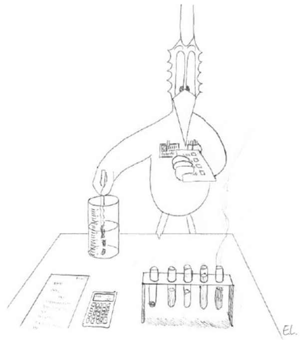
PERFORMANCE EVALUATION EXERCISES

TESTS

With Solutions Jean-Yves Le Boudec, Spring 2021



1. If the data is in the critical region we...

   (a) ☐ Accept H_0
   (b) ☒ Reject H_0
   (c) ☐ It depends on the nature of the test
   (d) ☐ It depends on the size of the test

2. Saying that a test is of size 5% means that...

   (a) ☐ The probability to accept H_0 when H_0 does not hold is ≤ 0.05
   (b) ☒ The probability to reject H_0 when H_0 it holds is ≤ 0.05
   (c) ☐ Both

3. If the $p-$value of a test is small we ...

   (a) ☐ Accept H_0
   (b) ☒ Reject H_0
   (c) ☐ It depends on the nature of the test
   (d) ☐ It depends on the size of the test

4. We have a collection of random variables X_i, Y_i which correspond to non paired simulation results with configuration 1 or 2. How can you test whether the configuration plays a role or not ?

   (a) ☐ With a Wilcoxon Rank Sum test
   (b) ☐ With an ANOVA test
   (c) ☒ With either
   (d) ☒ With none

   Solution. (a) is robust and can be used if we can ensure that the simulation runs are independent and that the two distributions differ by a location shift, i.e. have same variance.

   (b) is applicable if, in addition, X_i and Y_i can be assumed gaussian with same variance.

5. We test whether a distribution is gaussian using a Kolmogorov-Smirnov test against the fitted distribution. We obtain a $p-$value.

   (a) ☒ The true $p-$value is smaller

(b) ☐ We have obtained the true $p-$value

(c) ☐ The true $p-$value is larger

(d) ☐ It depends on the data

Solution. The KS test applies if we are testing against a fixed, non fitted distribution F. By using a fitted distribution, we are biasing the test, we are making it more likely than should be to accept the distribution F, i.e. to accept H_0. The $p-$value should therefore be higher (since a small $p-$value means rejecting H_0).

6. We have two data sets X_i and Y_j believed to be iid and from one exponential distribution each. We want to test whether the parameter of their exponential distribution is the same.

   Give the design of a corresponding likelihood ratio test. Give a formula for the $p-$value when m, n are large.

   Solution. The generic model is $X_i \sim$ iid Exp(λ), $Y_j \sim$ iid Exp(μ) and $X_i.Y_i$ independent for $i = 1 : m, j = 1 : n$ and for some $\lambda, \mu > 0$. The parameter is $\theta = (\lambda, \mu)$ and $\Theta = \{(\lambda, \mu) : \lambda > 0 \text{ and } \mu > 0\}$ H_0 corresponds to $\lambda = \mu$.

   The PDF of the Exp(λ) distribution is $f_X(x) = \lambda e^{-\lambda x}$ hence the likelihood of the data $(\vec{x}, \vec{y})$ is

$$f_{\vec{X},\vec{Y}}(\vec{x}, \vec{y}) = \prod_{i=1}^{m} \left(\lambda e^{-\lambda x_i}\right) \times \prod_{j=1}^{n} \left(\mu e^{-\mu y_j}\right) \tag{1}$$

   and the log-likelihood is

$$m \log \lambda + n \log \mu - \lambda \sum_{i=1}^{m} x_i - \mu \sum_{j=1}^{n} y_j \tag{2}$$

   We use the following notation: $x^* = \frac{1}{m} \sum_i x_i$, $y^* = \frac{1}{n} \sum_j y_j$ and $z^* = \frac{1}{m+n} \left(\sum_i x_i + \sum_j y_j\right)$.

   We compute the optimal value of the log likelihood under H_0, namely under the constraint $\lambda = \mu$:

$$\min_{\lambda > 0} \left(m \log \lambda + n \log \lambda - \lambda \sum_{i=1}^{m} x_i - \lambda \sum_{j=1}^{n} y_j \right)$$

   Taking the derivative we find that the optimum is for $\lambda = \mu = \frac{\sum_i x_i + \sum_j y_j}{m+n} = \frac{1}{z^*}$ and the value of the optimal log-likelihood under H_0 is

$$\begin{aligned} \ell_0 &= (m+n) \log \frac{1}{z^*} - m - n \\ &= -(m+n) \left(\log z^* + 1\right) \end{aligned}$$

   Under H_1 the optimal is for $\lambda = \frac{1}{x}$ and $\mu = \frac{1}{y}$. The value of the optimal log-likelihood under H_1 is

$$\begin{aligned} \ell_1 &= m \left(\log \frac{1}{x^*} - 1\right) + n \left(\log \frac{1}{y^*} - 1\right) \\ &= -m \left(\log x^* + 1\right) - n \left(\log y^* + 1\right) \end{aligned}$$

   The log-likelihood ratio statistic is

$$lrs = \ell_1 - \ell_0 = (m+n) \log z^* - m \log x^* - n \log y^*$$

Apply theorem 4.3: $q_2 = 2 - 1 = 1$ hence the distribution of $2lrs$ under H_0 is approximately χ_1^2 and the p value is approximately

$$p = 1 - F_{\chi_1^2}(2lrs)$$

where $F_{\chi_1^2}$ is the CDF of χ_1^2.

7. We have some data set $\vec{Y} = Y_{i=1:I}$ modelled with a parametric model with $\theta \in \Theta$. Let $f_{\vec{Y}}(\vec{y}|\theta)$ be the PDF of the observation $\vec{y} = y_{1:I}$. We assume that we have a method to compute $\hat{\theta}(\vec{y})$, the maximum likelihood estimator of θ for value of the data set $\vec{y}$.

   (a) Give a likelihood ratio test for the test

   $H_0 : \theta = \theta_0$ versus $H_1 : \theta \in \Theta$

   (b) Give the pseudo-code of an algorithm to compute the $p-$value of this test using Monte-Carlo simulation with R runs.

   (c) We run this algorithm with $R = 10'000$ and find $p = 0$. Give a 99% confidence for the true $p-$value. What can we conclude at a size of 5% ?

Solution.

(a) Under H_0 the MLE is $\hat{\theta} = \theta_0$ (a fixed value) and the log-likelihood is

$$\ell_0 = \log f_{\vec{Y}}(\vec{y}|\theta_0)$$

Under H_1 the optimal likelihood is obtained for $\theta = \hat{\theta}(\vec{y})$ and thus

$$\ell_1 = \log f_{\vec{Y}}(\vec{y}|\hat{\theta}(\vec{y}))$$

and

$$lrs(\vec{y}) = \log f_{\vec{Y}}(\vec{y}|\hat{\theta}(\vec{y})) - \log f_{\vec{Y}}(\vec{y}|\theta_0)$$

The test has a rejection region of the form $lrs(\vec{y}) > C$ for some constant C that depends on the size of the test. The $p-$value is thus

$$p = \mathbb{P}\left(lrs(\vec{Y}) > lrs(\vec{y})\right) \tag{3}$$

where $\vec{Y}$ is a random sequence that has PDF $f_{\vec{Y}}\left(\cdot|\theta_0\right)$

(b) To compute the $p-$value by Monte-Carlo simulation with R runs we draw R replicates of the sequence $\vec{Y}$ from the PDF $f_{\vec{Y}}\left(\cdot|\theta_0\right)$ and evaluate p as the empirical mean:

tot =0
for $r = 1 : R$ do
    draw $\vec{y}^{\,r}$ from the distribution with PDF $f_{\vec{Y}}\left(\cdot|\theta_0\right)$
    if $\{lrs\left(\vec{y}^{\,r}\right) > lrs(\vec{y})\}$ tot = tot+1
end
return$\left(\frac{\text{tot}}{R}\right)$

(c) By Theorem 2.4, a 99% confidence interval for p is $[0, p_0]$ with $p_0 = 5.3 \times 10^{-4}$. Since the upper bound on the estimated $p-$value is very small, much smaller than 5%, we reject H_0.

8. We consider again the case in the previous question. Using Monte-Carlo simulation, we have obtained a 99% confidence interval $[\ell(\vec{y}), u(\vec{y})]$ for the $p-$value. We reject H_0 if the true p is small, but since we don't know the true $p-$value, we use the rejection condition $u(\vec{y}) < \alpha$. What value of α should we chose to ensure that this way of doing provides a test of size 5% ?

Solution. Let Z represent the sequence of all random numbers we have used to estimate the $p-$value $p(\vec{y})$. The sequence Z is independent of the data $\vec{y}$. The upper bound $u(\vec{y}, Z)$ is true with confidence 99%, therefore

$$\mathbb{P}\left(u(\vec{y}, Z) \leq p(\vec{y})\right) \leq 1\% \tag{4}$$

Furthermore, let $\vec{Y}_0$ be a random variable that has the distribution of $\vec{Y}$ under H_0. By definition of a $p-$value we know that:

$$\mathbb{P}\left(p(\vec{Y}_0) < \alpha\right) = \alpha \tag{5}$$

Also, observe that (4) should hold for any realization of $\vec{y}$, therefore

$$\mathbb{P}\left(u(\vec{Y}_0, Z) \leq p(\vec{Y}_0)\right) \leq 1\% \tag{6}$$

Now our test consists in rejecting H_0 whenever $u(\vec{y}, Z) < \alpha)$. The probability p_I of a type-I error is the probability of rejection when $\vec{y}$ is a sample of $\vec{Y}_0$:

$$\begin{aligned}
p_I &= \mathbb{P}\left(u(\vec{Y}_0, Z) < \alpha\right) \\
&= \mathbb{E}\left(\mathbf{1}_{\{u(\vec{Y}_0,Z)<\alpha\}}\right) \\
&= \mathbb{E}\left(\mathbf{1}_{\{u(\vec{Y}_0,Z)<\alpha\}}\mathbf{1}_{\{u(\vec{Y}_0,Z)>p(\vec{Y}_0)\}} + \mathbf{1}_{\{u(\vec{Y}_0,Z)<\alpha\}}\mathbf{1}_{\{u(\vec{Y}_0,Z)\leq p(\vec{Y}_0)\}}\right) \\
&= \mathbb{E}\left(\mathbf{1}_{\{u(\vec{Y}_0,Z)<\alpha\}}\mathbf{1}_{\{u(\vec{Y}_0,Z)>p(\vec{Y}_0)\}}\right) + \mathbb{E}\left(\mathbf{1}_{\{u(\vec{Y}_0,Z)<\alpha\}}\mathbf{1}_{\{u(\vec{Y}_0,Z)\leq p(\vec{Y}_0)\}}\right) \\
&\leq \mathbb{E}\left(\mathbf{1}_{\{p(\vec{Y}_0)<\alpha\}}\right) + \mathbb{E}\left(\mathbf{1}_{\{u(\vec{Y}_0,Z)\leq p(\vec{Y}_0)\}}\right) \\
&= \mathbb{P}\left(p(\vec{Y}_0) < \alpha\right) + \mathbb{P}\left(u(\vec{Y}_0, Z) \leq p(\vec{Y}_0)\right) \\
&\leq \alpha + 1\%
\end{aligned}$$

Therefore, we need to take $\alpha = 4\%$ to obtain a test of size 5%. The uncertainty due to the Monte Carlo estimation needs to be added to α.