UNIT – 1 INFERENTIAL STATISTICS

QUESTION BANK

PART - A

1. Define the term Population by giving relevant example .

Ans : Any complete set of observations (or potential observations) may be characterized as a population.

Example: population might be described as “attitudes toward abortion of currently Enrolled students at Bucknell University” or as “SAT critical reading scores of currently Enrolled students at Rutgers University.

2. What are the different types of population ?

Ans : The types of population are,

(i) REAL POPULATION and
(ii) HYPOTHETICAL POPULATION.

3. Explain the concept sample ?

Ans : Any subset of observations from a population may be characterized as a sample. In typical

applications of inferential statistics, the sample size is small relative to the population size.

Example : less than 1 percent of all U.S. worksites are included in the Bureau of Labor Statistics’ monthly survey to estimate the rate of unemployment. And although, only 1475 likely voters had been sampled in the final poll for the 2012 presidential election by the NBC News/Wall Street Journal.

4. What do you mean by the term Optimal sample size ?

Ans : Optimal sample size depends on the answers to a number of questions, including “What is the estimated variability among observations?” and “What is an acceptable amount of error in our conclusion?”

5. Explain Random sampling with appropriate example

Ans : Random sampling occurs if, at each stage of sampling, the selection process guarantees that all potential observations in the population have an equal chance of being included in the sample.

Example : The names of 25 employees being chosen out of a hat from a company of 250 employees. In this case, the population is all 250 employees, and the sample is random because each employee has an equal chance of being chosen.

6. Why casual or hazapard sample are not random? give reason.

Ans : A casual or haphazard sample doesn't qualify as a random sample because for instance in the names of 25 employees being chosen out of a hat from a company of 250 employees, not everyone as the equal chance of being sampled, if the company chooses only the fresher's.

7. Define probability and statistics .

Ans :

Probability:

Probability refers to the proportion or fraction of times that a particular event Is likely to occur.

Example : the probability of flipping a coin and it being heads is ½, because there is 1 way of getting a head and the total number of possible outcomes is 2 (a head or tail). We write P(heads) = ½ .

Statistics:

Statistics is the discipline that concerns the collection, organization, analysis, interpretation, and presentation of data. In applying statistics to a scientific, industrial, or social problem.

Example : a report of numbers saying how many followers of each religion there are in a particular country.

8. Define the terms i). Common outcome

Ii). Rare outcome .

Ans : (i). Common outcomes signify, most generally, a lack of evidence that something Special has occurred, in the language of statistics, it lacks statistical significance.

(ii). Rare outcomes are outcomes signify that something special has occurred and it has statistical significance.

9. What is common or rare ?

Ans As an aid to determining whether observed results should be viewed as common or Rare, statisticians, we interpret different proportions of area under theoretical curves, such as the normal curve as probabilities of random outcomes is called common or rare.

10. What is sampling distribution?

Ans : sampling distribution of the mean refers to the probability distribution of Means for all possible random samples of a given size from some population.

11. List the properties of mean of the sample mean.

Ans : 1. The sample mean is a statistic obtained by calculating the arithmetic average of the values of a variable in a sample**.**

2. If the sample is drawn from probability distributions having a common expected value, then the sample mean is an estimator of that expected value.

12. Derive the formula of mean of a sampling distribution .

Ans : The mean of the sampling distribution of the mean always equals the mean of the Population.

MEAN OF SAMPLING DISTRUBTION

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

Where µX represents the mean of the sampling distribution and µ represents the mean of the Population.

13. Explain standard error of the mean.

Ans : The distribution of sample means also has a standard deviation, referred to as the standard error of the mean.

The standard error of the mean equals the standard deviation of the population divided by the square root of the sample size.

$$\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}}$$

Where X represents the standard error of the mean; σ represents the standard deviation of the population; And n represents the sample size.

14. What are special type of standard deviations?

Ans : The standard error of the mean serves as a special type of standard deviation that measures Variability in the sampling distribution.

It supplies us with a standard, much like a yardstick, that describes the amount by which Sample means deviate from the mean of the sampling distribution or from the population Mean.

The error in standard error refers not to computational errors, but to errors in generalizations Attributable to the fact that, just by chance, most random samples aren't exact replicas of the Population.

15. Elucidate the effect of sample size.

Ans : A modest demonstration of this effect where the means of all possible samples cluster closer to the population mean (equal to 3.5) than do the four original observations in the population. A more dramatic demonstration occurs with larger sample size.

$$\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}} = \frac{110}{\sqrt{100}} = \frac{110}{10} = 11$$

16. Define Hypothesis testing.

Ans : A statistical hypothesis test is a method of statistical inference used to decide whether the data at hand sufficiently support a particular hypothesis.

17. Explain hypothesized sampling.

Ans : If the null hypothesis is true, then the distribution of sample means—that is, the sampling Distribution of the mean for all possible random samples, each of size 100, from the local

Population of freshmen—will be centered about the national average of 500.

This sampling distribution is referred to as the hypothesized sampling distribution, since its mean equals 500, the hypothesized mean reading score for the local population of freshmen.

18. Define the term Z-test .

Ans : A Z-test is any statistical test for which the distribution of the test statistic under the null hypothesis can be approximated by a normal distribution. Z-tests test the mean of a distribution

19. Give the formula for Z-test for population mean.

Ans : Z Test = $(\bar{x} – \mu) / (\sigma / \sqrt{n})$

20. Derive the ratio for single population mean.

Ans :

Where z indicates the deviation of the observed sample mean in standard error units, above or below

The hypothesized population mean.

$$z = \frac{\overline{X} - \mu_{\text{hyp}}}{\sigma_{\overline{X}}}$$

21. List the assumptions of a Z -teat.

Ans : When a hypothesis test evaluates how far the observed sample mean deviates, in standard error units, from the hypothesized population mean, as in the present example, it is referred to as a z test or, more accurately, as a z test for a population mean.

This z test is accurate only when (1) the population is normally distributed or the sample size is large enough to satisfy the requirements of the central limit theorem and (2) the population standard deviation is known.

22. Define Null hypothesis (H0) with example.

Ans : A null hypothesis is a type of statistical hypothesis that proposes that no statistical significance exists in a set of given observations or in simple words the null hypothesis is that two possibilities are the same.

Example : Rohan will win at least Rs.100000 in lucky draw.

23. Define alternative Hypothesis (H1) with example.

Ans : An alternative hypothesis, states that there is statistical significance between two variables. It is denoted by H1 or Ha.

Example : Rohan will win less than Rs.100000 in lucky draw.

24. What is decision rule ?

Ans : A decision rule is a fun A decision rule specifies precisely when H0 should be rejected a decision rule specifies precisely when H0 should be rejected

25. What is level of significance alpha ?

Ans : Total area that is identified with rare outcomes. Often referred to as the level of significance of

the statistical test, this proportion is symbolized by the Greek letter α (alpha).

26. What is the significance of statistics ?

Ans : Statistical significance is a determination made by an analyst that the results in the data are not explainable by chance alone or Statistical significance refers to the claim that a result from data generated by testing or experimentation is not likely to occur randomly or by chance but is instead likely to be attributable to a specific cause.

27. Explain probability sampling with relevant example.

Ans : Probability sampling refers to the selection of a sample from a population, when this selection is based on the principle of randomization, that is, random selection or chance. Probability sampling is more complex, more time-consuming and usually more costly than non-probability sampling.

Example : if you had a population of 100 people, each person would have odds of 1 out of 100 of being chosen.

28. What is one of the distinct differences between a population parameter and a sample statistic?

Ans : A sample statistic changes each time you try to measure it, but a population parameter Remains fixed.

A parameter is a number describing a whole population (e.g., population mean), while a statistic is a number describing a sample (e.g., sample mean).

29What is central theorem?

Ans : In probability theory, the central limit theorem (CLT) states that the distribution of a sample variable approximates a normal distribution (i.e., a “bell curve”) as the sample size becomes larger, assuming that all samples are identical in size, and regardless of the population's actual distribution shape.

30. What is the sampling Distribution of the sample mean.

Ans : The overall shape of the distribution is symmetric and approximately normal.

There are no outliers or other important deviations from the overall pattern.

The center of the distribution is very close to the true population mean.

PART – B

1.Indicate whether the following statements are True or False. The Mean of all sample means, μ X, . . .

(a). equals the value of a particular sample mean.

(b). 100 if, in fact, the population mean equals 10

(C). Usually equals the value of a particular sample mean.

(d). interchangeable with the population mean.

Ans

(a) False. It always equals the value of the population mean.
(b) True

(c) False. Because of chance, most sample means tend to be either larger or smaller Than the mean of all sample means.

(d) as it is interchangeable with the population mean.

2. Indicate whether the following statements are True or False. The Standard error of the mean, X , . . .

(a) measures the average amount by which sample means deviate from the population Mean.

(b) Measures variability in a particular sample.

(c) Increase in value with larger sample sizes.

(d) Equals 5, given that $\sigma = 40$ and $n = 64$.

Ans:

(a) True measures the average amount by which sample means deviate from the population mean.
(b) False. It measures variability among sample means
(c) False. It decreases in value with larger sample sizes.
(d) True as

$$\mathrm{SE} = \frac{\sigma}{\sqrt{n}}$$

3. Indicate whether the following statements are True or False. The central limit theorem

(a) States that, with sufficiently large sample sizes, the shape of the population is normal.
(b) States that, regardless of sample size, the shape of the sampling distribution of the mean is Normal.

(C) Ensures that the shape of the sampling distribution of the mean equals the shape of the Population.

(d) to the shape of the sampling distribution—not to the shape of the population and not to The shape of the sample.

Ans:

(a) False. The shape of the population remains the same regardless of sample size.
(b) False. It requires that the sample size be sufficiently large—usually between 25 And 100.

(C) False. It ensures that the shape of the sampling distribution approximates a normal Curve, regardless of the shape of the population (which remains intact).

(d) true According to the central limit theorem, regardless of the shape of the population, The shape of the sampling distribution approximates a normal curve if the sample Size is sufficiently large.

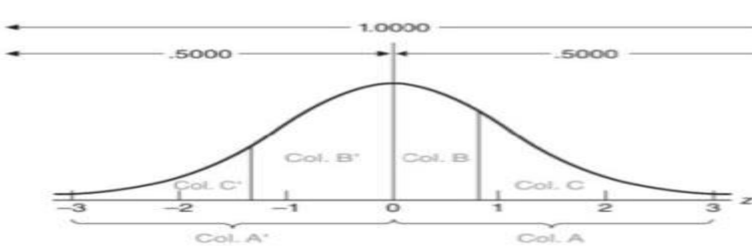
4. Referring to the standard normal table (Table A, Appendix C), find The probability that a randomly selected z score will be

(a) 1.96

(b) above 1.96 or below –1.96

(C) between –1.96 and 1.96

(d) either above 2.58 or below –2.58.



Ans:

(a) .0250
(b) .0250 + .0250 = .0500
(c) .4750 + .4750 = .9500
(d) .0049 + .0049 = .0098

5 The probability of a boy being born equals .50, or ½, as does the probability of a girl Being born. For a randomly selected family with two children, what's the probability of

(a) Two boys, that is, a boy and a boy? (Reminder: Before using either the addition or Multiplication rule, satisfy yourself that the various events are either mutually exclusive Or independent, respectively.)
(b) Two girls?
(c) Either two boys or two girls?

Ans :

(a) $\left(\frac{1}{2}\right)\left(\frac{1}{2}\right)=\frac{1}{4}$

(b) $\left(\frac{1}{2}\right)\left(\frac{1}{2}\right)=\frac{1}{4}$

(c) $\left(\frac{1}{4}\right)+\left(\frac{1}{4}\right)=\frac{2}{4}$

6.. imagine a very simple population consisting of only five observations: 2, 4, 6, 8, 10.

(a) list all possible samples of size two.

(b) construct a relative frequency table showing the sampling distribution of the mean.

Ans :

(a)

(1) 2,2	(6) 4,2	(11) 6,2	(16) 8,2	(21) 10,2
(2) 2,4	(7) 4,4	(12) 6,4	(17) 8,4	(22) 10,4
(3) 2,6	(8) 4,6	(13) 6,6	(18) 8,6	(23) 10,6
(4) 2,8	(9) 4,8	(14) 6,8	(19) 8,8	(24) 10,8
(5) 2,10	(10) 4,10	(15) 6,10	(20) 8,10	(25) 10,10

(b)

$\overline{X}$	PROBABILITY
10	1/25
9	2/25
8	3/25
7	4/25
6	5/25
5	4/25
4	3/25
3	2/25
2	1/25

7.Indicate whether each of the following statements is True or False. A random selection of 10 playing cards from a deck of 52 cards implies that

(a) random sample of 10 cards accurately represents the important features of the Whole deck.

(b) card in the deck has an equal chance of being selected.

(c) it is impossible to get 10 cards from the same suit (for example, 10 hearts).

(d) Any outcome, however unlikely, is possible.

Ans:

(a)False. Sometimes, just by chance, a random sample of 10 cards fails to represent The important features of the whole deck.

(b). True they have a good and equal chance of getting selected.

(c) False. Although unlikely, 10 hearts could appear in a random sample of 10 cards.

(d) True

8 .Describe how you would use the table of random numbers to take

(a) random sample of five statistics students in a classroom where each of nine rows Consists of nine seats.

(b) random sample of size 40 from a large directory consisting of 3041 pages, with 480 Lines per page.

Ans:

(a)there are many ways. For instance, consult the tables of random numbers, using The first digit of each 5-digit random number to identify the row (previously labeled 1, 2, 3, and so on), and the second digit of the same random number to locate a Particular student's seat within that row. Repeat this process until five students Have been identified. (If the classroom is larger, use additional digits so that every Students can be sampled.)

(b)once again, there are many ways. For instance, use the initial 4 digits of each random Number (between 0001 and 3041) to identify the page number of the telephone directory and the next 3 digits (between 001 and 480) to identify the particular line on that page. Repeat the Process, using 7-digit numbers, until 40 telephone numbers have been identified.

9 .Assume that 12 subjects arrive, one at a time, to participate in an Experiment. Use random numbers to assign these subjects in equal numbers to group A And group B. Specifically, random numbers should be used to identify the first subject as Either A or B, the second subject as either A or B, and so forth, until all subjects have Been identified.There should be six subjects identified with A and six with B.
(a) formulate an acceptable rule for single-digit random numbers. Incorporate into this Rule a procedure that will ensure equal numbers of subjects in the two groups. Check Your answer in Appendix B before proceeding.
(b) from left to right in the top row of the random number page (Table H, Appendix C), use the random digits of each random number in conjunction with your Assignment rule to determine whether the first subject is A or B, and so forth. List the Assignment for each Subject.
Ans :
(a) For instance, if the first digit is odd (1, 3, 5, 7, or 9), the first subject is assigned to group A, and if the first digit is even (0, 2, 4, 6, or 8), the first subject is assigned to group B. To ensure equal groups, the second subject is assigned automatically to **t**he group opposite that for the first subject. Repeat this procedure for the remaining five pairs of subjects. There are other acceptable rules, all involving pairs of subjects (to ensure **e**qual group sizes). For instance, if the first digit equals 0, 1, 2, 3, or 4, the first subject is assigned to group A; otherwise, the first subject is assigned to group B, and so on.
(b) Answer shows two possible assignment rules. In practice only one assignment rule actually would be used.

SUBJECT#	RANDOM NUMBER (TOP ROW, TABLE H)	ASSIGNMENT RULE 1*	OR	ASSIGNMENT RULE 2**
1	1	A		A
2	–	automatically B		automatically B
3	0	B		A
4	–	automatically A		automatically B
5	0	B		A
6	–	automatically A		automatically B
7	9	A		B
8	–	automatically B		automatically A
9	7	A		B
10	–	automatically B		automatically A
11	3	A		A
12	–	automatically B		automatically B

**Odd digits = group A; even digits = group B.*
***Digits 0, 1, 2, 3, 4 = group A; digits 5, 6, 7, 8, 9 = group B.*

10.. Assuming that people are equally likely to be born during any one of the months, what is the probability of Jack being born during
(a) June?
(b) Any month other than June?
(C)either May or June?
Ans:
(a) 1 /12
(b) 11 /12
(c) 2/ 12

11. Assuming that people are equally likely to be born during any of the months, and also assuming (possibly over the objections of astrology fans) that the birthdays of married couples are independent, what's the probability of

(a) The husband being born during January and the wife being born during February?

(b) Both husband and wife being born during December?

(C) both husband and wife being born during the spring (April or May)? (Hint: First, find the probability of just one person being born during April or May.

Ans:

(a) 1/12 * 1/12= 1/144

(b) 1/12 * 1/12 = 1/144

(c) 2/12 * 2/12 = 4/144.

12.. Among 100 couples who had undergone marital counseling, 60 couples described their relationships as improved, and among this latter group, 45 couples had children. The remaining couples described their relationships as unimproved, and among this group, 5 couples had children. (Hint: Using a frequency analysis, begin with the 100 couples, first branch into the number of couples with improved and unimproved relationships, then under each of these numbers, branch into the number of couples with children and without children. Enter a number at each point of the diagram before proceeding.)

(a) what is the probability of randomly selecting a couple who described their relationship as improved?

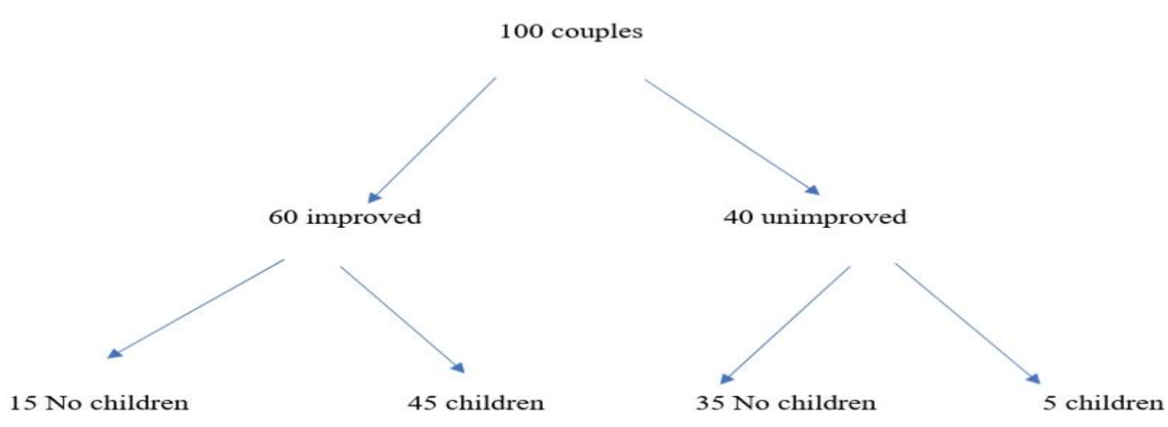
(b) What is the probability of randomly selecting a couple with children?

(c) What is the conditional probability of randomly selecting a couple with children, given that their relationship was described as improved?

(d) what is the conditional probability of randomly selecting a couple without children, given that their relationship was described as not improved?

(e) What is the conditional probability of an improved relationship, given that a couple has children?

Ans :



(a)60 /100 = .60

(b)45 +5 /100 = 50/100 = .50

(c) 45 /60 = .75

(d)35 /40 = .875
(e) 45/45+5 = 45/50 = .90

13. Calculate the value of the z test for each of the following situations:

(a) $\overline{X} = 566; \sigma = 30; n = 36; \mu_{\text{hyp}} = 560$
(b) $\overline{X} = 24; \sigma = 4; n = 64; \mu_{\text{hyp}} = 25$
(c) $\overline{X} = 82; \sigma = 14; n = 49; \mu_{\text{hyp}} = 75$
(d) $\overline{X} = 136; \sigma = 15; n = 25; \mu_{\text{hyp}} = 146$

Ans :

(a) $z = \dfrac{566 - 560}{30/\sqrt{36}} = \dfrac{6}{5} = 1.20$

(b) $z = \dfrac{24 - 25}{4/\sqrt{64}} = \dfrac{-1}{.5} = -2.00$

(c) $z = \dfrac{82 - 75}{14/\sqrt{49}} = \dfrac{7}{2} = 3.50$

(d) $z = \dfrac{136 - 146}{15/\sqrt{25}} = \dfrac{-10}{3} = -3.33$

14. Indicate what’s wrong with each of the following statistical Hypothesis:

(a) $H_0: \mu = 155$
$H_1: \mu \neq 160$

(b) $H_0: \overline{X} = 241$
$H_1: \overline{X} \neq 241$

Ans :

(a) Different numbers appear in H_0 and H_1.
(b) Sample means (rather than population means) appear in H_0 and H_1

15. First using words, then symbols, identify the null hypothesis for Each of the following situations. (Don’t concern yourself about the precise form of the alternative hypothesis at this point.)
(a) A school administrator wishes to determine whether sixth-grade boys in her school district Differ, on average, from the national norms of 10.2 pushups for sixth-grade boys.

(b) A consumer group investigates whether, on average, the true weights of packages of Ground Beef sold by a large supermarket chain differ from the specified 16 ounces.

(c) marriage counselor wishes to determine whether, during a standard conflict-resolution Session, his

(a) Retain H_0 at the .05 level of significance because $z = 1.74$ is less positive than 1.96.
(b) Retain H_0 at the .05 level of significance because $z = 0.13$ is less positive than 1.96.
(c) Reject H_0 at the .05 level of significance because $z = -2.51$ is more negative than -1.96.

clients differ, on average, from the 11 verbal interruptions reported for "well- Adjusted couples."

Ans :

(a) Sixth-grade boys in her school district average 10.2 pushups. $H_0: \mu = 10.2$
(b) On average, weights of packages of ground beef sold by a large supermarket chain equal 16 ounces. $H_0: \mu = 16$
(c) The marriage counselor's clients average 11 interruptions per session. $H_0: \mu = 11$

16. For each of the following situations, indicate whether H0 Should Be retained or rejected and Justify your answer by specifying the precise relationship between Observed and critical z Scores. Should H0 Be retained or rejected, given a hypothesis test with Critical z scores of ± 1.96 ans

(a) z = 1.74 (b) z = 0.13 (c) z = –2.51.

Ans :

17. According to the American Psychological Association, members With a doctorate and a full-time Teaching appointment earn, on the average, $82,500 per year, With a standard deviation of $6,000. An investigator wishes to determine whether $82,500 is Also the mean salary for all female Members with a doctorate and a full-time teaching appointment. Salaries are obtained for a random Sample of 100 women from this population, and the Mean salary equals $80,100.

(a) Someone claims that the observed difference between $80,100 and $82,500 is large Enough by Itself to support the conclusion that female members earn less than male Members. Explain why It is important to conduct a hypothesis test.

(b) investigator wishes to conduct a hypothesis test for what population?

(C) What is the null hypothesis, H0

(d) is the alternative hypothesis, H1?

(e) Specify the decision rule, using the .05 level of significance.

(f) Calculate the value of z. (Remember to convert the standard deviation to a standard error.)

(g) What is your decision about H0?

(h) Using words, interpret this decision in terms of the original problem.

Ans :

(a) The observed difference between $80,100 and $82,500 cannot be interpreted at face value, as it could have happened just by chance. A hypothesis test permits us to evaluate the effect of chance by measuring the observed difference relative to the standard error of the mean.

(b) All female members of the APA with a Ph.D. degree and a full-time teaching appointment.

(c) H0: µ = 82,500

(d) H1: µ 82,500

(e) Reject H0 at the .05 level of significance if z 1.96 or z –1.96

$$\textbf{(f)}\quad z=\frac{80{,}000-82{,}500}{\frac{6000}{\sqrt{100}}}=\frac{-2{,}400}{600}=-4.00$$

(g) Reject H0 at the .05 level of significance because z = –4.00 is more negative than –1.96.
(h) The average salary of all female APA members (with a Ph.D. and a full-time teaching appointment) is less than $82,5face value, as it could have happened just by chance. A hypothesis test permits us to evaluate the effect of chance by measuring the observed difference relative to the standard error of the mean.

18. For the population at large, the Wechsler Adult Intelligence Scale is designed to yield A Normal distribution of test scores with a mean of 100 and a standard deviation of 15. School district officials wonder whether, on the average, an IQ score different From 100 Describes the intellectual aptitudes of all students in their district. Wechsler IQ scores Are obtained for a random sample of 25 of their students, and the mean IQ Is found to Equal 105. Using the step-by-step procedure described in this chapter, Test the null Hypothesis at the .05 level of significance.

Ans : Research Problem
Does the mean IQ of all students in the district differ from 100?
Statistical Hypotheses
H0: µ = 100
H1: µ 100
Decision Rule
Reject H0 at the .05 level of significance if z equals or is more positive than 1.96 or if z equals or is more negative than –1.96.
Calculations

$$\text{Given that } \overline{X}=105;\quad \sigma_{\overline{X}}=\frac{15}{\sqrt{25}}=\frac{15}{5}=3$$

$$z=\frac{105-100}{3}=\frac{5}{3}=1.67$$

Decision
Retain H0 at the .05 level of significance because z = 1.67 is less positive than 1.96.
Interpretation
There is no evidence that the mean IQ of all students differs frames the mean IQ of all students in the district differ from 100?
IQ of all students differs from 100.