5) The time taken to assemble a car in a certain plant is a random variable having a normal distribution of 20 hours and a standard deviation of 2 hours. What is the probability that a car can be assembled at this plant in a period of time
   a) less than 19.5 hours?   b) between 20 and 22 hours?

5 STATISTICAL INFERENCES

5.1 Introduction

In research, one always has some fixed ideas about cetain population parameters based on say, prior experiments, surveys or experience. However, these are only ideas. There is therefore a need to ascertain whether these ideas /claims are correct or not.
The ascertaining of claims is done by first collecting information in the form of sample data. We then decide whether our sample observations (statistic) have come from a postulated population or not.

Definitions

A **hypothesis** is a claim (assumption) about a population parameters such as the population mean, the population proportion or the population standard deviation is a postulated or a stipulated value of a parameter

Example: The mean monthly cell phone bill in this city is μ = $42

The proportion of adults in this city with cell phones is $\pi = 0.68$

On the basis of observation data, one then performs a test to decide whether the postulated hypothesis should be accepted or not. However, we note that the decision aspect is prone to error/risk.

Null Hypothesis (denoted H_0): Statement of zero or no change and is the hypothesis which is to be actually tested for acceptance or rejection. If the original claim includes equality (<=, =, or >=), it is the null hypothesis. If the original claim does not include equality (<, not equal, >) then the null hypothesis is the complement of the original claim. The null hypothesis *always* includes the equal sign. The decision is based on the null hypothesis.

Eg: The average number of TV sets in U.S. Homes is equal to three ($H_0 : \mu = 3$)

It's always about a population parameter, and not about a sample statistic

Ie $H_0 : \mu = 3$ but **NOT** $H_0 : \bar{x} = 3$

We begin with the assumption that the null hypothesis is true

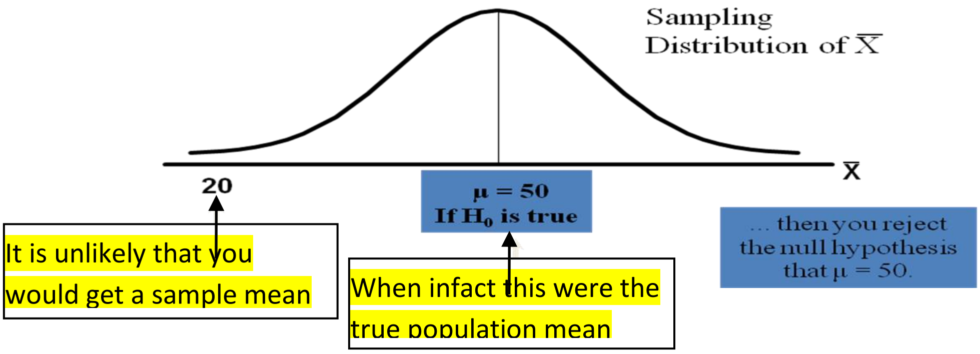
- Similar to the notion of innocent until proven guilty

Alternative Hypothesis (denoted H_1 or H_a): Statement which is true if the null hypothesis is false. it Challenges the status quo. It Is generally the hypothesis that the researcher is trying to prove and it is accepted when H0 is rejected and vice versa. The type of test (left, right, or two-tail) is based on the alternative hypothesis.

5.2 The Hypothesis Testing Process

Claim: The population mean age is 50. Ie Hypothesis H0: $\mu = 50$, vs H_1: $\mu \neq 50$

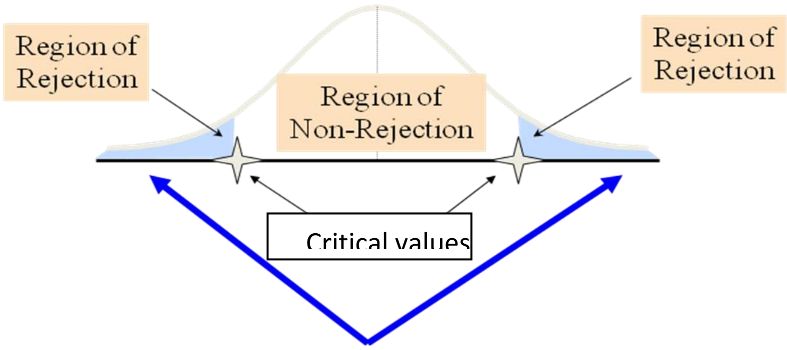
Sample the population and find sample mea. Suppose the sample mean age was $\bar{x} = 20$. This is significantly lower than the claimed mean population age of 50. If the null hypothesis were true, the probability of getting such a different sample mean would be very small, so you reject the null hypothesis .In other words, getting a sample mean of 20 is so unlikely if the population mean was 50, you conclude that the population mean must not be 50.



- If the sample mean is close to the assumed population mean, the null hypothesis is not rejected.
- If the sample mean is far from the assumed population mean, the null hypothesis is rejected.

How far is "far enough" to reject H_0? The critical value of a test statistic creates a "line in the sand" for decision making -- it answers the question of how far is far enough.

Sampling Distribution of the test statistic



"Too Far Away" From Mean of Sampling Distribution

5.2.1 Possible Errors in Hypothesis Test Decision Making

When taking a decision about the acceptance or rejection of a null hypothesis/ alternative hypothesis, there is a risk of committing an error. These errors are of two types:

Type I error; Mistake of rejecting the null hypothesis when it is true (saying false when true). It is usually the more serious error.

The probability of a Type I Error is (denoted α) is Called the level of significance of the test and it is Set by researcher in advance. α= 0.05 and α = 0.01 are common. If no level of significance is given, use α = 0.05. The level of significance is the complement of the level of confidence in estimation.

Type II error: Mistake of failing to reject the null hypothesis when it is false (saying true when false). The probability of a Type II Error is denoted by **β**

Remarks

1) The confidence coefficient $(1-\alpha)$ is the probability of not rejecting H_0 when it is true.
2) The confidence level of a hypothesis test is $100\,(1-\alpha)\%$.
3) The power of a statistical test $(1-\beta)$ is the probability of rejecting H_0 when it is false

Possible Hypothesis Test Outcomes		
	Actual Situation	
Decision	H_0 True	H_0 False
Do Not Reject H_0	No Error Probability $1-\alpha$	Type II Error Probability β
Reject H_0	Type I Error Probability α	No Error Probability $1-\beta$

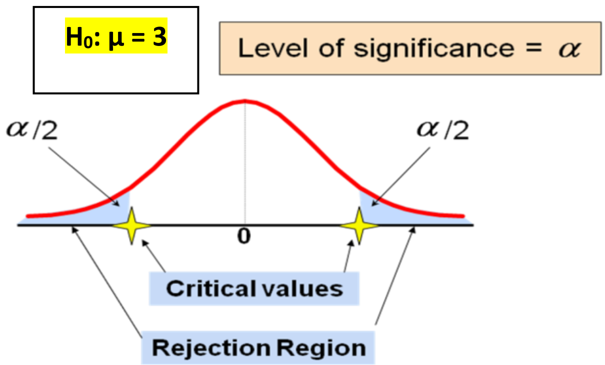
5.2.1 Relationship between Type I & Type II Error

Type I and Type II errors cannot happen at the same time
- A Type I error can only occur if H_0 is true
- A Type II error can only occur if H_0 is false

If Type I error probability (α) increases, then Type II error probability (β) decreases

5.2.3 Level of Significance and the Rejection Region

Critical region: Set of all values which would cause us to reject H_0
Critical value(s): The value(s) which separate the critical region from the non-critical region. The critical values are determined independently of the sample statistics.



This is a two-tail test because there is a rejection region in both tails

Test statistic: Sample statistic used to decide whether to reject or fail to reject the null hypothesis
Probability Value (P-value): The probability of getting the results obtained if the null hypothesis is true. If this probability is too small (smaller than the level of significance), then we reject the null hypothesis. If the level of significance is the area beyond the critical values, then the probability value is the area beyond the test statistic.
Decision: A statement based upon the null hypothesis. It is either "reject the null hypothesis" or "fail to reject the null hypothesis". We will never accept the null hypothesis.

Conclusion: A statement which indicates the level of evidence (sufficient or insufficient), at what level of significance, and whether the original claim is rejected (null) or supported (alternative).

5.2.4 Steps in Hypothesis Testing

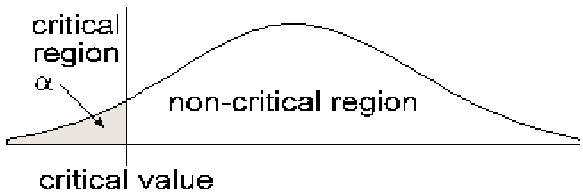
Any hypothesis testing is done under the assumption that the null hypothesis is true.
Here are the steps to performing hypothesis testing

a) Write the null and alternative hypothesis.
b) Use the alternative hypothesis to identify the type of test.
c) specify the level of significance, α and find the critical value using the tables
d) Compute the test statistic
e) Make a decision to reject or fail to reject the null hypothesis.
f) Write the conclusion

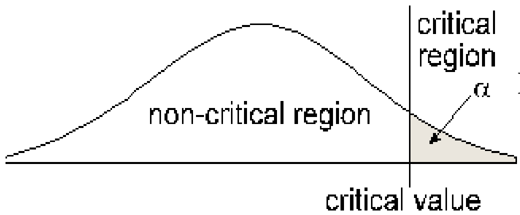
Remarks

The first thing to do when given a claim is to write the claim mathematically (if possible), and decide whether the given claim is the null or alternative hypothesis. If the given claim contains equality, or a statement of no change from the given or accepted condition, then it is the null hypothesis, otherwise, if it represents change, it is the alternative hypothesis.
The type of test is determined by the *Alternative Hypothesis* (H_1)

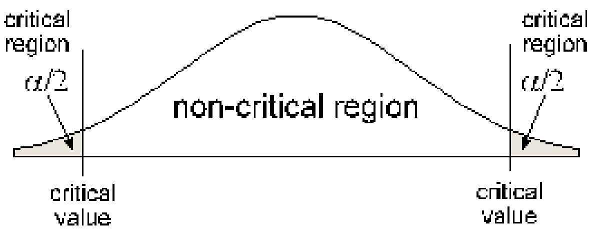
Left Tailed Test
H_1: parameter < value
Notice the inequality points to the left
Decision Rule: Reject H_0 if t.s. < c.v.



Right Tailed Test
H_1: parameter > value
Notice the inequality points to the right
Decision Rule: Reject H_0 if t.s. > c.v.



Two Tailed Test
H_1: parameter **not equal** to a value
Notice the inequality points to both sides
Decision Rule: Reject H_0 if t.s. < c.v. (left) or t.s. > c.v. (right)



If the test statistic falls into the non rejection region, do not reject the null hypothesis H_0. If the test statistic falls into the rejection region, reject the null hypothesis. Express the managerial conclusion in the context of the problem
Conclusions are sentence answers which include whether there is enough evidence or not (based on the decision) and whether the original claim is supported or rejected. Conclusions are based on the original claim, which may be the null or alternative hypotheses.

5.3 Approaches to Hypothesis Testing

There are three approaches to hypothesis testing namely Classical Approach, p vale approach and the confidence interval approach

5.3.1 The Classical Approach

The Classical Approach to hypothesis testing is to compare a test statistic and a critical value. It is best used for distributions which give areas and require you to look up the critical value (like the Student's t distribution) rather than distributions which have you look up a test statistic to find an area (like the normal distribution).

The Classical Approach also has three different decision rules, depending on whether it is a left tail, right tail, or two tail test.

One problem with the Classical Approach is that if a different level of significance is desired, a different critical value must be read from the table.

5.3.2 P-Value Approach

The P-Value Approach, short for Probability Value, approaches hypothesis testing from a different manner. Instead of comparing z-scores or t-scores as in the classical approach, you're comparing probabilities, or areas.

The level of significance (alpha) is the area in the critical region. That is, the area in the tails to the right or left of the critical values.

The p-value is the area to the right or left of the test statistic. If it is a two tail test, then look up the probability in one tail and double it.

If the test statistic is in the critical region, then the p-value will be less than the level of significance. It does not matter whether it is a left tail, right tail, or two tail test. This rule always holds.

Reject the null hypothesis if the p-value is less than the level of significance.

You will fail to reject the null hypothesis if the p-value is greater than or equal to the level of significance.

The p-value approach is best suited for the normal distribution when doing calculations by hand. However, many statistical packages will give the p-value but not the critical value. This is because it is easier for a computer or calculator to find the probability than it is to find the critical value.

Another benefit of the p-value is that the statistician immediately knows at what level the testing becomes significant. That is, a p-value of 0.06 would be rejected at an 0.10 level of significance, but it would fail to reject at an 0.05 level of significance. Warning: Do not decide on the level of significance after calculating the test statistic and finding the p-value.

Here are a couple of statements to help you keep the level of significance the probability value straight.

The Level of Significance is pre-determined before taking the sample. It **does not** depend on the sample at all. It is the area in the critical region,that is the area beyond the **c**ritical values. It is the probability at which we consider something unusua**l**.

The Probability-Value can only be found after taking the sample. It depends on the sample. It is the area beyond the test statistic**.** It is the probability of getting the results we obtained if the null hypothesis is true.

5.3.3 Confidence Intervals as Hypothesis Tests

Using the confidence interval to perform a hypothesis test only works with a two-tailed test.

a) If the hypothesized value of the parameter lies within the confidence interval with a 1-alpha level of confidence, then the decision at an alpha level of significance is to fail to reject the null hypothesis.

b) If the hypothesized value of the parameter lies outside the confidence interval with a 1-alpha level of confidence, then the decision at an alpha level of significance is to reject the null hypothesis.

However, it has a couple of problems.

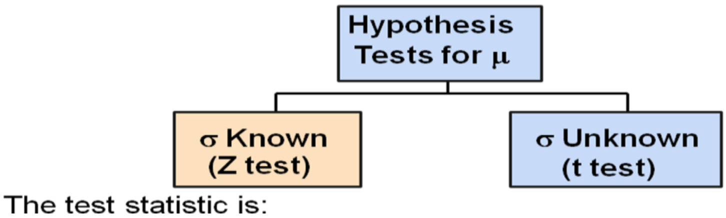
- It only works with two-tail hypothesis tests.
- It requires that you compute the confidence interval first. This involves taking a z-score or t-score and converting it into an x-score, which is more difficult than standardizing an x-score.

5.4 Testing a Single Mean

The value for all population parameters in the test statistics come from the null hypothesis. This is true not only for means, but all of the testing we're going to be doing.

The following hypotheses are to be tested: $\mathrm{H}_0 : \mu = \mu_0$ vs $\mathrm{H}_1 : \mu \neq \mu_0$ 0r $\mathrm{H}_0 : \mu > \mu_0$ 0r $\mathrm{H}_0 : \mu < \mu_0$ Where μ_0 is some hypothesised value.

The statistic and the critical values depends on whether **σ**, is known or unknown.



5.4.1 Population Standard Deviation Known

If the population standard deviation **σ**, is known, then the population mean has a normal distribution, and you will be using the z-score formula for sample means. The test statistic is the standard formula you've seen before. $z = \dfrac{\bar{x} - \mu}{\sigma / \sqrt{n}}$

The critical value is obtained from the normal table.

Example Test at 5% level the claim that the true mean # of TV sets in US homes is equal to 3. Suppose the sample results are n = 100, $\bar{x} = 2.84$ (σ = 0.8 is assumed known**)**

Solution

State the appropriate null and alternative hypotheses

- H_0: μ = 3 H_1: μ ≠ 3 (This is a two-tail test)

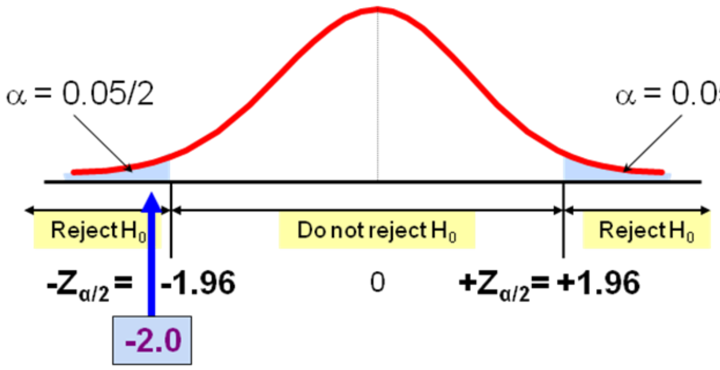
Determine the appropriate technique

- σ is assumed known so this is a Z test.

Determine the critical values

- For α = 0.05 the critical Z values are ±1.96

Compute the test statistic Z_{STAT} so the test statistic is: $Z_{STAT} = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} = \frac{2.84 - 3}{0.8/\sqrt{100}} = -2.0$



Since $Z_{STAT} = -2.0 < -1.96$, reject the null hypothesis and conclude there is sufficient evidence that the mean number of TVs in US homes is not equal to 3

Exercise

1. A simple random sample of 10 people from a certain population has a mean age of 27. Can we conclude that the mean age of the population is less than 30? The variance is known to be 20. Let α = .05.
2. Bon Air Elementary School has 300 students. The principal of the school thinks that the average IQ of students at Bon Air is at least 110. To prove her point, she administers an IQ test to 20 randomly selected students. Among the sampled students, the average IQ is 108. Assuming variance is known to be 100, should the principal accept or reject her original hypothesis? at 5% level of significance
3. Central bank believes that if consumer confidence is too high, the economy risks over heating. Low confidence is a warning that rcession might be on the way. In either case, the bank may choose to intervene by altering interest rates. The ideal value for the bank's chosen measure is 50. We may assume the measure is normally distributed with standard deviation 10. The bank takes a survey of 25 people. Which returned a sample mean of 54 for the index. What would you advice the bank to do? Use α = .05.
4. A manager will switch to a new technology if the production process exceeds 80 units per hour. The manager asks the company statistician to test the null hypothesis: H_0: $\mu = 80$ against the alternative hypothesis: H_1: $\mu > 80$ If there is strong evidence to reject the null hypothesis then the new technology will be adopted. Past experience has shown that the standard deviation is 8. A data set with n = 25 for the new technology has a sample mean of 83 Does this justify adoption of the new technology?

5.4.2 Population Standard Deviation Unknown

If the population standard deviation **σ**, is unknown, then the population mean has a student's t distribution, and you will be using the t-score formula for sample means. The test statistic is very similar to that for the z-score, except that sigma has been replaced by s and z has been replaced by t. ie $t = \dfrac{\bar{x} - \mu}{s/\sqrt{n}}$

The critical value is obtained from the t-table. The degrees of freedom is n-1.

Example 1 A fertilizer mixing machine is set to give 12 kg of nitrate for every 100kg bag of fertilizer. Ten 100kg bags are examined. The percentages of nitrate are as follows: 11, 14, 13, 12, 13, 12, 13, 14, 11, 12. Is there reason to believe that the machine is defective at 5% level of significance?

Solution

Hypothesis H$_0$: $\mu = 12$ H$_1$: $\mu \neq 12$ (This is a two-tail test)

σ is unknown so this is a t test. Use the unbiased estimator ie $s = \sqrt{\dfrac{\sum(x-\bar{x})^2}{n-1}}$

Critical Region based on $\alpha = 0.05$ and 9 degrees freedom

$t_{9,0.025} = 2.262$ ie reject H$_0$: $\mu = 12$ if $|t_c| \geq 2.262$

From calculator $\bar{x} = 12.5$ and $s = 1.0801$

Test statistic $t_c = \dfrac{\bar{x}-\mu}{s/\sqrt{n}} = \dfrac{12.5-12}{1.0801/\sqrt{10}} = 1.4639$

Decision since $|t_c| = 1.2639 < 2.262$, we fail to reject H$_0$ and conclude that the machine is not defective.

Example 2 The following figures give the end of year profits of ten randomly selected Chemists in Nairobi county.

Profit(in million shillings)	21.8	24.8	27.3	29.3	30.8	31.8	32.8	32.5	32.1	31.3

On the basis of this data, test whether the average profit is greater than 30M KSH at 1% level of significance

Solution

Hypothesis H$_0$: $\mu = 30$ H$_1$: $\mu > 30$ (This is a 1-tail test)

σ is unknown so this is a t test. Use the unbiased estimator ie $s = \sqrt{\dfrac{\sum(x-\bar{x})^2}{n-1}}$

Critical Region based on $\alpha = 0.01$ and 9 degrees freedom

$t_{9,0.01} = 2.82$ ie reject H$_0$: $\mu = 12$ if $|t_c| \geq 2.82$

From calculator $\bar{x} = 29.415$ and $s = 3.6601$

Test statistic $t_c = \dfrac{\bar{x}-\mu}{s/\sqrt{n}} = \dfrac{29.415-30}{3.6601/\sqrt{10}} = -0.51$

Decision since $|t_c| = 0.51 < 2.82$, we don't reject H$_0$ and conclude that the average profit is not greater than 30M KSH.

Exercise

1. Identify the critical *t* value for each of the following tests:
   a. A two-tailed test with α=0.05 and 11 degrees of freedom
   b. A one-tailed test with α=0.01 and n=17
2. Consider a sample with $n = 20$ $\bar{x} = 8.0$ and $s = 2$ Do the following hypothesis tests.
   a) . H$_0$: μ =8.7 H$_1$: $\mu > 8.7$ at α=0.01
   b) H$_0$: μ =8.7 H$_1$: $\mu \neq 8.7$ at α=0.05
3. It is widely believed that the average body temperature for healthy adults is 98.6 degrees Fahrenheit. A study was conducted a few years go to examine this belief. The body temperatures of *n* = 130 healthy adults were measured (half male and half female). The average temperature from the sample was found to be $\bar{x} = 98.249$ with a standard

deviation s = 0.7332. Do these statistics contradict the belief that the average body temperature is 98.6? test at 1% level of significance

4. A study is to be done to determine if the cognitive ability of children living near a lead smelter is negatively impacted by increased exposure to lead. Suppose the average IQ for children in the United States is 100. From a pilot study, the mean and standard deviation were estimated to be $\bar{x} = 89$ and s = 14.4 respectively. Test at 5% level whether there is a negative impact.
5. The average cost of a hotel room in New York is said to be \$168 per night. To determine if this is true, a random sample of 25 hotels is taken and resulted in $\bar{x} = \$172.5$ and $s = \$15.40$. Test the appropriate hypotheses at α = 0.05.
6. A sample of eleven plants gave the following shoot lengths
   Shoot length (cm) 10.1 21.5 11.7 12.9 14.8 11.0 19.2 11.4 22.6 10.8 10.2
   An earlier study reported that the mean shoot length is 15cm. Test whether the experimental data confirms the old view at 5% level of significance.
7. A simple random sample of 14 people from a certain population gives body mass indices as shown in Table 7.2.1. Can we conclude that the BMI is not 35? Let α = .05.

subject	1	2	3	4	5	6	7	8	9	10	11	12	13	14
BM	23	25	21	37	39	21	23	24	32	57	23	26	31	45

8. A company selling licenses for a franchise operation claims that, in the first year, the yield on an initial investment is 10%. How should a hypothesis test be stated? If there is strong evidence that the mean return on the investment is below 10% this will give a cautionary warning to a potential investor. Therefore, test the null hypothesis: H0 : μ = 10 against the alternative hypothesis: H1 : μ < 10 From a sample of n = 10 observations, the sample statistics are: $\bar{x} = 8.82$ x and s = 2.40
9. We know the distance that an athlete can jump is normally distributed but we do not know the standard deviation. We record 15 jumps: 7.48 7.34 7.97 5.88 7.48 7.67 7.49 7.48 8.51 5.79 7.13 6.80 6.19 6.95 5.93 Test whether these values are consistent with a mean jump length of 7m. Do you have any reservations about this test?
10. The manufacturing process should give a weight of 20 ounces. Does the data show evidence that the process is operating correctly? Test the null hypothesis: H0 : μ = 20 the process is operating correctly against the alternative: H1 : μ ¹ 20 the process is not operating correctly From the data set, the sample statistics are: n = 9, $\bar{x} = 20.356$ (ounces) and s = 0.6126