CHAPTER 7

Point Estimation

Education is man's going forward from cocksure ignorance to thoughtful uncertainty.

DON CLARKS' SCRAPBOOK

7-1 POPULATIONS AND SAMPLES

In Table 7-1, we review the concepts of population and sample. It is essential to remember that the population mean μ and variance σ^2 are constants (though generally unknown). These are called population *parameters*.

By contrast, the sample mean $\overline{X}$ and sample variance s^2 are random variables. Each varies from sample to sample, according to its sampling distribution. For example, the distribution of $\overline{X}$ was found to be approximately normal in (6-9). A random variable such as $\overline{X}$ or s^2, which is calculated from the observations in a sample, is given the technical name *sample statistic*. In Table 7-1 and throughout the rest of the text, we shall leave the *population gray* and make the *sample colored* in order to keep the distinction clear, just as we did in Chapter 6.

Now we can address the problem of statistical inference that we posed in Chapter 1: How can the population be estimated by the sample? Suppose, for example, that to estimate the nation's mean household income μ, we take a random sample of 100 incomes. Then the sample mean $\overline{X}$ surely is a reasonable estimator of μ. By the normal approximation rule (6-9), we know that $\overline{X}$ fluctuates about μ; sometimes it will be above μ, sometimes below. Even better than estimating μ with the single *point* estimate $\overline{X}$ would be to construct an *interval* estimate about $\overline{X}$ that is likely to bracket μ—a task we shall leave to Chapter 8.

For now, this chapter will concentrate on point estimates. How good is the sample mean $\overline{X}$ as an estimator of μ? Would the sample median be better? To answer such questions, we now develop criteria for judging a good estimator.

7-2 EFFICIENCY OF UNBIASED ESTIMATORS

A—UNBIASED ESTIMATORS

We already have noted that the sample mean $\overline{X}$ is, on average, exactly on its target μ. We therefore call $\overline{X}$ an *unbiased* estimator of μ.

To generalize, we consider any population parameter θ (Greek theta) and denote its estimator by U. If, on average, U is exactly on target as

TABLE 7-1 Review of Population versus Sample

A Random Sample is *a Random Subset of the Population*	
Relative frequencies f/n are used to compute	Probabilities $p(x)$ are used to compute
$\overline{X}$ and s^2	μ and σ^2
These random variables are examples of statistics or estimators.	These fixed constants are examples of parameters or targets.

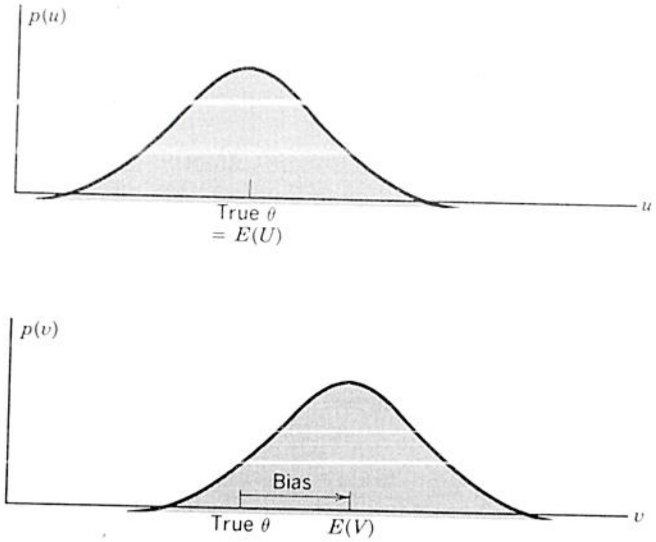


FIGURE 7-1
Comparison of (a) unbiased estimator, and (b) biased estimator.

shown in Figure 7-1*a*, it is called an unbiased estimator. More formally, we define:

$$\boxed{\textbf{U is an unbiased estimator of } \theta \textbf{ if } E(U) = \theta} \qquad (7\text{-}1) \text{ like } (6\text{-}5)$$

Of course, an estimator V is called biased if $E(V)$ is different from θ. In fact, bias is defined as this difference:

$$\boxed{\text{Bias} \equiv E(V) - \theta} \qquad (7\text{-}2)$$

Bias is illustrated in Figure 7-1*b*, where the distribution of V is off-target. Since $E(V)$ is greater than θ, the bias given by (7-2) is positive—reflecting the tendency of V to be too high.

As we stressed already, to avoid bias we have to randomly sample from the whole population. To show the difficulty we can encounter if we fail to follow this fundamental principle, consider an example of nonresponse bias.

EXAMPLE 7-1

Let us give a concrete example of the sample survey mentioned at the beginning of Chapter 6. Suppose each of the 200,000 adults in a city under study has eaten a number X of fast-food meals in the past week. However,

TABLE 7-2 Target Population, and Subpopulation Who Would Respond

X = Number of Meals	Whole Target Population		Subpopulation Responding	
	Freq. f	Rel. Freq. f/N	Freq. f	Rel. Freq. f/N
0	100,000	.50	38,000	.76
1	40,000	.20	6,000	.12
2	40,000	.20	4,000	.08
3	20,000	.10	2,000	.04
	200,000	1.00	50,000	1.00

a residential phone survey on a week-day afternoon misses those who are working—the very people most likely to eat fast foods. As shown in Table 7-2 below, this leaves a small subpopulation who would respond, especially small for higher values of X.

a. What is the mean μ of the whole target population, and the mean μ_R of the subpopulation who would respond?

b. A random sample of 200 phone calls will bring a response of about 50, whose average $\bar{R}$ will be used to estimate μ. What is its bias?

SOLUTION

a. Using the probabilities (rel. freq.) for the whole target population, we obtain $\mu = 0(.50) + 1(.20) + \ldots = .90$.

Similarly, using the probabilities for the responding subpopulation, we obtain $\mu_R = 0(.76) + 1(.12) + \ldots = .40$.

b. The sample mean $\bar{R}$ has an obvious nonresponse bias. Since heavy buyers of fast foods are far less likely to respond, $\bar{R}$ will tend to be much too small. To calculate just how serious the bias is, note that $\bar{R}$ is the average of a random sample drawn from the subpopulation with mean μ_R. Therefore, $E(\bar{R}) = \mu_R$ according to (6-5), and consequently,

$$\text{Bias} = E(\bar{R}) - \mu \quad \text{like (7-2)}$$
$$= \mu_R - \mu$$
$$= .40 - .90 = -.50 \quad (7\text{-}3)$$

Thus the bias is indeed very large—an underestimate of .50 meals per week.

EFFICIENT ESTIMATORS (MINIMUM VARIANCE)

As well as being on target on the average, we also would like the distribution of an estimator to be highly concentrated—that is, to have a small variance. This is the notion of *efficiency*, shown in Figure 7-2. We describe the estimator V in panel (*a*) as more efficient than the estimator W in panel (*b*) because it has smaller variance. More formally, we define the relative efficiency of two unbiased estimators:

$$\text{Efficiency of } V \text{ compared to } W \equiv \frac{\text{var } W}{\text{var } V} \tag{7-4}$$

For example, in the rare case when the population being sampled is exactly symmetric, its center can be estimated without bias by either the sample mean $\overline{X}$ or median $\check{X}$. For some populations, $\overline{X}$ is more efficient; for others, $\check{X}$ is more efficient. In sampling from a normal population for instance, we show in Appendix 7-2 that for large samples:

$$\text{var } \check{X} \simeq 1.57\ \sigma^2/n \tag{7-5}$$

Since $\overline{X}$ has variance σ^2/n, as given in (6-6), this smaller variance makes it

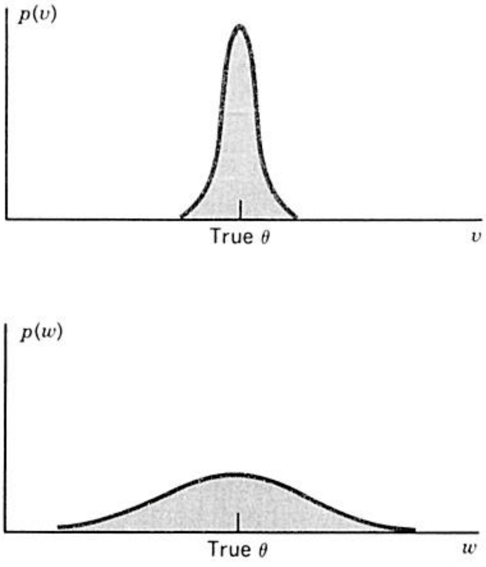


FIGURE 7-2
A comparison of (a) efficient estimator, and (b) inefficient estimator.

more efficient. Specifically, for normal populations,

$$\text{Efficiency of } \overline{X} \text{ relative to } \overset{|}{X} \equiv \frac{\text{var } \overset{|}{X}}{\text{var } \overline{X}} \qquad \text{like (7-4)}$$

$$\simeq \frac{1.57\sigma^2/n}{\sigma^2/n}$$

$$= 1.57 = 157\% \qquad (7\text{-}6)$$

The greater efficiency of the sample mean is nicely confirmed by the Monte Carlo study in Figure 6-11—where the ratio of variances was also 1.57 ($\simeq 1.26^2/1.01^2$). We conclude that, in estimating the center of a *normal* population, the sample mean $\overline{X}$ is about 57% more efficient than the sample median $\overset{|}{X}$. (In fact, it could be proved that the sample mean is more efficient than *every* other estimator of the center of a normal population.)

Of course, by increasing sample size n, we can reduce the variance of either the sample mean or median. This provides an alternative way of looking at the greater efficiency of the sample mean (in sampling from normal populations). The sample median will yield as accurate an estimate only if we take a 57% larger sample. Hence the sample mean is more efficient because it costs less to sample. Note how the economic and statistical definitions of efficiency coincide in this case.

Is the sample mean always more efficient than the median? An example will give the answer.

EXAMPLE 7-2

One of the population models with thicker tails than the normal is called the Laplace. In this case, we show in Appendix 7-2 that for large samples:

$$\text{var } \overset{|}{X} \simeq .50\ \sigma^2/n \qquad (7\text{-}7)$$

Now what is the efficiency of $\overline{X}$ relative to $\overset{|}{X}$?

SOLUTION

Again, $\overline{X}$ has variance σ^2/n, as given in (6-6). In this case it has larger variance than $\overset{|}{X}$, and so $\overline{X}$ is now less efficient. Specifically, for Laplace populations,

$$\text{Efficiency of } \overline{X} \text{ relative to } \overset{|}{X} \equiv \frac{\text{var } \overset{|}{X}}{\text{var } \overline{X}} \qquad \text{like (7-4)}$$

$$\simeq \frac{.50\ \sigma^2/n}{\sigma^2/n} \qquad (7\text{-}8)$$

$$= .50 = 50\% \qquad (7\text{-}9)$$

Although the Laplace distribution is a mathematical rarity, it nicely illustrates a very practical point: If a population has thick tails, so that outlying observations are likely to occur, then the sample mean has larger variance—because it takes into account all the observations, even the distant outliers that the sample median ignores. In Chapter 16 we will pursue this issue further.

PROBLEMS

7-1 Assuming as usual that samples are random, answer True or False; if False, correct it.

a. Samples are used for making inferences about the population from which they are drawn.

b. μ is a random variable (varying from sample to sample), and is an unbiased estimator of the parameter $\overline{X}$.

c. If we double the sample size, we halve the standard error of $\overline{X}$, and consequently double its accuracy in estimating the population mean.

d. The sample proportion P is an unbiased estimator of the population proportion π.

7-2 Based on a random sample of 2 observations, consider two competing estimators of the population mean μ:

$$\overline{X} \equiv \frac{1}{2}X_1 + \frac{1}{2}X_2$$

$$\text{and } U \equiv \frac{1}{3}X_1 + \frac{2}{3}X_2$$

a. Are they unbiased?

b. Which estimator is more efficient? How much more efficient?

7-3 An economist gathers a random sample of 500 observations, and loses the records of the last 180. This leaves only 320 observations from which to calculate the sample mean. What is the efficiency of this, relative to what could have been obtained from the whole sample?

7-4 What is the efficiency of the sample median relative to the sample mean in estimating the center of a normal population? [Hint: Recall from (7-6) that the efficiency of the mean relative to the median was 157%.]

7-5 a. Answer True or False; if False, correct it.

In both Problems 7-3 and 7-4 we have examples of estimates that are only 64% efficient. In Problem 7-3, this inefficiency was obvious, because 36% of the observations were lost in calculating $\overline{X}$. In Problem 7-4, the inefficiency was more subtle, because it was caused merely by using the sample median instead of the sample mean. However, in terms of results—producing an estimate with more variance than necessary—both inefficiencies are equally damaging.

b. In view of part **a**, what advice would you give to a researcher who spends \$100,000 collecting data, and \$100 analyzing it?

7-6 Suppose that a surveyor is trying to determine the area of a rectangular field, in which the measured length X and the measured width Y are independent random variables that fluctuate widely about the true values, according to the following probability distributions:

x	$p(x)$	y	$p(y)$
8	1/4	4	1/2
10	1/4	6	1/2
11	1/2		

The calculated area $A = XY$ of course is a random variable, and is used to estimate the true area. If the true length and width are 10 and 5, respectively,

a. Is X an unbiased estimator of the true length?

b. Is Y an unbiased estimator of the true width?

c. Is A an unbiased estimator of the true area? (Hint: see Problem 5-38)

7-7 **a.** To guide long-term planning, an automobile executive commissioned two independent sample surveys to estimate the proportion π of car owners who intend to buy a smaller car next time. The first survey showed a proportion $P_1 = 60/200 = 30\%$. The second and larger survey showed a proportion $P_2 = 240/1000 = 24\%$. To get an overall estimate, the simple average $P^* = 27\%$ was taken. What is the variance of this estimate? [Hint: You may assume simple random sampling, so that var $P \simeq P(1 - P)/n$]

b. The first poll is clearly less reliable than the second. So it was proposed to just throw the first away, and use the estimate $P_2 = 24\%$. What is the variance of this estimate? What then is its efficiency relative to P^*?

c. The best estimate of all, of course, would count each observation equally (not each *sample* equally). That is, take the overall proportion in favor, $P = (60 + 240)/(200 + 1000) = 25\%$. What is the variance of this estimate? Then what is its efficiency relative to P^*?

d. True or False? If False, correct it:
It is important to know the reliability of your sources. For exam-

ple, if an unreliable source is not discounted appropriately, using it can be worse than simply throwing it away.

7-3 EFFICIENCY OF BIASED AND UNBIASED ESTIMATORS

In comparing unbiased estimators, we chose the one with minimum variance. Now suppose we are comparing both biased and unbiased estimators, as in Figure 7-3. It may no longer be appropriate to select the estimator with least variance: W qualifies on that score, but is unsatisfactory because it is so badly biased. Nor do we necessarily pick the estimator with least bias: U has zero bias, but seems unsatisfactory because of its high variance. Instead, the estimator that seems to be closest to the target overall is V, because it has the best *combination* of small bias and small variance.

How can we make precise the notion of being "closest to the target overall"? We are interested in how an estimator, V let us say, is spread around its *true target* θ:

$$\textbf{Mean squared error (MSE)} \equiv E(V - \theta)^2 \tag{7-10}$$

like (4-34)

This is similar to the variance, except that it is measured around the true target θ rather than around the (possibly biased) mean of the estimator. Then, as Appendix 7-4 proves, MSE does indeed turn out to be a combination of variance and bias:

$$\textbf{MSE} = \textbf{(variance of estimator)} + \textbf{(its bias)}^2 \tag{7-11}$$

We choose the estimator that minimizes this MSE.

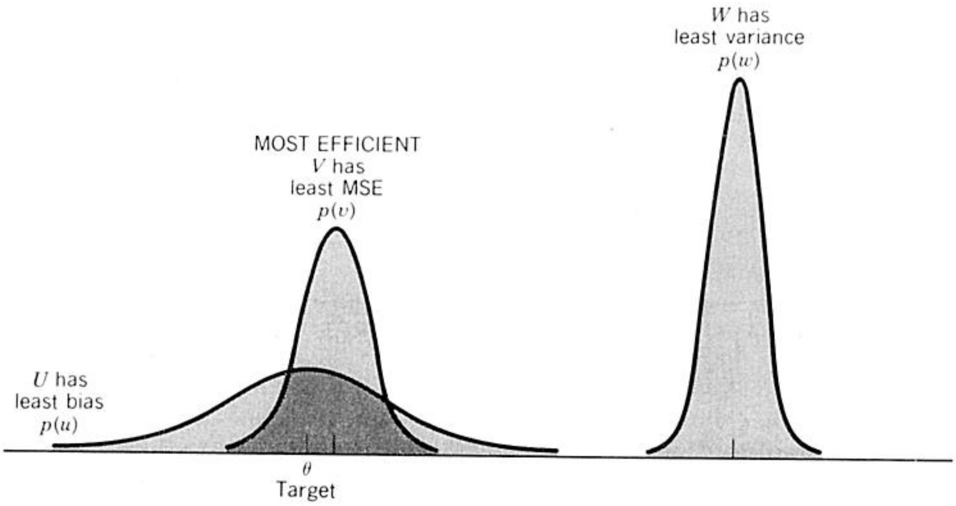


FIGURE 7-3
V is the estimator with the best combination of small bias and variance.

This confirms two earlier conclusions: if two estimators with equal variance are compared (as in Figure 7-1), the one with less bias is preferred; and if two unbiased estimators are compared (as in Figure 7-2) the one with smaller variance is preferred. In fact, if two estimators are unbiased, it is evident from (7-11) that the MSE reduces to the variance. Thus MSE may be regarded as a general kind of variance, applying to either unbiased or biased estimators. This leads to a general definition of the relative efficiency of two estimators:

For any two estimators—whether biased or unbiased—

$$\textbf{Efficiency of } V \textbf{ compared to } W \equiv \frac{\text{MSE}(W)}{\text{MSE}(V)} \tag{7-12}$$

like (7-

To sum up, because it combines the two attractive properties of small bias and small variance, the concept of minimum MSE (or maximum efficiency) becomes the single most important criterion for judging between estimators. An example will illustrate.

EXAMPLE 7-3

In Example 7-1, recall the phone survey of 50 responses from 200 calls, that had a serious nonresponse bias. In addition, the average response $\bar{R}$ has variability too.

a. To measure how much $\bar{R}$ fluctuates around its target μ overall, calculate its MSE.

b. If the sample size was increased fivefold, how much would the MSE be reduced?

c. A second statistician takes a sample survey of only $n = 20$ phone calls, with persistent follow-up until he gets a response. Let this small but unbiased sample have a sample mean denoted by $\bar{X}$. What is its MSE?

d. In trying to publish his results, the second statistician was criticized for using a sample only 1/10 as large as the first. In fact, his sample size $n = 20$ was labeled "ridiculous." What defense might he offer?

SOLUTION

a. Since $\bar{R}$ is the sample mean for $n = 50$ observations drawn only from the subpopulation who would respond, this is the population whose moments are relevant:

Subpopulation Who Would Respond (from Table 7-2)

r	$p(r)$	$rp(r)$	$(r-\mu_R)$	$(r-\mu_R)^2$	$(r-\mu_R)^2p(r)$
0	.76	0	−.4	.16	.1216
1	.12	.12	.6	.36	.0432
2	.08	.16	1.6	2.56	.2048
3	.04	.12	2.6	6.76	.2704
	confirmed: $\mu_R = .40$				$\sigma_R^2 = .64$

Thus we can confirm that the bias in $\bar{R}$ was:

$$\text{Bias} = .40 - .90 = -.50 \qquad \text{(7-3) repeated}$$

Also, from the subpopulation variance $\sigma_R^2 = .64$ in the table, we can deduce the variance of $\bar{R}$:

$$\text{var}(\bar{R}) = \frac{\sigma_R^2}{n} = \frac{.64}{50} = .013 \qquad \text{like (6-6)}$$

Now we can put the bias and variance together to get the MSE:

$$\begin{aligned} \text{MSE} &= \text{var} + \text{bias}^2 \qquad \text{(7-11) repeated} \\ &= .013 + .25 = .263 \end{aligned}$$

b. A sample five times as large would reduce the variance:

$$\text{var}(\bar{R}) = \frac{\sigma_R^2}{n} = \frac{.64}{5 \times 50} = .003$$

Unfortunately it would not reduce the bias—the same nonresponse would merely be repeated more often. Thus,

$$\begin{aligned} \text{MSE} &= \text{var} + \text{bias}^2 \\ &= .003 + .25 = .253 \end{aligned} \qquad \text{(7-13)}$$

Since the predominant term is the bias, which is unaffected by sample size, the MSE was reduced hardly at all (from .263 to .253).

c. By persistent follow up, the statistician gets everybody to reply. He is therefor sampling from the whole target population, whose distribution now becomes the relevant one:

Total Target Population (from Table 7-2)

x	p(x)	xp(x)	$(x - \mu)$	$(x - \mu)^2$	$(x - \mu)^2 p(x)$
0	.50	0	−.90	.81	.405
1	.20	.20	.10	.01	.002
2	.20	.40	1.10	1.21	.242
3	.10	.30	2.10	4.41	.441
	confirmed: $\mu = .90$				$\sigma^2 = 1.09$

Since this sample is drawn from the whole population, $\overline{X}$ is now unbiased, with

$$\text{var}(\overline{X}) = \frac{\sigma^2}{n} = \frac{1.09}{20} = .055$$

Thus we can calculate the MSE:

$$\text{MSE} = \text{var} + \text{bias}^2 \qquad \text{(7-11) repeated}$$

$$= .055 + 0 = .055 \qquad \text{(7-14)}$$

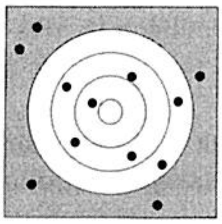
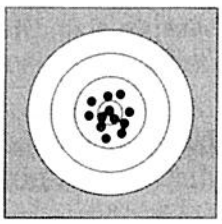
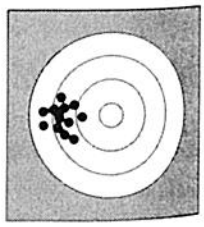
d. His defense would simply be that his estimator is far better because it has a far smaller MSE—four times smaller! Or, in less mathematical terms, "it's the quality of the sample that counts, not mere quantity." He might even point out the lesson of part **b**: Increasing the sample size (even by 5 times) without dealing with its bias would provide little practical improvement.

PROBLEMS

7-8 Each of three guns is being tested by firing 12 shots at a target from a clamped position. Gun *A* was not clamped down hard enough, and wobbled. Gun *B* was clamped down in a position that pointed slightly to the left, due to a misaligned sight. Gun *C* was clamped down correctly.

a. Which of the following patterns of shots belongs to gun *A*? gun *B*? gun *C*?

b. Which guns are biased? Which gun has minimum variance? Which has the largest MSE? Which is most efficient? Which is least efficient?

7-9 A large chain of shops specializing in tuneups has to choose one of four gauges to measure the gap in a spark plug. When tested, each gauge showed a slight error (in hundredths of mm.):

Gauge	*A*	*B*	*C*	*D*
bias	none	−10	5	2
standard dev.	10	none	5	8

Which gauge has the smallest MSE (greatest accuracy)?

7-10 A market survey of young business executives was undertaken to determine what sort of computer would suit a combination of their professional and personal needs. Since those with more children were thought to be more likely to buy a home computer, one of the questions each executive was asked was, "How many children do you have?"

Unfortunately, those with more children tend to have less time and inclination to reply to the survey, as the following table shows:

	Total Population (Target)		*Subpopulation Who Would Respond*	
x = Number of Children Over 5 Years Old	*Frequency f*	*Rel. Frequency f/N*	*Frequency f*	*Rel. Frequency f/N*
0	20,000	.40	6,200	.62
1	12,000	.24	2,100	.21
2	10,000	.20	1,200	.12
3	6,000	.12	400	.04
4	2,000	.04	100	.01
	N = 50,000	1.00	N = 10,000	1.00

Two types of sample survey were proposed:

i. *High volume*, with 1000 executives sampled, and with no follow up. Their overall response rate would be 10,000/50,000 = 20% as given by the table, yielding 200 replies.

ii. *High quality*, with 25 executives sampled, and enough follow-up to get a 100% response rate.

a. Calculate the mean number of children in the population μ.

b. In estimating μ, does either survey have a sample mean $\overline{X}$ that is unbiased?

c. Which survey has the smallest MSE (greatest accuracy)?

***7-11** In Problem 7-10, note how the response rate of executives drops as the number of children X increases. For example, when $X = 0$, the response rate is 6200/20,000 = 31%, while for $X = 4$, the response rate drops to $100/2000 = 5\%$. This is what causes the nonresponse bias, of course.

Now suppose a sample survey with enough follow-up to guarantee 100% response was prohibitively expensive. So a compromise was suggested: Sample 100 executives, with enough follow-up to get a response rate of $12{,}000/20{,}000 = 60\%$ when $X = 0$; and then, for $X = 1, 2, 3, 4$, response rates of 40%, 30%, 30%, and 30%, respectively.

How would the MSE of this compromise survey compare to the other two surveys in Problem 7-10?

*7-4 CONSISTENT ESTIMATORS

A—CONSISTENCY: EVENTUALLY ON TARGET

Like efficiency, consistency is one of the desirable properties of estimators. But consistency is more abstract, because it is defined as a limit: A consistent estimator is one that concentrates in a narrower and narrower band around its target as sample size n increases indefinitely. This is sketched in Figure 7-4, and made more precise in Appendix 7-4.

One of the conditions that makes an estimator consistent is if its MSE approaches zero in the limit. In view of (7-11), this may be reexpressed as follows.

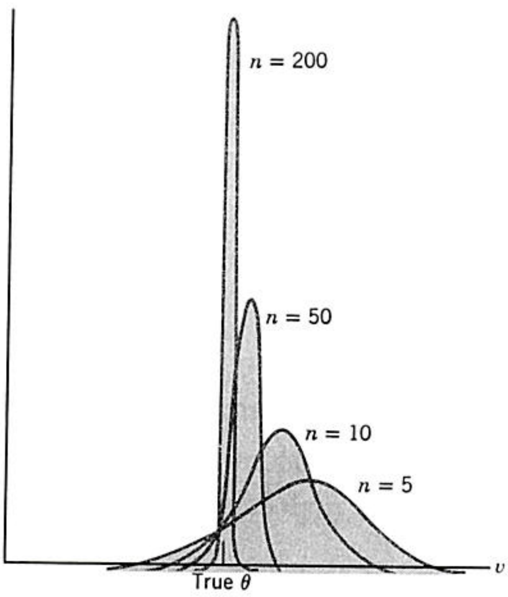


FIGURE 7-4
A consistent estimator, showing how the distribution of V concentrates on its target θ as sample size n increases.

One of the conditions that makes an estimator consistent is: if its bias and variance *both* approach zero.	(7-15)

EXAMPLE 7-4

a. Is $\overline{X}$ a consistent estimator of μ?
b. Is P a consistent estimator of π?
c. Is the average response $\overline{R}$ in Example 7-1 (based on a 25% response rate) a consistent estimator of μ?

SOLUTION

a. From the normal approximation rule (6-9), we know that $\overline{X}$ has:

$$\text{Bias} = 0 \quad \text{for all } n$$

$$\text{var} = \frac{\sigma^2}{n}, \text{ which approaches zero}$$

Thus (7-15) assures us that $\overline{X}$ is a consistent estimator of μ.

b. From the normal approximation rule for P given in (6-12), we similarly see that P is a consistent estimator of π.

c. Recall the nonresponse bias: The estimator $\overline{R}$ concentrated around the value $\mu_R = .40$, which is far below the target $\mu = .90$. So $\overline{R}$ is inconsistent.

B—ASYMPTOTICALLY UNBIASED ESTIMATORS

Sometimes an estimator has a bias that fortunately tends to zero as sample size n increases. Then it is called *asymptotically unbiased*. If its variance also tends to zero, (7-15) then assures us it will be consistent. An example will illustrate.

EXAMPLE 7-5

Consider the mean squared deviation:

$$\text{MSD} = \frac{1}{n}\Sigma(X - \overline{X})^2 \qquad (7\text{-}16)\ \ (2\text{-}10) \text{ repeated}$$

This is a biased estimator of the population variance σ^2. Specifically, on the average it will underestimate, as can be seen very easily in the case of

$n = 1$. Then $\overline{X}$ coincides with the single observed X, so that (7-16) gives MSD = 0, no matter how large the population variance σ^2 may be.

However, if we inflate MSD by dividing by $n - 1$ instead of n, we obtain the sample variance:

$$s^2 \equiv \frac{1}{n-1} \Sigma(X - \overline{X})^2 \qquad \text{(7-17)} \quad \text{(2-11) repeated}$$

It can be proved (Lindgren, 1976) that this slight adjustment inflates s^2 just enough to make it perfectly unbiased. [And in the extreme case above, where $n = 1$, the zero divisor in (7-17) makes s^2 undefined. This provides a simple warning that σ^2 cannot be estimated with a single X, since a single isolated observation gives us no idea whatsoever how spread out the underlying population may be.]

If you were puzzled earlier by the divisor $n - 1$ used in defining s^2, you now can see why. It is to ensure that s^2 will be an *unbiased* estimator of the population variance.

a. Although we have seen that MSD is biased, is it nevertheless asymptotically unbiased?

b. Suppose we used an even larger divisor, $n + 1$, to obtain the following estimator:

$$s_*^2 \equiv \frac{1}{n+1} \Sigma(X - \overline{X})^2 \qquad \text{(7-18)}$$

Is s_*^2 asymptotically unbiased?

SOLUTION

a. Let us write MSD in terms of the unbiased s^2:

$$\text{MSD} = \left(\frac{n-1}{n}\right) s^2 = \left(1 - \frac{1}{n}\right) s^2 \qquad \text{like (2-22)}$$

$$\text{E(MSD)} = \left(1 - \frac{1}{n}\right) \text{E}(s^2)$$

Finally, since s^2 is an unbiased estimator of σ^2,

$$E(\text{MSD}) = \left(1 - \frac{1}{n}\right) \sigma^2 = \sigma^2 - \left(\frac{1}{n}\right) \sigma^2$$

Since $1/n$ tends to zero, the last term—the bias—also tends to zero. So the MSD is indeed asymptotically unbiased.

b. Similarly, we may write:

$$s_*^2 = \left(\frac{n-1}{n+1}\right) s^2 = \left(1 - \frac{2}{n+1}\right) s^2$$

And since $2/(n+1)$ tends to zero, this is also asymptotically unbiased.

REMARKS

We have shown that both MSD and s_*^2 are asymptotically unbiased. And s^2 itself is unbiased for *any* sample size n. It could further be shown that all three estimators have variance that approaches zero, so that they are all consistent.

Which of the three estimators should we use? Since all three are consistent, we need a stronger criterion to make a final choice, such as efficiency. For many populations, including the normal, it turns out that s_*^2 is most efficient.

C—CONCLUSIONS

Although consistency has an abstract definition, it often provides a useful preliminary criterion for sorting out estimators.

Nevertheless, to finally sort out the best estimator, a stronger criterion such as efficiency is required—as we saw in Example 7-5. Another familiar example will illustrate: In estimating the center of a normal population, both the sample mean and median satisfy the consistency criterion. To choose between them, efficiency is the criterion that will finally select the winner (the sample mean).

*PROBLEMS

7-12 The population of American personal incomes is skewed to the right (as we saw in Figure 2-5, for men in 1975, for example). Which of the following will be consistent estimators of the population mean μ?

a. From a random sample of incomes, the sample mean? The sample median? The sample mode?

b. Repeat part **a**, for a sample of incomes drawn at random from the cities over one million.

7-13 When S successes occur in n trials, the sample proportion $P = S/n$ customarily is used as an estimator of the probability of success π. However, sometimes there are good reasons to use the estimator

$P^* \equiv (S + 1)/(n + 2)$. Alternatively, P^* can be written as a line combination of the familiar estimator P:

$$P^* = \frac{nP + 1}{n + 2} = \left(\frac{n}{n + 2}\right) P + \left(\frac{1}{n + 2}\right)$$

a. What is the MSE of P? Is it consistent?

b. What is the MSE of P^*? Is it consistent? (Hint: Calculate th mean and variance of P^*, in terms of the familiar mean an variance of P.)

c. To decide which estimator is better, P or P^*, does consistenc help? What criterion *would* help?

d. Tabulate the efficiency of P^* relative to P, for example when $n = 10$ and $\pi = 0, .1, .2, \ldots, .9, 1.0$.

e. State some possible circumstances when you might prefer t use P^* instead of P to estimate π.

CHAPTER 7 SUMMARY

7-1 Statistics such as $\overline{X}$ from random samples (colored blue) are used tc estimate parameters such as μ from populations (gray).

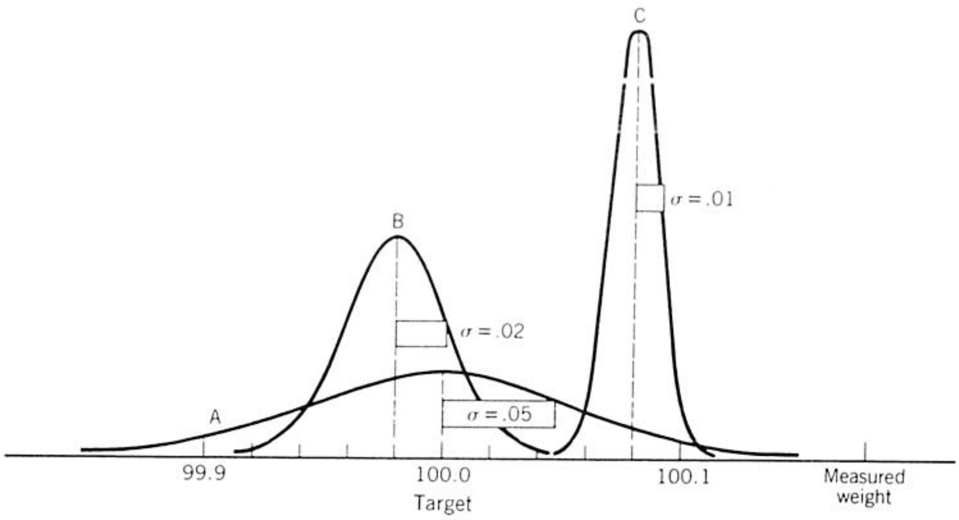
7-2 An estimator is called unbiased if, on average, it is exactly on target. An unbiased estimator is called efficient if it has the smallest variance.

7-3 For estimators with bias as well as variance, minimum MSE (mean squared error) is the appropriate measure of efficiency. MSE remains disappointingly high for estimators with persistent bias, such as nonresponse bias.

***7-4** A consistent estimator is one that eventually is on target. (Not only on target on average, but the whole sampling distribution gets squeezed onto the target, as the sample size n increases infinitely.)

REVIEW PROBLEMS

7-14 An estimator that has small variance (but may be biased) is called *precise*. An estimator that has small MSE is called *accurate*. To illustrate: A standard 100-gm mass was weighed many many times on a scale A, and the distribution of measurements is graphed below. A similar distribution was obtained for scale B, and finally for scale C.



Scale A: $\mu = 100.00$, $\sigma = .05$
Scale B: $\mu = 99.98$, $\sigma = .02$
Scale C: $\mu = 100.08$, $\sigma = .01$

a. Which scale is most precise? Most accurate?

b. What is the relative efficiency of scale A relative to B? Of scale C relative to B? Do these answers agree with part **a**?

c. Which is more important: for an estimator to be precise or accurate?

7-15 **a.** Continuing Problem 7-14, since the scales were not perfect, it was decided in each case to weigh an object 25 times and take the average as the best estimate of the true weight. When used this way, which scale gives the most accurate $\overline{X}$?

b. Answer True or False; if False, correct it:

If a single measurement is taken, the random part (σ) and the systematic part (bias) are equally important.

When several measurements are averaged, the random part of the error gets averaged out, while the systematic part persists. Then it is particularly important to have little bias.

7-16 Suppose that two economists estimate μ (the average expenditure of American families on food), with two unbiased (and statistically independent) estimates U and V. The second economist is less careful than the first—the standard deviation of V is 3 times as large as the standard deviation of U. When asked how to combine U and V to get a publishable overall estimate, three proposals are made:

i. $W_1 = \frac{1}{2}U + \frac{1}{2}V$ (simple average)

ii. $W_2 = \frac{3}{4}U + \frac{1}{4}V$ (weighted average)

iii. $W_3 = 1U + 0V$ (drop the less accurate estimate)

a. Which are unbiased?

b. Intuitively, which would you guess is the best estimator? The worst?

c. Check out your guess in part **b** by making the appropriate calculations.

*d. Intuitively, W_2 works well because it gives only ⅓ as much weight to the component (V) that has 3 times the standard deviation.

Is it possible to do even better than W_2? Suggest some possibilities, and then check them out.

7-17 A processor of sheet metal produces a large number of square plates, whose size must be cut within a specified tolerance. To measure the final product, a slightly worn gauge is used: Its measurement error is normally distributed with a mean $\mu = 0$ and standard deviation $\sigma = .10$ inch. To improve the accuracy, and to protect against blunders, two independent measurements of a plate's length are taken with this gauge, say X_1 and X_2. To find the area of a plate, the quality control manager is in a dilemma:

i. Should he square first, and then average:

$$\frac{X_1^2 + X_2^2}{2}$$

ii. Should he average first, and then square:

$$\left(\frac{X_1 + X_2}{2}\right)^2$$

a. Are methods **i** and **ii** really different, or are they just two different ways of saying the same thing? (Hint: Try a simulation. Suppose, for example, the two measured lengths are $X_1 = 5.9$ and $X_2 = 6.1$.)

b. Which has less bias? [Hint: See equation (4-36).]

c. As an alternative estimator of the area, what is the bias of X_1X_2? (Hint: See Problem 5-38.)

***7-18** A free-trade agreement has opened up a new market of 50 million potential customers for personal computers, and a market survey of these customers is being planned. People with higher incomes are more likely to buy a computer within the next 6 months, and also more likely to respond to a phone survey, as the following table shows:

Income Level	Proportion Who Will Buy	Total Population (Target) Frequency f (millions)	Subpopulation Who Would Respond Frequency f (millions)
\$0–20,000	2%	40	7
20–40,000	4%	5	1
40–80,000	10%	3	1
over 80,000	20%	2	1
		N = 50	N = 10

a. In the 50 million population, how many will buy a computer? Answer as a total figure, and then as a percentage.

b. A market survey of 1000 random phone calls would bring about how many replies? Among these replies, the percentage P who will buy is a natural estimator of the population percentage in **a**. What is the bias and MSE of P?

c. A smaller survey was also considered, with just 100 calls but enough follow-up to get a 100% response. What is the bias and MSE of the resulting estimator P^*?

7-19 To interpret MSE concretely, we could take its square root to get the "typical" error (more precisely, the Root-Mean-Square or RMS error—just like we took the square root of the variance to get the standard deviation).

In Problem 7-18, calculate this RMS error:

a. For P and P^*, the two competing estimators of the percentage of the population who will buy.

b. For the two corresponding estimators of the *total* number in the population who will buy (the "market size").

7-20 ***A Final Challenge: How Much Follow-Up Should a Survey Use?***

A market survey was being planned to estimate the number of drug circulars physicians have read in the past seven days. Physicians who read more were also more likely to respond to the survey, as the following table shows:

X = Number of Circulars Read	Whole Target Population Frequency f	Subpopulation Responding to First Contact Frequency f	Subpopulation Responding to First or Second Contact Frequency f
0	40,000	2,000	14,000
1	5,000	1,000	2,000
2	5,000	2,000	4,000
	N = 50,000	N = 5,000	N = 20,000

To get an accurate estimate of the total number of circulars read, determine which of the following surveys would be better.

i. A large survey of 1000 physicians contacted just once.

ii. A small survey of only 25 physicians, with relentless follow-up to get a 100% response rate.

iii. A compromise survey of 100 physicians contacted a second time if necessary, that would obtain the response rate given in the final column.