# *Hypothesis testing for two populations*

*Chrysafis Vogiatzis*

*Lecture 28-29*

Learning objectives

After lectures 28–29, we will be able to:

- Accept or reject hypotheses for two populations:
  - for the difference between their means.
  - for the difference between their proportions.
  - for the ratio of their variances.
- Use this statistical tool to compare two populations and make decisions about them.

## *Motivation: weather differences*

You may have heard people say something along the lines "The weather is so different nowadays!" or "It used to snow during Halloween when I was a kid!" or even something like "Last year, it was much warmer/colder!". How can we employ statistics and probability theory to **reject** or **fail to reject** such claims? Could we somehow compare the mean temperature/snow/humidity/etc. from one year to the next?

## *Motivation: electoral considerations*

During an election, political parties and candidates would like to know how specific populations behave. Do farmers overwhelmingly care about one item versus another? How about first-generation college students? We then would like to check and compare different populations, hopefully to find common ground that can help us address as many issues as possible, without alienating one group or another.

## *Motivation: online education and audiovisual tools*

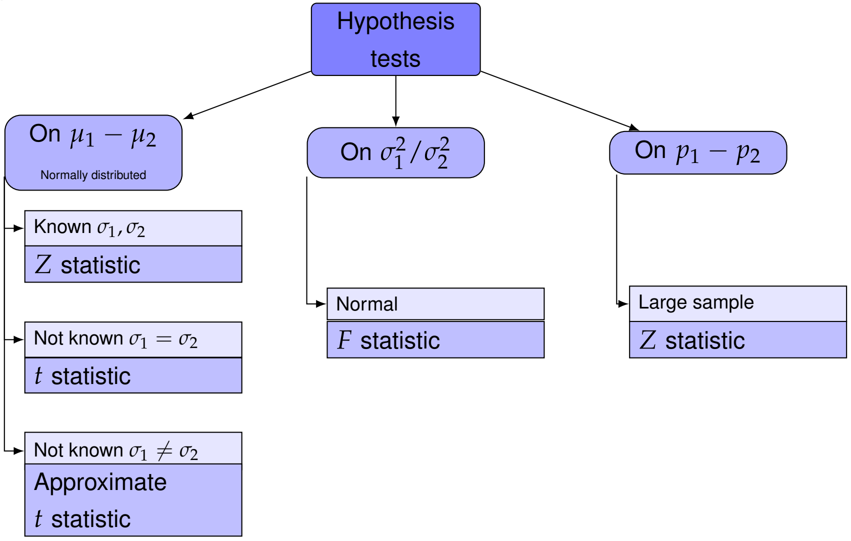
It is easy to completely demonize or completely agree with online education and its tools. What is more difficult is to quantify what happens with the variability of the performance of students in an online setting with the audiovisual tools at our disposal. Can we test the claim whether properly designed online courses lead to *lower variability* in the learning of students?

## *Hypothesis testing for two populations*

I know we just discussed some motivating examples. Let me state them here in a more specific setting:

- Is the weather in Chicago significantly different than it was 10 years ago?
- Do students who have access to audiovisual aids for a class perform better than students who do not?
- Does the variability in the duration of a call decrease when the signal reception is improved?
- Do voters from one group overwhelmingly prefer one candidate over another in a local election compared to another group of voters?

All of the above examples have one thing in common: they deal with two populations and how they **compare** and **contrast**. Like we did in the past, we will again deal here with hypothesis testing for means, variances, and proportions. Visually, we discuss the following:


Hypothesis tests
On $\mu_1 - \mu_2$
Normally distributed
On $\sigma_1^2/\sigma_2^2$
On $p_1 - p_2$
Known $\sigma_1, \sigma_2$
$Z$ statistic
Not known $\sigma_1 = \sigma_2$
$t$ statistic
Not known $\sigma_1 \neq \sigma_2$
Approximate $t$ statistic
Normal
$F$ statistic
Large sample
$Z$ statistic

## *Hypothesis testing for means of two normally distributed populations*

We have three cases (consult the earlier figure). They are:

1. normally distributed populations with known variances $\sigma_1^2, \sigma_2^2$.
2. normally distributed populations with unknown variances that are known to be equal, that is unknown $\sigma_1^2 = \sigma_2^2$.
3. normally distributed populations with unknown variances that are not known to be equal, that is unknown $\sigma_1^2 \neq \sigma_2^2$.

Their derivation is again based on their confidence intervals, so we simply provide a summary of their results.

### *Normally distributed populations with known variances $\sigma_1^2, \sigma_2^2$*

A quick review before getting started.

- Assume two normally distributed populations $X, Y$ with mean $\mu_1, \mu_2$ and standard deviations $\sigma_1, \sigma_2$. Then:
  1. Pick a sample of $n_1$ elements from $X$: $\overline{X} \sim \mathcal{N}\left(\mu_1, \sigma_1^2/n_1\right)$.
  2. Pick a sample of $n_2$ elements from $Y$: $\overline{Y} \sim \mathcal{N}\left(\mu_2, \sigma_2^2/n_2\right)$.

- Additionally, for combinations of the two populations, we have:
  1. $X + Y \sim \mathcal{N}\left(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2\right)$.
  2. $X - Y \sim \mathcal{N}\left(\mu_1 - \mu_2, \sigma_1^2 + \sigma_2^2\right)$.
  3. $aX + bY \sim \mathcal{N}\left(a\mu_1 + b\mu_2, a^2\sigma_1^2 + b^2\sigma_2^2\right)$.

- Finally, consider we pick a sample $n_1$ from $X$ and a sample $n_2$ from $Y$:
  1. Pick a sample of $n_1$ elements from $X$: $\overline{X} \sim \mathcal{N}\left(\mu_1, \sigma_1^2/n_1\right)$.
  2. Pick a sample of $n_2$ elements from $Y$: $\overline{Y} \sim \mathcal{N}\left(\mu_2, \sigma_2^2/n_2\right)$.
  3. Combine to get that

$$\boxed{\overline{X} - \overline{Y} \sim \mathcal{N}\left(\mu_1 - \mu_2, \sigma_1^2/n_1 + \sigma_2^2/n_2\right).}$$

Means with known $\sigma_1^2, \sigma_2^2$

Null hypothesis: Test statistic: Distribution:

$$H_0 : \mu_1 - \mu_2 = \Delta_0. \quad Z_0 = \frac{\left(\overline{X}_1 - \overline{X}_2\right) - \Delta_0}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}}. \quad Z_0 \sim \mathcal{N}\left(0, 1\right).$$

| $H_1$ | Rejection region | $P$-value |
|---|---|---|
| $\mu_1 - \mu_2 \neq \Delta_0$ | $\lvert Z_0 \rvert > z_{\alpha/2}$ | $2 \cdot (1 - \Phi(\lvert Z_0 \rvert))$ |
| $\mu_1 - \mu_2 > \Delta_0$ | $Z_0 > z_\alpha$ | $1 - \Phi(Z_0)$ |
| $\mu_1 - \mu_2 < \Delta_0$ | $Z_0 < -z_\alpha$ | $\Phi(Z_0)$ |

Like in the single population cases, we should reject the null hypothesis under the following conditions:

1. Check whether the observed sample average $\overline{X}$ or the $Z_0$ statistic fall in the rejection region.

2. Calculate the $P$-value and compare to $\alpha$.

Vaping

Two vaping products are being tested for their relationship with the outbreak of lung injury (see this CDC link). The first product has been responsible for more illnesses, so we are interested in seeing whether the nicotine content is at least 0.2 milligrams higher than in the second product. We have found that $n_1 = 50$ products of the first kind had an average nicotine content of $\overline{X}_1 = 2.61$ milligrams and $n_2 = 40$ products of the second kind had $\overline{X}_2 = 2.38$ milligrams. Using $\alpha = 0.05$, can we claim that the first product has 0.2 milligrams of difference or is it higher? Assume that standard deviations per product are known and equal to $\sigma_1 = 0.8$ and $\sigma_2 = 1.1$ milligrams, respectively.

We want to compare two population means: more specifically we want to see whether the difference is $\Delta_0 = 0.2$. We then have:

$$H_0 : \mu_1 - \mu_2 = 0.2 \quad H_1 : \mu_1 - \mu_2 > 0.2.$$

We pick:

- from the first population: $n_1 = 50, \overline{X}_1 = 2.61, \sigma_1 = 0.8$
- from the second population: $n_2 = 40, \overline{X}_2 = 2.38, \sigma_2 = 1.1$

**Vaping**

Now, calculate the test statistic as:

$$Z_0 = \frac{2.61 - 2.38 - 0.2}{\sqrt{\frac{0.8^2}{50} + \frac{1.1^2}{40}}} = \frac{0.03}{0.21} = 1/7 = 0.14.$$

It is one-sided, so find critical value $z_\alpha = z_{0.05} = 1.645$. Seeing as $Z_0 \leq z_\alpha$, we *fail to reject*.

We did not end up needing this, but we could have used the corresponding confidence intervals to decide whether we want to reject a hypothesis or not. How? First, calculate $\overline{X}_1 - \overline{X}_2$. Then, check the CI region: if the difference of the sample averages falls within or the confidence interval, then we fail to reject the null hypothesis.

| $H_1$ | CI region |
|---|---|
| $\mu_1 - \mu_2 \neq \Delta_0$ | $(\mu_1 - \mu_2) \pm z_{\alpha/2}\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}$ |
| $\mu_1 - \mu_2 > \Delta_0$ | $\left(-\infty, (\mu_1 - \mu_2) + z_\alpha\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}\,\right]$ |
| $\mu_1 - \mu_2 < \Delta_0$ | $\left[(\mu_1 - \mu_2) - z_\alpha\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}, +\infty\right)$ |

### *Normally distributed populations with unknown, but equal, variances* $\sigma_1^2 = \sigma_2^2$

Let's try to derive the procedure now! First, assume that $\sigma_1 = \sigma_2 = \sigma$. Then the test statistic can be written as:

$$Z_0 = \frac{\left(\overline{X}_1 - \overline{X}_2\right) - \Delta_0}{\sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}.$$

You've guessed the next step! If $\sigma$ is unknown, I need to somehow estimate it.. Can we use a sample standard deviation? Recall that the sample standard deviation $s$ can be a good estimator for the unknown population standard deviation $\sigma$. However, that was for a single population. What can we do here?

Since we have two samples from two populations, each with their own (possibly different) sample standard deviations $s_1, s_2$, we use the so-called **pooled estimator**, where we treat both as if they are one population. We then get:

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

A couple of notes:

- $s_p^2$ is a weighted average of the two variances $s_1, s_2$.
- Letting $n_1 = n_2$ leads to $s_p^2 = \left(s_1^2 + s_2^2\right)/2$.

Finally, as we are moving from *known* variances to *unknown* ones, we also need to account for it by moving from a normal distribution (and its $z$ values) to a Student's $T$ distribution (and the corresponding $t$ values). Note that the distribution has $n_1 + n_2 - 2$ degrees of freedom. Overall we have:

Means with unknown but equal $\sigma_1^2 = \sigma_2^2$

| Null hypothesis: | Test statistic: | Distribution: |
|---|---|---|
| $H_0 : \mu_1 - \mu_2 = \Delta_0.$ | $T_0 = \dfrac{\left(\overline{X}_1 - \overline{X}_2\right) - \Delta_0}{s_p\sqrt{1/n_1 + 1/n_2}}.$ | $T_0 \sim T_{n_1+n_2-2}.$ |

| $H_1$ | Rejection region | $P$-value |
|---|---|---|
| $\mu_1 - \mu_2 \neq \Delta_0$ | $\lvert T_0\rvert > t_{\alpha/2,n_1+n_2-2}$ | $2 \cdot \left(1 - T_{n_1+n_2-2}(\lvert T_0\rvert)\right)$ |
| $\mu_1 - \mu_2 > \Delta_0$ | $T_0 > t_{\alpha,n_1+n_2-2}$ | $1 - T_{n_1+n_2-2}(T_0)$ |
| $\mu_1 - \mu_2 < \Delta_0$ | $T_0 < -t_{\alpha,n_1+n_2-2}$ | $T_{n_1+n_2-2}(T_0)$ |

Let us not forget that we may also decide to reject or not based on whether $\overline{X}_1 - \overline{X}_2$ falls within or outside the corresponding confidence interval!

| $H_1$ | CI region |
|---|---|
| $\mu_1 - \mu_2 \neq \Delta_0$ | $(\mu_1 - \mu_2) \pm t_{\alpha/2,n_1+n_2-2} s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$ |
| $\mu_1 - \mu_2 > \Delta_0$ | $\left(-\infty, (\mu_1 - \mu_2) + t_{\alpha,n_1+n_2-2} s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}\ \right]$ |
| $\mu_1 - \mu_2 < \Delta_0$ | $\left[(\mu_1 - \mu_2) - t_{\alpha,n_1+n_2-2} s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}, +\infty\right)$ |

Catalyst comparison

Two catalysts, $A$ and $B$, are being compared to see how they affect the mean yield of a chemical process. We have devised a pilot operation and results using the two catalysts are shown below for 8 runs. Using $\alpha = 0.05$ and assuming unknown but equal standard deviations, can we deduce that the two catalysts affect the yield differently?

| Run | $A$ | $B$ | Run | $A$ | $B$ |
|---|---|---|---|---|---|
| 1 | 91.5 | 89.19 | 5 | 91.79 | 97.19 |
| 2 | 94.18 | 90.95 | 6 | 89.07 | 97.04 |
| 3 | 92.18 | 90.46 | 7 | 94.72 | 91.07 |
| 4 | 95.39 | 93.21 | 8 | 89.21 | 92.75 |

We have:

- Population 1 for Catalyst A with: $n_1 = 8, \overline{X}_1 = 92.255, s_1 = 2.39$
- Population 2 for Catalyst B with: $n_2 = 8, \overline{X}_2 = 92.7325, s_2 = 2.98$

Recall that we know that $\sigma_1 = \sigma_2$, but this will not imply that the sample standard deviations will be equal too!
Now, on to formulating the hypothesis. We have:

$$H_0 : \mu_1 - \mu_2 = 0 \quad H_1 : \mu_1 - \mu_2 \neq 0.$$

The pooled standard deviation is:

$$s_P = \sqrt{\frac{(n_1 - 1)\, s_1^2 + (n_2 - 1)\, s_2^2}{n_1 + n_2 - 2}} = \sqrt{\frac{7 \cdot 2.39^2 + 7 \cdot 2.98^2}{14}} = 2.7.$$

With all that, we get the corresponding test statistic as:

$$T_0 = \frac{92.255 - 92.7325 - 0}{2.7\sqrt{\frac{1}{8} + \frac{1}{8}}} = \frac{-0.4775}{1.35} = -0.35.$$

Since $\alpha = 0.05$ and we have a two-sided hypothesis, we need to identify the proper critical value as $t_{\alpha/2,14} = t_{0.025,14} = 2.145$. As $-t_{\alpha/2,14} \leq t_0 \leq t_{\alpha/2,14}$, we *fail to reject*.

### *Normally distributed populations with unknown, and not necessarily equal, variances $\sigma_1^2 \neq \sigma_2^2$*

In this case, things get a little more complicated. Had we known what $\sigma_1, \sigma_2$ were:

$$Z_0 = \frac{\left(\overline{X}_1 - \overline{X}_2\right) - \Delta_0}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}} \sim \mathcal{N}(0,1)\,.$$

Replacing $\sigma_1$, $\sigma_2$ with their sample counterparts $s_1, s_2$, we get:

$$T_0 = \frac{\left(\overline{X}_1 - \overline{X}_2\right) - \Delta_0}{\sqrt{s_1^2/n_1 + s_2^2/n_2}} \sim T_v.$$

We then say that $T_0$ is distributed *approximately* as the $T$ distribution, but with degrees of freedom equal to $v$:

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

This number will usually be fractional – we typically round down when needing to consult a $t$-table.

Means with unknown and not necessarily equal $\sigma_1^2 \neq \sigma_2^2$

| Null hypothesis: | Test statistic: | Distribution: |
|---|---|---|
| $H_0 : \mu_1 - \mu_2 = \Delta_0.$ | $T_0 = \frac{\left(\overline{X}_1 - \overline{X}_2\right) - \Delta_0}{\sqrt{s_1^2/n_1 + s_2^2/n_2}}.$ | $T_0 \sim T_v.$ |

| $H_1$ | Rejection region | $P$-value |
|---|---|---|
| $\mu_1 - \mu_2 \neq \Delta_0$ | $\lvert T_0\rvert > t_{\alpha/2,v}$ | $2 \cdot (1 - T_v(\lvert T_0\rvert))$ |
| $\mu_1 - \mu_2 > \Delta_0$ | $T_0 > t_{\alpha,v}$ | $1 - T_v(T_0)$ |
| $\mu_1 - \mu_2 < \Delta_0$ | $T_0 < -t_{\alpha,v}$ | $T_v(T_0)$ |

In the above, we calculate the approximate degrees of freedom $v$ as:

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

Let's put that to the use. We will follow the same example as before, however now we will drop the assumption that the two variances are equal.

### Catalyst comparison

Two catalysts, $A$ and $B$, are being compared to see how they affect the mean yield of a chemical process. We have devised a pilot operation and results using the two catalysts are shown below for 8 runs. Using $\alpha = 0.05$ and assuming unknown standard deviations, can we deduce that the two catalysts affect the yield differently?

| Run | $A$ | $B$ | Run | $A$ | $B$ |
|---|---|---|---|---|---|
| 1 | 91.5 | 89.19 | 5 | 91.79 | 97.19 |
| 2 | 94.18 | 90.95 | 6 | 89.07 | 97.04 |
| 3 | 92.18 | 90.46 | 7 | 94.72 | 91.07 |
| 4 | 95.39 | 93.21 | 8 | 89.21 | 92.75 |

We follow a very similar logic to earlier. However, we now will need the approximate degrees of freedom before proceeding (plus the $T_0$ statistic calculation changes slightly). We have the same hypothesis $H_0 : \mu_1 - \mu_2 = 0 \quad H_1 : \mu_1 - \mu_2 \neq 0$ and the same $n_1 = 8, \overline{X}_1 = 92.255, s_1 = 2.39, n_2 = 8, \overline{X}_2 = 92.7325, s_2 = 2.98$. Here is where things change now:

1. Calculate test statistic:

$$T_0 = \frac{92.255 - 92.7325 - 0}{\sqrt{\frac{2.39^2}{8} + \frac{2.98^2}{8}}} = \frac{-0.4775}{1.35} = -0.35.$$

2. Calculate approximate degrees of freedom:

$$v = \frac{\left(s_1^2/n_1 + s_2^2/n_2\right)^2}{\frac{\left(s_1^2/n_1\right)^2}{n_1-1} + \frac{\left(s_2^2/n_2\right)^2}{n_2-1}} = \frac{(0.714 + 1.11)^2}{\frac{0.714^2}{7} + \frac{1.11^2}{7}} =$$

$$= \frac{1.824^2}{0.073 + 0.176} = 13.361 \to 13.$$

Finally, we find $t_{\alpha/2,v} = t_{0.025,13} = 2.16$. Becase $-t_{\alpha/2,13} \leq t_0 \leq t_{\alpha/2,13}$, we *fail to reject*.

## *Hypothesis testing for the ratio of the variances of two normally distributed populations*

As must be obvious by now, we are taking each two population confidence interval and adapting it to the hypothesis testing procedure. Next up is the ratio of the two unknown variances of two **normally**

**distributed** populations.

Ratio of variances

Null hypothesis: Test statistic: Distribution:

$$H_0 : \sigma_1^2 = \sigma_2^2. \qquad F_0 = \frac{s_1^2}{s_2^2}. \qquad F_0 \sim F_{n_1-1,n_2-1}.$$

| $H_1$ | Rejection region |
|---|---|
| $\sigma_1^2 \neq \sigma_2^2$ | $F_0 > f_{\alpha/2,n_1-1,n_2-1}$ or<br>$F_0 < f_{1-\alpha/2,n_1-1,n_2-1}$ |
| $\sigma_1^2 > \sigma_2^2$ | $F_0 > f_{\alpha,n_1-1,n_2-1}$ |
| $\sigma_1^2 < \sigma_2^2 0$ | $F_0 < f_{1-\alpha,n_1-1,n_2-1}$ |

Variability in thickness

The variability in the thickness of oxide layers in semiconductor wafers is a critical characteristic, where low variability is desirable. A company is investigating two different ways to mix gases so as to reduce the variability of the oxide thickness. We produce 16 wafers with each gas mixture and our results indicate that the standard deviation is $s_1 = 1.96$Å and $s_2 = 2.13$Å for the two mixtures. Using $\alpha = 0.05$, is there evidence to indicate that either gas is preferable for better wafers?

We have two populations: i) population 1 with: $n_1 = 16, s_1 = 1.96$ and ii) population 2 with: $n_2 = 16, s_2 = 2.13$. As always, we begin by formulating our hypothesis as

$$H_0 : \sigma_1^2 = \sigma_2^2 \quad H_1 : \sigma_1^2 \neq \sigma_2^2.$$

Proceed to calculate our test statistic, based on the sample variances as:

$$F_0 = \frac{s_1^2}{s_2^2} = \frac{1.96^2}{2.13^2} = 0.8467.$$

Since this is a two-sided hypothesis test, we need two critical values (recall that the $F$ distribution is not symmetric!): $f_{\alpha/2,n_1-1,n_2-1} = f_{0.025,15,15} = 2.86$ and $f_{1-\alpha/2,n_1-1,n_2-1} = \frac{1}{f_{\alpha/2,n_2-1,n_1-1}} = \frac{1}{2.86} = 0.35$. Seeing as $f_{1-\alpha/2,n_1-1,n_2-1} \leq F_0 \leq f_{\alpha/2,n_1-1,n_2-1}$, we *fail to reject*: that is, we do not have enough evidence to claim that the two variances are not equal.

## *Hypothesis testing for the difference in the proportions of two populations*

We finish our venture in two population hypothesis testing with proportions. It should come as no surprise that this also emulates the discussion of the two population proportions confidence interval we had seen earlier in the class! A few definitions and assumptions before we start:

1. *Large* samples from the two populations ($n_i p_i \geq 30$ and $n_1(1-p_i) \geq 30$ for both populations $i = 1, 2$).
2. sample size and observed proportion from population 1: $n_1, \hat{p}_1$, and sample size and observed proportion from population 2: $n_2, \hat{p}_2$.
3. a (hypothesized) difference $p_1 - p_2 = \Delta_0$.
4. a **pooled proportion estimator** in the form of

$$\hat{p} = \frac{n_1\hat{p}_1 + n_2\hat{p}_2}{n_1 + n_2}.$$

   - Much like a *weighted average* of the two observed proportions.

With these available, we may derive the hypothesis testing procedure as follows:

Proportions of two populations, $p_1, p_2$

Null hypothesis: Test statistic: Distribution:

$H_0 : p_1 - p_2 = \Delta_0.$

$$Z_0 = \frac{(\hat{p}_1 - \hat{p}_2) - \Delta_0}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}.$$

$Z_0 \sim \mathcal{N}(0,1)$.

| $H_1$ | Rejection region | $P$-value |
|---|---|---|
| $p_1 - p_2 \neq \Delta_0$ | $\lvert Z_0 \rvert > z_{\alpha/2}$ | $2 \cdot (1 - \Phi(\lvert Z_0 \rvert))$ |
| $p_1 - p_2 > \Delta_0$ | $Z_0 > z_\alpha$ | $1 - \Phi(Z_0)$ |
| $p_1 - p_2 < \Delta_0$ | $Z_0 < -z_\alpha$ | $\Phi(Z_0)$ |

Politicians favorability ratings

A recent survey asked people in Urbana and Champaign whether they like their elected officials. Out of 118 Urbana residents, 37 said yes; for Champaign citizens there were 135 residents, among whom 61 said yes. Is there significant evidence (using $\alpha = 0.05$) to assume that Champaign's citizens showcase higher approval rates for their elected officials?

First collect our information:

- Urbana: $n_1 = 118, \hat{p}_1 = \frac{37}{118} = 0.314$.
- Champaign: $n_2 = 135, \hat{p}_2 = \frac{61}{135} = 0.452$.

We formulate the hypothesis as

$$H_0 : p_1 = p_2 \quad H_1 : p_1 < p_2.$$

Then, calculate the pooled proportion estimator as $\hat{p} = \frac{37+61}{118+135} = \frac{98}{253} = 0.387$. We are now ready to calculate the test statistic:

$$\begin{aligned} Z_0 &= \frac{(\hat{p}_1 - \hat{p}_2) - \Delta_0}{\sqrt{\hat{p}\,(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \\ &= \frac{0.138}{\sqrt{0.387 \cdot 0.613 \cdot (1/118 + 1/135)}} = 2.25. \end{aligned}$$

To reject or fail to reject, we need the critical value for $\alpha = 0.05$:$z_\alpha = z_{0.05} = 1.64$. Because $Z_0 > z_\alpha$, we have to *reject the hypothesis*: hence, we deduce that Champaign does indeed have higher approval rates for elected officials (under $\alpha = 0.05$)!