Practice Problems Notebook

MAT241 Class

Table of contents

In this notebook, we'll work through several practice problems. The problems appear in no particular order, and each one is best addressed using an inferential tool that we've encountered so far. As a reminder, we've seen

- Confidence Intervals for...

  - One population proportion
  - A difference between two population proportions
  - One population mean
  - A difference between two population means

- Hypothesis Tests for...

  - One population proportion (one sample z-test)
  - A difference between two population proportions (two sample z-test)
  - Chi-Squared Goodness of Fit Tests for the distribution of a multiclass categorical variable
  - Chi-Squared Tests for Independence between two potentially multiclass categorical variables
  - One population mean (one sample t-test)
  - A difference between two population means (two sample t-test)

You may benefit by referring back to the General Strategy for Constructing Confidence Intervals, General Strategy for Conducting Hypothesis Tests, and Standard Error Decision Tree documents.

Example Scenarios

Scenario 1: Average Commute Time

An urban planning committee wants to verify that the average commute time for citizens residing in a section of their city is higher than the national average of 30 minutes. They survey 40 residents, whose average commute time is 32 minutes with a standard deviation of 5 minutes. Test if the average commute time for city residents in this location is significantly greater than 30 minutes.

Scenario 2: Psychology Study on Sleep Quality

A psychology researcher wants to investigate whether a new sleep therapy program is effective in improving sleep quality. They collect data from 50 patients who completed the program and 50 who did not use the program. Among the program participants, 60% reported significantly improved sleep quality, while 48% of the control group reported improvement. Test for a significant difference in sleep improvement rates between the two groups.

Scenario 3: Effectiveness of Two Math Tutoring Methods

A school tests two methods for teaching math to see which is more effective. In Method A, 15 students achieve an average improvement of 12 points on a standardized test with a standard deviation of 4 points. In Method B, 18 students improve by an average of 14 points with a standard deviation of 3 points. Test if the two methods yield different average improvements at the $\alpha = 0.10$ level of significance.

Scenario 4: Marketing Campaign Success Rate

A company runs a social media campaign targeting two age groups: under 30 and over 30. Out of 150 users under 30, 45 clicked on the ad, while 40 out of 100 users over 30 clicked on it. Construct a 95% confidence interval for the difference in the click-through rates between the two age groups.

Scenario 5: Free Throw Accuracy

A basketball coach tracks the free throw accuracy of two players. Both players shot 100 free throws under the same conditions. Player A made 85 while Player B made 77. Is there a statistically significant difference in free throw accuracy between the two players?

Scenario 6: Exercise Preferences

A fitness company wants to understand if there is an association between preferred exercise type and age group among their clients. They randomly survey 296 clients and categorize their preferences into three exercise types: Strength Training, Cardio, and Yoga/Pilates. The clients are also grouped by age: Under 30, 30-50, and Over 50. They summarize the results in the table below and ask you to determine whether an association between age group and exercise preference exists.

Age Group	Strength Training	Cardio	Yoga/Pilates	Total
Under 30	43	55	25	123
30-50	42	37	33	112
Over 50	21	22	18	61
Total	106	114	76	296

Scenario 7: Sociology Survey on Pet Ownership

A survey of 300 randomly identified households in the US found that 200 households own pets. Construct a 90% confidence interval for the proportion of households that own pets.

Scenario 8: Customer Waiting Time in a New Restaurant

A new restaurant claims that customers wait an average of 10 minutes to receive their orders. A random sample of 25 customers has an average waiting time of 12 minutes with a standard deviation of 3 minutes. Conduct a hypothesis test to determine if the actual average wait time at this location exceeds the restaurant's claim.

Scenario 9: Recidivism Rates

In a study comparing recidivism rates for two rehabilitation programs, Program A has a recidivism rate of 22% (n = 120), while Program B has a rate of 18% (n = 100). Test for a difference in recidivism rates between the two programs at the 5% significance level.

Scenario 10: Distribution of Car Colors

A car dealership wants to know if their sales align with national color preferences. Nationally, 30% of cars are white, 25% are black, 20% are gray, 15% are blue, and 10% are other colors. A sample of 500 cars sold at the dealership shows 150 white, 130 black, 110 gray, 80 blue, and 30 other cars. Test whether the dealership's color distribution matches the national distribution.

Scenario 11: Plant Growth

A botanist tests a new fertilizer's effect on plant growth. A sample of 25 plants grown with the fertilizer has an average height of 15 cm with a standard deviation of 3 cm. Construct a 95% confidence interval for the average height of plants grown with the fertilizer.

Scenario 12: Blood Pressure Treatment Effectiveness

A study evaluates the effectiveness of two drugs in lowering blood pressure. Drug A shows an average reduction of 10 mmHg (n = 35, SD = 4 mmHg), while Drug B shows an average reduction of 8 mmHg (n = 40, SD = 5 mmHg). Construct a 99% confidence interval for the difference in blood pressure reduction between the two drugs.

Scenario 13: Housing Market Survey

In a city, 60% of a sample of 200 renters express interest in purchasing a home in the next year. Construct a 95% confidence interval for the proportion of renters interested in home ownership.

Scenario 14: Ice Cream Flavor Preferences at a Festival

A food festival organizer wants to know if their customers' preferences for ice cream flavors match the flavors ordered nationwide. Nationally, the distribution of favorite flavors is 35% vanilla, 25% chocolate, 20% strawberry, 10% mint chocolate chip, 4% cookie dough, and 6% other flavors. During the festival, a random sample of 750 ice cream orders shows 267 for vanilla, 196 for chocolate, 153 for strawberry, 71 for mint chocolate chip, 24 cookie dough, and 39 for other flavors. Conduct a chi-squared goodness of fit test to determine if the flavor preferences at the festival differ from the national distribution.

Scenario 15: Vaccination Rates

A city health department wants to estimate the vaccination rate for a specific disease. In a sample of 400 residents, 312 have been vaccinated. Construct a 90% confidence interval for the vaccination rate.

Scenario 16: Porch Pirating

Porch pirating is a term used to describe an unattended package being stolen from a doorstep, porch, etc. An analyst compares rates of porch pirating in two neighborhoods. In Neighborhood A, 24 of 100 households report at least one porch pirating incident within the last 12 months, while in Neighborhood B, 16 of 80 households reported these events over the same time period. Test for a difference in burglary rates between the two neighborhoods.

Scenario 17: Athlete Recovery Time

Sports scientists measure the average heart rate recovery time (in seconds) for a group of 23 NCAA Division II athletes after a high-intensity workout. The average time to return to an individual's normal heart rate after the workout is recovery time is 117 seconds with a standard deviation of 9 seconds. Test if the recovery time is significantly less than two minutes.

Scenario 18: Running Speed Comparison

A coach tests the average running speed of two groups of athletes. Group A has an average speed of 7.2 m/s (n = 20, SD = 0.5), and Group B has an average speed of 7.0 m/s (n = 25, SD = 0.6). Construct a 95% confidence interval for the difference in average speed between the two groups.

Scenario 19: Study Habit Survey

A school surveys 500 students to see if study habits are associated with grade level (freshman, sophomore, junior, senior). Test for independence between grade level and whether students study regularly using a chi-squared test.

Grade Level	Regular Studying (ie. Weekly)	Cramming or No Studying	Total
Freshman	70	55	125
Sophomore	80	45	125
Junior	95	30	125
Senior	70	55	125
Total	315	185	500

Scenario 20: Memory Retention

A psychologist investigates the effect of caffeine on memory retention. In a sample of 30 students who consumed caffeine, the average recall score was 70% (SD = 8%). Construct a 99% confidence interval for the mean recall score for students who consumed caffeine.

Scenario 21: Customer Satisfaction Survey

A store conducts a satisfaction survey with 400 customers, finding that 85% are satisfied. Construct a 95% confidence interval for the proportion of satisfied customers.

Scenario 22: Online Course Completion Rates

A college compares completion rates between students enrolled in online and in-person classes. Out of 153 online students, 114 completed the course, while 147 out of 161 in-person students completed it. Test whether course completion rates for in-person students exceeds that for online students.

Scenario 23: Streaming Service Usage

A media company surveys 350 randomly identified households from its member population and finds that 48% of them subscribe to a music streaming service. Construct a 90% confidence interval for the proportion of households that subscribe to music streaming.

Scenario 24: College Students' Weekly Study Hours

A university administrator wants to see if students at their institution study more than the national average of 15.5 hours per week. A random sample of 19 students from their school has an average weekly study time of 16.25 hours with a standard deviation of 2.3 hours. Test if the university's students study more on average than students nationally.

Scenario 25: Comparing Productivity in Two Office Settings

A company measures productivity (number of tasks completed) in two different office layouts, Layout A and Layout B. A sample of 20 employees working in Layout A completes an average of 30 tasks with a standard deviation of 6 tasks, while a sample of 25 employees in Layout B completes an average of 27 tasks with a standard deviation of 5 tasks. Use the $\alpha = 0.10$ level of significance to test if there is evidence of a difference in productivity between the two office layouts.

Scenario 26: Customer Loyalty Program Impact

A company tests a new loyalty program. In a sample of 211 program participants, 124 made repeat purchases, while in a control group of 196 non-participants, 97 made repeat purchases. Conduct a test to determine whether the loyalty program is successful in generating repeat purchases.

Scenario 27: Traffic Violations by Age Group

A traffic study looks at the frequency of violations among two age groups. Among 18-25 year-olds, 43 of 155 drivers have received a violation in the past year, compared to 32 out of 127 drivers aged 26-35. Test for a difference in violation rates between the two age groups.

Scenario 28: Exercise and Blood Pressure

A health study evaluates blood pressure reductions from exercise programs. The average blood pressure reduction for 25 participants in Program A was 12 mmHg (SD = 3 mmHg), and for 30 participants in Program B, it was 10 mmHg (SD = 4 mmHg). Construct a 95% confidence interval for the difference in blood pressure reduction between the two programs.