Tufts University
Department of Mathematics
Spring 2022

MA 166: Statistics

Homework 6 (v1.1) [1]

Assigned Monday 28 February 2022
Due Monday 7 March 2022 at 11:59 pm EDT.

1. **Larsen & Marx, Section 6.4, Problem 6.4.4, page 373: Construct a power curve for the $\alpha = 0.05$ test of $H_0: \mu = 60$ versus $H_1: \mu \neq 60$ if the data consist of a random sample of size 16 from a normal distribution having $\sigma = 4$.**

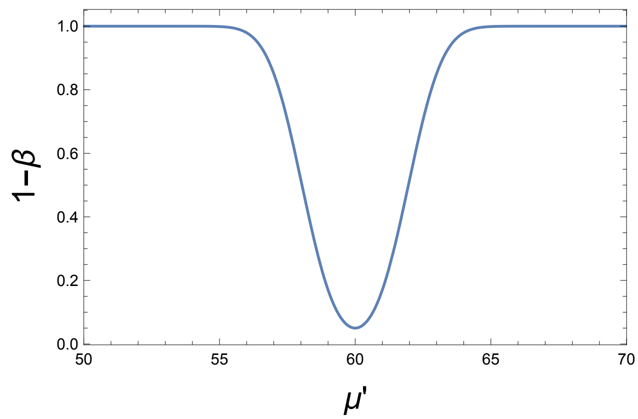
   The upper and lower cutoffs $\mu_{c\pm}$ for incurring a Type I error by rejecting a valid H_0 are given by
$$\frac{\mu_{c\pm} - \mu_0}{\sigma/\sqrt{n}} = \pm z_{\alpha/2},$$
   so that
$$\mu_{c\pm} = \mu_0 \pm z_{\alpha/2}\frac{\sigma}{\sqrt{n}}.$$
   In other words, assuming that H_0 is true, a Type I error will be made if the sample mean is greater than μ_{c+} or less than μ_{c-}.

   If H_0 is not true, on the other hand, and the actual mean is located at some value μ', a Type II error will be made with probability
$$\begin{aligned}
\beta =& P\left(\mu_{c-} \leq \overline{\mu} \leq \mu_{c+} \mid \overline{\mu} \text{ is } N(\mu', \sigma/\sqrt{n}) \text{ r.v.}\right)\\
=& \int_{\mu_{c-}}^{\mu_{c+}} d\xi\ \frac{1}{\sqrt{2\pi}\ \sigma/\sqrt{n}} \exp\left[-\frac{(\xi-\mu')^2}{2\sigma^2/n}\right]\\
=& \int_{\frac{\mu_{c-}-\mu'}{\sigma/\sqrt{n}}}^{\frac{\mu_{c+}-\mu'}{\sigma/\sigma n}} dz\ \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right)\\
=& \int_{\frac{\mu_{c-}-\mu'}{\sigma/\sqrt{n}}}^{\infty} dz\ \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right) - \int_{\frac{\mu_{c+}-\mu'}{\sigma/\sqrt{n}}}^{\infty} dz\ \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right)\\
=& \int_{\frac{\mu_0-\mu'}{\sigma/\sqrt{n}}-z_{\alpha/2}}^{\infty} dz\ \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right) - \int_{\frac{\mu_0-\mu'}{\sigma/\sqrt{n}}+z_{\alpha/2}}^{\infty} dz\ \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right).
\end{aligned}$$
   Going forward, let us define the function Z^{-1} by $\forall\beta \in [0,1]:\ Z^{-1}(z_\beta) = \beta$. Then we can satisfy the above equation by taking
$$\begin{aligned}
Z^{-1}\left(\frac{\mu_0-\mu'}{\sigma/\sqrt{n}} - z_{\alpha/2}\right) =& \frac{\gamma+\beta}{2}\\
Z^{-1}\left(\frac{\mu_0-\mu'}{\sigma/\sqrt{n}} + z_{\alpha/2}\right) =& \frac{\gamma-\beta}{2},
\end{aligned}$$

whence

$$\beta = Z^{-1}\left(\frac{\mu_0 - \mu'}{\sigma/\sqrt{n}} - z_{\alpha/2}\right) - Z^{-1}\left(\frac{\mu_0 - \mu'}{\sigma/\sqrt{n}} + z_{\alpha/2}\right).$$

Given α, σ, n and μ_0, the above equation allows us to make a plot of $1-\beta$ as a function of μ', which is the power curve. In our case, $\alpha = 0.05$, $\sigma = 4$, $n = 16$ and $\mu_0 = 60$, and the above relationship between $1-\beta$ and μ' is plotted below



Note that, unlike the power curve shown in Fig. 6.4.4 of the Larsen and Marx text, the graph of the above power curve is symmetric about μ_0, owing to the two-sided nature of the test. On the other hand, it remains true that $\lim_{\mu' \to \mu_0}(1-\beta) = \alpha$, since

$$\begin{aligned}
1 - \beta =& 1 - Z^{-1}\left(-z_{\alpha/2}\right) + Z^{-1}\left(z_{\alpha/2}\right)\\
=& 1 - Z^{-1}\left(z_{1-\alpha/2}\right) + Z^{-1}\left(z_{\alpha/2}\right)\\
=& 1 - (1-\alpha/2) + (\alpha/2)\\
=& \alpha.
\end{aligned}$$

2. **Larsen & Marx, Section 6.5, Problem 6.5.2, page 377: Let $y_1, y_2, \ldots, y_{10}$ be a random sample from an exponential pdf with unknown parameter λ. Find the form of the GLRT for $H_0 : \lambda = \lambda_0$ versus $H_1 : \lambda \neq \lambda_0$. What integral would have to be evaluated to determine the critical value if α were equal to 0.05?**

The exponential distribution is $f_Y(y;\lambda) = \lambda e^{-\lambda y}$ for $y > 0$. Here we must have $\lambda > 0$, else f_Y will not be normalizable. The likelihood function is then

$$L(\lambda) = \prod_{j=1}^{n} \lambda e^{-\lambda y_j} = \lambda^n e^{-n\lambda\overline{y}},$$

and the log likelihood is

$$\log L(\lambda) = n \log\lambda - n\lambda\overline{y}.$$

The maximum likelihood occurs when

$$0 = \frac{\partial \log L(\lambda)}{\partial\lambda} = \frac{n}{\lambda} - n\overline{y}$$

or $\lambda = \lambda_e := 1/\overline{y}$.

The set of λ values consistent with the null hypothesis is $\omega = \{\lambda_0\}$, while the set of all possible λ values is $\Omega = \mathbb{R}^+$. Hence we have

$$\max_{\lambda \in \omega} L(\lambda) = L(\lambda_0) = \lambda_0^n e^{-n\lambda_0 \overline{y}} = \lambda_0^n e^{-n\lambda_0/\lambda_e}$$

and

$$\max_{\lambda \in \Omega} L(\lambda) = L(\lambda_e) = \lambda_e^n e^{-n}$$

The GLR is usually denoted by λ, but we can not use that here because we are already using λ to denote the parameter. So let us denote the GLR by the next letter in the Greek alphabet, μ. We have

$$\mu = \frac{\max_{\lambda \in \omega} L(\lambda)}{\max_{\lambda \in \Omega} L(\lambda)} = \left(\frac{\lambda_0}{\lambda_e}\right)^n \exp\left[n\left(1 - \frac{\lambda_0}{\lambda_e}\right)\right] = (\lambda_0 \overline{y})^n \exp\left[n\left(1 - \lambda_0 \overline{y}\right)\right],$$

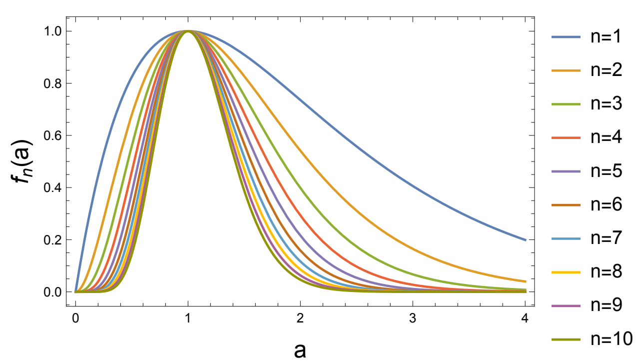
and so the GLRT is that we reject H_0 whenever

$$\mu = (\lambda_0 \overline{y})^n \exp\left[n\left(1 - \lambda_0 \overline{y}\right)\right] \leq \mu^*.$$

If we define $a = \lambda_0 \overline{y}$, this criterion becomes

$$f_n(a) := a^n \exp\left[n(1-a)\right] \leq \mu^*.$$

The function $f_n(a)$ is plotted against a for various values of n below.



For low values of n, it is seen that $f_n(a)$ is a skewed distribution, but for large values of n, it begins to resemble something more familiar. To see what it becomes, note that $f_n(a) = e^{F_n(a)}$, where

$$F_n(a) = \log\left[f_n(a)\right] = n \log a + n - na,$$

and that

$$\begin{aligned} F_n'(a) &= \frac{n}{a} - n \\ F_n''(a) &= -\frac{n}{a^2}. \end{aligned}$$

From this, we see that

$$\begin{aligned} F_n(1) =& 0 \\ F_n'(a) =& 0 \\ F_n''(a) =& -n, \end{aligned}$$

and, in fact, this much is evident in the plots of $f_n(a)$ provided above. Hence, the leading term in the Taylor expansion of $F_n(a)$ about $a = 1$ is

$$F_n(a) \approx -\frac{n}{2}(a-1)^2,$$

whence

$$f_n(a) \approx C_n \exp\left[-\frac{n}{2}(a-1)^2\right],$$

where we have allowed for a proportionality constant C_n. To interpret f_n as a pdf for a, we need to choose the C_n so that $f_n(a)$ is normalized, and this yields

$$f_n(a) \approx \frac{1}{\sqrt{2\pi/n}} \exp\left[-\frac{(a-1)^2}{2/n}\right],$$

This is a normal distribution with mean 1 and variance $1/n$. In this limit then, the GLRT becomes

$$-z_{\alpha/2} \le \sqrt{n}(a-1) \le +z_{\alpha/2},$$

or

$$1 - \frac{z_{\alpha/2}}{\sqrt{n}} \le a \le 1 + \frac{z_{\alpha/2}}{\sqrt{n}},$$

or, since $a = \lambda_0 \overline{y}$, we reject H_0 if

$$\boxed{1 - \frac{z_{\alpha/2}}{\sqrt{n}\,\lambda_0} \le \overline{y} \le 1 + \frac{z_{\alpha/2}}{\sqrt{n}\,\lambda_0}.}$$

It is seen that the GLRT reduces to a standard Z test in the large n limit.

3. **Larsen & Marx, Section 7.3, Problem 7.3.2, page 388: Find the moment-generating function for a chi square random variable and use it to show that $\boldsymbol{E(\chi_n^2) = n}$ and $\mathbf{Var}\boldsymbol{(\chi_n^2) = 2n}$.**

   The chi squared pdf with n degrees of freedom is

$$f_U(u) = \frac{1}{2^{n/2}\Gamma\left(\frac{n}{2}\right)} u^{(n/2)-1} e^{-u/2}$$

for $u \geq 0$. The moment generating function for this is given by

$$
\begin{aligned}
M_U(t) &= \int_0^\infty du\ e^{tu} f_U(u) \\
&= \int_0^\infty du\ e^{tu} \frac{1}{2^{n/2}\Gamma\left(\frac{n}{2}\right)} u^{(n/2)-1} e^{-u/2} \\
&= \frac{1}{2^{n/2}\Gamma\left(\frac{n}{2}\right)} \int_0^\infty du\ u^{(n/2)-1} e^{-(1/2-t)u} \\
&= \frac{1}{2^{n/2}(1/2-t)^{n/2}} \left[\frac{(1/2-t)^{n/2}}{\Gamma\left(\frac{n}{2}\right)} \int_0^\infty du\ u^{(n/2)-1} e^{-(1/2-t)u} \right].
\end{aligned}
$$

The quantity in square brackets is the integration of a gamma pdf with parameters $n/2$ and $(1/2 - t)$, and hence is equal to one. (Really, this last conclusion is valid only for $t < 1/2$, but the variable in a moment generating function is usually regarded as a formal variable and not associated with a numerical value.) In any case, we are left with the moment generating function

$$
\boxed{M_U(t) = \frac{1}{(1-2t)^{n/2}},}
$$

for which the binomial expansion in t to second order is

$$
M_U(t) = 1 + nt + \frac{n(n+2)}{2} t^2 + \cdots
$$

From the definition of the moment generating function, we see that

$$
\begin{aligned}
M_U(t) &= \int_0^\infty du\ e^{tu} f_U(u) \\
&= \int_0^\infty du\ \left(\sum_{j=0}^\infty \frac{t^j u^j}{j!} \right) f_U(u) \\
&= \sum_{j=0}^\infty \frac{t^j}{j!} \int_0^\infty du\ u^j f_U(u) \\
&= \sum_{j=0}^\infty \frac{t^j}{j!} E(u^j).
\end{aligned}
$$

Comparing this with the binomial expansion of $M_U(t)$ given above, we identify

$$
\boxed{E(U) = n,}
$$

and

$$
E(U^2) = n(n+2) = n^2 + 2n.
$$

It follows that

$$
\mathrm{Var}(U) = E(U^2) - E(U)^2 = n^2 + 2n - n^2,
$$

or

$$\boxed{\mathrm{Var}(U) = 2n,}$$

as was to be shown.