Inferential Statistics **First semester, 2022**

Exercises on Confidence sets and Hypothesis testing

Instructor: Erlis Ruli (ruli@stat.unipd.it), Department of Statistical Sciences

Example 6.1 *Let $X > 0$ be the number of defective items in a shipment of a very large number of items and suppose that $X \sim F_\theta$, with p.d.f. $f(x;\theta) = c_\theta \theta^x$, $x = 1, 2, \dots,$, with $\theta \in (0, 1)$ the unknown parameter and $c_\theta > 0$ a real depending on θ. Consider a sample of size 1 from this model.*

(a) *Discuss a test statistic for $H_0 : \theta \leq \theta_0$ against $H_1 : \theta > \theta_0$ of level or size α; compute the test with $\theta_0 = 1/2$ and $\alpha = 0.0625$. Is this a test of size α?*

(b) *Compute the power of the test when $\theta = 0.9$ and $\theta_0 = 1/2$ and $\alpha = 0.0625$.*

(c) *Derive the Wald confidence interval (CI) for θ of level $1 - \alpha$ with a sample of size n; compute the CI with $\alpha = 0.9$ given an observed sample with $\overline{x} = 10$ and $n = 30$.*

(d) *Derive the likelihood-based confidence set for θ, with $\alpha = 0.9$ given an observed sample with $\overline{x} = 10$ and $n = 30$.*

Solution.

(a) We may use NP Lemma but before that we have to check that the monotone likelihood ratio (MLR) property holds. Thus consider the pair (θ_1, θ_2), with $\theta_1, \theta_2 \in (0, 1)$ and $\theta_1 < \theta_2$, then

$$\frac{L(\theta_2)}{L(\theta_1)} = \frac{1-\theta_2}{1-\theta_1} \left(\frac{\theta_2}{\theta_1}\right)^{X-1}$$

Since $\theta_2/\theta_1 > 1$ and $1 - \theta_i > 0$, then $\frac{L(\theta_2)}{L(\theta_1)}$ is an increasing function of X, thus the distribution satisfies the MLR property.

Note that $\Theta_0 = (0, \theta_0]$ and $\Theta_1 = (\theta_0, 1)$. By the NP Lemma, the size α critical region R must be the set of all samples X such that for any two values $\theta_1 \in \Theta_0$ and $\theta_2 \in \Theta_1$

$$\frac{L(\theta_2)}{L(\theta_1)} = \left(\frac{1-\theta_2}{1-\theta_1}\right) \left(\frac{\theta_2}{\theta_1}\right)^{X-1} \geq k \text{ for all } X \in R,$$

and

$$\frac{L(\theta_2)}{L(\theta_1)} = \left(\frac{1-\theta_2}{1-\theta_1}\right)\left(\frac{\theta_2}{\theta_1}\right)^{X-1} < k \text{ for all } X \notin R.$$

Thus for all samples in R we must have

$$\left(\frac{\theta_2}{\theta_1}\right)^{X-1} \geq k\left(\frac{1-\theta_1}{1-\theta_2}\right) \iff X \geq \log\left(k\left(\frac{1-\theta_1}{1-\theta_2}\right)\right) - \log(\theta_2/\theta_1) + 1$$
$$\iff X \geq c,$$

for some c. The test is thus

$$\text{reject } H_0 \text{ if } X \in R, \text{ where } R = \{X : X \geq c\}$$

and c is the value that solves

$$\alpha \geq \sup_{\theta\in\Theta_0} P_\theta(X \in R) = \sup_{\theta\in\Theta_0} \sum_{x=c}^{\infty}(1-\theta)\theta^{x-1}$$
$$\geq \sup_{\theta\in\Theta_0}\{\theta^{c-1}\}.$$

For $\theta_0 = 1/2$ and $\alpha = 0.065$ we have to solve

$$0.0625 \geq \sup_{\theta\leq 1/2}\{\theta^{c-1}\}.$$

Since c is discrete, we can solve this equation by trying each possible value of c. Thus

- *with* $c = 1$, $\sup_{\theta\leq 1/2}\{\theta^{c-1}\} = 1 > 0.0625$,
- *with* $c = 2$, $\sup_{\theta\leq 1/2}\{\theta\} = 1/2 > 0.0625$,
- *with* $c = 3$, $\sup_{\theta\leq 1/2}\{\theta^2\} = (1/2)^2 > 0.0625$
- *with* $c = 4$, $\sup_{\theta\leq 1/2}\{\theta^3\} = (1/2)^3 > 0.0625$
- *with* $c = 5$, $\sup_{\theta\leq 1/2}\{\theta^4\} = (1/2)^4 = 0.0625$.

A test of size α is thus to reject H_0 if $X \geq 5$. Note that if α was 0.05 then we could only build a test of level α.

(b) The power function is $\gamma(\theta) = P_\theta(X \in R)$ *for all* $\theta \in \Theta_1$*, thus*

$$\gamma(0.9) = \sum_{x=5}^{\infty} 0.1 \cdot 0.9^{x-1} = 0.6561$$

(c) It can be checked that the MLE is $\widehat{\theta} = (\overline{X} - 1)/\overline{X}$ *and by the large sample property of the MLE*

$$\widehat{\theta} \xrightarrow{d} N(\theta, I_n(\theta)^{-1}),$$

with $I_n(\theta) = \theta(1-\theta)^2/n$ *we have the pivot* $\frac{\widehat{\theta}-\theta}{se(\widehat{\theta})} \sim N(0,1)$*, with* $se(\widehat{\theta}) = \sqrt{I_n(\theta)^{-1}}$*. Although this pivot is correct, it is not easy to use since* $se(\widehat{\theta})$ *depends on the unknown parameter* θ*. Thus, instead of* $se(\widehat{\theta})$ *we may prefer to use* $\widehat{se}(\widehat{\theta}) = \sqrt{I_n(\widehat{\theta})^{-1}}$*, so*

$$P_\theta\left(-z_{\alpha/2} \leq \frac{\widehat{\theta}-\theta}{\widehat{se}(\widehat{\theta})} \leq z_{\alpha/2}\right) = P_\theta\left(\widehat{\theta} - z_{\alpha/2}\widehat{se}(\widehat{\theta}) \leq \theta \leq \widehat{\theta} + z_{\alpha/2}\widehat{se}(\widehat{\theta})\right)$$
$$\to 1 - \alpha,$$

and the desired confidence interval is thus $[\theta \pm z_{\alpha/2}\widehat{se}(\widehat{\theta})]$*. With the observed sample we get* $\widehat{\theta} = 0.9$*, and* $\widehat{se}(\widehat{\theta}) \doteq 0.0173$ *and the CI is* $[0.9 \pm 1.645 \cdot 0.0173] = [0.872, 0.928]$*. With the given sample we are 90% confident that the true value* θ *is between 0.871 and 0.928.*

(d) Since there is only one parameter, under the hypothesis $H_0 : \theta = \theta_0$ *against* $H_1 : \theta \neq \theta_0$ *we have the pivot*

$$2 \log \frac{\sup_{\theta \in (0,1)} L(\theta)}{L(\theta_0)} = 2(\ell(\widehat{\theta}) - \ell(\theta_0)) \xrightarrow{d} \chi^2_1 \text{ for all } \theta_0 \in (0,1).$$

We see that $2(\ell(\widehat{\theta}) - \ell(\theta)) = 2((n\overline{X}-1)(\log(\overline{X}-1) - \log\theta) + (1-n-n\overline{X})\log\overline{X} - n\log(1-\theta))$ *thus the* $1-\alpha$ *likelihood-based confidence set is*

$$\{\theta \in (0,1) : 598 \log 9 - 658 \log 10 - 60 \log(1-\theta) - 598 \log\theta \leq 2.701.\}$$

Example 6.2 *Let* $Y_i \overset{\text{iid}}{\sim} \text{Poi}(\theta_1)$, $i = 1, \ldots, n$ *and* $X_j \overset{\text{iid}}{\sim} \text{Poi}(\theta_2)$, $j = 1, \ldots, m$, *with* Y_i, X_j *being independent for all* i, j.

(a) Construct the log-likelihood ratio test for testing $H_0 : \theta_1 = \theta_2$ *against* $H_1 : \theta_1 \neq \theta_2$ *at the level* α. *Compute the test with observed samples* $\overline{y} = 3$, $\overline{x} = 5$,$m = 15, n = 10$, *and get the associated p-value.*

(b) Test the above hypothesis by constructing a suitable Wald CI and apply it to the observed samples given above.

(c) Test the above hypothesis by constructing a suitable Wald CI in case of dependent samples; assuming $n = m = 15$, *can you apply this CI to the above observed samples?*

Solution.
(a) Note that $\theta_1, \theta_2 \in \mathbb{R}_{>0}$, *thus the pair* $(\theta_1, \theta_2) \in \mathbb{R}_{>0} \times \mathbb{R}_{>0} = \Theta$. *Since* Y_i *and* X_j *are independent, the joint distribution of the two samples is*

$$\begin{aligned} f_{Y_1,\ldots,Y_m,X_1,\ldots,X_n}(y_1, \ldots, y_m, x_1, \ldots, x_n; \theta_1, \theta_2) &= f_{Y_1}(y_1; \theta_1) \cdots f_{Y_m}(y_m; \theta_1) f_{X_1}(x_1; \theta_2) \cdots f_{X_n}(x_n; \theta_2) \\ &= \left(\prod_{i=1}^{m} f_{Y_i}(y_i; \theta_1) \right) \left(\prod_{j=1}^{n} f_{X_j}(x_j; \theta_2) \right), \end{aligned}$$

where $f_{X_j}(x_j; \theta_2)$ *is the p.d.f. of* $\text{Poi}(\theta_2)$. *But when we see it as a function of* (θ_1, θ_2), *this is the likelihood function* $L(\theta_1, \theta_2)$. *In this problem we see that the likelihood is actually the product to two functions with separate arguments*

$$L(\theta_1, \theta_2) = L_1(\theta_1) L_2(\theta_2).$$

Under H_1 *we take the maximum of the likelihood over* Θ. *But as the likelihood factorises, the sup is much simpler*

$$\sup_{(\theta_1,\theta_2)\in\Theta} L(\theta_1, \theta_2) = \left(\sup_{\theta_1 \in \mathbb{R}_{>0}} L_1(\theta_1) \right) \left(\sup_{\theta_2 \in \mathbb{R}_{>0}} L_2(\theta_2) \right) = L_1(\widehat{\theta}_1) L_2(\widehat{\theta}_2),$$

where $\widehat{\theta}_1 = \overline{Y}$ *and* $\widehat{\theta}_2 = \overline{X}$. *Under* H_0 *we have the constraint* $\theta_1 = \theta_2$ *thus let* $\Theta_0 = \{(\theta_1, \theta_2) \in \Theta : \theta_1 = \theta_2\}$. *We have to find* $\sup_{(\theta_1,\theta_2)\in\Theta_0} L(\theta_1, \theta_2)$, *which is equivalent to* $\sup_{\theta_1 \in \mathbb{R}_{>0}} L(\theta_1, \theta_1)$. *For, note that*

$$
\begin{aligned}
L(\theta_1,\theta_1) = L_1(\theta_1)L_2(\theta_1) &= \left(\prod_{i=1}^{m}\frac{e^{-\theta_1}\theta_1^{Y_i}}{Y_i!}\right)\left(\prod_{j=1}^{n}\frac{e^{-\theta_1}\theta_1^{X_j}}{X_j!}\right)\\
&= e^{-(n+m)\theta_1}\theta_1^{\sum_i Y_i+\sum_j X_j}\frac{1}{\left(\prod_i Y_i!\right)\left(\prod_j X_j!\right)}.
\end{aligned}
$$

You can check that the maximum of this function is at $\widehat{\theta}_0 = \frac{\sum_i Y_i+\sum_j X_j}{n+m}$*, and thus* $\sup_{\theta_1\in\mathbb{R}_{>0}} L(\theta_1,\theta_1) = L(\widehat{\theta}_0,\widehat{\theta}_0)$*. The likelihood ratio statistic is then*

$$
\Lambda_n = \frac{\overline{Y}^{m\overline{Y}}\,\overline{X}^{n\overline{X}}}{\left(\frac{n\overline{Y}+m\overline{X}}{n+m}\right)^{n\overline{Y}+m\overline{X}}}.
$$

The LRT is thus to

reject H_0 *if* λ_n^{obs}*, the observed value of* $\lambda_n = 2\log\Lambda_n$*, is greater than some threshold* c*,*

defined such that $\alpha \geq \sup_{(\theta_1,\theta_2)\in\Theta_0} P_{(\theta_1,\theta_2)}(\lambda_n \geq c)$*. In this case, such a probability is not computable since there is no exact distribution for* Λ_n *or for* λ_n*. We may appeal to the large sample property of* λ_n*. In particular, in our case we have that*

$$
\lambda_n \xrightarrow{d} \chi_1^2,
$$

since $dim(\Theta) - \dim(\Theta_0) = 1$*, thus the choice* $c = \chi^2_{1,\alpha}$ *will lead to a test with rejection region of size approximately equal to* α*.*

With the observed samples as above we obtain $\lambda_n^{obs} = 2\log\frac{3^{45}5^{50}}{\left(\frac{45+50}{15+10}\right)^{45+50}} \doteq 70.01$*. With* $\alpha = 0.05$*,* $\chi^2_{1,0.05} = 3.84$*, so we reject* H_0*. Furthermore, the p-value is* $p(\chi^2_1 \geq 70.01) < 0.001$*, so there is a strong evidence against* H_0*.*

(b) The above hypotheses can be restated as $H_0 : \delta = 0$ *against* $H_1 : \delta \neq 0$*, where* $\delta = \theta_1 - \theta_2$*. We can thus test* H_0 *by computing a CI for* δ *and applying the test:*

if 0 *is inside the CI we do not reject* H_0*, otherwise we reject* H_0*.*

If the CI has confidence level $1-\alpha$ *then the test has level or size* α*.*

By the equivariance principle of the MLE we can estimate δ *by the estimator* $\widehat{\delta} = \widehat{\theta}_1 - \widehat{\theta}_2$*.*

By the large sample properties of the MLE, we have $\widehat{\theta}_1 \xrightarrow{d} N(\theta_1, \widehat{se}(\theta_1)^2)$ *and* $\widehat{\theta}_2 \xrightarrow{d} N(\theta_2, \widehat{se}(\theta_2)^2)$ *and since the samples* Y_i *and* X_j *are independent also* $\widehat{\theta}_1$ *is independent from* $\widehat{\theta}_2$*; here* $\widehat{se}(\widehat{\theta}_1) = \sqrt{I_n(\widehat{\theta}_1)^{-1}}$.

Thus

$$\widehat{\delta} = \widehat{\theta}_1 - \widehat{\theta}_2 \xrightarrow{d} N\left(\theta_1 - \theta_2, \widehat{se}(\widehat{\theta}_1)^2 + \widehat{se}(\widehat{\theta}_2)^2\right),$$

and so

$$\frac{\widehat{\delta}-(\theta_1-\theta_2)}{\sqrt{\widehat{se}(\widehat{\theta}_1)^2+\widehat{se}(\widehat{\theta}_2)^2}} \xrightarrow{d} N(0,1).$$

Therefore

$$\begin{aligned}
1-\alpha &= P_\theta\left(\left|\frac{\widehat{\delta}-\delta}{\sqrt{\widehat{se}(\widehat{\theta}_1)^2+\widehat{se}(\widehat{\theta}_2)^2}}\right| \leq z_{\alpha/2}\right)\\
1-\alpha &= P_\theta\left(-z_{\alpha/2} \leq \frac{\widehat{\delta}-\delta}{\sqrt{\widehat{se}(\widehat{\theta}_1)^2+\widehat{se}(\widehat{\theta}_2)^2}} \leq z_{\alpha/2}\right)\\
1-\alpha &= P_\theta\left(\widehat{\delta} - z_{\alpha/2}\sqrt{\widehat{se}(\widehat{\theta}_1)^2 + \widehat{se}(\widehat{\theta}_2)^2} \leq \delta \leq \widehat{\delta} + z_{\alpha/2}\sqrt{\widehat{se}(\widehat{\theta}_1)^2 + \widehat{se}(\widehat{\theta}_2)^2}\right),
\end{aligned}$$

and the desired confidence interval with approximate level $1-\alpha$ *is* $[\widehat{\delta} \pm z_{\alpha/2}\sqrt{\widehat{se}(\widehat{\theta}_1)^2 + \widehat{se}(\widehat{\theta}_2)^2}]$. *With the observed sample we get the observed confidence interval [-3.64,-0.36]. Since* $\delta = 0$ *is not inside the interval we reject* H_0.

(c) We build a Wald CI by finding a suitable pivot. Let $D_i = Y_i - X_i$*, for* $i = 1, \dots, n$, $\overline{D} = n^{-1}\sum_i D_i = \overline{Y} - \overline{X}$ *and* $S_D^2 = (n-1)^{-1}\sum_i (D_i - \overline{D})^2$ *and note that* D_i *are independent; let* $\delta = E(D_i) = \theta_1 - \theta_2$ *and* $\sigma_D^2 = \mathrm{var}(D_i)$.

Now by the LLN

$$\frac{\sqrt{n}(\overline{D}-\delta)}{\sqrt{\sigma_D^2}} \xrightarrow{d} N(0,1).$$

But from previous theory we know that S_D^2 *is a consistent estimator for* σ_D^2*, thus the above formula is asymptotically equivalent to*

$$\frac{\sqrt{n}(\overline{D}-\delta)}{\sqrt{S_D^2}} \xrightarrow{d} N(0,1),$$

which is a pivot for

Example 6.3 *Let* $X_i \overset{\text{iid}}{\sim} \mathrm{N}(\mu, \sigma^2)$, $i = 1, \ldots, n$. *Suppose that* μ *and* σ^2 *are unknown.*

(a) *Construct a test for* $H_0 : \mu \leq \mu_0$ *against* $H_1 : \mu > \mu_0$*; apply it to an observed sample with* $\overline{x} = 1.5$, $s^2 = 1.1$, $n = 23$, *given* $\mu_0 = 1$ *and* $\alpha = 0.05$.

(b) *For (a) determine the sample size* n *needed in order to achieve* β *as close as possible to but not exceeding* 0.01 *given the true value (under* H_1*) is equal to* $\mu_1 = 1.2$.

(c) *Construct an upper confidence set for* σ^2 *of level* $1 - \alpha$ *and apply it to the observed sample in (a).*

Solution.

Since σ^2 *is unknown we cannot apply the NP Lemma (the MLR property cannot be applied either) but we can resort to the* t*-test. The pivot to be used is*

$$T_n = \tfrac{\sqrt{n}(\overline{X}-\mu)}{S} \sim t_{n-1}.$$

Intuitively we should reject H_0 *if from our sample we get an average that is "too high", thus the rejection region should be of the form*

$$R = \{X_1, \ldots, X_n : \overline{X} \geq c\};$$

the threshold c *is determined in such a way that the rejection region has size* α*. For any* μ *under* H_0,

$$\begin{aligned}
\alpha &= \sup_{\mu\leq\mu_0} P_\mu(\overline{X} \geq c) = \sup_{\mu\leq\mu_0} P_\mu\left(\tfrac{\sqrt{n}(\overline{X}-\mu)}{S} \geq \tfrac{\sqrt{n}(c-\mu)}{S}\right) \\
&= \sup_{\mu\leq\mu_0} P\left(t_{n-1} \geq \tfrac{\sqrt{n}(c-\mu)}{S}\right) \\
&= 1 - F_{t_{n-1}}\left(\tfrac{\sqrt{n}(c-\mu_0)}{S}\right),
\end{aligned}$$

where $F_{t_{n-1}}$ *is the d.f. of the* t_{n-1} *random variable (see also Example 6.4, in Lecture 6.). By definition then* $t_{n-1,\alpha} = \frac{\sqrt{n}(c-\mu_0)}{S}$ *and* $c = t_{n-1,\alpha}S/\sqrt{n} + \mu_0$*. The test is thus to reject* H_0 *whenever* $\overline{X} \geq t_{n-1,\alpha}S/\sqrt{n} + \mu_0$.

Since $\overline{x} = 1.05 < 1.38 = t_{22,0.05}\sqrt{1.1/23} + 1$, *then we do not reject* H_0.

(b) The type size of II error is

$$\beta(\mu) = 1 - P\left(t_{n-1} \geq \frac{\sqrt{n}(t_{n-1,\alpha}-\mu)}{S}\right) \quad \textit{for all } \mu > \mu_0,$$

but by assumption $\beta(\mu_1) = 0.01$ *for* $\mu_1 = 1.2$. *Therefore* n *is found by solving*

$$\beta(\mu_1) = 0.1 = 1 - P\left(t_{n-1} \geq \frac{\sqrt{n}(t_{n-1,\alpha}-2)}{S}\right),$$

from which we find that $n = \left(\frac{SF_{t_{n-1}}^{-1}(0.1)}{t_{n-1,\alpha}-1.2}\right)^2 \doteq 26$.

(c) Using the fact that $\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$, *then we have that*

$$1 - \alpha = P_\theta\left(\frac{(n-1)S^2}{\sigma^2} \leq \chi^2_{n-1,\alpha}\right) = P_\theta\left(\frac{(n-1)S^2}{\chi^2_{n-1,\alpha}} \leq \sigma^2\right)$$

where $\chi^2_{n-1,\alpha}$ *is the upper* α*th quantile of the* χ^2_{n-1} *distribution, thus a* $1-\alpha$ *upper confidence interval for* σ^2 $\left[\frac{(n-1)S^2}{\chi^2_{n-1,\alpha}}, \infty\right)$. *With the observed sample at hand we have the 0.95 CI* $[0.713, \infty)$.