Expected Value of an Estimator

The statistical expectation of an estimator is useful in many instances. Expectations are an "average" taken over all possible samples of size n. The process is fairly simple when working with discrete random variables. As an example, we examine a population of 4 rats (rat A, B, C, and D) each with a number of ticks (exact counts of the number of ticks on each of the rats are: A=2 ticks, B=4 ticks, C=2 ticks and D=8 ticks). We are interested in the mean number of ticks per rat, say μ (a parameter). It seems reasonable to use the sample mean $\hat{\mu}$ = # ticks/n as an estimator of μ. We decide to take a sample of size 2 for the example. Using the binomial coefficient we find that there are 6 ways to choose a sample of 2 rats from a population of 4 rats ("4 choose 2" = 6). The sample data are summarized below:

Sample	No. ticks	Sample mean
AB	6	3
AC	4	2
AD	10	5
BC	6	3
BD	12	6
CD	10	5 .

This covers all possible samples of size 2 (n=2) and the corresponding estimates, $\hat{\mu}$. The mean of these values is the expected value of the estimator $\hat{\mu}$:

$$(3+2+5+3+6+5)/6 = 24/6 = 4.$$

Thus, the expected value of the estimator $\hat{\mu}$ is 4; this is denoted as $E(\hat{\mu})$. The population total = 16 ticks (i.e., 2+4+2+8= 16) for the 4 individual rats, then the population mean $\mu = 16/4 = 4$. In this case, the expected value of the estimator $\hat{\mu}$ (denoted $E(\hat{\mu})$) is **unbiased**. In general, bias is written

$$\text{bias} = E(\hat{\theta}) - \theta,$$

where θ is some parameter and $\hat{\theta}$ is its estimator. Recall, θ is often used as a generic symbol for a parameter; θ could be a survival probability, a mean, population size, resighting probability, etc.

It is important to separate two kinds of bias:

> "**small sample bias**". Given a model, this bias goes to 0 as sample size goes to ∞. This is often a trivial concern and assumes, with real data, that one knows the exact model to use. The magnitude of the bias is often approximately 1/n. Thus, if sample size is 90, the bias is about 0.011; a trivial consideration when all the other issues are considered. If one assumes *the* model and has but a small sample

size, then often bias-adjusted estimators can be found. Brownie et al. (1985) has many examples of this, denoted with a tilde, e.g., $\tilde{S}$. Such adjustments are easy only if the estimator is in closed form.

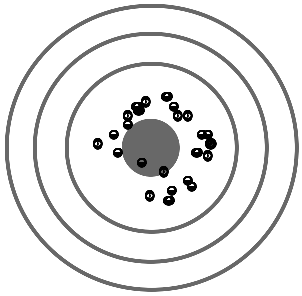
Far more important is the idea of "**model bias**." Here the model for the data is poor and substantial bias can often be expected. This is one reason why the issue of model selection is so important.

A note: the whole notion of small sample bias has been perhaps overblown in the statistical and applied literature. One reads that an estimator is "unbiased" and implies that everything is fine with all aspects of the study. This statement only reveals that *if* the model is the true model, then on average, in repeated sampling, the estimator equals the parameter. This is pretty shallow. It is a worthwhile concept, but one must understand its (minor) relevance.

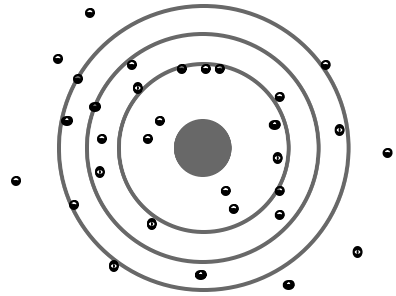
Note also that the bias-adjusted estimator $\tilde{\theta}$ is not the MLE anymore and puts one in a position of using a slightly less likely value as a estimate of the parameter. This is done to achieve a property that has relevance to many other samples that were never taken! A trade-off.

It would be instructive to compute the expectation of the estimator $\hat{\mu}$ for sample size of 3. This is computationally easy, but think hard to fully understand the concept here. Perhaps you should add a rat to the population (then $N = 5$) and let it have 16 ticks. Then, compute μ and $\hat{\mu}$ and assess the bias in the estimator. This would be a good exercise.

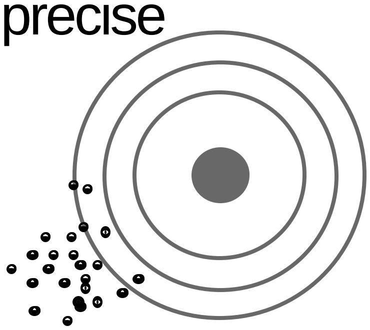
There are many practical cases where it is essentially impossible to compute the exact expectation of an estimator (we will see many such cases in FW663). Often we can *estimate* the expected value by Monte Carlo simulation; denote this as $\hat{\mathrm{E}}(\hat{\theta})$. Here, "computer intensive" methods draw a very large number of samples (say, 1000 or 10,000), compute the value of the estimate, based on a particular estimator, and then average these. If done carefully, this procedure provides a good estimate of the expected value of an estimator. Improved estimates of such expected values can be had by further increasing the number of simulation repetitions.



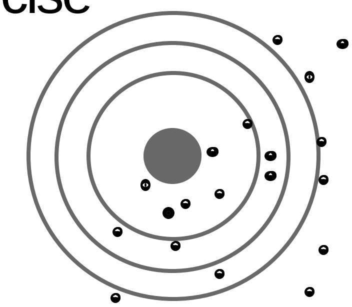
Unbiased and precise



Unbiased but not precise

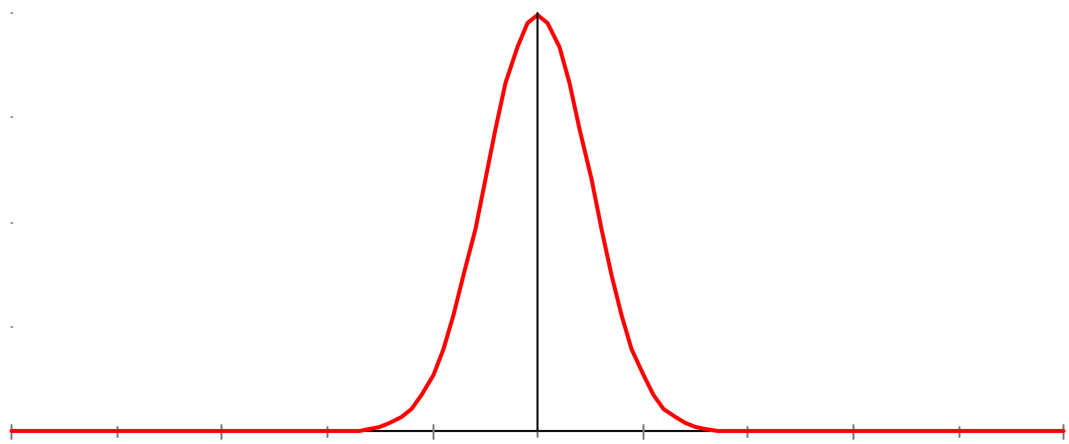


Biased but precise



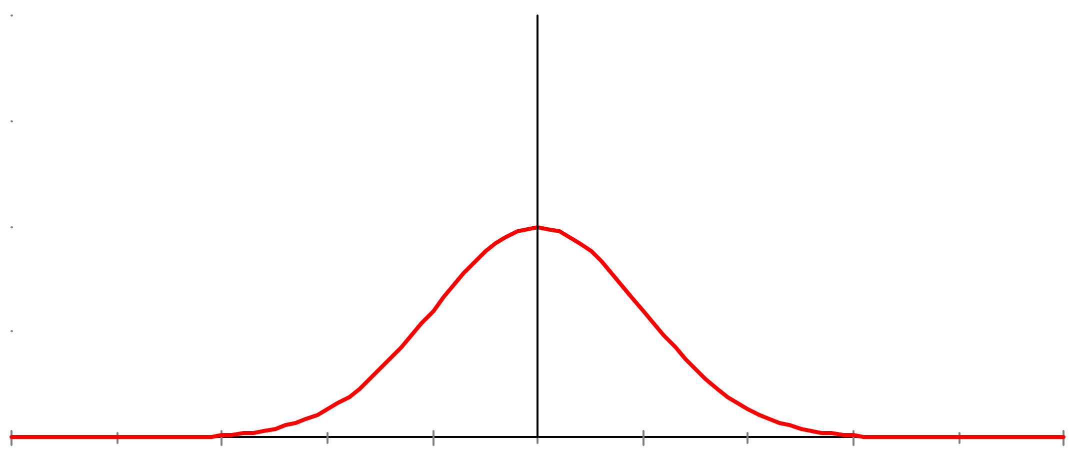
Biased and not precise

No Bias and Precise



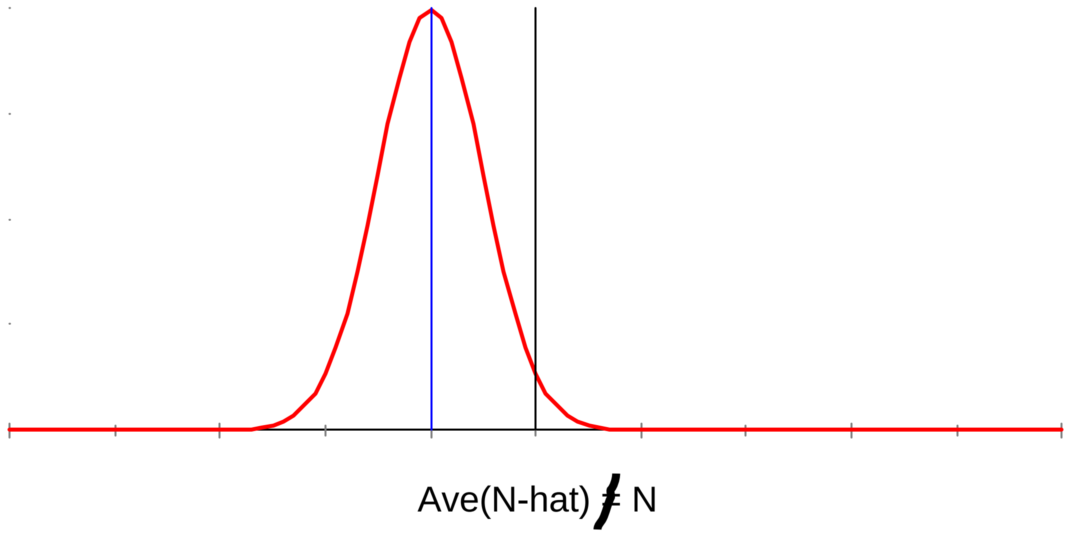
Ave(N-hat) = N

No Bias and not Precise



Ave(N-hat) = N

Biased and Precise



Biased and not Precise