Statistics for Data Science -1

Lecture 9.1: Discrete Random Variable: Application-credit card

Usha Mohan

Indian Institute of Technology Madras

Learning objectives

Learning objectives

1. Define what is a random variable.

Learning objectives

1. Define what is a random variable.
2. Types of random variables: discrete and continuous.

Learning objectives

1. Define what is a random variable.
2. Types of random variables: discrete and continuous.
3. Probability mass function, graph, and examples.

Learning objectives

1. Define what is a random variable.
2. Types of random variables: discrete and continuous.
3. Probability mass function, graph, and examples.
4. Cumulative distribution function, graphs, and examples.

Learning objectives

1. Define what is a random variable.
2. Types of random variables: discrete and continuous.
3. Probability mass function, graph, and examples.
4. Cumulative distribution function, graphs, and examples.
5. Expectation and variance of a random variable.

Application: Credit cards

Introduction

We are interested in analyzing the number of credit cards owned by people. Towards this, we collect data on a number of people and ask them how many credit cards they own. Their response is recorded.

Introduction

We are interested in analyzing the number of credit cards owned by people. Towards this, we collect data on a number of people and ask them how many credit cards they own. Their response is recorded.

▶ Random experiment: Consider the random experiment of selecting an adult at random from the sample.

Introduction

We are interested in analyzing the number of credit cards owned by people. Towards this, we collect data on a number of people and ask them how many credit cards they own. Their response is recorded.

▶ Random experiment: Consider the random experiment of selecting an adult at random from the sample.

▶ Random variable: The number of credit cards owned by the person.

Introduction

We are interested in analyzing the number of credit cards owned by people. Towards this, we collect data on a number of people and ask them how many credit cards they own. Their response is recorded.

▶ Random experiment: Consider the random experiment of selecting an adult at random from the sample.

▶ Random variable: The number of credit cards owned by the person.

    ▶ This is a discrete random variable.

Introduction

We are interested in analyzing the number of credit cards owned by people. Towards this, we collect data on a number of people and ask them how many credit cards they own. Their response is recorded.

- ▶ Random experiment: Consider the random experiment of selecting an adult at random from the sample.
- ▶ Random variable: The number of credit cards owned by the person.
  - ▶ This is a discrete random variable.
- ▶ The following table summarizes the probability distribution of the number of credit cards per person based on the relative frequencies.

Introduction

We are interested in analyzing the number of credit cards owned by people. Towards this, we collect data on a number of people and ask them how many credit cards they own. Their response is recorded.

▶ Random experiment: Consider the random experiment of selecting an adult at random from the sample.

▶ Random variable: The number of credit cards owned by the person.

    ▶ This is a discrete random variable.

▶ The following table summarizes the probability distribution of the number of credit cards per person based on the relative frequencies.

▶

X	0	1	2	3	4
$P(X = x_i)$	0.42	0.36	0.14	0.06	0.02

Probability mass function
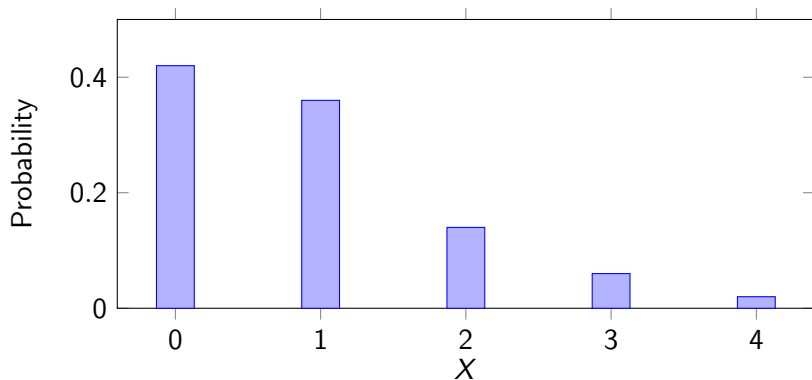
Questions

▶ Describe the distribution

Questions

▶ Describe the distribution
- ▶ The distribution is skewed right with a peak at 0.
- ▶ Number of credit cards owned by people vary between 0 to 4 credit cards.

Questions

- ▶ Describe the distribution
  - ▶ The distribution is skewed right with a peak at 0.
  - ▶ Number of credit cards owned by people vary between 0 to 4 credit cards.
- ▶ Choose an adult at random. Is he/she more likely to have 0 credit cards or 2 or more credit cards?

Questions

- ▶ Describe the distribution
  - ▶ The distribution is skewed right with a peak at 0.
  - ▶ Number of credit cards owned by people vary between 0 to 4 credit cards.
- ▶ Choose an adult at random. Is he/she more likely to have 0 credit cards or 2 or more credit cards?
  - ▶ The probability that an adult has no credit cards is 0.42, while the probability of having 2 or more credit cards is about 0.22, so the probability of having no credit cards is higher.

Questions-continued

- ▶ You take a random sample of 500 people and ask them how many credit cards they own. Would you be surprised at the following:

Questions-continued

- ▶ You take a random sample of 500 people and ask them how many credit cards they own. Would you be surprised at the following:
    - ▶ Everyone owns a credit card.

Questions-continued

- You take a random sample of 500 people and ask them how many credit cards they own. Would you be surprised at the following:
  - Everyone owns a credit card.
    YES: 42% of adults do not own credit cards. Hence, it is unlikely that every one of the 500 would own credit cards.

Questions-continued

- You take a random sample of 500 people and ask them how many credit cards they own. Would you be surprised at the following:
  - Everyone owns a credit card.
    YES: 42% of adults do not own credit cards. Hence, it is unlikely that every one of the 500 would own credit cards.
  - 72 people respond that they own two credit cards.
    NO: 14% own two credit cards. 14% of 500 = 70, so it is likely that 72 people from sample of 500 own two credit cards.

Questions-continued

- You take a random sample of 500 people and ask them how many credit cards they own. Would you be surprised at the following:
  - Everyone owns a credit card.
    YES: 42% of adults do not own credit cards. Hence, it is unlikely that every one of the 500 would own credit cards.
  - 72 people respond that they own two credit cards.
    NO: 14% own two credit cards. 14% of 500 = 70, so it is likely that 72 people from sample of 500 own two credit cards.
- Again choose an adult at random. How many credit cards would you "Expect" that person to own?

Section summary

- ▶ Probability mass function- tabular form and graph.
- ▶ Cumulative distribution function- definition and graph.
- ▶ Key ideas:
    - ▶ Shape of distribution: skewed, symmetric, constant, etc.
    - ▶ Answer questions about distribution of random variable.