Probability-2 - Assignment #2

Daniil Buchko - dbuchko@nes.ru

November 17, 2021

Question 1

Suppose that you have a random sample $X_1, \dots, X_n$ from a distribution with the density

$$f(x) = \begin{cases} \dfrac{1}{\theta^2} x e^{-x/\theta}, & 0 \leq x < \infty \\ 0 \quad \text{else} \end{cases}$$

(a) Construct the Maximum Likelihood estimator $\hat{\theta}_{ML}$
(b) Compute Fisher information for your estimator. Compute the variance of $\hat{\theta}_{ML}$
(c) Using the result of the previous step, design a test for the hypothesis $H_0 : \theta \leq \theta_0$ against $H_1 : \theta > \theta_0$

(a) Defining likelihood function:

$$\mathcal{L} = \prod_{i=1}^{n} \frac{1}{\theta^2} x_i e^{-x_i/\theta} \to \max_{\theta}$$

Loglikelihood function:

$$l = -2n \ln \theta + \sum_{i=1}^{n} \ln x_i - \frac{x_i}{\theta} \to \max_{\theta}$$

Finding derivatives:

$$\frac{\partial l}{\partial \theta} : \quad \frac{-2n}{\hat{\theta}} + \sum_{i=1}^{n} \frac{x_i}{\hat{\theta}^2} = 0$$

$$\hat{\theta}_{ML} = \overline{X}/2$$

(b)

$$\ln f(x,\theta) = -2\ln\theta + \ln x - \frac{x}{\theta}$$

$$\frac{\partial \ln f(x,\theta)}{\partial\theta} = \frac{-2}{\theta} + \frac{x}{\theta^2} \implies \frac{\partial \ln^2 f(x,\theta)}{\partial\theta^2} = \frac{2}{\theta^2} - \frac{2x}{\theta^3}$$

$$I(\theta) = -\mathbb{E}_\theta\left[\frac{\partial \ln^2 f(x,\theta)}{\partial\theta^2}\right] = \frac{2}{\theta^2} - \frac{2}{\theta^3}\mathbb{E}(X) = \frac{2}{\theta^2}$$

It is clear that $Var(\hat{\theta}) = \theta^2/2n$

(c) Using the asympotical normality of $\hat{\theta}_{ML}$ we can write that

$$\mathbb{P}\left(\frac{\overline{X}}{2} > \underbrace{A\frac{\sigma}{\sqrt{n}} + \theta_0}_{c}\right) = 1 - \Phi(A) \implies A = \frac{c-\theta_0}{\sigma/\sqrt{n}}$$

And thus we may claim that

$$\mathbb{P}\left(\frac{\overline{X}}{2} > c\right) = 1 - \Phi\left(\frac{c-\theta_0}{\sigma/\sqrt{n}}\right)$$

where c depends on the significance level. Thus we reject H_0 if $\overline{X}/2 > c$ and not reject otherwise.

Question 2

Suppose that you have a random sample from a normal distribution $\mathcal{N}(\mu,\sigma^2)$ where μ and σ are parameters

(a) If we know $\sigma = 1$ and:

$$\begin{cases} H:\ 0 & \mu = 1 \\ H:\ 1 & \mu > 1 \end{cases}$$

$$\mathbb{P}\left(\overline{X} > \underbrace{\frac{A}{\sqrt{n}} + \mu}_{c}\right) = 1 - \Phi(A)$$

$$\overline{X} > \frac{z_a}{\sqrt{n}} + 1$$

(b) Power function:

$$\gamma(\mu) = \mathbb{P}\left(\overline{X} > \frac{z_a}{\sqrt{n}} + 1\right) = \mathbb{P}\left(\frac{\overline{X}-\mu}{1/\sqrt{n}} > \sqrt{n}\left[\frac{z_a}{\sqrt{n}} + 1 - \mu\right]\right) \implies 1 - \Phi\left(z_a + \sqrt{n} - \mu\sqrt{n}\right)$$

Question 4

Suppose that your friend has collected some data in hope to reject a certain hypothesis H_0 at the conventional confidence level 5%. However, the data did not let him to reject it with the p-value of 11%. Then he collected another sample (that is independent from the first one), obtained the p-value of 4%, and reported that H_0 is rejected at the 5% level. If you think of the procedure that he implemented, what is the p-value that should be have been reported? Explain.

In order to avoid doing such a mistake we need to determine the proper sample size before doing the trial. We can do Power analysis, which is done before the experiment and tells us how many replicates we need in order to do have a relatively high probability of correctly rejecting the null hypothesis.