Sampling distribution model for comparing 2 proportions

Let p_1, p_2 denote the true proportions of something in a population.

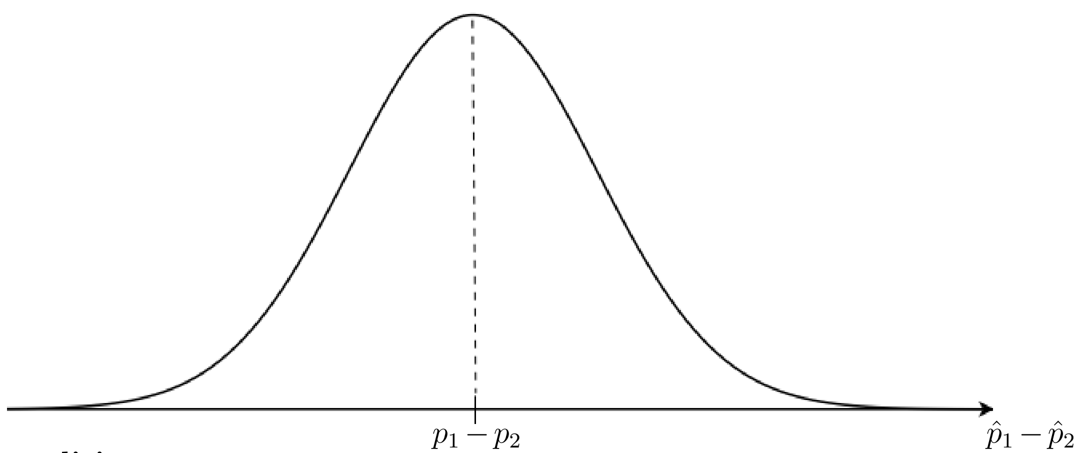
(Note: That means p_1 and p_2 are both population parameters.)

Example: $p_1, p_2 =$ true proportion of student-atheletes among first-years, sophomores.

Suppose the respective sampled values are: $\hat{p}_1$, $\hat{p}_2$.
The respective sample/group sizes are: n_1, n_2.

The sampling distribution model for $\hat{p}_1 - \hat{p}_2$:

$$N\left(p_1 - p_2, \quad \sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}\right)$$



The conditions:

a. The data in each sample (or group) must be independent.

b. The samples/groups must be independent of each other.

c. Each sample/group must be large enough.

Sampling distribution model for comparing 2 means

Let μ_1, μ_2 denote the true mean values of something in a population.

(As before, that would make both μ_1 and μ_2 population parameters.)

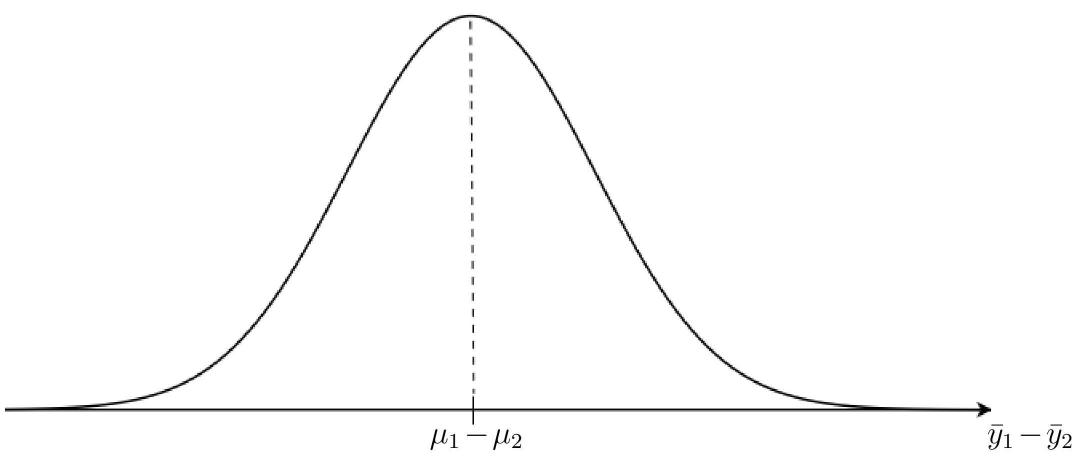
Example: μ_1, μ_2 = mean # of hours of on-campus employment for first-years, sophomores.

Suppose the respective sampled means are: $\bar{y}_1$, $\bar{y}_2$.
The respective sampled standard deviations are: s_1, s_2.
The respective sample/group sizes are: n_1, n_2.

The sampling distribution model for $\bar{y}_1 - \bar{y}_2$:

$$T_{df}\left(\mu_1 - \mu_2, \quad \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}\right)$$

The degrees of freedom (*df*) of the T-distribution are given by a messy formula that students won't be required to know. You will be given the *df* when needed.



The conditions:

a. The data in each sample (or group) must be independent.

b. The samples/groups must be independent of each other.

c. Each sample/group must be approximately normally distributed.