Ans : A, A, B, B

1. The end points of a confidence interval are called:

* **Confidence limits**
* Confidence coefficient
* Error of estimation
* Parameters

2. Estimation can be classified into

* **Point estimation and interval estimation**
* Biased and Unbiased
* Type-I and Type-II
* One sided and sided testing

3. A range (set) of values within which the population parameter is expected to occur is called:

* Level of significance
* **Confidence interval**
* Confidence coefficient
* Confidence limits

4. If θ^ is the estimator of the parameter θ, then θ^ is called unbiased if:

* E(θ^)≠0
* **E(θ^)>θ**
* E(θ^)>0
* E(θ^)<0

5. The probability associated with confidence interval is called

* Confidence coefficient
* Degrees of freedom
* Level of confidence
* **Confidence limits**

6. ‘Statistic’ is an estimator and its computer values is called:

* Interval estimate
* Estimation
* Estimator
* **Estimate**

7. The process of using sample data to estimate the values of unknown population parameters is called

* Estimator
* Interval estimate
* **Estimation**
* Estimate

8. The process of making estimates about the population parameter from a sample is called

* Statistical decision
* Statistical hypothesis
* **Statistical inference**
* Statistical independence

9. A single value used to estimate a population value is called:

* **Point estimate**
* Interval estimate
* Level of significance
* Confidence limits

10. The numerical value which we determine from the sample for a population parameter is called

* Confidence Coefficient
* **Estimator**
* Estimator
* Estimate

11. The estimator is said to be \_\_\_\_\_\_\_\_\_\_\_ if the mean of the estimator is not equal to the mean of the population parameter.

* Negatively biased
* Positively biased
* **Biased**
* Unbiased

12. A set (range) of values calculated from the sample data and it is likely to contain the true value of the parameter with some probability is called:

* Level of confidence
* Confidence limits
* Point estimate
* **Interval estimate**

13. The estimate is the observed value of an:

* Interval estimation
* Estimation
* **Estimator**
* Unbiased estimator

14. There are two main branches of statistical inference, namely

* Point estimate and interval estimate
* **Estimation of parameter and testing of hypothesis**
* Biased estimator and unbiased estimator
* Level of significance and degree of freedom

15. A formula or rule used for estimating the parameter of interest is called:

* Interval Estimate
* Estimation
* Estimate
* **Estimator**

Top of Form

Your answers are shown below:

1. The set of all possible outcomes (sample points) is called

* a sample
* an event
* **the sample space**
* a population

2. If P(A|B)=0.3

* **P(B|A)=0.7**
* P(Ac|B)=0.7
* P(A|Bc)=0.7
* P(Ac|Bc)=0.7

3. If events A and B are independent events with P(A)=0.1 and P(B)=0.4, then

* P(A∩B)=0
* **P(A∩B)=0.04**
* P(A∪B)=0.5
* P(A∩B)=0.25

4. The range of probability is

* any value greater than zero
* any value less than one
* **zero to one**
* any value between −1 to 1

5. If events A and B are independent events with P(A)=0.2 and P(B)=0.6, then P(A∪B)=?

* 0.62
* 0.12
* 0.60
* **0.68**

P(A∪B)=P(A)+P(B)–P(A∩B)

6. If P(A)=0.6, P(B)=0.3, and P(A∩B)=0.2, then P(B|A)=?

* 0.33
* 0.5
* 0.67
* **0.9**

7. If P(A)=0.5 and P(B)=0.5, then P(AcapB) is

* 0
* 0.25
* **1**
* Cannot be determined from the given information

8. One a December day, the probability of snow is 0.30. The probability of a “cold” day is 0.50. The probability of snow and a “cold” is 0.15. Do snow and “cold” weather are independent events?

* Only if given that it snowed
* **no**
* yes
* Only when events are mutually exclusive

9. If P(A|B)=0.3 and P(B)=0.8, then

* P(A)=0.24
* P(B|A)=0.7
* P(A∪B)=0.5
* P(A∩B)=0.24

Question was not answered

10. The probability of an intersection of two events is computed by using the

* Law of Addition
* **Law of Multiplication**
* Law of Subtraction
* Law of Division

11. The Complement of P(A|B) is

* P(Ac|B)
* P(A|Bc)
* P(Ac|Bc)
* **P(B|A)**
* P(A∩B)

12. If A and B are independent events with P(A)=0.6 and P(B)=0.6, then P(A∩B)=?

* 0.76
* 1
* **0.36**
* 0.2

13. Events A and B are mutually exclusive. Which of the following statements is true?

* P(A∪B)=P(A)P(B)
* A and B are independent
* **P(A∪B)=P(A)+P(B)**
* P(A∩B)=P(A)+P(B)

14. An experiment consists of tossing 4 coins successively. The number of sample points in this experiment is

* **16**
* 8
* 4
* 2

15. A six-sided die is tossed three times. The probability of observing three ones in a row is

* 1/3
* 1/6
* 1/27
* **1/216**

16. The sample space (experimental outcomes) refers to

* any particular experimental outcome
* The sample size minus one
* **the set of all possible experimental outcomes**
* both (a) and (c)

17. If A and B are independent events with P(A)=0.4 and P(B)=0.25, then P(A∪B=?

* **0.65**
* 0.55
* 0.1
* given information is not enough

18. In a statistical experiment, each time the experiment is repeated

* the same outcome must occur
* the same outcome cannot occur again
* **a different outcome may occur**
* None of the above

19. If two events A and B are mutually exclusive events, then

* P(A)+P(B)=0
* P(A)+P(B)=1
* P(A∪B)=0
* **$P(A \cap B)=0**

20. If events A and B are independent events with P(A)=0.05 and P(B)=0.65, then P(A|B)=?

* 0.05
* **0.0325**
* 0.65
* 0.8

Random Var

. A variable whose value is determined by the outcome of a random experiment is called

* Random
* Random variable
* Constant
* None of these

Question was not answered

2. If X and Y are random variables then E(X+Y) is equal to

* E(X)+E(Y)
* E(X)–E(Y)
* E(X)+Y
* None of these

Question was not answered

3. Random numbers can be generated mechanically by

* Use of random numbers table
* Use of digital computers
* Ordinary calculators
* None of these

Question was not answered

4. The sum of probabilities of a discrete random variable is

* Zero
* Four
* Three
* One

Question was not answered

5. The number of students in a class is an example of

* Continuous variable
* Discrete variable
* Definite variable
* None of these

Question was not answered

6. Height measurements of 50 students studying in a college

* Discrete variable
* Random variable
* Constant
* Continuous variable

Question was not answered

7. A variable which can assume each and every value within a given range is called

* Discrete variable
* Random variable
* Qualitative variable
* Continuous variable

Question was not answered

8. A random variable is also called

* Constant
* Variable
* Attribute
* Chance variable

Question was not answered

9. A quantity which can vary from one individual to another is called

* Constant
* Variable
* Data
* None of these

Question was not answered

10. The lifetime of a car tire is

* Continuous variable
* Discrete variable
* Random variable
* Qualitative variable

Question was not answered

11. A random variable assuming an infinite number of values is called

* Absolute variable
* Discrete random variable
* Continuous random variable
* None of these

Question was not answered

12. If x is a discrete random variable, the function f(x) is

* Distribution function
* Probability function
* Density function
* None of these

Question was not answered

13. A random variable assuming only a finite number of values is called:

* Discrete random variable
* Continuous random variable
* Random variable
* None of these

Question was not answered

14. Suppose, four coins are tossed, the value of a random variable H (No. of heads) is:

* 1, 2, 3, 4
* 0, 1, 2, 3, 4
* 0, 1, 2, 3
* 0, 1

Question was not answered

15. The speed of the car is an example of

* Continuous variable
* Discrete variable
* Absolute variable
* None of these

Bottom of Form

#### Probability Distributions

Which of the following probability distributions is/are NOT valid, complete probability distribution(s)?

A -

|  |  |
| --- | --- |
| X | Probability |
| 0 | 0.1 |
| 2 | 0.2 |
| 3 | 0.3 |
| 4 | 0.4 |

B -

|  |  |
| --- | --- |
| X | Probability |
| 0 | 0.15 |
| 0.5 | 0.15 |
| 1 | 0.20 |
| 1.5 | 0.25 |
| 2 | 0.25 |

C -

|  |  |
| --- | --- |
| X | Probability |
| 0 | -0.05 |
| 1 | 0.20 |
| 2 | 0.25 |
| 3 | 0.35 |
| 4 | 0.40 |
|  |  |

#### Expected Value of Player's Winnings

Suppose we change the game’s rules to the following:

|  |  |
| --- | --- |
| **Outcome** | **Prize** |
| 4 red balls | +150 |
| 4 blue balls | -150 |
| Any other outcome | -10 |

What will be the expected value now for X (the amount of money won by a player after playing the game once)?

Recall the problem you saw earlier, where we were asked by a company to suggest whether it should invest in a given project or not. We had made this probability distribution for X, the net revenue of the project.

|  |  |
| --- | --- |
| X (Net Revenue of Project, in ₹ crores) | P(x) |
| -305 | 0.1 |
| +15 | 0.7 |
| +95 | 0.2 |

Now, we are in a position to find the expected value for X, the return of the project. This is called the **expected return.** If it comes out to be negative, we can say that the project is not worth investing in.

#### Expected Return

Find the expected value of X if its probability distribution is as given in the table above. (Answer in terms of crores.)

Let's say that **Rajya Laxmi Bank** has given **student loans** to 10,000 people. However, the government wants to ensure that the bank is not giving away very risky loans and, hence, wants to know the **“Expected Loss”** of the bank’s student loan portfolio.

The expected loss is basically the expected value of the money lost by the bank due to people defaulting on their loans, i.e. not paying their EMIs.

The data that the bank uses to calculate the expected loss, looks like this -

|  |  |  |  |
| --- | --- | --- | --- |
| Customer No. | Exposure at Default (in ₹ lakh) | Recovery (%) | Probability of Default |
| 00001 | 11.50 | 20% | 0.007 |

Here,

* Exposure at default(**EAD**) is the total money owed by the customer in case of default
* Recovery (**R**) is the percentage of the exposed money that the bank would be able to recover.
  + For example, in the above example, the bank would recover 20% of the exposed money, i.e. 20%\*11.5 = ₹2.3 lakh.
* Probability of default (**PD**) is the probability that the customer will default. This is calculated for each customer using a number of factors such as family income, university attended, etc.

#### Expected Loss

Now, to find the expected loss, the bank has created the following table. As you already know, the probability of customer 1 being a defaulter is 0.007 and that of the customer not being a defaulter is 0.993. Now, in the case of default, what would be the value of X, i.e. the money lost by the bank?

|  |  |  |
| --- | --- | --- |
| Defaulter/Non-defaulter | Probability | X (Money Lost by Bank) |
| Defaulter | 0.007 | - |
| Non-defaulter | 0.993 | 0 |

#### Expected Loss

Now, the bank has the following data -

|  |  |  |
| --- | --- | --- |
| Defaulter/Non-defaulter | Probability | X (Money Lost by Bank) |
| Defaulter | 0.007 | ₹9.2 lakh |
| Non-defaulter | 0.993 | 0 |

What would be the expected loss (expected value of the money lost by the bank)?

#### Expected Loss

So, in general, what would be the expected loss?

#### Expected Loss

Now, the bank wants to find the total expected loss. This is done by calculating the expected losses for each of the 10,000 customers and then adding them up.

The data for all the customers of the bank is given in the file Inferential Statistics - Student Loan.csv. What would be the value of the total expected loss?

#### Probability Without Experiment

What is the probability that you would get this combination of balls after 4 trials? (One trial = taking out a ball, noting its colour, and putting it back in the bag.)

Note that the bag contains 2 blue and 3 red balls.



i.e., the combination Blue-Red-Red-Red.

[You’re not expected to know the right answer at this point. This question is given just to get you thinking in the right direction.]

#### Probability Without Experiment

What is the probability that you will get 3 red balls after 4 trials? (One trial = taking out a ball, noting its colour, and putting it back in the bag.)

[You’re not expected to know the right answer at this point. This question is given just to get you thinking in the right direction.]

#### Probability Without Experiment

What is the probability that you would get the combination of balls shown below after 4 trials? (One trial = taking out a ball, noting its colour, and putting it back in the bag).



Note that we have changed the contents of the bag — now it has 4 red balls and 1 blue ball.

#### Probability Without Experiment

What is the probability that you would get exactly 3 red balls after 4 trials? (One trial = taking out a ball, noting its colour, and putting it back in the bag.)

#### Binomial Distribution

A student attempts a test with 10 multiple choice questions, where each question has 4 possible options. She does not know the correct answer to any of the questions and just randomly selects one of the 4 options for each of them. The probability that she would still get 3 questions correct is:

#### Binomial Distribution

Now, the test is modified by adding 5 true and false questions to the set of 10 MCQs. The student does not know the answer to the new questions either and randomly selects true or false for each of them too.

Now, according to binomial distribution, the probability that she would still get 5 of the 15 questions (5 true/false Qs, and 10 MCQs with 4 options) correct is:

#### Cumulative Probability

Let’s define X as the number of wickets Ishant Sharma would take in the next T20 match he plays. Also, the following is an incomplete table for cumulative probability based on previous experience:

|  |  |
| --- | --- |
| **x(Number of wickets taken by Ishant Sharma in a T20 Match)** | **F(x)** |
| 0 | 0.35 |
| 1 | 0.55 |
| 2 | 0.75 |

What is the probability that he would take more than 2 wickets in the next  T20 match he plays?

Let’s say you get a job as a **senior analyst** at a food regulatory body such as the **FSSAI**. One of your tasks is to check if the **amount of lead content** in certain food products is **within the permissible limit**.

Recently, the company Kwick Foods filed a complaint against Sunshine, its competitor, that Sunshine’s ready-to-cook pasta contains excess quantities of lead.

Suppose you take **10 random pasta packets** from the market and get it tested for the amount of lead. It just so happens that you have picked out packets from a very defective batch, and there is a **5% probability** that any pasta packet you select is going to be defective.

#### Discrete Probability Distributions

What is the probability that 2 of the 10 packets tested would turn out to be defective?

(Use this nCr [calculator](http://www.calculatorsoup.com/calculators/discretemathematics/combinations.php).)

#### Discrete Probability Distributions

What is the probability that, after testing these 10 packets, not more than 2 packets would turn out to be defective?

(Use this nCr [calculator](http://www.calculatorsoup.com/calculators/discretemathematics/combinations.php).)

#### Discrete Probability Distributions

Let’s define X as the number of packets found to be defective after the 10 packets have been tested. What will be the expected value of X?

(Hint: You can Use Excel for the calculations for this part. Use "COMBIN(n,r)" in Excel )

Suppose a new **cancer treatment** has been discovered, claiming to increase the **one year survival rate** for pancreatic cancer to **40%.** In other words, the probability that a patient suffering from pancreatic cancer would survive for at least one year after receiving this treatment is 40%.

Suppose a  hospital is planning to use this treatment for its pancreatic cancer patients.

#### Pancreatic Cancer Hospital

The hospital has a total of 10 patients suffering from pancreatic cancer. What is the probability that exactly 4 of these patients would survive the first year after receiving this treatment?

(Use this nCr [calculator](http://www.calculatorsoup.com/calculators/discretemathematics/combinations.php).)

#### Pancreatic Cancer Hospital

What is the probability that the number of patients that survive the first year after receiving the treatment would not be more than 2?

(Use this nCr [calculator](http://www.calculatorsoup.com/calculators/discretemathematics/combinations.php).)

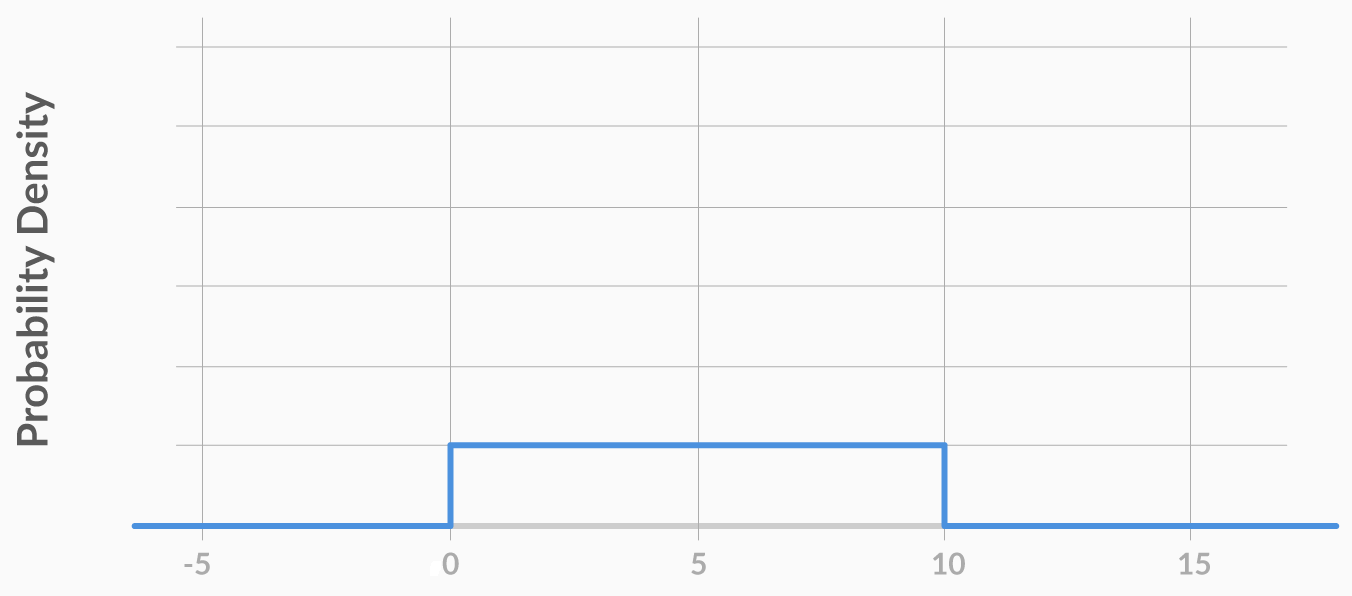
#### Cumulative Probability of Continuous Distributions

Recall that, in the above video, we talked about a fictional company with 3,000 employees in a local office. Now, let’s say you went to that office, and asked those 3,000 employees to play our UpGrad red ball game (from the previous sessions 1 and 2). Based on the data on these 3,000 people, if you create the probability distribution for X, then the values of P(X=0), P(X=1), P(X=2), P(X=3) and P(X=4) would all be:

[X = number of red balls drawn by a player after playing the game once]

#### Uniform Distribution

In a uniform PDF, all the possible values have the same probability density. The figure below shows such a uniform PDF, where the possible values are 0 to 10.



For this graph, what is the value of the probability density from X = 0 to X = 10?

#### Cumulative Probability of Continuous Variables

Suppose you work at a sports analysis company and you want to analyse the effect a bowler’s height has on his/her performance. So, you create a list of all 5 wicket hauls in the last decade. Based on this data, they created a cumulative probability distribution for X, where X = height of the bowler who took the 5 wicket haul.

Now, based on the data, you conclude that the cumulative probability, F(175.3 cm) = 0.3. In this case, which of the following statements is correct?

1. P(X<175.3 cm) = 0.3
2. P(X<175.3 cm) = 0.3

(Remember that height is a continuous variable.)

#### Pobability of Normal Random Variables

Let’s say that you need to find the cumulative probability for a random variable X which is normally distributed. You do not know what the value of X is or, for that matter, what the value of µ and σ is. You only know that X = µ + σ. Can you find the cumulative probability, i.e. the probability of the variable being less than µ + σ?

#### Normal Variables

What is the probability of a normally distributed random variable lying within 1.65 standard deviations of the mean?

[You can use the Z table [here](https://www.math.arizona.edu/~rsims/ma464/standardnormaltable.pdf).]

Let’s say you **work as an analyst** at a **pharma company** which manufactures an antipyretic drug (tablet form) with **paracetamol** as the active ingredient. The amount of paracetamol specified by the drug regulatory authorities is **500 mg** with a **permissible error** of **10%**. Anything below 450 mg would be a quality issue for your company since the drug will be ineffective, while above 550 mg would be a serious regulatory issue.

#### Cumulative Probability Distributions

The regulatory authority selects a random tablet from Batch Z2. Based on previous knowledge, you know that Batch Z2 has a mean paracetamol level of 510 mg, and its standard deviation is 20 mg.

What is the probability that the tablet that has been selected by the authority has a paracetamol level below 550 mg?

#### Continuous Probability Distributions

Now, the company’s QC (Quality Control) department comes and selects a tablet at random from Batch Z2. It is interested in finding if the paracetamol level is above 450 mg or not.

What is the probability that the tablet selected by QC has a paracetamol level above 450 mg?

#### Continuous Probability Distributions

Now, let’s say that QC decides to sample one more tablet. This time, it selects a tablet from Batch Y4. Based on previous knowledge, you know that Batch Y4 has a mean paracetamol level of 505 mg, and its standard deviation is 25 mg. This time, QC wants to check both the upper limit and the lower limit for the paracetamol level.

What is the probability that the tablet selected by QC has a paracetamol level between 450 mg and 550 mg?

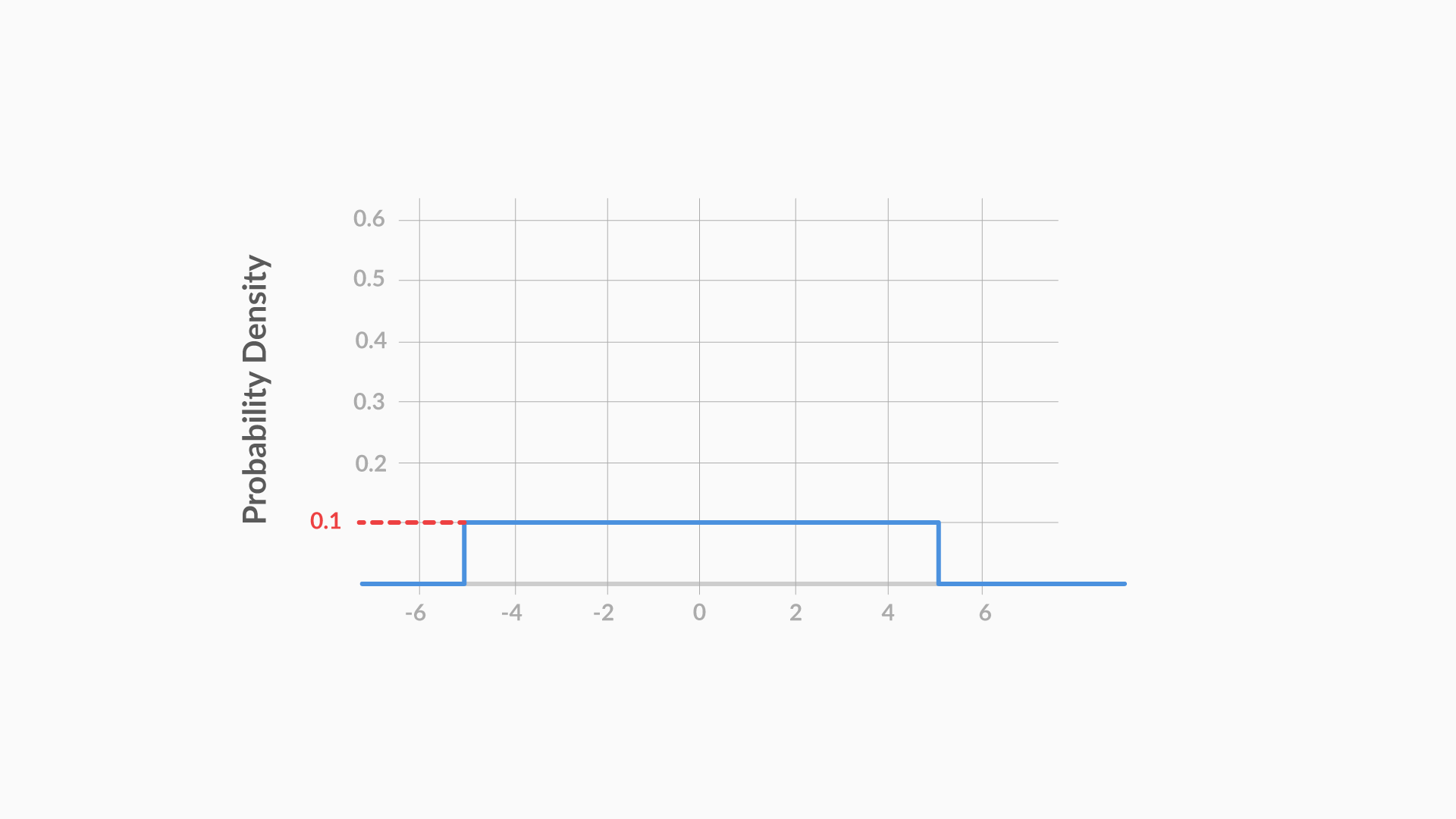


Figure 17 - Uniformly Distributed Variable (p = 0.1)

The graph you see above represents the PDF of a uniformly distributed random variable X. As you can see, the probability density is equal for all the possible values of X (-5 to +5).

**Question 1/1**

Mandatory

#### Uniform Distribution

What is the probability of the random variable X lying between -1.5 and +2.5, i.e. P(-1.5<X<2.5)?