

7. Examples, Ups and Downs for UMVUE

GULP Winter 2015

March 3, 2015

1 Common Examples

We start with the sample mean. The sample mean is the UMVUE for the parameter when sample is Bernoulli distributed, Poisson distributed and Normally distributed. The way to show this is to use the C-R bound and check that the sample mean actually reaches the C-R bound.

The sample variance, which is

$$S^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n - 1}$$

is the UMVUE for the normal sample when the true mean is unknown. The way to prove that is using the fact that $(\sum_i X_i, \sum_i X_i^2)$ is sufficient and complete, and use Lehmann-Scheffe theorem as follows:

$$S^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n - 1} = \frac{\sum_{i=1}^n X_i^2 - n(\bar{X})^2}{n - 1}.$$

This is an unbiased estimator for the parameter based on the sufficient and complete statistic. So it is UMVUE.

The order statistics are widely used in the text when the support is related to the parameter. Take uniform distribution $(0, \theta)$ as an example. The unbiased estimator based on the sample mean $2\bar{X}$ has a variance of $\theta^2/3$, while the UMVUE, $(n+1)/n \cdot X_{(n)}$, have a variance of $\theta^2/((n+2)n)$. You may see that the advantage of the UMVUE is huge comparing to the estimator based on the sample mean.

2 Comparing to MLE

The question is that, is UMVUE always good? Not really. It is optimal in the range of the unbiased estimators, but not anymore when we expand the range of the estimators, even when we keep the form of the loss function.

Consider the square loss and the problem of estimating the normal sample variance when the population mean is unknown. The UMVUE gives the total risk as $2\sigma^4/(n-1)$ by using the fact that $\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{\sigma^2}$ follows χ^2 distribution with degree of freedom $n-1$.

The MLE of σ^2 has the form of

$$\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n}$$

which is biased, with a bias of σ^2/n . Using the same fact before we can get the variance of this estimator as $2(n-1)\sigma^4/n^2$. Therefore the total loss equals

$$\left(\frac{\sigma^2}{n}\right)^2 + \frac{2(n-1)\sigma^4}{n^2} = \frac{(2n-1)\sigma^4}{n^2}.$$

This is strictly less than the risk of UMVUE. This tell us that UMVUE is even strictly worse than the MLE.