27 Hypothesis testing

27.1 Layout of a hypothesis test

Step 1: Identify null and alternative hypothesis

Our first step in every hypothesis testing question is to identify the **null hypothesis** (H_0) and **alternative hypothesis** (H_1).

To do that, we must first identify the (claim for the) **population mean** μ. For example, a question may state that "Tau wishes to test a claim that the population mean height of plants is 28.5 cm". In that case, our null hypothesis is $H_0 : \mu = 28.5$.

Our null hypothesis in our syllabus will always be of the form $\mu = \mu_0$, where μ_0 is the (claimed) population mean to be determined from the question.

Our next step is to identify the alternative hypothesis, which can take three forms: the left-tail test, the two-tailed test or the right-tailed test. The following table summarizes some common key words associated with each of the H_1 options:

$H_0 : \mu = \mu_0$ $H_1 : \mu < \mu_0$	$H_0 : \mu = \mu_0$ $H_1 : \mu \neq \mu_0$	$H_0 : \mu = \mu_0$ $H_1 : \mu > \mu_0$
less than decreased smaller	is (equal to) is not (equal to) changed	more than increased larger
overstated		understated
at least not less than		at most not more than

For example, a question may state that "Tau wishes to test a claim that the population mean height of plants has increased". In that case, our null hypothesis and alternative hypothesis will be $H_0 : \mu = 28.5, \quad H_1 : \mu > 28.5$.

Step 2: Calculate unbiased estimates if necessary

For question with summarized data such as $\sum x = 5928, \quad \sum x^2 = 587000$ and a sample of size 60, we will often calculate the **sample mean** (which is also the unbiased estimate of population mean) $\overline{x}$.

$$\overline{x} = \frac{\sum x}{n}$$

Some questions will directly give us the **population variance** σ^2. In that case we will just use that value for our subsequent calculations. For questions where that is not provided, we will then calculate the **unbiased estimate of population variance** s^2 with our formulas:

$$\begin{aligned} s^2 &= \frac{n}{n-1}\,(\text{sample variance}) \\ &= \frac{n}{n-1}\left(\frac{\sum(x-\overline{x})^2}{n}\right) \\ &= \frac{1}{n-1}\left(\sum x^2 - \frac{(\sum x)^2}{n}\right) \end{aligned}$$

Step 3: Present the distribution of the sample mean

Step 3 in our hypothesis testing template is a presentation step, where we write down the distribution in preparation to use our GC later. Recall from the sampling distribution topic that

$$\overline{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

We can present this, along with the value of μ from step 1 and σ^2 (or s^2 if population variance is unknown). To make the template more standardized across different question types, however, we recommend standardizing our distribution instead.

$$Z = \frac{\overline{X} - \mu}{\sqrt{\frac{\sigma^2}{n}}} \sim N(0,1)$$

If X is not known to be of normal distribution, we may also need to invoke CLT to justify the use of a normal distribution, if n is large. Refer to template A in the next section to see this in action.

Step 4: Key into GC to get p-value

We now have all the information we need written down and it's now time to transfer into our GC. This is under $\boxed{\text{STATS}} \to \boxed{\text{TESTS}} \to \boxed{\text{Z-TEST...}}$.

We will use "STATS" for the "INPT" option, μ_0 from our step 1 in H_1, σ as our population standard deviation (using s if necessary), sample mean $\overline{x}$, sample size n and the type of test (left/right/two-tailed).

We then key the calculate button and then copy down the p-value that is shown to us.

Step 5: Conclude

We compare the p-value with the **level of significance** $\alpha\%$ given to us.

If $p \leq \alpha\%$, we **reject** H_0 and conclude that there is **sufficient** evidence that H_0 is false and H_1 is true (phrased in context of the question).

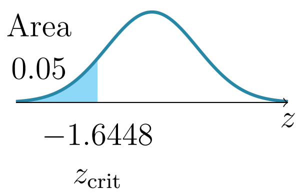
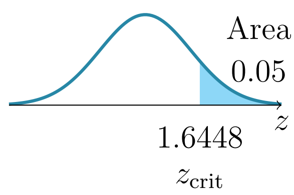
If $p > \alpha\%$, we **do not reject** H_0 and conclude that there is **insufficient** evidence to conclude whether H_0/H_1 is true or false (phrased in context of the question).

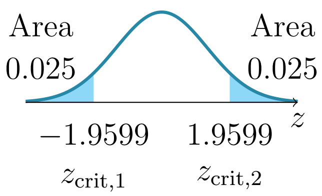
Variations: using invNorm

For template C, we will work backwards where we are given a conclusion and then asked for the range of values of $\mu, \overline{x}, \sigma, n$, etc. In such a case we will use invNorm on the level of significance to get the **critical value**.

In the following table/example we have used a level of significance of 5% to obtain the critical value of ± 1.6448 or ± 1.9599 from our GC via the invNorm calculation (depending on the type of test).

For 5% level of significance,

<table>
<tr><td>Tail(s)</td><td>Left</td><td>Right</td></tr>
<tr><td>Picture</td><td>
</td><td>
</td></tr>
<tr><td>Reject H_0</td><td>$\dfrac{\overline{x}-\mu}{\frac{\sigma}{\sqrt{n}}} \leq z_{\text{crit}}$</td><td>$\dfrac{\overline{x}-\mu}{\frac{\sigma}{\sqrt{n}}} \geq z_{\text{crit}}$</td></tr>
<tr><td>Do not reject H_0</td><td>$\dfrac{\overline{x}-\mu}{\frac{\sigma}{\sqrt{n}}} > z_{\text{crit}}$</td><td>$\dfrac{\overline{x}-\mu}{\frac{\sigma}{\sqrt{n}}} < z_{\text{crit}}$</td></tr>
</table>

<table>
<tr><td>Tail(s)</td><td>Two-tailed</td></tr>
<tr><td>Picture</td><td>
</td></tr>
<tr><td>Reject H_0</td><td>$\dfrac{\overline{x}-\mu}{\frac{\sigma}{\sqrt{n}}} \leq z_{\text{crit},1}$ or $\dfrac{\overline{x}-\mu}{\frac{\sigma}{\sqrt{n}}} \geq z_{\text{crit},2}$</td></tr>
<tr><td>Do not reject H_0</td><td>$z_{\text{crit},1} < \dfrac{\overline{x}-\mu}{\frac{\sigma}{\sqrt{n}}} < z_{\text{crit},2}$</td></tr>
</table>

27.2 Template A1: Carry out a test (rejection)

Example 1. A baker states that a certain type of small cake that he produces has a mean mass of 100 grams. A food inspector wishes to investigate whether the mean mass of these cakes is actually less than 100 grams. The food inspector selects a random sample of 60 cakes of this type. The masses, x, in grams, are summarised by

$$\sum x = 5928, \qquad \sum x^2 = 587\,000.$$

Determine the conclusion the food inspector should reach if she carries out a test at the 5% significance level.

Solution 1.

$H_0 : \mu = 100$
$H_1 : \mu < 100$

Unbiased estimate of population mean $\overline{x} = \frac{5928}{60} = 98.8$.
Unbiased estimate of population variance
$s^2 = \dfrac{1}{59}\left(587000 - \dfrac{5928^2}{60}\right) = 22.264.$

Under H_0,
Test statistic $Z = \dfrac{\overline{X} - 100}{\sqrt{\frac{22.264}{60}}} \sim N(0,1)$ approx by CLT since $n = 60$ is large.

We perform a hypothesis test at 5% level of significance.

Using GC, using $\mu_0 = 100, `\sigma' = \sqrt{22.264}, \overline{x} = 98.8, n = 60, \mu < \mu_0$,
p-value $= 0.024421 \leq 0.05 \quad \Rightarrow \quad H_0$ rejected.

Hence there is sufficient evidence at the 5% level of significance for the food inspector to conclude that the mean mass of these cakes is actually less than 100 grams. ■

27.3 Template A2: Non-rejection

Example 2. A baker states that a certain type of small cake that he produces has a mean mass of 100 grams. A food inspector wishes to investigate whether the the baker is wrong about the mean mass of the cakes. The food inspector selects a random sample of 60 cakes of this type. The masses, x, in grams, are summarised by

$$\sum x = 5928.$$

It is known that the mass of cakes is normally distributed and the population variance is 19.0 grams.

Determine the conclusion the food inspector should reach if she carries out a test at the 2% significance level.

Solution 2.

$H_0 : \mu = 100$
$H_1 : \mu \neq 100$

Unbiased estimate of population mean $\overline{x} = \frac{5928}{60} = 98.8$.
Population variance $\sigma^2 = 19.0$

Under H_0,
Test statistic $Z = \dfrac{\overline{X} - 100}{\sqrt{\frac{19.0}{60}}} \sim N(0, 1)$.

We perform a hypothesis test at 2% level of significance.

Using GC, using $\mu_0 = 100, \sigma = \sqrt{19.0}, \overline{x} = 98.8, n = 60, \mu \neq \mu_0$,
p-value $= 0.032969 > 0.02 \quad \Rightarrow \quad H_0$ not rejected.

Hence there is insufficient evidence at the 2% level of significance for the food inspector to conclude whether the baker is wrong about the mean mass of the cakes. ∎

27.4 Template B: Unknown level of significance

Example 3. A baker states that a certain type of small cake that he produces has a mean mass of 100 grams. A food inspector wishes to investigate whether the the mean mass of cakes has increased. The food inspector selects a random sample of 60 cakes of this type and found that the mean mass of his sample is 101.1 grams.

It is known that the mass of cakes is normally distributed and the population variance is 19.0 grams.

Given that the food inspector concludes that the mean mass of cakes has increased when he carries out a hypothesis test at $\alpha\%$ significance level, find the range of possible values of α.

Solution 3.

$H_0 : \mu = 100$
$H_1 : \mu > 100$

Unbiased estimate of population mean $\overline{x} = 102.1$.
Population variance $\sigma^2 = 19.0$

Under H_0,
Test statistic $Z = \dfrac{\overline{X} - 100}{\sqrt{\frac{19.0}{60}}} \sim N(0, 1)$.

We perform a hypothesis test at $\alpha\%$ level of significance.

Using GC, using $\mu_0 = 100, \sigma = \sqrt{19.0}, \overline{x} = 101.1, n = 60, \mu > \mu_0$,
p-value $= 0.025306$.

Since the food inspector concludes that the mean mass of cakes has increased, H_0 is rejected so
p-value $\leq \alpha\%$
$0.025306 \leq \frac{\alpha}{100}$
$\alpha \geq 2.53$. ■

27.5 Template C: Unknown $\overline{x}, \sigma, n, \mu_0$

Example 4. A baker states that a certain type of small cake that he produces has a mean mass of 100 grams. A food inspector wishes to investigate whether the the mean mass of cakes has changed. The food inspector selects a random sample of 60 cakes of this type.

It is known that the population variance is 19.0 grams.

Given that a hypothesis test indicates that there is sufficient evidence to reject the baker's statement at the 3% level of significance, find the set of values for the mean mass of this sample, correct to 2 decimal places.

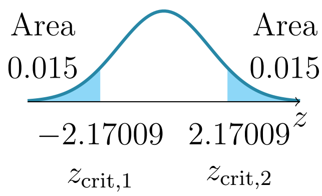
Solution 4.

$H_0 : \mu = 100$
$H_1 : \mu \neq 100$

Population variance $\sigma^2 = 19.0$

Under H_0,
Test statistic $Z = \dfrac{\overline{X} - 100}{\sqrt{\frac{19.0}{60}}} \sim N(0, 1)$ approx by CLT since $n = 60$ is large.

We perform a hypothesis test at 3% level of significance.



Using GC, critical values = ± 2.17009.

Since there is sufficient evidence to reject H_0,

$$\frac{\overline{x} - 100}{\sqrt{\frac{19.0}{60}}} \leq -2.17009 \quad \text{or} \quad \frac{\overline{x} - 100}{\sqrt{\frac{19.0}{60}}} \geq 2.17009$$

$$\overline{x} \leq 98.78 \quad \text{or} \quad \overline{x} \geq 101.22$$

Hence the set of values of the mean mass is $(-\infty, 98.78] \cup [101.22, \infty)$. ■

Remark: the same method can be done to handle left/right tailed tests, or if μ, σ or n is unknown.