Minimal Sufficient Statistics

Primary goal:

Reduce the sample to the smallest set of sufficient statistics.
(i.e. minimal set of suff. stats).

Def a set of statistics is called a minimal sufficient set if the members of the set are jointly sufficient for the parameters, and if they are a fct of every other set of jointly suff. statistics.

The definition is not very helpful when it comes to finding a set of minimal sufficient stats. Lehman and Scheffé devise a technique for finding ((9.66)). Typically, sufficient stats found using the factorization criterion will also be a set of minimal sufficient stats. (at least for the examples in this book).

9.5 Rao-Blackwell Thm

Thm 9.5 The Rao-Blackwell Thm

Let $\hat{\theta}$ be an unbiased estimator for θ such that $V(\hat{\theta}) < \infty$.
If u is a suff. stat for θ, define $\hat{\theta}^* = E(\hat{\theta} \mid u)$. Then
for all θ,

$$E(\hat{\theta}^*) = \theta \quad \text{and} \quad V(\hat{\theta}^*) \le V(\hat{\theta})$$

Pf:

Since u is suff for θ

$\Rightarrow f(\hat{\theta} \mid u)$ does not depend on θ

$\Rightarrow \hat{\theta}^* = E(\hat{\theta} \mid u)$ does not depend on θ

i.e. $\hat{\theta}^*$ is a statistic and is a fct of u

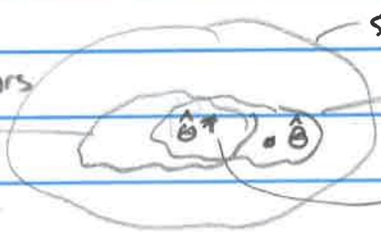
$$E(\hat{\theta}^*) = E_u\left[E(\hat{\theta} \mid u)\right] = E(\hat{\theta}) = \theta \quad \checkmark$$

$$\text{var}(\hat{\theta}) = \text{var}_u\left[E(\hat{\theta} \mid u)\right] + E_u\left[\text{var}(\hat{\theta} \mid u)\right]$$

$$= \text{var}(\hat{\theta}^*) + E_u\underbrace{\left[\text{var}(\hat{\theta} \mid u)\right]}_{\ge 0}$$

$$\Rightarrow \text{var}(\hat{\theta}^*) \le \text{var}(\hat{\theta}) \quad \checkmark$$

What does all this mean? Well, if we want an unbiased estimator of θ with small variance, we should start w/ a fct of a suff stat.



note:

① we know that $var(\hat{\theta}^*) \le var(\hat{\theta})$. However, it is possible that if we used another suff. stat, U_2, that for

$$\hat{\theta}^*_2 = E[\hat{\theta} / U_2] \quad , \quad var(\hat{\theta}^*_2) < var(\hat{\theta}^*)$$

Nevertheless, unbiased fcts. of suff. stats is a good place to start.

② To establish that $\hat{\theta}^*$ has the minimum variance among all unbiased estimators (uniformly minimum-variance unbiased estimator or UMVUE or MVUE or UMVU) we need to consider unbiased fcts of suff. stats from a family of complete (uniqueness) density fcts. In this book, we will only consider complete family density fcts so we will actually be finding UMVUEs although we haven't really verified it.

9.68

Rao-Blackwell Example

Let Y_1, Y_2, Y_3 be a r.s. from

$$f(y/\theta) = \frac{1}{\theta} \quad , \quad 0 < y < \theta \quad , \quad \theta > 0 \qquad F(y) = y/\theta$$

0)

We have already shown that $Y_{(3)}$ is a suff. stat for θ

Consider $Y_{(2)}$

$$g(y_2) = \frac{3!}{1!\,1!}\left[\frac{y_2}{\theta}\right]^1\left[1 - \frac{y_2}{\theta}\right]\frac{1}{\theta}$$

$$= \frac{6}{\theta^3} y_2(\theta - y_2) \quad , \quad 0 < y_2 < \theta$$

$$E(Y_2) = \int_0^{\theta} \frac{6}{\theta^3} y_2^2 (\theta - y_2)\, dy_2 = \frac{6}{\theta^3}\left[\frac{y_2^3 \theta}{3} - \frac{y_2^4}{4}\right]\Big|_0^{\theta}$$

$$= \frac{6}{\theta^3}\left[\frac{\theta^4}{12}\right] = \frac{\theta}{2}$$

$\Rightarrow 2Y_{(2)}$ is unbiased for θ

Let $\hat{\theta}^* = E[2Y_{(2)} \mid Y_{(3)}]$

$$g(y_2, y_3) = \frac{3!}{1!\,0!\,0!}\left[\frac{y_2}{\theta}\right]^1 \cdot \frac{1}{\theta^2} \qquad 0 < y_2 < y_3 < \theta$$

$$= \begin{cases} \frac{6}{\theta^3} y_2 & , 0 < y_2 < y_3 < \theta \\ 0 & , \text{---} \end{cases}$$

$$g(y_3) = \frac{3}{\theta^3} y_3^2 \quad , 0 < y_3 < \theta$$

$$g(y_2 \mid y_3) = \frac{\frac{6}{\theta^3} y_2}{\frac{3}{\theta^3} y_3^2} = \begin{cases} \frac{2y_2}{y_3^2} & , 0 < y_2 < y_3 \\ 0 & , \text{---} \end{cases}$$

note: conditional pdf is independent of θ!!! (as it should because $Y_{(3)}$ is suff. for θ)

$$\hat{\theta}^* = E[2Y_2 \mid Y_3] = \int_0^{y_3} 2y_2 \cdot \frac{2y_2}{y_3^2}\, dy_2 = \frac{4}{y_3^2} \cdot \frac{y_2^3}{3}\Big|_0^{y_3}$$

$$= \boxed{\frac{4}{3} y_3}$$

note: ① we have seen $\frac{n+1}{n} Y_{(n)}$ is unbiased for θ

② $\hat{\theta}^*$ is a fct of $Y_{(3)}$

Check Rao-Blackwell:

[$var(\frac{4}{3}Y_{(3)})$ should be less than $var(2Y_{(2)})$]

$$var\left(\frac{4}{3}Y_{(3)}\right) = E\left[\left(\frac{4}{3}Y_{(3)}\right)^2\right] - \theta^2$$

$$= \frac{16}{9}\int_0^\theta y_3^2 \cdot \frac{3y_3^2}{\theta^3}\,d\theta - \theta^2 = \frac{16}{9}\int_0^\theta \frac{3y_3^4}{\theta^3}\,dy_3 - \theta^2$$

$$= \frac{16}{9}\cdot\frac{3y_3^5}{5\theta^3}\Big|_0^\theta - \theta^2 = \frac{16\theta^2}{15} - \theta^2 = \frac{\theta^2}{15}$$

$$var(2Y_{(2)}) = E\left[(2Y_{(2)})^2\right] - \theta^2$$

$$= 4\int_0^\theta y_2^2 \cdot \frac{3!}{1!1!}\left(\frac{y_2}{\theta}\right)\left(1-\frac{y_2}{\theta}\right)\frac{1}{\theta}\,dy_2 - \theta^2$$

$$= 4\int_0^\theta \frac{6}{\theta^3}\, y_2^3(\theta - y_2)\,dy_2 - \theta^2$$

$$= \frac{24}{\theta^3}\left[\frac{\theta y_2^4}{4} - \frac{y_2^5}{5}\right]\Big|_0^\theta - \theta^2$$

$$= \frac{24}{\theta^3}\left[\frac{\theta^5}{20}\right] - \theta^2 = \frac{\theta^2}{5}$$

In practice, it can be difficult to find the conditional expectation $E(\hat{\theta} \mid u)$. However, if we find u, and then a fct $h(u)$ where $E[h(u)] = \theta$, we know we have a pretty good estimator.

Ideally, we would like $\hat{\theta}^{\#}$ to be a UMVUE for θ. Our book

says (1) use factorization thm to find u
(2) Find $h(u)$ where $E[h(u)] = \theta$.

$\Rightarrow$ conclude that $\hat{\theta}^* = h(u)$ is MVUE.

note: this will be true for the specific examples we look at, but to establish this in general, we need completeness

e.g. Let $Y_1 \cdots Y_n$ be a r.s. from $Y \sim POI(\lambda)$

$$L(\lambda) = \prod_{i=1}^{n} \frac{e^{-\lambda} \lambda^{y_i}}{y_i!} = \frac{e^{-n\lambda} \lambda^{\Sigma y_i}}{\prod_{i=1}^{n} y_i!} = \underbrace{e^{-n\lambda} \lambda^{\Sigma y_i}}_{g(u,\lambda)} \underbrace{\left(\frac{1}{\prod_{i=1}^{n} y_i!} \right)}_{h(y_1 \cdots y_n)}$$

as before, the factorization thm tells us that $u = \Sigma y_i$ is suff. for λ.

Now, we need to find $h(u)$ where $E[h(u)] = \lambda$

$E(u) = E(\Sigma y_i) = n\lambda$

$\Rightarrow$ let $h(u) = \frac{\Sigma y_i}{n} = \boxed{\bar{Y}}$

$\bar{Y}$ is an unbiased estimator of λ, and it is a fct. of a suff. stat. Rao-Blackwell Thm $\Rightarrow$ $\bar{Y}$ is a very good estimator. as it turns out, $\bar{Y}$ is MVUE for λ. (Technically we haven't shown this)

Example 9.7

$Y_1 \cdots Y_n$ is a r.s. from

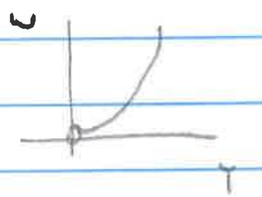
$$f(y|\theta) = \begin{cases} \frac{2y}{\theta} e^{-y^2/\theta}, & y > 0 \\ 0, & \text{—} \end{cases}$$

Find an MVUE for θ

$$L(\theta) = \prod_{i=1}^{n} \frac{2y_i}{\theta} e^{-y_i^2/\theta} = \left(\frac{2}{\theta}\right)^n \prod_{i=1}^{n} y_i \cdot e^{-\sum y_i^2/\theta}$$

$$= \underbrace{\left(\frac{2}{\theta}\right)^n e^{-\sum y_i^2/\theta}}_{g(u,\theta)} \cdot \underbrace{\prod_{i=1}^{n} y_i}_{h(y_1,\dots,y_n)} \Rightarrow u = \sum y_i^2 \text{ is suff. for } \theta$$

Now, $E(u) = n \cdot E(Y^2)$ let $w = Y^2 \rightarrow Y = \sqrt{w}$ $\frac{dy}{dw} = \frac{1}{2\sqrt{w}}$



1-1 trans.

$$f(w) = f_Y(\sqrt{w}) \cdot \frac{1}{2\sqrt{w}} = \frac{2\sqrt{w}}{\theta} \cdot e^{-w/\theta} \cdot \frac{1}{2\sqrt{w}} = \frac{1}{\theta} e^{-w/\theta}, \ w > 0$$

$\Rightarrow w = Y^2 \sim EXP(\theta) \Rightarrow E(Y^2) = \theta$

so,

$E(\sum Y_i^2) = n\theta$ so let $\boxed{\hat{\theta}^* = h(u) = \frac{\sum y_i^2}{n}}$ (This turns out to be MVUE for θ)

Example 9.8

Let $Y_1 \cdots Y_n$ be a r.s. from $Y \sim N(\mu, \sigma^2)$

$$L(\mu, \sigma^2) = \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n \exp\left[-\frac{1}{2\sigma^2} \sum_{i=1}^{n} (y_i - \mu)^2\right]$$

$$= \underbrace{\left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n \exp\left[-\frac{1}{2\sigma^2}\left(\sum y_i^2 - 2\mu \sum y_i + n\mu^2\right)\right]}_{g(\sum y_i, \sum y_i^2, \mu, \sigma^2)} \cdot \underbrace{1}_{h(y_1,\dots,y_n)}$$

$\Rightarrow \sum y_i, \sum y_i^2$ are jointly suff. for μ, σ^2

We also know that $\bar{Y} = \frac{\sum y_i}{n}$, $E(\bar{Y}) = \mu$

$$s^2 = \frac{\sum (y-\bar{y})^2}{n-1} = \frac{\sum y_i^2 - \frac{(\sum y_i)^2}{n}}{n-1}, \quad E(s^2) = \sigma^2$$

$\Rightarrow$ $\bar{Y}$ and s^2 are unbiased fcts of sufficient stats. As it turns out, they are also MVUE for μ, σ^2.

Example 9.9

Let $Y_1, \dots, Y_n$ be a r.s. from $f(y|\theta) = \begin{cases} \frac{1}{\theta} e^{-y/\theta} & , y > 0 \\ 0 & , \text{—} \end{cases}$

find an MVUE for $V(Y_i)$

$Y \sim \text{Exp}(\theta) \Rightarrow var(Y) = \theta^2$

$$L(\theta) = \underbrace{\frac{1}{\theta^n} e^{-\frac{1}{\theta}\sum y_i}}_{g(u,\theta)} \cdot \underbrace{1}_{h(y_1,\dots,y_n)} \qquad \Rightarrow u = \sum y_i \text{ is suff. for } \theta$$

We need a fct. of $\sum y_i$ that is unbiased for θ^2 (not θ)

$E(\sum y_i) = n\theta \quad \Rightarrow E(\bar{y}) = \theta$

Try $E(\bar{Y}^2)$?

$$Var(\bar{Y}) = E(\bar{Y}^2) - E(\bar{Y})^2$$

$$\frac{\theta^2}{n} = E(\bar{Y}^2) - \theta^2 \quad \Rightarrow E(\bar{Y}^2) = \frac{(n+1)\,\theta^2}{n}$$

$$\Rightarrow \text{let } \hat{\theta}^* = h(u) = \frac{n}{n+1}\left(\frac{\sum Y_i}{n}\right)^2 = \boxed{\frac{(\sum Y_i)^2}{n(n+1)}}$$

9.61 $Y_1 \cdots Y_n$ is a r.s. from $Y \sim \text{unif}(0, \theta)$

Find an MVUE for θ.

in 9.49 we showed $Y_{(n)}$ is suff. for θ

$$f(y_n) = \frac{n}{\theta^n} y_n^{n-1}, \quad 0 < y_n < \theta$$

$$E[Y_{(n)}] = \int_0^\theta \frac{n}{\theta^n} y_n^n \, dy_n = \frac{y_n^{n+1}}{n+1} \cdot \frac{n}{\theta^n} \Big|_0^\theta = \theta\left(\frac{n}{n+1}\right)$$

let $\hat{\theta}^* = h(u) = \boxed{\frac{n+1}{n} Y_{(n)}}$ $\qquad E(\hat{\theta}^*) = \frac{n+1}{n} \cdot \frac{n}{n+1} \cdot \theta = \theta$

$\Rightarrow \frac{n+1}{n} Y_{(n)}$ is an unbiased estimator of θ, and a fct.

of a suff. stat.

9.6 Method of Moments

We use the method to estimate the k^{th} moment of a RV,
$\mu_k' = E(Y^k)$

Let $m_k' = \dfrac{\sum_{i=1}^{n} y_i^k}{n}$ be the k^{th} sample moment

Process: If we are interested in estimating $\theta_1 \cdots \theta_k$ then set:
$\mu_1' = m_1' \quad \mu_2' = m_2' \quad \cdots \quad \mu_k' = m_k'$
Use k equs to solve for $\hat{\theta}_1, \ldots, \hat{\theta}_k$ → method of moments estimators (MMEs)

e.g. let $Y_1 \cdots Y_n$ be a r.s. from $Y \sim POI(\lambda)$
Find the MME for λ:
$\mu_1' = E(Y) = \lambda \qquad m_1' = \dfrac{\sum_{i=1}^{n} y_i}{n} = \bar{Y}$

$\boxed{\hat{\lambda} = \bar{Y}}$ That was easy. Is this estimator any good?

Is it unbiased? $E(\bar{Y}) = \dfrac{n\lambda}{n} = \lambda$ ✓

Is it consistent? $var(\hat{\lambda}) = var(\bar{Y}) = \sigma^2/n = \lambda/n$

$\lim_{n\to\infty} var(\hat{\lambda}) = \lim_{n\to\infty} \dfrac{\lambda}{n} = 0$

by Thm 9.1, $\bar{Y} \xrightarrow{P} \lambda$ i.e. $\hat{\lambda}$ is consistent

Is it a fct. of a suff. stat?

$L(\lambda) = \prod_{i=1}^{n} \dfrac{e^{-\lambda}\lambda^{y_i}}{y_i!} = \underbrace{e^{-n\lambda}\cdot\lambda^{\sum y_i}}_{g(u,\lambda)} \underbrace{\left(\dfrac{1}{\prod_{i=1}^{n} y_i!}\right)}_{h(y_1,\ldots,y_n)}$ $u = \sum y_i$ is suff.

⇒ $\bar{Y}$ is a fct. of a suff. stat

These are all good signs that $\bar{Y}$ is a good estimator for λ.

Spse we want to estimate $p = P(Y=0) = e^{-\lambda}$ so $\lambda = -\ln p$

If we set $\bar{Y} = -\ln p$

$\Rightarrow \boxed{\hat{p} = e^{-\bar{Y}}}$ is the MME for $p = P(Y=0)$

This mean that MMEs have an invariance property:

i.e. MME for $h(\theta)$ is $h(\hat{\theta})$

Example 9.11

Let $Y_1 \cdots Y_n$ be a r.s from $Y \sim unif(0, \theta)$

Find the MME for θ:

$\mu_1' = E(Y) = \frac{\theta}{2}$

$\frac{\theta}{2} \overset{SET}{=} \bar{Y} \Rightarrow \boxed{\hat{\theta} = 2\bar{Y}}$

Is $2\bar{Y}$ consistent for θ?

- $E(2\bar{Y}) = 2E(\bar{Y}) = 2\left(\frac{\theta}{2}\right) = \theta$ unbiased ✓

- $var(2\bar{Y}) = 4 \cdot var(\bar{Y}) = 4 \frac{var(Y)}{n} = \frac{4\theta^2}{12n} = \frac{\theta^2}{3n}$

$\lim_{n\to\infty} \frac{\theta^2}{3n} = 0 \Rightarrow \hat{\theta} = 2\bar{Y}$ is consistent

Is it a fct. of a suff. stat?

$L(\theta) = \prod_{i=1}^{n} \frac{1}{\theta} \cdot I_{(0,\theta)}(y_i) = \prod_{i=1}^{n} \frac{1}{\theta} \cdot I_{(0,\infty)}(y_{(1)}) \cdot I_{(-\infty,\theta)}(y_{(n)})$

$= \left[\frac{1}{\theta^n} \cdot I_{(-\infty,\theta)}(y_{(n)})\right]\left[I_{(-\infty,0)}(y_{(1)})\right], \quad -\infty < y_i < \infty$

$\Rightarrow Y_{(n)}$ is suff for θ. Recall that we showed $\hat{\theta} = \frac{n+1}{n} Y_{(n)}$ is unbiased for θ.

So, $2\bar{Y}$ is consistent, but $\frac{n+1}{n} Y_{(n)}$ is probably a better estimator because it is an unbiased fct. of a suff. stat.

Example 9.13

$Y_1 \cdots Y_n$ is a r.s. from $Y \sim GAM(\alpha, \beta)$

Find the MMEs for α, β:

For 2 parameters, we need 2 eqns:

$$\mu_1' = \alpha\beta \quad , \quad \mu_2' = \alpha\beta^2 + (\alpha\beta)^2 = \alpha\beta^2(1+\alpha)$$

$$\bar{Y} = \alpha\beta \qquad \frac{\sum Y_i^2}{n} = \alpha\beta^2 + \alpha^2\beta^2$$

$$\hat{\beta} = \frac{\bar{Y}}{\hat{\alpha}} \rightarrow \frac{\sum Y_i^2}{n} = \hat{\alpha}\left(\frac{\bar{Y}}{\hat{\alpha}}\right)^2 + \hat{\alpha}^2\left(\frac{\bar{Y}^2}{\hat{\alpha}^2}\right)$$

$$\frac{\sum Y_i^2}{n} = \frac{\bar{Y}^2}{\hat{\alpha}} + \bar{Y}^2$$

$$\Rightarrow \frac{\bar{Y}^2}{\hat{\alpha}} = \frac{\sum Y_i^2}{n} - \bar{Y}^2 \Rightarrow \hat{\alpha} = \frac{\bar{Y}^2}{\frac{\sum Y_i^2}{n} - \bar{Y}^2} = \boxed{\frac{n\bar{Y}^2}{\sum(Y_i - \bar{Y})^2}}$$

$$\Rightarrow \hat{\beta} = \frac{\bar{Y}}{\frac{n\bar{Y}^2}{\sum(Y_i-\bar{Y})^2}} = \boxed{\frac{\sum(Y_i - \bar{Y})^2}{n\bar{Y}}}$$

---

are $\hat{\alpha}$ and $\hat{\beta}$ consistent?

$$E(\bar{Y}) = \alpha\beta \text{ and } var(\bar{Y}) = \frac{\alpha\beta^2}{n} \text{ and } \lim_{n\to\infty} \frac{\alpha\beta^2}{n} = 0$$

$$\Rightarrow \bar{Y} \xrightarrow{P} \alpha\beta$$

$$E(Y_i^2) = \alpha\beta^2 + \alpha^2\beta^2 \Rightarrow E\left[\frac{\sum Y_i^2}{n}\right] = \alpha\beta^2 + \alpha^2\beta^2$$

$$var(Y_i^2) = E(Y_i^4) - E(Y_i^2)^2$$

$$E(Y^4) = \int_0^\infty \frac{1}{\Gamma(\alpha)\beta^\alpha} y^{\alpha+4-1} e^{-y/\beta}\, dy$$

$$= \frac{1}{\Gamma(\alpha)\beta^\alpha} \cdot \Gamma(\alpha+4)\cdot\beta^{\alpha+4} \cdot \cancel{\int_0^\infty \frac{1}{\Gamma(\alpha+4)\beta^{\alpha+4}} y^{\alpha+4-1} e^{-y/\beta}\, dy}^{\,1}$$

$$= \frac{(\alpha+3)!}{(\alpha-1)!}\beta^4 = (\alpha+3)(\alpha+2)(\alpha+1)\alpha\cdot\beta^4$$

$$\Rightarrow var(Y^2) = (\alpha+3)(\alpha+2)(\alpha+1)\alpha\cdot\beta^4 - [\alpha\beta^2 + \alpha^2\beta^2]$$

This looks awful, but all we need to do is show $Var\left(\frac{\sum Y_i^2}{n}\right)$ goes to 0 as $n \to \infty$

$$var\left(\frac{\sum Y_i^2}{n}\right) = \frac{Var(Y^2)}{n^2}$$ which goes to 0 as $n \to \infty$

$$\Rightarrow \frac{\sum Y_i^2}{n} \xrightarrow{P} \alpha\beta^2 + \alpha^2\beta^2$$

Now, by Thm 9.2 (d) if $\hat{\theta}_n \xrightarrow{P} \theta$ then $g(\hat{\theta}_n) \xrightarrow{P} g(\theta)$

$$\hat{\alpha} = \frac{\bar{Y}^2}{\frac{\sum Y_i^2}{n} - \bar{Y}^2} \xrightarrow{P} \frac{(\alpha\beta)^2}{\alpha\beta^2 + \alpha^2\beta^2 - (\alpha\beta)^2} = \frac{\alpha^2\beta^2}{\alpha\beta^2} = \alpha$$

$$\hat{\beta} = \frac{\bar{Y}}{\hat{\alpha}} \xrightarrow{P} \frac{\alpha\beta}{\alpha} = \beta$$

$\hat{\alpha}, \hat{\beta}$ are consistent estimators. are they fcts. of suff. stats?

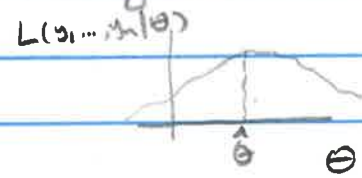
$$L(\alpha,\beta) = \underbrace{\left(\frac{1}{\Gamma(\alpha)\beta^{\alpha}}\right)^{n}\left(\prod_{i=1}^{n} y_i\right)^{\alpha-1} \cdot e^{-\sum y_i/\beta}}_{g(u_1, u_2, \alpha, \beta)} \cdot \underbrace{1}_{h(y_1, \dots y_n)}$$

$\Rightarrow \prod_{i=1}^{n} y_i$, $\sum y_i$ are jointly sufficient for α, β

$\Rightarrow$ MMEs are not fcts of suff. stats, so we could probably find better estimators.

9.7 Method of Maximum Likelihood

Spse that the likelihood fct depends on k parameters $\theta_1, \dots, \theta_k$. Choose $\hat{\theta}_1 \cdots \hat{\theta}_k$ to be values that maximize the likelihood. The estimates are usually referred to as MLEs



Conceptual example: (p. 476)

contains 3 balls that are either red or white

We need to guess what the contents are, 2 red/1 white, 3 red, etc. Let τ = total no. of red balls

We draw 2 balls w.o.r. and they are both red

$$L(\tau=2) = \frac{\binom{2}{2}\binom{1}{0}}{\binom{3}{2}} = 1/3$$

$$L(\tau=3) = \frac{\binom{3}{2}}{\binom{3}{2}} = 1$$

$\Rightarrow L(\tau)$ is maximized if $\tau = 3$

$\Rightarrow \hat{\tau} = 3$ is the MLE for τ.

---

e.g. Let $Y_1 \cdots Y_n$ be a r.s. from $POI(\lambda)$

$$L(\lambda) = \frac{e^{-n\lambda} \lambda^{(\sum y_i)}}{\prod_{i=1}^{n} y_i!}$$

We could try to maximize $L(\lambda)$ directly, but it would be easier to maximize $\ln L(\lambda)$

note: $\ln L(\lambda)$ is a monotone increasing fct of $L(\lambda)$

$$\ln L(\lambda) = -n\lambda + (\Sigma y_i)\ln\lambda - \ln\left(\prod_{i=1}^{n} y_i!\right)$$

we want to maximize $L(\lambda)$ w.r.t. λ so differentiate w.r.t. λ (y_i treated as fixed constants - sample has been taken)

$$\frac{\partial \ln L(\lambda)}{\partial \lambda} = -n + \frac{\Sigma(y_i)}{\lambda} \overset{\text{SET}}{=} 0 \Rightarrow \boxed{\hat{\lambda} = \frac{\Sigma y_i}{n} = \bar{Y}}$$

This matches the MME which was consistent and a fct of a suff. stat. It is also MVUE.

Like the MME, the MLE also has the invariance property (9.94)

i.e. MLE for $\tau(\theta)$

is $\boxed{\hat{\tau}(\theta) = \tau(\hat{\theta})}$

Find MLE for $P(Y=0) = e^{-\lambda} = \boxed{e^{-\bar{Y}}}$

MLEs also usually turn out to be fcts. of suff. stats (unlike MMEs) which also means they tend to be MVUEs

Thm Let $Y_1 \cdots Y_n$ be a r.s. from a dist. w/ density fct $f(y|\theta)$. If a suff. stat $u(Y_1, \dots, Y_n)$ for θ exists and if an MLE $\hat{\theta}$ of θ also exists uniquely, then $\hat{\theta}$ is a fct. of the suff. stat $u(Y_1, \dots, Y_n)$.

Pf:

If u is suff., by factorization thm

$$L(\theta) = L(y_1, \dots, y_n | \theta) = \underbrace{g(u, \theta)}_{\text{only a fct of } u \text{ and } \theta} \cdot \underbrace{h(y_1, \dots, y_n)}_{\text{does not depend on } \theta}$$

$\ln[L(\theta)] = \ln[g(u,\theta)] + \ln[h(y_1,\dots,y_n)]$

maximizing $\ln[L(\theta)]$ relative to θ is equivalent to maximizing $\ln[g(u,\theta)]$ relative to θ

Since $\ln[g(u,\theta)]$ is only a fct of u and θ, $\hat{\theta}$ must be a fct of u

$\Rightarrow$ If we find a unique MLE $\hat{\theta}$ and adjust it to be unbiased, if necessary, it will most likely be an MVUE.

e.g. Let $Y_1 \cdots Y_n$ be a r.s. from $Y \sim N(\mu, \sigma^2)$

$$L(\mu,\sigma^2) = \frac{1}{(2\pi)^{n/2}\sigma^n} e^{-\frac{1}{2\sigma^2}\Sigma(y_i-\mu)^2}$$

$$\ln L = -\frac{n}{2}\ln(2\pi) - n\ln(\sigma) - \frac{1}{2\sigma^2}\Sigma(y_i-\mu)^2$$

1 $$\frac{\partial \ln L}{\partial \mu} = +\frac{\Sigma(y_i-\mu)}{\sigma^2} \overset{\text{SET}}{=} 0$$

2 $$\frac{\partial \ln L}{\partial \sigma} = \frac{-n}{\sigma} + \frac{1}{\sigma^3}\Sigma(y_i-\mu)^2 \overset{\text{SET}}{=} 0$$

1 $\Sigma(y_i-\mu)=0 \Rightarrow \Sigma y_i = n\cdot\mu \Rightarrow \boxed{\hat{\mu}=\bar{y}}$

2 $\frac{\Sigma(y_i-\bar{y})^2}{\sigma^3} = \frac{n}{\sigma} \Rightarrow \boxed{\hat{\sigma}^2 = \frac{\Sigma(y_i-\bar{y})^2}{n}}$

From prior work ... $\hat{\mu} = \bar{Y}$ is unbiased, fct of suff. stat.

$\Rightarrow$ MVUE for μ

$\hat{\sigma}^2 = \frac{\Sigma(y_i - \bar{y})^2}{n}$ is biased, but is a fct of a suff. stat. (s^2)

$\hat{\sigma}^2 = \frac{(n-1)s^2}{n}$. a simple adjustment $\frac{n}{n-1} \cdot \hat{\sigma}^2 = s^2$ will be MVUE for σ^2

---

e.g. Let $Y_1 \cdots Y_n$ be a r.s. from unif $(0, \theta)$

note: having θ in the support can cause problems trying to find an MLE

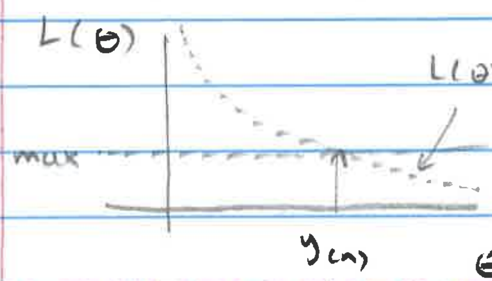
$L(\theta) = \frac{1}{\theta^n}$ $\quad \ln L(\theta) = -n \ln \theta$

$\frac{\partial \ln L}{\partial \theta} = \frac{-n}{\theta} \overset{\text{SET}}{=} 0$ $\quad$ no soln.

Re-examine $L(\theta)$:

$$L(\theta) = \begin{cases} \frac{1}{\theta^n}, & 0 < y_i < \theta \quad i = 1 \text{ to } n \\ 0, & \text{———} \end{cases}$$

or $$L(\theta) = \begin{cases} \frac{1}{\theta^n}, & y_{(n)} < \theta \\ 0, & \text{———} \end{cases}$$



L(θ) only exists out here

$\Rightarrow$ $\boxed{\hat{\theta} = Y_{(n)}}$ is the MLE for θ

We showed before that $Y_{(n)}$ is suff for θ, but $E(Y_{(n)}) = \frac{n\theta}{n+1}$

a simple adjustment $\frac{n+1}{n} \cdot Y_{(n)}$ will be an MVUE for θ.

e.g. let $Y_1 \cdots Y_n$ be a r.s. from $Y \sim unif(\theta, 2\theta)$, $\theta > 0$
Find the MLE for θ

$L(\theta) = \frac{1}{\theta^n}$ so a gum $\frac{\partial \ln L(\theta)}{\partial \theta} = \frac{-n}{\theta} \stackrel{\text{set}}{=} 0$ has no soln.

We know $\theta < Y_{(1)}$ and $Y_{(n)} < 2\theta$

$$\Rightarrow \frac{Y_{(n)}}{2} < \theta < Y_{(1)}$$

What value of θ will maximize $L(\theta) = \frac{1}{\theta^n}$? The smallest value

$\Rightarrow$ let $\boxed{\hat{\theta} = \frac{Y_{(n)}}{2}}$

Is this a fct of a suff stat?

$L(\theta) = \frac{1}{\theta^n}$, $0 < y_i < \theta$

$$= \underbrace{\frac{1}{\theta^n} \cdot I_{(\theta,\infty)}(y_{(1)}) \cdot I_{(-\infty,2\theta)}(y_{(n)})} \cdot \underset{h(y_1, \ldots y_n)}{1}$$

$\Rightarrow Y_{(1)}$ and $Y_{(n)}$ are jointly suff. for θ

$\Rightarrow \hat{\theta} = \frac{Y_{(n)}}{2}$ is a fct. of a suff. stat

$$F(y) = \frac{y-\theta}{\theta} \Rightarrow F_{Y_{(n)}}(y) = \left(\frac{y-\theta}{\theta}\right)^n$$

$$\Rightarrow f_{(n)}(y) = n\left(\frac{y-\theta}{\theta}\right)^{n-1} \cdot \frac{1}{\theta} = \frac{n(y-\theta)^{n-1}}{\theta^n}$$

$$E(Y_{(n)}) = \int_{\theta}^{2\theta} \frac{n \cdot y(y-\theta)^{n-1}}{\theta^n} \qquad \text{let } u = y-\theta, \quad du = dy$$

$$= \int_0^{\theta} \frac{n(u+\theta)\cdot u^{n-1}}{\theta^n} = \frac{n}{\theta^n}\int_0^{\theta}\left[u^n + \theta u^{n-1}\right] du$$

$$= \frac{n}{\theta^n}\left[\frac{u^{n+1}}{n+1} + \theta \cdot \frac{u^n}{n}\right]\Bigg|_0^{\theta}$$

$$= \frac{n}{\theta^n}\left[\frac{\theta^{n+1}}{n+1} + \frac{\theta^{n+1}}{n}\right] = \theta\left[1 + \frac{n}{n+1}\right]$$

$$\Rightarrow E\left[\frac{Y_{(n)}}{2}\right] = \frac{\theta}{2}\left[\frac{2n}{n+1}\right] = \frac{\theta n}{n+1}$$

Slight adjustment to MLE: $\boxed{\frac{n+1}{n}\left[\frac{Y_{(n)}}{2}\right]}$