TEL502E – Homework 1

Due 01.03.2016

1. Suppose X_1, X_2 are independent variables uniformly distributed over $[0, \theta]$, where $\theta > 0$ is an unknown constant. In order to estimate θ, two estimators are proposed.

$$\theta_1 = X_1 + X_2, \quad \theta_2 = \max(X_1, X_2).$$

   (a) Determine whether θ_1 is unbiased or not. If it is biased, modify it to obtain an unbiased estimator.

   (b) Determine whether θ_2 is unbiased or not. If it is biased, modify it to obtain an unbiased estimator.

   (c) Compare the variances of the estimators you found in parts (a) and (b). Which would you prefer to use?

Solution. (a) Notice that $\mathbb{E}(X_1) = \mathbb{E}(X_2) = \theta/2$. Therefore, $\mathbb{E}(\theta_1) = \theta$ and θ_1 is unbiased.

(b) Let us find the pdf of θ_2 first. We will use the pdf for part (c) also. Note that the cdf of θ_2 is given as,

$$\begin{aligned} F_{\theta_2}(t) &= P(\theta_2 \leq t) \\ &= P((X_1 \leq t) \cap (X_2 \leq t)) \\ &= P((X_1 \leq t))P((X_2 \leq t)) \\ &= \begin{cases} 0, & \text{if } t < 0, \\ t^2/\theta^2, & \text{if } 0 \leq t \leq \theta, \\ 1, & \text{if } \theta < t. \end{cases} \end{aligned}$$

Differentiating, we find f_{θ_2} as,

$$f_{\theta_2}(t) = \begin{cases} 2t/\theta^2, & \text{if } 0 \leq t \leq \theta, \\ 0, & \text{otherwise.} \end{cases}$$

We now compute $\mathbb{E}(\theta_2)$ as,

$$\mathbb{E}(\theta_2) = \int_0^\theta t \, \frac{2t}{\theta^2} \, dt = \frac{2}{3} \, \theta.$$

Therefore θ_2 is biased, but $\tilde{\theta}_2 = \dfrac{3}{2}\theta_2$ is unbiased.

(c) First, note that by independence of X_1, X_2, we have,

$$\mathbb{E}(\theta_1^2) = \mathbb{E}(X_1)^2 + \mathbb{E}(X_1)^2 + 2\mathbb{E}(X_1)\,\mathbb{E}(X_2) = \frac{\theta^2}{3} + \frac{\theta^2}{3} + \frac{\theta^2}{2} = \frac{7}{6}\,\theta^2.$$

Therefore, $\operatorname{var}(\theta_1) = \dfrac{1}{6}\,\theta^2$.

Notice now that,

$$\mathbb{E}(\tilde{\theta}_2^2) = \frac{9}{4}\mathbb{E}(\theta_2^2) = \frac{9}{4}\int_0^\theta t^2 \, \frac{2t}{\theta^2} \, dt = \frac{9}{8}\theta^2.$$

Therefore $\operatorname{var}(\tilde{\theta}_2) = \dfrac{1}{8}\,\theta^2$.

Since

$$\operatorname{var}(\tilde{\theta}_2) < \operatorname{var}(\theta_1), \quad \text{for all } \theta,$$

I would prefer $\tilde{\theta}_2$.

2. Show that if $\operatorname{var}(X) = 0$ for a random variable, then $X = \mathbb{E}(X)$ (i.e., X is a constant).

Solution. This is an application of Chebyshev's inequality. But let us show it for a special case, using the Cauchy-Schwarz inequality (CSI).

Suppose var $X = 0$ and X is a contiunous random variable with a pdf $f_X(t)$ that is non-zero in some interval I. Using the decomposition

$$t\, f_X(t) = \left[t\sqrt{f_X(t)}\right]\left[\sqrt{f_X(t)}\right],$$

we have, by CSI,

$$\Big(\mathbb{E}(X)\Big)^2 = \left[\int t f_X(t)\, dt\right]^2 \le \left[\int t^2\, f_X(t)\, dt\right] \underbrace{\left[\int f_X(t)\, dt\right]}_{=1}.$$

Recall that, in order for this to hold with equality, we must have,

$$t\sqrt{f_X(t)} = c\sqrt{f_X(t)}, \qquad \text{for all } t,$$

where c is a non-zero constant. But since $f_X(t)$ is a non-zero function that integrates to 1, this is not possible (why?). Therefore, $\big[\mathbb{E}(X)\big]^2 < \mathbb{E}(X^2)$. Thus, $\text{var}(X) = \mathbb{E}(X^2) - \big[\mathbb{E}(X)\big]^2 > 0$, which contradicts the assumption.

3. Suppose X is distributed as $\mathcal{N}(0, \sigma^2)$. Notice that X^2 is an unbiased estimator for σ^2. But suppose we are interested in σ and not σ^2. Is $|X|$ an unbiased estimator for σ?
(Hint : You do not need to evaluate $\mathbb{E}(|X|)$ to answer this question.)

   Solution. Suppose $Z = |X|$ is an unbiased estimator of σ, that is, $\mathbb{E}(|X|) = \sigma$. Observe that $\mathbb{E}(Z^2) = \mathbb{E}(X^2) = \sigma^2$. Thus, $\text{var}(Z) = \mathbb{E}(Z^2) - \big[\mathbb{E}(Z)\big]^2 = 0$. But this means, by Q2 that $Z = |X| = 0$, which is clearly not the case. Therefore, $\mathbb{E}(|X|) \neq \sigma$ and $|X|$ is biased as an estimator σ. In fact, we have $\mathbb{E}(|X|) < \sigma$ (why?).

4. Suppose X_1 and X_2 are independent random variables distributed as $\mathcal{N}(\sqrt{\theta}, \theta)$. Two estimators are proposed for θ :

   $$\theta_1 = X_1\, X_2, \quad \theta_2 = X_1^2 + X_2^2.$$

   (a) Determine whether θ_1 is unbiased or not. If it is biased, modify it to obtain an unbiased estimator.
   (b) Determine whether θ_2 is unbiased or not. If it is biased, modify it to obtain an unbiased estimator.
   (c) Compare the variances of the estimators you found in parts (a) and (b). Which would you prefer?
   (Note : You might need the fourth moments of a Gaussian random variable for this part.)

   Solution. (a) Observe that

   $$\mathbb{E}(\theta_1) = \mathbb{E}(X_1\, X_2) = \mathbb{E}(X_1)\, \mathbb{E}(X_2) = \theta.$$

   Thus θ_1 is unbiased.

   (b) First notice that $\text{var}(X_1) = \mathbb{E}(X_1^2) - \big(\mathbb{E}(X_1)\big)^2 = \theta$. Since $\mathbb{E}(X_1) = \sqrt{\theta}$, it follows that $\mathbb{E}(X_1^2) = 2\theta$. Similarly, $\mathbb{E}(X_2^2) = 2\theta$. Now,

   $$\mathbb{E}(\theta_2) = \mathbb{E}(X_1^2 + X_2^2) = \mathbb{E}(X_1^2)\, \mathbb{E}(X_2^2) = 4\theta.$$

   Thus θ_2 is biased, but we can derive an unbiased estimator from θ_2 as, $\tilde{\theta}_2 = \theta_2/4$.

   (c) Notice that

   $$\mathbb{E}(\theta_1^2) = \mathbb{E}(X_1^2)\, \mathbb{E}(X_2^2) = 4\theta^2.$$

   Thus, $\text{var}(\theta_1) = 3\theta^2$.
   For the second estimator, we have (check this!),

   $$\mathbb{E}(\tilde{\theta}_2^2) = \mathbb{E}\left(\frac{1}{16}(X_1^4 + X_2^4 + 2\, X_1^2\, X_2^2\right) = \frac{7}{4}\, \theta^2.$$

   Thus, $\text{var}(\tilde{\theta}_2) = \frac{3}{4}\, \theta^2$. Observe that

   $$\text{var}(\tilde{\theta}_2) < \text{var}(\theta_1), \quad \text{for all } \theta.$$

   Thus, I would prefer $\tilde{\theta}_2$.

5. Suppose X_1 and X_2 are independent random variables distributed as $\mathcal{N}(2\theta, 1)$ and $\mathcal{N}(3\theta, 1)$ respectively, where θ is an unknown parameter.

   (a) Write down the joint pdf of X_1 and X_2.

   (b) Compute the Fisher information for θ, that is,

   $$I(\theta) = \mathbb{E}\left(\left[\partial_\theta\left(\ln f(X_1, X_2; \theta)\right)\right]^2\right),$$

   where $f(X_1, X_2; \theta)$ denotes the joint pdf of X_1 and X_2.

   (c) Find the UMVUE for θ in terms of X_1 and X_2.

Solution. (a) The joint pdf is,

$$f(x_1, x_2; \theta) = \frac{1}{2\pi} \exp\left(-\frac{1}{2}\left((x_1 - 2\theta)^2 - (x_2 - 3\theta)^2\right)\right).$$

(b) Notice that

$$\partial_\theta\left(\ln f(x_1, x_2; \theta)\right) = 2(x_1 - 2\theta) + 3(x_2 - 3\theta) = 13\left(\frac{2x_1 + 3x_2}{13} - \theta\right).$$

Observe now that if $Z = (2X_1 + 3X_2)/13$, then $\mathbb{E}(Z) = \theta$. Thus, $I(\theta) = 13^2 \operatorname{var}(Z)$. But since X_1 and X_2 are independent, $\operatorname{var}(Z) = (2/13)^2 \operatorname{var}(X_1) + (3/13)^2 \operatorname{var}(X_2) = 1/13$. Thus $I(\theta) = 13$.

(c) In part (b), we found that,

$$\partial_\theta\left(\ln f(x_1, x_2; \theta)\right) = I(\theta)(z - \theta),$$

where $z = \dfrac{2x_1 + 3x_2}{13}$. Thus, the unbiased estimator $Z = \dfrac{2X_1 + 3X_2}{13}$ satisfies the CRLB and must be the UMVUE.