Probability and Statistics

Topic 12 - Discrete Probability Distributions: Discrete Random Variables

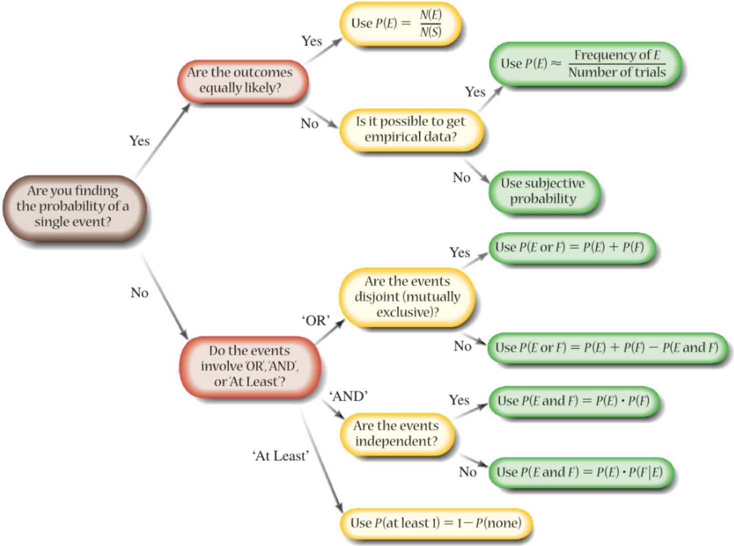
Aamir Alaud Din, PhD

November 14, 2024

TABLE OF CONTENTS

TABLE OF CONTENTS

RECAP



Are you finding the probability of a single event?
Yes
Are the outcomes equally likely?
Yes
Use $P(E) = \frac{N(E)}{N(S)}$
No
Is it possible to get empirical data?
Yes
Use $P(E) \approx \frac{\text{Frequency of } E}{\text{Number of trials}}$
No
Use subjective probability
No
Do the events involve 'OR', 'AND', or 'At Least'?
'OR'
Are the events disjoint (mutually exclusive)?
Yes
Use $P(E \text{ or } F) = P(E) + P(F)$
No
Use $P(E \text{ or } F) = P(E) + P(F) - P(E \text{ and } F)$
'AND'
Are the events independent?
Yes
Use $P(E \text{ and } F) = P(E) \cdot P(F)$
No
Use $P(E \text{ and } F) = P(E) \cdot P(F|E)$
'At Least'
Use $P(\text{at least } 1) = 1 - P(\text{none})$

RECAP

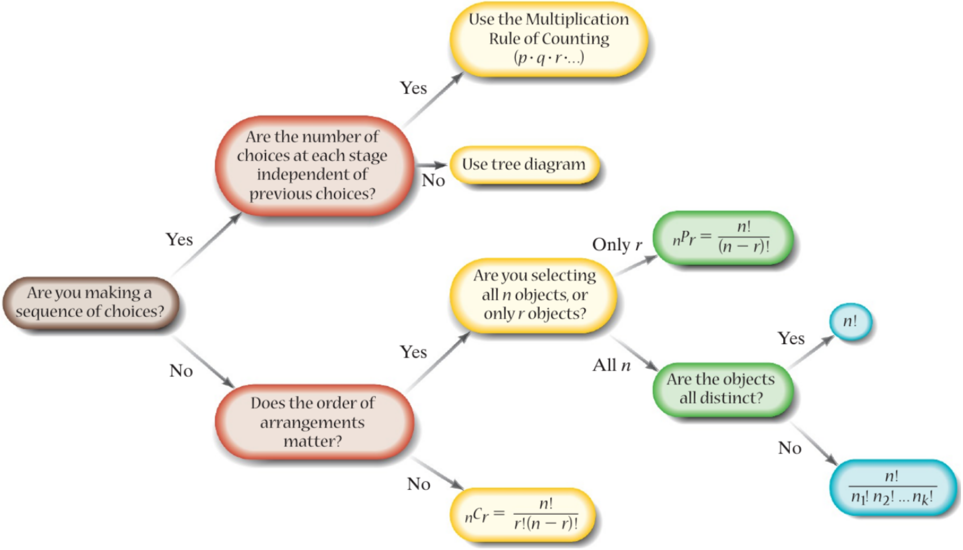

Are you making a sequence of choices?
Yes
Are the number of choices at each stage independent of previous choices?
Yes
Use the Multiplication Rule of Counting $(p \cdot q \cdot r \cdot \ldots)$
No
Use tree diagram
No
Does the order of arrangements matter?
Yes
Are you selecting all n objects, or only r objects?
Only r
${}_nP_r = \frac{n!}{(n-r)!}$
All n
Are the objects all distinct?
Yes
$n!$
No
$\frac{n!}{n_1!\, n_2! \ldots n_k!}$
No
${}_nC_r = \frac{n!}{r!(n-r)!}$

TABLE OF CONTENTS

OBJECTIVES

After learning this topic and studying, you should be able to:

1. Distinguish between discrete and continuous random variables
2. Identify discrete probability distributions
3. Graph discrete probability distributions
4. Compute and interpret the mean of a discrete random variable
5. Interpret the mean of a discrete random variable as an expected value
6. Compute the standard deviation of a discrete random variable

TABLE OF CONTENTS

- Based on our previous knowledge, we already know how to compute the mean and standard deviation of a data set.
- What if the data set is a probability model in which data points are discrete and the probability of every discrete value is different from the other as shown in Table 1.

Table 1

x	$P(x)$
0	0.01
1	0.10
2	0.38
3	0.51

- Our objective is to know how to compute the mean and standard deviation of probability models with discrete data set.

TABLE OF CONTENTS

DISCRETE AND CONTINUOUS RANDOM VARIABLES

- Consider a probability experiment in which we flip a coin two times.
- The outcomes of the experiment are $\{HH, HT, TH, TT\}$.
- Rather than being interested in a particular outcome, we might be interested in the number of heads.
- If the outcome of a probability experiment is a numerical result, we say the outcome is a random variable.

Random Variable

A random variable is a numerical measure of the outcome of a probability experiment, so its value is determined by chance. Random variables are typically denoted using capital letters such as X.

DISCRETE AND CONTINUOUS RANDOM VARIABLES

- So, in our coin-flipping example, if the random variable X represents the number of heads in two flips of a coin, the possible values of X are $x = 0, 1,$ or 2..
- Notice that we follow the practice of using a capital letter, such as X, to identify the random variable and a lowercase letter, x, to list the possible values of the random variable, or the sample space of the experiment.
- As another example, consider an experiment that measures the time between arrivals of cars at a drive-through.
- The random variable T describes the time between arrivals, so the sample space of the experiment is $t > 0$.
- There are two types of random variables, discrete and continuous.

DISCRETE AND CONTINUOUS RANDOM VARIABLES

Discrete Random Variable

A discrete random variable has either a finite or countable number of values. The values of a discrete random variable can be plotted on a number line with space between each point. See Figure (a).

Continuous Random Variable

A continuous random variable has infinitely many values. The values of a continuous random variable can be plotted on a line in an uninterrupted fashion. See Figure (b).

(a) Discrete Random Variable

(b) Continuous Random Variable

DISCRETE AND CONTINUOUS RANDOM VARIABLES

EXAMPLE 1

(a) The number of As earned in a section of statistics with 15 students enrolled is a discrete random variable because its value results from counting. If the random variable X represents the number of As, then the possible values of X are $x = 0, 1, 2, \ldots, 15$.

(b) The number of cars that travel through a McDonald's drive-through in the next hour is a discrete random variable because its value results from counting. If the random variable X represents the number of cars, the possible values of X are $x = 0, 1, 2, \ldots$.

(c) The speed of the next car that passes a state trooper is a continuous random variable because speed is measured. If the random variable S represents the speed, the possible values of S are all positive real numbers; that is, $s > 0$.

TABLE OF CONTENTS

DISCRETE PROBABILITY DISTRIBUTIONS

- Because the value of a random variable is determined by chance, we may assign probabilities to the possible values of the random variable.

Probability Distribution

The probability distribution of a discrete random variable X provides the possible values of the random variable and their corresponding probabilities. A probability distribution can be in the form of a table, graph, or mathematical formula.

Table 1

x	$P(x)$
0	0.01
1	0.10
2	0.38
3	0.51

EXAMPLE 2

Suppose we ask a basketball player to shoot three free throws. Let the random variable X represent the number of shots made, so $x = 0, 1, 2,$ or 3. Table 1 shows a probability distribution for the random variable X.

We denote probabilities using the notation $P(x)$, where x is a specific value of the random variable. We read $P(x)$ as "the probability that the random variable X equals x." For example, $P(3) = 0.51$ is read "the probability that the random variable X equals 3 is 0.51."

- Recall from Section 5.1 that probabilities must obey certain rules.
- Below are the rules for a discrete probability distribution using the notation just introduced.

Rules for a Discrete Probability Distribution

Let $P(x)$ denote the probability that a random variable X equals x; then

1. $\sum P(x) = 1$
2. $0 \leq P(x) \leq 1$

EXAMPLE 3

The figure on the next slide shows three data tables. Which of them is a discrete probability distribution?

(a)

x	$P(x)$
1	0.20
2	0.35
3	0.12
4	0.40
5	−0.07

(b)

x	$P(x)$
1	0.20
2	0.25
3	0.10
4	0.14
5	0.49

(c)

x	$P(x)$
1	0.20
2	0.25
3	0.10
4	0.14
5	0.31

TABLE OF CONTENTS

GRAPHING DISCRETE PROBABILITY DISTRIBUTIONS

- In the graph of a discrete probability distribution, the horizontal axis represents the values of the discrete random variable and the vertical axis represents the corresponding probability of the discrete random variable.
- When graphing a discrete probability distribution, we want to emphasize that the data are discrete.
- Therefore, the graph of discrete probability distributions is drawn using vertical lines above each value of the random variable to a height that is the probability of the random variable.
- Graphs of discrete probability distributions help determine the shape of the distribution.
- Recall that we describe distributions as skewed left, skewed right, or symmetric.
- The graph in Figure below is skewed left.

EXAMPLE 4

Graph the discrete probability distribution given in Table 1 from Example 2.

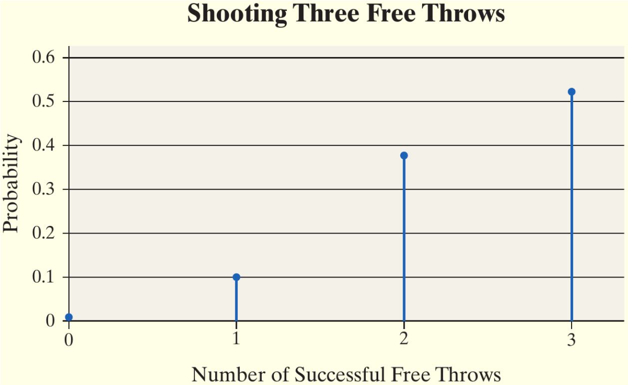


TABLE OF CONTENTS

MEAN OF A DISCRETE RANDOM VARIABLE

- Remember, when describing the distribution of a variable, we describe its center, spread, and shape.
- We will use the mean to describe the center and use the standard deviation to describe the spread.
- Let's see where the formula for computing the mean of a discrete random variable comes from.
- One semester I asked a small statistics class of 10 students to disclose the number of people living in their households.
- I obtained the following data:

$$2, 4, 6, 6, 4, 4, 2, 3, 5, 5$$

MEAN OF A DISCRETE RANDOM VARIABLE

Table 2

x	$P(x)$
2	$\frac{2}{10} = 0.2$
3	$\frac{1}{10} = 0.1$
4	$\frac{3}{10} = 0.3$
5	$\frac{2}{10} = 0.2$
6	$\frac{2}{10} = 0.2$

MEAN OF A DISCRETE RANDOM VARIABLE

- What is the mean number of people in the 10 households?
- We could find the mean by adding the observations and dividing by 10, but we will take a different approach.
- Letting the random variable X represent the number of people in the household, we obtain the probability distribution in Table 2.
- Now compute the mean as follows:

$$\mu = \frac{\sum x_i}{N} = \frac{2+4+6+6+4+4+2+3+5+5}{10}$$

MEAN OF A DISCRETE RANDOM VARIABLE

$$
\begin{aligned}
\mu = \frac{\sum x_i}{N} &= \frac{2+4+6+6+4+4+2+3+5+5}{10} \\
&= \frac{\overbrace{2+2}^{2}+\overbrace{3}^{1}+\overbrace{4+4+4}^{3}+\overbrace{5+5}^{2}+\overbrace{6+6}^{2}}{10} \\
&= \frac{2\cdot 2+3\cdot 1+4\cdot 3+5\cdot 2+6\cdot 2}{10} \\
&= 2\cdot\frac{2}{10}+3\cdot\frac{1}{10}+4\cdot\frac{3}{10}+5\cdot\frac{2}{10}+6\cdot\frac{2}{10} \\
&= 2\cdot P(2)+3\cdot P(3)+4\cdot P(4)+5\cdot P(5)+6\cdot P(6) \\
&= 2(0.2)+3(0.1)+4(0.3)+5(0.2)+6(0.2) \\
&= 4.1
\end{aligned}
$$

MEAN OF A DISCRETE RANDOM VARIABLE

- We conclude that the mean of a discrete random variable is found by multiplying each possible value of the random variable by its corresponding probability and then adding these products.

Mean of a Discrete Random Variable

The mean of a discrete random variable is given by the formula

$$\mu_X = \sum [x.P(x)]$$

where X is the value of the random variable and $P(x)$ is the probability of observing the value x.

EXAMPLE 5

Compute the mean of the discrete random variable given in Table 1 from Example 2.

MEAN OF A DISCRETE RANDOM VARIABLE

- We will follow the practice of rounding the mean and standard deviation to one more decimal place than the values of the random variable.

How to Interpret the Mean of a Discrete Random Variable

- The mean of a discrete random variable can be thought of as the mean outcome of the probability experiment if we repeated the experiment many times.
- If we repeated the experiment in Example 5 of shooting three free throws many times, we would expect the mean number of free throws made to be around 2.4.

Interpretation of the Mean of a Discrete Random Variable

Suppose an experiment is repeated n independent times and the value of the random variable X is recorded.

MEAN OF A DISCRETE RANDOM VARIABLE

Interpretation of the Mean of a Discrete Random Variable (Contiued ...)

As the number of repetitions of the experiment increases, the mean value of the n trials will approach μ_X, the mean of the random variable X. In other words, let x_1 be the value of the random variable X after the first experiment, x_2 be the value of the random variable X after the second experiment, and so on. Then

$$\bar{x} = \frac{x_1 + x_2 + \cdots + x_n}{n}$$

The difference between $\bar{x}$ and μ_X gets closer to 0 as n.

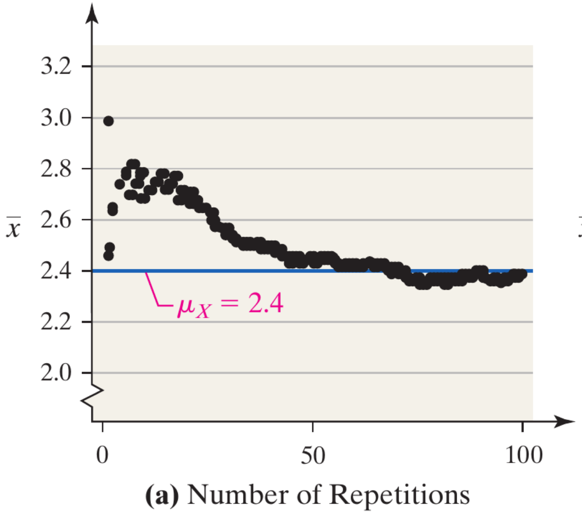
EXAMPLE 6

The basketball player from Example 2 is asked to shoot three free throws 100 times. Compute the mean number of free throws made.

MEAN OF A DISCRETE RANDOM VARIABLE

- Figures (a) and (b) further illustrate the mean of a discrete random variable.
- Figure (a) shows the mean number of free throws made versus the number of repetitions of the experiment for the data in Table 4.
- Figure (b) shows the same information when the experiment of shooting three free throws is conducted a second time for 100 repetitions.
- In both plots the player starts "hot," since the mean number of free throws made is above the theoretical level of 2.4.
- However, both graphs approach the theoretical mean of 2.4 as the number of repetitions of the experiment increases.

MEAN OF A DISCRETE RANDOM VARIABLE



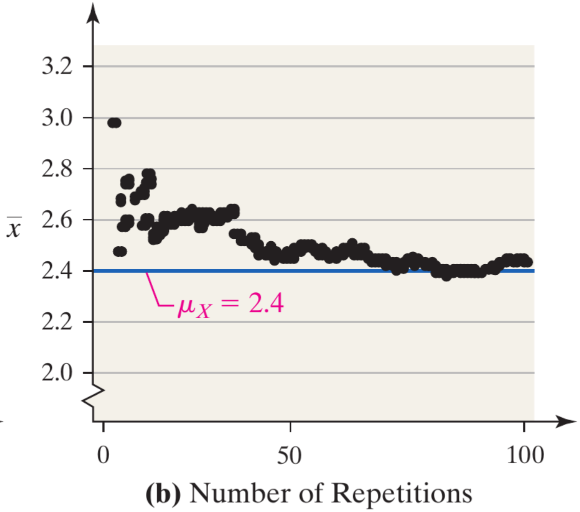


TABLE OF CONTENTS

MEAN OF A DISCRETE RANDOM VARIABLE AS AN EXPECTED VALUE

- Because the mean of a random variable represents what we would expect to happen in the long run, it is also called the expected value, $E(X)$.
- The interpretation of expected value is the same as the interpretation of the mean of a discrete random variable.

EXAMPLE 7

A term life insurance policy will pay a beneficiary a certain sum of money upon the death of the policyholder. These policies have premiums that must be paid annually. Suppose an 18-year-old male buys a $250,000 1-year term life insurance policy for $350. According to the National Vital Statistics Report, Vol. 58, No. 21, the probability that the male will survive the year is 0.998937. Compute the expected value of this policy to the insurance company.

TABLE OF CONTENTS

STANDARD DEVIATION OF A DISCRETE RANDOM VARIABLE

- We now introduce a method for computing the standard deviation of a discrete random variable.

Standard Deviation of a Discrete Random Variable

The standard deviation of a discrete random variable X is given by

$$\sigma_X = \sqrt{\sum[(x - \mu_X)^2.P(x)]} = \sqrt{\sum[x^2.P(x)] - \mu_X^2}$$

EXAMPLE 8

Find the standard deviation of the discrete random variable given in Table 1 from Example 2.

STANDARD DEVIATION OF A DISCRETE RANDOM VARIABLE

- The variance of the discrete random variable is the value under the square root in the computation of the standard deviation.
- The variance of the discrete random variable in Example 8 is

$$\sigma_X^2 = 0.4979 \approx 0.5$$

EXAMPLE 9

Use statistical software or a graphing calculator to find the mean and the standard deviation of the random variable whose distribution is given in Table 1.

TABLE OF CONTENTS

SUMMARY

- The distribution of a discrete random variable can't be continuous but lines at discrete values.
- The mean of a discrete random variable is given by the formula

$$\mu_X = \sum [x.P(x)]$$

- The standard deviation of a discrete random variable is given by the equation

$$\begin{aligned}\sigma_X &= \sqrt{\sum [(x - \mu_X)^2.P(x)]} \\ &= \sqrt{\sum [x^2.P(x)] - \mu_X^2}\end{aligned}$$

- Mean of discrete random variable is also called the expected value as it tells the expected value of the random variable that would happen in the long run.


Thank You!