10.1: Sampling Distribution of the Difference Between Two Means

Often, we do not have enough information to hypothesize a value for the population mean. Instead of comparing the mean to a hypothesized value, we want to compare means for two different groups.

For example, we might want to compare the mean tumor shrinkage in two sets of cancer patients, one group receiving conventional treatment and the other group receiving a new treatment.

A marketing executive wishing to measure the success of a new marketing campaign could compare the mean sales for a sample of stores not using the marketing strategy to the mean sales for stores that used the marketing strategy.

Hypothesis testing for the difference between two means (critical-value approach):

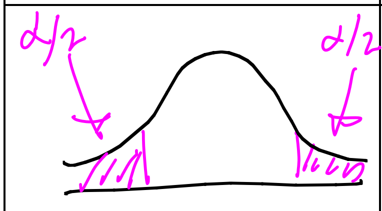
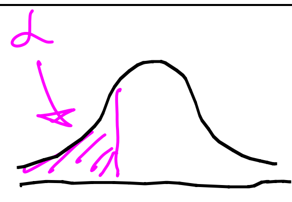
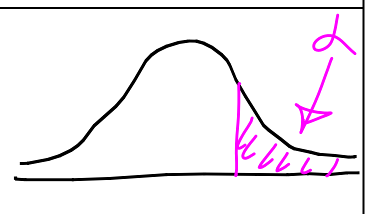
When comparing the means of two samples, we assume each sample was taken from a different population. In other words, Sample 1 comes from a population with mean μ_1; Sample 2 comes from a population with mean μ_2.

To determine whether the two means are the same, we set up null and alternative hypotheses:

$$H_0 : \mu_1 = \mu_2$$
$$H_a : \mu_1 \neq \mu_2$$

Note: This is equivalent to:

$$H_0 : \mu_1 - \mu_2 = 0$$
$$H_a : \mu_1 - \mu_2 \neq 0$$

Two-Tailed Test (most common)	Left-Tailed Test (rare)	Right-Tailed Test (rare)
$H_0 : \mu_1 = \mu_2$ $H_a : \mu_1 \neq \mu_2$	$H_0 : \mu_1 = \mu_2$ $H_a : \mu_1 < \mu_2$	$H_0 : \mu_1 = \mu_2$ $H_a : \mu_1 > \mu_2$
 Rejection Region	 Rejection Region	 Rejection Region

Sampling distribution of the difference between sample means for two independent samples:

Consider samples of size n_1 from a population with mean μ_1, and SD σ_1
and samples of size n_2 from a population with mean μ_2 and σ_2
The difference between the sample means, $\bar{X}_1 - \bar{X}_2$, has the following mean and standard deviation:

$$\mu_{\bar{X}_1-\bar{X}_2} = \mu_1 - \mu_2$$

$$\sigma_{\bar{X}_1-\bar{X}_2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

(standard error)

In addition, if the variable x is normally distributed in both populations, then $\bar{X}_1 - \bar{X}_2$ is also normally distributed.

Recall:
$$\sqrt{\frac{\sigma^2}{n}} = \frac{\sqrt{\sigma^2}}{\sqrt{n}} = \frac{\sigma}{\sqrt{n}}$$

Familiar? this is the standard error for 1 sample

Example 1: Suppose a variable has a mean of 23.1 and a standard deviation of 7.4 in Population A, and has a mean of 26.0 and a standard deviation of 6.8 in Population B. If a sample of size 17 is taken from Population A, and a sample of size 22 is taken from Population B, find the (a) mean and (b) standard deviation of $\bar{X}_1 - \bar{X}_2$. (c) Can we conclude that $\bar{X}_1 - \bar{X}_2$ is normally distributed?

Population A
$\mu_1 = 23.1$
$\sigma_1 = 7.4$
$n_1 = 17$

Population B
$\mu_2 = 26.0$
$\sigma_2 = 6.8$
$n_2 = 22$

(a) mean of $\bar{x}_1 - \bar{x}_2$:

$$\mu_{\bar{x}_1-\bar{x}_2} = \mu_1 - \mu_2 = 23.1 - 26.0 = \boxed{-2.9}$$

(b) SD of $\bar{x}_1 - \bar{x}_2$

$$\sigma_{\bar{x}_1-\bar{x}_2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} = \sqrt{\frac{(7.4)^2}{17} + \frac{(6.8)^2}{22}} = \sqrt{5.323} \approx \boxed{2.307}$$

(c) We can't assume that $\bar{x}_1 - \bar{x}_2$ is normally distributed, because it didn't say that x is normally distributed in Pop A, B.

Example 2: Suppose we are interested in whether the mean of Population 1 is greater than the mean of Population 2. What are the null and alternative hypotheses?

H_0: $\mu_1 = \mu_2$

H_a: $\mu_1 > \mu_2$

Example 3: Suppose we are interested in whether the mean of Population 1 differs from the mean of Population 2. What are the null and alternative hypotheses?

H_0: $\mu_1 = \mu_2$

H_a: $\mu_1 \neq \mu_2$