eman ta zabal zazu

Universidad del País Vasco — Euskal Herriko Unibertsitatea

Activity 2

1 Synopsis.

What this activity is about. In the previous seminar you saw how to assess estimators using the Monte Carlo method. You now have to use this technique again in conjunction with the theory to examine the properties of several estimators.

What you need. You need to be fully acquainted with the content of previous seminars and practice assignment. You will also need access to a computer equiped with R.

2 Background

In this activity we will explore further the $U(0, \theta)$, one of the distributions that was repeatedly used in class as an example (of lack of regularity and consequent inapplicability of the Cramér-Rao lower bound, of an inefficient moment estimator, of a biased maximum likelihood estimator…).

You will be required to check, both analytically and using Monte Carlo simulation, that:

1. Even though the Cramér-Rao theorem cannot be used to establish efficiency of the MLE of θ, $\hat{\theta}_{MLE}$, it is still true that in terms of mean square error (MSE) it dominates moment estimators[1].

2. An unbiased estimator may not be optimal in terms of mean square error.

[1] When the Cramér-Rao inequality is not applicable due to lack of regularity, as in this case, there are alternative results which provide, when they exist, optimal estimators in some sense (for instance minimum variance unbiased estimators). One such result is the Rao-Blackwell theorem. If you are curious, you can check any of the manuals [1], [3], [2], [4] or [5], all of which are of higher level than this course and suitable as continuation texts.

3 Problems

Consider $X \sim U(0,\theta)$. Our goal is to obtain a good estimator of θ from a sample of independent observations $X_1,\ldots,X_n$.

1. Consider the usual moment estimator, obtained from matching the first order population and sample moments, $\hat{\theta} = 2n^{-1}\sum_{i=1}^n X_i = 2\overline{X}$. Obtain the mean and variance of said estimator. Is it unbiased?

2. The previous moment estimator was obtained by equating $\alpha_1 = E[X]$ to $\overline{X}$, which is the usual and recommended practice as it leads to the simplest estimator. But we could equate $\alpha_k(= E[X^k])$ and $n^{-1}\sum_{i=1}^n X_i^k$ to obtain an alternative equation from which to derive an alternative moment estimator.

   (a) Check that the k-th order moment is $\alpha_k = \dfrac{\theta^k}{k+1}$.

   (b) Equate α_k to $n^{-1}\sum_{i=1}^n X_i^k$ to obtain an alternative moment estimator $\hat{\theta}_k$ of θ. Show that
   $$\hat{\theta}_k = \left(\frac{k+1}{n}\sum_{i=1}^n X_i^k\right)^{1/k}.$$
   (You can check that for $k=1$ we have $\hat{\theta}_k = \hat{\theta} = 2\overline{X}$, the ordinary moment estimator obtained before.)

   (c) Is it easy to check unbiasedness for $\hat{\theta}_k$? What stands in your way when you try to establish unbiasedness?

3. Remember that the MLE of θ is $\hat{\theta}_{MLE} = X_{(n)} = \max(X_1, X_2, \ldots, X_n)$.

   (a) Find its mean and variance.

   (b) Is it unbiased?

   (c) What is its mean square error?

4. Generate $N = 1000$ samples of size $n = 50$ from the $U(0,\theta)$ with $\theta = 2$.

   (a) For each sample compute: i) The ordinary moment estimator $\hat{\theta}$; ii) The MLE, $\hat{\theta}_{MLE}$; and iii) The alternative moment estimator $\hat{\theta}_k$ for values of $k=5$ and $k=10$. Save your results.

   (b) Estimate the bias of all estimators. It should be close to the theoretical values that you have obtained for $\hat{\theta}$ and $\hat{\theta}_{MLE}$.

   (c) Estimate the mean square error (MSE) of all estimators. Which estimator comes out best?

5. *(optional)* It would seem that $\hat{\theta}_k$ improves for larger k. What would happen if we let k take a very large value? (You can obtain some insight if you take the limit

$$\lim_{k\to\infty} \hat{\theta}_k = \lim_{k\to\infty} \left(\frac{k+1}{n} \sum_{i=1}^{n} X_i^k \right)^{1/k}$$

and see what you obtain.)

6. Compute the efficiency of the moment estimator $\hat{\theta}$ relative to the MLE estimator corrected of bias.

7. *(optional)* If you get this far, you might want to play with your program. What happens if you correct the bias of the MLE estimator? Does the MSE improve? What happens with the relative performance of the estimators for different sample sizes n? Remember that the MLE estimator is guaranteed (in regular cases, which is not the case here) to be asymptotically efficient, but for small sample sizes we might find others which are vastly superior. Report any results that you find interesting.

4 Hints and comments

1. It is very easy to obtain the k-th order moment of the $U(0,\theta)$ distribution. All that is required is to solve the integral

$$\alpha_k = \int_0^\theta x^k \left(\frac{1}{\theta}\right) dx$$

which is immediate.

2. To obtain the mean and variance of $\hat{\theta}_{MLE} = X_{(n)}$ you need first its distribution. Notice that,

$$F_{X_{(n)}}(x) = P(X_{(n)} \leq x) = \left(\frac{x}{\theta}\right)^n.$$

(The probability that the largest sample value is below x is the probability that all n independent sample values are below x. Taking the derivative of $F_{X_{(n)}}(x)$ will give you the density and allow easy, tough tedious, computation of mean and variance.)

3. When comparing biased estimators, variance alone is of no help; MSE (mean square error, $E(\hat{\theta}-\theta)^2$) makes more sense.

4. A possible work flow for question 4 would be:

   (a) Create a matrix with 1000 rows and 4 columns to hold the values of all four tested estimators for each of the $N = 1000$ samples. Fill with zeros or whatever and name the columns. You can use code as:

```
> results <- matrix(0, 1000, 4)
> colnames(results) <- c("Moment-1", "Moment-5", "Moment-10", "MLE")
```

(b) Open a loop, `for (i in 1:1000) { }` inside which you generate the samples, compute the values of all estimators and save them in row `i` of matrix `results`

(c) When you exit the loop, estimate the bias and/or the mean square for all columns. For instance, to estimate the means of the columns you might use

```
> means <- colMeans(results)
```

This enables you to estimate the bias. To compute the mean square error you might use:

```
> MSE <- colMeans( (results - 2)^2 )
```

(d) To quickly see what is going on, you might want to draw boxplots of the values of each estimator. Do for instance,

```
> boxplot(as.data.frame(results))
```

(Function `boxplot` requires as its argument a list or data frame; we can turn our matrix of results into a data frame on the fly, as shown.)

5. To answer the optional question 5 above, you might want to notice that:

$$\lim_{k\to\infty}\left(\frac{k+1}{n}\right)^{1/k} = 1$$

Also notice that

$$X_1^k + X_2^k + \ldots + X_n^k \approx X_{(n)}^k \qquad \text{as } k \to \infty$$

where $\approx$ means "asymptotically equivalent" and $X_{(n)} = \max_i(X_1, \ldots, X_n)$. In other words, when we have a sum of non-negative numbers all raised to a large power, the powers of all but the largest become negligible in the sum.

References

[1] P. J. Bickel and K. A. Doksum. *Mathematical Statistics*. Holden-Day, Inc., San Francisco, 1977.

[2] P. H. Garthwaite, I. T. Jolliffe, and B. Jones. *Statistical Inference*. Prentice Hall, London, 1995.

[3] J. C. Kiefer. *Introduction to Statistical Inference*. Springer-Verlag, New York, 1987 edition, 1983.

[4] E. L. Lehmann. *Theory of Point Estimation*. Wiley, New York, 1983.

[5] V. Spokoiny and T. Dickhaus. *Basics of Modern Mathematical Statistics*. Springer Verlag, 2014.