**BNAD/ECON/MGMT 276. STATISTICAL INFERENCE IN MANAGEMENT**

**Presession. Summer 2016.**

**FINAL EXAM. DATE: 3 JUNE 2016**

**Time: 120 minutes (2 hours)**

NOTE:

1. There are total 28 questions of which there are 80 sub-questions.
2. You can use calculator. Your calculator may be programmable but you are not allowed to use the programming tools.
3. There are 3 parts:

Part 1: Mandatory. You must answer all questions in this part, i.e. question 1 to question 16. The total score of this part is 53.

Part 2: Optional: question 17 to question 23. Several questions have sub-questions, eg. 17.1, 17.2, etc. There are 23 sub-questions in this part. Each sub-question worths 1 point. **You can choose to answer how many sub-questions (N)** and which sub-questions. For example, you can choose to answer only 17.1 and 18.2. In the end, **your total score of this part is the correct credits within the N sub-questions you've chosen.** When you're about to submit your exam, please write down explicitly here which sub-questions you chose to be counted into the final score (or clearly state "counted" in those sub-questions):

Your N = …….. (from 0 to 23), they are …….

Part 3: Extra credits. There are 7 sub-questions in this part. Each worths 1 point.

**You final score of the exam = (mandatory score + optional score + extra credits) / (53 + N)**

Good luck!

---

# PART I. MANDATORY PART. YOU MUST ANSWER ALL QUESTIONS IN THIS PART.

**Question 1.** (3 points) You are given 6 formulas/descriptions/notations on the left and 9 statistical concepts on the right column. Please match each formulas/description to its corresponding statistical concepts. Notice there are **5 redundant concepts**, i.e. 5 concepts that do not have the corresponding descriptions.

Hence, you need to **match 6 formulas/descriptions to the most correct 6 concepts**. Each correct match worths 0.5 point.

1. $\frac{s_{xy}}{s_x s_y}$
2. $s_{xy} = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{n-1}$
3. $A \cup B$
4. $\frac{\sum X_i}{n}$
5. $P(A^c) = 1 - P(A)$
6. $s_X^2 = \frac{\sum(X_i - \bar{X})^2}{n-1}$

A. Sample mean
B. Mean
C. Sample variance
D. Variance
E. Standard deviation
F. Sample covariance
G. Correlation coefficient
H. Union event
I. Intersection event
J. Conditional event
K. Complement event

Please write down your answer as specific as 1A, 2B, 3C… below:

1.

2.

3.

4.

5.

6.

**Question 2.** **(4 points)** You are given 8 formulas/descriptions/notations on the left and 11 statistical concepts on the right column. Please match each formulas/description to its corresponding statistical concepts. Notice there are **3 redundant concepts**, i.e. 3 concepts that do not have the corresponding descriptions.

Hence, you need to **match 8 formulas/descriptions to the correct 8 concepts**. Each correct match worths 0.5 point.

1. $f(x) = \frac{1}{b-a} \ for \ a \leq x \leq b$
2. $bell\ curve/bell\ shape$
3. $\sum xf(x)$
4. $\sigma$
5. $\sum x^2 f(x) - \mu^2$
6. $f(x) = \binom{n}{x} p^x (1-p)^{n-x}$
7. $\sum x^2 f(x)$
8. Number of occurrences in a time period

A. Bernoulli distribution
B. Binomial distribution
C. Poisson distribution
D. Normal distribution
E. Uniform distribution
F. Standard deviation
G. Sample variance
H. Mean
I. Sample mean
J. Variance
K. Second moment, $E(X^2)$

Please write down your answer as specific as 1A, 2B, 3C… below:

1.

2.

3.

4.

5.

6.

7.

8.

**Question 3.** (1 pt) The measure of dispersion which is not measured in the same units as the original data is the… (Choose one correct answer among the belows)

a. correlation coefficient
b. standard deviation
c. coefficient of determination
d. variance

## Question 4. (3 pts) The following data is used for questions 4.1, 4.2, 4.3

The closing stock price of 7 corporations on Friday is shown below.

| Corporation | Stock Price |
|---|---|
| A | 83 |
| B | 87 |
| C | 83 |
| D | 89 |
| E | 85 |
| F | 90 |
| G | 91 |

4.1 (1 point) What is the type of this data set? (Choose a or b or c or d)

a. Cross-sectional data
b. Time-series data
c. Panel data
d. Longitudinal data

4.2 (1 point) What is the **mean** value of the stock price?

4.3 (1 point) What is the **median** value of the stock price?

**Question 5. (6 pts) The following data set is used for questions 5.1, 5.2, 5.3, 5.4, 5.5, 5.6**

The ABC Company has reported the number of commercial advertisings and the sale revenues as follows:

| Month | Advertisings (campaigns) | Sale Revenues($) |
|---|---|---|
| 1 | 33 | 40 |
| 2 | 27 | 30 |
| 3 | 7 | 10 |
| 4 | 15 | 20 |
| 5 | 20 | 25 |

You may find the below mysterious table useful to answer questions.

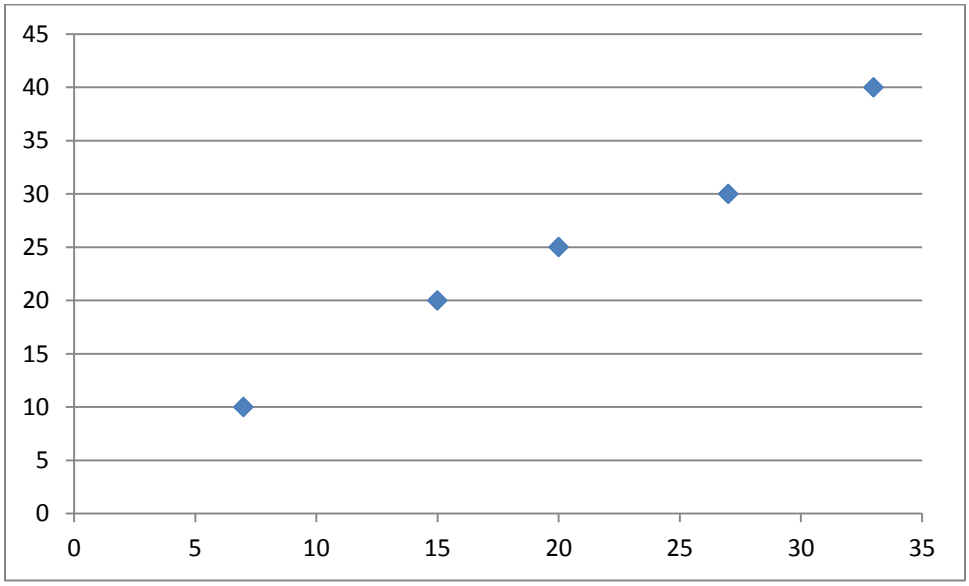
| | | | | | | |
|---|---|---|---|---|---|---|
| 33 | 40 | 12.6 | 15 | 189 | 158.76 | 225 |
| 27 | 30 | 6.6 | 5 | 33 | 43.56 | 25 |
| 7 | 10 | -13.4 | -15 | 201 | 179.56 | 225 |
| 15 | 20 | -5.4 | -5 | 27 | 29.16 | 25 |
| 20 | 25 | -0.4 | 0 | 0 | 0.16 | 0 |
| | | | | | | |

5.1. What is the **sample variance** of the **number of advertisings**? (1 point)

5.2. What is the **sample variance** of the **sales revenue**? (1 point)

5.3. Calculate the **sample covariance** of the Advertising Campaigns and the Sales Revenues? (1 point).

5.4. What is the name of the plot below? (1 point)



5.5. (1 pt) What do you think about the **sign and the magnitude** of the **sample correlation coefficient** of the Price Per Gallon and the Number of Gallons? Specifically, is the sample correlation coefficient negative, positive? Is it near to - 1, 0, or 1? (You can either calculate the correlation coefficient to answer this question, or you can look at the plot in the above question to answer this question.)

5.6. (1 pt) Which of the following statement is the **most complete correct**?

a. The sample correlation coefficient of the two variables (number of advertisings and sales revenues) is positive. This means that an increase in the number of advertising is associated with an increase in the sales revenue.
b. The sample covariance of the two variables (number of advertisings and sales revenues) is positive. This means that the two variables have a positive relationship.
c. The sample correlation coefficient of the two variables (number of advertisings and sales revenues) is positive and close to 1. This means that an increase in the number of advertising leads to an increase in the sales revenue (causal relationship).
d. The sample correlation coefficient of the two variables (number of advertisings and sales revenues) is negative and close to - 1. This means that an increase in the number of advertising leads to a decrease in the sales revenue (causal relationship).

**Question 6. (3 points)** Suppose Z is standard normal random variable, and let $\sigma$ be the standard deviation of Z.

6.1. (1 pt) What is the probability that P(Z < 1.12)?

**6.2.**(1 pt) What is the probability that P( Z > 2.11)?

6.3.(1pt) What is the probability that Z is between $-2\sigma$ and $2\sigma$ ?

**Question 7. (3 points)** Suppose that a random variable X has the following probability density function:

$$f(x) = \frac{1}{40\sqrt{\pi}} e^{\frac{-(x-90)^2}{1600}}$$

7.1. (1 pt) What is the name of the distribution above? What is the mean value of X?

7.2. (1 pt) What is the standard deviation of X?

7.3. (1 pt) Draw the pdf $f(x)$above. Put $x$ on the horizontal axis and $f(x)$ on the vertical axis. Indicate clearly the **location of the mean**. (You don't need to draw a very precise graph. You only need to show how it looks like and the location of the mean on the graph)

**Question 8.** (1pt) Is the following a valid probability mass function? Explain why.

| f(x) | x | |
|---|---|---|
| 0.2 | 0.2 | |
| −0.2 | 0 | |
| 0.5 | 0.4 | |
| 0.5 | 0.4 | |
| | | Total |

**Question 9.** (1pt) A machine is designed to fill toothpaste tubes with 5.8 ounces of toothpaste. The manufacturer does not want any underfilling or overfilling. The correct hypotheses to be tested are

a. $H_0$: $\mu \neq 5.8$ $H_a$: $\mu = 5.8$

b. $H_0$: $\mu = 5.8$ $H_a$: $\mu \neq 5.8$

c. $H_0$: $\mu > 5.8$ $H_a$: $\mu \leq 5.8$

d. $H_0$: $\mu \geq 5.8$ $H_a$: $\mu < 5.8$

**Question 10.** (1 pt) Is this probability mass function valid? Explain why?

| f(x) | x | |
|---|---|---|
| 0.2 | −1 | |
| 0.1 | 0 | |
| 0.6 | 2 | |
| | | Total |

**Question 11.** (3pts) An experimental diet to induce weight loss was followed for three months by a randomly selected group of 6 students with the following results.

| Student | Loss in Pounds |
|---|---|
| 1 | 15 |
| 2 | 16 |
| 3 | 18 |
| 4 | 15 |
| 5 | 14 |
| 6 | 12 |

**11.1.**(1pt) Find a point estimate for the average amount lost after three months on this diet.

**11.2.** (1pt) Find a point estimate for the variance of the amount lost on this diet. (Hint: you are given the mysterious table below. Please check carefully before using the table.)

| Student | Loss in Pounds | | | | |
|---|---|---|---|---|---|
| 1 | 15 | 0 | 0 | 0 | 0 |
| 2 | 16 | 1 | 1 | 1 | 1 |
| 3 | 18 | 9 | 3 | 81 | 27 |
| 4 | 15 | 0 | 0 | 0 | 0 |
| 5 | 14 | 1 | -1 | 1 | -1 |
| 6 | 12 | 9 | -3 | 81 | -27 |
| | | | | | |

**11.3.** (1pt) Find a point estimate for the standard deviation of the amount lost on this diet.

**Question 12.** (2pts) Assume that the population variance is 40, the population mean is 30, and the sample size is 150.

**12.1.** (1pt) What is the sampling distribution of the sample mean? Choose one answer:

a. Normal distribution with mean 30 and standard deviation $\frac{40}{150}$

b. Normal distribution with mean 30 and standard deviation $\sqrt{\frac{40}{150}}$

c. Normal distribution with mean 30 and variance $\frac{40}{\sqrt{150}}$

d. Normal distribution with mean 30 and variance $\frac{\sqrt{40}}{150}$

e. Normal distribution with mean 30 and variance $\sqrt{\frac{40}{150}}$

**12.2.** (1pt) What is the probability that the sample mean will be less than 50? Let Z follow standard normal distribution. Choose one answer:

a. $P(Z < \frac{30-50}{\frac{40}{150}})$

b. $P(Z < \frac{30-50}{\sqrt{\frac{40}{150}}})$

c. $P(Z < \frac{30-50}{\frac{40}{\sqrt{150}}})$

d. $1 - P(Z < \frac{30-50}{\frac{40}{\sqrt{150}}})$

e. $P(Z < \frac{30-50}{\frac{\sqrt{40}}{150}})$

**Question 13.** (2 pts) It is stated that 60% of primary care doctors think their patients receive unnecessary medical care (Reader's Digest, December 2011/January 2012). Suppose a sample of 200 primary care doctors were taken.

**13.1.** (1pt) Show the sampling distribution of the sample proportion of the doctors who think their patients receive unnecessary medical care. Clearly state the name of the distribution, the mean and the variance of the distribution.

**13.2.** (1pt) What is the probability that the sample proportion will be greater than 0.3? Let Z be the standard normal distribution. Choose one answer among the followings:

a. $P(Z < \frac{0.6-0.3}{\frac{\sqrt{0.6\times0.4}}{200}})$ since N = 200, p = 0.6

b. $P(Z < \frac{0.6-0.3}{0.0346})$

c. $1 - P(Z < \frac{0.6-0.3}{0.0346})$

d. $1 - P(Z < \frac{0.6-0.3}{\sqrt{0.6\times0.4}})$

e. $P(Z < \frac{0.6-0.3}{\frac{0.6\times0.4}{200}})$ since N = 200, p = 0.6

**Question 14.** (2 pts) In order to determine the average weight of carry-on luggage by passengers in airplanes, a sample of 15 pieces of carry-on luggage was weighed. The average sample weight was 20 pounds.

**14.1.** (1pt) Determine a 90% confidence interval estimate for the mean weight of the carry-on luggage. Assume that we know the standard deviation of the population to be 8 pounds. (The final result must be an interval [a,b]; you can leave your final result like the form in question 14.2, i.e. without giving a single specific number in the end.)

**14.2.** (1pt) Determine a 90% confidence interval estimate for the mean weight of the carry-on luggage. Assume that the sample standard deviation is 8 pounds.

a. $\left[20 - 1.960 \times \frac{4}{\sqrt{15}};\ 20 + 1.960 \times \frac{4}{\sqrt{15}}\right]$

b. $\left[20 - 1.645 \times \frac{4}{\sqrt{15}};\ 20 + 1.645 \times \frac{4}{\sqrt{15}}\right]$

c. $\left[20 - 1.761 \times \frac{4}{\sqrt{14}};\ 20 + 1.761 \times \frac{4}{\sqrt{14}}\right]$

d. $\left[20 - 1.761 \times \frac{4}{\sqrt{15}};\ 20 + 1.761 \times \frac{4}{\sqrt{15}}\right]$

e. $\left[20 - 1.753 \times \frac{4}{\sqrt{14}};\ 20 + 1.753 \times \frac{4}{\sqrt{14}}\right]$

f. $\left[20 - 1.753 \times \frac{4}{\sqrt{15}};\ 20 + 1.753 \times \frac{4}{\sqrt{15}}\right]$

**Question 15.** (4 pts) A student believes that the average grade on the statistics final examination was 85. A sample of 36 final examinations was taken. The average grade in the sample was 80. Assume the population standard deviation is 12.

15.1.(1pt) State the null and alternative hypotheses testing the student's belief.

15.2 (1pt) Compute the test statistic. Clearly state whether it is the test statistic t or test statistic z.

15.3. (2pts) Test the hypotheses at the level of significance of 5%. You can use whatever approach.

**WARNING: The multiple choice sub-questions in the question 16 have more than one correct answer. Read the sub-questions carefully.**

**Question 16.** (14 pts) Suppose X follows the normal distribution with the mean $\mu = 50$ and the standard deviation $\sigma = 40$

**16.1.**(1pt) Calculate $\mathrm{P}(\mathrm{X} < 62.8)$. Show your work, not just the final number.

**16.2.**(4pts) What is equal to $P(-30.4 < X < 58.4)$? **Choose 4 correct answers** among the followings:

a. $P(X < 58.4) + P(X < -30.4)$
b. $P(Z < 0.21) - P(Z < -2.01)$
c. $P(X < 58.4) - P(X > -30.4)$
d. $0.5832 + 0.0222$
e. $0.5832 - 0.0222$
f. $P(X > 58.4) - P(X < -30.4)$
g. $P(X < 58.4) - P(X < -30.4)$
h. $P(X > -30.4) - P(X > 58.4)$
i. $P(X < 58.4) + P(X > -30.4)$

**16.3.**(3pts) What is $P(X > 58.4 \quad or \quad X < -26)$? **Choose 3 correct answers** among the followings:

a. $P(X > 58.4) + P(X < -26)$
b. $P(Z < 0.21) + P(Z < -1.90)$
c. $P(Z > 0.21) + P(Z < -1.90)$
d. $0.5832 + 0.0287$
e. $0.5832 + 0.9713$
f. $0.4168 + 0.0287$
g. $0.5832 - 0.0287$

**16.4.**(3pts) We want to find the value V such that the probability that X is greater than V is 0.025. **Which three** of the following statements are correct?

a. $P(X < V) = 0.025$
b. $P(X > V) = 0.025$
c. $P\left(Z < \frac{V-\mu}{\sigma}\right) = 0.975$
d. $P\left(Z > \frac{V-\mu}{\sigma}\right) = 0.975$
e. $P\left(Z < \frac{X-\mu}{\sigma}\right) = 0.975$
f. $\frac{X-\mu}{\sigma} = 1.96$
g. $\frac{V-\mu}{\sigma} = 1.96$

**16.5.**(3pts) What is the probability $P(X > -10.4 \ \ or \ \ X < -2.4)$? **Choose 3 correct answers** among the followings:

a. $P(X > -10.4) + P(X < -2.4)$
b. $P(X < -10.4) + P(X < -2.4)$
c. $P(Z > -1.51 \ \ or \ \ Z < -1.31)$
d. $0.0655 + 0.0951$
e. $1 - 0.0655 + 0.0951$
f. $P(Z < -1.51) + P(Z < -1.31)$
g. $P(X < 10.4 \quad or \quad X > 2.4)$
h. 1

## PART II: OPTIONAL PART. YOU CAN CHOOSE TO ANSWER HOW MANY AND WHICH SUB-QUESTIONS. FOR EXAMPLE, YOU MAY ONLY CHOOSE SUB-QUESTION 17.1 AND 18.2

**Question 17.** (3 pts) Bank of America has conducted a survey on how frequently consumers use credit cards in making a purchase. The results of the study provide the following information.

37% of consumers use a credit card when making a purchase.

**Among** those consumers who use a credit card, 60% are female.

55% of people interviewed for the survey are male.

Let F = event that a consumer is female.

Let U (use a credit card) = event that a consumer uses a credit card in making a purchase.

17.1. Determine the value of $P(F)$? (1 pt)

17.2. Determine the value of $P(F|U)$? (1 pt)

17.3. What is the probability that you see your friend, Susan, uses a credit card? You, of course, know that Susan is female. (1 point)

**Question 18. (6 pts).** We have a discrete random variable X.

The possible values that X can take are - 1, 0, 1, 2. The probability function of X is given as follows:

$$f(x) = \begin{cases} \frac{x^2}{10} \; for \; x = -1 \\ 0.3 \; for \; x = 0, 1, 2 \end{cases}$$

18.1. (1 pt) What is the probability that X is **greater than or equal to 1?**

18.2. (1 pt) What is the expected value of X?

18.3. (1 pt) What is the variance of X?

Suppose that we construct a new random variable $Y = 2 + 3X$ based on the above random variable *X*. (X follows the distribution on the previous page. We are still in question 18.)

18.4. (1 pt) What is the **probability** that Y is less than or equal to 5?

18.5. (1 pt) What is the mean of Y?

18.6. (1 pt) What is the variance of Y?

**Question 19. (4 pts)**

Suppose X follows the binomial distribution with p = 0.3 and the number of trials n = 10.

19.1. (1 pt) What is the probability that X = 2? Choose one answer among the followings:

a. $\binom{10}{2} 0.3^2(1-0.3)^8$

b. $\binom{10}{2} 0.3^8(1-0.3)^2$

c. $0.3^2(1-0.3)^8$

d. $0.3^8(1-0.3)^2$

19.2. (1 pt) What is the probability that $X \geq 2$? Choose one answer among the followings:

a. $\binom{10}{1} 0.3^1(1-0.3)^9 + \binom{10}{2} 0.3^2(1-0.3)^8$

b. $1 - \binom{10}{0} 0.3^0(1-0.3)^{10} - \binom{10}{1} 0.3^1(1-0.3)^9$

c. $1 - \binom{10}{1} 0.3^1(1-0.3)^9 - \binom{10}{2} 0.3^2(1-0.3)^8$

d. $\binom{10}{0} 0.3^0(1-0.3)^{10} + \binom{10}{1} 0.3^1(1-0.3)^9$

19.3. (1 pt) What is the expected value of X?

19.4. (1 pt) What is the variance of X?

**Question 20. (3 pts)** Suppose X follows a Poisson distribution with $\mu = 4$

**20.1.** (1pt) What is the probability that X = 2? Choose one correct answer.

a. $\frac{4^2 e^{-4}}{2!}$

b. $\frac{2^4 e^{-2}}{4!}$

c. $\frac{2^4 e^{-4}}{2!}$

**20.2.** (1pt) What is the probability that $X \geq 2$?

a. $\frac{4^1 e^{-4}}{1!} + \frac{4^2 e^{-4}}{2!}$

b. $\frac{4^0 e^{-4}}{0!} + \frac{4^1 e^{-4}}{1!}$

c. $1 - \frac{4^0 e^{-4}}{0!} - \frac{4^1 e^{-4}}{1!}$

d. $1 - \frac{4^1 e^{-4}}{1!} - \frac{4^2 e^{-4}}{2!}$

**20.3.** (1 pt) What is the probability that $X \leq 1$?

a. $1 - 4e^{-4}$

b. $4e^{-4}$

c. $1 - 5e^{-4}$

d. $5e^{-4}$

**Question 21.** (4 pts)

Consider the continuous random variable X, which has the uniform distribution over the interval from 20 to 28.

21.1. (1pt) Write down the probability density function f(x)

21.2. (1pt) Draw the probability density function f(x)

21.3. (1pt) What is the probability that X will take on a value between 21 and 25?

21.4. (1pt) What is the expected value/mean of X?

**Question 22.** (2 pts) On average, 7 cars arrive at the drive up window of a bank every hour. Define the random variable X to be the number of cars arriving in any hour.

22.1. (1pt) What is the probability that exactly 3 cars arrive in the next hour?

22.2. (1pt) Let Y be the number of cars arriving at the window from 11am to 11.30am. What is the distribution of Y? (Give the name of the distribution, and the probability density function f(y) with clearly stating the expected value E(Y) = ? )

**Question 23. (1 pt)** Suppose that a random variable X has the uniform distribution with the support from 10 to H, where 10 < H

$$f(x) = \begin{cases} 0.25 & for \;\; 10 \leq x \leq H \\ 0 \;\; , & otherwise \end{cases}$$

What is the appropriate value for H? Explain how you found it.

## PART III: EXTRA CREDITS. YOU GET EXTRA CREDITS IF YOUR ANSWER IS CORRECT.

**Question 24.** (2 pts) The Bureau of Labor Statistics reported that the average yearly income of dentists in the year 2009 was $110,000. A sample of 100 dentists, which was taken in 2010, showed an average yearly income of $120,000. Assume the standard deviation of the population of dentists in 2010 is $36,000. We want to test to determine if there has been a significant increase in the average yearly income of dentists between the year 2009 and 2010.

24.1.(1pt) Provide the null and the alternative hypotheses.

24.2.(1pt) Compute the test statistic. Give the name of the sampling distribution we are going to use to test the hypotheses in this case. (You don't need to conduct the complete test. I just need a name of the distribution we're going to use if we do testing.)

**Question 25.** (1pt) The age of students at a university are normally distributed with a mean of 21. What percentage of the students in the university is at least 21 years old? Choose one answer among the followings:

a. It could be any value, depending on the magnitude of the standard deviation

b. 50%

c. 21%

d. 42%

e. $\sigma$ %

**Question 26.** (2pts)

26.1. (1pt) Toss a coin, roll a die, and toss a coin again. We are interested in the number that appears in an outcome (for the coin, assume Head = 1, Tail = 0). Is this experiment a binomial experiment? Explain.

26.2. (1pt) Roll a fair die 6 times. We are interested in the number that appears in an outcome. Is this experiment a binomial experiment? Explain.

**Question 27.** (1pt) 40% of the parts produced by a machine are defective. We select 6 parts randomly. Let X be the number of parts that are not defective. Write down the probability function f(x) of X.

**Question 28. (1 pt)** There is a mathematical operation called Integration. To denote this operation, the symbol $\int$ is used and it is called the Integral. What this operation does is that it gives some area under a function. For example, if we calculate $\int_a^b f(x)dx$, it merely gives us the area under $f(x)$ and between $a$ and $b$. Knowing this, calculate the following expression:

$$\int_{-\infty}^{50} \frac{1}{\sqrt{50\pi}} e^{\frac{-(x-40)^2}{50}} dx$$