Inferences Involving Two Populations (Independent Samples)

Inferences about $\mu_1 - \mu_2$ will be based on the difference between the observed sample means $\overline{x_1} - \overline{x_2}$ instead of the mean of paired differences. Since the samples are independent we may not be able to formed paired differences because the sample sizes may be different.

We can form the sampling distribution for $\overline{x_1} - \overline{x_2}$. If the two populations are normally distributed then so is this sampling distribution

In this case $\mu_{\overline{x_1}-\overline{x_2}} = \mu_1 - \mu_2$

The standard error will be estimated using

$$s_{\overline{x_1}-\overline{x_2}} = \sqrt{\frac{{s_1}^2}{n_1} + \frac{{s_2}^2}{n_2}}$$

where s_1 and s_2 are sample standard deviations from the the two populations using sample sizes n_1 and n_2 respectively.

To form a confidence interval estimate we use

$$E = t(\alpha/2) \cdot s_{\overline{x_1}-\overline{x_2}} = t(\alpha/2) \cdot \sqrt{\frac{{s_1}^2}{n_1} + \frac{{s_2}^2}{n_2}}$$

For hypothesis tests we use the test statistic

$$t = \frac{(\overline{x_1} - \overline{x_2}) - (\mu_1 - \mu_2)}{\sqrt{\frac{{s_1}^2}{n_1} + \frac{{s_2}^2}{n_2}}}$$

But we have two sample sizes so what do we use for degrees of freedom?

It is clear that

$$\leq df \leq$$

↑

most conservative approach

An often used approximation for the degrees of freedom is the **Welch–Satterthwaite equation**.

$$df = \frac{\left(\frac{{s_1}^2}{n_1} + \frac{{s_2}^2}{n_2}\right)^2}{\frac{{s_1}^4}{{n_1}^2(n_1-1)} + \frac{{s_2}^4}{{n_2}^2(n_2-1)}}$$

When using this equation we always round down to the nearest integer. (Why?)

Example: Heights were measured from 20 randomly selected women and 30 randomly selected men, taken in order to estimate the difference in their mean heights. Assume that heights are normally distributed. Construct a 95% confidence interval estimate for the difference between the mean heights, $\mu_m - \mu_f$.

Sample	Number	Mean	Standard Deviation
Female	20	63.8	2.18
Male	30	69.8	1,92

Example: Two models of photocopiers were tested to compare repair times after failures. Evaluate whether model B takes significantly more time to repair at 0.01 level of significance.

	Model A	Model B
n	40 failures	45 failures
$\overline{x}$	84.2min	91.6min
s	19.4	18.8

Example: The breaking distances for two brands of winter tires were compared. Test at 5% significance whether one tire performs better than the other.

	Brand 1	Brand 2
n	12	10
Σx	207.6m	141.8m
Σx^2	3654.28m^2	2120.84m^2