Probability 2

Random variables, probability distributions, binomial random variable

Example 1 : Consider the experiment of flipping a fair coin three times. The number of tails that appear is noted as a discrete random variable :

X= number of tails that appear in 3 flips of a fair coin.

There are 8 possible outcomes of the experiment : namely the sample space consists of

$$\Omega=\{HHH, HHT, HTH, HTT, THH, TTT, THT, TTH\} \quad \text{where}$$
$$X= \quad 0, \quad 1, \quad 1, \quad 2, \quad 1, \quad 3, \quad 2, \quad 2$$

are the corresponding values taken by the random variable X.

Definition : A **random variable** is a function from outcomes $\omega\in\Omega$ to numbers*. For the above example with $\omega=HTT$, $X(\omega)=X(HTT)=2$ counts the number of tails in the three coin flips. A **discrete** random variable is one which only takes a *finite* or *countable* set of values as opposed to a **continuous** random variable which can take say any real number value in an interval of real numbers. (There are *uncountably* many real numbers in an interval of positive length.)

a) What are the possible values that X takes on and what are the probabilities of X taking a particular value? From the above we see that the possible values of X are the 4 values

$$X = 0, 1, 2, 3.$$

Said differently the sample space is a disjoint union of the 4 events $\{X=j\}$ for short ($j=0,1,2,3$)where **what is meant by the event {X=j}** is :

$\{\omega : X(\omega)=j\}$ $j=0,1,2,3$ Specifically in our example :

$\{X=0\}=\{HHH\}$,
$\{X=1\}=\{THH, HTH, HHT\}$,
$\{X=2\}=\{TTH, HTTT, THT\}$,
$\{X=3\}=\{TTT\}$.

Since for a fair coin we assume that each element of the sample space is equally likely (with probability 1/8) we find that the probabilities for the various values of X, called the **probability distribution** $f(x)=f_X(x)=P(X=x)$ of X or for X discrete, also commonly called the **probability mass function** (or **pmf**) $f(x)=p(x)=p_X(x)=P(X=x)$ can be summarized in the following table listing the possible values beside the probability of that value :

x	0	1	2	3
$p(x)$	1/8	3/8	3/8	1/8

We can say that this pmf places mass 3/8 on the value X=2 .
The numbers (=”masses” or probabilities) for a pmf should be between 0 and 1.
The total mass (i.e. total probability) must add up to 1.

b) What is the expected value of the random variable X ?

this is the sum of the values times their probabilities :

$$E[X] = 0\cdot(1/8)+1\cdot(3/8)+2\cdot(3/8)+3\cdot(1/8) = 12/8 = 3/2 \quad .$$

This is not surprising since for a fair coin on average half the time we expect to see tails, and the answer is half of the n=3 flips = np where p=1/2 is the probability of a tail occurring.

Formula for the expected value of a discrete r.v. : Given a pmf for a discrete random variable X, its expected value is given by the formula :

$$E[X] = \sum_x x\cdot p_X(x)$$ where the sum is over all possible values of the random variable.

Permutations and Combinations

Consider the following problem: In how many ways can 8 horses finish in a race (assuming there are no ties)? We can look at this problem as a decision consisting of 8 steps. The first step is the possibility of a horse to finish first in the race, the second step the horse finishes second, ... , the 8th step the horse finishes 8th in the race. Thus, by the Fundamental Principle of counting there are

$$8 \cdot 7 \cdot 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 40,320 \text{ ways}$$

This problem exhibits an example of an ordered arrangement, that is, the order the objects are arranged is important. Such ordered arrangement is called a **permutation**. Products such as $8 \cdot 7 \cdot 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1$ can be written in a shorthand notation called factoriel. That is, $8 \cdot 7 \cdot 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 8!$ (read "8 factoriel"). In general, we define **n factoriel** by

$$n! = \begin{cases} n(n-1)(n-2)\cdots 3 \cdot 2 \cdot 1, & \text{if } n \geq 1 \\ 1, & \text{if } n = 0. \end{cases}$$

where n is a whole number n.

Example

Evaluate the following expressions:

(a) $6!$ (b) $\frac{10!}{7!}$.

Solution.

(a) $6! = 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 720$

(b) $\frac{10!}{7!} = \frac{10 \cdot 9 \cdot 8 \cdot 7 \cdot 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1}{7 \cdot 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1} = 10 \cdot 9 \cdot 8 = 720 \blacksquare$

Using factoriels we see that the number of permutations of n objects is $n!$.

Example

There are 6! permutations of the 6 letters of the word "square." In how many of them is r the second letter?

Solution.
Let r be the second letter. Then there are 5 ways to fill the first spot, 4 ways to fill the third, 3 to fill the fourth, and so on. There are 5! such permutations.■

Example

Five different books are on a shelf. In how many different ways could you arrange them?

Solution.
The five books can be arranged in $5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 5! = 120$ ways.

Counting Permutations

We next consider the permutations of a set of objects taken from a larger set. Suppose we have n items. How many ordered arrangements of r items can we form from these n items? The number of permutations is denoted by $P(n, r)$. The n refers to the number of different items and the r refers to the number of them appearing in each arrangement. This is equivalent to finding how many different ordered arrangements of people we can get on r chairs if we have n people to choose from. We proceed as follows.

The first chair can be filled by any of the n people; the second by any of the remaining $(n-1)$ people and so on. The rth chair can be filled by $(n-r+1)$ people. Hence we easily see that

$$P(n,r) = n(n-1)(n-2)...(n-r+1) = \frac{n!}{(n-r)!}.$$

Example

How many ways can gold, silver, and bronze medals be awarded for a race run by 8 people?

Solution.
Using the permuation formula we find $P(8,3) = \frac{8!}{(8-3)!} = 336$ ways.

Example

How many five-digit zip codes can be made where all digits are unique? The possible digits are the numbers 0 through 9.

Solution.
$P(10,5) = \frac{10!}{(10-5)!} = 30,240$ zip codes

Combinations

Let $C(n,r)$ denote the number of ways in which r objects can be selected from a set of n distinct objects. Since the number of groups of r elements out of n elements is $C(n,r)$ and each group can be arranged in $r!$ ways then $P(n,r) = r!C(n,r)$. It follows that

$$C(n,r) = \frac{P(n,r)}{r!} = \frac{n!}{r!(n-r)!}.$$

Example

How many ways are there to select a committee to develop a discrete mathematics course at a school if the committee is to consist of 3 faculty members from the mathematics department and 4 from the computer science department, if there are 9 faculty members of the math department and 11 of the CS department?

Solution.
There are $C(9,3) \cdot C(11,4) = \frac{9!}{3!(9-3)!} \cdot \frac{11!}{4!(11-4)!} = 27,720$ ways.

Finding Probabilities Using Combinations and Permutations

Combinations can be used in finding probabilities as illustrated in the next example.

Example

Given a class of 12 girls and 10 boys.

(a) In how many ways can a committee of five consisting of 3 girls and 2 boys be chosen?
(b) What is the probability that a committee of five, chosen at random from the class, consists of three girls and two boys?
(c) How many of the possible committees of five have no boys?(i.e. consists only of girls)
(d) What is the probability that a committee of five, chosen at random from the class, consists only of girls?

Solution.

(a) First note that the order of the children in the committee does not matter. From 12 girls we can choose $C(12,3)$ different groups of three girls. From the 10 boys we can choose $C(10,2)$ different groups. Thus, by the Fundamental Principle of Counting the total number of committee is

$$C(12,3)\cdot C(10,2)=\frac{12!}{3!9!}\cdot\frac{10!}{2!8!}=\frac{12\cdot 11\cdot 10}{3\cdot 2\cdot 1}\cdot\frac{10\cdot 9}{2\cdot 1}=9900$$

(b) The total number of committees of 5 is $C(22,5)=26,334$. Using part (a), we find the probability that a committee of five will consist of 3 girls and 2 boys to be

$$\frac{C(12,3)\cdot C(10,2)}{C(22,5)}=\frac{9900}{26,334}\approx 0.3759.$$

(c) The number of ways to choose 5 girls from the 12 girls in the class is

$$C(10,0)\cdot C(12,5)=C(12,5)=\frac{12\cdot 11\cdot 10\cdot 9\cdot 8}{5\cdot 4\cdot 3\cdot 2\cdot 1}=792$$

(d) The probability that a committee of five consists only of girls is

$$\frac{C(12,5)}{C(22,5)} = \frac{792}{26,334} \approx 0.03 \blacksquare$$

Exercise:

The license plates in the state of Utah consist of three letters followed by three single-digit numbers.

(a) If Edward's initials are EAM, what is the probability that his license plate will have his initials on it (in any order)?

An example of the Binomial distribution

Suppose, we have an unfair coin for which the probability of getting a head is $\frac{2}{3}$ and the probability of a tail is $\frac{1}{3}$. Consider tossing the coin five times in a row and counting the number of times we observe a head. We can denote this number as

$$X = \text{No. of heads in 5 coin tosses}$$

X can take on any of the values 0, 1, 2, 3, 4 and 5.

X is a **discrete random variable**

Some values of X will be more likely to occur than others. Each value of X will have a probability of occuring. What are these probabilities? Lets consider the probability of obtaining just one head in 5 coin tosses, i.e. $X = 1$.

One possible way of obtaining one head is if we observe the pattern HTTTT. The probability of obtaining this pattern is

$$P(HTTTT) = \frac{2}{3} \times \frac{1}{3} \times \frac{1}{3} \times \frac{1}{3} \times \frac{1}{3}$$

There are 32 possible patterns of heads and tails we might observe. 5 of the patterns contain just one head

HHHHH	THHHH	HTHHH	HHTHH	HHHTH	HHHHT	TTHHH	THTHH
THHTH	THHHT	HTTHH	HTHTH	HTHHT	HHTTH	HHTHT	HHHTT
TTTHH	TTHTH	TTHHT	THTTH	THTHT	THHTT	HTTTH	HTTHT
HTHTT	HHTTT	$\boxed{\text{HTTTT}}$	$\boxed{\text{THTTT}}$	$\boxed{\text{TTHTT}}$	$\boxed{\text{TTTHT}}$	$\boxed{\text{TTTTH}}$	TTTTT

The other 5 possible combinations all have the same probability so the probability of obtaining one head in 5 coin tosses is

$$P(X = 1) = 5 \times \left(\frac{2}{3} \times \left(\frac{1}{3}\right)^4\right) = 0.0412 \text{ (to 4dp)}$$

What about P(X = 2)? This probability can be written as

$$\begin{aligned} P(X=2) = &\quad \text{No. of patterns} &\times&\quad \text{Probability of pattern} \\ = &\quad {}^5C_2 &\times&\quad \left(\frac{2}{3}\right)^2 \times \left(\frac{1}{3}\right)^3 \\ = &\quad 10 &\times&\quad \frac{4}{243} \\ = &\quad 0.165 \end{aligned}$$

It's now just a small step to write down a formula for this situation specific situation in which we toss a coin 5 times

$$P(X = x) \quad = \quad {}^5C_x \times \left(\frac{2}{3}\right)^x \times \left(\frac{1}{3}\right)^{(5-x)}$$

We can use this formula to tabulate the probabilities of each possible value of X.

$$\begin{aligned}
P(X = 0) &= {}^5C_0 \times \left(\tfrac{2}{3}\right)^0 \times \left(\tfrac{1}{3}\right)^5 &= 0.0041 \\
P(X = 1) &= {}^5C_1 \times \left(\tfrac{2}{3}\right)^1 \times \left(\tfrac{1}{3}\right)^4 &= 0.0412 \\
P(X = 2) &= {}^5C_2 \times \left(\tfrac{2}{3}\right)^2 \times \left(\tfrac{1}{3}\right)^3 &= 0.1646 \\
P(X = 3) &= {}^5C_3 \times \left(\tfrac{2}{3}\right)^3 \times \left(\tfrac{1}{3}\right)^2 &= 0.3292 \\
P(X = 4) &= {}^5C_4 \times \left(\tfrac{2}{3}\right)^4 \times \left(\tfrac{1}{3}\right)^1 &= 0.3292 \\
P(X = 5) &= {}^5C_5 \times \left(\tfrac{2}{3}\right)^5 \times \left(\tfrac{1}{3}\right)^0 &= 0.1317
\end{aligned}$$

These probabilities are plotted in Figure 1 against the values of X. This shows the **distribution** of probabilities across the possible values of X. This situation is a specific example of a Binomial distribution.

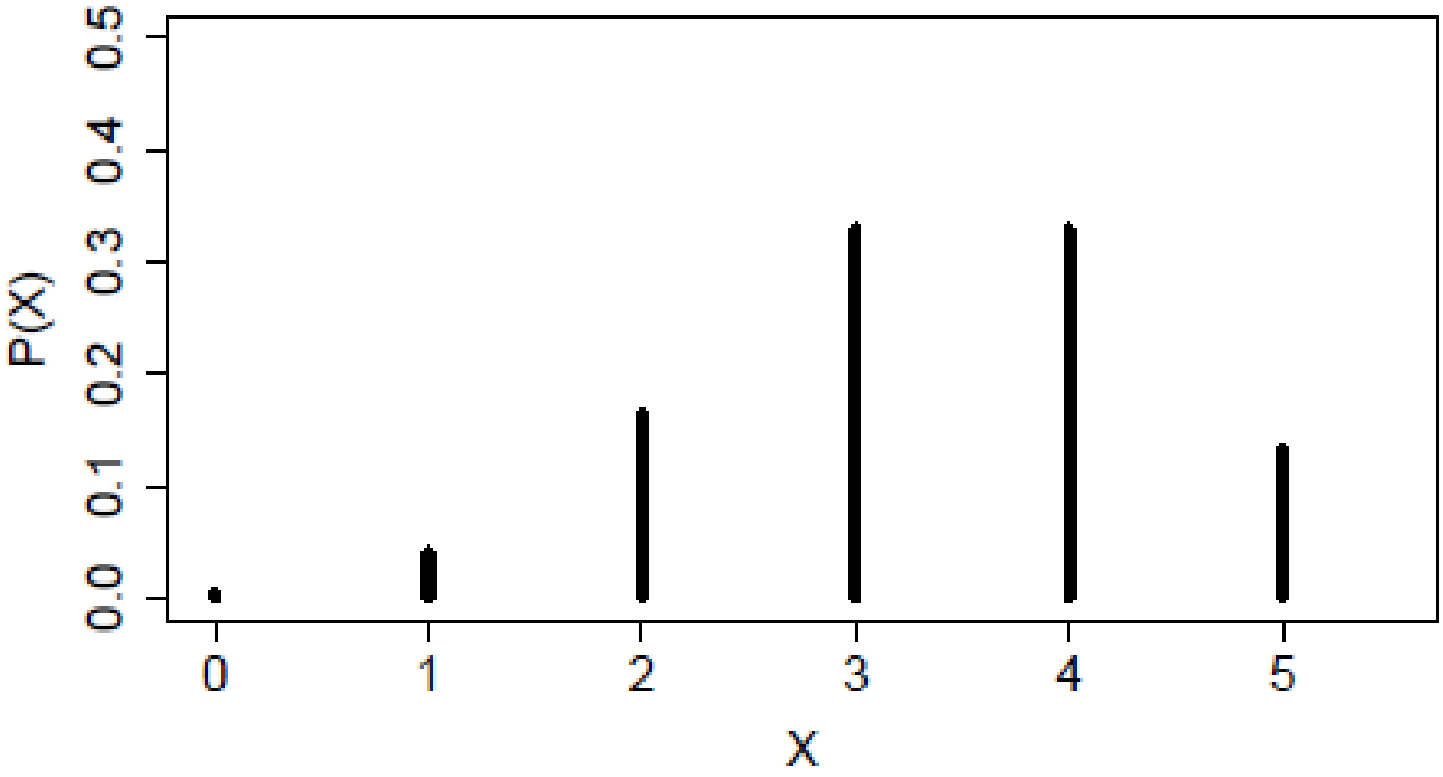


Figure 1: A plot of the Binomial(5, 2/3) probabilities.

The Binomial distribution

The key components of a Binomial distribution

In general a Binomial distribution arises when we have the following 4 conditions

- n identical trials, e.g. 5 coin tosses
- 2 possible outcomes for each trial "success" and "failure", e.g. Heads or Tails
- Trials are independent, e.g. each coin toss doesn't affect the others
- P("success") = p is the same for each trial, e.g. P(Head) = 2/3 is the same for each trial

Binomial distribution probabilities

If we have the above 4 conditions then if we let

$$X = \text{No. of "successes"}$$

then the probability of observing x successes out of n trials is given by

$$\boxed{P(X = x) \quad = \quad {}^nC_x p^x (1-p)^{(n-x)} \qquad x = 0, 1, \ldots, n}$$

If the probabilities of X are distributed in this way, we write

$$\boxed{X \sim \text{Bin}(n, p)}$$

n and p are called the **parameters** of the distribution. We *say* X follows a binomial distribution with parameters n and p.

Examples

1. Suppose $X \sim \text{Bin}(10, 0.4)$, what is $P(X = 7)$?

2. Suppose $Y \sim \text{Bin}(8, 0.15)$, what is $P(Y < 3)$?

3. Suppose $W \sim \text{Bin}(50, 0.12)$, what is $P(W > 2)$?

Solutions:

1. $$\begin{aligned} P(X=7) &= {}^{10}C_7(0.4)^7(1-0.4)^{(10-7)} \\ &= (120)(0.4)^7(0.6)^3 \\ &= 0.0425 \end{aligned}$$

2. $$\begin{aligned} P(Y<3) &= P(Y=0)+P(Y=1)+P(Y=2) \\ &= {}^8C_0(0.15)^0(0.85)^8 + {}^8C_1(0.15)^1(0.85)^7 + {}^8C_2(0.15)^2(0.85)^6 \\ &= 0.2725 + 0.3847 + 0.2376 \\ &= 0.8948 \end{aligned}$$

3. $$\begin{aligned} P(W>2) &= P(W=3)+P(W=4)+\ldots+P(W=50) \\ &= 1-P(W\leq 2) \\ &= 1-\Big(P(W=0)+P(W=1)+P(W=2)\Big) \\ &= 1-\Big({}^{50}C_0(0.12)^0(0.88)^{50} + {}^{50}C_1(0.12)^1(0.88)^{49} + {}^{50}C_2(0.12)^2(0.88)^{48}\Big) \\ &= 1-\Big(0.00168+0.01142+0.03817\Big) \\ &= 0.94874 \end{aligned}$$

Expected Values

- **Example:** Let X be the random variable equal to the sum of the numbers that appear when a pair of dice are rolled.
- There are 36 outcomes (=pairs of numbers from 1 to 6).
- The **range** of X is {2,3,4,5,6,7,8,9,10,11,12}
- Are the 36 outcomes equally likely?
- Yes, if the dice are not biased.
- Are the 11 values of X equally likely?
  No, the probabilities vary across values.

Expected Values

- $P(X=2) = 1/36$
- $P(X=3) = 2/36 = 1/18$
- $P(X=4) = 3/36 = 1/12$
- $P(X=5) = 4/36 = 1/9$
- $P(X=6) = 5/36$
- $P(X=7) = 6/36 = 1/6$
- $P(X=8) = 5/36$
- $P(X=9) = 4/36 = 1/9$
- $P(X=10) = 3/36 = 1/12$
- $P(X=11) = 2/36 = 1/18$
- $P(X=12) = 1/36$

Expected Values

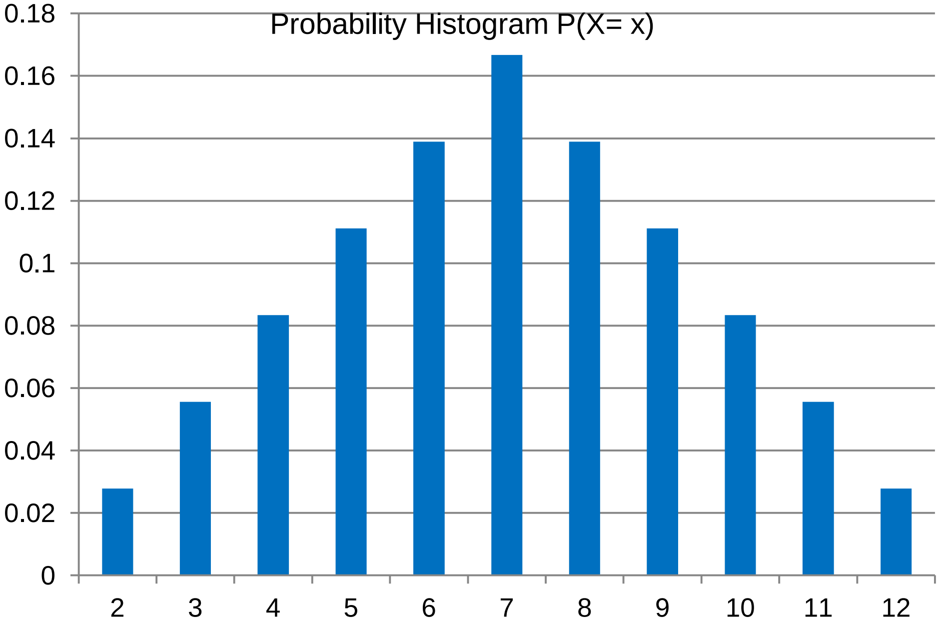
- $E(x) = 2\,(1/36) + 3(1/18)+4(1/12) + 5(1/9) + 6(5/36) + 7(1/6) + 8(5/36) + 9*(1/9) + 10*(1/12)+ 11*(1/18) + 12*(1/36)$
- $E(x) = 7$

- This means that if we roll 2 dice many times, sum all the numbers that appear and divide the sum by the number of trials, we expect to find a value of 7.

Probability Histogram

- The above random variable X, and the associated set of probabilities for each possible value, is called a discrete **probability histogram** or **Probability Distribution.**
- It gives the probabilities of all the possible values of the discrete r.v.
- Note: sum of p(X) is equal to 1.
- Can illustrate histogram with a diagram.

Probability Histogram



- Probabilities for a discrete random variable.
- Note: The sum of all probabilities is = 1.
- In the limit of a large number of values for the random variable, the histogram becomes a smooth continuous function.

Discrete Random Variable

- A discrete random variable is one whose set of possible values is countable (finite).
- A random variable is discrete if and only if its **cumulative probability distribution function** is a stair-step function; *i.e.*, if it is piecewise constant and only increases by jumps.
- **Examples:** no. of people in a family, no. of records in a file, age of a person to nearest year.

Continuous Random Variable

- A **random variable** X is also called *continuous* if its set of possible values is very large (infinite, in theory), and the chance that it takes any particular value is zero (in symbols, if $P(X = y) = 0$ for every real number y).
- A random variable is continuous if and only if its **cumulative probability distribution function** is a continuous function (a function with no jumps). **Examples**: temperature, exact height, exact age.

Continuous Random Variable

- When X is a continuous r.v. then the probability histogram → a continuous smooth function called the **probability density function.** The chance that a continuous r.v. is in any range of values can be calculated as the area under a curve over that range of values. This curve is the **probability density function** of the r.v.
- That is, if X is a continuous random variable, there is a function $f(x)$ such that for every pair of numbers $a<=b$,
    - $P(a<= X <=b)$ = (area under f between a and b)

Exercises

Past exam questions:

1. Define mathematically what is meant by Independent events. In an experiment, 2 dice are tossed. Let the random variable X denote the sum of the 2 numbers appearing.
   1. Find the probability distribution of X
   2. Compute E[X], the expected value of X
2. Define mathematically (using formulas)
   1. Sample Space
   2. Conditional Probability, P(AlB)
   3. Mutually exclusive events

3. Let A and B be events with P(A)=1/3, P(B)=1/4 and P(AUB)=1/2. Compute (using fractions)

(i) P(AIB) and P(BIA)

(ii) Are A and B independent?

4. Whenever horses A, B and C race together, their respective probabilities of winning are 0.3, 0.5 and 0.2. They race 3 times.

(i) Compute the prob. that A wins all 3 races.

(ii) Compute the prob. That A wins the first 2 races and loses the 3rd race.

(iii) Compute the prob. That A, B and C each win one race.