# Contents

# Day 17

**This material is not on the exam but will tie into different concepts that will be present**

# Review

## Neyman-Pearson Hypothesis Testing

Step 1: Define a parameter and its value under $H_0$ & $H_1$

Step 2: Define a test statistic and its sampling distributions under the two hypothesis.

Step 3: Define our alpha level ($\alpha$) which is $P(Type\,1\,Error)$. This translates to P(accept $H_1$ | $H_0$ is true). Use it to compute a critical region.

Step 4: Collect sample data and compute observed value of test statistic.

Step 5: If value in critical region, then accept $H_1$. If value not in critical region, then accept $H_0$.

## What changes now?

We are using stand in variables for instead of using categories. This is still using the NPHT framework but uses numerical data.

Step 1: Parameter is $\mu$ = population mean. Now we are talking about numerical data now, not categorical data.

- $H_0 = \mu = \mu_0$
- $H_1 = \mu = \mu_1$
- This value of $\mu$ is cc

Step 2: Statistic $\bar{X}$ = sample mean. Assuming Central Limit Theorem holds, we will have an approximately normal sampling distribution and will be centered at $\mu_0$ under the null and $\mu_1$ under the alternative hypothesis.

- $H_0, \bar{X} \sim N(\mu_0, \frac{\sigma}{\sqrt{n}})$
- $H_1, \bar{X} \sim N(\mu_1, \frac{\sigma}{\sqrt{n}})$

## Sonnet Example

For an old author, sonnets are known to contain an average of 8.9 "new" words, with a standard deviation of 2.5 words. New meaning that they are unique to each of the sonnet and are not used in the others.

A new set of 6 sonnets has been discovered and authorship has been disputed. The sonnets contain an average of 10.2 "new" words. Suppose we believe a different author would use on average 10.9 "new" words.

Using $\alpha = 0.05$, what should we conclude about the authorship of the new sonnets?

### Response

Parameter: $\mu$ = population mean which in this context will be the number of "new" words

- Under $H_0$: $\mu = 8.9$
    - Step 2: $\bar{X} \sim N(8.9, \frac{2.5}{\sqrt{6}})$
- Under $H_1$: $\mu = 10.9$
    - Step 2: $\bar{X} \sim N(10.9, \frac{2.5}{\sqrt{6}})$
- Here we are stating that it is worse to give credit to someone who didn't write the sonnets.
- We then accept the null hypothesis because the value is not in the critical region of 10.58. This means the original author wrote the sonnets.
- The way we set up the problem will affect the conclusion we come to.

### R-Code

```
alpha <- 0.05
mu <- 8.9
sigma_from_sample <- 2.5/sqrt(6)

qnorm(alpha, mu, sigma_from_sample, lower.tail = FALSE)
```

## Neyman-Pearson Power Analysis

<u>Step 1</u>: Define a parameter and its value under $H_0$ & $H_1$

<u>Step 2</u>: Define a test statistic and its sampling distributions under the two hypothesis.

<u>Step 3</u>: Define our alpha level ($\alpha$) which is $P(Type\,1\,Error)$. This translates to P(accept $H_1$ | $H_0$ is true). Use it to compute a <u>critical region</u>.

<u>Step 4</u>: Compute our power = P(test statistic in critical region | $H_1$ is true)

<u>Step 5</u>: [OPTIONAL] Compare power to 80% and/or compute $\beta$ to $\alpha < \beta \le 0.2$

### Sonnet Example with Power Analysis

For an old author, sonnets are known to contain an average of 8.9 "new" words, with a standard deviation of 2.5 words. New meaning that they are unique to each of the sonnet and are not used in the others.

A new set of 6 sonnets has been discovered and authorship has been disputed. The sonnets contain an average of 10.2 "new" words. Suppose we believe a different author would use on average 10.9 "new" words.

Using $\alpha = 0.05$, what is the power of our test to detect the "new" author?

You are trying to compute a probability here so this R code will suffice

### R-Code

```
alpha <- 0.05
old_mu <- 8.9
mu <- 10.9
sigma_from_sample <- 2.5/sqrt(6)

critical_value <- qnorm(alpha, old_mu, sigma_from_sample, lower.tail = FALSE)

pnorm(critical_value, mu, sigma_from_sample, lower.tail = FALSE)
[1] 0.6235198
```

## NHST Framework

Step 1: Define a parameter and its value under $H_0$

Step 2: Define an interval representing an inequality (under $H_0$, parameter in that interval)

Step 3: Define a test statistic and its sampling distribution under $H_0$

Step 4: Collect data and compute the observed value of the test statistic.

Step 5: Compute the P-Value which is P(observe a test statistic as or more favorable to $H_1$ | $H_0$ is true)

- P(test statistic $\geq$ observed value) OR
- P(test statistic $\leq$ observed value) OR
- P(test statistic "further" from parameter value compared to observed value)

Find either of the three listed above ↑ and double the smaller one.

Step 6: We define our significance level.

**If**

- P-Value $\leq$ significance level $\implies$ reject $H_0$ & accept $H_1$
- P-Value $>$ significance level $\implies$ fail to reject $H_0$

## Sonnet Example with NHST

For an old author, sonnets are known to contain an average of 8.9 "new" words, with a standard deviation of 2.5 words. New meaning that they are unique to each of the sonnet and are not used in the others.

A new set of 6 sonnets has been discovered and authorship has been disputed. The sonnets contain an average of 10.2 "new" words. Suppose we believe a different author would use on average 10.9 "new" words.

Theory suggest a different author would use more "new" words.

What should we conclude about authorship at the 5% significance level?

Parameter: $\mu$ = population mean number of new words

- $H_0 : \mu = 8.9$
- $H_1 : \mu > 8.9$
    - can only be $>, <, \neq$

We assume the new author is going to have more words than the original author.

$$\bar{X} \sim N(8.9, \frac{2.5}{\sqrt{6}})$$

## R-Code

```
x_bar <- 10.2
mu <- 8.9
sigma_from_sample <- 2.5/sqrt(6)

pnorm(x_bar, mu, sigma_from_sample, lower.tail = FALSE)
[1] 0.1013787
```

Since the outcome is 10 percent, we fail to reject the null hypothesis. We need to default back onto the original assumption. We do not have any definitive truth to accept the null hypothesis but since we have nothing else to fall back onto be need to say its likely to still be true.

## Two-sided Tests

Use this when theory does not definitively give an "alternative".

For Neyman-Pearson tests: critical region is half in left tail and half in right tail of sampling distribution under $H_0$.

This means that our power will decrease ($\downarrow$).

For NHST: Find the "one-sided" p-value & double it.