

Hypothesis testing for means and variances

Chrysafis Vogiatzis

Lecture 26-27

Learning objectives

After lectures 26–27, we will be able to:

- Accept or reject hypotheses for means:
 - normally distributed population with known variance.
 - normally distributed population with unknown variance.
 - not normally distributed population, but with a large enough sample.
- Accept or reject hypotheses for normally distributed population variances.

P-values

Before we get to the means and variances, we begin from our previous worksheet. The last two exercises asked you to compute a so called *P*-value. Let us see the reason behind investigating this quantity.

Definition 1 (P-values) *In hypothesis testing, the P-value is the largest probability that still leads to the null hypothesis being correct (failing to reject).*

We begin by comparing α to the *P*-value. The probability α is a **rigid, pre-specified limit** to the risk we are willing to take. The risk, of course, translates to $P(\text{reject } H_0 | H_0 \text{ is true})$. No matter how useful α is fails to reveal the whole picture of statistical hypothesis testing.

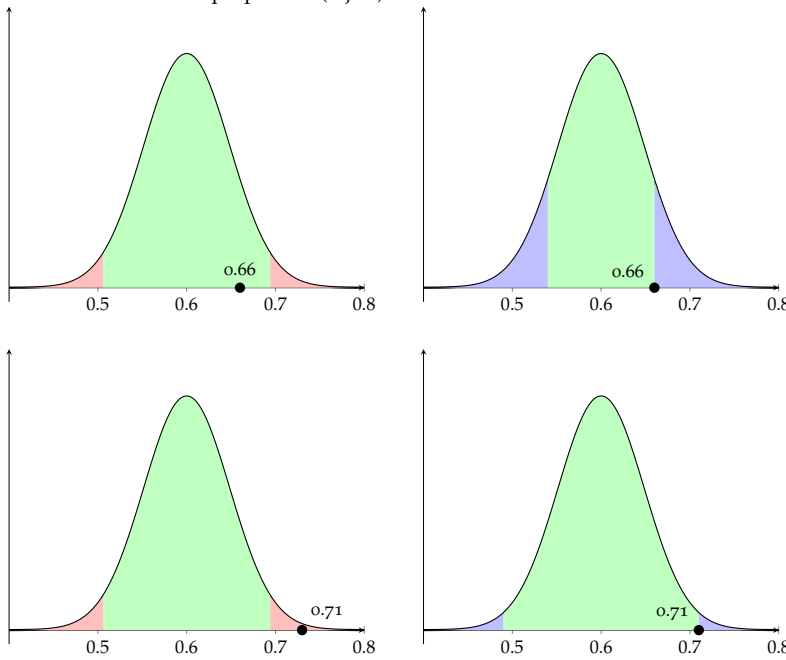
On the other hand, the *P*-value is an **observed significance level**, that depends on the observed (obtained) sample. As it is the largest probability that would still allow us to fail to reject H_0 , we immediately get the following as a consequence:

- if we accept the null hypothesis, then $\alpha \leq P\text{-value}$;
- if we reject the null hypothesis, then $\alpha > P\text{-value}$.

We show this consequence visually in Figure ??.

So, how to calculate a *P*-value? It depends on whether we are testing two sides or one side. Recall that for a given observed proportion

Figure 1: In the first pair, we have that the observed proportion is $\hat{p} = 0.66$. We then show α (as in the rejection areas, in red) and the observed P -value (in blue). In the second pair, we show the same values but now the observed proportion is $\hat{p} = 0.71$. Note how $P\text{-value} \geq \alpha$ in the first observed proportion (fail to reject) and $P\text{-value} < \alpha$ in the second observed proportion (reject).



\hat{p} , we may compute the test statistic $Z_0 = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$. Then, we have for the P -values: ¹

¹ See Worksheet 24-25 for the details.

- Two-sided hypothesis: $P = 2(1 - \Phi(|Z_0|))$
- One-sided (upper) hypothesis: $P = 1 - \Phi(Z_0)$
- One-sided (lower) hypothesis: $P = \Phi(Z_0)$

Based on our discussion here, we have **two ways to recommend rejection of a null hypothesis**:

1. Check whether the observed proportion \hat{p} or the Z_0 statistic fall in the rejection region.
2. Calculate the P -value and compare to α .

In summary, we have for proportion hypothesis testing:

Proportion hypothesis testing

Null hypothesis: Test statistic: Distribution:

$$H_0 : p = p_0. \quad Z_0 = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}. \quad Z_0 \sim \mathcal{N}(0, 1).$$

H_1	Rejection region	P -value
$p \neq p_0$	$ Z_0 > z_{\alpha/2}$	$2 \cdot (1 - \Phi(Z_0))$
$p > p_0$	$Z_0 > z_{\alpha}$	$1 - \Phi(Z_0)$
$p < p_0$	$Z_0 < -z_{\alpha}$	$\Phi(Z_0)$

Reject if Z_0 falls in the rejection region or if P -value $< \alpha$.

Polling for a law

We surveyed 100 people on whether they support a new proposed law that will be on the ballot. 58% said that they do. Our hypothesis is that our county is evenly divided and hence 50% actually do support the law. What is the observed P -value? Should we reject the hypothesis that there are 50% in support of the law when $\alpha = 0.05$?

We have that

$$Z_0 = \frac{0.58 - 0.5}{0.05} = 1.6.$$

Now, we calculate $\Phi(Z_0) = \Phi(1.6) = 0.9452$. Hence, we get that P -value $= 2 \cdot (1 - 0.9452) = 0.1096$. Because P -value $\geq \alpha$, we fail to reject the hypothesis.

Polling for a law

Assume that in a different county, we surveyed 100 people on whether they support the law: we now got that 38% said that they do. Our hypothesis is that our city is again that the county supports it by 50%; but now our alternative hypothesis is the lower side only (i.e., $H_1 : p < 0.5$). What is the observed P -value in this case? Should we reject the hypothesis that there are 50% in support of the law when $\alpha = 0.05$?

We have that

$$Z_0 = \frac{0.38 - 0.5}{0.05} = -2.4.$$

Now, we calculate $\Phi(Z_0) = \Phi(-2.4) = 1 - \Phi(2.4) = 0.0082$. This is also the P -value. Note that $P\text{-value} < \alpha$, and thus we should reject the hypothesis in favor of the alternative (that is, less than 50% support the law).

Hypothesis testing for means

We now do the same for means. Recall the three cases we are interested in!

1. normally distributed population with known variance σ^2 .
2. normally distributed population with unknown variance.
3. not normally distributed population, but we have a large enough sample.

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

Normally distributed population with known variance σ^2

Mean with known variance hypothesis testing

Null hypothesis:	Test statistic:	Distribution:
$H_0 : \mu = \mu_0.$	$Z_0 = \frac{\bar{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}}.$	$Z_0 \sim \mathcal{N}(0, 1).$

H_1	Rejection region	P-value
$\mu \neq \mu_0$	$ Z_0 > z_{\alpha/2}$	$2 \cdot (1 - \Phi(Z_0))$
$\mu > \mu_0$	$Z_0 > z_{\alpha}$	$1 - \Phi(Z_0)$
$\mu < \mu_0$	$Z_0 < -z_{\alpha}$	$\Phi(Z_0)$

Like earlier, we may reject if:

1. Check whether the observed sample average \bar{X} or the Z_0 statistic fall in the rejection region.
2. Calculate the P -value and compare to α .

A life expectancy example

We select a random sample of 100 recorded deaths in the city of Urbana. The sample average is 71.8 years old. Assuming that life expectancy is normally distributed with a (known) standard deviation of 9 years, can we claim that life expectancy in Urbana is 70 years or is it higher? Use $\alpha = 5\%$.

First, state the null and alternate hypotheses. In our case:

$$H_0 : \mu = 70 \quad H_1 : \mu > 70$$

Now, calculate $Z_0 = \frac{71.8 - 70}{9/\sqrt{100}} = 2$. Since it is one-sided, check $z_{\alpha} = z_{0.05} = 1.645$. We finally reject the hypothesis, as $Z_0 > z_{\alpha}$.

In this example, we could have built the upper confidence interval (for $\alpha = 5\%$) around the hypothesized mean as:

$$[0, U] = \left[0, \mu_0 + z_{\alpha} \frac{\sigma}{\sqrt{n}} \right] = [0, 71.48].$$

We can see that even doing it this way, we still note that the observed sample average ($\bar{X} = 71.8$) is outside the confidence interval, and hence we should reject the null hypothesis that life expectancy is at 70 years in favor of the alternative that it is higher than 70 years.

Normally distributed population with unknown variance σ^2

When the variance is unknown, we recall from the confidence interval discussion that the sampling distribution is no longer the normal one; instead we used the so-called Student's T distribution.

Mean with unknown variance hypothesis testing

Null hypothesis:	Test statistic:	Distribution:
$H_0 : \mu = \mu_0.$	$T_0 = \frac{\bar{X} - \mu_0}{\frac{s}{\sqrt{n}}}.$	$T_0 \sim T_{n-1}.$

H_1	Rejection region	P-value
$\mu \neq \mu_0$	$ T_0 > t_{\alpha/2, n-1}$	$2 \cdot (1 - T_{n-1}(T_0))$
$\mu > \mu_0$	$T_0 > t_{\alpha, n-1}$	$1 - T_{n-1}(T_0)$
$\mu < \mu_0$	$T_0 < -t_{\alpha, n-1}$	$T_{n-1}(T_0)$

A quick note about the notation. With lower case t we typically refer to the T distribution critical values (e.g., $t_{\alpha, n-1}$). On the other hand, with upper case T we typically refer to the cumulative distribution of the T distribution (e.g., $T_{n-1}(t) = P(T \leq t)$): for these values we would typically consult a cumulative distribution function for the T distribution table.

A life expectancy example

We select a random sample of 16 recorded deaths in the city of Urbana. The sample average is 71.8 years old and the sample standard deviation is 9 years. Assuming that life expectancy is normally distributed but with no known standard deviation, can we claim that life expectancy in Urbana is 70 years or is it higher? Use $\alpha = 5\%$.

We have the same null and alternate hypotheses as in the previous case, because it is again one-sided. However, now, we have a different test statistic:

$$H_0 : \mu = 70 \quad H_1 : \mu > 70$$

$$T_0 = \frac{71.8 - 70}{9/\sqrt{16}} = 0.8.$$

The corresponding critical value we want to find is $t_{\alpha, n-1} = t_{0.05, 15} = 1.753$. Due to that, we cannot reject the hypothesis, as $T_0 \leq t_{\alpha, n-1}$.

Much like what we did earlier, we will again build a confidence

interval around the unknown mean. We'd have:

$$[0, U] = \left[0, \mu_0 + t_{\alpha, n-1} \frac{s}{\sqrt{n}} \right] = [0, 73.94].$$

Note how the observed sample average is $\bar{X} = 71.8$ years and it totally is part of the confidence interval. This is another way we could have deduced that we cannot reject the hypothesis.

Not normally distributed population

Not normally distributed population mean hypothesis testing

Null hypothesis:	Test statistic:	Distribution:
$H_0 : \mu = \mu_0.$	$Z_0 = \frac{\bar{X} - \mu_0}{\frac{s}{\sqrt{n}}}.$	$Z_0 \sim \mathcal{N}(0, 1).$

H_1	Rejection region	P-value
$\mu \neq \mu_0$	$ Z_0 > z_{\alpha/2}$	$2 \cdot (1 - \Phi(Z_0))$
$\mu > \mu_0$	$Z_0 > z_{\alpha}$	$1 - \Phi(Z_0)$
$\mu < \mu_0$	$Z_0 < -z_{\alpha}$	$\Phi(Z_0)$

The only difference from the first case? We need a bigger sample size (say, $n \geq 30$) and we do not necessarily need the variance. Instead, we may estimate it (if unknown) as the sample variance s^2 and use it instead.

Hypothesis testing for normally distributed population variances

We are ready to show the last hypothesis testing procedure for a single population! For the variance of a normally distributed population we may reject or fail to reject a hypothesis on its true value following (again) the same logic as for its confidence interval, which was based on the χ^2 distribution.

Normally distributed population variance hypothesis testing

Null hypothesis:	Test statistic:	Distribution:
$H_0 : \sigma^2 = \sigma_0^2.$	$\chi_0^2 = \frac{(n-1)s^2}{\sigma_0^2}.$	$\chi_0^2 \sim \chi_{n-1}^2.$

H_1	Rejection region	CI region
$\sigma^2 \neq \sigma_0$	$\chi_0^2 > \chi_{\alpha/2, n-1}^2$ $\chi_0^2 < \chi_{1-\alpha/2, n-1}^2$	$\left[\frac{(n-1)\sigma_0^2}{\chi_{\alpha/2, n-1}^2}, \frac{(n-1)\sigma_0^2}{\chi_{1-\alpha/2, n-1}^2} \right]$
$\sigma^2 > \sigma_0$	$\chi_0^2 > \chi_{\alpha, n-1}^2$	$\left[\frac{(n-1)\sigma_0^2}{\chi_{\alpha, n-1}^2}, +\infty \right)$
$\sigma^2 < \sigma_0$	$\chi_0^2 < \chi_{1-\alpha, n-1}^2$	$\left(-\infty, \frac{(n-1)\sigma_0^2}{\chi_{1-\alpha, n-1}^2} \right]$

Hence, we would reject the null hypothesis whenever the χ_0^2 statistic is inside the rejection region, or whenever the hypothesized σ_0^2 is outside the confidence interval region.

A life expectancy example

We select a random sample of 16 recorded deaths in the city of Urbana. The sample average is 71.8 years old and the sample standard deviation is 9 years. Assuming that life expectancy is normally distributed but with no known standard deviation, can we claim that the standard deviation is equal to 7 years? Or is it different than that? Use $\alpha = 5\%$.

We have:

$$H_0 : \sigma^2 = 49 \quad H_1 : \sigma^2 \neq 49$$

The test statistic is:

$$\chi_0^2 = \frac{15 \cdot 81}{49} = 24.796.$$

The corresponding critical value is $\chi_{\alpha/2, n-1}^2 = \chi_{0.025, 15}^2 = 27.488$ and $\chi_{1-\alpha/2, n-1}^2 = \chi_{0.975, 15}^2 = 6.262$. Hence, we cannot reject the hypothesis, as $\chi_{1-\alpha/2, n-1}^2 \leq \chi_0^2 \leq \chi_{\alpha/2, n-1}^2$.

Like we did earlier, note that building the confidence interval we would have gotten:

$$[L, U] = \left[\frac{15 \cdot 49}{27.488}, \frac{15 \cdot 49}{6.262} \right] = [26.74, 117.37],$$

which includes the sample variance $s^2 = 81$.