Research method - 10004

Mind, Brain And Behaviour 2 (University of Melbourne)

1. Inference

A. Introduction to quantitative psychological research
B. Inference as the goal of psychological research
C. Populations and samples
D. Examples
E. Summary

A. intro to quantitative psychological research

Quantitative psychological research addresses a broad range of topics. Even though there are many types of questions asked, they tend to fall into one of three categories:

1. Difference - **is one group of people different to another in some way**? MBB2 online modules focus on this.

2. Association – **is one construct related to another?** A MBB2 tutorial class will address this

3. Prediction – does one construct influence another? You will learn about this in future psychology subjects

B. The Goal of Psychological Research

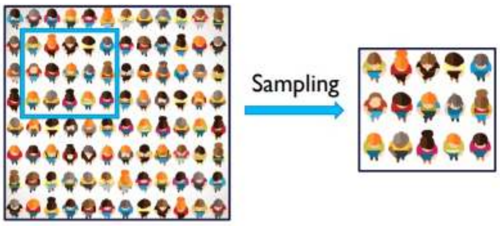
- When we conduct a psychological research project, our aim is to make ***inferences*** (in other words, suggestions or claims) about a **population**. Put simply, we want to say something about a ***population***
- A population is **everyone of interest to a research question**. In other words, it is the research question that defines the population.

c. samples & populations

Taking Samples from Populations

It is usually not possible to recruit all people in a population to participate in a study.

Instead use a ***sample***: a group of people taken from the population to participate in a study.



Making Inferences Based on Samples

We can then make ***inferences*** about the population based on what happens with ***measurement*** of our sample.

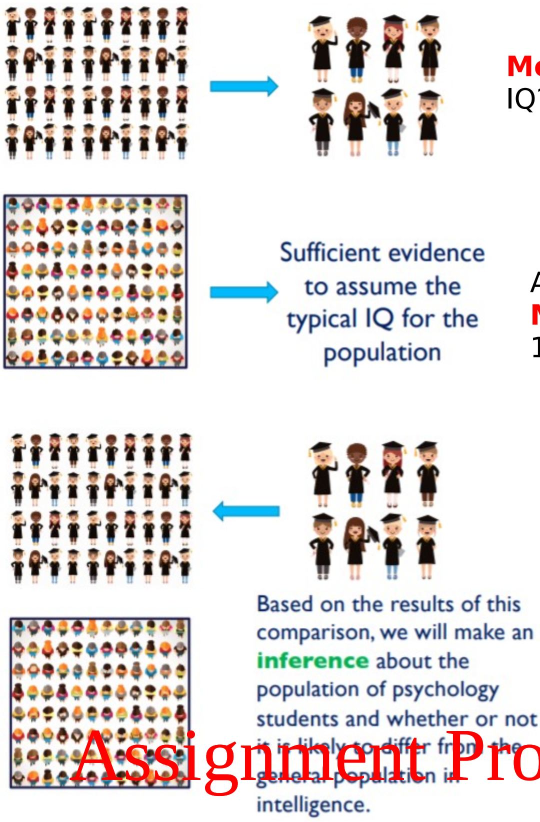
We aim to infer that what is typical for our sample should also be typical for the population.

D. Example

Example 1: Are Psychology Students Smarter?

- **Research Question**:
  Are psychology students smarter than the general population?
- Answering this question involves knowing the **typical IQ for**:
  1. Psychology students 2. The general population.

Once we know **what is typical** for each of these groups, then we can compare them and assess the evidence for a difference.



Mean IQ?

Assume **Mean** IQ is 100

In quantitative terms, **mean scores** will be our indicator of what is typical.

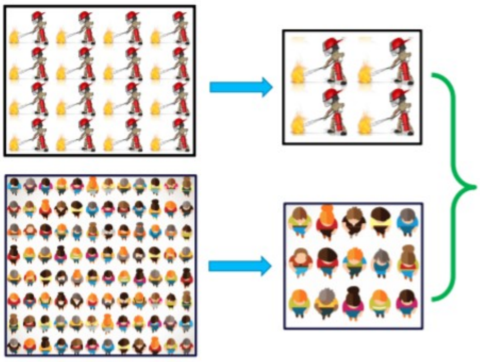
We know the population mean in this example. We could compare our sample mean to this value to assess evidence for a **difference**.

Based on the results of this comparison, we will make an **inference** about the population of psychology students and whether or not it is likely to differ from the general population in intelligence.

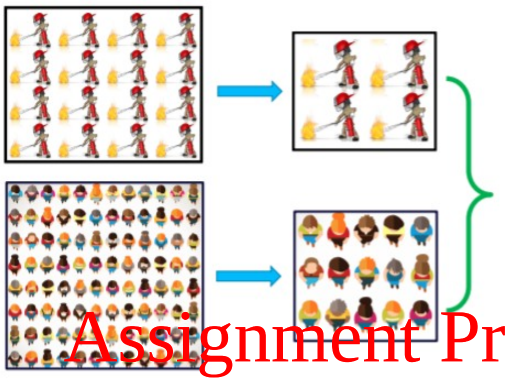
EXAMPLE 2: Anxiety in Firefighters

- Research Question:
  Do Firefighters differ from the general population in their experience of anxiety?
- Answering this question involves knowing the **typical level of anxiety** for:
  1. Firefighters 2. People who are not Firefighters

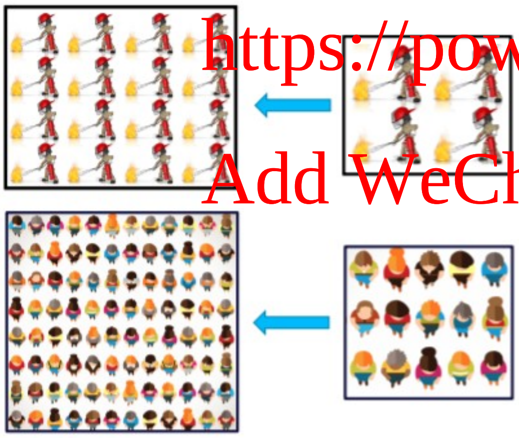
Once we know what is **typical** for each of these groups, then we can **compare** them and **assess the evidence for a difference.**



Is there a **difference** between groups in the **typical level** of anxiety?



We would **<u>compare</u>** the **mean** anxiety score for our sample of firefighters and the **mean** anxiety score for our **sample** of people who are not firefighters.
Is there a difference between mean scores?



All along, we are aiming to **infer** that the result of <u>the sample mean comparison</u> is likely to be the <u>same for the populations.</u>

E, Summary

- Quantitative psychological research aims to generate knowledge about **populations**.
- A population is **everyone** of interest to a research question.
- Because we usually can't measure a population in its entirety, **<u>a sample is drawn from the population and measured.</u>**
- We can **<u>make inferences about the population based on the evidence observed in the sample.</u>**

- In these MBB2 modules, we will focus on inference about **<u>differences in means scores.</u>**

2, Typical and Extreme Scores

A. Recap on distributions of individuals' data
B. The Normal Distribution
C. Typical and extreme scores
D. The 2s rule of thumb
E. Summary

A. Recap on distributions of individuals' data

Distribution of Data:

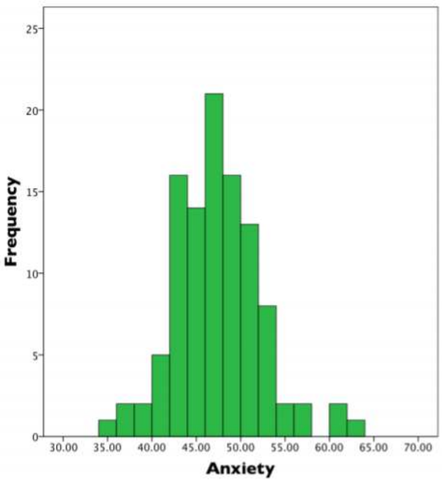
- To understand a psychological construct, we need to know how it is **distributed across a population**.
- When measured, constructs takes on different values for different people in a sample.

- Collectively, those different values form a **distribution** of data, which can be described in terms of **central tendency (mean) and variability (**the standard deviation)

  ➔ SD = A common measure of average spread (or variability) around the centre score (i.e., the mean)

Typically we estimate σ using s, the standard deviation of a sample

-

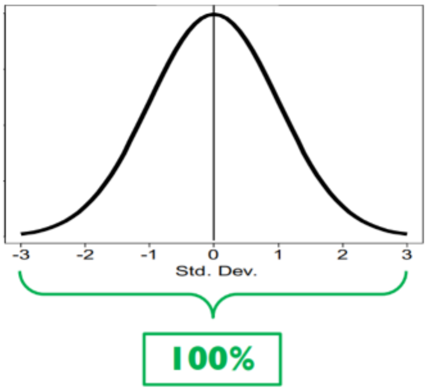


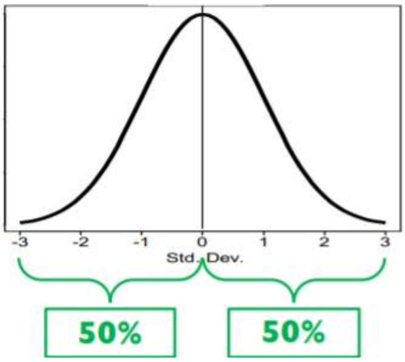
A **histogram** is one way to represent a distribution of data

Distributions of data can be described according to their **central tendency (*m*)** and **variability (*s*)**
m = 46.87 s = 4.84

➔ a **normal** shape.
- Most of the people are in the middle. The peak of the graph.
- Relatively fewer people are on the outsides. The tails of the graph.

B. Normal distribution

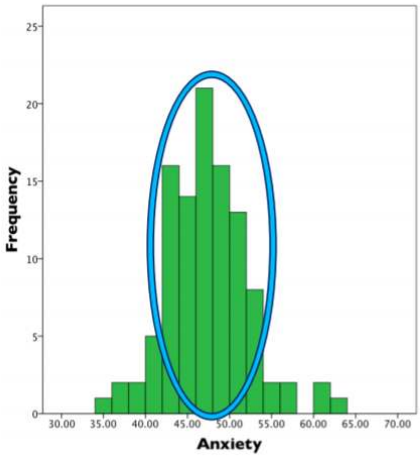




Majority of observations are in the **middle**.
Observations **reduce in frequency towards the tails.**
The distribution is **symmetrical**.

C. Typical and Extreme Scores

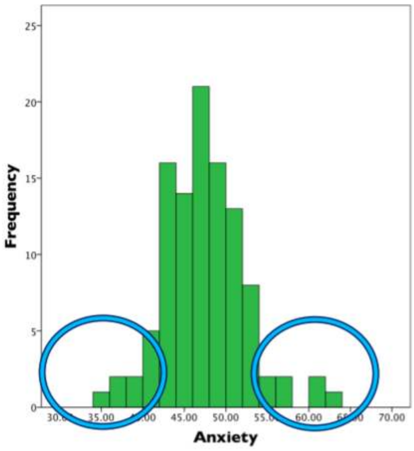
Typical scores



In this distribution, it is expected or typical to find scores around 40 - 55 in this distribution. Why? These scores **occur frequently**

Typical scores are expected, occur with high frequency/probability

Extreme scores



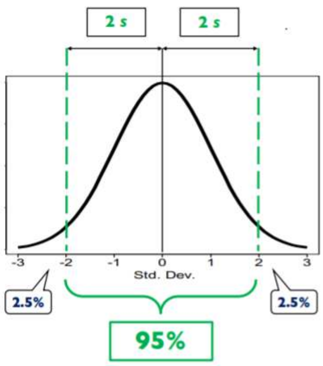
it is relatively **unusual** to find extreme scores – ones that are very low or very high.
Why? They occur **infrequently** in this distribution.

They indicate a **difference** to typical scores in terms of whatever is being measured.

Typical VS Extreme

How can we more reliably decide what counts as **typical vs extreme?**

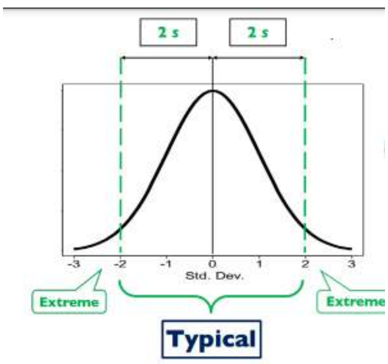
D. The 2s Rule of Thumb



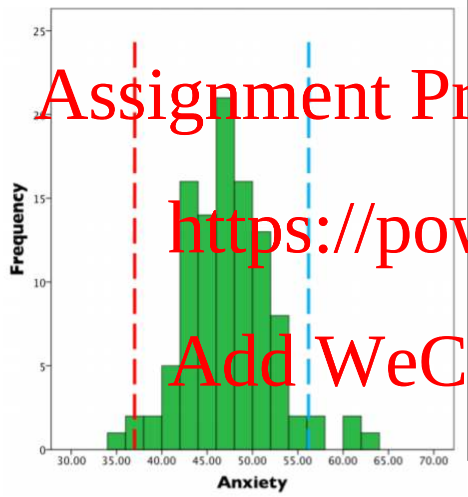
In a distribution with a normal shape, **95%** of scores fall within approximately **2 standard deviations (s)** of the mean.

those scores o**utside** **2 standard deviations (s)** of the mean as being typical. They are expected as they occur frequently in this distribution

those scores outside 2 standard deviations (s) of the mean as being **extreme**. They are not expected as they occur infrequently in this distribution



Applying the 2s Rule of Thumb



We can apply our **2s rule of thumb** to decide what might be **typical** and **extreme** in this real distribution of data.

m = 46.87, s = 4.84

s = 4.84 x 2 = 9.68

Lower limit

46.87 - 9.68 = 37.19

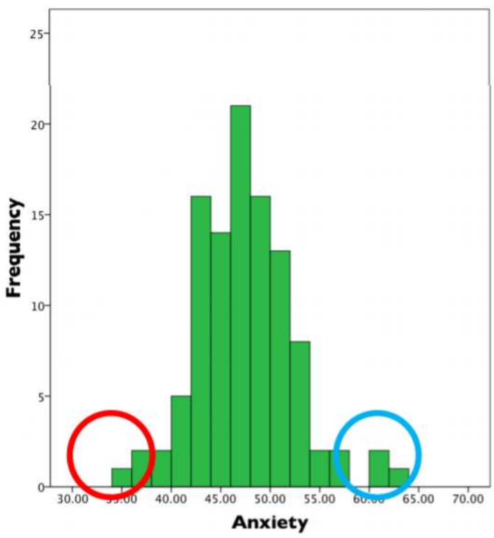
Upper limit

46.87 9.68 = **56.55**

➔ More extreme than lower limit: 2.02%(m-lower l)

More extreme than upper limit: 4.04%

➔ Within 2s of distribution mean: **93.94%**



The extreme cases are noteworthy.

They are unusual in terms of reported level of anxiety and indicate a difference from what is typical.

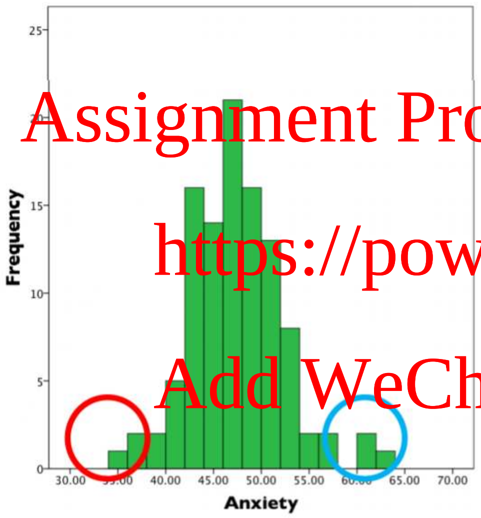
E. Summary

- A distribution of data can be fundamentally described according to its central tendency (m) and variability (s).
- In a normal distribution, most observations are close to ***m***, and they reduce in frequency towards the distribution's tails.
- We can use the **2s Rule of Thumb** to define a **critical limit,** beyond which, cases are considered extreme.
- We are most interested in extreme cases. These cases are unusual and indicate a difference in terms of what is being measured

3, Distribution of Sample Means

A. Recap: Distributions of Individuals’ Data
B. The Distribution of Sample Means
C. Central Limit Theorem
D. Summary

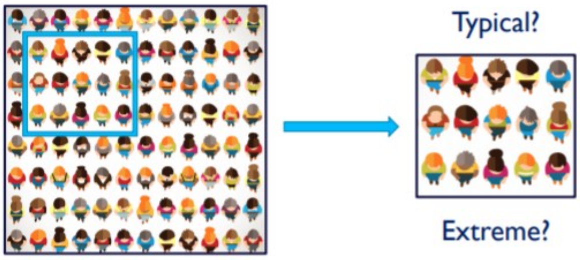
A. Recap: Distributions of Individuals’ Data



We previously determined if **an individual’s score** was typical or if it was extreme according to the characteristics of a **distribution of OTHER individuals’ scores.**

B. What about an Entire Sample – the distribution of Sample Means

What if we wanted to know if not just an individual, but a **sample**, was typical or if it was extreme?



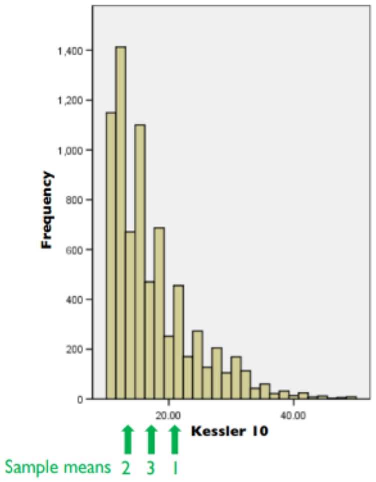
One population, many sample

- When we conduct research, we usually **recruit one sample** from each population of interest.
- However, there are **many samples** that could possibly be recruited from any population.
- How can we tell if our one sample is typical of a certain population, or if it is extreme and indicates a difference?
    - ➔ We previously examined **individual scores** in a **distribution of other individual's scores**. We can do much the same thing to determine if a sample is typical or if it is extreme.
    - ➔ ① we would need a score for our one sample. This would be **the sample mean.** ② we would need a **distribution** made up of **sample means**, within which, we could examine our **one sample mean**.

THE DISTRIBUTION OF SAMPLE MEANS

Collection of all the possible random sample of a particular size (n) that can be obtained from a population

Examples:



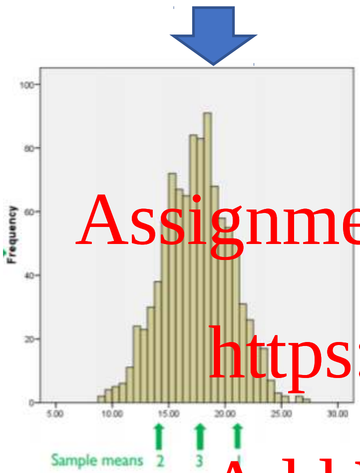
Suppose this is a **population**

Let's take some samples of 5 people at random.

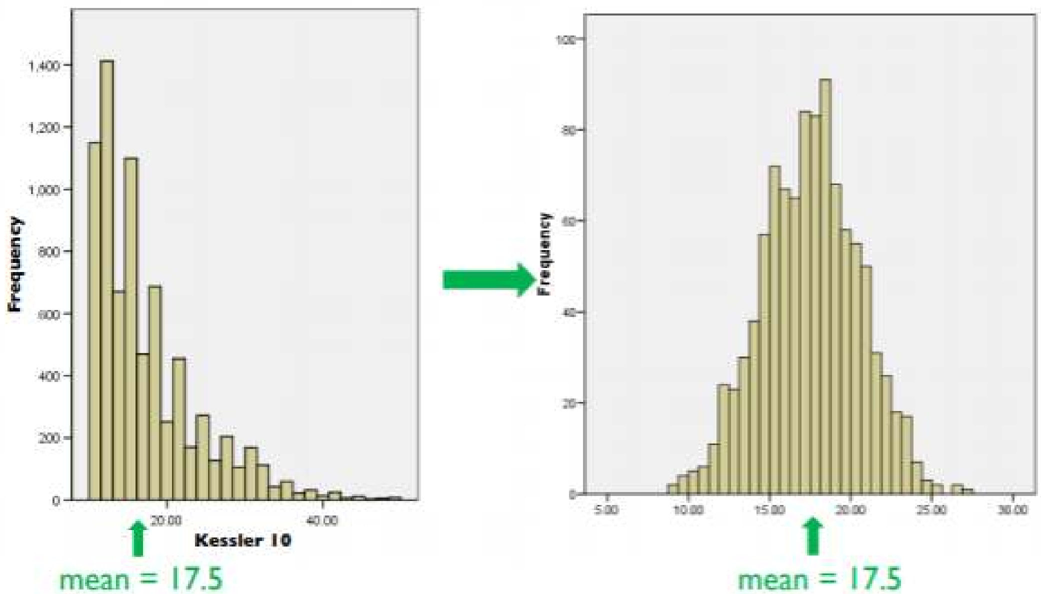
Scores – sample 1
13, 16, 18, 28, 29 Mean = 20.8

Scores – sample 2
12, 13, 13, 14, 18 Mean = 14.0

Scores – sample 3
15, 15, 15, 20, 21 Mean = 17.2



A **Distribution of Sample Means** with 1000 random samples, each with n = 5.



C. Central Limit Theorem

- In real research, it is not pragmatically possible to collect all possible random samples from a population. However, we don't need to.
- The **Distribution of Sample Means** is a **theoretical distribution** governed by a mathematical theorem the **Central Limit Theorem**.
- **The Central Limit Theorem** tells us the precise characteristics of the distribution of any distribution of sample means.

Central Limit Theorem Tells Us

- The precise characteristics of a distribution of sample means for samples of any size (n).
- The distribution of sample means is the **SAME as the population mean.**
- For **large sample sizes (≥ 30)**, the distribution of sample means will have a **NORMAL** shape, regardless of shape, mean, SD of population
- so we can work out what is 2SDs from the mean

- the standard deviation of the sampling distribution means is called the ***standard error*** of the mean

formula of standard error

The standard error formula: $\sigma_M = \frac{\sigma}{\sqrt{n}}$

As sample size ↑, standard error ↓.

In turn, estimation of the population mean becomes more precise. When a sample is large enough, its mean provides a reliable estimate of the population mean.

What does this all mean?

- We know if our sample is large enough, the distribution of sample means will be **normal**.
- We know the mean of the distribution of sample means, and we can calculate its **standard error.**
- Therefore, we can use our **2s rule of thumb** to test if our sample mean is typical or if it is extreme

D. Summary

- We can decide if individuals' scores are typical or extreme by comparing one score with a distribution of other individuals' scores.

- We can apply a similar process to determine if a sample is typical or extreme, with the use of our **sample mean** and a **distribution of sample means**.
- **Central Limit Theorem** tells us:

1) ***the distribution of sample means*** **=** ***population mean***;

2) the details of standard error; how the standard error is related to the population standard deviation

3) as sample size increases, standard error decreases.

infer the population mean from the sample mean

4: Null Hypothesis Significance Testing

A) Hypotheses
B) The Null Hypothesis
C) Null Hypothesis Significance Testing
D) Determining the probability of a sample mean
E) Summary

A) Hypotheses

- A hypothesis is a predictive statement. When we conduct psychological research, we pose and test an **experimental hypothesis.**
- An experimental hypothesis is a statement that **predicts an EFFECT**. This effect might be one of…DIFFERENCE / ASSOCIATION
- **Experimental** hypotheses = **alternative** hypotheses. But, alternative to what?

B) Null Hypotheses

- A **null hypothesis** is also a prediction. The null always predicts the one basic concept regardless of what is being investigated – that **nothing is happening**.
- States that in the general population there is no change, no difference, or no relationship
- In other words, the null hypothesis is a hypothesis of **NO effect**
  - ➔ NO DIFFERENCE / NO ASSOICATION
- The **experimental** hypothesis is the alternative to the null hypothesis. Only one of these hypotheses can be supported by the research data

Statistical Notation:

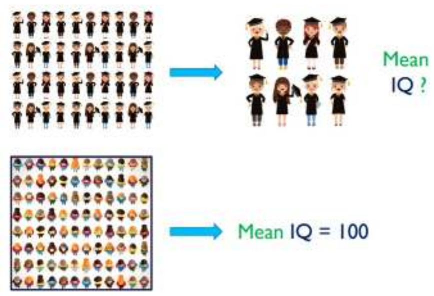
Null Hypotheses: H0
Alternative (Experimental): H1

C) Null Hypothesis Significance Testing

1. Propose a null hypothesis that a population parameter (mean) has a particular value.
2. Proceed assuming the null hypothesis is true.
3. Determine the probability of the sample mean occurring if the null hypothesis is true.
4. If the probability of the sample mean occurring is small, reject the null hypothesis.
   If the probability is large, do not reject the null hypothesis.

D) Example: Null Hypothesis Significance Testing

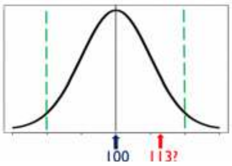
Are Psychology Students Smarter?



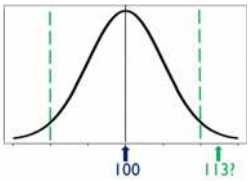
Step 1:
Pose a null hypothesis that a **population parameter** has a certain value.
H0=100
Step 2:
Proceed assuming H0 = 100 is true. Collect the data. Observe the sample mean.
H0 *= 100, Observed* ***m*** *= 113*

Step 3:

- Determine the **probability** of the sample mean occurring if the null hypothesis is true. Given H0 = 100, is our sample mean of 113 typical, or is it extreme?
- Involves a statistical test based on a distribution of sample means with a mean of 100 (the same as the null hypothesized population mean). We know the shape will be normal. We can also calculate standard error.
  - ➔ If our sample mean is typical, then it does not provide evidence for a difference.



  - ➔ If our sample mean is extreme, then it does provide evidence for a difference.



Step 4: Either reject or do not reject the null hypothesis.

Extreme sample mean? Evidence for a difference. Reject null hypothesis.

➔ Probability is low

Typical sample mean? No evidence for a difference. Do not reject null hypothesis.

➔ Probability is HIGH

if the null hypothesis is true we expect our sample mean to be within 2 SDs of the null hypothesis mean

We can calculate the probability using z-scores

E) Summary

When we conduct research, we pose an **experimental hypothesis** – one of effect.

The opposite of this is a **null hypothesis** – one of **no effect.**

Null Hypothesis Significance Testing involves

1) ***assigning a value*** to the null hypothesis.
2) then conduct ***a statistical test*** to determine the ***probability of our sample mean occurring if the null hypothesis is true***.
3) If the probability is low, we reject the null hypothesis
   if the probability is high, we do not reject it

5 The Single Sample z-Test

A. Null Hypothesis Significance Testing Recap

1. Propose a null hypothesis that a population parameter (mean) has a particular value.
2. Proceed assuming the null hypothesis is **true**.
3. Determine the **probability** of the sample mean occurring if the null hypothesis is true.
   - ➔ If the probability of the sample mean occurring is small, reject the null hypothesis. If the probability is large, do not reject the null hypothesis.

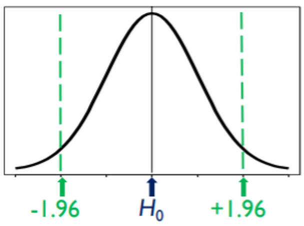
Probability: What Counts as Small?

5%: Alpha Level / Level of Significance

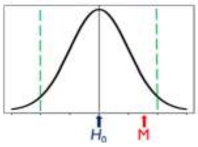
B. 5% Alpha level / Level of significance

The alpha level defines which sample means in a distribution of sample means are expected or typical, and which are unlikely or extreme, if the null hypothesis is true

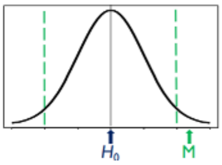
When the comparison distribution is perfectly normal, the critical limits set by the 5% Alpha Level are precisely +/-1.96 standard errors from the mean of the distribution.



- ➔ If our sample mean is inside these limits, the probability is greater than 5%, and therefore, high. Do not reject the null hypothesis.



➔ If our sample mean is outside these limits, the probability is lower than 5%, and therefore, low. Reject the null hypothesis



C.The Single Sample z-Test

Now that an **Alpha Level of 5%** has been set, we must determine <u>the probability of our sample mean occurring.</u>
We can do this by performing a **single sample z-test.**

In other words, we will **<u>calculate a z-score for our sample mean</u>**.
In this context, a z-score will express **how many standard errors our sample mean is away from H0**

For example, a z-score of $z = 1.5$ would indicate that our sample mean is 1.5 standard errors above the mean of the distribution of sample means.

a z-score of **$z = -1.5$** would indicate that our sample mean is **1.5 standard errors below** the mean of the distribution of sample means.

A single sample z-test is calculated using the following formula

$$z = \frac{M - \mu}{\sigma_M}$$

- $\sigma_M = \sigma / \sqrt{n}$

z = the z-score for our sample mean M = our sample mean μ = population mean σ_M = standard error of the mean. Note that population standard deviation is known.

It is a single sample test: compares the sample mean with a given number and asks is there a difference

D. Summary

When determining if our sample mean provides evidence to support the null hypothesis or not, we…

Set an alpha level. Here we have used 5%. This defines which sample means are expected and which are unlikely if the null hypothesis is true.
In a normal distribution of sample means, the 5% alpha level corresponds to +/-1.96 standard errors from the null hypothesised mean value.

Calculate a z-score for our sample mean. Standard error is based on population standard deviation in this instance.
If the z-score is less than 1.96, do not reject the null hypothesis.
If the z-score is greater than 1.96, reject the null hypothesis.

6 The Sigle Sample t-test

A. z-test recap

$$z = \frac{M - \mu}{\sigma_M}$$

z-score: **how many standard errors our sample mean is away from H0.**

z = the z-score for our sample mean
M = our sample mean
μ = population mean
σ_M = standard error of the mean. $\sigma_M = \sigma / \sqrt{n}$

Note that population standard deviation is known.

B. Handling the case of unknown population standard deviation

However, population standard deviation is rarely known.

In instances when we don't know the population standard deviation, we can't use a **z-test** and the normal distribution to assess our sample mean.
That's okay! Instead, we can use a **t-test** and the '**t-distribution'.**

We use **SAMPLE standard deviation** as an **ESTIMATE of POPULATION standard deviation**. Single sample t-tests and z-tests are very similar otherwise.

Z-Test VS T-Test

Z-Test standard error (σ)	T-Test (estimate **σ with s)**

$\sigma_M = \frac{\sigma}{\sqrt{n}}$	$\boxed{s_M} = \frac{\boxed{s}}{\sqrt{n}}$
Z-Test Formula	**t-test formula**
$z = \frac{M - \mu}{\sigma_M}$ M = sample mean µ = population	$t = \frac{M - \mu}{s_M}$

Almost all aspects of the process are the **same** when conducting either a single sample t-test or z-test.

We still use **Null Hypothesis Significance Testing**. We still assign a value that indicates no effect to the null hypothesis, and proceed assuming the null is true.

We still apply an alpha level of 5% and determine the probability of our mean occurring. The result determines whether the null hypothesis is rejected or not.

One Difference of t-distribution

- In the t-distribution, the critical limit corresponding to our alpha level of 5% will not be fixed at +/- 1.96 as it was with the z-test and normal distribution.

- Instead, the t-distribution requires that we consider sample size and **degrees of freedom (df)**, when determining the probability of the sample mean.
  - ➔ Degrees of Freedom are one less than our sample size for a single sample t-test (**n-1)**. n = sample size

- The critical limit varies along with df

Why do we use n-1 when calculating sample standard deviation?

- We use n-1 as the process of inferring our sample is a representative of the population is an ***estimation*** – imprecise process
  - This process is biased because samples are often less variable (less spread out) than the population they come from
  - **The underestimation bias is corrected by using n-1 degrees of freedom**
  - Dividing by n-1 gives us **a larger result**

- Both tests have very similar calculations to ***get a probability for a sample mean***
- The t-score gives the probability for the sample mean, just like the z-score
  - ➔ We reject the null hypothesis if the probability is too small, just as before
- However:
  - ➔ T distribution requires degrees of freedom, df, to calculate the correct probability
    df = n – 1, where n= sample size

Example Does Head Injury Affect IQ?

- Imagine that we have a sample of 10 people with an acquired head injury. Has IQ been affected?
- The **null hypothesis** would be that **head injury has not affected IQ**. **H0: μ = 100**.
- Sample data: **M** = 94.20, **s** = 6.16, **n** = 10
- T-score $t = \frac{M - \mu}{s_M}$

Step 1: Calculate s_M (standard error)

$$t = \frac{M - \mu}{\boxed{s_M}} \qquad s_M = \frac{s}{\sqrt{n}}$$

$$s_M = \frac{6.16}{\sqrt{10}}$$

$$s_M = \frac{6.16}{3.16}$$

$$\boxed{s_M = 1.95}$$

Step 2: Calculate the t-score for our sample mean

$$t = \frac{94.2 - 100}{1.95}$$

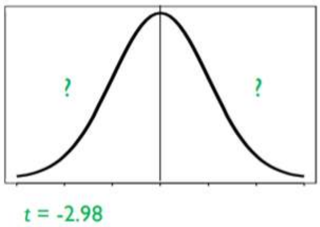
$$t = \frac{-5.8}{1.95}$$

$$\boxed{t = -2.98}$$

$$t = \frac{M - \mu}{s_M}$$

Is the t-score of **-2.98** for our sample mean more extreme than the critical limit, taking **df** into account?

- Yes! A t-score of -2.98 is extreme for df =n-1=10-1= 9.



This means that the probability of our sample mean occurring, assuming the null hypothesis is true, is less than the alpha level of 5%. In fact, it is just 1.6%

Summary