TEACHING INTRO STATISTICS USING INQUIRY BASED AND PROJECT BASED LEARNING STRATEGIES

Presenter: Jae Ki Lee, Ed.D

Associate Professor at Borough of Manhattan Community College
· Explore students’ achievements of the course after learning the course based on Inquiry Based and Project Based Learning strategies.
· Those alternative instructional methods enhance students’ learning deeply and conceptually.
· Students’ satisfaction
Each student gets assigned in a group and completed group project.

At the beginning of the semester, each group decides their project topic:
- Topics must be able to analyze numerically; Height, shoe size, Weight, Commuting time
- Both genders

Students get assigned a total of 8 labs:
- Between the Lab 1 through 4: Review the contents
- Between the Lab 5 through 8: Hypothesis Test Project.
THIS PRESENTATION

- Demonstrate
  - Labs
  - Activities
    - Creating frequency tables
    - Binomial Probability
    - Normal Distribution / Central Limit Theorem

- Hypothesis Project
- Activity sheets
- Students’ Presentation
- Sample students’ projects
Lab 1:

Throughout the Lab1, you would choose one quantitative topic, such as height, weight, foot size, age, or etc. You would write a short paragraph based on the following steps, and submit by next class.

Step 1: Determine your quantitative topic

Step 2: Collect 30 data at BMCC campus: Your data must include gender.

Step 3: Create Frequency table with 5, 6, or 7 class. Justify why you choose the number of class

Step 4: Create histogram

Step 5: Find the central Tendency

Step 6: Create Box-Plot

Step 7: Find the variance and standard deviation.
Lab 2

Throughout the Lab 2, you would create probability contents based on your data from the Lab 1. Write a short paragraph based on answering the following steps.

Step 1: Create the conditional probability table based on your data.

<table>
<thead>
<tr>
<th></th>
<th>Male</th>
<th>Female</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Greater than or equal</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>to the mean</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Less than the mean</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Step 2: Answer the following questions:

a. Find the probability of male.
b. Find the probability of female.
c. Find the probability “Greater than or equal to the mean”
d. Find the probability “Less than mean”
e. Find the probability of male given “Greater than or equal to the mean”
f. Find the probability of “Less than the mean” given that female
g. Find the probability of female given that “Less than the mean”
h. Are events “Male and Female” mutually exclusive?
i. Are events “Greater than or equal to the mean” and “Female” mutually exclusive?
j. Are Less than the mean and female independent event?
Lab 3

Using your data from the Lab 1, write a short paragraph based on answering the following questions.

Question 1: Find the probability to select male and female

Question 2: Find the probability to select “Greater than or equal to the mean”, and “Less than mean”

Question 3: Find the probability to select
   a. exactly three females selected
   b. less than four females selected
   c. between 6 and 7 females

Question 4: Find the probability to select
   a. exactly six people selected “Greater than or equal to the mean”
   b. between four and six people selected “Less than the mean”
Lab 4: Summary of the previous three labs.

Using your own words, create essay paragraphs of your projects. This short summary must be at least 6 pages with including

1. Frequency Table with the description

2. Minimum three displaying graphs: Histogram, Pie-chart, Bar-charts, line graph and etc.

3. Box-plot with as detail description as possible

4. Input a conditional probability table, and describe at least three conditional probabilities.

5. Define at least one mutually exclusive events

6. Demonstrate two of independent/dependent event, and explain why.
### ACTIVITY: FREQUENCY TABLE

<table>
<thead>
<tr>
<th>Number of Siblings (x)</th>
<th>Frequency (f)</th>
<th>( x \cdot f )</th>
<th>( x^2 \cdot f )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1. Find the mean of the data using \( f \)
2. Find the variance and standard deviation

\[
s = \sqrt{\frac{n(\sum x^2 f) - (\sum xf)^2}{n(n-1)}}
\]
# ACTIVITY: FREQUENCY TABLE

<table>
<thead>
<tr>
<th>Number of Siblings (x)</th>
<th>Frequency (f)</th>
<th>$x \cdot f$</th>
<th>$x^2 \cdot f$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>4</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>6</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>6</td>
<td>12</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>6</td>
<td>18</td>
</tr>
<tr>
<td>4</td>
<td>6</td>
<td>24</td>
<td>96</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>5</td>
<td>25</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>22</strong></td>
<td><strong>47</strong></td>
<td><strong>157</strong></td>
</tr>
</tbody>
</table>

1. Find the mean of the data using \( f \)

\[
\bar{x} = \frac{47}{22} = 2.14
\]

2. Find the variance and standard deviation

\[
s^2 = \frac{22 \cdot 157 - 47^2}{22 \cdot 21} = 2.694
\]

\[
s = \sqrt{2.694} = 1.642
\]

\[
s = \sqrt{\frac{n(\sum (x^2 \cdot f) - (\sum x \cdot f)^2)}{n(n-1)}}
\]
Prof. Lee gives a multiple choice quiz, which is composed to a total of four choice, and only one answer is correct. James is one of his student, but he did not study, thus, he doesn’t understand any question.
1. If the quiz has only one question, what is the probability that he chooses the correct answer?

All students found the answer is $\frac{1}{4}$ easily.
2. If the quiz has two questions,
   A. What is the probability that he choose all correct answers?
   B. What is the probability that he guess one question correctly?
The most of the group thought
- The answer is 2/8, means 1/4

Instructor: Is the probability getting one question correct among one question and two questions?
- Students realized that it should not be the same.
One group found all the possibilities, and found the answer

- \( P(\text{All questions correct}) = \frac{1}{16} \)
- \( P(\text{one question correct}) = \frac{6}{16} = \frac{3}{8} \)

Instructor: Can you find a different way?
3. If the quiz has three questions,
   A. What is the probability that he guesses three questions correctly?
   B. What is the probability that he guesses one question correctly?
   C. What is the probability that he guesses two questions correctly?
   D. What is the probability that he guesses at least one question correctly?
After discussing using the probability, students started using the probabilities.

Some students started thinking a shortcut using the combination property: \( \text{nCr} \)

Instructor: Why do you think to use the combination for this problem?

Student A: Because \( P(C,I,I) \), \( P(I,C,I) \), and \( P(I,I,C) \) are all the same cases, so the question did not matter with the order.
After the activity, introducing the binomial probability theorem:

\[ P(x) = \binom{n}{x} p^x q^{n-x} = \frac{n!}{(n-x)!x!} p^x q^{n-x} \]
Sixty five percent of BMCC students responds that they are the fan of New York Yankees. You randomly select 10 students and ask each if she/he considers herself/himself a New York Yankee’s fan.

1. Find the probability that exactly 8 students responds “Yes”
2. Find the probability that between 6 and 8 students responds “Yes”
3. Find the probability that at least 2 students responds “Yes”
Students understood the Binomial Probability
Realized that Binomial Probability generate one case each time
Finding more than one case, they need to find all different cases, and combined them.
Showing the complementary event:

\[ P(b) = 1 - P(\text{not } b) \]
SOLVING PROBLEMS USING BINOMIAL PROBABILITY THEOREM

MAT 150 Introduction to Statistics

Scenario 2
Sixty-five percent of BMCC students responds that they are the fan of New York Yankees. You randomly select 10 students and ask each if she/he considers herself/himself a New York Yankee's fan.

4. Find the probability that exactly 8 students responds "Yes"
5. Find the probability that between 6 and 8 students responds "Yes"
6. Find the probability that at least 2 students responds "Yes"

4. \[ P(X = 8) = \binom{10}{8} \left( \frac{65}{100} \right)^8 \left( \frac{35}{100} \right)^2 \]
5. \[ P(6 \leq X \leq 8) = \sum_{x=6}^{8} \binom{10}{x} \left( \frac{65}{100} \right)^x \left( \frac{35}{100} \right)^{10-x} \]
6. \[ P(X \geq 2) = 1 - P(X < 2) = 1 - \left( \binom{10}{0} \left( \frac{65}{100} \right)^0 \left( \frac{35}{100} \right)^{10} + \binom{10}{1} \left( \frac{65}{100} \right)^1 \left( \frac{35}{100} \right)^9 \right) \]

Example:
\[ P(X = 3) = \binom{10}{3} \left( \frac{65}{100} \right)^3 \left( \frac{35}{100} \right)^7 \]

Analyze the results and discuss the implications of the probabilities calculated.
Claire and Susan are very close friends with each other. They are going to the same college, and have been taking many classes together. Although, they are best friends, on the other hand, they are always rivals. One semester, they are taking statistics classes, but different classes with different professors. After the first exam, Claire got 75, and Susan got 77 on their first exam. The average of the Claire’s class is 68 with the standard deviation 8.2, and the average of the Susan’s class is 70 with the standard deviation 6.7

Questions: **Who got the higher grade on the first exam based on the data distribution?**
Susan got higher grade: her score is greater．

Their score is the same because the mean deviation is equal to each other.

What about the standard deviation?
NORMAL DISTRIBUTION
NORMAL DISTRIBUTION

One group began to use the z-score

Watch the online lecture video
Students found out that

- A smaller standard deviation generates
  - narrower and taller normal curve
  - Greater z-score
CLASSROOM ACTIVITY
CLASSROOM ACTIVITY
CLASSROOM ACTIVITY
HYPOTHESIS PROJECT

Topic: Number of Siblings

Based on the previous Prof. Lee’s statistic project, the mean number of siblings is 3, with the standard deviation of 0.85.

Day 1
1. Do you agree with the claim?
2. Set up the Hypothesis set up.
3. Collect 15 data: You can search it using the internet, taking survey, or other resources.

Day 2
4. Decide the Hypothesis Test method (z-test, P-value test, or t-test)
5. Using the Hypothesis Test, Conclude the result

Day 3
Create the Power Point Slide and Hand Out paper

Day 4
Review the Project Document
Topic: Number of Siblings

Based on the previous Prof. Lee’s statistic project, the mean number of siblings is 3, with the standard deviation of 0.85.

Day 1

1. Do you agree with the claim?
2. Set up the Hypothesis set up.
3. Collect 15 data: You can search it using the internet, taking survey, or other resources.

Day 2

4. Decide the Hypothesis Test method (z-test, P-value test, or t-test)
5. Using the Hypothesis Test, Conclude the result

Day 3

Create the Power Point Slide and Hand Out paper

Day 4

Review the Project Document
HYPOTHESIS
PROJECT
The mean number of siblings is 3, with the standard deviation of 0.85.
YES WE AGREE WITH THE CLAIM
$H_0: \mu = 3$

Claim

$H_a: \mu \neq 3$
We decided to use the t-test method because our data set is smaller than 30.
\[ N < 15 \]

\[ \mu = 3 \]

\[ \bar{x} = 2.9333 \]

\[ S = 0.85 \]

\[ df = 14 - 2.145 \]
- Fail to reject Ho due to the lack of evidence to reject the claim
- There is not enough evidence to support the claim
CONCLUSION

• The mean number of siblings is 3
STUDENTS’ PRESENTATION

- Students must use
  - Power Point Slide show
  - Wear Formal dress code
  - Divide role to present, so that all group members present.
STUDENTS’ PRESENTATION
STUDENTS’ PRESENTATION
P-VALUE TEST

- We used excel to calculate mean, variance and standard deviation.
- N=40  X=7.7  s=8.7853  α=0.05
- P-value test: Z = \frac{X - \mu}{\sigma/\sqrt{n}} = \frac{7.7 - 8.5}{8.7853/\sqrt{40}} = -0.64
- Z(-0.64) - P = 0.2611 > α = 0.05
- Ho: μ ≥ 8.5
- Fail to Reject Ho, Claim is incorrect
- Ha: μ < 8.5

Z-TEST

- Z = -0.64  α = 0.05 = -1.645  n = ∞  Ho ↓
- Because -0.64 belongs to Ho are,
- Fail to Reject Ho
- Claim is Incorrect
- Ho: μ ≥ 8.5
- Ha: μ < 8.5  Claim
Results

- \( \mu = 20 \), \( \sigma = 0.10 \), \( \sigma = 1.46507 \), \( 1.4651 \), \( n = 40 \)

- \( \bar{x} = 24.20 \), \( s = 1.72672 \)

- \( z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \frac{24.20 - 20}{1.4651/\sqrt{40}} \)

- Z-test: \( z = -7.2672 < -2.326 \)
- P-test: \( p = 0.0002 < 0.10 \)
# Grading Distribution

<table>
<thead>
<tr>
<th>Category</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>A- ~ A</td>
<td>11</td>
</tr>
<tr>
<td>B- ~ B+</td>
<td>4</td>
</tr>
<tr>
<td>C- ~ C+</td>
<td>5</td>
</tr>
<tr>
<td>D or Below D</td>
<td>2</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>22</strong></td>
</tr>
</tbody>
</table>

**Chart Title**

- **Series 1**: Frequency distribution of grades
- **Column1**: A- ~ A
- **Column2**: B- ~ B+, C- ~ C+, D or below D

![Chart showing distribution of grades]
STUDENTS’ FEEDBACKS

- Enjoyed your class!!
- I did not even realized the semester is over.
- Keep doing this professor.
- I truly learned statistics. I can analyze other hypothesis test.
# STATISTIC ANALYSIS

<table>
<thead>
<tr>
<th>Group Statistics</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>group</td>
<td>N</td>
<td>Mean</td>
<td>Std. Deviation</td>
</tr>
<tr>
<td>Pre and post</td>
<td>posttest</td>
<td>24</td>
<td>62.0033</td>
<td>28.39851</td>
</tr>
<tr>
<td></td>
<td>pretest</td>
<td>24</td>
<td>45.7429</td>
<td>18.86852</td>
</tr>
</tbody>
</table>

## Independent Samples Test

<table>
<thead>
<tr>
<th></th>
<th>Levene's Test for Equality of Variances</th>
<th>t-test for Equality of Means</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>F</td>
<td>Sig.</td>
</tr>
</tbody>
</table>
Do you think the Hypothesis group project improve your understanding of the Hypothesis Test?

<table>
<thead>
<tr>
<th>A. STRONGLY AGREE</th>
<th>B. AGREE</th>
<th>C. SOMEWHAT AGREE</th>
<th>D. DISAGREE</th>
<th>E. STRONGLY DISAGREE</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>5</td>
<td>9</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>
After the Hypothesis Project, do you feel to conduct a hypothesis project if any situation is given to you?

- a. I am very confident to do it: 30%
- b. Maybe I can do it: 44%
- c. Somewhat I still need help: 22%
- d. No way: 4%

Legend:
- a. I am very confident to do it
- b. Maybe I can do it
- c. Somewhat I still need help
- d. No way
CONCLUSION

• Combining between Inquiry Based Learning and Project Based Learning is a great tool to teach Intro Statistics Class
  • IBL provided students to explore questions deeply and conceptually.
  • They developed own approaches: A student thought the combination nCr for Binomial Probability before the class discussed the topic.

• Students are engaged, and improved grade significantly
CONCLUSION

- Group Project provides them to explore statistical contents, and enhance to understand them
- Students had an experience to apply all the class materials into their own project
- The Hypothesis Test Project Provided all the students to be ready to answer for their presentation.
- Collaborative work
- Role play: All students must participate in group presentation.
- All group members must understand their presentation material thoroughly.
QUESTIONS?

Thank You

Contact Information
Jae Ki Lee, Ed.D
Associate Professor
Borough of Manhattan Community College
jaeelee@bmcc.cuny.edu