 & \text{if } X = 2, \\ 4 & \text{if } X = 3, \\ 3 & \text{if } X = 4, \\ 2 & \text{if } X = 5, \\ 1 & \text{if } X = 6, \\ \frac{79}{94} & \text{if } X = 7. \end{cases}$$

Under f_0, the random variable $\frac{f_1(X)}{f_0(X)}$ has the following probability mass function:

$$\begin{aligned} \mathbb{P}_{f_0}\left(\frac{f_1(X)}{f_0(X)} = 6\right) &= \mathbb{P}_{f_0}(X=1) = 0.01, \\ \mathbb{P}_{f_0}\left(\frac{f_1(X)}{f_0(X)} = 5\right) &= \mathbb{P}_{f_0}(X=2) = 0.01, \\ \mathbb{P}_{f_0}\left(\frac{f_1(X)}{f_0(X)} = 4\right) &= \mathbb{P}_{f_0}(X=3) = 0.01, \\ \mathbb{P}_{f_0}\left(\frac{f_1(X)}{f_0(X)} = 3\right) &= \mathbb{P}_{f_0}(X=4) = 0.01, \\ \mathbb{P}_{f_0}\left(\frac{f_1(X)}{f_0(X)} = 2\right) &= \mathbb{P}_{f_0}(X=5) = 0.01, \\ \mathbb{P}_{f_0}\left(\frac{f_1(X)}{f_0(X)} = 1\right) &= \mathbb{P}_{f_0}(X=6) = 0.01, \\ \mathbb{P}_{f_0}\left(\frac{f_1(X)}{f_0(X)} = \frac{79}{94}\right) &= \mathbb{P}_{f_0}(X=7) = 0.94. \end{aligned}$$

Thus, the cdf $y \mapsto \mathbb{P}_{f_0}(f_1(X)/f_0(X) \le y) = F_0(y)$ is given by

$$F_0(y) = \begin{cases} 0 & \text{if } y < \frac{79}{94}, \\ 0.94 & \text{if } \frac{79}{94} \le y < 1, \\ 0.95 & \text{if } 1 \le y < 2, \\ 0.96 & \text{if } 2 \le y < 3, \\ 0.97 & \text{if } 3 \le y < 4, \\ 0.98 & \text{if } 4 \le y < 5, \\ 0.99 & \text{if } 5 \le y < 6, \\ 1 & \text{if } y \ge 6, \end{cases}$$

For $\alpha = 0.04$, the 0.96-quantile of F_0 is 2. Thus, q_α should satisfy

$$\mathbb{P}_{f_0}\left(\frac{f_1(X)}{f_0(X)} > 2\right) + q_\alpha \mathbb{P}_{f_0}\left(\frac{f_1(X)}{f_0(X)} = 2\right) = 0.04.$$

Since $\mathbb{P}_{f_0}(f_1(X)/f_0(X) > 2) = 1 - F_0(2) = 0.04$ and $\mathbb{P}_{f_0}(f_1(X)/f_0(X) = 2) = 0.01 \neq 0$, it follows that $q_\alpha = 0$. Hence,

$$\phi_{\text{NP}}(X) = \begin{cases} 1 & \text{if } \frac{f_1(X)}{f_0(X)} > 2, \\ 0 & \text{if } \frac{f_1(X)}{f_0(X)} \leq 2. \end{cases}$$

Now, note that $f_1(X)/f_0(X) > 2$ if and only if $X \leq 4$, so

$$\phi_{\text{NP}}(X) = \begin{cases} 1 & \text{if } X \leq 4, \\ 0 & \text{if } X > 4. \end{cases}$$

(b) Let us denote this power by β. By definition, we have

$$\beta = \mathbb{E}_{f_1}[\phi_{\text{NP}}(X)] = \mathbb{P}_{f_1}(X \leq 4) = \mathbb{P}_{f_1}(X = 1) + \mathbb{P}_{f_1}(X = 2) + \mathbb{P}_{f_1}(X = 3) + \mathbb{P}_{f_1}(X = 4) = 0.18.$$