Many-sample problems

Sometimes we need to test whether several populations have the same means, or if at least one pair of means is different.

The null and research hypotheses are

\[ H_0 : \mu_1 = \mu_2 = \ldots = \mu_t \]  
\[ H_a : \text{At least one pair } \mu_i, \mu_j \text{ is different,} \]

where \( t \) is the number of populations, and \( \mu_1, \mu_2, \ldots, \mu_t \) are the population means.

This is a *many-sided* hypothesis test.

**Example 1: A bad idea.** Suppose that we have a list of 10 countries and we want to test whether or not the average lifespans in those countries are all the same.

1. Carry out the following hypothesis tests:

\[ H_{i,j}^0 : \mu_i = \mu_j \]
\[ H_{i,j}^a : \mu_i \neq \mu_j, \]

where \( \mu_i \) = mean lifespan in country \( i \), \( i \neq j \), \( \alpha = 0.05 \).

2. Reject the \( H_0 \) in (1) if we reject at least one of the \( H_{i,j}^0 \).

Why is this a bad idea?

- Even if \( \mu_1 = \mu_2 = \ldots = \mu_t \), we are likely to reject at least one of the \( H_{i,j}^0 \).

**Reason:**

- The probability of rejecting a \( H_{i,j}^0 \) if \( H_{i,j}^0 \) is true is ______________.
- Let \( X \) = the number of \( H_{i,j}^0 \) s that get rejected.
  - \( X \sim \) ______________
  - \( P(X \geq 1) = \)

**Example 2: Another bad idea.**

- Suppose that we want to see whether the average heights of the populations of men of 3 different countries are the same.
- Select random samples of size 10 from each country. Let \( \overline{y}_i \) = average of sample \( i \), \( i = 1, 2, 3 \).

- Reject \( H_0 \) if \( \overline{y}_1, \overline{y}_2, \overline{y}_3 \) are too spread apart.

The problem here is, how do we define "too" spread apart?
In the worksheet "May 10 lecture" are height data from 6 populations. Populations 1-3 are \(N(69, 2.5)\). Populations 4-6 are \(N(69, 5)\).

(a) Select 3 random samples of size 10 from populations 1-3, then find the spread of the sample means.

(b) Select 3 random samples of size 10 from populations 4-6, then find the spread of the sample means.

(c) Which set of means is more spread out? Why?

The right idea
Instead of computing the spread of the sample means, we have to compute the spread of the sample means compared to the combined spread of the individual samples.

To carry out the test (1), we (or Minitab) first compute the test statistic

\[
F = \frac{\text{variability of sample means}}{\text{variation within samples}}.
\]

To say precisely what \(F\) is we need some notation:

- \(t\) = number of populations.
- \(n_i\) = size of \(i\)th sample, \(i = 1, ..., t\).
- \(N\) = Total sample size = \(n_1 + n_2 + ... + n_t\).
- \(y_{ij}\) = \(j\)th sample observation from population \(i\).
- \(\bar{y}_i\) = average of \(i\)th sample.
- \(\bar{y}\) = average of all sample observations.

\[
MST = \text{"Mean square treatment"} = \frac{n \sum_{i=1}^{t} (\bar{y}_i - \bar{y})^2}{t - 1}.
\]

\(MST\) measures the variability of samples means.
\[ \text{MSE} = \text{"Mean square error"} = \frac{\sum_i \sum_j (y_{ij} - \bar{y}_i)^2}{N - t} \]

\( \text{MSE} \) measures the variation within samples.

The test statistic:
\[
F = \frac{\text{MST}}{\text{MSE}}.
\]

When to reject \( H_0 \):
Reject \( H_0 \) when
\[
F > F_{\alpha, t-1, N-t}.
\]

This testing procedure is called Analysis of Variance, ANOVA, or AOV.

Assumptions for ANOVA:
1. The samples are all independent, i.e. measurements in one sample are not connected in any way
2. All populations are normally distributed.
3. The variances of all the populations are the same.

Testing equality of several variances using Minitab
1. First, stack all data in one column, say \( c_r \). This column is called the response column.
2. In another column, say \( c_f \). In each row of \( c_f \), place a numeric tag that identifies which population the corresponding entry in \( c_r \) came from.
3. Select
   \[
   \text{Stat} \to \text{ANOVA} \to \text{Test for equal variances}
   \]
4. In the "Response" window, enter \( c_r \). In the "Factors" window, enter \( c_f \).
5. Minitab will test \( H_0 \) : All variances are equal vs. \( H_a \) : Not all variances are equal. Use \( \alpha = 0.05 \) and the \( p \)-value from Levine’s test. If \( p \leq \alpha \), you cannot use ANOVA.

Example 3a: A researcher wants to try three different techniques to lower blood pressure. One group of subjects takes medication, the second group exercises, and the third group follows a special diet. After four weeks, the reduction in blood pressure for each subject is recorded. The data is in the worksheet "May 10 lecture.mws". Test the equality of the variances of the three populations. What is the \( p \)-value? What do you conclude?
Performing ANOVA test using Minitab

Now that you have stacked your data and you have made a "factors" column, you can use Minitab to use ANOVA to test equality of population means.

**Example 3b:** Test the equality of the means of the populations from part 3a with $\alpha = 0.05$, using Minitab (don’t forget to test the normality assumption). What is the $F$–statistic and $p$–value?

What is your conclusion? Also, using the confidence intervals given by Minitab, which treatment seems most effective?

To begin: select

Stat $\rightarrow$ ANOVA $\rightarrow$ One-way

In the "Response" window, enter the column with the blood pressure reductions.
In the "Factors" window, enter the column with the sample numbers.