Statistical Methods

Nels Grevstad

Metropolitan State University of Denver

ngrevsta@msudenver.edu

September 2, 2019

Topics

Objectives

Objectives:

- Carry out a pooled two-sample t test for two population means.
- Compute and interpret a pooled two-sample t CI for the difference between two population means.

Pooled t Test for Two Population Means μ_1 and μ_2

- Suppose again we have random samples of sizes m and n from **two populations**, but now suppose also that the **population standard deviations** σ_1 and σ_2 are **equal.**

Pooled t Test for Two Population Means μ_1 and μ_2

- Suppose again we have random samples of sizes m and n from **two populations**, but now suppose also that the **population standard deviations** σ_1 and σ_2 are **equal.**

  The appropriate test for deciding if the **population means** μ_1 and μ_2 are different is called the ***pooled two-sample t test for $\mu_1 - \mu_2$.***

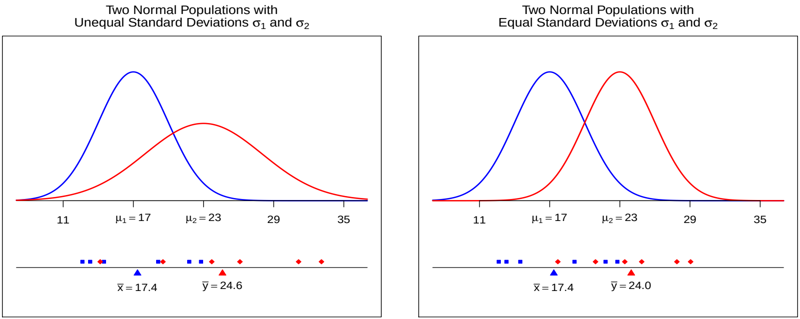


Figure: Normal populations with unequal standard deviations (left), and equal standard deviations (right).

- The **null hypothesis** is that no difference between the population means μ_1 and μ_2:

Null Hypothesis:

$$H_0 : \mu_1 - \mu_2 \; = \; 0$$

- The **alternative hypothesis** will depend on what we're trying to "prove":

Alternative Hypothesis: The alternative hypothesis will be one of

1. $H_a : \mu_1 - \mu_2 > 0$ (**one-sided, upper-tailed**)
2. $H_a : \mu_1 - \mu_2 < 0$ (**one-sided, lower-tailed**)
3. $H_a : \mu_1 - \mu_2 \neq 0$ (**two-sided, two-tailed**)

depending on what we're trying to verify using the data.

- Letting σ be the **common population standard deviation**, if the samples are from **normal** populations (or m and n are both **large**),

$$\bar{X} - \bar{Y} \sim N\left(\mu_1 - \mu_2, \sqrt{\frac{\sigma^2}{m} + \frac{\sigma^2}{n}}\right),$$

and

$$Z = \frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{\sqrt{\sigma^2/m + \sigma^2/n}} \sim N(0,1). \tag{1}$$

- The ***pooled estimator of σ^2*** is defined as

$$S_p^2 = \frac{\sum_{i=1}^{m}(X_i - \bar{X})^2 + \sum_{i=1}^{n}(Y_i - \bar{Y})^2}{n+m-2}$$

- The ***pooled estimator of σ^2*** is defined as

$$
\begin{aligned}
S_p^2 &= \frac{\sum_{i=1}^m (X_i - \bar{X})^2 + \sum_{i=1}^n (Y_i - \bar{Y})^2}{n+m-2} \\
&= \frac{(m-1)S_1^2 + (n-1)S_2^2}{n+m-2}
\end{aligned}
$$

- The ***pooled estimator of σ^2*** is defined as

$$\begin{aligned} S_p^2 &= \frac{\sum_{i=1}^{m}(X_i - \bar{X})^2 + \sum_{i=1}^{n}(Y_i - \bar{Y})^2}{n+m-2} \\ &= \frac{(m-1)S_1^2 + (n-1)S_2^2}{n+m-2} \\ &= \frac{(m-1)}{n+m-2}S_1^2 + \frac{(n-1)}{n+m-2}S_2^2 \,, \end{aligned} \tag{2}$$

- The ***pooled estimator of σ^2*** is defined as

$$\begin{aligned} S_p^2 &= \frac{\sum_{i=1}^m (X_i - \bar{X})^2 + \sum_{i=1}^n (Y_i - \bar{Y})^2}{n+m-2} \\ &= \frac{(m-1)S_1^2 + (n-1)S_2^2}{n+m-2} \\ &= \frac{(m-1)}{n+m-2} S_1^2 + \frac{(n-1)}{n+m-2} S_2^2 \, , \end{aligned} \tag{2}$$

(a **weighted average** of the **sample variances** S_1^2 and S_2^2).

Proposition

If $X_1, X_2, \ldots, X_m$ are a random sample from a $N(\mu_1, \sigma)$ distribution and $Y_1, Y_2, \ldots, Y_n$ are a random sample from a $N(\mu_2, \sigma)$ distribution, and the two samples are drawn *independently* of each other, then

$$T = \frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{\sqrt{S_p^2/m + S_p^2/n}} \sim t(n + m - 2)\,. \tag{3}$$

Furthermore, (3) still holds (at least approximately) even if the samples are from **non-normal** populations as long as the sample sizes m and n are both **large**.

Pooled Two-Sample t Test Statistic for $\mu_1 - \mu_2$:

$$T = \frac{\bar{X} - \bar{Y} - 0}{\sqrt{S_p^2/m + S_p^2/n}}$$

1. **Large positive** values of t provide **evidence against H_0 in favor of**
   $H_a : \mu_1 - \mu_2 > 0$.
2. **Large negative** values of t provide **evidence against H_0 in favor of**
   $H_a : \mu_1 - \mu_2 < 0$.
3. **Large positive *and* large negative** values of t provide **evidence against H_0 in favor of**
   $H_a : \mu_1 - \mu_2 \neq 0$.

Sampling Distribution of the Test Statistic Under H_0:
If t is the pooled two-sample t test statistic, then when

$$H_0 : \mu_1 - \mu_2 = 0$$

is true,

$$t \sim t(m + n - 2).$$

- The $t(m + n - 2)$ curve gives us:

- The $t(m+n-2)$ curve gives us:
  - The ***rejection region*** as the **extreme 100α% of t values** (in the direction(s) specified by H_a).

- The $t(m+n-2)$ curve gives us:
  - The ***rejection region*** as the **extreme 100α% of t values** (in the direction(s) specified by H_a).
  - The ***p-value*** as the **tail area(s) beyond the observed t value** (in the direction(s) specified by H_a).

Exercise

A computer simulation of a complex semiconductor manufacturing line was carried out to examine the sensitivity of the cycle time (time in seconds from beginning to completion of the manufacturing process) to the number of operators stationed at one point along the line.

Exercise

A computer simulation of a complex semiconductor manufacturing line was carried out to examine the sensitivity of the cycle time (time in seconds from beginning to completion of the manufacturing process) to the number of operators stationed at one point along the line.

The model was run separately **16** times under each of two conditions: Using just **one operator** at the station, and using **two**. For each run, the cycle time was recorded.

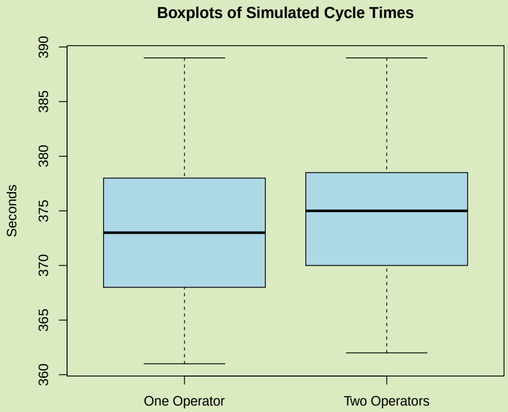
Here are the summary statistics:

One Operator			**Two Operators**		
m	$=$	16	n	$=$	16
$\bar{x}$	$=$	373.6	$\bar{y}$	$=$	374.8
s_1	$=$	7.8	s_2	$=$	7.3

Here are the summary statistics:

One Operator			Two Operators		
m	$=$	16	n	$=$	16
$\bar{x}$	$=$	373.6	$\bar{y}$	$=$	374.8
s_1	$=$	7.8	s_2	$=$	7.3

Side-by-side boxplots are shown on the next slide.


Boxplots of Simulated Cycle Times
Seconds
360
365
370
375
380
385
390
One Operator
Two Operators

The two **sample standard deviations** S_1 and S_2 are roughly the same, so it's reasonable to assume that σ_1 and σ_2 are **equal**.

The two **sample standard deviations** S_1 and S_2 are roughly the same, so it's reasonable to assume that σ_1 and σ_2 are **equal**.

Carry out the **pooled two-sample** t **test** to decide if it makes **any difference** whether **one** or **two operators** are used. Use a **level of significance** $\alpha = 0.05$.

The two **sample standard deviations** S_1 and S_2 are roughly the same, so it's reasonable to assume that σ_1 and σ_2 are **equal**.

Carry out the **pooled two-sample t test** to decide if it makes **any difference** whether **one** or **two operators** are used. Use a **level of significance** $\alpha = 0.05$.

Hint: You should end up with $S_p^2 = 57.1$, $t = -0.5$, and **p-value = 0.620**.

Pooled t Confidence Interval for $\mu_1 - \mu_2$

Pooled Two-Sample CI for $\mu_1 - \mu_2$: Suppose $X_1, X_2, \ldots, X_m$ and $Y_1, Y_2, \ldots, Y_n$ are independent random samples from populations whose means are μ_1 and μ_2 and whose *standard deviations are equal*. Then a $100(1-\alpha)\%$ ***pooled two-sample t confidence interval for $\mu_1 - \mu_2$*** is

$$\bar{X} - \bar{Y} \pm t_{\alpha/2,\, m+n-2} \cdot \sqrt{\frac{S_p^2}{m} + \frac{S_p^2}{n}}.$$

- The CI is valid if the samples are from **populations** whose **standard deviations** are **equal** and either the populations **normal** or m and n **are large.**

- The CI is valid if the samples are from **populations** whose **standard deviations** are **equal** and either the populations **normal** or m and n **are large**.
- In either case, we can be $100(1-\alpha)\%$ confident that $\mu_1 - \mu_2$ will be contained in the CI.

Exercise

Using the simulated cycle times data of the last exercise, compute and interpret a **95% pooled t confidence interval for $\boldsymbol{\mu_1 - \mu_2}$.**

Exercise

Using the simulated cycle times data of the last exercise, compute and interpret a **95% pooled t confidence interval for $\mu_1 - \mu_2$.**

Hint: Using $t_{\alpha/2,\, m+n-2} = 2.042$, you should get $(-6.66,\ 4.26)$

To Pool or Not to Pool: That's the Question (9.2)

- Although the **pooled** (two-sample t test and CI) procedures outperform the **non-pooled** procedures by a bit (smaller Type II error probability and narrower CI) when $\sigma_1 = \sigma_2$, ...

To Pool or Not to Pool: That's the Question (9.2)

- Although the **pooled** (two-sample t test and CI) procedures outperform the **non-pooled** procedures by a bit (smaller Type II error probability and narrower CI) when $\sigma_1 = \sigma_2$, ...

  the former can lead to erroneous conclusions when $\sigma_1 \neq \sigma_2$, ...

To Pool or Not to Pool: That's the Question (9.2)

- Although the **pooled** (two-sample t test and CI) procedures outperform the **non-pooled** procedures by a bit (smaller Type II error probability and narrower CI) when $\sigma_1 = \sigma_2$, ...

  the former can lead to erroneous conclusions when $\sigma_1 \neq \sigma_2$, ...

  i.e. they're **not *robust*** to violations of the equal population standard deviations assumption.

To Pool or Not to Pool: That's the Question (9.2)

- Although the **pooled** (two-sample t test and CI) procedures outperform the **non-pooled** procedures by a bit (smaller Type II error probability and narrower CI) when $\sigma_1 = \sigma_2$, ...

  the former can lead to erroneous conclusions when $\sigma_1 \neq \sigma_2$, ...

  i.e. they're **not *robust*** to violations of the equal population standard deviations assumption.

- Therefore, the **non-pooled** procedures are **preferred** unless there's compelling evidence for doing otherwise.