MAST7866
Random variables and probability distributions

Jian Zhang

University of Kent

Randomness

- Randomness plays an important part in our lives. Most things happen randomly.
- For example, consider the following events: winning a lottery, your car breaking down, getting sick, getting involved in an accident, losing a job, making money in the stock market, etc.

- We cannot predict when, where and to whom these things can or will happen. All these events are uncertain, and they happen randomly to people.

Random Variables

- Similarly, consider the following events:
    - how many customers will visit a bank, a grocery store or a petrol station on a given day?
    - How many cars will pass a bridge on a given day?
    - How many students will be absent from class on a given day?

- In all these examples, the number of customers, cars or students are random; that is, each of these can assume any value within a certain interval.

- This table gives the frequency and relative frequency distributions of the number of vehicles owned by all 2000 families living in a small town.

Number of vehicles owned	Frequency	Relative Frequency
0	30	30/2000 = 0.015
1	320	320/2000 = 0.160
2	910	910/2000 = 0.455
3	580	580/2000 = 0.290
4	160	160/2000 = 0.080
	N = 2000	**Sum = 1.000**

- Suppose one family is randomly selected from this population. The process of randomly selecting a family is called a random experiment. Let x denote the number of vehicles owned by the selected family. Then x can assume any of the five possible values (0, 1, 2, 3 and 4). The value of x depends on which family is selected. Thus, this value depends on the outcome of a random experiment. Consequently, x is called a random variable.

Discrete Random Variables

- A **discrete random variable** assumes values that can be counted. The earlier example of the number of cars owned by a family is an example of a discrete random variable. Some other examples include:
  - The number of cars sold at a dealership during a given month
  - The number of houses in a certain block
  - The number of fish caught on a fishing trip
  - The number of complaints received at the office of an airline on a given day

Continuous Random Variables

- A random variable whose values are not countable is called a **continuous random variable**. A continuous random variable can assume any value over an interval or intervals. As the number of values contained in an interval is infinite, the possible number of values that a continuous random variable can assume is also infinite.

- Consider the life of a battery. We can measure it as precisely as we want. For instance, the life of a battery might be 40 hours or 40.25 hours or 40.247 hours. Assume that the maximum life of a battery is 200 hours. Let x denote the life of a randomly selected battery of this kind. Then x can assume any value in the interval 0 to 200.

Continuous examples

- A few other examples of **continuous random variables** include:

  - The length of a room
  - The time taken to commute from home to work
  - The weight of a letter
  - The price of a house

- Note that money is often treated as a continuous random variable, specifically when there are a large number of unique values.

Probability Distribution of a random variable

- Let x be a discrete random variable. The probability distribution of x describes how the probabilities are distributed over all the possible values of x.

- Return to the previous example about car ownership. Because the relative frequencies represent the population, they give the actual probabilities of outcomes.

Number of vehicles owned, x	Probability P(x)
0	0.015
1	0.160
2	0.455
3	0.290
4	0.080
Total	**1.000**

Properties of probability distributions

- The **probability distribution of a discrete random variable** possesses the following two characteristics:

- The probabilities lie between 0 and 1, i.e. $0 \leq P(x) \leq 1$ for each value of x.
- The probabilities sum to 1, i.e. $\sum P(x) = 1$

Binomial distribution

- The binomial probability distribution is one of the most widely used discrete probability distributions. It is applied to find the probability that an outcome will occur x times in n performances of an experiment.

- For example, if 75% of students at a University use Instagram, we may want to find the probability that in a random sample of five students at this University that exactly three use Instagram.

Conditions for Binomial

- A binomial experiment must satisfy the following conditions:
  - There are n identical trials.
  - Each trial has only two possible outcomes (success and failure).
  - The probabilities of the two outcomes remain constant.
  - The trials are independent.

Probability function for Binomial

For a binomial experiment, the probability of exactly x successes in n trials is given by the formula:

$$P(x) = \binom{n}{x} p^x (1-p)^{n-x}$$

where n denotes the number of trials and p denotes the probability of success.

Poisson Distribution

- The Poisson probability distribution, named after the French mathematician, Simeon-Denis Poisson, is another important probability distribution of a discrete random variable that has a large number of applications.
- Suppose a washing machine in a launderette breaks down an average of three times a month, we may want to find the probability of exactly two break downs in the next month. This is an example of a Poisson probability distribution problem. Each breakdown is called an occurrence.

Conditions for Poisson

- The following three conditions must have satisfied to apply the Poisson probability distribution:

  - x is a discrete random variable
  - The occurrences are random
  - The occurrences are independent

Probability function for Poisson

According to the **Poisson probability distribution**, the probability of x occurrences in an interval is

$$P(x) = \frac{\lambda^x e^{-\lambda}}{x!}$$

where λ is the mean number of occurrences in that interval.