

## UMVU estimators

## Lecture 8

Recall that a uniformly best estimator does not exist. Solutions:

- a) restrict to unbiased estimators
- b) restrict to estimators that are invariant under scale and location transformation
- c) minimax theory, Bayes theory, using an "overall" measure of the risk of an estimator.

UMVU uniformly minimum variance unbiased.

Def. An estimator  $\delta$  of  $g(\theta)$  is called unbiased if

$$E_{\theta}(\delta(z)) = g(\theta) \quad \forall \theta \in \mathcal{J}$$

#

Do unbiased estimators always exist?

If so, if  $\exists$  unbiased estimator of  $g(\theta)$  we call  $\delta$ , or  $g(\theta)$ , U-estimable.

(1)

Ex:  $X \in \text{Bin}(n, p)$ . Want to estimate

$$g(p) = \frac{1}{p}.$$

$\exists$  unbiased estimator? Assume  $\delta(X)$  is unbiased. This means that

$$\underset{p}{E}(\delta(X)) = \sum_{k=0}^n \delta(k) \binom{n}{k} p^k (1-p)^{n-k}$$

$$= g(p) = \frac{1}{p}$$

we want this.

for all  $0 < p < 1$ .

But: If  $p \rightarrow 0$  then

$$g(p) \rightarrow \infty$$

$$E(\delta(X)) \rightarrow \delta(0)$$

Not possible! Therefore there is no unbiased estimator of  $g(p) = \frac{1}{p}$ .

#

Now suppose  $f(\theta)$  estimable and let us look at all unbiased estimators of  $g(\theta)$ .

(2)

There can be characterized by

$$\{\text{unbiased est. of } g(\theta)\} = \text{(*)}$$

$$\rightarrow \mathcal{J}_0 - \{\text{unbiased est. of } \theta\}$$

fixed unbiased est. of  $g(\theta)$ .

or

$$\mathcal{J} = \mathcal{J}_0 - U \leftarrow \begin{array}{l} \text{arbitrary unbiased} \\ \text{est. of } \theta. \end{array}$$

$\nearrow$

$\nwarrow$

arbitrary unbiased est. of  $g(\theta)$       fixed unbiased est. of  $f(\theta)$

---

To show this (\*) one does the standard proof of the sets  $A, B$  being equal  $A=B$  by showing  $A \subseteq B$  and  ~~$B \subseteq A$~~   $A \supseteq B$

( $\subseteq$ )  $\mathcal{J}$  arbitrary unbiased est. of  $g(\theta)$   
so in the LHS, and  $\mathcal{J}_0$  a fixed unbiased est. of  $g(\theta)$ . Then  $U = \mathcal{J}_0 - \mathcal{J}$  is unbiased est. of  $\theta$  and  $\mathcal{J} = \mathcal{J}_0 - U$  so that  $\mathcal{J}$  is in the RHS.

( $\supseteq$ )  $\mathcal{J}_0$  fixed unbiased estimator of  $\theta$   
and  $U$  arbitrary unbiased of  $\theta$ , so

(3) that  $\mathcal{J}_0 - U = \mathcal{J}$  is in RHS. Then  $\mathcal{J}$  is unbiased est. of  $g(\theta)$  so  $\mathcal{J}$  is LHS.

If we let

$$\mathcal{U} = \{U \text{ est. of } \phi\}$$

mean

$$E_\theta(U(\bar{x})) = \phi \quad \forall U \in \mathcal{U}.$$

We want to find the UMVU. If  $\hat{\sigma}$  is unbiased estimator of  $g(\phi)$ , we want to minimize

$$\text{Var}_\theta(\hat{\sigma}(\bar{x}))$$

i.e. want to find the  $\hat{\sigma}$  that minimizes this (looking over all unbiased est. of  $g(\phi)$ ). We see that by character. (\*)

$$\text{Var}_\theta(\hat{\sigma}(\bar{x})) = \cancel{\text{Var}_\theta}((\hat{\sigma}_0(\bar{x}) - U(\bar{x}))^2)$$

$$= E_\theta((\hat{\sigma}_0(\bar{x}) - U(\bar{x}))^2)$$

$$- (E_\theta[\hat{\sigma}_0(\bar{x}) - U(\bar{x})])^2$$

$$= E_\theta((\hat{\sigma}_0(\bar{x}) - U(\bar{x}))^2) - (g(\theta))^2$$

(4)

Want to minimize

$$\sum_{i=1}^n (x_i - a)^2$$

over  $a$ . Done by

$$\hat{a} = \frac{1}{n} \sum_{i=1}^n x_i$$

Want to minimize

$$\sum (x_i - (c_0 - b))^2$$

over  $b$ . Done by

$$c_0 - \hat{b} = \underbrace{(c_0 - b)}_{\text{fixed}} = \frac{1}{n} \sum x_i$$

i.e.

$$\hat{b} = c_0 - \frac{1}{n} \sum x_i = c_0 - \hat{a}$$

i.e.

$$\underset{b}{\operatorname{argmin}} \sum (x_i - (c_0 - b))^2$$

$$= c_0 - \underset{a}{\operatorname{argmin}} \sum (x_i - a)^2$$

Therefore

$$\underset{\begin{array}{l}\text{argmin} \\ \delta \in \text{unbiased} \\ \text{est. of } g(\theta)\end{array}}{\text{Var}_{\theta}}(\delta(x)) =$$

$$= \delta(x) - \underset{\begin{array}{l}\text{argmin} \\ \delta \in \text{unbiased} \\ \text{est. of } \theta\end{array}}{\text{E}}((\delta_0(x) - \delta(x))^2)$$

Def. An unbiased estimator  $\delta$  of  $g(\theta)$  is called UMVU (uniformly minimum variance unbiased) if

$$\cancel{\text{Var}_{\theta} \delta(x)} \leq \cancel{\text{Var}_{\theta} \delta'(x)}$$

$$\text{Var}_{\theta} \delta(x) \leq \text{Var}_{\theta} \delta'(x) \quad \forall \theta \in \Omega$$

for any other unbiased estimator  $\delta'$  of  $g(\theta)$ . It is called LMVU (locally minimum variance unbiased) at  $\theta_0$  if

$$\text{Var}_{\theta_0} \delta(x) \leq \text{Var}_{\theta_0} \delta'(x)$$

for any other unbiased estimator of  $g(\theta)$ .

Problem 1.12 Show that if an UMVU exists it is unique.

Example 1.8 Read on your own.

Characterization of the UMVU estimators

Let

$$\Delta = \{ \delta : E_\theta (\delta(X))^2 < \infty \quad \forall \theta \in \Omega \}$$

$$U = \{ U : E_\theta (U(X)) = 0, \quad \dots, \quad E_\theta (U^2(X)) < \infty, \quad \forall \theta \in \Omega \}$$

Theorem 1.7 .

Assume  $X \sim F_\theta$  for  $\theta \in \Omega$ . Let

$U \in U$ ,  $\delta \in \Delta$ . Then if  $E_\theta (\delta(X)) = g(\theta)$   
 $\forall \theta \in \Omega$

$$\begin{matrix} \delta \text{ is UMVU} \\ \Leftrightarrow \end{matrix}$$

$$\textcircled{6} \quad E_U (\delta(X) U(X)) = 0 \quad \forall U \in U \quad \forall \theta \in \Omega$$

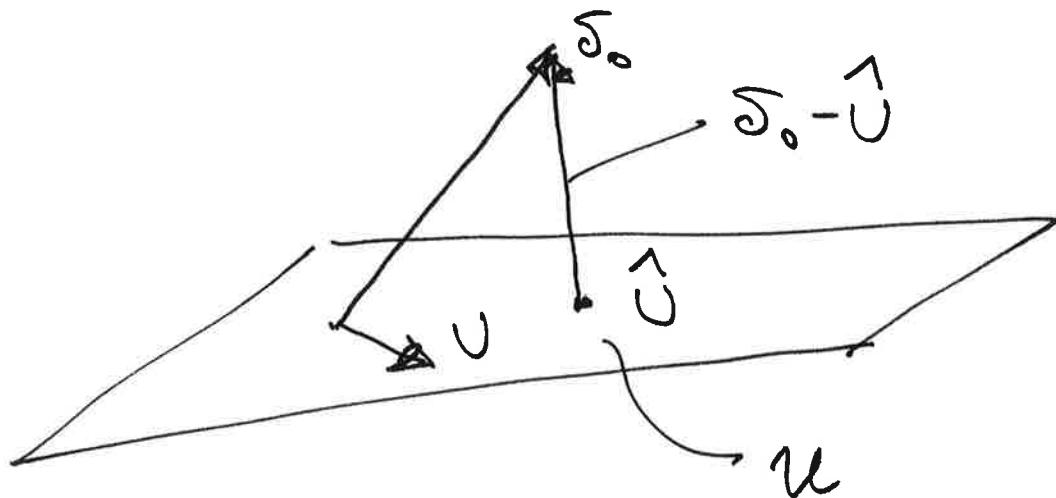
Proof

We know that  $\hat{\sigma} = \sigma_0 + \hat{U}$  is obtained  
as

$$\hat{\sigma} = \sigma_0 - \hat{U}$$

where

$$\hat{U} = \underset{U \in \mathcal{U}}{\operatorname{argmin}} E_{\theta} ((\sigma_0(\tilde{x}) - U(\tilde{x}))^2).$$



$$U \perp \sigma_0 - \hat{U} = \hat{\sigma} \quad \forall U \in \mathcal{U}$$

But

$$\hat{U} = \underset{U \in \mathcal{U}}{\operatorname{argmin}} E_{\theta} ((\sigma_0 - U)^2)$$

$$= \underset{U \in \mathcal{U}}{\operatorname{argmin}} \| \sigma_0 - U \|_{F_{\theta}}^2$$

(7)

where

$$\|h(x)\|_{F_\theta}^2 = \int h(x)^2 dF_\theta(x)$$

and we can see this as plus by  
a scalar product

$$\begin{aligned}\langle h(x), u(x) \rangle_{F_\theta} &= \\ &= \int h(x) u(x) dF_\theta(x)\end{aligned}$$

so that

$$\|h(x)\|_{F_\theta}^2 = \langle h(x), h(x) \rangle$$

Then we know that

$$\hat{U} = \underset{U \in \mathcal{U}}{\text{argmin}} \quad E_S((\mathcal{S}_0(x) - U(x))^2)$$

$$= \underset{U \in \mathcal{U}}{\text{argmin}} \quad \| \mathcal{S}_0 - U \|_{F_\theta}^2$$

i.e.

$$\langle \mathcal{S}_0 - \hat{U}, U \rangle_{F_\theta} = 0 \quad \forall U \in \mathcal{U}$$

and

$$\begin{aligned}
 & \langle \delta_0 - \hat{U}, U \rangle = \langle \hat{\mathcal{J}}, U \rangle_{F_\Theta} \\
 &= \int_{\mathbb{R}} \hat{\mathcal{J}}(x) U(x) dF_\Theta(x) \\
 &= \underline{E_\Theta(\hat{\mathcal{J}}(\mathbf{x}) U(\mathbf{x}))}
 \end{aligned}$$

$U$  is a closed space and it is  
linear.

#

Corollary.

If  $\mathbf{x} \sim F_\Theta$ ,  $\theta \in \Omega$ , then if  $E_\Theta(\delta(\mathbf{x})) = f(\theta)$   $\forall \theta \in \Omega$  then

$\mathcal{J}$  is LMVU at  $\theta_0$   
i.f.t.

$$E_{\theta_0}(\delta(\mathbf{x}) U(\mathbf{x})) = 0 \quad \forall U \in \mathcal{U}$$

#

Classification of scalar product, linearity and closed-ness of  $\mathcal{U}$ .

Note that

$$E_{\Phi}(h^2(\xi)) = \frac{1}{2} \|h^2\|_{F_\Phi}^2$$

defines a norm on the space of all functions  $h$  s.t.  $E_{\Phi} h^2(\xi) < \infty$   $\forall \xi \in \mathcal{X}$ , since

$$\langle \mathbb{1}, h \rangle_{F_\Phi} := \int k(x) h(x) dF_\Phi(x)$$

defines a scalar product.

(i) (bilinear) (linear in each argument)

$$\langle a_1 \mathbb{1} + a_2 \mathbb{K}_\xi, h \rangle_{F_\Phi} =$$

$$a_1 \langle \mathbb{1}, h \rangle_{F_\Phi} + a_2 \langle \mathbb{K}_\xi, h \rangle_{F_\Phi}$$

$$(ii) \quad \langle k, h \rangle = \overline{\langle h, k \rangle}$$

(iii) the norm arises as

$$\|h\| = \langle h, h \rangle^{1/2}$$

Note that angle  $\theta$  is defined as between  $h, k$

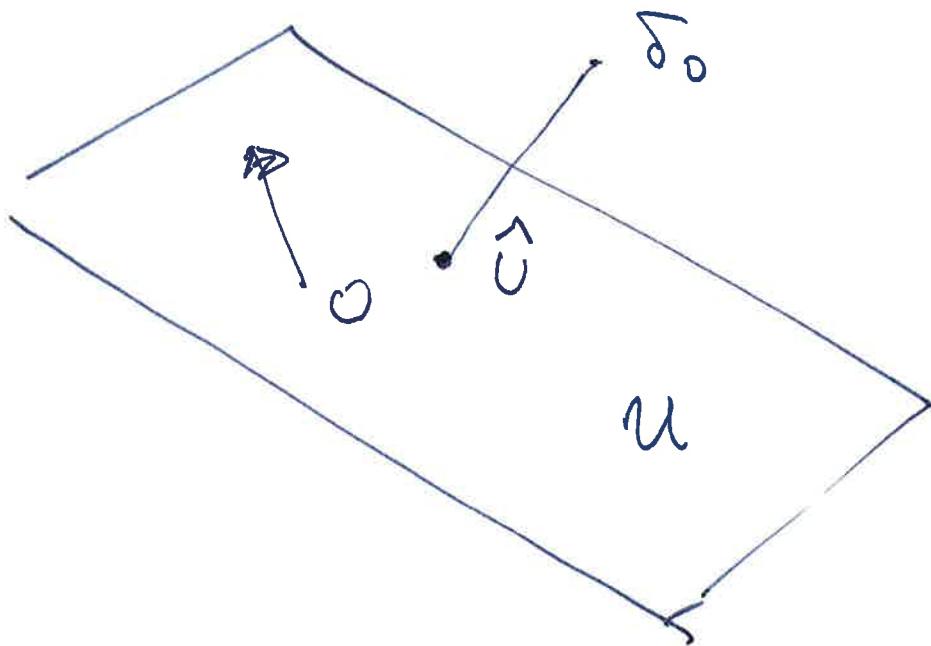
$$\cos \theta = \frac{\langle h, k \rangle}{\langle h, h \rangle^{1/2} \langle k, k \rangle^{1/2}}$$

9a

$$= \frac{\int h(\xi) k(\xi) dF_\Phi(\xi)}{\left( \int h^2(\xi) dF_\Phi(\xi) \right)^{1/2} \left( \int k^2(\xi) dF_\Phi(\xi) \right)^{1/2}}$$

$$\text{if } E_{\Phi}(h(\xi)) = \left( \int h^2(\xi) dF_\Phi(\xi) \right)^{1/2} \text{ and } \left( \int k^2(\xi) dF_\Phi(\xi) \right)^{1/2}$$

$$E_{\Phi}(h(\xi)k(\xi)) = \frac{\text{Cov}(h(\xi), k(\xi))}{D(h(\xi)) \cdot D(k(\xi))} = \text{Corr}(h(\xi), k(\xi))$$



$U$  is a linear space:  $u_1, u_2 \in U, c_1, c_2 \in \mathbb{R}$

$$\begin{aligned} E_\theta(c_1 u_1(\xi) + c_2 u_2(\xi)) &= \\ &= c_1 E_\theta(u_1(\xi)) + c_2 E_\theta(u_2(\xi)) \\ &= c_1 0 + c_2 0 = 0 \\ &\Rightarrow \\ c_1 u_1 + c_2 u_2 &\in U \end{aligned}$$

$U$  is a closed space. Then

$$\hat{u} = \underset{u \in U}{\operatorname{argmin}} E_\theta((\delta_0(\xi) - u(\xi))^2)$$

$$= \underset{\text{i.e.}}{\operatorname{argmin}} \| \delta_0 - u \|_{F_\theta}^2$$

$$\langle \underbrace{\delta_0 - \hat{u}}_{f}, u \rangle_{F_\theta} = 0 \quad \forall u \in U.$$

(9b)

~~9b~~

Example 1.f Read on your own.

Estimators that are consistent are not interesting. If  $f$  not a constant

Case 1- No U-estimable function  
has an UMVU, for the family

Problem 1.g.  $\{F_\theta : \theta \in \Omega\}$

Case 2. Some but not all non-consistent U-estimable functions have UMVU-estimators.

Case 3 Every U-estimable function has an UMVU.

Lemma.

Let  $X \sim F_\theta \in \mathcal{F}$ . Assume that  $T$  is complete sufficient statistic for  $\mathcal{F}$ .

Then every U-estimable function  $g(\theta)$  has a unique unbiased estimator that is a function of  $T$ .

Proof.

$\exists$   $\delta$  unbiased estimator of  $g(\theta)$  (since  $g(\theta)$  is U-estimable). Then define

$$\eta(T) = \underset{\theta}{E}(\delta(X)|T)$$

(i) Then

$$\begin{aligned} E_{\theta}(\eta(T)) &= E_{\theta}(E_{\theta}(\delta(X)|T)) \\ &= E_{\theta}(\delta(X)) = g(\theta) \quad \forall \theta \in \Omega \end{aligned}$$

so existence is clear.

(ii) Suppose  $\delta_1, \delta_2$  two unbiased estimators both are functions of  $T$ .

Then

$$E_{\theta}(\underbrace{\delta_1(T) - \delta_2(T)}_{f(T)}) = 0 \quad \forall \theta \in \Omega$$

so therefore by completeness

$$0 = f(T) = \delta_1(T) - \delta_2(T) \quad (\text{a.e. } f)$$

(11)

then by the Rao-Blackwell theorem  
 $\eta(T)$  is the UMVU, since it has  
smaller risk than any other unbiased  
estimator.

But Rao-Blackwell need only convexity  
for the loss function. Therefore

### Theorem

$X \sim F_\theta \in \mathcal{F}$ . Assume  $T$  complete  
sufficient statistic. Assume  $L(\theta, d)$   
(loss function) is convex in  $d$

For every  $U$ -optimal function  $g(\theta)$   
there is an unbiased estimator that  
uniformly ~~minimizes~~ minimizes the risk.

(In particular for quadratic  $\overset{\text{loss}}{\cancel{\text{loss}}}$  the  
estimator is UMVU).

If function strictly convex the estimator  
is unique.