

Lecture 9

For an exponential family \mathcal{F} of full rank the statistic (T_1, \dots, T_S) is complete sufficient, therefore by considering unbiased estimator on it, one obtains the UMVU.

How to do this in practical situations?

If \bar{T} is already a function of the complete sufficient statistic, we are done.

If not...

1) Solving - equations

If T is complete sufficient, then finding the \bar{T} that satisfies

$$\mathbb{E}_\theta \bar{T}(T) = g(\theta) \quad \forall \theta \in \Omega$$

gives us the UMVU of $g(\theta)$.

Ex. If $\bar{X} \in \text{Bin}(n, p)$, then $T = \bar{X}_1$ and we want to estimate $g(p) = p(1-p) = pq$. ($g = 1-p$).

We are to solve

①

$$E_p(\delta(\tau)) = \sum_{t=0}^n \binom{n}{t} \delta(t) p^t q^{n-t} = pq \quad (0 < p < 1)$$

$$\begin{aligned} \Leftrightarrow \sum_{t=0}^n \binom{n}{t} \delta(t) \left(\frac{p}{q}\right)^t &= pq^{1-n} = \\ &= \frac{p}{q} \cdot \left(\frac{1}{q}\right)^{n-2} \\ &= \frac{p}{q} \left(1 + \frac{p}{q}\right)^{n-2} \end{aligned}$$

$$\left(p = \frac{p}{q}\right) \Leftrightarrow \sum_{t=0}^n \binom{n}{t} \delta(t) p^t = p \sum_{k=0}^{n-2} \binom{n-2}{k} p^k$$

$$= \sum_{t=1}^{n-1} \binom{n-2}{t-1} p^t$$

Identify LHS and RHS coefficients

$$\binom{n}{0} \delta(0) + \binom{n}{1} \delta(1) p + \binom{n}{2} \delta(2) p^2 t \dots$$

$$\dots + \binom{n}{n} \delta(n) p^n = \binom{n-2}{0} p + \binom{n-2}{1} p^2 t \dots$$

$$\dots + \binom{n-2}{n-1} p^{n-1}$$

i.e.

(2)

$$\left. \begin{array}{l} \binom{n}{0} \delta(0) = 0 \\ \binom{n}{1} \delta(1) p = \binom{n-2}{0} p \\ \vdots \\ \binom{n}{2} \delta(2) p^2 = \binom{n-2}{1} p^2 \\ \vdots \\ \binom{n}{n-1} \delta(n-1) p^{n-1} = \binom{n-2}{n-1} p^{n-1} \\ \binom{n}{n} \delta(n) p^n = 0 \end{array} \right\}$$

\Leftrightarrow

$$\left. \begin{array}{l} \binom{n}{k} \delta(k) = \binom{n-2}{k-1} \Leftrightarrow \delta(k) = \binom{n-2}{k-1} / \binom{n}{k} \\ = \frac{k(n-k)}{n(n-1)} \\ k=1, \dots, n-1 \\ \delta(0) = \delta(n) = 0 \end{array} \right\}$$

#

(3)

2) Conditioning on a complete sufficient statistic.

Take any unbiased estimator δ of $g(\theta)$ and condition on T .

Ex. 1.14 Read on your own, do exercise 1.6.2.

#

We know that if loss function $L(\theta, \delta)$ is convex in δ , then there is an UMVU of an U-estimable $g(\theta)$.

→ (and even claim that there is an estimator that uniformly minimizes the risk over all unbiased estimators).

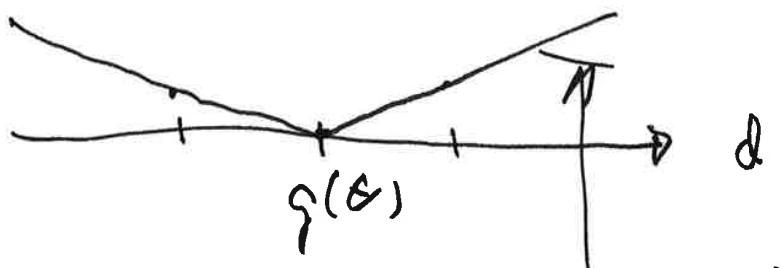
What happens if $L(\theta, \delta)$ not convex in δ ?

Theorem (Berger)

Assume that the loss function $L(\theta, d)$ satisfies

$$L(\theta, g(\theta)) = 0 \quad (*)$$

$$L(\theta, d) \leq M < \infty$$



this function
is convex, does not
satisfy the
restrictions (*)

Assume that $g(\theta)$ is V -estimable and θ_0 fixed but arbitrary. Then there exists a sequence of unbiased estimators $\hat{\theta}_n$ such that

$$R(\theta_0, \hat{\theta}_n) \rightarrow 0 ; n \rightarrow \infty.$$

Proof

Let $\hat{\delta}(x)$ be an unbiased estimator of $g(\theta)$. Define

$$\hat{\delta}'_p(x) = \begin{cases} g(\theta_0) & \text{with prob } 1-p \\ \frac{1}{p}(\hat{\delta}(x) - g(\theta_0)) + g(\theta_0) & \text{with prob } p \end{cases}$$

for every p . Then

$$\begin{aligned} E_{\theta}(\hat{\delta}'_p(x)) &= g(\theta_0)(1-p) + \\ &\quad + \left(\frac{1}{p} (E_{\theta}(\hat{\delta}(x)) - g(\theta_0)) + g(\theta_0) \right) p \end{aligned}$$

Note that $\hat{\delta}'_p(x)$ is a randomized estimator, i.e. we can write it as

$$\hat{\delta}'_p(x) = \tilde{\delta}(x, Y)$$

with Y a random number.

Here we can write

$$Y = \begin{cases} 0 & \text{with } 1-p \\ 1 & \text{with prob p.} \end{cases}$$

and

$$\tilde{\delta}(x, 0) = g(\theta_0)$$

$$\tilde{\delta}(x, 1) = \frac{1}{p} (\delta(x) - g(\theta_0)) + g(\theta_0)$$

Therefore

$$\begin{aligned} \bar{E}_{\theta_0}(\delta_p'(\bar{x})) &= \bar{E}_{\theta_0}(\tilde{\delta}(\bar{x}, Y)) \\ &= \bar{E}_{\theta_0}(\bar{E}_{\theta_0}(\tilde{\delta}(\bar{x}, Y)) | Y) \\ &= \bar{E}_{\theta_0} \left[g(\theta_0)(1-p) + \right. \\ &\quad \left. + \cancel{g(\theta_0)(\delta(\bar{x}))} \cancel{+ \left(\frac{1}{p} (\delta(\bar{x}) - g(\theta_0)) \right)} \right. \\ &\quad \left. + g(\theta_0) \right] \cdot p \\ &= g(\theta_1)(1-p) + \left(\bar{E}_{\theta_0}(\bar{E}_{\theta_0}(\delta(\bar{x}) | Y=1)) \right. \\ &\quad \left. - g(\theta_0) \right) + g(\theta_0)p \end{aligned}$$

⑦

$$= g(\theta_0)$$

And with

$$\begin{aligned}
 R(\theta_0, \tilde{\delta}_p') &= \\
 &= \bar{E}_{\theta} (\bar{E}_{\theta_0} (\lambda(\theta_0, \tilde{\delta}(x, y)) | Y)) \\
 &= \bar{E}_{\theta} (\underbrace{\bar{E}_{\theta} (\lambda(\theta_0, \tilde{\delta}(x, 0)) | Y=0)}_{=0} \cdot (1-p) \\
 &\quad + \bar{E}_{\theta} (\bar{E}_{\theta} (\lambda(\theta_0, \tilde{\delta}(x, 1)) | Y=1) p) \\
 &\leq M \cdot p \rightarrow 0 ; p \rightarrow 0
 \end{aligned}$$

Take

$$n = \left[\frac{1}{p} \right]$$

recall M is
the bound on
the loss
function

(8)

For any fixed θ_0 the size of δ_0
 can be made arbitrarily small, and
 unbiasedness does not help here.

Examples to continuous and discrete problems.

Types of continuous problems...

Assume x_1, \dots, x_n i.i.d. data from
 location family

$$P_\theta(\bar{x}_i \leq x) = F(x - \theta) \quad (1)$$

~~Different levels~~

Different levels of complexity

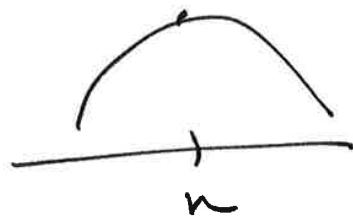
a) F completely unspecified.

(b) $P_\theta(\bar{x}_i \leq x) = F_\theta\left(\frac{x - \theta}{\sigma}\right)$
 F_θ completely specified.)

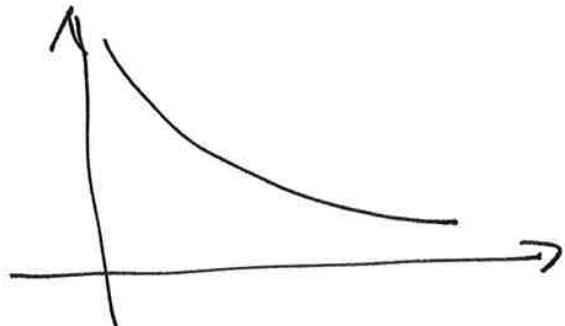
⑨ not a special case of (1)

b) different forms of symmetry
or shape

- (i) F "symmetric" centered Θ
- (ii) F has finite variance
- (iii) $F' = f$ exists and is
monotone (e.g. Φ or IR_+)
unimodal



- (iv) $F' = f$ exists and is convex
- (v) f analytic.



- (vi) $f = F'$ exists and lies
in Besov, Sobolev...
spaces

The multivariate normal one-sample problem

Assume we have n p -vectors that are i.i.d. p -variate normal $p=2$. Thus (\bar{x}_i, y_i) , $i=1, \dots, n$ (\bar{x}_i, y_i) are 2-dim r.v. Gaussian for each i .

Joint density of the sample is

$$\left(\frac{1}{2\pi\sigma\sqrt{1-\rho^2}} \right)^n \exp\left(-\frac{1}{2(1-\rho^2)} \cdot \left[\frac{1}{\sigma^2} \sum (\bar{x}_i - \xi)^2 - \frac{2\rho}{\sigma\tau} \sum (\bar{x}_i - \xi)(y_i - \eta) + \frac{1}{\tau^2} \sum (y_i - \eta)^2 \right] \right)$$

with $\xi = E(\bar{x}_i)$, $\eta = E(y_i)$, σ^2 , τ^2 variances and $\rho = \frac{\text{Cov}(\bar{x}_i, y_i)}{\sigma \cdot \tau}$ is correlation.

Five parameter exponential family (11)

of full rank and

$$T = (\bar{x}, \bar{y}, s^2_{xx}, s^2_y, s^2_{xy})$$

is the complete sufficient statistic.

To find the UMVU:

The marginals of (X_i, Y_i) are;

$$X_i \in N(\mu, \sigma^2)$$

$$Y_i \in N(\eta, \tau^2),$$

and unbiased estimators of μ, σ^2 are

$$\begin{array}{lll} \bar{x} & \text{unbiased est. of } \mu \\ \frac{s^2_x}{n-1} & -/- & \sigma^2 \end{array}$$

and, since functions of T , they are the UMVUs.

Problem 2.10. Show that

$$\frac{s^2_{xy}}{n-1}$$

is unbiased of $\rho\sigma^2$ (the covariance)

(12)

and therefore the UMVU of the covariance.

Want to estimate correlation $\frac{\sigma_{xy}}{\sigma_x \sigma_y}$

$$\rho = \frac{\text{cov}(X, Y)}{\sigma_x \sigma_y} = h(\text{cov}, \sigma_x, \sigma_y)$$

then can use estimate $\hat{\rho} = h(\hat{\text{cov}}, \hat{\sigma}_x, \hat{\sigma}_y)$

$$R = \frac{s_{xy}^2}{\sqrt{s_x^2 s_y^2}} = h(\hat{\text{cov}}, \hat{s}_x, \hat{s}_y)$$

(note this is actually ~~the~~ ^a plug-in estimator

$$\hat{\theta} = h(\gamma)$$

where θ interesting parameter, γ another parameter that I already can estimate, with $\hat{\gamma}$, then

$$\hat{\theta} = h(\hat{\gamma})$$

is the plug-in estimator.)

However R not unbiased.

Can adjust R to bias-correct.

(13)

Read more in the book.

We can get
 $G(R)$

unbiased estimator, and

$$G(R) = R \left(1 + \frac{1-R^2}{2(n-1)} + G\left(\frac{1}{n}\right) \right)$$

exact calculation of R or computer.

Extension to $p > 2$, need on your own.

#

(4)