STATS300A - Lecture 6

Dominik Rothenhaeusler
Scribed by Michael Howes

10/05/21

Contents

1 Recap

- Sums of UMVUEs are UMVUE.

- Cramer - Rao lower bound.

- Equivariance and invariance.

2 Equivariance

Assume $\mathcal{X} = \mathbb{R}^n$, $d\mu(x) = dx$.

Definition 1. A decision problem is *location invariant* if

- The densities satisfy $p_{\theta+c}(x + c) = p_\theta(x)$.

- The loss satisfies $L(\theta + c, d + c) = L(\theta, d)$.

In a location invariant decision problem we can always write $L(\theta, d) = \rho(\theta - d)$ for some function ρ.

Definition 2. An estimator δ is *location equivariant* if

$$\delta(X_1 + c, \ldots, X_n + c) = \delta(X) + c.$$

Last time we saw

Theorem 1. *If δ is location equivariant for a location invariant decision problem, then the risk of δ does not depend on θ.*

This gives rise to the following definition.

Definition 3. The *minimum risk equivariant estimator (MRE)* is the equivariant estimator with the smallest risk.

Lemma 1. *Let δ_0 be location equivariant, then δ is location equivariant if and only if $U = \delta - \delta_0$ is location invariant, that is $U(X_1 + c, \ldots, X_n + c) = U(X)$.*

Lemma 2. *An estimator U is location invariant if and only if*

$$U(X) = V(X_1 - X_n, \ldots, X_{n-1} - X_n),$$

for some function $V : \mathbb{R}^n \to D$.

Proof. If $U(X) = v(X_1 - X_n, \ldots, X_{n-1} - X_n)$ we immediately have that U is location invariant. Conversely, if U is location invariant we can define $V(Y_1, \ldots, Y_{n-1}) = U(Y_1, \ldots, Y_{n-1}, 0)$. $\qquad\square$

Thus, we can decompose any location equivariant estimator δ as

$$\delta(X) = \delta_0(X) - v(Y_1, \ldots, Y_{n-1}),$$

where $\delta_0(X)$ is a fixed location equivariant estimator and $Y_i = X_i - X_n$. By varying $v : \mathbb{R}^{n-1} \to D$ we can study all equivariant estimators.

Theorem 2. *Suppose we have a location invariant decision problem and δ_0 is any equivariant estimator with finite risk. Let $v^*(y)$ be the minimizer of*

$$\mathbb{E}_0[\rho(\delta_0(X) - v)|Y = y],$$

where $Y = (X_1 - X_n, \ldots, X_{n-1} - X_n)$, then $\delta^* = \delta_0 - v^*(y)$ is the MRE.

Proof. We can see immediately that δ^* is equivariant. Let δ be any other equivariant estimator, then we have

$$\delta(X) = \delta_0(X) - v(Y),$$

for some function v. Since δ and δ^* are both equivariant, their risk functions are constant. Thus, we can simply compare the risk when $\theta = 0$. We then have

$$
\begin{aligned}
R(0, \delta) &= \mathbb{E}_0[\rho(\delta(X))] \\
&= \mathbb{E}_0[\rho(\delta_0(X) - v(Y))] \\
&= \mathbb{E}_0[\mathbb{E}_0[\rho(\delta_0(X) - v(y))|Y = y]] \\
&\geq \mathbb{E}_0[\mathbb{E}_0[\rho(\delta_0(X) - v^*(y))|Y = y]] \\
&= \mathbb{E}_0[\rho(\delta_0(X) - v^*(Y))] \\
&= R(0, \delta^*).
\end{aligned}
$$

Thus δ^* is the MRE. $\qquad\square$

For existence/uniqueness of MRE see scribed notes.

Example 1. Suppose $X_1, \ldots, X_n \overset{\text{iid}}{\sim} p(x; \theta)$ where

$$p(x; \theta) = \exp(-(x - \theta))\mathbb{I}(x > \theta),$$

and suppose $L(\theta, d) = |d - \theta|$. Note that $X_{(1)} = \min\{X_i\}$ is complete and sufficient for this model. Also note that $\delta_0(X) = X_{(1)}$ is location equivariant. Furthermore, note that $Y = (X_1 - X_n, \ldots, X_{n-1} - X_n)$ is ancillary. Thus, by Basu's theorem

$$
\begin{aligned}
v^*(y) &= \arg\min_v \mathbb{E}_0[|\delta_0(X) - v||Y = y] \\
&= \arg\min_v \mathbb{E}_0[|\delta_0(X) - v|].
\end{aligned}
$$

We want $\partial_v \mathbb{E}_0 |\delta_0(X) - v| = 0$. One can check that $\partial_v \mathbb{E}_0 |\delta_0(X) - v| = \mathbb{E}_0[\mathbb{I}(v \geq \delta_0(X)) - \mathbb{I}(v < \delta_0(X))]$. Thus setting the above derivative equal to 0 and evaluting the expectation we see that we must have $\mathbb{P}(v > \delta_0(X)) = \mathbb{P}(v < \delta_0(X))$. That is v is the median of $\delta_0(X) = X_{(1)}$. One can check that this implies $v^* = \frac{\log(2)}{n}$ and thus the MRE is

$$\delta(X) = X_{(1)} - \frac{\log(2)}{n}.$$

Compared to UMRUEs, MREs are more likely to have bias and more likely to depend on the choice of loss function.

3   MREs for squared error loss

Suppose $\rho(t) = t^2$ (i.e. $L(\theta, d) = (d - \theta)^2$).

Theorem 3. *For any location equivariant δ_0, we have*

$$v^*(y) = \arg\min_y \mathbb{E}_0[(\delta_0(X) - v)^2 | Y = y]$$
$$= \mathbb{E}_0[\delta_0(X) | Y = y].$$

Thus, the MRE is $\delta(X) = \delta_0(X) - \mathbb{E}_0[\delta_0(X)|Y]$.

Theorem 4. *Suppose $p(x; \theta) = f(x_1 - \theta, \ldots, x_n - \theta)$ where $\theta \in \Omega = \mathbb{R}$ is unknown and f is known. Then under mean square error, the MRE of θ is*

$$\delta^*(X) = \frac{\int_{-\infty}^{\infty} u f(x_1 - u, \ldots, x_n - u) du}{\int_{-\infty}^{\infty} f(x_1 - u, \ldots, x_n - u) du}.$$

The estimator $\delta^*(X)$ is called Pitman's estimator of location.

Proof. Take $\delta_0(X) = X_n$. We can see that $\delta_0(X)$ is equivariant. We wish to compute $\mathbb{E}_0[X_n|Y = y]$. We have the density of $(X_1, \ldots, X_n)$ and we want the density of $Z = (Y_1, \ldots, Y_{n-1}, X_n)$. Note that

$$\begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_{n-1} \\ X_n \end{bmatrix} = \begin{bmatrix} Y_1 + X_n \\ Y_2 + X_n \\ \vdots \\ Y_{n-1} + X_n \\ X_n \end{bmatrix} = \begin{bmatrix} 1 & 0 & \ldots & 0 & 1 \\ 0 & 1 & \ldots & 0 & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 1 & 1 \\ 0 & 0 & \ldots & 0 & 1 \end{bmatrix} \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_{n-1} \\ X_n \end{bmatrix}.$$

Thus, $\det \frac{\partial X}{\partial Z} = 1$. Hence, the density of $Z = (Y_1, \ldots, Y_{n-1}, X_n)$ is

$$p_Z(z_1, \ldots, z_n) = p_X(z_1 + z_n, \ldots, z_{n-1} + z_n, z_n).$$

And hence

$$p(z_n | z_1, \ldots, z_{n-1}) \alpha f(z_1 + z_n, \ldots, z_{n-1} + z_n, z_n).$$

Thus, we can conclude that

$$\mathbb{E}_0[X_n | Y] = \mathbb{E}_0[Z_n | Z_1, \ldots, Z_{n-1}]$$
$$= \frac{\int_{-\infty}^{\infty} t f(z_1 + t, \ldots, z_{n-1} + t, t) dt}{\int_{-\infty}^{\infty} f(z_1 + t, \ldots, z_{n-1} + t, t) dt}.$$

Performing the change of variable $t = x_n - u$ we get

$$\mathbb{E}_0[X_n|Y] = \frac{\int_{-\infty}^{\infty} (x_n - u) f(x_1 - u, \ldots, x_{n-1} - u, x_n - u) du}{\int_{-\infty}^{\infty} f(x_1 - u, \ldots, x_{n-1} - u, x_n - u) du}$$

$$= x_n - \frac{\int_{-\infty}^{\infty} u f(x_1 - u, \ldots, x_{n-1} - u, x_n - u) du}{\int_{-\infty}^{\infty} f(x_1 - u, \ldots, x_{n-1} - u, x_n - u) du}.$$

Thus, the MRE is

$$\delta^*(X) = X_n - \mathbb{E}_0[X_n|Y] = \frac{\int_{-\infty}^{\infty} u f(x_1 - u, \ldots, x_{n-1} - u, x_n - u) du}{\int_{-\infty}^{\infty} f(x_1 - u, \ldots, x_{n-1} - u, x_n - u) du},$$

as claimed. □

Example 2. Suppose $X_1, \ldots, X_n \sim \text{Unif}[\theta - 1/2, \theta + 1/2]$. Then

$$f(x_1, \ldots, x_n) = \prod_{i=1}^{n} \mathbb{I}(-1/2 \le x_i \le 1/2) = \mathbb{I}(-1/2 \le x_{(1)}; x_{(n)} \le 1/2).$$

Thus, the Pitman estimator of location is

$$\delta^*(X) = \frac{\int_{-\infty}^{\infty} u \mathbb{I}(-1/2 \le x_{(1)} - u; x_{(n)} - u \le 1/2) du}{\int_{-\infty}^{\infty} \mathbb{I}(-1/2 \le x_{(1)} - u; x_{(n)} - u \le 1/2) du}$$

$$= \frac{\int_{-\infty}^{\infty} u \mathbb{I}(u \le x_{(1)} + 1/2; x_{(n)} - 1/2 \le u) du}{\int_{-\infty}^{\infty} \mathbb{I}(u \le x_{(1)} + 1/2; x_{(n)} - 1/2 \le u)) du}$$

$$= \frac{\int_{x_{(n)}-1/2}^{x_{(1)}+1/2} u \, du}{\int_{x_{(n)}-1/2}^{x_{(1)}+1/2} du}$$

$$= \frac{1}{2} \cdot \frac{\left((x_{(1)} - 1/2)^2 - (x_{(n)} + 1/2)\right)}{x_{(1)} - x_{(n)}}$$

$$= \frac{1}{2}(x_{(1)} + x_{(n)}).$$

4 UMVUEs and MREs

Lemma 3. [TPE 3.1.23] *Under square error loss*

(a) *If $\delta(X)$ is location equivariant with bias $b \ne 0$, then $\delta(X) - b$ is an estimator with strictly smaller risk. (we can subtract b because $\delta(X)$ is location equivariant and thus the bias of $\delta(X)$ does not depend on θ).*

(b) *The MRE is unbiased.*

(c) *If the UMVUE exists and is location equivariant then it is the MRE. This is because MRE implies unbiased and thus the UMVUE and MSE have the same risk. Under strictly convex loss (such as square error) the UMVUE is unique and thus UMVUE = MRE.*