25UCU305: DISCRETE MATHEMATICS WITH PROBABILITY AND HYPOTHESIS TESTING

Year: I (B.Sc., CS, IT, AI-DS, DS &V, BCA, DA, Cyber Security, DevOps, FSWD)

Semester: II

Unit V Section A

1. Define t test and give the formula for t test statistic for single mean.
2. State any four properties of student's 't' distribution.
3. A sample of 10 students has an average test score of 72, and the sample standard deviation is 8. Test whether the mean score is different from the hypothesized population mean of 75. Compute the t-test statistic.
4. A food-delivery company claims that its average delivery time in a metropolitan city is 30 minutes or less. To verify this claim, a quality-control team randomly selected 16 orders and recorded the following delivery times (in minutes): 32, 28, 35, 39, 27, 31, 33, 29, 36, 30, 34, 31, 38, 26, 33, 30. At $\alpha = 0.05$, perform a one-sample t-test to determine whether the true mean delivery time is significantly greater than 30 minutes. Critical t-value from t-table: $t_{0.05,\ 15} = 1.753$.
5. State t-test formula for difference between two means. (K)
6. What is chi-square test? State its uses.
7. Give the formula for goodness-of-fit in chi-square test. (K).
8. Give the formula for independence of attributes in chi-square test. (K)
9. State any four properties of χ^2 distribution.
10. Explain F-test and its degrees of freedom. (C)
11. State the formula for finding degrees of freedom (Welch–Satterthwaite approximation).
12. A data science team is comparing the effectiveness of two machine learning models on a small validation dataset.

Parameter	Model A (Baseline)	Model B (New)
Sample Size (n)	24	19
Mean Accuracy (%)	78.5	82.3
Sample Standard Deviation (%)	4.8	6.2

Compute the degrees of freedom using the Welch–Satterthwaite approximation. (Typical Scaler Data Science / AI-ML Certification Question).

13. Two laboratories measure the concentration of a chemical in the same solution. Samples taken from each lab give the following results:

    Lab X: Sample size $n_1 = 15$ Sample variance $s_1^2 = 36$

    Lab Y: Sample size $n_2 = 13$Sample variance $s_2^2 = 16$

    Test at the 5% significance levelwhether the two population variances are equal. Critical value $F_{0.025,14,12} = 3.206$.

14. A quality engineer wants to determine if two production lines (Line P and Line Q) are operating with the same process variability in terms of product weight (in grams). Independent random samples were taken with the following results: Calculate the F-statistic.

Unit V Section B

15. Give the procedure for finding significant difference between population mean and sample mean for t test.
16. Give the procedure for finding significant difference between two sample means for t test.
17. A manufacturer claims that the average lifetime of a LED bulb is5000 hours. A sample of10 bulbsgives an average lifetime of4800 hours, with a sample standard deviation of 300 hours. Using a one-sample t-test, compute thet-test statistic.

18. In a data science project analyzing employee performance, a sample of 16 analysts from a firm yields an average daily output of 42 reports, with a standard deviation of 3.8 reports. The industry standard mean output is 40 reports per day. Conduct a one-sample t-test at the 1% significance level to assess if the firm's analysts exceed the standard.Employee Productivity Analysis (Representative of Scaler Data Science Course Exercises).

19. A marketing team tests a new ad campaign by surveying 20 customers, finding an average satisfaction score of 7.8 out of 10, with a standard deviation of 1.2. The historical population mean satisfaction score is 7.0. Using a 5% significance level, perform a one-sample t-test to determine if the campaign has improved satisfaction. Marketing Campaign Effectiveness (Representative of NxtWave Data Analytics Course Exercises).

20. A pharmaceutical company is evaluating a new formulation of a drug intended to lower systolic blood pressure more effectively than the current placebo-controlled baseline.

Parameter	Placebo Group	New Drug Group
Sample Size (n)	15	19
Mean reduction in systolic BP (mmHg)	2.80	7.40
Sample Standard Deviation (mmHg)	3.10	4.20

Conduct a two-tailed Welch's t-test at $\alpha = 0.05$ to determine whether the new drug produces a statistically significant difference in mean blood-pressure reduction compared to the placebo Assume unequal population variances.

21. An e-commerce platform conducted an A/B test to evaluate a new recommendation engine expected to increase user engagement.

Parameter	Control Group (Old Algorithm)	Treatment Group (New Algorithm)
Sample Size (n)	22	18
Mean session duration (minutes)	6.20	8.10
Sample Standard Deviation (minutes)	1.80	2.30

Using Welch's two-sample t-test at $\alpha = 0.05$, test the hypothesis that the new recommendation algorithm leads to a statistically significant increase in average session duration. Assume population variances are unequal.(tcrit≈1.694). (Scaler Data Science, Scaler AI/ML, NxtWave CCBP 4.0)

22. A retail chain manager claims that customer footfall in their store is evenly distributed across the six days it is open (Monday to Saturday), i.e., approximately 16.67% of weekly customers visit on each day. Over a period of 8 weeks, a total of 720 customers were recorded with the following observed distribution:

Day of the Week	Monday	Tuesday	Wednesday	Thursday	Friday	Saturda
Observed Frequency (O_i)	98	102	108	110	142	160
Expected Frequency (E_i)	?	?	?	?	?	?

Using a Chi-Square Goodness-of-Fit test at $\alpha = 0.05$, test whether the customer footfall is uniformly distributed across the six days. (Critical value $\chi^2_{0.05,\ 5} = 11.070$) (Scaler Data Science,Scaler AI/ML, NxtWave CCBP 4.0/Intensive, Great Learning, UpGrad, and most international proctored exams).

23. A retail bank wants to test whether Gender (Male / Female) is independent of Preference for Digital Banking Channel (Mobile App / Net Banking / Branch Visit).A random sample of 500 customers produced the following contingency table:

Gender \ Channel	Mobile App	Net Banking	Branch Visit	Row Total

Male	120	85	45	250
Female	140	60	50	250
Column Total	260	145	95	500

Perform a Chi-Square Test of Independence at $\alpha = 0.05$ to determine if Gender and Preferred Banking Channel are independent. critical value = 5.991. (Very commonly repeated in Scaler DS & AIML 2023–2025 batches).

24. An ed-tech company wants to check whether Student Employment Status (Full-time Employed / Part-time Employed / Unemployed / Student) is independent of Course Completion (Completed / Dropped Out). From a cohort of 600 learners:

Employment Status \ Course Outcome	Completed	Dropped Out	Row Total
Full-time Employed	110	40	150
Part-time Employed	95	55	150
Unemployed	120	30	150
Student	105	45	150
Column Total	430	170	600

Using $\alpha = 0.05$, Conduct a Chi-Square Test of Independence. (df = 3, critical value = 7.815). (Standard in NxtWave CCBP 4.0 & Scaler final exams).

25. A machine learning team is evaluating two versions of a sentiment-analysis model on separate small test sets. The following summary statistics were obtained:

Parameter	Model X (Current)	Model Y (Proposed)
Sample Size (n)	26	21
Mean F1-Score (%)	84.6	88.9
Sample Standard Deviation (%)	5.4	7.1

Assuming unequal population variances, compute the approximate degrees of freedom for Welch's two-sample t-test using the Welch–Satterthwaite approximation. (Scaler Data Science, Scaler AI-ML, NxtWave CCBP 4.0/Intensive).

26. A retail chain wants to test whether the proportion of customers who purchase "Product X" is the same across its three geographical regions (North, East, and West). Random samples of customers were taken from each region with the following results:

Region	Purchased Product X (Yes)	Did Not Purchase (No)	Row Total
North	85	115	200
East	72	128	200
West	96	104	200
Column Total	253	347	600

Using a Chi-Square test at $\alpha = 0.05$, determine whether the proportion of customers purchasing Product X is homogeneous (the same) across the three regions. (df = 2, critical value = 5.991).
(Extremely Common in Scaler &NxtWave 2023–2025 Batches).

27. An ed-tech platform wants to check whether the course completion rate is the same across four different marketing channels (Google Ads, Facebook Ads, Email Campaign, and Referral)

Marketing Channel	Completed Course	Dropped Out	Row Total
Google Ads	120	80	200
Facebook Ads	105	95	200
Email Campaign	135	65	200
Referral	92	108	200
Column Total	452	348	800

Perform a Chi-Square test forhomogeneity at $\alpha = 0.05$ to test if the course completion rate is identical across the four marketing channels. (df = 3, critical value = 7.815). (Standard in Almost Every Certification Final Exam)

28. A company wants to check whether two machines produce items with thesame variability. Two samples are taken:
    Machine A:Sample size $n_1 = 10$ Sample variance $s_1^2 = 25$
    Machine B: Sample size $n_2 = 12$Sample variance $s_2^2 = 10$
    Test whether the population variances are equal at the 5% level of significance. Critical value $F_{0.025,9,11} = 4.03$.

Unit V Section C

29. A data science team is analyzing user engagement for a mobile app. The product manager claims that the average daily time spent by users is at least 35 minutes. To check this claim, a data analyst collects a random sample of 12 users, and finds the following *time spent per day (in minutes):* Sample data: 34, 30, 28, 36, 40, 32, 38, 29, 31, 33, 37, 35. Using a one-sample t-test, test whether the mean daily usage is less than 35 minutes at the 5% significance level. State H_0andH_1. 2. Compute the sample mean and sample standard deviation. Calculate the t-test statistic. State whether the claim is supported or rejected. Critical t-value from t-table: $t_{0.05,\ 11} = 1.796$. (similar to SCALAR AIML, SCALAR Data Science, NxtWave, etc.).

30. An e-commerce company wants to test whether a new website design (Version B) leads to a statistically significant increase in average time spent on the site compared to the old design (Version A).

Parameter	Version A (Old Design)	Version B (New Design)

Sample Size (n)	30	35
Sample Mean ($\bar{x}$)	4.8 minutes	5.7 minutes
Sample Standard Deviation (s)	1.4 minutes	1.9 minutes
Sample Variance (s^2)	1.96	3.61

Perform a two-sample t-test at $\alpha = 0.05$ to determine whether the new design significantly increases the average time spent. Assume the population variances are unequal. ($t_{crit} \approx 1.6699$)(Frequently used in Scaler DS & AIML batches).

31. An online retail company introduced a redesigned checkout interface to reduce transaction abandonment by shortening completion time

Parameter	Old Checkout Layout (A)	New Checkout Layout (B)
Sample Size (n)	25	21
Mean checkout time (seconds)	78.40	69.20
Sample Standard Deviation (seconds)	12.50	15.80

At a 5% significance level, use Welch's two-sample t-test (unequal variances assumed) to test whether the new layout significantly reduces the average checkout time (one-tailed test). Provide the full solution including the t-statistic, approximate degrees of freedom, decision rule, and practical interpretation.($t_{crit} \approx t_{0.95, df \approx 37.84} \approx 1.68613$) (Scaler Data Science, Scaler AI/ML, and NxtWave certification programs).

32. A data science team at an e-commerce company believes that customer complaints received through their support portal follow a uniform distribution across the five working days (Monday to Friday), i.e., 20% of complaints each day.They recorded the complaints over 10 weeks (total 300 complaints) and obtained the following observed frequencies:

Day	Monday	Tuesday	Wednesday	Thursday	Friday	T
Observed (O_i)	72	54	48	51	75	3
Expected (E_i)	?	?	?	?	?	3

At $\alpha = 0.05$, perform a Chi-Square Goodness-of-Fit test to determine whether the complaints are uniformly distributed across the five weekdays. (From Chi-Square table: $\chi^2_{0.05,4} = 9.488$). (Scaler Data Science, Scaler AI/ML, NxtWave CCBP 4.0, and other certification final exams and mock tests).

33. An e-commerce company wants to test whether Customer Segment (New vs Returning) is independent of Preferred Payment Method (Credit Card, Debit Card,

UPI, Cash on Delivery). A random sample of 400 recent transactions yielded the following contingency table:

Customer Segment \ Payment Method	Credit Card	Debit Card	UPI	Cash on Delivery	Row Total
New Customer	48	62	95	45	250
Returning Customer	72	38	25	15	150
Column Total	120	100	120	60	400

At $\alpha = 0.05$, perform a Chi-Square Test of Independence to determine whether Customer Segment and Preferred Payment Method are independent. (Critical value $\chi^2_{0.05,3} = 7.815$). (Regularly asked in Scaler Data Science, Scaler AI/ML, NxtWave CCBP 4.0/Intensive, and most international certification final assessments). Representative Proctored-Exam Problem.

34. A manufacturing company is comparing the consistency of two production machines (Machine A and Machine B) by measuring the diameter (in mm) of 10 randomly selected parts from each machine. The company wants to test whether the variances in diameters are significantly different at $\alpha = 0.05$.

    Machine A diameters: 50.1, 49.8, 50.3, 49.9, 50.2, 50.0, 50.4, 49.7, 50.1, 50.0

    Machine B diameters: 50.5, 49.5, 50.7, 49.6, 50.8, 49.4, 50.6, 49.3, 50.9, 49.2

    Upper critical value: (Upper critical value: $F_{0.975,9,9} \approx 4.02599$.

35. A manufacturing engineer wishes to determine whether two CNC machines (Machine X and Machine Y) produce parts with the same precision (variability in diameter). Random samples of parts are measured with the following results:

Machine	Sample Size (n)	Sample Variance (s^2)
Machine X	15	0.028
Machine Y	21	0.073

(a) State the null and alternative hypotheses. (b) Compute the F-statistic (show the formula and calculation). (c) State the critical value(s) from the F-table. (d) State the decision and conclusion. (Upper critical value: $F_{0.025,\ 20,\ 14} \approx 2.57$).

Prepared By

1. Dr.K.M. Manikandan

Assistant Professor and Head

Department of Mathematics

Dr SNS rajalakshmi College of Arts and Science (Autonomous)

Coimbatore – 641049

2. Mr. P. Murugan

Assistant Professor

Department of Mathematics

Dr SNS Rajalakshmi College of Arts and Science (Autonomous)

Coimbatore – 641049

Covering Letter

From

Dr.K.M. Manikandan

Assistant Professor and Head

Department of Mathematics

Dr SNS Rajalakshmi College of Arts and Science (Autonomous)

Coimbatore – 641049

To

The Principal

Dr SNS Rajalakshmi College of Arts and Science (Autonomous)

Coimbatore – 641049

Respected Madam

Sub: Submission of Question Bank for the EVEN SEMESTER 2025 – 2026 Reg:

Here with have submitted the Question Bank of the Course **25UCU305: "DISCRETE MATHEMATICS WITH PROBABILITY AND HYPOTHESIS TESTING**" for your kind notice. Thanking You

With Regards

Yours Sincerely

Date of Submission: 09.12.2025

Dr.K.M. Manikandan