Two-sample t-test
(not pooled)

Setup

Consider independent random variables

$$Y_{ij}, i = 1,2; j = 1,2, \dots, n_i,$$

such that

$$Y_{ij} \sim N\left(\mu_i, \sigma_i^2\right).$$

Sample 1: $y_{11}, y_{12}, \dots, y_{1n_1}$ from $N(\mu_1, \sigma_1^2)$

Sample 2: $y_{21}, y_{22}, \dots, y_{2n_2}$ from $N(\mu_2, \sigma_2^2)$

where $\sigma_1^2 \neq \sigma_2^2$

Estimation of $\mu_1 - \mu_2$

Let

$$\bar{Y}_i = \frac{1}{n_i}\sum_{j=1}^{n} Y_{ij}\,, \text{for } i = 1, 2,$$

then

$\bar{Y}_1 - \bar{Y}_2$ is still the BLUE for $\mu_1 - \mu_2$

$$\boxed{\bar{Y}_1 - \bar{Y}_2 \sim N\left(\mu_1 - \mu_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right).}$$

Standardizing $\bar{Y}_1 - \bar{Y}_2$:

$$Z = \frac{\bar{Y}_1 - \bar{Y}_2 - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0, 1)$$

This pivotal quantity can be used to construct a hypothesis test and confidence interval for $\mu_1 - \mu_2$.

Estimation of σ_i^2

We can use

$$S_i^2 = \frac{1}{n_i - 1}\sum_{j=1}^{n}(Y_{ij} - \bar{Y}_i)^2.$$

which is an unbiased estimator of σ_i^2, but what about

$$\boxed{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}?$$

We can use $\boxed{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}$.

What is the distribution of this?

Test statistic

We could use the test statistic

$$T = \frac{\bar{Y}_1 - \bar{Y}_2 - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}.$$

but T does not have a t_k–distribution for any value of k.

Strictly speaking, T is not a pivotal quantity.

This is because the distribution of T depends on $\frac{\sigma_1^2}{\sigma_2^2}$.

Instead, we choose a t_k distribution that approximates the true distribution of T.

Behrens - Fisher problem : finding a test statistic with known distribution for testing $\mu_1 = \mu_2$ under the setting in the first slide

Hypothesis test, p-value, and CI

Using t_k distribution as an approximation of the true distribution of T, we can construct a hypothesis test:

$H_0: \mu_1 - \mu_2 = 0$ vs $H_a: \mu_1 - \mu_2 \neq 0$

The test statistic is $T = \dfrac{\bar{Y}_1 - \bar{Y}_2 - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$.

Under H_0, $T \sim t_k$ approximately.

We reject H_0 if

$|T| \geq t_{k, \frac{\alpha}{2}}$.

The P-value is $P(|T| > |t|)$.

$T \sim t_k$

observed test statistic

Choosing k

How to choose an k such that t_k approximates the true distribution of T well.

Method 1

Choose

$$k = \min(n_1 - 1, n_2 - 1).$$

Method 2

Use

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

Choosing k - Method 1

Choosing $k = \min(n_1 - 1, n_2 - 1)$:

- Rationale of this approximation is the worse case scenario where either $\sigma_1^2 = 0$ or $\sigma_2^2 = 0$.

If $\sigma_1^2 = 0$, then $\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$ becomes $\frac{\sigma_2^2}{n_2}$.

We then have $T = \frac{(\bar{Y}_1 - \bar{Y}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_2^2}{n_2}}} \sim t_{n_2-1}$.

If $\sigma_2^2 = 0$, then $\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$ becomes $\frac{\sigma_1^2}{n_1}$.

Hence $T = \frac{(\bar{Y}_1 - \bar{Y}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1}}} \sim t_{n_1-1}$.

- This method is recommended for manual calculations
- It is conservative in the sense that the term $t_{k,\alpha/2}$ will either be correct or too large

Choosing k - Method 2

Using Welch's approximation: (Welch- Satterthwaite approximation)

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

$$T = \frac{\bar{Y}_1-\bar{Y}_2-(\mu_1-\mu_2)}{\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}}$$

$$= \frac{\bar{Y}_1-\bar{Y}_2-(\mu_1-\mu_2)}{\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}}\sqrt{\frac{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}$$

$$= \underbrace{\boxed{\frac{\bar{Y}_1-\bar{Y}_2-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}}}_{N(0,1)} \Bigg/ \sqrt{\underbrace{\boxed{\frac{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}}_{V}}$$

We want to find k such that

$$V \approx \frac{X}{k}$$

where $X \sim \chi^2_k$.

Choosing k - Method 2

We try to match the moments of V with that of $\frac{X}{k}$.

i.e. Solve $E[V] = E\left[\frac{X}{k}\right]$ for k.

$$E[V] = E\left[\frac{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}\right]$$

$$= \frac{\frac{E[S_1^2]}{n_1} + \frac{E[S_2^2]}{n_2}}{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

$$= \frac{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

$$= 1$$

Recall $X \sim \chi^2_k$

$$E[X] = k$$

$$\text{var}(X) = 2k$$

$$E\left[\frac{X}{k}\right] = \frac{1}{k}E[X] = \frac{k}{k} = 1$$

We have both $E[V]$ and $E\left[\frac{X}{k}\right]$ equal 1. This is not helpful in finding k. We will look at the second moment instead, which is equivalent to this.

$$\boxed{\text{Solve } \text{var}(V) = \text{var}\left(\frac{X}{k}\right) \text{ for } k.}$$

Choosing k - Method 2

$$\mathrm{var}(V) = \mathrm{var}\left(\frac{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}\right)$$

$$= \frac{\mathrm{var}\left(\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}\right)}{\left(\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}\right)^2}$$

$$= \frac{\mathrm{var}\left(\frac{S_1^2}{n_1}\right)+\mathrm{var}\left(\frac{S_2^2}{n_2}\right)}{\left(\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}\right)}$$

$$= \frac{\frac{2\sigma_1^4}{n_1^2(n_1-1)}+\frac{2\sigma_2^4}{n_2^2(n_2-1)}}{\left(\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}\right)}$$

Recall $\frac{(n_1-1)S_1^2}{\sigma_1^2} \sim \chi^2_{n_1-1}$

$$\mathrm{var}\left[\frac{(n_1-1)S_1^2}{\sigma_1^2}\right] = 2(n_1-1)$$

$$\left(\frac{n_1-1}{\sigma_1^2}\right)^2\mathrm{var}(S_1^2) = 2(n_1-1)$$

$$\mathrm{var}(S_1^2) = \frac{2\sigma_1^4}{n_1-1}$$

$$\mathrm{var}\left(\frac{S_1^2}{n_1}\right) = \frac{2\sigma_1^4}{n_1^2(n_1-1)}$$

(by independence of S_1^2 and S_2^2)

$$\mathrm{var}\left(\frac{S_2^2}{n_2}\right) = \frac{2\sigma_2^4}{n_2^2(n_2-1)}$$

$$\mathrm{var}\left(\frac{X}{k}\right) = \frac{1}{k^2}\mathrm{var}(X) = \frac{1}{k^2}(2k) = \frac{2}{k}.$$

Choosing k - Method 2

$$E[V] = E\left[\frac{X}{k}\right]$$

$$\frac{\frac{2\sigma_1^4}{n_1^2(n_1-1)} + \frac{2\sigma_2^4}{n_2^2(n_2-1)}}{\left(\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right)^2} = \frac{2}{k}$$

$$k = \frac{\cancel{2}\left(\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right)^2}{\frac{\cancel{2}\sigma_1^4}{n_1^2(n_1-1)} + \frac{\cancel{2}\sigma_2^4}{n_2^2(n_2-1)}}$$

$$= \frac{\left(\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right)^2}{\frac{\sigma_1^4}{n_1^2(n_1-1)} + \frac{\sigma_2^4}{n_2^2(n_2-1)}}$$

$$\approx \frac{\left(\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}\right)^2}{\frac{S_1^4}{n_1^2(n_1-1)} + \frac{S_2^4}{n_2^2(n_2-1)}}$$

If we use this approximation for our test statistic T, the t-test is called the Welch t-test.

The range of k under Welch's approximation is

$$\min(n_1-1,\ n_2-1) \leq k \leq n_1+n_2-2$$

(method 1) (pooled t-test)

Pooled versus not-pooled

'Rule of thumb'
Use a pooled two-sample t-test if

$$\frac{\max(s_1, s_2)}{\min(s_1, s_2)} < 2.$$

We may also do a preliminary F-test to test for $\sigma_1^2 = \sigma_2^2$.

Recommendation:

Use the non-pooled t-test, unless there are some compelling evidence to suggest $\sigma_1^2 = \sigma_2^2$.

Example 2.13

Summary statistics of the time required for a random sample of women and men to complete a test is shown in the table below. Do the data represent sufficient evidence to suggest a difference in the true mean time between men and women? Assume inequality of population variance and use $\alpha = 0.05$.

Men	Women
$n_1 = 18$	$n_2 = 12$
$\bar{y}_1 = 20.17$ seconds	$\bar{y}_2 = 19.23$ seconds
$s_1 = 4.3$	$s_2 = 3.8$

Example 2.13 Solution

Assume the populations are normally distributed.

Let $\mu_1 =$ true mean time for men

$\mu_2 =$ true mean time for women

$H_0: \mu_1 = \mu_2$ vs $H_a: \mu_1 \neq \mu_2$

$$T = \frac{\bar{y}_1 - \bar{y}_2 - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{(20.17 - 19.23) - 0}{\sqrt{\frac{4.3^2}{18} + \frac{3.8^2}{12}}} \approx 0.6294$$

$$k = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{s_1^4}{n_1^2(n_1-1)} + \frac{s_2^4}{n_2^2(n_2-1)}} = \frac{\left(\frac{4.3^2}{18} + \frac{3.8^2}{12}\right)^2}{\frac{4.3^4}{18^2(17)} + \frac{3.8^4}{12^2(11)}} \approx 25.685$$

$t_{25.685,\, 0.025} \approx 2.057$ qt(0.975, 25.685)

Critical region is $|T| \geq 2.057$. So T is not in the critical region.

There is insufficient evidence to reject H_0 at 0.05 significance level.