This or That Supplement

A list of some general concepts or terms you will start to see and use in this course
as we explore inferential techniques such as
confidence interval estimation and hypothesis testing.

Concept/Term	General notation for this course	Other notation options and description
population mean	μ **(mu)**	The mean computed for a quantitative response or variable based on using the entire population of values or a model that describes the population of values. Since it is for the population, it is a parameter.
Sample mean	$\bar{x}$ **(x-bar)**	$\hat{\mu}$ The mean computed for a quantitative response based on using a sample of values selected from the population. Since it is for the sample, it is a statistic.
Population standard deviation	σ **(sigma)**	The standard deviation computed for a quantitative response or variable based on using the entire population of values or a model that describes the population of values. Since it is for the population, it is a parameter.
Sample standard deviation	s	$\hat{\sigma}$ The standard deviation computed for a quantitative response based on using a sample of values selected from the population. Since it is for the sample, it is a statistic.
Population proportion	p	π The proportion of *successes* computed for a categorical response or variable based on using the entire population of values (each either a *success* or *failure)* or a model that describes the population of such values. Since it is for the population, it is a parameter.
Sample proportion	$\hat{p}$ **(*p*-hat)**	$\hat{\pi}$ The mean computed for a quantitative response based on using a sample of values selected from the population. Since it is for the sample, it is a statistic.
normal distribution	**N(mean, variance)**	**N(<u>mean</u>, <u>standard deviation</u>)** There is a family of normal distributions or bell-curves, each indexed by two parameters, the mean and the standard deviation (or the square of the standard deviation, known as the variance). Sometimes the second term in the parentheses represents the variance, and sometimes you will see a preference to put the standard deviation in the second spot as we often report standard deviations).
Multiplier for forming 95% margin of error	**z* = 1.96**	In a bell curve, about 95% of the observations are expected to be within 2 standard deviations of the mean. So when forming a 95% confidence interval, the margin of error is generally computed as 2 (estimated) standard errors. The exact multiplier for 95% is actually 1.96 when forming confidence intervals for population proportions or for population means, when the sample sizes are reasonably large. In some cases, for smaller sample sizes, the multiplier may need to be adjusted, generally found from a t-distribution and thus referred to as a t* value.

This or That Supplement

Concept/Term	General notation for this course	Other notation options and description
Sampling distribution of the sample mean* * this can be extended more generally to the standard error of any *statistic*	$\bar{x}$ **is (approximately) normal with mean of** μ **and standard deviation** $\sigma/\sqrt{n}$	The sample mean is a statistic. As a statistic, if repeated samples were selected of the same size from the same population, different values of the sample mean would arise. If we were to have all possible values of the sample mean and made a graph of them to view their distribution, we would be viewing the *sampling* distribution of the sample mean. Many sampling distributions for many different statistics end up being approximately bell-shaped, often centered at the corresponding parameter value.
standard error of the sample mean* * this can be extended more generally to the standard error of any *statistic*	$\sigma_{\bar{x}}$ **or** $se(\bar{x})$ $\sigma_{\bar{x}} = \sigma/\sqrt{n}$	The sample mean is a statistic. As a statistic, if repeated samples were selected of the same size from the same population, different values of the sample mean would arise. The standard error of the sample mean represents the true standard deviation of all the possible values for the sample mean, the variability of the distribution for the sample mean.
estimated standard error of the sample mean	$\hat{\sigma}_{\bar{x}}$ **or** $estimated\ se(\bar{x})$ $\hat{\sigma}_{\bar{x}} = s/\sqrt{n}$	As we cannot generally compute the true standard deviation of all the possible values for the sample mean, we are able to use the results from our one sample to come up with an *estimate* of that variability.
standard error of the sample proportion	$\sigma_{\hat{p}}$ **or** $se(\hat{p})$ $\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}}$	The sample proportion is a statistic. As a statistic, if repeated samples were selected of the same size from the same population, different values of the sample proportion would arise. The standard error of the sample proportion represents the true standard deviation of all the possible values for the sample proportion, the variability of the distribution for the sample proportion.
estimated standard error of the sample proportion	$\hat{\sigma}_{\hat{p}}$ **or** $estimated\ se(\hat{p})$ $\hat{\sigma}_{\hat{p}} = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$	As we cannot generally compute the true standard deviation of all the possible values for the sample proportion, we are able to use the results from our one sample to come up with an *estimate* of that variability.
Multiplier for forming 95% margin of error	**z* = 1.96**	In a bell curve, about 95% of the observations are expected to be within 2 standard deviations of the mean. So when forming a 95% confidence interval, the margin of error is generally computed as 2 (estimated) standard errors. The exact multiplier for 95% is actually 1.96 when forming confidence intervals for population proportions or for population means, when the sample sizes are reasonably large. In some cases, for smaller sample sizes, the multiplier may need to be adjusted, generally found from a t-distribution and thus referred to as a t* value.