GATE

CS & DA

DPP: 01

Probability and Statistics

Sampling Theory & Distribution

Q1 If a sample of 400 male workers has a mean height of 67.47 inches, is it reasonable to regard the sample as a sample from a large population with a mean height of 67.39 inches and a standard deviation of 1.30 inches at a 5% level of significance?

Q2 A quality engineer wants to check whether there is a difference between population and sample proportion for the rejection rate for parts manufactured on a production line. The rejected part proportion is 5% during production, whereas it was 8% when we selected 50 random samples.
Is this difference in proportion statistically significant if the Level of Significance is 0.05?

Q3 Let's say you're testing two flu drugs A and B. Drug A works on 41 people out of a sample of 195. Drug B works on 351 people in a sample of 605. Are the two drugs comparable? Use a 5% alpha level.

Q4 The mean weekly sales of soap bars in departmental stores was 146.3 bars per store. After an advertising campaign the mean weekly sales in 22 stores for a typical week increased to 153.7 and showed a standard deviation of 17.2. Was theadvertising campaign successful ?

Q5 A random sample of 10 boys had the following I.Q.'s: 70, 120, 110, 101, 88, 83, 95, 98, 107, 100. Do these data support the assumption of a population mean I.Q. of 100? Find a reasonable range in which most of the mean I.Q. values of samples of 10 boys lie.

Q6 Samples of two types of electric light bulbs were tested for length of life and following data were obtained:

	Type I	Type II
Sample No.	$n_1 = 8$	$n_2 = 7$
Sample Means	$\bar{x}_1 = 1,234$ hrs	$\bar{x}_2 = 1,234$ hrs.
Sample S.D.'s	$s_1 = 36$ hrs	$s_2 = 40$ hrs

Is the difference in the means sufficient to warrant that type 1 is superior to type II regarding length of life ?

Q7 A random sample of size 20 from a normal population gives the sample standard deviaton of 6. Test the hypothesis that the population standard deviation is 9.

Q8 A sample of 20 observation gave a standard deviation 3.72. Is this compatible with the hypothesis that the sample is from a normal population with variance 4.35?

Q9 Weigths in kg. of 10 students are given 38, 40, 45, 53, 47, 43, 55, 48, 52, 49.
can you say that variance of distribution of weights of all students from which the above sample of 10 students was drawn is equal to 20 square kg?

Q10 A random sample of size 10 drawn from normal population gave the following values:
65, 72, 68, 74, 77, 61, 63, 69, 73, 71.
Test the hypothesis that the population variance is 32.

Q11 A dog trainer wants to know if golden retrievers and French bulldogs are equally good at learning how to skateboard. She tries to train 40 golden retrievers and 60 French bulldogs to skateboard and finds the following:

	Skateboards	Can't skateboard
Golden	20	20

retrievers		
French bulldogs	50	10

Should she reject the null hypothesis that the dog's breed is unrelated to their skateboarding ability?

a) She should reject the null hypothesis.

b) She should fail to reject the null hypothesis

Q12 A restaurant reviewer wants to know if three popular burger restaurants are equally recommended by their customers. At each of the three restaurants, he asks 25 random customers whether they would recommend the restaurant to a friend. He finds the following:

	Would recommend	Would not recommend
Tasty Burgers	10	5
Burger Prince	22	3
Burger Town	18	7

Should he reject the null hypothesis that the proportion of customers recommending the restaurant is the same for the three restaurants?

a) He should reject the null hypothesis.

b) He should fail to reject the null hypothesis,

Q13 You work at a nut factory and you're in charge of quality control. The nut factory produces a nut mix that's supposed to be 50% peanuts, 30% cashews, and 20% almonds.

To check that the nut mix proportions are acceptable, you randomly sample 1000 nuts and find the following frequencies:

Nut	Frequency
Peanuts	621
Cashew	189
Almonds	190

Should you reject the null hypothesis that the nut mix has the desired proportions of nuts?

a) I should reject the null hypothesis.

b) I should fail to reject the null hypothesis.

GATE

Answer Key

Q1	0.5
Q2	0
Q3	0
Q4	= 9.03
Q5	= 0.62
Q6	9.39
Q7	0
Q8	4.35
Q9	20
Q10	32
Q11	She should reject the null hypothesis.
Q12	b
Q13	a

GATE

Hints & Solutions

Q1 Text Solution:

Taking the null hypothesis that the mean height of the population is equal to 67.39 inches, we can write:

H_0: μ = 67.39"

H_0 μ ≠ 67.39"

x = 67.47", σ = 1.30", n = 400

Assuming the population to be normal, we can work out the test statistic z as under

$Z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$

Z = 1.231

Population Mean (μ): 67.39

Population Variance (σ^2): 1.69

Sample Mean (M) : 67.47

Sample Size (N) : 400

Z Scroe Calculations

$Z = \frac{M - \mu}{\sigma^2 / n}$

$Z = \frac{67.47 - 67.39}{\sqrt{1.69 / 400}}$

Z = 0.0799999999999983/0.065

Z = 1.23077

Significance Level:

0.01

0.05

0.10

One-tailed or two-tailed hypothesis?:

One-tailed

One-tailed or two-tailed hypothesis?:

One-tailed

Two-tailed

The value of z is 1.23077. The value of p is. 10935. The result is not significant at p <.05.

Q2 Text Solution:

Step-1: Collect the required data

Population Proportion (Po) = 0.05, Sample Proportion (P) = 0.08, number of samples (n) = 50, alpha = 0.05

Step-2: Check if we can use the z test

The next step to to check for the following prior conditions to conduct the z-test.

- Normal distribution of data: Test of normality is found ok
- All data points are independent: Yes independent samples are considered.
- Standard deviation is known: Yes, the value is given in the problem statement
- Sample Size ≥ 30: yes
- Equal sample variance: Yes

Step-3: Define null and alternative hypotheses

Null Hypothesis: There is no difference between the sample and production parts.

Alternate Hypothesis: Sample and production parts are different.

Step-4: Finalize alpha

Considering cur application, we are considering alpha = 0.05

Step-5: Calculate the z-statistic

Z-Statistics = $\frac{P - Po}{Po1 - Po / n} = \frac{0.08 - 0.05}{0.05 * 1 - 0.05 / 50}$

Z-Statistics $\frac{0.03}{\sqrt{0.000095}} = \frac{0.03}{0.0308} = 0.974$

Step-6: Calculate the critical value

p-value for z-statistic(0.974) = 0.330

Step-7: Evaluate the results

Since:

Calculated p-value (0.330) > alpha (0.05)

The test results are not statiscally significant and we can not reject the null hypothesis.

Q3 Text Solution:

Step 1: Find the two proportions:

- $P_1 = \frac{41}{195}$ =0.21(that's prime s 21%)
- $P_2 = \frac{351}{605}$ = 0.58 that's 58%).

Step 2: Find the overall sample proportion. The numerator will be the total number of "positive" results for the two samples and the denominator is the total number of people in the two samples.

- $p = \frac{41 + 351}{195 + 605} = 0.49 = 0.49$

Set this number aside for a moment.

Step 3: Insert the numbers from Step 1 and Step 2 into the test statistic formula:

$$z = \frac{\hat{p}_1 - \hat{p}_2 - 0}{\sqrt{\hat{p}1 - \hat{p}\frac{1}{n_1} + \frac{1}{n_2}}}$$

$$z = \frac{0.58 - 0.21 - 0}{\sqrt{0.491 - 0.49\frac{1}{195} + \frac{1}{605}}}$$

Solving the formula, we get:

Z = 8.99

We need to find if the z-score falls into the "rejection region".

Step 4: Find the z-score associated with $\frac{\alpha}{2}$. I will use teh following table of known values:

Confindence level	Alpha	Alpha/2	zalpha/2
90%	10%	5.0%	1.645
95%	5%	2.5%	1.96
98%	2%	1.0%	2.326
99%	1%	0.5%	2.576

The z-score associated with a 5% alpha level/2 is 1.96.

Step 5: compare the calculated z-score from step 3 with the table z-score from step 4. If the calculated z-score is larger you can reject the null hypothesis.

8.99 > 1.96, so we can reject the null hypothesis.

Q4 Text Solution:

We are given : n = 22, $\bar{x} = 153.7$, s = 17.2.

Null Hypothesis. The advertising campaign is not successful, i.e.,

$H_0 : \mu = 146.3$

Alternative Hypothesis. $H_1 : \mu > 146.3$. (Right - tail).

Test Statistic. Under the null hypothesis, the test statistic is :

$t = \frac{\bar{x} - \mu}{\sqrt{s^2 / n - 1}} \sim t_{22-1} = t_{21}$

Now, $t = \frac{153.7 - 146.3}{\sqrt{17.2^2 / 21}} = \frac{7.4 \times \sqrt{21}}{17.2} = 9.03$

Conclusion. Tabulated value of t for 21 d . f. at 5% level of significance for single - tailed test is 1.72, Since calculated value is much greater than the tabulated value, it is highly significant. Hence we reject the null hypothesis and conclude that the advertising campaign was definitely successful in promoting sales.

Q5 Text Solution:

Null hypothesis, H0 : The data are consistent with the assumption of a mean I.Q. of 100 in the population, i. e., $\mu = 100$.

Alternative hypothesis , H1 :$\mu \neq 100$.

Test Statistic. Under H0, the test statistic is :

$t = \frac{\bar{x} - \mu}{\sqrt{S^2 / n}} \sim t_{n-1}$

where $\bar{x}$ and S^2 are to be computed from the sample values of I.Q. 's.

CALCULATIONS FOR SAMPLE MEAN AND S.D.

X	$X - \bar{x}$	$X - \bar{x}^2$
70	- 27.2	739.84
120	22.8	519.84
110	12.8	163.84
101	3.8	14.44
88	- 9.2	84.64
83	- 14.2	201.64
95	- 2.2	4.84
98	0.8	0.64
107	9.8	96.04
100	2.8	7.84
Total 972		1833.60

Hence n = 10,

$\bar{x} = \frac{972}{10} = 97.2$ and $S^2 = \frac{1833.60}{9} = 203.73$

$\therefore\ t = \frac{97.2 - 100}{\sqrt{203.73 / 10}} = \frac{2.8}{\sqrt{20.37}} = \frac{2.8}{4.514} = 0.62$

Tabulated $t_{0.05}$ for (10 - 1) i.e., 9 d. f, for two - tailed test is 2.262

Conclusion. Since calculated is less than tabulated $t_{0.05}$ for 9 d.f., Ho may be accepted at 5% level of significance and we may conclude that the data are consistent with the assumption of mean I.Q. of 100 in the population.

The 95% confidence limits within which the meah I.Q. values of samples of 10 boys will lie are given by :

$\bar{x} \pm t_{0.05} S / \sqrt{n} = 97.2 \pm 2.262 \times 4.514$

.

Q6 Text Solution:

Null Hypothesis, $H_0 : \mu_X = \mu_Y$. i.e.., the two types I and II of electric bulbs are indentical.

Alternative Hypothesis, $H_1 : \mu_X = \mu_Y$. i.e.., type 1 is superior to type II, Test statistic. Under H_0, the test statistic is :

$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{S^2 \frac{1}{n_1} + \frac{1}{n_2}}} \sim t_{n1 + n2 - 2} = t_{13}$

where $S^2 = \frac{1}{n_1 + n_2 - 2} \sum x_1 - \bar{x}_1{}^2 + \sum x_1 - \bar{x}_2{}^2$

$=$

$\frac{1}{n_1 + n_2 - 2} \quad n_1 s_1^2 + n_2 s_2^2 \quad = \frac{1}{13} \quad 8 \times 36^2 + 7 \times 40^2 = 1659.08$

$\therefore \quad t = \dfrac{1234 - 1036}{\sqrt{1659.08\frac{1}{8} + \frac{1}{7}}} = \dfrac{198}{\sqrt{1659.08 \times 0.2679}} = 9.39$

Tabulated value of t for 13 d.f. at 5% level of significance for right (single) tailed test is 1 - 77. [This is the value of $t_{0.10}$ for 13 df. from two-tail tables given in Appendix).

Conclusion. Since calculated 'r' is much greater than tabulated 't', it is highly significant and H_0 is rejected. Hence the two types of electric bulbs differ significantly. Further since $\bar{x}_1$, is much greater than $\bar{x}_2$, we conclude that type I is definitely superior to type II.

Q7 Text Solution:

Step 1: Null Hypthesis: H_0: σ = 9

Alternative Hypthesis: H_1: σ >

Step 2: Given that n = 20; s = 6

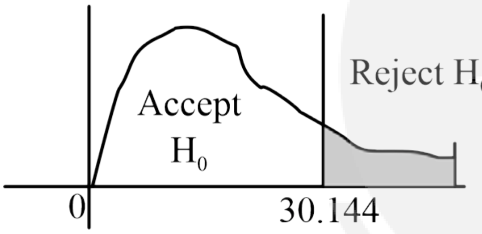
under H_0:

$x^2 = \frac{ns^2}{\sigma^2}$

$= \frac{206^2}{9^2} = 8.89$

Step 3: d.f. is 20 – 1 = 19

Tabulated value $x_{19}^2 0.05 = 30.144$



since 8.89 < 30.144

So, we may H_0 or we fail to reject H_0.

Thus population standard deviation may be considered as 9 at 5% level of significance.

Q8 Text Solution:

Step 1: Null Hypthesis: H_0: $\sigma^2 = 4.35$

Alternative Hypthesis: H_1: $\sigma^2 > 4.35$

Step 2: Given that n = 20; s = 3.72

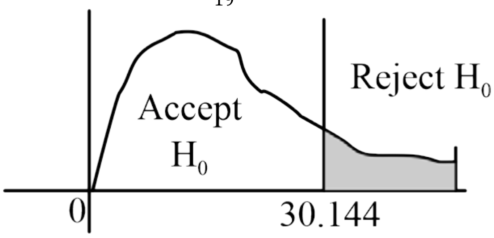
under H_0:

$x^2 = \frac{ns^2}{\sigma^2}$

$= \frac{203.72^2}{4.35} = 63.62$

Step 3: difference is 20 – 1 = 19

Tabulated value $x_{19}^2 0.05 = 30.144$



since 63.62 < 30.144

So, we may H_0

Thus, we conclude that the random sample is definitely NOT from the population variance 4.35.

Q9 Text Solution:

Null Hypothesis: $H_0 = \sigma^2 = 20$

Alternative Hypothesis: $H_1 = \sigma^2 > 20$

Step 2: $\bar{x} = \frac{\sum x_i}{n} = \frac{470}{10} = 47$

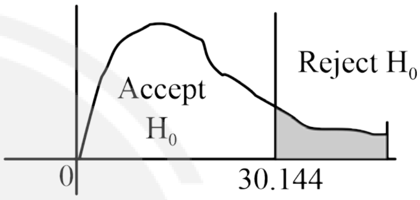
$s^2 = \frac{\sum x_i - \bar{x}^2}{n} = \frac{280}{10} = 28$

Under $H_0 = \frac{ns^2}{\sigma^2}$

$= \frac{1028}{20} = 14$

Step 3: difference is 10 – 1 = 9

Tabulated value $x_9^2 0.05 = 16.92$



since 14 < 16.92,

so we MAY ACCEPT H_0.

Thus, we conclude that the variance of distribution of weights of all the student in the population is 20 sqaure kgs.

Q10 Text Solution:

Null Hypothesis: H_0 : $\sigma^0 = 32$

Alternative Hypothesis : H_1 : $\sigma^2 = 32$

Step 2: $\bar{x} = \frac{\sum x_i}{n} = \frac{693}{10} = 69.3$

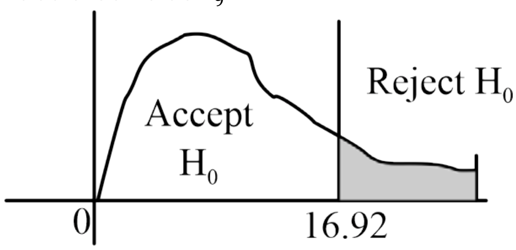
$s^2 = \frac{\sum x_i - \bar{x}^2}{n} = \frac{23.4.1}{10} = 23.41$

Under H_0: $x^2 = \frac{ns^2}{\sigma^2}$

$= \frac{10236.41}{32} = 73.3156$

Step 3: difference is 10 – 1 = 9

Tabulated value $x_9^2 0.05 = 16.92$



Since 7.3156 < 16.92,

so we MAY ACCEPT H_0.

Thus, we conclude that the variance of the random sample is taken from the population variance 32.

Q11 Text Solution:

Step 1: Calculate the expected frequencies

	Skateboards	Can't skateboard	Row total
Golden retrievers	20 (40 × 70) / 100 = 28	20 (40 × 30) / 100 = 12	40
French bulldogs	50 (60 * 70) / 100 = 42	10 (60 * 30) / 100 = 18	60
Column total	70	30	N = 100

Step 2: Calculate chi-square

Intervention	Outcome	Observed	Expected	O – E	$(O - E)^2$	$(O - E)^2 / E$
Golden retrievers	Skateboards	20	28	−8	64	2.29
	Can't skateboard	20	12	8	64	5.33
French bulldogs	Skateboards	50	42	8	64	1.52
	Can't skateboard	10	18	−8	64	3.56

$X^2 = 2.29 + 5.33 + 1.52 + 3.56 = 12.7$

Step 3: Find the critical chi-square value

Since there are two dog breed and two outcomes there is (2 − 1) × (2 − 1) = 1 degree of freedom.

For a test of significance at α = .05 and df = 1, the X^2 critical value is 3.84.

Step 4: Compare the chi-square value to the critical value

$X^2 = 12.7$

Critical value = 3.84

The X^2 value is greater than the critical value.

Step 5: Decide whether the reject the null hypothesis

The X^2 value is greater than the critical value. Therefore, the dog trainer should reject the null hypothesis that a dog's breed is unrelated to whether they can learn to skateboard. Her data suggests that a larger proportion of french bulldogs can learn to skateboard than golden retrievers.

Q12 Text Solution:

Step 1: Calculate the expected frequencies

	Would recommend	Would not recommend	Row total
Tasty Burgers	20 (25 * 60) / 75 = 20	5 (25 * 15) / 75 = 5	25
Burger Prince	22 (25 * 60) / 75 = 20	3 (25 * 15) / 75 = 5	25
Burger Town	18 (25 * 60) / 75 = 20	7 (25 * 15) / 75 = 5	25
Column total	60	14	75

Step 2: Calculate chi-square

Intervention	Outcome	Observed	Expected	O – E	$(O - E)^2$	$(O - E)^2 / E$
Tasty Burgers	Would recommend	20	20	0	0	0
	Would not recommend	5	5	0	0	0
Burger Prince	Would recommend	22	20	2	4	0.2
	Would not recommend	3	5	−2	4	0.8
Burger Town	Would recommend	18	20	−2	4	0.2

Android App | iOS App | PW Website

GATE

	Would not recommend	7	5	2	4	0.8

$X^2 = 0 + 0 + 0.2 + 0.8 + 0.2 + 0.8 = 2$

Step 3: Find the critical chi-square value

Since there are three restaurants and two outcomes there are $(3 - 1) \times (2 - 1) = 2$ degrees of freedom.

For a test of significance at $\alpha = .05$ and df = 2, the X^2 critical value is 5.99.

Step 4: Compare the chi-square value to the critical value

$X^2 = 2$

Critical value = 5.99

The X^2 value is less than the critical value.

Step 5: Decide whether the reject the null hypothesis

The X^2 value is less than the critical value. Therefore, the restaurant reviewed should not reject the null hypothesis the proportion of customers recommending the restaurant is the same for the three restaurants

Q13 Text Solution:

Step 1: Calculate the expected frequencies

Nut	Frequency	Expected
Peanuts	621	1000 * 0.5 = 500
Cashew	189	1000 * 0.3 = 300
Almonds	190	1000 * 0.2 = 200

Step 2: Calculate chi-square

Phenotype	Observed	Expected	O − E	$(O - E)^2$	$(O - E)^2$ / E
Peanuts	621	500	121	14641	29.28
Cashew	189	300	−111	12 321	41.07
Almonds	190	200	−10	100	0.5

$X^2 = 29.28 + 41.07 + 0.5 = 70.85$

Step 3: Find the critical chi-square value

Since there are three groups, there are two degrees of freedom.

For a test of significance at $\alpha = .05$ and df = 2, the X2 critical value is 5.99.

Step 4: Compare the chi-square value to the critical value

$X^2 = 70.85$

Critical value = 5.99

The X^2 value is greater than the critical value.

Step 5: Decide whether the reject the null hypothesis

The X^2 value is greater than the critical value, so you should reject the null hypothesis that the nut mix has the desired proportions of nuts. The data suggests that there's a problem with the nut mix.

Android App | iOS App | PW Website