Session 13: Hypothesis Testing

Agenda

- **Define Hypothesis Testing**
- **Null Hypothesis Vs Alternative Hypothesis**
- **Type I Vs Type II Errors**
- **P-value And Significance Level**
- α **Vs** β
- **Steps Of Hypothesis Testing**
- **Determine A P-value When Testing A Null Hypothesis**
- **Question And Answer**
- **Upper-Tailed, Lower-Tailed, Two-Tailed Tests**
- **Problems And Answer**
- **Types Of Hypothesis Testing**

Define Hypothesis Testing

- What is Hypothesis?
  - An educated guess
  - A claim or statement about a property of a population

- What is the goal of Hypothesis Testing?
  - To analyze a sample in an attempt to distinguish between population characteristics, that are likely to occur and population characteristics that are unlikely to occur.

Null Hypothesis Vs Alternative Hypothesis

Null Hypothesis

- Statement about the value of a population parameter
- Represented by H0
- Always stated as an Equality

Alternative Hypothesis

- Statement about the value of a population parameter that must be true if the null hypothesis is false
- Represented by H1
- Stated in on of three forms
  - $>$
  - $<$
  - $\neq$

Type I Vs Type II Errors

- A type I error is the incorrect rejection of a true null hypothesis (a "false positive"), while a type II error is the failure to reject a false null hypothesis (a "false negative")

Examples of type I errors include a test that shows a patient to have a disease when in fact the patient does not have the disease, a fire alarm going off indicating a fire when in fact there is no fire.

- A type II error (or error of the second kind) is the failure to reject a false null hypothesis.

Examples of type II errors would be a blood test failing to detect the disease it was designed to detect, in a patient who really has the disease, a fire breaking out and the fire alarm does not ring.

Type I vs Type II Errors(Contd..)

	Condition of null hypothesis	
Possible action	True	False
Fail to reject H_0	Correct $(1-\alpha)$	Type II error β
Reject H_0	Type I error α	Correct $(1-\beta)$

P-value And Significance Level

P-Value

In statistical significance testing, the p-value is the probability of obtaining a test statistic at least as extreme as the one that was actually observed, assuming that the null hypothesis is true.

Significance level

The probability of rejecting the null hypothesis is when it is called the significance level α

Note: If the *p*-value is equal to or smaller than the significance level (α), it suggests that the observed data is inconsistent with the assumption that the null hypothesis is true and thus this hypothesis must be rejected (but this does not automatically mean the alternative hypothesis can be accepted as true).

α Vs β

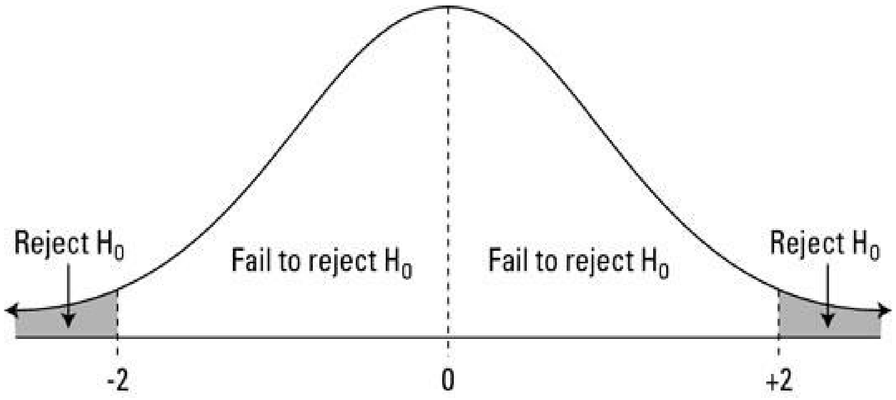
- α is the probability of Type I error
- β is the probability of Type II error
- The experimenters (you and I) have the freedom to set the α-level for a particular hypothesis test.
- That level is called the level of significance for the test. Changing α can (and often does) affect the results of the test—whether you reject or fail to reject H_0.
- As α increases, β decreases and vice versa.
- The only way to decrease both α and β is to increase the sample size. To make both quantities equal zero, the sample size would have to be infinite- you would have to sample the entire population.

Steps Of Hypothesis Testing

1. Describe the population characteristic about which hypotheses are to be tested
2. State the null hypothesis, Ho
3. State the alternative hypothesis, H1 or Ha
4. Display the test statistic to be used
5. Identify the rejection region
6. Is it an upper, lower, or two-tailed test?
7. Determine the critical value associated with α, the level of significance of the test
8. Compute all the quantities in the test statistic, and compute the test statistic itself
9. State the conclusion. That is, decide whether to reject the null hypothesis, Ho, or not to reject the null hypothesis. The conclusion depends on the level of significance of the test. Also, remember to state your result in the context of the specific problem

Determine A p-Value When Testing A Null Hypothesis

- If the alternative hypothesis is the less-than alternative, you reject H0 only if the test statistic falls in the left tail of the distribution (below –2). Similarly, if Ha is the greater-than alternative, you reject H0 only if the test statistic falls in the right tail (above 2)



Determine A P-value When Testing A Null Hypothesis (Contd.)

- If Ha contains a less-than alternative, find the probability that Z is less than your test statistic
- If Ha contains a greater-than alternative, find the probability that Z is greater than your test statistic (look up your test statistic on the Z-table, find its corresponding probability, and subtract it from one). The result is your p-value.
- If Ha contains a not-equal-to alternative, find the probability that Z is beyond your test statistic and double it.
  - If your test statistic is negative, first find the probability that Z is less than your test statistic (look up your test statistic on the Z-table and find its corresponding probability). Then double this probability to get the p-value.
  - If your test statistic is positive, first find the probability that Z is greater than your test statistic (look up your test statistic on the Z-table, find its corresponding probability, and subtract it from one). Then double this result to get the p-value.

Question

- True/false: First you decide on the null hypothesis. Then you analyze the data and calculate the probability value. You look at this probability value, and depending on what it is, choose an appropriate alpha level. Then you decide whether you can reject the null hypothesis.

Answer

Ans – False

- You have to select the alpha level before you calculate the probability value. You compare your probability value to your previously selected alpha level when deciding whether or not you can reject the null hypothesis.

Upper-Tailed, Lower-Tailed, Two-Tailed Tests

- H1: $\mu > \mu 0$, where $\mu 0$ is the comparator or null value (e.g., $\mu 0 = 191$ in the example about weight in men in 2006) and an increase is hypothesized - this type of test is called an upper-tailed test
- H1: $\mu < \mu 0$, where a decrease is hypothesized and this is called a lower-tailed test
- H1: $\mu \neq \mu 0$, where a difference is hypothesized and this is called a two-tailed test.

Problem

- Suppose the seller says that the mean life of a fan is more than 15,000 hours. In a sample of 40 fans, it was found that they only last 14,900 hours on average. Assume the population standard deviation is 110 hours. At .05 significance level, can we reject the claim by the seller?
- The null hypothesis is μ ≥ 15000. Begin with computing the test statistic.
-
```
> xbar = 14900           # sample mean
> mu0 = 15000            # hypothesized value
> sigma = 110            # population standard deviation
> n = 40                 # sample size
> z = (xbar−mu0)/(sigma/sqrt(n))
> z                      # test statistic
[1] -5.749091
```

Answer

- The critical value at .05 significance level

```
> alpha = .05
> z.alpha = qnorm(1−alpha)
> −z.alpha            # critical value
[1] −1.6449
```

- The test statistic -4.5644 is less than the critical value of -1.6449. Hence, at .05 significance level, we reject the claim that mean life of a fan is above 15,000 hours.
- Or we can use pnorm on z value

```
> pnorm(-5.749091)
>4.486226e-09
```

Problem: Upper Tail Test Of Population Mean With Known Variance

- Suppose the chocolate wrapper states that there is at most 4 grams of saturated fat in a single chocolate. In a sample of 70 chocolates , it is found that the mean amount of saturated fat per chocolate is 4.2 grams. Assume that the population standard deviation is 0.50 grams. At .05 significance level, can we reject the claim on wrapper?

Answer

- The null hypothesis is that $\mu \le 4$. We begin with computing the test statistic.

```
> xbar = 4.2            # sample mean
> mu0 = 4               # hypothesized value
> sigma = 0.50          # population standard deviation
> n = 70                # sample size
> z = (xbar−mu0)/(sigma/sqrt(n))
> z                     # test statistic
[1] 3.344
```

- We then compute the critical value at .05 significance level.

```
> alpha = .05
> z.alpha = qnorm(1−alpha)
> z.alpha               # critical value
[1] 1.6449
```

Answer (Contd.)

- The test statistic 3.344 is greater than the critical value of 1.6449. Hence, at .05 significance level, we reject the claim that there is at most 4 grams of saturated fat in a chocolate.

Or

- We apply the pnorm function to compute the upper tail p-value of the test statistic. As it turns out to be less than the .05 significance level, we reject the null hypothesis that $\mu \leq 2$.

```
> pval = pnorm(z, lower.tail=FALSE)
> pval                  # upper tail p−value
[1] 0.0004128987
```

Problem - Two-tailed Test Of Population Mean With Known Variance

- Suppose the mean weight of King Penguins found in an Antarctic colony last year was 15.4 kg. In a sample of 35 penguins same time this year in the same colony, the mean penguin weight is 14.6 kg. Assume the population standard deviation is 2.5 kg. At .05 significance level, can we reject the null hypothesis that the mean penguin weight does not differ from last year?

Answer

- The null hypothesis is that μ = 15.4. We begin with computing the test statistic.

```
> xbar = 14.6            # sample mean
> mu0 = 15.4             # hypothesized value
> sigma = 2.5            # population standard deviation
> n = 35                 # sample size
> z = (xbar−mu0)/(sigma/sqrt(n))
> z                      # test statistic
[1] −1.8931
```

- We then compute the critical values at .05 significance level.

```
> alpha = .05
> z.half.alpha = qnorm(1−alpha/2)
> c(−z.half.alpha, z.half.alpha)
[1] −1.9600  1.9600
```

Answer (Contd.)

- The test statistic -1.8931 lies between the critical values -1.9600 and 1.9600. Hence, at .05 significance level, we do not reject the null hypothesis that the mean penguin weight does not differ from last year.
- We apply the pnorm function to compute the two-tailed p-value of the test statistic. It doubles the lower tail p-value as the sample mean is less than the hypothesized value. Since it turns out to be greater than the .05 significance level, we do not reject the null hypothesis that μ = 15.4.

```
> pval = 2 * pnorm(z)    # lower tail
> pval                   # two−tailed p−value
[1] 0.058339
```

Types Of Hypothesis Testing

- Large Sample Tests, Population Mean (known population standard deviation)
- Large Sample Tests, Population Proportion (unknown population standard deviation)
- Small Sample Tests, Mean of a Normal Population

Next Class
Analysis Of Variance (ANOVA)

Sl. No.	Topics For The Agenda
1.	Comparing More Than Two Groups
2.	Difference Between ANOVA And T-Test
3.	Variance For Means Comparison
4.	Between/Within Groups
5.	So What Does ANOVA Do?
6.	One Way ANOVA, Example
7.	Group Comparison
8.	Variance Within And Between Groups
9.	Adjusting For Group Sizes

Sl. No.	Topics For The Agenda
10.	So What's The Deal With F ?
11.	Summary Of ANOVA
12.	ANOVA Table
13.	Some Facts About ANOVA
11.	Summary Of ANOVA
12.	ANOVA Table
13.	Some Facts About ANOVA
14.	R^2 Statistic
15.	R Squared And Adjusted R Squared
16.	Exercises

Contact Info:

- **Website** : http://www.datasciencenirvana.com/
- **LinkedIn** : https://www.linkedin.com/in/gautham111/
- **Email** : egautham@gmail.com