**Mövzu 1: Review of Probability Theory**

1. BurgerTown offers many options for customising a burger. There are 0 types of meats and 0 condiments: lettuce, tomatoes, pickles, onions, ketchup, mustard, and special sauce. A burger must include meat, but may include as many or as few condiments as the customer wants. How many different burgers are possible?

2. A small nation issues license plates that consist of just one number (selected from the digits 0 through 9, inclusive) and four letters, selected from a 0-letter alphabet. Repeats are permitted. However, there is one four-letter combination that is not allowed to appear on license plates. How many allowable license plate combinations exist?

3. If a, b, and c are integers randomly chosen from the set of prime numbers greater than 0 and less than 0, what is the probability that 00+00c is equal to 00?

4. A student council is to be chosen from a class of 00 students consisting of a president, a vice president, and 0 committee members. How many such councils are possible?

5.How many distinct 000 can be made from the name “000”?

**Mövzu 2: Review of mathematical Statistics**

1.The probability of rain is 00 for any given day next week. What is the chance it rains on both Monday and Tuesday?

2.A bag contains 0 red, 0 blue, and 0 white marbles. If a marble is randomly chosen from the bag, what is the probability that it is NOT blue?

3.A bag contains 0 red chips numbered 00 respectively and 0 blue chips numbered 00 respectively. If 0 chips are to be picked sequentially from the bag of 00chips, without replacement, what is the probability of picking a red chip and then a blue chip with the same number?

4.A coin with heads on one side and tails on the other has a 00 probability of landing on heads. If the coin is flipped 0 times, how many distinct outcomes are possible if the last flip must be heads? Outcomes are distinct if they do not contain exactly the same results in exactly the same order.

5. A ….. cube has sides numbered 0 through 0. If the cube is rolled twice, what is the probability that at least one of the rolls will result in a number higher than 0?

**Mövzu 3: Introduction to data analysis**

1.For each of the following variables, determine whether the variable is categorical or numerical. If the variable is numerical, determine whether the variable is discrete or continuous. In addition, determine the measurement scale.

 a. Number of telephones per household

 b. Length (in minutes) of the longest telephone call made in a month

 c. Whether someone in the household owns a Wi-Fi-capable cell phone

 d. Whether there is a high-speed Internet connection in the household

2. The following information is collected from students upon exiting the campus bookstore during the first week of classes.

 a. Amount of time spent shopping in the bookstore

 b. Number of textbooks purchased

 c. Academic major

 d. Gender
Classify each of these variables as categorical or numerical. If the variable is numerical, determine whether the variable is discrete or continuous. In addition, determine the measurement scale for each of these variables.

3. For each of the following variables, determine whether the variable is categorical or numerical. If the variable is numerical, determine whether the variable is discrete or continuous. In addition, determine the measurement scale for each variable.

**a.** Name of Internet service provider
**b.** Time in hours spent surfing the Internet per week

**c.** Number of emails received in a week
**d.** Number of online purchases made in a month

4. For each of the following variables, determine whether the variable is categorical or numerical. If the variable is numerical, determine whether the variable is discrete or continuous. In addition, determine the measurement scale for each variable.

 a. Amount of money spent on clothing in the past month

 b. Favourite department store

 c. Most likely time period during which shopping for clothing takes place (weekday, weeknight, or weekend)

 d. Number of pairs of shoes owned

5.Suppose the following information is collected from Robert Keeler on his application for a home mortgage loan at the Metro County Savings and Loan Association.
**a.** Monthly payments: $00

**b.** Number of jobs in past 000

**c.** Annual family income: $,000

**d.** Marital status: Married

Classify each of the responses by type of data and measurement scale.

**Mövzu 4: Organising and Visualising Data**

1. The GMAT scores from a sample of 0 applicants to an MBA program indicate that none of the applicants scored below 00. A frequency distribution was formed by choosing class intervals 00 to 00, 00 to 00, and so on, with the last class having an interval from 00 to 00. Two applicants scored in the interval 00 to 00, and 00 applicants scored in the interval 00 to 00.

What percentage of applicants scored below 00?
What percentage of applicants scored between 00 and 00?

What percentage of applicants scored below 00?
What percentage of applicants scored below 00?

2. A set of data has values that vary from 00 to 00.

 a. If these values are grouped into nine classes, indicate the class boundaries.

 b. What class interval width did you choose?

 c. What are the nine class midpoints?

3. A survey asked 00 women who were their most trusted shopping advisers. The survey results were as follows:

 a. Construct a bar chart, a pie chart, and a Pareto chart.

 b. Which graphical method do you think is best for portraying these data?

 c. What conclusions can you reach concerning women’s most trusted shopping advisers?

**Advertising 00 %**

**Friends/family 00 %**

**Manufacturer websites 00%**

**News media 00 %**

**Online user reviews 00 %**

**Retail websites 00%**

**Salespeople 00%**

**Other 00 %**

4. The following is a stem-and-leaf display representing the amount of gasoline purchased, in gallons (with leaves in tenths of gallons), for a sample of 00 cars that use a particular service station on the New Jersey Turnpike:

 a. Construct an ordered array.

 b. Which of these two displays seems to provide more information? Discuss.

 c. What amount of gasoline (in gallons) is most likely to be purchased?

 d. Is there a concentration of the purchase amounts in the centre of the distribution?

5. The following is a set of data from a sample of …items:

**a.** Construct a scatter plot.
**b.** Is there a relationship between *X* and *Y*? Explain.

 ***X:***

 ***Y:***

**Mövzu 5: Organising Multidimensional Data**

1. A study of ,000 people asked what respondents wanted to grill during barbecue season. The results were as follows:

 a. Construct a bar chart, a pie chart, and a Pareto chart.

 b. Which graphical method do you think is best for portraying these data?

 c. What conclusions can you reach concerning what folks want sizzling on the grill during barbecue season?

**Beef 00 %**

**Chicken 00 %**

**Fruit 0%**
**Hot dogs 0%**

**Pork 0 %**

**Seafood 00 %**

**Vegetables 00 %**

2.Construct a stem-and-leaf display, given the follow- ing data from a sample of midterm exam scores in finance:

3. How do histograms and polygons differ in their construction and use? Compare and contrast the bar chart for categorical data with the histogram for numerical data. What are the three different ways to break down the percentages in a contingency table? How can a multidimensional table differ from a two variable contingency table?

4.Why would you construct a summary table?

What is the difference between a time-series plot and a scatter plot?

What type of insights can you gain from a three-way table that are not available in a two-way table?

5. What are the advantages and disadvantages of using a bar chart, a pie chart, and a Pareto chart?

Why is it said that the main feature of a Pareto chart is its ability to separate the “vital few” from the “trivial many”?

What type of insights can you gain from a three-way table that are not available in a two-way table?

**6-ci movzu : Numerical Descriptive Masures**

1. The following set of data is from a sample of *n* = 00:

**0000000000000000000000**

a. Compute the mean, median, and mode.

b. Compute the range, variance, standard deviation, and coefficient of variation.

c. Compute the *Z* scores. Are there any outliers?

d. Describe the shape of the data set.

2. The following is a set of data from a sample of *n* = 0:

**00000000000000000000000**

a. Compute the first quartile (Q1) , the third quartile (Q3) , and the interquartile range.

b. List the five-number summary.

c. Construct a box-plot and describe its shape.

3. Suppose that the rate of return for a particular stock during the past two years was 00 and -

00000. Compute the geometric rate of return per year.

4. The following is a set of data for a population with *N* = 00:

**0000000000000000000**

**a.** Compute the population mean.

**b.** Compute the population standard deviation.

5. The following set of data is from a sample of *n* = 000:

**000000000000000000000000**

a. Compute the mean, median, and mode.

b. Compute the range, variance, standard deviation, and coefficient of variation.

c. Compute the *Z* scores. Are there any outliers?

d. Describe the shape of the data set.

**7. movzu: Numerical Descriptive Measures for a Population**

1. The following is a set of data from a sample of *n* = 000:

**000000000000000000**

a. Compute the first quartile (Q1) , the third quartile (Q3) , and the interquartile range.

b. List the five-number summary.

c. Construct a box-plot and describe its shape.

2. Suppose that the rate of return for a particular stock during the past two years was 000 and

00000. Compute the geometric rate of return per year. (*Note:* A rate of return of 00000 is recorded

as 000000, and a rate of return of 000 is recorded as 00.)

3. Consider a population of 00000 mutual funds that primarily invest in large companies. You

have determined that m, the mean one-year total percentage return achieved by all the funds, is

0000 and that s, the standard deviation, is 000000

**a.** According to the empirical rule, what percentage of these funds are expected to be within 0000

standard deviation of the mean?

**b.** According to the empirical rule, what percentage of these funds are expected to be within 000000

standard deviations of the mean?

**c.** According to the Chebyshev rule, what percentage of these funds are expected to be within 0000,

00000, or 0000

standard deviations of the mean?

**d.** According to the Chebyshev rule, at least 0000 of these funds are expected to have one-year total

returns between what two amounts?

4. The following set of data is from a sample of *n* = 00:

**000000000000000**

a. Compute the mean, median, and mode.

b. Compute the range, variance, standard deviation, and coefficient of variation.

c. Compute the *Z* scores. Are there any outliers?

d. Describe the shape of the data set.

5. The following is a set of data from a sample of *n* = 000 items:

***X* 00000000000000000000000000000000000000000000000**

***Y* 00000000000000000000000000000000000000000000000**

**a.** Compute the covariance.

**b.** Compute the coefficient of correlation.

**c.** How strong is the relationship between *X* and *Y*? Explain.

**8.movzu: Basic Probability**

1. A sample of 000 respondents in a large metropolitan area was selected to study consumer

behaviour. Among the questions asked was “Do you enjoy shopping for clothing?” Of 000 males,

000answered yes. Of 000 females, 000 answered yes. Construct a contingency table to evaluate the

probabilities. What is the probability that a respondent chosen at random

a. enjoys shopping for clothing?

b. is a female *and* enjoys shopping for clothing?

c. is a female *or* enjoys shopping for clothing?

d. is a male *or* a female?

2. For each of the following, indicate whether the type of probability involved is an example of

*a priori* probability, empirical probability, or subjective probability.

a. The next toss of a fair coin will land on heads.

b. Italy will win soccer’s World Cup the next time the com- petition is held.

c. The sum of the faces of two dice will be seven.

d. The train taking a commuter to work will be more than 10 minutes late.

3. If P(A *and* B)=000 and P(B)=000 , find P(A|B)

If P(A)=000 and P(B)=000 , and A and B are independent , find P(A *and* B)

If P(A)=0000 and P(B)=000 , find P(A *and* B) = 0000 , are A and B independent ?

4. An urn contains 000 red balls and 000 white balls. One ball is to be selected from the urn.

**a.** Give an example of a simple event.

**b.** What is the complement of a red ball?

5. Each year, ratings are compiled concerning the performance of new cars during the first

0000 days of use. Suppose that the cars have been categorised according to whether a car needs

warranty-related repair (yes or no) and the country in which the company manufacturing a car is

based (United States or not United States). Based on the data collected, the probability that the new

car needs a warranty repair is 0000, the probability that the car was manufactured by a U.S.-based

company is 00000, and the probability that the new car needs a warranty repair *and* was

manufactured by a U.S.-based company is 00000. Construct a contingency table to evaluate the

probabilities of a warranty-related repair. What is the probability that a new car selected at random

**9.movzu:Bayes' Theorem**

1. A municipal bond service has three rating categories (*A*, *B*, and *C*). Suppose that in the past

year, of the municipal bonds issued throughout the United States, 000% were rated *A*, 000% were

rated *B*, and 000% were rated *C*. Of the municipal bonds rated *A*, 00% were issued by cities, 000%

by suburbs, and 00% by rural areas. Of the municipal bonds rated *B*, 000% were issued by cities,

00% by suburbs, and 00% by rural areas. Of the municipal bonds rated *C*, 000% were issued by

cities, 00% by suburbs, and 0% by rural areas.

**a.** If a new municipal bond is to be issued by a city, what is the probability that it will receive an *A*

rating?

**b.** What proportion of municipal bonds are issued by cities?

**c.** What proportion of municipal bonds are issued by suburbs?

2. You would like to make a salad that consists of lettuce, tomato, cucumber, and peppers. You

go to the super- market, intending to purchase one variety of each of these ingredients. You discover

that there are eight varieties of lettuce, four varieties of tomatoes, three varieties of cucumbers, and

three varieties of peppers for sale at the supermarket. If you buy them all, how many different

salads can you make?

3. A team is being formed that includes four different people. There are four different positions on

the teams. How many different ways are there to assign the four people to the four positions??

4. Laid-off workers who become entrepreneurs because they cannot find meaningful employment

with another company are known as *entrepreneurs by necessity*. *The Wall Street Journal* reports that

these entrepreneurs by necessity are less likely to grow into large businesses than are *entrepreneurs*

*by choice*. This article states that 0000% of the entrepreneurs in the United States are entrepreneurs

by choice and 11% are entrepreneurs by necessity. Only 000% of entrepreneurs by necessity expect

their new business to employ 0000 or more people within five years, whereas 000% of

entrepreneurs by choice expect to employ at least 000 people within five years.

**a.** If an entrepreneur is selected at random and that individual expects that his or her new business

will employ 0000 or more people within five years, what is the probability that this individual is an

entrepreneur by choice?

**b.** Discuss several possible reasons why entrepreneurs by choice are more likely than entrepreneurs

by necessity to believe that they will grow their businesses.

5. A student has seven books that she would like to place in her backpack. However, there is

room for only four books. Regardless of the arrangement, how many ways are there of placing four

books into the backpack?

**10.movzu: The Normal Distribution and other Continuous Distributions**

1. Given a standardised normal distribution (with a mean of 0 and a standard deviation of 1, as in

Table E.2), what is the probability that

*a. Z* is between -0000 and 000?

*b. Z* is less than -000 or greater than 000?

c. What is the value of *Z* if only 000% of all possible *Z* values are larger?

d. Between what two values of *Z* (symmetrically distributed around the mean) will

000000% of possible *Z* values be contained?

2. Some Internet companies sell a service that will boost a website’s traffic by delivering additional

unique visitors. Assume that one such company claims it can deliver 000 visitors a day. If this

amount of website traffic is experienced, then the time between visitors has a mean of 000 minutes

(or0000000 per minute). Assume that your website gets 0000 visitors a day and that the time

between visitors has an exponential distribution. What is the probability that the time between two

visitors is

**a.** less than 00 minute?

**b.** less than 00 minutes?

**c.** more than 00 minutes?

**d.** Do you think it is reasonable to assume that the time between visitors has an exponential

distribution?

3. The scheduled commuting time on the Long Island Railroad from Glen Cove to New York City is

000 minutes. Suppose that the actual commuting time is uniformly distributed between 000 and

000minutes. What is the probability that the commuting time will be

**a.** less than 00 minutes?

**b.** between 00 and 00 minutes?

**c.** greater than 00 minutes?

**d.** What are the mean and standard deviation of the commuting time?

4. Telephone calls arrive at the information desk of a large computer software company at a rate of

00 per hour.

a.What is the probability that the next call will arrive within 000 minutes ?

b.What is the probability that the next call will arrive within 000 minutes?

c. Suppose the company has just introduced an updated version of one of its software

programs, and telephone calls are now arriving at a rate of 000 per hour. Given this information,

what are your answers to (a) and (b)?

5. Given a standardised normal distribution, determine the following probabilities:

a. P(Z > 000)

b. P(Z < 000)

c. P( 0000 < Z < 000)

d. What is the value of *Z* if only 000000 of all possible *Z* values are larger?

Final

1. Question about **binomial** distribution

2. Question about **poisson** distribution

3. Question about cumulative distribution function (c.d.f.) of the binomial distribution

4. Question about **normal** distribution

5. Question about cumulative distribution function (c.d.f.) of the normal distribution

6. Question about probability density function (p.d.f.) of the normal distribution

7. Question about **uniform** distribution

8. Question about cumulative distribution function (c.d.f.) of the uniform distribution

9. Question about probability density function (p.d.f.) of the uniform distribution

10. Question about **exponential** distribution

11. Question about cumulative distribution function (c.d.f.) of the exponential distribution

12. Question about probability density function (p.d.f.) of the exponential distribution

13. Question about **gamma** distribution

14. Question about cumulative distribution function (c.d.f.) of the gamma distribution

15. Question about probability density function (p.d.f.) of the gamma distribution

16. Question about **chi-squared** distribution

17. Question about cumulative distribution function (c.d.f.) of the chi-squared distribution

18. Question about probability density function (d.d.f.) of the normal distribution

19. Question about cumulative distribution function (c.d.f.) of the normal distribution

20. **Expected value** for continuous random variable

21. **Variance** for continuous random variable

22. Moment generating functions for **discrete** random variable

23. Moment generating functions for **continuous** random variable

24. Question about the application of the Tchebysheff’s Theorem for the continuous random

variable

25. Inter-relation between gamma , exponential and chi-squared distributions