The distribution of the difference between two populations.

Sometimes, we LOOK for information about the difference between two sample means or two sample variances.

For instance, we might be interested in the difference between the mean incomes of two cities, or between the mean grades in a statistics course using in-class teaching and on-line teaching.

The distribution of the difference between two sample means.

Assume that a sample of size n is drawn from the first population and a sample of size m is drawn from the second population. We have two cases:-

Case 1: If $X_1, X_2, \ldots, X_n \overset{r.s.}{\sim} N(\mu_1, \sigma_1^2)$

and $Y_1, Y_2, \ldots, Y_m \overset{r.s.}{\sim} N(\mu_2, \sigma_2^2)$,

then $\overline{X} \sim N(\mu_1, \sigma_1^2/n)$ and $\overline{Y} \sim N(\mu_2, \sigma_2^2/m)$.

In this case, we have two subcases:

<u>Subcase 1a</u>: If σ_1 and σ_2 are known, then

$$\bar{X}-\bar{Y} \sim N\left(\mu_1-\mu_2,\ \sigma_1^2/n + \sigma_2^2/m\right)$$

or

$$Z = \frac{(\bar{X}-\bar{Y})-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{n}+\frac{\sigma_2^2}{m}}} \sim N(0,1).$$

<u>Subcase 1b</u>: If $\sigma_1=\sigma_2=\sigma$ (unknown, but identical)

then

$$T = \frac{(\bar{X}-\bar{Y})-(\mu_1-\mu_2)}{S_p\sqrt{\frac{1}{n}+\frac{1}{m}}} \sim T(n+m-2),$$

$$\sqrt{\frac{S_p^2}{n}+\frac{S_p^2}{m}}$$

where $S_p^2 = \dfrac{(n-1)S_1^2+(m-1)S_2^2}{n+m-2}$ is the pooled (combined) variance (see chapter 1 for more inf.).

<u>Case 2</u>. If the population distributions are not normal and $n, m \geq 30$, then Z defined below has <u>approximately</u> the standard normal distribution.

i.e.,

$$Z = \frac{(\bar{X}-\bar{Y})-(\mu_1-\mu_2)}{\sqrt{\frac{S_1^2}{n}+\frac{S_2^2}{m}}} \sim N(0,1).$$

Ex: Suppose that the grades of female and male students in Calculus 101 are normally distributed with means 70 and 65, and standard deviations 8 and 10, respectively. In samples of 15 female and 20 male students, find the probability that the female students will have an average more than the male students average.

Soln. $X_1, X_2, \dots, X_{15} \overset{r.s.}{\sim} N(70, 8^2)$.

$Y_1, Y_2, \dots, Y_{20} \overset{r.s.}{\sim} N(65, 10^2)$.

Then $\bar{X} \sim N(70, 8^2/15)$ and $\bar{Y} \sim N(65, 10^2/20)$

and hence $\bar{X} - \bar{Y} \sim N\left(5, \frac{8^2}{15} + \frac{10^2}{20}\right)$.

70 − 65

$$P(\bar{X} > \bar{Y}) = P(\bar{X} - \bar{Y} > 0)$$

$$= P\left(Z > \frac{0-5}{\sqrt{\frac{64}{15} + \frac{100}{20}}}\right)$$

$$= P(Z > -1.64)$$

$$= 1 - P(Z \leq -1.64)$$

$$= 1 - 0.0505$$

$$= 0.9495.$$

The distribution of the difference between two sample proportions.

If $\hat{P}_1 \sim N(P_1, P_1(1-P_1)/n)$

and $\hat{P}_2 \sim N(P_2, P_2(1-P_2)/m)$,

then $\hat{P}_1 - \hat{P}_2 \sim N\left(P_1 - P_2, P_1(1-P_1)/n + P_2(1-P_2)/m\right)$

or $$Z = \frac{(\hat{P}_1 - \hat{P}_2) - (P_1 - P_2)}{\sqrt{\frac{P_1(1-P_1)}{n} + \frac{P_2(1-P_2)}{m}}} \sim N(0, 1).$$

Ex. Suppose that 50% of population A own cars, while 35% of population B own cars. If a sample of size 100 is drawn from population A and a sample of size 80 is drawn from population B, what is the probability that the difference between the sample proportions $\hat{P}_A - \hat{P}_B$ will be between 0.1 and 0.2?

Soln.

$P_1 = 0.50$	$P_2 = 0.35$
$1 - P_1 = 0.50$	$1 - P_2 = 0.65$
$n = 100$	$m = 80$

$$\hat{P}_1 - \hat{P}_2 \sim N\left(0.5 - 0.35, \frac{(0.5)(0.5)}{100} + \frac{(0.35)(0.65)}{80}\right)$$

$$\therefore \hat{P}_1 - \hat{P}_2 \sim N(0.15, 0.73^2).$$

$$P(0.1 < \hat{P}_1 - \hat{P}_2 < 0.2) = P\left(\frac{0.1 - 0.15}{0.073} < Z < \frac{0.2 - 0.15}{0.073}\right)$$

$$= P(-0.08 < Z < 0.68)$$

$$= P(Z < 0.68) - P(Z < -0.68)$$

$$= 0.7517 - 0.2483$$

$$= 0.5034.$$

Searching keywords:

- Sampling distribution of the difference between two sample means
- Sampling distribution of the difference between two sample proportions
- Find the probability of
- The University of Jordan الجامعة الأردنية
- Principles of Statistics مبادئ الإحصاء
- Baha Alzalg بهاء الزالق

References: See the textbook in the course website

http://sites.ju.edu.jo/sites/Alzalg/Pages/131.aspx

For any comments or concerns, please use my email to contact me.

د. بهاء محمود الزالق
The University of Jordan
Dr. Baha Alzalg
baha2math@gmail.com

B. Alzalg, 2020, Amman, Jordan