# Maths & Stats Pre-Sessional Tutorial

SOLUTIONS to exercise for topic 2: Probability Distributions

---

**Exercise 1**

(a) $P(A) = P(r > 10\%) = P(d \cup e) = 0.33 + 0.21 = 0.54$

(b) $P(B) = P(r < 0\%) = P(a \cup b) = 0.04 + 0.14 = 0.18$

(c) $A$ complement is the event that the rate of return is not more than 10%.

(d) $P(\hat{A}) = P(a) + P(b) + P(c) = 0.04 + 0.14 + 0.28 = 0.46$.

(e) The intersection between more than 10% and return will be negative is the null or empty set.

(f) $P(A \cap B) = 0$.

(g) The union of $A$ and $B$ is the event that the rates of return are either more than 10% or less than 0%.

(h) $P(A \cup B) = P(A) + P(B) - P(A \cap B) = P(A) + P(B) = 0.54 + 0.18 = 0.72$.

(i) $A$ and $B$ are mutually exclusive because their intersection is the null set.

(j) $A$ and $B$ are not collectively exhaustive because their union does not equal 1.

**Exercise 2**

(a) $P(A \cap B) = P(A) + P(B) - P(A \cup B) = 0.80 + 0.10 - 0.82 = 0.08$

(b)

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{0.08}{0.10} = 0.80$$

(c) $A$ and $B$ are independent since $P(A|B) = P(A)$.

**Exercise 3**

a) P(X = 1, Y = 1) = 0.30

b)

| | | X | | P(y) |
|---|---|---|---|---|
| | | 1 | 2 | |
| Y | 0 | 0.20 | 0.25 | 0.45 |
| | 1 | 0.30 | 0.25 | 0.55 |
| P(x) | | 0.50 | 0.50 | 1 |

c)

Covariance and correlation are given by:

$$Cov(X,Y) = \sum_x \sum_y xyP(x,y) - \mu_x \mu_y$$

and

$$\rho = Corr(X,Y) = \frac{Cov(X,Y)}{\sigma_x \sigma_y}.$$

Hence, start by computing the mean of $X$ and $Y$ as follows:

$$\mu_x = 1 \cdot 0.5 + 2 \cdot 0.5 = 1.5$$

$$\mu_y = 0 \cdot 0.45 + 1 \cdot 0.55 = 0.55$$

Then, compute the first term of the covariance formula using the table above:

$$\sum_x \sum_y xyP(x,y) = 1 \cdot 0 \cdot 0.2 + 1 \cdot 1 \cdot 0.3 + 2 \cdot 0 \cdot 0.25 + 2 \cdot 1 \cdot 0.25 = 0.80.$$

and obtain the covariance as follows:

$$Cov(X,Y) = 0.80 - 1.5 \cdot 0.55 = -0.025$$

Now, compute the variance of $X$ and $Y$:

$$\sigma_x^2 = E(X^2) - \mu_x^2 = 1^2 \cdot 0.5 + 2^2 \cdot 0.5 - (1.5)^2 = 0.5 + 2 - 2.25 = 0.25$$

$$\sigma_y^2 = E(Y^2) - \mu_y^2 = 0^2 \cdot 0.45 + 1^2 \cdot 0.55 - (0.5)^2 = 0.55 - 0.3025 = 0.2475$$

and obtain the correlation between $X$ and $Y$:

$$\rho = Corr(X,Y) = \frac{Cov(X,Y)}{\sigma_x \sigma_y} = \frac{-0.025}{\sqrt{0.25}\sqrt{0.2475}} = -0.1005.$$

**Exercise 4**

$P = 6000 - 3X$. Using the properties of expected values and variances, we can write:

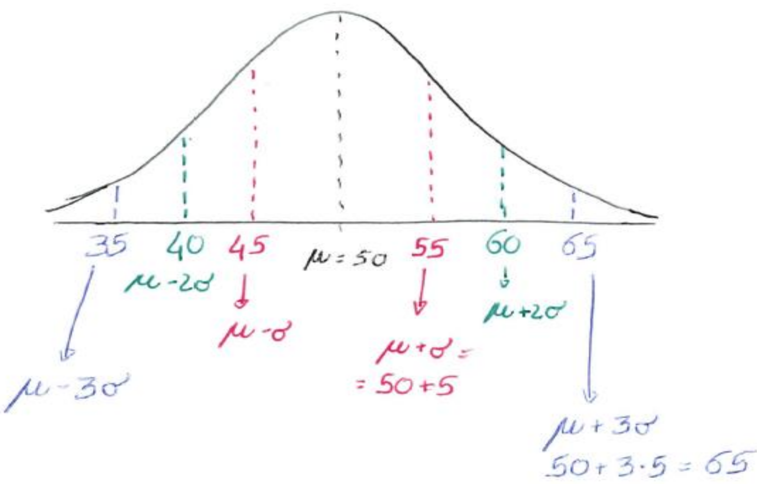
$$\mu_P = 6000 - 3\mu_X = 6000 - 3 \cdot 1000 = 3000.$$

and:

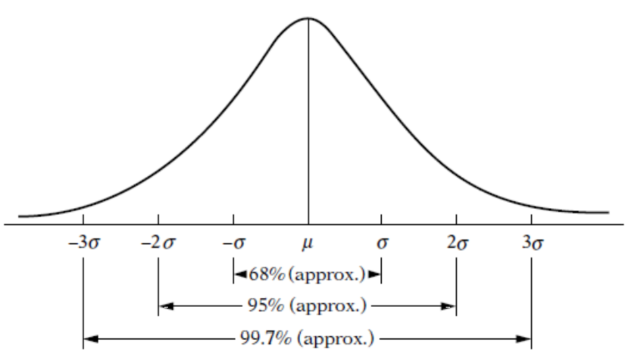
$$\sigma_P^2 = (-3)^2 \cdot \sigma_X^2 = 9 \cdot 900 = 8100.$$

**Exercise 5**

a)




b) The empirical rule, or three-sigma rule, states that about 68% of the area under the normal curve lies between $\mu \pm \sigma$, 95% of the area between $\mu \pm 2\sigma$, and about 99.7% of the area between $\mu \pm 3\sigma$. Therefore, the probability the random variable will assume a value between 45 and 55 is 68%




c) Similar to point b). The probability the random variable will assume a value between 40 and 60 is 95%

## Check your knowledge:

Test your knowledge with the following multiple-choice questions.

For each question, select the correct answer. Explain your decision.

### Question 1

The qualities of discrete data can be:

a) Measured
**b) Counted**
c) Both
d) None

**Explanation:** Discrete data refers to data that can only take on specific values and cannot be measured on a continuous scale. It can only be counted or enumerated. This means that the values of discrete data can be expressed as whole numbers or integers, such as the number of students in a class or the number of cars in a parking lot. Therefore, the correct answer is "Counted."

### Question 2

The qualities of continuous data can be:

**a) Measured**
b) Counted
c) Both
d) None

**Explanation:** Continuous data refers to data that can take on any value within a specific range. It is not limited to specific values or categories. The term "measured" implies that continuous data can be quantitatively measured, such as temperature, height, or weight. On the other hand, counting is more suitable for discrete data, where values can only be whole numbers or specific categories. Therefore, the correct answer is "Measured" because continuous data can be measured rather than counted.

### Question 3

Which of these is NOT continuous data?

a) An item's weight
b) The time spent every day on Instagram
**c) The number of students in a classroom**
d) None of these

## Question 4

What is the difference between a discrete random variable and a continuous random variable?

a) A discrete random variable takes only negative numbers while a continuous random variable takes both positive and negative numbers.
b) A discrete random variable takes both positive and negative numbers while a continuous random takes only negative numbers.
c) A discrete random variable takes all values in an interval of numbers while a continuous random variable has a fixed set of possible values with gaps between.
**d) A discrete random variable has a fixed set of possible values with gaps between while a continuous random variable takes all values in an interval of numbers.**

## Question 5

A marketing survey compiled data on the number of cars in households. If X = the number of cars, and we omit the rare cases of more than 5 cars, then X has the following probability distribution:

| $x$ | 0 | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|---|
| $P(X=x)$ | 0.24 | 0.37 | 0.20 | 0.11 | 0.05 | 0.03 |

What is the probability that a randomly chosen household has at least two cars?

(a) 0.20
(b) 0.29
**(c) 0.39**
(d) 0.81

**Explanation:** P(X >= 2) = P(X=2)+P(X=3)+P(X=4)+ P(X=5) = 0.2+0.11+0.05+0.03 = 0.39

## Question 6

A stockbroker estimates that at the end of the year, there is a 40% chance a stock will be worth $50, a 35% chance it will be worth $60 and a 25% chance it will be worth $70.

What expected value does this broker assign to this stock's end-of-the-year price?

**(a) $58.50**
(b) $60.00
(c) $62.50
(d) $65.00

**Explanation:** E(X) = 50*P(X=50) + 60*P(X=60) +70*P(X=70) = 50*0.4 + 60*0.35 +70*0.25 = 58.5

## Question 7

Find the missing value for P(X=5) in the probability distribution:

Number of Toys | Probability

**0| 0.03**

**1| 0.16**

**2| 0.30**

**3| 0.23**

**4| 0.17**

**5| ?**

a) **0.11**
b) 0.21
c) 0.01
d) 0.31

**Explanation:** the sum of probabilities is equal to 1. Therefore P(X=5) = 1-0.03-0.16-0.30-0.23-0.17=0.11

## Question 8

What is the probability that a student does play a sport given they do not play an instrument?

Middle School Music and Sports Survey

| | Plays Team Sport | Does Not Play Team Sport | Total |
|---|---|---|---|
| Plays Instrument | 8 | 3 | 11 |
| Does Not Play Instrument | 2 | 7 | 9 |
| Total | 10 | 10 | 20 |

a) 0.20
b) 0.2222
c) **0.10**
d) 0.50

**Explanation:** 2/20=0.1

### Question 9

What is the probability that a student plays an instrument?

| Middle School Music and Sports Survey | | | |
|---|---|---|---|
| | Plays Team Sport | Does Not Play Team Sport | Total |
| Plays Instrument | 8 | 3 | 11 |
| Does Not Play Instrument | 2 | 7 | 9 |
| Total | 10 | 10 | 20 |

- **a) 0.55**
- b) 0.50
- c) 0.45
- d) 0.40

**Explanation:** 11/20=0.55

### Question 10

Which of these values below is a marginal frequency?

| Concession Stand Sales | | | | |
|---|---|---|---|---|
| | Soda | Water | No Drink | Total |
| Hot Dog | **50** | 62 | 46 | **158** |
| Pizza | **120** | 58 | **4** | **182** |
| No Food | 30 | 20 | **10** | 60 |
| Total | **200** | **140** | 60 | **400** |

- **a) 182**
- b) 120
- c) 46
- d) 10

**Explanation:** Marginal Frequency is the sums of the rows and columns. The Marginal Relative Frequency isthe sum of the joint relative frequencies in a row or column.

## Question 11

Which of the following mentioned standard Probability density functions is applicable to discrete Random Variables?

a) Gaussian Distribution (Normal Distribution)
**b) Poisson Distribution**
c) t-student Distribution
d) Exponential Distribution

## Question 12

What is the area under a conditional Cumulative density function?

a) 0
b) Infinity
**c) 1**
d) Changes with CDF

**Explanation:** Area under any conditional CDF is 1.

## Question 13

When do the conditional density functions get converted into the marginally density functions?

a) Only if random variables exhibit statistical dependency
**b) Only if random variables exhibit statistical independency**
c) Only if random variables exhibit deviation from its mean value
d) If random variables do not exhibit deviation from its mean value

## Question 14

Mutually Exclusive events ___________

a) Contain all sample points.
b) Contain all common sample points.
c) Does not contain any sample point.
**d) Does not contain any common sample point.**

**Explanation:** Events are said to be mutually exclusive if they do not have any common sample point.

**Question 15**

What would be the probability of an event 'G' if H denotes its complement, according to the axioms of probability?

a) P (G) = 1 / P (H)
**b) P (G) = 1 – P (H)**
c) P (G) = 1 + P (H)
d) P (G) = P (H)

**Explanation:** According to the given statement P(G) + P(H) = 1.

**Question 16**

A variable that can assume any value between two given points is called ___________

**a) Continuous random variable**
b) Discrete random variable
c) Irregular random variable
d) Uncertain random variable

**Question 17**

If a variable can certain integer values between two given points is called ___________

a) Continuous random variable
**b) Discrete random variable**
c) Irregular random variable
d) Uncertain random variable

**Question 18**

The expected value of a discrete random variable 'x' is given by ___________

a) P(x)
b) ∑ P(x)
**c) ∑ x P(x)**
d) 1

**Explanation:** Expected value refers to mean which is given by http://mathurl.com/zqymzn7 in case of discrete probability distribution.

## Question 19

If 'X' is a continuous random variable, then the expected value is given by __________

a) P(X)
b) ∑ x P(x)
**c) ∫ X P(X)**
d) No value such as expected value

## Question 20

Out of the following values, which one is not possible in probability?

a) P(x) = 1
b) ∑ x P(x) = 3
c) P(x) = 0.5
**d) P(x) = – 0.5**

**Explanation:** In probability P(x) is always greater than or equal to zero.

## Question 21

If E(x) = 2 and E(z) = 4, then E(z – x) =?

**a) 2**
b) 6
c) 0
d) We cannot calculate it due to Insufficient data

**Explanation:** E(z – x) = E(z) – E(x) = 4 – 2 = 2.

## Question 22

The expected value of a random variable is its __________

**a) Mean**
b) Standard Deviation
c) Mean Deviation
d) Variance

**Explanation:** Expected value of a random variable is its mean.

### Question 23

The covariance of two independent random variable is ___________

a) 1
**b) 0**
c) -1
d) Undefined

**Explanation:** Two random variables are said to be independent if their covariance is zero

### Question 24

If P(x) = 0.5 and x = 4, then E(x) = ?

a) 1
b) 0.5
c) 4
**d) 2**

**Explanation:** E(x) = x P(x) = 0.5 * 4 = 2.

### Question 25

If the values taken by a random variable are negative, the negative values will have ___________

**a) Positive probability**
b) Negative Probability
c) May have negative or positive probabilities
d) Insufficient data

**Explanation:** Probabilities are always positive and not greater than 1.

### Question 26

How is called a variable that assigns a real number value to an event in a sample space ?

**Answer:** Random variable

### Question 27

Binomial Distribution is a ___________

a) Continuous distribution
**b) Discrete distribution**
c) Irregular distribution
d) Not a Probability distribution

### Question 28

In a Binomial Distribution, if ‘n’ is the number of trials and ‘p’ is the probability of success, then the mean value is given by ___________

**a) np**
b) n
c) p
d) np(1-p)

**Explanation:** For a discrete probability function, the mean value or the expected value is given by Mean $(\mu)=\sum xp(x)$. For Binomial Distribution we substitute in above equation and solve to get $\mu = np$.

### Question 29

The shape of the Normal Curve is ___________

**a) Bell Shaped**
b) Flat
c) Circular
d) Spiked

**Explanation:** Due to the nature of the Probability Mass function, a bell shaped curve is obtained.

### Question 30

Normal Distribution is symmetric around the ___________

a) Variance
**b) Mean**
c) Standard deviation
d) Covariance

**Explanation:** Due to the very nature of Normal Distribution, the graph appears such that it is symmetric about its mean.

### Question 31

The standard normal curve is symmetric about the value ___________

a) 0.5
b) 1
c) ∞
**d) 0**

**Explanation:** Normal curve is always symmetric about mean, for standard normal curve or variate mean = 0.

## Question 32

For a standard normal distribute variable, the value of Standard Deviation is ___________

a) 0
**b) 1**
c) ∞
d) not defined

**Explanation:** If the mean and standard deviation of a normal distributed variable are 0 and 1 respectively, it is called as standard normal variable.

## Question 33

The variance of a random variable X ,Var(X), is defined by ___________

**a)** $\boldsymbol{Var(X) = E(X^2) - (E(X))^2}$
b) $Var(X) = E(X) - E(X)$
c) $Var(X) = E(X^2) - E(X)$
d) $Var(X) = E(X^2) - E(X^2)$

**Explanation:** $Var((X - m)^2)$

= $E((X^2) - 2Xm + m^2)$

= $E(X^2) - 2mE(X) + m^2$

= $E(X^2) - 2mm + m^2$

= $E(X^2) - m^2)$

= $E(X^2) - (E(X))^2$

## Question 34

For a random variable X, $Var(aX + b) = a^2Var(X)$, is true or false?

**a) True**
b) False

### Question 35

The variances of two independent random variables X and Y are 0.2 and 0.5 respectively.

Let Z= 5X-2Y. The variance of Z is?

a) 3
b) 4
c) 5
**d) 7**

**Explanation:**

Let Z = 5X – 2Y. Then, Var(Z) = Var(5X-2Y)= Var(5X) + Var(2Y)= 25Var(X) + 4Var(Y) = 7.

### Question 36

A continuous random variable X has uncountable many values in the interval [a, b]. If C is a values in the interval [a, b], then P{ X = C }

a) Is zero
b) Is strictly non-zero
c) depends on the limit [a,b]
d) is less than one, but non-zero

**Explanation:** the probability density function (PDF) of a continuous random variable describes the relative likelihood of the random variable taking on different values, and for any specific value, the area under the PDF curve at that value is zero. Therefore, P{ X = C } = 0 for any C in the interval [a, b].