

# Point Estimation

## Estimators and Estimates

- Populations can be at least partially described by population parameters.
- Population parameters include mean, proportion, variance, etc.
- Because populations are often very large (maybe infinite, like the output of a process) or otherwise hard to investigate, we often have no way to know the exact values of the parameters.
- *Statistics* or *point estimators* are used to estimate population parameters.

1

- An *estimator* is a function of the sample, i.e., it is a rule that tells you how to calculate an estimate of a parameter from a sample.
- An estimate is a value of an estimator calculated from a sample.
- Different estimators are possible for the same parameter.
- Therefore, we need to find what are “good” estimators.
- To evaluate “goodness”, it’s important to understand facts about the estimator’s sampling distribution, its mean, its variance, etc.

2

## Criteria for Goodness of an Estimator

- We need to evaluate the estimators based on some criteria to determine which is best.
- Complication: An estimator may be better than another estimator with respect to a criterion, but worse than the other estimator with respect to another criterion.
- Some important criteria for goodness of an estimator are based on the following properties:
  - Bias
  - Variance
  - Mean Square Error (MSE)

3

## Repeated Estimation Yields Sampling Dist.

- If you use an estimator once, and it works well, is that enough proof for you that you should always use that estimator for that parameter?
- Visualize calculating an estimate over and over with different samples from the same population.
- This process yields sampling distribution for the estimator.
- We look at the mean of this sampling distribution to see what value our estimates are centered around.
- We look at the spread of this sampling distribution to see how much our estimates vary.

4

## Bias

- We may want to make sure that the estimates are centered around the parameter of interest (the population parameter that we're trying to estimate).
- One measurement of center is the mean, so may want to see how far the mean of the estimates is from the parameter of interest  $\rightarrow$  bias.
- $Bias(\hat{\theta}) = E(\hat{\theta}) - \theta$  if  $\hat{\theta}$  is an estimator for the parameter  $\theta$ . If  $Bias(\hat{\theta}) = 0$ , then  $\hat{\theta}$  is *unbiased*.
- Two common unbiased estimators are:
  1. Sample mean  $\bar{X}$  for population mean  $\mu$ .
  2. Sample proportion  $\hat{p}$  for population proportion  $p$ .

5

## Bias of the Sample Variance

As an estimator of population variance  $\sigma^2$ , what is the bias of the sample variance,  $s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$ ? Contrast this case with that of the estimator  $s^{*2} = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$ , which looks more like the formula for population variance.

6

## Variance of an Estimator

- If we consider two possible estimators for the same population parameter, and both are unbiased, variance might help you choose between them.
- It's desirable to have the most precision possible when estimating a parameter, would prefer the estimator with smaller variance if both are unbiased.
- The standard deviation of an estimator is also called the *standard error* of that estimator.
- Variances of two common estimators:
  1.  $Var(\hat{p}) = \frac{p(1-p)}{n}$
  2.  $Var(\bar{X}) = \frac{\sigma^2}{n}$

7

## Mean Square Error of an Estimator

- If one or more of the estimators are biased, it may be harder to choose between them.
- For example, one estimator may have a very small bias and a small variance, while another is unbiased but has a very large variance. In this case, you may prefer the biased estimator over the unbiased one.
- Mean square error (MSE) is a criterion which tries to take into account concerns about both bias and variance of estimators.
- $MSE(\hat{\theta}) = E[(\hat{\theta} - \theta)^2] \rightarrow$  the expected size of the squared error, which is the difference between the estimate  $\hat{\theta}$  and the actual parameter  $\theta$ .

8

## Another Expression for MSE

Show that the MSE of an estimate can be re-stated in terms of its variance and its bias, so that

$$MSE(\hat{\theta}) = Var(\hat{\theta}) + [Bias(\hat{\theta})]^2$$

## Comparing Two Populations

- Parameters and sample statistics that have been discussed so far only apply to one population. What if we want to compare two populations?
- Example: We want to estimate the difference in the mean income in the year after graduation between economics majors and other social science majors  
 $\rightarrow \mu_1 - \mu_2$
- Example: We want to estimate the difference in the proportion of students who go on to grad school between economics majors and other social science majors  $\rightarrow p_1 - p_2$

- Try to develop a point estimate for these quantities based on estimators we already have.
- For the difference between two means,  $\mu_1 - \mu_2$ , we try the estimator  $\bar{x}_1 - \bar{x}_2$ .
- For the difference between two proportions,  $p_1 - p_2$ , we try the estimator  $\hat{p}_1 - \hat{p}_2$ .

We want to evaluate the “goodness” of these estimators.

- What do we know about the sampling distributions for these estimators?
- Are they unbiased?
- What are their variances?

## Mean and Variance of $\bar{x}_1 - \bar{x}_2$

Show that  $\bar{x}_1 - \bar{x}_2$  is an unbiased estimator for  $\mu_1 - \mu_2$ .

Also show that the variance of this estimator is  $\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$

if those two samples are independent.

### Mean and Variance of $\hat{p}_1 - \hat{p}_2$

Show that  $\hat{p}_1 - \hat{p}_2$  is an unbiased estimator for  $p_1 - p_2$ .

Also show that the variance of this estimator is

$\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}$  if those two samples are independent.

### Summary of Two Sample Estimators

- We have just shown that  $\bar{x}_1 - \bar{x}_2$  and  $\hat{p}_1 - \hat{p}_2$  are unbiased estimators, as were  $\bar{x}$  and  $\hat{p}$ .
- The CLT doesn't directly apply to these estimators since they are not sample means - they are differences of sample means.
- But the CLT separately applies to each sample mean, and since the difference of sample means is a linear combination of the sample means, the difference is also approximately normal when both sample sizes are large (at least 30).

### Error of Estimation

- Even with a good point estimate  $\hat{\theta}$ , there is very likely to be some error ( $\hat{\theta} = \theta$  not likely).
- We can express this error of estimation, denoted  $\varepsilon$ , as  $\varepsilon = |\hat{\theta} - \theta|$ .
- This is the number of units that our estimate,  $\hat{\theta}$ , is off from  $\theta$  (doesn't take into account the direction of the error).
- Since  $\hat{\theta}$  is a random quantity, so is the error of estimation.
- We can use the sampling distribution of  $\hat{\theta}$  to help place some probabilistic bounds on the error of estimation.