Statistics (R.237) Patrick Finke, 651932001

Exercise 2.17 Def (Median) A value m such that
$P(X \leq m) \geq \frac{1}{2}$ and $P(X \geq m) \geq \frac{1}{2}$
or for X continuous
$\int_{-\infty}^{m} f(x)\,dx = \frac{1}{2} = \int_{m}^{\infty} f(x)\,dx$.

We only need to check one integral because $\int_{\mathbb{R}} f(x)\,dx = 1$.

(a) $f(x) = 3x^2, \quad 0 < x < 1$.

$\frac{1}{2} \overset{!}{=} \int_0^m f(x)\,dx = 3\int_0^m x^2\,dx = 3\left(\frac{1}{3}x^3\right]_0^m = m^3$

$\Rightarrow m = \frac{1}{2^{1/3}}$.

(b) $f(x) = \frac{1}{\pi(1+x^2)}, \quad -\infty < x < \infty$.

$\frac{1}{2} \overset{!}{=} \frac{1}{\pi}\int_{-\infty}^{m} \frac{1}{1+x^2}\,dx = \frac{1}{\pi}\left[\arctan(x)\right]_{-\infty}^{m} = \frac{1}{\pi}\arctan(m) + \frac{1}{2}$

$\Rightarrow \arctan(m) = 0 \Rightarrow m = \tan(0) = 0$.

Theorem 5.26 $X_1, \dots, X_n$ i.i.d random sample
$EX_i = \mu$, $\operatorname{Var} X_i = \sigma^2 < \infty \quad \forall i = 1, \dots, n$.
Then for
$\bar{X} := \frac{1}{n}\sum_{i=1}^{n} X_i$ (sample mean) and
$S^2 := \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$ (sample variance)
there holds:

(a) $E\bar{X} = \mu$ (b) $\operatorname{Var}\bar{X} = \frac{\sigma^2}{n}$ (c) $ES^2 = \sigma^2$

Proof (a) $E\bar{X} = E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n}\sum_{i=1}^{n} EX_i = \frac{1}{n} n\mu = \mu$.

(b) $\operatorname{Var}\bar{X} = \operatorname{Var}\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}\operatorname{Var}\left(\sum_{i=1}^{n} X_i\right)$

$= \frac{1}{n^2} E\left[\sum_{i=1}^{n} X_i - E\left(\sum_{i=1}^{n} X_i\right)\right]^2$

$= \frac{1}{n^2} E\underbrace{\left[\sum_{i=1}^{n}(X_i - EX_i)\right]^2}_{(*)}$.

Observing (*), there are two types of terms:

(i) $\underline{(X_i - EX_i)^2}$:

$$E(X_i - EX_i)^2 = \operatorname{Var} X_i$$

(ii) $\underline{(X_i - EX_i)(X_j - EX_j) \quad (i \neq j)}$:

$$E(X_i - EX_i)(X_j - EX_j) = \operatorname{Cov}(X_i, X_j) \overset{\text{indep.}}{=} 0.$$

Thus,

$$\operatorname{Var} \overline{X} = \frac{1}{n^2} \sum_{i=1}^n \operatorname{Var} X_i = \frac{1}{n^2} n \cdot \sigma^2 = \frac{\sigma^2}{n}.$$

(c) First observe,

$$\sum_{i=1}^n X_i^2 = \sum_{i=1}^n \left((X_i - \overline{X}) + \overline{X} \right)^2$$

$$= \sum_{i=1}^n (X_i - \overline{X})^2 + 2 \underbrace{\sum_{i=1}^n (X_i - \overline{X})}_{\substack{= \sum_{i=1}^n X_i - n\overline{X} \\ = n\overline{X} - n\overline{X} = 0.}} \overline{X} + \sum_{i=1}^n \overline{X}^2$$

$$= \sum_{i=1}^n (X_i - \overline{X})^2 + \sum_{i=1}^n \overline{X}^2$$

$$\Rightarrow \sum_{i=1}^n (X_i - \overline{X})^2 = \sum_{i=1}^n X_i^2 - \sum_{i=1}^n \overline{X}^2.$$

Now,

$$ES^2 = E\left[\frac{1}{n-1} \sum_{i=1}^n (X_i - \overline{X})^2 \right]$$

$$\overset{\text{above}}{=} E\left[\frac{1}{n-1} \left(\sum_{i=1}^n X_i^2 - \sum_{i=1}^n \overline{X}^2 \right) \right]$$

$$\overset{\text{i.i.d.}}{=} \frac{1}{n-1} \left(n\, EX_1^2 - n\, E\overline{X}^2 \right)$$

For a r.v. X we have

$$EX^2 = EX^2 - (EX)^2 + (EX)^2 = \operatorname{Var} X + (EX)^2.$$

Thus, using (a) and (b),

$$ES^2 = \frac{1}{n-1} \left(n(\sigma^2 + \mu^2) - n\left(\frac{\sigma^2}{n} + \mu^2 \right) \right)$$

$$= \frac{1}{n-1} \left(n\sigma^2 - \sigma^2 \right) = \sigma^2.$$

651932001
Patrick Finke

Example 7.2.7 $X_1, \dots, X_n$ iid. Bernoulli(p).

Then the maximum likelihood function is

$$
\begin{aligned}
L(p|\underline{x}) &= f_{\underline{X}}(\underline{x}|p) \\
&\overset{\text{iid.}}{=} \prod_{i=1}^{n} f_{X_i}(x_i|p) \\
&\overset{\text{Bernoulli}}{=} \prod_{i=1}^{n} p^{x_i}(1-p)^{1-x_i} \\
&= p^{\sum_{i=1}^{n} x_i}(1-p)^{\sum_{i=1}^{n}(1-x_i)} \\
&= p^{y}(1-p)^{n-y} \quad \text{for } y := \sum_{i=1}^{n} x_i, \ 0 \le y \le n
\end{aligned}
$$

To find the MLE, we maximize the log likelihood

$$\log L(p|\underline{x}) = y \log p + (n-y)\log(1-p)$$

and distinguish the cases

$$
\log L(p|\underline{x}) = \begin{cases} n \log(1-p) & y = 0 \quad (1) \\ y \log p + (n-y)\log(1-p) & 0 < y < n \quad (2) \\ n \log p & y = n \quad (3) \end{cases}
$$

Case (2): $\frac{d}{dp} \log L(p|\underline{x}) = \frac{y}{p} - \frac{n-y}{1-p}$

$$= \frac{(1-p)y - p(n-y)}{p(1-p)} = \frac{y - pn}{p(1-p)} \overset{!}{=} 0$$

$$\Leftrightarrow y = pn \Leftrightarrow p = y/n.$$

Case (1): $\log(p)$ monotonic increasing

$\Rightarrow \log(1-p)$ monotonic decreasing

$\Rightarrow$ max. at $\hat{p} = 0 \underset{(y=0)}{=} y/n$.

Case (3): $\log(p)$ monotonic increasing

$\Rightarrow$ max. at $\hat{p} = 1 \underset{(y=n)}{=} y/n$

Because log is monotonic increasing, $\hat{p} = y/n = \frac{1}{n}\sum_{i=1}^{n} X_i$ also maximizes $L(p|\underline{x})$, i.e. $\hat{p}$ is MLE.

Example 8.2.3 $X_1, -, X_n$ i.i.d random sample from exp. population.
with pdf

$$f(\underset{\sim}{x}|\theta) = \begin{cases} e^{-(x-\theta)}, & x \geq \theta \\ 0, & x < \theta \end{cases} \qquad (-\infty < x < \infty)$$

Then, the maximum likelihood function is

$$L(\theta|\underline{x}) = f_{\underline{x}}(\underline{x}|\theta) \overset{\text{i.i.d}}{=} \prod_{i=1}^{n} f_{X_i}(x_i|\theta)$$
$$= \begin{cases} e^{-\sum x_i + n\theta}, & \forall i: x_i \geq \theta \\ 0, & \exists i: x_i < \theta \end{cases}$$
$$= \begin{cases} e^{-\sum x_i + n\theta}, & \min x_i = x_{(1)} \geq \theta \\ 0, & \min x_i = x_{(1)} < \theta. \end{cases}$$

We consider testing

$$H_0 : \theta \leq \theta_0 \qquad , \qquad H_1 : \theta > \theta_0$$

with given θ_0, and now search for

$$\lambda(\underline{x}) = \frac{\sup_{\theta \leq \theta_0} L(\theta|\underline{x})}{\sup_{\theta} L(\theta|\underline{x})}.$$

For $-\infty < \theta \leq x_{(1)}$, $L(\theta|\underline{x})$ is increasing (and 0 after that), obtains it's maximum at $x_{(1)}$. Thus

$$\sup_{\theta} L(\theta|\underline{x}) = L(x_{(1)}|\underline{x})$$

For the restricted supremum, we distinguish cases:

1) $\underline{x_{(1)} \leq \theta_0}$: $\sup_{\theta \leq \theta_0} L(\theta|\underline{x}) = L(x_{(1)}|\underline{x})$, because here $x_{(1)}$ is in the set of θ's over which the supremum is taken.

2) $\underline{x_{(1)} > \theta_0}$: On the set $\{\theta \leq \theta_0\} \subset \{\theta \leq x_{(1)}\}$ $L(\theta|\underline{x})$ is increasing, thus

$$\sup_{\theta \leq \theta_0} L(\theta|\underline{x}) = L(\theta_0|\underline{x})$$

Therefore,

$$\lambda(\underline{x}) = \begin{cases} \dfrac{L(x_{(1)}|\underline{x})}{L(x_{(1)}|\underline{x})} = 1 & x_{(1)} \leq \theta_0 \\ \dfrac{L(x_{(1)}|\underline{x})}{L(\theta_0|\underline{x})} = e^{-n(x_{(1)} - \theta_0)} & x_{(1)} > \theta_0 \end{cases}$$

Finally, we determine the rejection region for an LRT that rejects H_0 if $\lambda(x) \leq c$ for a given $c \in (0,1)$.

Patrick Finke 651932001

$\lambda(\underline{x}) \leq c$

$\Leftrightarrow \quad e^{-n(x_{(1)}-\theta_0)} \leq c \quad \Leftrightarrow -n(x_{(1)}-\theta_0) \leq \log c$

$\Leftrightarrow \quad \theta_0 - x_{(1)} \leq \frac{\log c}{n} \quad \Leftrightarrow \theta_0 - \frac{\log c}{n} \leq x_{(1)}.$

Thus, the rejection region is given by

$R = \{ \underline{x} : \lambda(\underline{x}) \leq c \} = \{ \underline{x} : x_{(1)} \geq \theta_0 - \frac{\log c}{n} \}$

only depending on $x_{(1)}$.

Theorem 8.2.4 Let $T(\underline{X})$ be a sufficient statistic for θ, $\lambda^*(t)$ and $\lambda(x)$ are the LRT statistics based on T and $\underline{X}$ respectively. Then

$$\lambda^*(T(x)) = \lambda(x)$$

for all x in the sample space.

Proof: We use the Factorization Theorem (6.2.6) which characterizes the property of sufficient statistic. Thus, using that T is sufficient for θ, the pdf/pmf $f(\underline{x}|\theta)$ of $\underline{X}$ can be written as

$$f(\underline{x}|\theta) = g(T(\underline{X})|\theta)\, h(\underline{X})$$

where $g(t|\theta)$ is the pdf/pmf of T.

Then,

$$\lambda(\underline{x}) = \frac{\sup_{\theta\in\Theta_0} L(\theta|\underline{x})}{\sup_{\theta\in\Theta} L(\theta|\underline{x})}$$

$$= \frac{\sup_{\theta\in\Theta_0} \cancel{L(\theta|\underline{x})}\, f(\underline{x}|\theta)}{\sup_{\theta\in\Theta} \cancel{L(\theta|\underline{x})}\, f(\underline{x}|\theta)}$$

$$= \frac{\sup_{\theta\in\Theta_0} g(T(\underline{X})|\theta)\, \cancel{h(\underline{X})}}{\sup_{\theta\in\Theta} g(T(\underline{X})|\theta)\, \cancel{h(\underline{X})}}$$

$$= \frac{\sup_{\theta\in\Theta_0} g(T(\underline{X})|\theta)}{\sup_{\theta\in\Theta} g(T(\underline{X})|\theta)}$$

g pdf/pmf of T.

$$= \frac{\sup_{\theta\in\Theta_0} L^*(\theta|T(\underline{X}))}{\sup_{\theta\in\Theta} L^*(\theta|T(\underline{X}))} = \lambda^*(T(\underline{X})).$$

Note: I did the correction of the distribution report in LaTeX but then forgot to send it! (more readable)

I wanted to hand it in now anyway, I also added some extra information.