INFERENCE FOR COMPARING TWO MEANS

There are two inference procedures for comparing means of two independent samples.

I. **Pooled Two-Sample t Procedures**

These are often emphasized in introductory mathematical statistics textbooks and in older statistics books. They make the assumptions that:

- A simple random sample of size n_1 is taken from a normal population with unknown mean μ_1, and an *independent* simple random sample of size n_2 is taken from another normal population with unknown mean μ_2.
- The two populations have the *same* (*unknown) standard deviation* σ.

The test statistic used is

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}},$$

where s_p is the *pooled standard deviation*

$$s_p = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}$$

Under the assumptions of independence and equal population standard deviations, this statistic has a t -distribution with $n_1 + n_2 -2$ degrees of freedom.

The pooled t procedures were common before statistical computer software was available because they are relatively easy to use. However, there are problems involved in using them:

- The assumption of equal standard deviations is hard to verify in most cases. (In particular, statistical tests for equal variances are not robust to departures from normality, even with large sample sizes.)
- When sample sizes are quite different, the pooled t procedures are not robust to unequal standard deviations.
- Unequal standard deviations are common in real data. (e.g., data sets with large means tend to have large standard deviations).

Thus unless there is good reason to believe that the standard deviations are equal or unless the sample sizes are very close, it is wise to abandon the pooled procedures in favor of the procedures described below.

II. **Unpooled Two-Sample t Procedures**

These make the assumptions that:

- A simple random sample of size n_1 is taken from a normal population with unknown mean μ_1, and an *independent* simple random sample of size n_2 is taken from another normal population with unknown mean μ_2.

No assumption about the standard deviations of the two populations is made.

The test statistic used is

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}},$$

where s_1 is the standard deviation of the first sample and s_2 is the standard deviation of the second sample.

This statistic does *not* have a t distribution. However, it has a approximate t distribution with degrees of freedom

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

This approximation is used by most statistical software for two-sample procedures. It is reasonably accurate when both sample sizes are at least 5. (Note: the df is usually not a whole number. However, t distributions are still defined in terms of their probability density functions for non-integer degrees of freedom.)