



CLINICAL EPIDEMIOLOGY AND POPULATION HEALTH

Key Points – Sampling Distribution

Inferential statistics: Making inferences about a population from a sample

Sampling Distribution of the Mean: the distribution of sample means. The sampling distribution has key properties that allow us to make inferences about the population. [Note: you can take the sampling distribution of other test statistics, like the standard deviation, but for the purposes of our course, whenever we say “sampling distribution” we are referring to sampling distribution of the mean.]

If you took repeated samples (of size n) from a population with mean (μ) and standard deviation (SD):

1. The mean of the sampling distribution is the same as the population mean (μ)
2. The standard deviation of the sampling distribution (which we call **Standard Error of the Mean, or Standard Error, or SE**), is equal to: SD/\sqrt{n} . (where n =the number of observations in any given sample)

Key things to know about SE:

- You can figure out key features of the sampling distribution if you know the mean, standard deviation and size of one sample
- Since we often don't know the population SD, we can use the sample SD to calculate SE
- SE is proportional to the sample SD
- SE is inversely proportional to the sample size (n)
- SE is smaller than SD (since you are dividing by SD/\sqrt{n})
- With sufficient sample size, the sampling distribution of the mean is normally distributed, *even if the sample data are not normally distributed*. This is key for inferential statistics (i.e., determining p-values and 95% Confidence Intervals) – thanks to the Central Limit Theorem

Standard Deviation vs. Standard Error: The **standard deviation (SD)** is based on measurements of individuals – it tells us how much variability can be expected among *individuals* in a population or sample. The **standard error of the mean (SE)** is the standard deviation of the means in a sampling distribution – it gives us the precision to which our sample mean estimates the population mean. This precision increases (i.e., SE decreases) when we increase sample size.