Two-Sample Hypothesis Testing

11/7/22

Case 1: Two Population Means (known σ)

$$Z = \frac{\bar{X} - \bar{Y} - \Delta_0}{\sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}}}$$

$$\sigma_{\bar{X}-\bar{Y}} = \sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}}$$

Square result for variance

Type II Error	Formula
$H_a: \mu_1 - \mu_2 > \Delta_0$	$\Phi\left(z_\alpha + \frac{\Delta_0 - \Delta'}{\sigma}\right)$
$H_a: \mu_1 - \mu_2 < \Delta_0$	$1 - \Phi\left(-z_\alpha + \frac{\Delta_0 - \Delta'}{\sigma}\right)$
$H_a: \mu_1 - \mu_2 \neq \Delta_0$	$\Phi\left(z_{\alpha/2} + \frac{\Delta_0 - \Delta'}{\sigma}\right) - \Phi\left(-z_{\alpha/2} + \frac{\Delta_0 - \Delta'}{\sigma}\right)$

Alt. Hypothesis	p-value
$H_a: \mu_1 - \mu_2 > \Delta_0$	Area to right
$H_a: \mu_1 - \mu_2 < \Delta_0$	Area to left
$H_a: \mu_1 - \mu_2 \neq \Delta_0$	2 (area to right)

Sample Size ($m=n$)	Formula
One-Tail	$\frac{(\sigma_1^2 + \sigma_2^2)(z_\alpha + z_\beta)^2}{(\Delta' - \Delta_0)^2}$
Two-Tail	$\frac{(\sigma_1^2 + \sigma_2^2)(z_{\alpha/2} + z_\beta)^2}{(\Delta' - \Delta_0)^2}$

Case 2: Non-normal, Unknown Variance, Large Sample ($m > 40$, $n > 40$)

$H_0: \mu_1 - \mu_2 = \Delta_0$

$H_a: \mu_1 - \mu_2 <, \neq, > \Delta_0$

$$z = \frac{\bar{X} - \bar{Y} - \Delta_0}{\sqrt{\frac{S_1^2}{m} + \frac{S_2^2}{n}}}$$

* Sub in σ if pop SD is known *

Confidence Interval: $\bar{X} - \bar{Y} \pm z_{\alpha/2} \sqrt{\frac{S_1^2}{m} + \frac{S_2^2}{n}}$

$H_a: \mu > \#$: Lower Confidence Bound $\bar{X} - \bar{Y} + z_\alpha \sqrt{\frac{S_1^2}{m} + \frac{S_2^2}{n}}$

$H_a: \mu < \#$: Upper Confidence Bound $\bar{X} - \bar{Y} - z_\alpha \sqrt{\frac{S_1^2}{m} + \frac{S_2^2}{n}}$

Alt. Hypothesis	p-value
$H_a: \mu_1 - \mu_2 > \Delta_0$	Area to right of z
$H_a: \mu_1 - \mu_2 < \Delta_0$	Area to left of z
$H_a: \mu_1 - \mu_2 \neq \Delta_0$	2 (area to right) of z

Case 3: Two-Sample t-test

$H_0: \mu_1 - \mu_2 = \Delta_0$
$H_a: \mu_1 - \mu_2 <, \neq, > \Delta_0$

Test Statistic

$$t = \frac{(\bar{X} - \bar{Y}) - \Delta_0}{\sqrt{\frac{S_1^2}{m} + \frac{S_2^2}{n}}}$$

Critical value

$t_{\alpha/2, v}$

Two-Sample CI: $\bar{X} - \bar{Y} \pm t_{\alpha/2, v} \sqrt{\frac{S_1^2}{m} + \frac{S_2^2}{n}}$

$H_a: \mu > \#$: Lower Confidence Bound
$H_a: \mu < \#$: Upper Confidence Bound

$$v = \frac{\left(\frac{S_1^2}{m} + \frac{S_2^2}{n}\right)^2}{\frac{(S_1^2/m)^2}{m-1} + \frac{(S_2^2/n)^2}{n-1}}$$

Alt. Hypothesis	p-value
$H_a: \mu_1 - \mu_2 > \Delta_0$	Area to right of $t_{\alpha,v}$
$H_a: \mu_1 - \mu_2 < \Delta_0$	Area to left of $t_{\alpha,v}$
$H_a: \mu_1 - \mu_2 \neq \Delta_0$	2 (area to right) of $t_{\alpha,v}$

Case 4: Pooled t-test (both populations have the same variance)

$$t = \frac{(\overline{X} - \overline{Y}) - (\mu_1 - \mu_2)}{\sqrt{S_p^2 \left(\frac{1}{m} + \frac{1}{n}\right)}}$$

σ^2 estimator $S_p^2 = \frac{m-1}{m+n-2} S_1^2 + \frac{n-1}{m+n-2} S_2^2$

CI: $\overline{X} - \overline{Y} \pm t_{\alpha/2,\, n+m-2} \sqrt{S_p^2 \left(\frac{1}{m} + \frac{1}{n}\right)}$

$H_a: \mu > \#$: Lower Confidence Bound
$H_a: \mu < \#$: Upper Confidence Bound

Alt. Hypothesis	p-value
$H_a: \mu_1 - \mu_2 > \Delta_0$	Area to right of t_{m+n-2}
$H_a: \mu_1 - \mu_2 < \Delta_0$	Area to left of t_{m+n-2}
$H_a: \mu_1 - \mu_2 \neq \Delta_0$	2 (area to right of t_{m+n-2})

Case 5: Paired t-test Procedure

$$t = \frac{\bar{d} - \Delta_0}{S_D / \sqrt{n}}$$

$H_0: \mu_D = \Delta_0$

$H_a: \mu_D <, \neq, > \Delta_0$

$$\bar{d} = \bar{x} - \bar{y}$$

$$S_D = \frac{1}{n-1} \sum_{i=1}^{n} (d_i - \bar{d})^2$$

CI: $\bar{d} \pm t_{\alpha/2, n-1} \, S_D / \sqrt{n}$

$H_a: \mu > \#$: Lower Confidence Bound

$H_a: \mu < \#$: Upper Confidence Bound

Alt. Hypothesis	p-value
$H_a: \mu_1 - \mu_2 > \Delta_0$	Area to right of t_{n-1}
$H_a: \mu_1 - \mu_2 < \Delta_0$	Area to left of t_{n-1}
$H_a: \mu_1 - \mu_2 \neq \Delta_0$	2(area to right of t_{n-1}

Paired vs Unpaired Experiments

1. If there is great heterogeneity between experimental units and a large correlation within experimental units (large positive ρ), then the loss in degrees of freedom will be compensated for by the increased precision associated with pairing, so a paired experiment is preferable to an independent-samples experiment.
2. If the experimental units are relatively homogeneous and the correlation within pairs is not large, the gain in precision due to pairing will be outweighed by the decrease in degrees of freedom, so an independent-samples experiment should be used.

Case 6: Difference in Proportions (Large Sample)

$$Z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}\hat{q}\left(\frac{1}{m} + \frac{1}{n}\right)}}$$

$$\hat{p} = \frac{m}{m+n}\hat{p}_1 + \frac{n}{m+n}\hat{p}_2$$

$$CI: \hat{p}_1 - \hat{p}_2 \pm z_{\alpha/2}\sqrt{\frac{\hat{p}_1\hat{q}_1}{m} + \frac{\hat{p}_2\hat{q}_2}{n}}$$

$$\text{Sample Size: } n = \frac{\left(z_{\alpha}\sqrt{(p_1+p_2)(q_1+q_2)/2} + z_{\beta}\sqrt{p_1q_1 + p_2q_2}\right)^2}{d^2}$$

Type II Error	Formula
$H_a: p_1 - p_2 > 0$	$\Phi\left(\frac{z_{\alpha}\sqrt{\bar{p}\bar{q}\left(\frac{1}{m}+\frac{1}{n}\right)} - (p_1 - p_2)}{\sigma}\right)$
$H_a: p_1 - p_2 < 0$	$1 - \Phi\left(\frac{-z_{\alpha}\sqrt{\bar{p}\bar{q}\left(\frac{1}{m}+\frac{1}{n}\right)} - (p_1 - p_2)}{\sigma}\right)$
$H_a: p_1 - p_2 \neq 0$	$\Phi\left(\frac{z_{\alpha/2}\sqrt{\bar{p}\bar{q}\left(\frac{1}{m}+\frac{1}{n}\right)} - (p_1 - p_2)}{\sigma}\right) - \Phi\left(\frac{-z_{\alpha/2}\sqrt{\bar{p}\bar{q}\left(\frac{1}{m}+\frac{1}{n}\right)} - (p_1 - p_2)}{\sigma}\right)$

$$\bar{p} = \frac{mp_1 + np_2}{m+n} \qquad \sigma = \sqrt{\frac{p_1q_1}{m} + \frac{p_2q_2}{n}}$$

$$\bar{q} = \frac{mq_1 + nq_2}{m+n}$$

Alt. Hypothesis	p-value
$H_a: \mu_1 - \mu_2 > \Delta_0$	Area to right of z
$H_a: \mu_1 - \mu_2 < \Delta_0$	Area to left of z
$H_a: \mu_1 - \mu_2 \neq \Delta_0$	2(area to right) of z

Key words from Exam 2 Review

5. The desired percentage of SiO_2 in a certain type of aluminous cement is 5.5. To test whether the true average percentage is 5.5 for a particular production facility, 16 independently obtained samples are analyzed. Suppose that the percentage of SiO_2 in a sample is normally distributed with $\sigma = 0.3$ and that $\bar{x} = 5.25$.

   a. Does this indicate conclusively that the true average percentage differs from 5.5?

   b. If the true average percentage is $\mu = 5.6$ and a level $\alpha = 0.01$ test based on n = 16 is used, what is the probability of detecting this departure from H_0?

   c. What value of n is required to satisfy $\alpha = 0.01$ and $\beta(5.6) = 0.01$?

5b: Probability of detecting departure = 1 - beta (1 - P(type II error))

1. Tensile-strength tests were carried out on two different grades of wire rod, resulting in the accompanying data.

Grade	Sample Size	Sample Mean (kg/mm^2)	Sample SD
AISI 1064	$m = 129$	$\bar{x} = 107.6$	$s_1 = 1.3$
AISI 1078	$n = 129$	$\bar{y} = 123.6$	$s_2 = 2.0$

Assume that we have large samples in both populations.

   a. Does the data provide compelling evidence for concluding that true average strength for the 1078 grade exceeds that for the 1064 grade by more than 10 kg/mm2? Test the appropriate hypotheses using a significance level of 0.01.

   b. Estimate a confidence interval on the difference between true average strengths for the two grades in a way that provides information about precision and reliability.

1b: Ha: mu > 10
Since the ">" sign was used in this test, the confidence interval required is a one-sided LOWER confidence bound

6. Many freeways have service (or logo) signs that give information on attractions, camping, lodging, food, and gas services prior to off-ramps. These signs typically do not provide information on distances. In a study was reported that in one investigation, six sites along Virginia interstate highways where service signs are posted were selected. For each site, crash data was obtained for a three-year period before distance information was added to the service signs and for a one-year period afterward. The number of crashes per year before and after the sign changes were as follows:

   Before: 15 26 66 115 62 64

   After: 16 24 42 80 78 73

   a. The study included the statement "A paired t test was performed to determine whether there was any change in the mean number of crashes before and after the addition of distance information on the signs." Carry out such a test. [*Note: Assume that data follows a normal distribution.*]

   b. If a seventh site were to be randomly selected among locations bearing service signs, between what values would you predict the difference in number of crashes to lie?

6b: "Between what values would you PREDICT = prediction interval!!!