COGS2020

TUTORIAL 8: Z TESTS AND T TESTS

Welcome back :D

- Tip for content to come – all the other tests you will learn (e.g. z test, t test, ANOVA, etc.) are variations on what we've learned so far
- If you can understand the framework of null hypothesis testing, you will see that all the tests are very similar!
- So far the R code you have learned calculates p values, etc. manually – useful for understanding how null hypothesis testing works (also problem sets!)
- You will now be introduced to code that runs the whole test for you

NHST Summary Recipe: Arbitrary Parameter

1. Specify the null and alternative hypotheses ($H0$ and $H1$) in terms of a population parameter θ.
2. Specify the type I error rate – denoted by the symbol α – you are willing to tolerate.
3. Specify the sample statistic $\theta\hat{}$ that you will use to estimate the population parameter θ in step 1 and state how it is distributed under the assumption that $H0$ is true.
4. Obtain a random sample and use it to compute the sample statistic from step 3. Call this value $\theta\hat{}obs$
5. If $\theta\hat{}obs$ or a **more extreme outcome** is very unlikely to occur under the assumption that $H0$ is true, then reject H.

What is θ and $\hat{\theta}$?

- θ is a population parameter that you are interested in estimating. E.g., in the case of a Normal test, θ is the population mean μ.

- $\hat{\theta}$ is a sample statistic that you use to estimate θ. E.g., in the case of a Normal test, $\hat{\theta}$ is the sample mean $\bar{x}$.

- The reason we bother to write the steps in terms of θ and $\hat{\theta}$ is that the steps are general and can be applied to any hypothesis test.

NHST Summary Recipe: Arbitrary Parameter

1. Specify the null and alternative hypotheses (H0 and H1) in terms of a population parameter θ.

Example:
H_0: θ = 0 (no difference)
H_1: θ ≠ 0 (difference exists)

NHST Summary Recipe: Arbitrary Parameter

2. Specify the type I error rate – denoted by the symbol α – you are willing to tolerate.

$$\alpha = 0.05$$

Content:

- **α (Alpha):** Probability of rejecting H_0 when H_0 is actually true.
- **Common choices:** 0.05, 0.01, or 0.10.
- **Interpretation:** If $\alpha = 0.05$, there is a 5% risk of incorrectly rejecting H_0.

NHST Summary Recipe: Arbitrary Parameter

3. Specify the sample statistic $\hat{\theta}$ that you will use to estimate the population parameter θ in step 1 and state how it is distributed under the assumption that H0 is true.

- **Sample Statistic ($\hat{\theta}$):** Estimate of the population mean, here denoted as $\hat{\mu}=\bar{x}$ (the sample mean).
- **Distribution under H_0:**

$$\widehat{\mu} = \overline{x} \sim \mathcal{N}(\mu_0, \frac{\sigma}{\sqrt{n}})$$

- where:
  - μ_0 is the population mean under the null hypothesis,
  - σ is the known population standard deviation,
  - n is the sample size.
- **Explanation:**
  Under H_0, the sample mean follows a normal distribution centered at μ_0 with standard error.

NHST Summary Recipe: Arbitrary Parameter

4. Obtain a random sample and use it to compute the sample statistic from step 3. Call this value θˆobs (μˆobs).

- Collect random sample → $x_{\text{obs}} = \{x_1, x_2, \ldots, x_n\}$
- Compute the Sample mean → $\widehat{\mu}_{\text{obs}} = \overline{x}_{\text{obs}} = \frac{1}{n}\sum_{i=1}^{n} x_i$

Meaning:

- μ^obs is the observed sample mean calculated from your data.
- This value will be compared to the expected distribution under H0 to decide whether to reject H0.

NHST Summary Recipe: Arbitrary Parameter

5. If θ ˆobs or a **more extreme outcome** is very unlikely to occur under the assumption that H0 is true, then reject H.

$$P(\overline{X} \leq \overline{x}_{\text{obs}} | H_0) < \alpha \rightarrow \text{reject } H_0$$
$$\text{Fail to reject } H_0 \text{ otherwise}$$

Symbol	Meaning	Type	Example
$\bar{X}$	Sample mean as a random variable	Theoretical	"What could happen"
$\bar{x}_{\text{obs}}$	Observed sample mean	Actual value	"What actually happened in my sample"

Normal Distribution

- A normal distribution is a symmetric, bell-shaped curve.
- It is fully described by:
  - **Mean** (μ): Center of the distribution
  - **Standard Deviation** (σ): Spread or width of the distribution

$X \sim N(\mu_X, \sigma)$

How do we create a hypothetical null model?

1. **Define the null and alterative** (what is considered no effect in the population?). (H0 and H1 in terms of the population parameter θ)
2. **What does the null look like in the population** (what would the population look like **if null is true** – define the key moments)

   X ~ *N*(parameters if null was true)

3. **How are we estimating the population mean (θ-hat)?** (eg. sample mean)
4. What is the **sampling distribution of this estimate** (given null is true)? **Building the sampling distribution null model**..

   x̄ ~ *distribution*(sampling dist parameters is null was true)

5. Do we have enough information about the population to create this null model distribution? If not, what alternative distribution could we use?
6. What is the **sample/test statistic (θ-hat$_{obs}$)** based on the distribution you use to model the null?
7. Set up your rejection zone(s). What is your alpha? What are/is your critical value(s)?
8. **Where does sample/test statistic lie when put into your null model (is it in the rejection zone or not)**? Is it likely to occur or unlikely?
   - If in rejection zone – it is unlikely we got our test statistic given this model, therefore we reject this null model
   - If not in rejection zone – it is likely we got our test statistic given this null model, therefore we do NOT reject this null model

Normal test broken down

1. **Define the null and alterative**

   H0: $\mu_X = c$

   H1: $\mu_X \neq c$, or $\mu_X > c$, or, $\mu_X < c$ (depends on your RQ, or what's given to you in the question)

2. **What does the null look like in the population** (what would the population look like **if null is true** – define the key moments)

   $X \sim N(\mu_X = c, \sigma = v)$

3. **How are we estimating the population mean (θ-hat)?**

   $\hat{\mu}_X = \bar{x}$ (little x bar is sample mean)

4. What is the **sampling distribution of this estimate** (given null is true)? **Building the sampling distribution null model.**

   $\bar{X} \sim N(\mu_{\bar{x}} = c, \sigma_{\bar{x}} = \sigma/\text{sqrt}(n))$

5. Do we have enough information about the population to create this null model distribution? If not, what alternative distribution could we use?

   Yes – we have all the key parameters needed to build our null model. We know the population mean (if null true) and also the population variance

6. What is the **test statistic ($\theta\text{-hat}_{obs}$)** based on the distribution you use to model the null?

   Given that we are using the normal distribution to model our population and make our sampling distribution (given null true), our test statistics will be the sample mean.

7. Set up your rejection zone(s).

   What is your alpha? What are/is your critical value(s)?

8. **Where does test statistic lie when put into your null model (is it in the rejection zone)**? Is it likely to occur or unlikely?
   - If in rejection zone – it is unlikely we got our test statistic given this model, therefore we reject this null model
   - If not in rejection zone – it is likely we got our test statistic given this null model, therefore we do NOT reject this null model

Normal test - practice

Let's say that based on previous data, on average, the temperature in Sydney during autumn is around **23°C**, with a standard deviation of **5°C**.

- I have recorded the temperature each day for 90 days this year's autumn.
- In the sample I collected, the sample mean was calculated to be 26°C and the sample standard deviation was 2°C.
- Research Question: Is the temperature this year's autumn significantly hotter than it has been previous years?

Work through each of the steps from the previous slide, and conduct a normal test. Highly recommend doing all the steps by hand on paper or on your devices. To calculate the exact p value – you can use R afterwards.

Normal test on temperature

1. **Define the null and alterative**

   H0: $\mu_X = 23$

   H1: $\mu_X > 23$

2. What does the null look like in the population (what would the population look like **if null is true** – define the key moments). Draw out the distribution and label the key moments.

   $X \sim N(\mu_X = 23, \sigma = 5)$

3. **How are we estimating the population mean (θ-hat)?**

   $\hat{\mu}_X = \bar{x} = 26$

4. What is the **sampling distribution of this estimate** (given null is true)? **Building the sampling distribution null model.** Draw out the distribution and label the key moments.

   $\bar{X} \sim N(\mu_{\bar{x}} = 23, \sigma_{\bar{x}} = 5 /\text{sqrt}(90))$

5. Do we have enough information about the population to create this null model distribution? If not, what alternative distribution could we use? Yes – we have both the key parameters to make our null model.

6. What is the **sample/test statistic (θ-hat$_{obs}$)** based on the distribution you use to model the null? $\bar{x} = 26$

7. Set up your rejection zone(s). Draw out the distribution, label the key moments and also the rejection zone.

   What is your alpha? What are/is your critical value(s)?

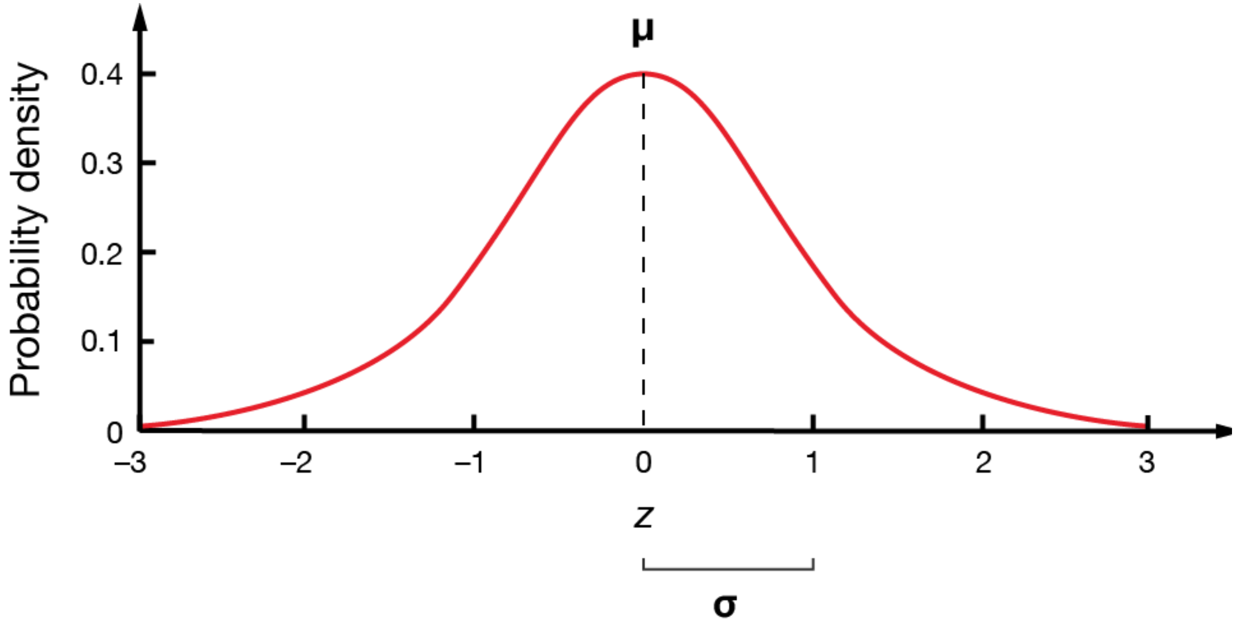
8. **Where does sample/test statistic lie when put into your null model (is it in the rejection zone)**? Is it likely to occur or unlikely? Draw out where your test statistic lies in your null model. Also label what areas/points represent the alpha, p value and critical value.

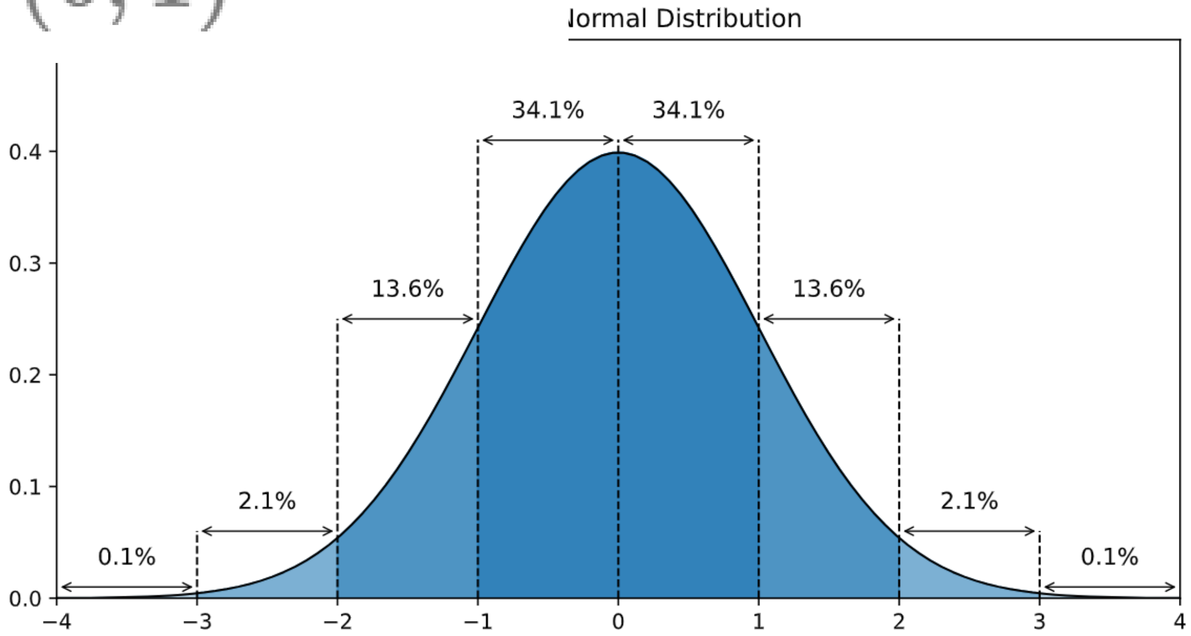
   - If in rejection zone – it is unlikely we got our test statistic given this model, therefore we reject this null model
   - If not in rejection zone – it is likely we got our test statistic given this null model, therefore we do NOT reject this null model

Z distribution

- The standardised version of a normal distribution
- This distribution consists of z scores – z scores are made by converting our original data via a formula

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

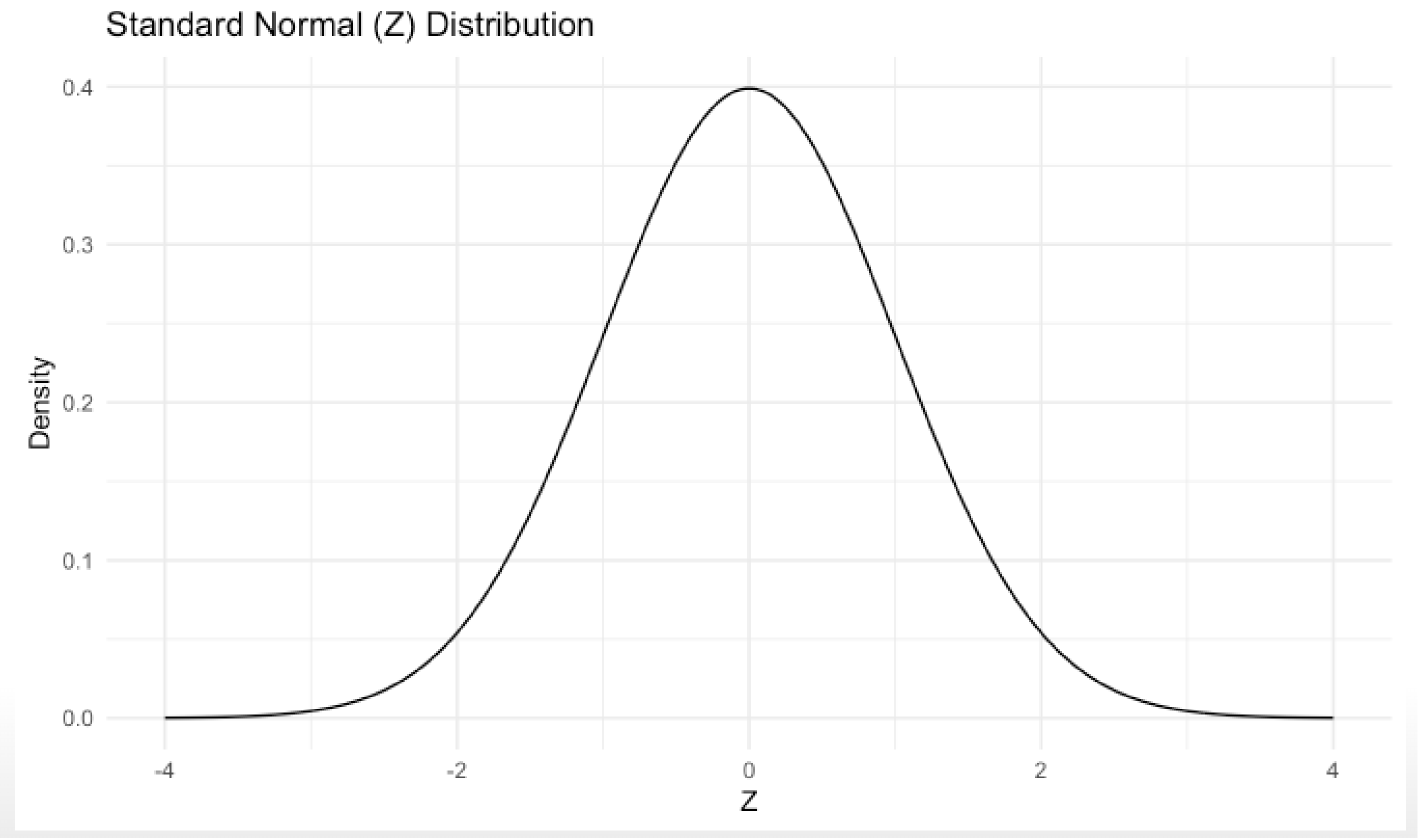




Characteristics of a Standard Normal (Z) Distribution

- **Standard Normal Distribution** has:
  - Mean = 0
  - Standard Deviation = 1

•A **Z-score** tells you how many standard deviations a value is from the mean.
•Positive Z: Above the mean.
•Negative Z: Below the mean.
Example:
A Z-score of 2.0 means the value is **2 standard deviations above** the mean.



How do we convert normal to standard normal (z scores)?

- We convert each data point into a z score via this formula:

$$Z = \frac{X - \mu_X}{\sigma_X}$$

Difference between your data point and the population mean

Population standard deviation

In the context of sample means:

$$z_{obs} = \frac{\bar{x}_{obs} - \mu_{\overline{X}}}{\sigma_{\overline{X}}}$$

Difference between your data point and the sampling distribution mean (of your null model)

Sampling distribution standard deviation

Remember from week 6?

$$\mu_{\overline{X}} = \mu_X$$

$$\sigma^2_{\overline{X}} = \frac{\sigma^2_X}{n}$$

Example: Turning Normal to Standard Normal

To **standardize** any normal variable X:

Step 1: Start with a Normal Distribution

- X∼N(μX,σX)
  - X is normally distributed with mean μX and standard deviation σX.
  - Example: Heights are normally distributed with mean μX=170 and σX=10 cm.

Step 2: Apply the Z Transformation

•Suppose someone is **185 cm** tall.
•Find their Z-score:

$$Z = \frac{185 - 170}{10} = \frac{15}{10} = 1.5$$

Step 3: Result: A Standard Normal Distribution

•After this transformation:
•Z∼N(0,1)
•Meaning:

- Mean = 0
- Standard Deviation = 1

$$X \sim \mathcal{N}(\mu_X, \sigma_X)$$

$$Z = \frac{X - \mu_X}{\sigma_X}$$

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

Z tests

- Instead of building the null model around the population directly, we “standardise” and “rescale” all the values to a standard normal curve (also called standard bell curve or z distribution)
- Why z tests when we have normal tests?
  - Back in the day we had no R or other fancy computer programs. P values were calculated by hand – had to look up corresponding p values to z scores in a huge table
  - Standardisation is good when you want to compare across data sets and studies

Z test practice

- Using the same temperature practice question from before, follow through the steps and conduct a z test.
- Hint: You will be using a **z distribution (*standard* normal distribution) to model your null**, NOT a normal distribution..

How will this change some of the steps?

Are your results any different?

Z test - practice

Let’s say that based on previous data, on average, the temperature in Sydney during autumn is around **23°C**, with a standard deviation of **5°C**.

- I have recorded the temperature each day for 90 days this year’s autumn.
- In the sample I collected, the sample mean was calculated to be 26°C and the sample standard deviation was 2°C.
- Research Question: Is the temperature this year’s autumn significantly hotter than it has been previous years?

Work through each of the steps from the previous slide, and conduct **a z test**. Highly recommend doing all the steps by hand on paper or on your devices. To calculate the exact p value – you can use R afterwards.

Z test on temperature, pt. 1

1. **Define the null and alterative**

    H0: $\mu_X = 23$

    H1: $\mu_X > 23$

2. What does the null look like in the population (what would the population look like **if null is true** – define the key moments). Draw out the distribution and label the key moments.

    $X \sim N(\mu_X = 23, \sigma = 5)$

3. **How are we estimating the population mean (θ-hat)?**

    $\hat{\mu}_X = \bar{x} = 26$

Note: everything is the exact same as the normal test so far

Z test on temperature, pt. 2

4. What is the **sampling distribution of this estimate** (given null is true)? **Building the sampling distribution null model.** Draw out the distribution and label the key moments.

$\bar{X} \sim N(\mu_{\bar{x}} = 23, \sigma_{\bar{x}} = 5 / \mathrm{sqrt}(90))$

BUT, the question asks us to use a z distribution. So, the sampling distribution in this case would be

$Z \sim N(\mu = 0, \sigma = 1)$

5. Do we have enough information about the population to create this null model distribution? If not, what alternative distribution could we use?

Yes – we have both the key parameters to make our null model. We also have the information we need to convert our data into z scores.

6. What is the **sample/test statistic (θ-hat$_{obs}$)** based on the distribution you use to model the null?

$$z_{obs} = \frac{\bar{x}_{obs} - \mu_{\overline{X}}}{\sigma_{\overline{X}}} \qquad \frac{26 - 23}{\frac{5}{\sqrt{9}}}$$

7. Set up your rejection zone(s). Draw out the distribution, label the key moments and also the rejection zone.

What is your alpha? What are/is your critical value(s)?

8. **Where does sample/test statistic lie when put into your null model (is it in the rejection zone)**? Is it likely to occur or unlikely? Draw out where your test statistic lies in your null model. Also label what areas/points represent the alpha, p value and critical value.

- If in rejection zone – it is unlikely we got our test statistic given this model, therefore we reject this null model
- If not in rejection zone – it is likely we got our test statistic given this null model, therefore we do NOT reject this null model

Z test vs normal test summary

- Z test is the standardised version of a normal test
- Results should be the same – the p value you calculate in a normal test and a z test should be identical
  - Why? Normal and Z distributions are defined by the exact same parameters (population mu and sigma) - this means that when you calculate "area under the curve" for your p value, that value should be the same.
- Z tests are great when you want to compare results across multiple experiments – it puts the data on the same “scale” and allows you to compare fairly

What if we don't know σX?

In real-world data:

- The population standard deviation σ_X is usually **unknown**.
- We **estimate** it from the sample → call this estimate s_X.

- When we use s_X instead of σX:
- The standardized value no longer follows a perfect normal distribution.

Moving from Z to t: Why Use t-distributions?

- If σ (population standard deviation) is unknown, estimate it using sample data.
- Replace σ with sample standard deviation σ^X.
- This turns the Z-score into a **t-score**:

$$t = \frac{X - \mu_X}{\hat{\sigma}_X} \quad \longrightarrow \quad t = \frac{\bar{x}_{\text{obs}} - \mu_X}{s_X / \sqrt{n}}$$

•This follows a **t-distribution**, not a normal distribution:

$$t \sim t(\text{df})$$

Degrees of freedom (for one sample and paired t test):

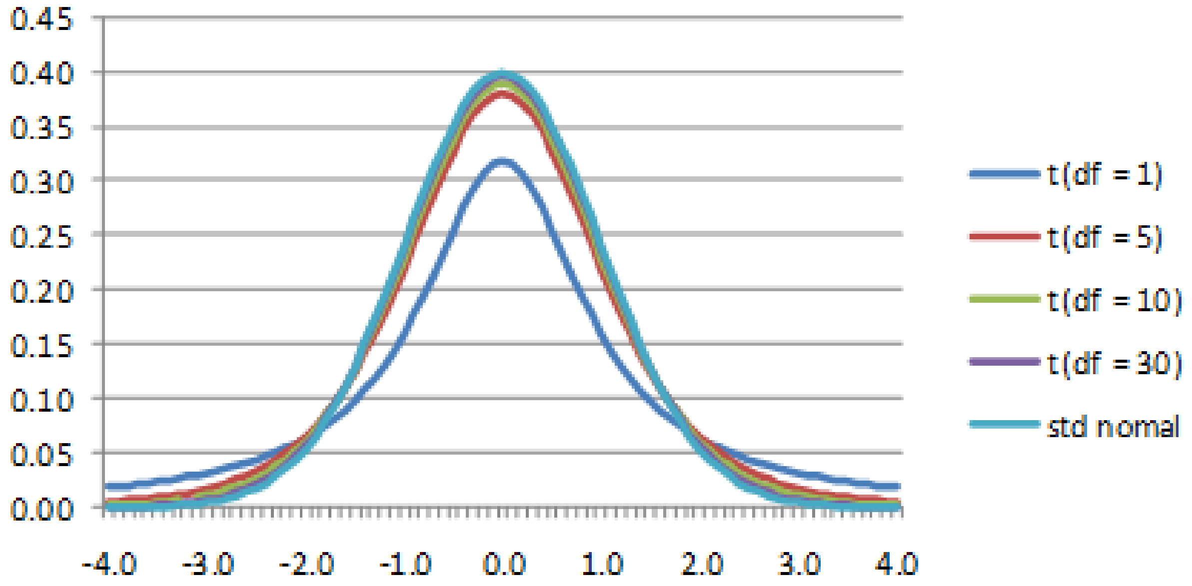
•df = n−1

where n = sample size.

T distribution

- A normal (ish) distribution when we don't know population variance
- T distribution are fully described by degrees of freedom (df) – df are generally calculated with sample size, but specific calculation depends on the type of t test you are running

$$t \sim t(df)$$



As **degrees of freedom ↑** , the t distribution peak becomes **taller** and its **tails** become **thinner**

Why Does This Matter?

Estimating σ_X introduces extra uncertainty.

The **t-distribution**:
- Is centered at 0 (like normal).
- Has **fatter tails** (more spread) — accounts for the extra uncertainty.

As sample size increases ($n \to \infty$):
- The t-distribution approaches the normal distribution.
- Why? T distributions are defined by degrees of freedom – degrees of freedom are calculated using sample size. So if n increases, degrees of freedom will increase, leading the shape of the t distribution to approximate a normal distribution

Example: Turning a Normal into a t-distribution

- You collect a small sample (n=9) of exam scores from students.
- You want to test if their average score differs from a known population mean (μ_X=75).
- However, you **don't know the population standard deviation** σ_X .

Sample Data:

- Sample mean: $\bar{x}$obs=78
- Sample standard deviation: s_X =5
- Sample size: n=9

Step 1: Compute the t-statistic

- Formula: $$t = \frac{\bar{x}_{\text{obs}} - \mu_X}{s_X/\sqrt{n}}$$

Plug in the numbers: $$t = \frac{78 - 75}{5/\sqrt{9}} = \frac{3}{5/3} = \frac{3}{1.6667} \approx 1.8$$

Step 2: What does this mean?

- The sample mean is about **1.8 estimated standard errors** above the hypothesized population mean.
- Because we used the **sample standard deviation** s_X (instead of the true σ_X), this statistic follows a **t-distribution — not** a normal distribution.

Creating a null model using the t distribution

Let's say that based on previous data, on average, the temperature in Sydney during autumn is around **23°C.** The variance is unknown.

- I have recorded the temperature each day for 90 days this year's autumn.
- In the sample I collected, the sample mean was calculated to be 26°C and the sample standard deviation was 2°C.
- Research Question: Is the temperature this year's autumn significantly hotter than it has been previous years?

Work through each of the steps from the previous slide, and conduct an appropriate statistical test. Highly recommend doing all the steps by hand on paper or on your devices. To calculate the exact p value – you can use R afterwards.

Types of t tests

- **One sample t test** – a single sample is compared to a known population mean (e.g. the example you just completed)
- **Independent samples t test** (or 2 sample t test) – compares the means of **two *independent* groups**
- **Paired t test** (repeated samples t test) – compares the means of two ***related*** groups, usually measuring the same subjects at two different points in time (before and after)

- I recommend in your own time, fully go through the null model steps for each of these tests to test your knowledge ☺

Independent Samples T test

- The independent samples t-test is used to compare the means of two random variabels that are independent of each other (i.e., samples come from different populations).
- Let X and Y be two **independent** random variables with means μ_X and μ_Y, respectively.
- An independent samples t-test is used to test the following hypotheses:

$$H_0 : \mu_X - \mu_Y = 0$$
$$H_1 : \mu_X - \mu_Y > 0$$

Independent Samples T test

- Test the hypothesis that the means of two random variables are equal.

```
x_obs <- rnorm(10, mean=10, sd=2)
y_obs <- rnorm(10, mean=12, sd=2)
res <- t.test(x=x_obs,
              y=y_obs,
              alternative="two.sided",
              mu=0,
              paired=F, # important to set this to False
              var.equal=T, # This flips between the two s_pooled caluclа
              conf.level=0.95)
```

Repeated Samples T test

- The repeated measures t-test is used to compare the means of two random variables that are drawn from the same population (e.g., samples come from the same subjects).
- Let X and Y be two random variables with means μ_X and μ_Y, respectively.
- An repeated measures t-test is used to test the following hypotheses:

$$H_0 : \mu_D = 0$$
$$H_1 : \mu_D > 0$$

Repeated Samples T test

```
res <- t.test(x=x_obs,
              y=y_obs,
              alternative="two.sided",
              mu=0,
              paired=T, ## this is important to set to True
              var.equal=T, ## this doesn't matter when paired=T
              conf.level=0.95)
```

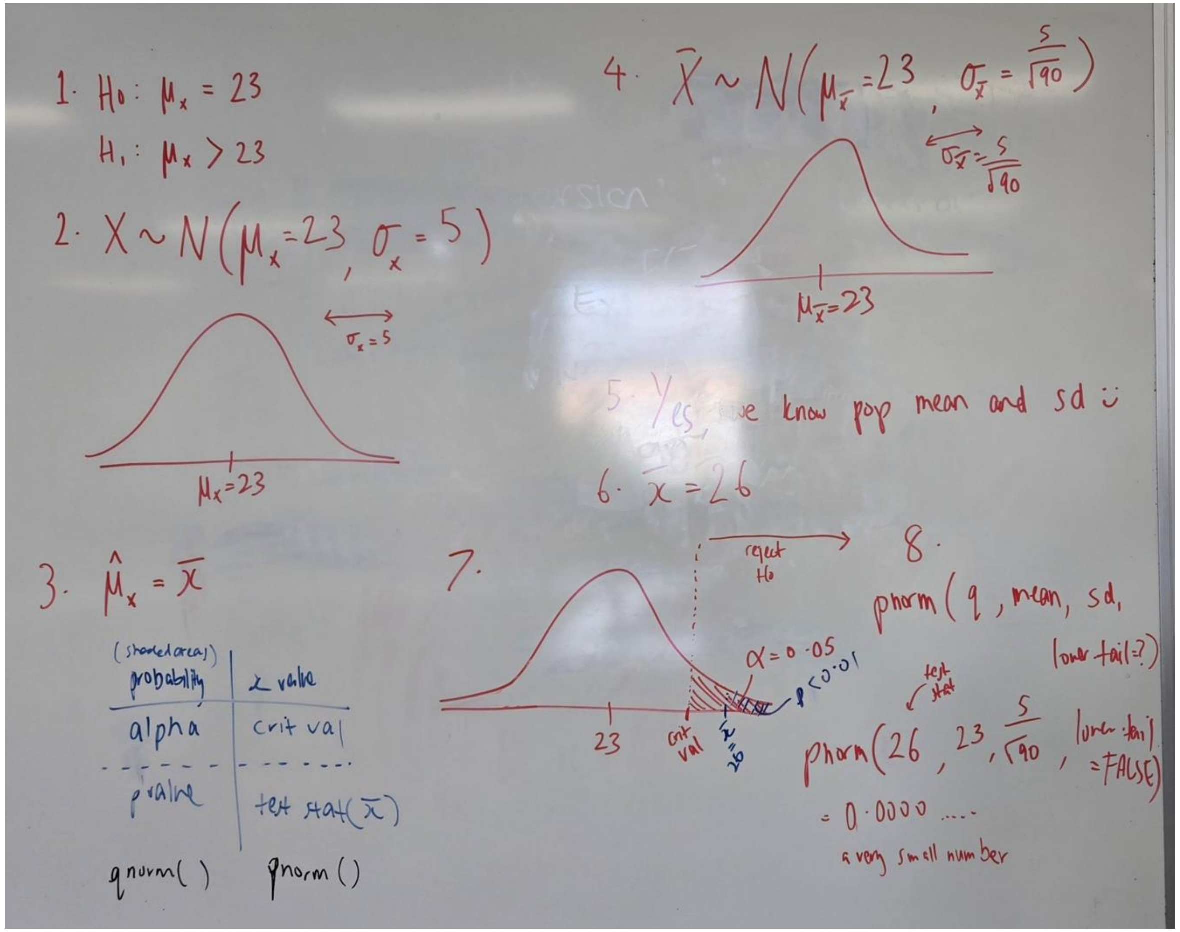
Practicing in R

- Use datasets from previous tutorials (e.g. the built in datasets in R) to practice running your different stats tests

- I will run through a simple example using the iris dataset

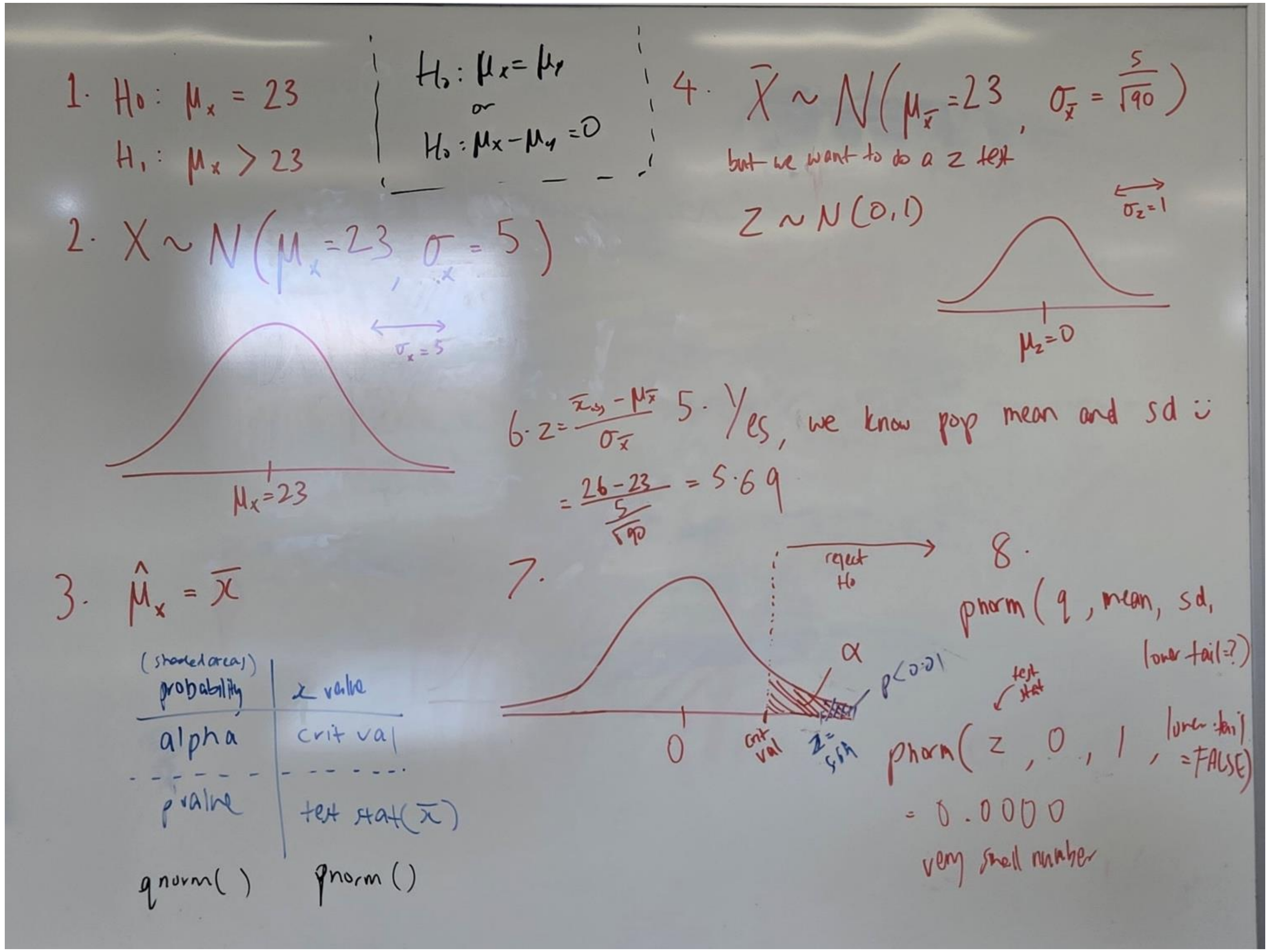
- Start planning for your final project:
  - Have you chosen a dataset?
  - Is the dataset appropriate for the tests you are required to run?
  - What are your research questions? What are your null/alternative hypotheses?
  - Start wrangling and cleaning early on, so if any issues, can ask me, or choose a different dataset

Whiteboard scribbles from Sophie’s classes – normal test for temperature example



Sorry about the glare :(

Whiteboard scribbles from Sophie's classes – z test for temperature example



Whiteboard scribbles from Sophie’s classes – independent t test for exam score example