Lecture 5: October 11

Lecturer: Joseph Romano

Scribes: Kenneth Tay

5.1 UMVU Examples

5.1.1 Nonparametric Family Example 1

Let $X_1, \ldots, X_n$ iid, family of distributions

$$\mathcal{F} = \{\text{all CDFs on the real line } F \text{ with } \mathrm{Var}_F X < \infty\}.$$

Say we want to estimate the mean $g(F) = \mathbb{E}_F X = \int x dF(x)$.

In this setting, we claim that the order statistics $X_{(1)} \leq \ldots \leq X_{(n)}$ are complete.

Proof:[Sketch, similar to Section 1 Problem 6.33] Consider the submodel of distributions with density

$$\propto \exp\left[\sum_{j=1}^n \theta_j x^j - x^{2n}\right].$$

This is an n-parameter exponential family within $\mathcal{F}$. Because it is full rank, we know that $\left(\sum X_i, \sum X_i^2, \ldots, \sum X_i^n\right)$ is complete sufficient.

Since the order statistics are equivalent to $\left(\sum X_i, \sum X_i^2, \ldots, \sum X_i^n\right)$, they are complete sufficient as well.

Generally speaking, if a statistic is complete for a smaller family of distributions, it is also complete for a larger family of distributions containing it (e.g. Section 1 Problem 6.32). Hence, the order statistics are complete sufficient for the original model $\mathcal{F}$. ■

Since the sample mean $\frac{1}{n}\sum X_i = \frac{1}{n}\sum X_{(i)}$, it is UMVU for $\mathbb{E}_F X$.

Similarly, for $n > 1$, the sample variance $\dfrac{\sum (X_i - \overline{X})^2}{n-1}$ is UMVU for $\mathrm{Var}_F X$.

(**Note**: In this setting, unbiasedness is a very restrictive condition!)

5.1.2 Multinomial Trials

Say we have n independent trials, with each trial resulting in outcome $y_1, \ldots y_p$ with probability $\theta_1, \ldots, \theta_p$.

Let X_i be the number of observations equal to y_i. Then for x_i's such that $\sum x_i = n$, we have

$$P\{X_1 = x_1, \ldots, X_n = x_n\} = \frac{n!}{x_1! \ldots x_p!}\theta_1^{x_1} \ldots \theta_p^{x_p}.$$

This is a multi-parameter exponential family based on $p-1$ parameters. Hence, we have $(X_1, \ldots X_{p-1})$ complete sufficient. Since the order statistics are equivalent to $(X_1, \ldots X_{p-1})$, they are complete sufficient in this setting too.

5.1.3 Nonparametric Family Example 2

Let $X_1, \ldots, X_n$ iid, family of distributions

$$\mathcal{F} = \{\text{all symmetric distributions } F \text{ about an unknown fixed value, with finite 2nd moments and having a density}\}.$$

(This is sometimes called a "semiparametric model".) We wish to estimate the center of symmetry, which is given by $\mathbb{E}_F X$.

Claim: There is no UMVU for $\mathbb{E}_F X$!

Proof: The intuition is as follows: If δ is UMVU for a family of distributions $\mathcal{F}$, then it must be UMVU for any subfamily of distributions of $\mathcal{F}$. If we can find 2 different subfamilies of $\mathcal{F}$ with different UMVUs, then there is no UMVU for $\mathcal{F}$.

We know that $\bar{X}$ is UMVU for the normal family of distributions. Hence, if an estimator δ is UMVU for $\mathcal{F}$, then δ must be equal to $\bar{X}$.

However, consider the family of distributions $\mathcal{G} = \{U(\theta - \sigma, \theta + \sigma)\}$, with θ and σ unknown. For this model, $(X_{(1)}, X_{(n)})$ is complete sufficient.

Consider the statistic $\dfrac{X_{(1)} + X_{(n)}}{2}$. It is unbiased for all distributions in the original model $\mathcal{F}$, and is a function of the complete sufficient statistic of $\mathcal{G}$. Hence, it is UMVU for $\mathcal{G}$.

Thus, if δ is UMVU for $\mathcal{F}$, it is also UMVU for $\mathcal{G}$ and must be equal to $\mathcal{G}$. But clearly we cannot have $\bar{X} = \frac{X_{(1)}+X_{(n)}}{2}$! Therefore there is no UMVU for this set-up. ■

Claim: The order statistics are not complete for this set-up.

Proof: To show that the order statistics are not complete, we need to find a function f of the order statistics which is not identically equal to zero, but $\mathbb{E}_F f(X) = 0$ for all F.

Look at $\dfrac{X_{(1)} + X_{(n)}}{2}$. It is an unbiased estimator of $\mathbb{E}_F X$. Hence, $\dfrac{X_{(1)} + X_{(n)}}{2} - \bar{X}$ is a function of the order statistics which expectation equal to zero under any distribution $F \in \mathcal{F}$. ■

5.2 General Points about UMVU Estimators

- UMVU estimators may not exist (e.g. Section 5.1.3).
- In fact, it is possible that no unbiased estimators exist at all!

  Example: $X \sim \text{Binom}(n, p)$. Then for an unbiased estimator we must have

  $$\mathbb{E}_p \delta(X) = \sum_{j=0}^{n} \delta(j) \binom{n}{j} p^j (1-p)^{n-j} = g(p)$$

  for all p. Notice that the LHS is a polynomial of p with degree $\leq n$. Hence, if g is not a polynomial of p with degree $\leq n$ (e.g. $g(p) = \frac{p}{1-p}$), then it does not have an unbiased estimator.
- We can have a UMVU estimator which is "bad", i.e. inadmissible. We make this concrete in the next section.

5.3 Admissibility

Definition 5.1 *An estimator δ is* ***inadmissible*** *if there exists another estimator δ' such that*

$$\begin{aligned} R(\theta, \delta') &\leq R(\theta, \delta) && \text{for all } \theta, \\ R(\theta, \delta') &< R(\theta, \delta) && \text{for some } \theta. \end{aligned}$$

An estimator δ is ***admissible*** *if it is not inadmissible.*

5.3.1 Example: Catastrophic UMVU

Let $X \sim \text{Poisson}(\lambda)$, estimate $e^{-a\lambda}$, where a is a constant not equal to 1.

Since X is complete sufficient in this set-up, for $\delta(X)$ to be UMVU, it just needs to be unbiased, i.e.

$$\begin{aligned} \mathbb{E}_\lambda \delta(X) = \sum \delta(j) \frac{e^{-\lambda}\lambda^j}{j!} &= e^{-a\lambda} && \forall \lambda, \\ \sum \delta(j) \frac{\lambda^j}{j!} &= e^{(1-a)\lambda} \\ &= \sum \frac{(1-a)^j \lambda^j}{j!} && \forall \lambda. \end{aligned}$$

Hence, the estimator $\delta(X) = (1-a)^X$ is UMVU for $e^{-a\lambda}$.

Now, consider the case of $a = 3$. Then the UMVU of $e^{-3\lambda}$ is $\lambda(X) = (-2)^X$, which is clearly ridiculous!

We show formally that δ is inadmissible:

Proof: Consider the estimator δ' given by

$$\delta'(X) = \begin{cases} (-2)^X & \text{if } X \text{ even}, \\ 0 & \text{if } X \text{ odd}. \end{cases}$$

Then

$$\begin{aligned} L(\lambda, \delta') &\leq L(\lambda, \delta) && \text{for even } X, \\ L(\lambda, \delta') &< L(\lambda, \delta) && \text{for odd outcomes } X. \end{aligned}$$

Since the above is true for the loss function L, it will be true for the risk function as well. Hence, δ is inadmissible. ∎

5.4 Two Sample Problems

Let $X_1, \ldots, X_m$ iid, $X_i \sim \mathcal{N}(\xi, \sigma^2)$, $Y_1, \ldots, Y_n$ iid, $Y_i \sim \mathcal{N}(\eta, \tau^2)$, and let the X_i's be independent of the Y_j's.

5.4.1 Case 1: All 4 parameters unknown

We have a 4-parameter exponential family of full rank. Hence, $\left(\bar{X}, \bar{Y}, \sum(X_i - \bar{X})^2, \sum(Y_j - \bar{Y})^2\right)$ is complete sufficient.

The UMVU of $\eta - \xi$ (difference of means) is $\bar{Y} - \bar{X}$.

What is the UMVU of $\frac{\sigma^2}{\tau^2}$ (ratio of variances)? It can be computed as the product of the UMVU of σ^2 (i.e. $\frac{1}{n-1}\sum(X_i - \bar{X})^2$) and the UMVU of $\frac{1}{\tau}^2$ (see Book for details).

5.4.2 Case 2: $\sigma^2 = \tau^2$ but still unknown

In this case, we have a 3-parameter exponential family of full rank, and nothing really changes from Case 1.

5.4.3 Case 3: $\eta = \xi$ but unknown, σ^2, τ^2 unknown

We claim that no UMVU for the mean exists!

Proof: Consider the submodel of distributions where $\sigma^2/\tau^2 = \gamma$, γ known. In this submodel, the UMVU for ξ is

$$\delta_\gamma = \alpha\bar{X} + (1-\alpha)\bar{Y}$$

with

$$\alpha = \frac{\tau^2/n}{\sigma^2/m + \tau^2/n}.$$

Note that α only depends on γ. (Intuitively, we are weighting inversely proportional to variance.)

If there was a UMVU for the bigger model, it must be UMVU for this submodel for all γ. But clearly δ_γ is not the same as we vary γ! Hence, there is no UMVU estimator of the mean in the original model. ∎

5.5 Non-Convex Loss Functions

In our discussion of UMVUs thus far, we have been looking at convex loss functions only. What happens if we consider non-convex loss functions?

Definition 5.2 *An estimator δ is a **locally minimum risk unbiased estimator** of $g(\theta)$ at θ_0 if:*

- *δ is unbiased for all θ, and*
- *For any other unbiased δ',*

$$R(\theta_0, \delta) \leq R(\theta_0, \delta').$$

Theorem 5.3 (Basu) *Assume we have a loss function L which is bounded by M, i.e. $0 \leq L(\theta, d) \leq M$, $L(\theta, g(\theta)) = 0$.*

Fix θ_0. Assume that there exists some unbiased estimator δ of $g(\theta)$. Then there exists a sequence of unbiased estimators whose risk at θ_0 goes to 0.

Proof: Let

$$\delta'_\pi = \begin{cases} g(\theta_0) & \text{with probability } 1-\pi, \\ \frac{1}{\pi}[\delta(x) - g(\theta_0)] + g(\theta_0) & \text{with probability } \pi. \end{cases}$$

Then δ'_π is unbiased, i.e. $\mathbb{E}_\theta \delta'_\pi = g(\theta_0)$ for all θ.

Because of the boundedness of the loss function, we have

$$R(\theta_0, \delta'_\pi) \leq \pi M.$$

Hence, as $\pi \to 0$, $R(\theta_0, \delta'_\pi) \to 0$. ■