# Module 4: Discrete Random Variables and Probability Distributions

## 4.1 Discrete RV's and Probability Mass Functions (*Weiss §5.1, 5.2*)

► It is frequent that interest lies in a function of the outcome, rather than on the outcome itself.

**Example 4.1.** Game of Craps

Playing the game of craps: at first, the player rolls two fair dice.

Then, depending on the sum of the two dice,

- if this sum is 7 or 11 : the player wins,
- if this sum is 2, 3 or 12 : the player loses,
- otherwise : the payer re-rolls until the same sum or 7 is obtained,
  $\rightarrow$ if same sum : the player wins,
  $\rightarrow$ if 7 : the player loses.

Focusing only on the initial sum of the two dice, the sample space of the experiment is

$$\Omega = \{\omega = (\omega_1, \omega_2) : 1 \le \omega_1 \le 6, 1 \le \omega_2 \le 6\}.$$

It should be noted that the specific outcome $\omega \in \Omega$ is unimportant.

Rather, what is important is the sum $\omega_1 + \omega_2$.

> **Definition 4.1.** Random Variable (RV)
>
> A **random variable** is a real-valued function whose domain is the sample space $\Omega$, i.e. it is a function $X : \Omega \rightarrow \mathbb{R}$.

**Example 4.1.** (cont'd)

To be more formal, we write

$$X(\omega) = \omega_1 + \omega_2,$$

so that X represents the sum of the two dice. Now, let

$$A_i = \{\omega \in \Omega : X(\omega) = i\},$$

which is abbreviated as $A_i = \{X = i\}$. Interpret the event $A_{10}$ and calculate $\mathbb{P}(A_{10})$.

**Example 4.2.** Red and Blue balls

An urn contains $n$ red balls and $n+1$ blue balls. ($n \geq 2$)

Two balls are to be drawn without replacement from the urn by a gambler who pays \$1 to play. Depending on the outcome of the draw, the player is then payed back

- \$2 if both balls are red,
- \$1 if one ball is red and the other is blue,
- nothing if both balls are blue.

As we are only interested in the result of the bet, we define

$X:$ gambler's net gain when playing the game.

Formally describe $\Omega$ for the random experiment consisting of drawing two balls from the urn and define $X(\omega)$ as a function of the outcome of this experiment. Also, interpret the events

$$A_i = \{\omega \in \Omega : X(\omega) = i\}$$

and calculate $\mathbb{P}(A_i)$ for relevant values of $i$.

**Definition 4.2.** Discrete Random Variable

A **discrete** random variable is one for which there exists a set $K$ that is finite or infinite countable such that

$$\mathbb{P}(X \in K) = 1.$$

As a consequence, if the set of possible values for $X$, called the **range** of $X$ and denoted $R_X$, is finite or infinite countable, then $X$ is a discrete random variable.

► There are other types or random variables!

(Recall continuous random variables and normal distributions...)

For example, heights, weights and waiting times are often continuous.

► In examples 4.1 and 4.2, $R_X$ is finite so that $X$ is a discrete RV.

**Example 4.3.** (See Example 3.9)

When rolling a fair die until a result of 6 is obtained, let

$X:$ nb. of rolls needed to see a 6 for the first time.

Argue that $X$ is a discrete random variable and find $\mathbb{P}(X = n)$ for $n \geq 1$.

**Definition 4.3.** Probability Mass Function

The **probability mass function** (PMF) of a discrete random variable X, denoted $P_X$, is the real-valued function

$$\begin{aligned} P_X(x) &= \mathbb{P}(X = x) \\ &= \mathbb{P}\Big(\big\{\omega \in \Omega : X(\omega) = x\big\}\Big). \end{aligned}$$

**Example 4.4.** Rolling two dice

Let $Y$ : maximum result.

Formally define Y as a function of the outcome of a random experiment and obtain its PMF.

**Proposition 4.4.** Basic Properties of Probability Mass Functions

The PMF of a discrete random variable X satisfies

1. $P_X(x) \geq 0$ for all $x \in R_X$,
2. $\sum_{x \in R_X} P_X(x) = 1$.

**Example 4.3.** (cont'd)

For

$X$ : number of rolls to get the first 6,

we have found that

$$\mathbb{P}(X = n) = \frac{1}{6}\left(\frac{5}{6}\right)^{n-1} \qquad \text{for } n \in R_X = \{1, 2, \ldots\}.$$

Deduce the PMF of X and verify the previous properties hold for it.

**Example 4.4.** (cont'd)

For

$Y$ : maximum result when rolling two fair dice,

we have found that

$$P_Y(k) = \frac{2k-1}{36} \qquad \text{for } k \in R_Y = \{1, 2, \ldots, 6\}.$$

Verify the previous properties hold for the PMF of Y.

**Proposition 4.5.** Fundamental Probability Formula

Let $X$ be a discrete random variable with PMF $P_X$. Then,

$$\mathbb{P}(X \in A) = \sum_{x \in A \cap R_X} P_X(x) \qquad \text{for any } A \subset \Omega.$$

In other words, we add the probabilities over all possible values of $X$ that are in $A$.

**Example 4.3.** (cont'd)

For

$X$ : number of rolls to get the first 6,

find the following probabilities:

- $\mathbb{P}(3 \le X \le 6)$,
- $\mathbb{P}(X \le k)$ for any $k \ge 1$,
- $\mathbb{P}(3 \le X \le 6)$ using the result obtained above.

## 4.2 Bernoulli Trials and Related RV's (*Weiss* §*5.3, 5.6, 5.7*)

**Definition 4.6.** Bernoulli Trials

**Bernoulli trials** are repeated trials of the same random experiment where

1. the outcomes of all trials are independent of one another,
2. each trial results in one of two possible outcomes,

   (Typically, outcomes are classified as "success" or "failure".)
3. the probability of the two outcomes ("success" and "failure") remain the same from trial to trial.

   (Typically, the probability of a "success" is denoted $p$ and the probability of a "failure" is denoted $q = 1 - p$.)

**Example 4.5.** Bernoulli trials

Which of the following can be seen as a sequence of Bernoulli trials?

- successive flips of a (not necessarily fair) coin,
- successive bets on "00" at the american roulette,
- successive draws in Pólya's urn model (*cf*. Example 3.4).

### 4.2.1 Binomial RV's (*Weiss* §5.3)

**Definition 4.7.** Binomial Random Variable

Assume a sequence of $n$ Bernoulli trials with success probability $p$ is to take place, and let

$X$ : number of successes in the $n$ trials.

Then, the PMF of $X$ is

$$P_X(x) = \begin{cases} \binom{n}{x} p^x (1-p)^{n-x} & \text{for } x = 0, 1, \ldots, n, \\ 0 & \text{otherwise,} \end{cases}$$

with range $R_X = \{0, 1, \ldots, n\}$ that is finite.

Also, $X$ is said to have the **Binomial distribution** with parameters $n$ and $p$, which we write as $X \sim B(n, p)$.

**Example 4.6.** Rolling a die

A fair die is rolled 5 times. Calculate the probability a result of 6 is obtained at least 4 times.

Assume now the die has been loaded so that the probability of getting a 6 on any roll is $7/10$. When this die is rolled 5 times, what is the probability a result of 6 is obtained at least 4 times?

**Example 4.7.** Gene inheritance

Suppose a trait of a person (such as eye color) is classified on the basis of one pair of genes. Further suppose $d$ represents a dominant gene gene and $r$ a recessive gene.

Then,
- $dd$ is pure dominance,
- $rr$ is pure recessive,
- $rd$ (or $dr$) is hybrid.

However, pure dominance and hybrid are alike in appearance and children receive one gene from each parent.

Finally, assume that each child, independently of other children, is equally likely to inherit either of the 2 genes from each parent.

If two hybrid parents have 4 children, what is the probability that at least 2 have the outward appearance of the dominant gene?

**Example 4.8.** Lotto $6/49$

If you buy a Lotto 6/49 ticket for every draw (i.e. twice a week), what is the probability you will win at least 2 prizes over one year?

▶ Note: the binomial PMF satisfies the basic properties of PMF's

- $P_X(x) \geq 0$ for all $x = 0, 1, \ldots, n$,
- $\displaystyle\sum_{x=0}^{n} P_X(x) = 1$.

**Proposition 4.8.** Property of the Binomial PMF

Assume that $X \sim B(n, p)$ and let

$$m = \lfloor (n+1)p \rfloor$$

where $\lfloor x \rfloor$ denotes the "floor", or "integer part" of $x$, i.e. the largest integer smaller then or equal to $x$.

1. If $(n+1)p$ is **not** an integer, then $P_X$ is
   - strictly increasing for $x = 0, 1, \ldots, m$,
   - strictly decreasing for $x = m, m+1, \ldots, n$.
2. If $(n+1)p$ is an integer, then $P_X$ is
   - strictly increasing for $x = 0, 1, \ldots, m-1$,
   - such that $P_X(m-1) = P_X(m)$,
   - strictly decreasing for $x = m, m+1, \ldots, n$.

In any case, $P_X$ is maximum for $x = m$.

**Example 4.6.** (cont'd)

When the fair and loaded dice are rolled 5 times, what is the most likely number of 6's in each case?

What if the fair die is rolled 15 times instead?

**Definition 4.9.** Bernoulli Random Variable

Assume *one* trial with success probability $p$ is to take place, and let

$$X = \begin{cases} 1 & \text{if the trial is a success,} \\ 0 & \text{if the trial is a failure.} \end{cases}$$

Then, $X$ is said to have a **Bernoulli distribution** with parameter $p$. Often, $X$ is referred to as an **indicator** random variable (indicating wether a success occurred or not).

We write $X \sim B(1, p)$ as this is the binomial distribution in the case of $n = 1$ trial.

**Example 4.8** (cont'd)

Assume that $n$ people each buy a randomly selected ticket for the same draw of Lotto 6/49. How large should $n$ be to make the probability that the jackpot is won larger than 50%, 75% and 90%?

### 4.2.2 Geometric RV's (*Weiss §5.4*)

**Definition 4.10.** Geometric Random Variable

Assume a sequence of Bernoulli trials with success probability $p$ is to take place, and let

$X$ : number of trials required to see a first success.

Then, the PMF of $X$ is

$$P_X(x) = \begin{cases} p(1-p)^{x-1} & \text{for } x = 1, 2, \ldots, \\ 0 & \text{otherwise,} \end{cases}$$

with range $R_X = \mathbb{N}$ that is infinite countable.

In this context, $X$ is said to have the **Geometric distribution** with parameter $p$ and we write $X \sim G(p)$.

**Example 4.3.** (cont'd)

When rolling a fair die, we have defined

$X$ : number of rolls to get the first.

What is the distribution of $X$?

► Note: the geometric PMF satisfies the basic properties of PMF's

- $P_X(x) \geq 0$ for all $x = 0, 1, \ldots, n$,
- $\sum_{x=1}^{\infty} P_X(x) = 1.$

**Proposition 4.11.** Tail Probabilities for Geometric RV's

Assume that $X \sim G(P)$. Then

$$\mathbb{P}(X > k) = (1-p)^k \qquad \text{for } k \in \mathbb{N}.$$

**Example 4.8.** (cont'd)

You buy a Lotto 6/49 ticket for every draw. What is the probability you do not win any prizes over a one-year period?

Assume now you have been playing Lotto 6/49 for a six-month period (i.e. 52 draws) without winning a prize even once. What is the probability you won't win a prize in the coming year?

**Proposition 4.12.** Lack-of-memory Property of Geometric RV's

Assume that $X \sim G(P)$. Then

$$\mathbb{P}(X = n + k \mid X > n) = \mathbb{P}(X = k) \qquad \text{for all } k, n \in \mathbb{N}.$$

Also,

$$\mathbb{P}(X > n + k \mid X > n) = \mathbb{P}(X > k) \qquad \text{for all } k, n \in \mathbb{N}.$$

This is known as the **lack-of-memory** property of the Geometric distribution.

Note that the Goemetric RV is the only discrete RV with this property.

**Example 4.8.** (cont'd)

Redo the second part of the previous example using the lack-of-memory property.

**Example 4.9.** Lifetimes of light-bulbs

It is assumed lifetimes (in hours) of light-bulbs of a certain type are randomly distributed according to a geometric distribution with parameter $p = 1/3000$. What is the probability a light-bulb of this type lasts for at least 1000 hours?

Given it has been in use for 3000 hours, what is the probability a light-bulb of this type lasts for at least another 1000 hours?

### 4.2.3 Negative Binomial RV's (*Weiss* §5.7)

**Definition 4.13.** Negative Binomial Random Variable

Assume a sequence of Bernoulli trials with success probability $p$ is to take place and let, for $r \geq 1$,

$X$ : number of trials required to obtain the $r^{\text{th}}$ success.

Then, the PMF of $X$ is

$$P_X(x) = \begin{cases} \binom{x-1}{r-1} p^r (1-p)^{x-r} & \text{for } x = r, r+1, \ldots, \\ 0 & \text{otherwise,} \end{cases}$$

with range $R_X = \{r, r+1, \ldots\}$ that is infinite countable.

In this context, $X$ is said to have the **Negative Binomial distribution** with parameters $r$ and $p$ and we write $X \sim NB(r, p)$.

► Note that the $NB(1, p)$ and $G(p)$ distributions are the same!

**Example 4.3.** (cont'd)

When repeatedly rolling a fair die, we have seen that for

$X$ : number of rolls to get the first 6,

we have that $X \sim G(1/6)$. Obviously, we can also write $X \sim NB(1, 1/6)$.

But, what is the probability that at least 4 rolls are needed to obtain two results of 6?

**Example 4.10.** Banach match problem

A pipe-smoking mathematician puts a matchbox containing N matches in each of his 2 pockets (right and left). Each time he needs a match, he is assumed to be equally likely to take it from either pocket.

At the moment when one matchbox is found to be empty, let X denote the number of matches left in the other matchbox. What is the PMF of X?

**Definition 4.14.** Generalized Binomial Coefficient

For $a \in \mathbb{R}$ and $k \in \mathbb{Z}$, define the **generalized binomial coefficient** $\binom{a}{k}$ by

1. $\displaystyle\binom{a}{0} = 1,$

2. $\displaystyle\binom{a}{k} = 0$ if $k < 0$,

3. $\displaystyle\binom{a}{k} = \frac{\overbrace{a(a-1)\cdots(a-k+1)}^{k \text{ terms}}}{k!}$ if $k > 0$.

- Note that
  - as before, we have that $\displaystyle\binom{a}{1} = a,$
  - if $a = n$ is a positive integer, then

$$\binom{n}{k} = \begin{cases} \dfrac{n!}{k!(n-k)!} & \text{if } 1 \le k \le n, \\ 1 & \text{if } k = 0, \\ 0 & \text{if } k < 0 \text{ or } k > n. \end{cases}$$

- Using these generalized coefficients, we can write for $X \sim NB(r, p)$

$$P_X(x) = \binom{-r}{x-r} p^r (1-p)^{x-r} \qquad \text{for } x \ge r,$$

hence, the name *negative* binomial.

## 4.3 Hypergeometric RV's (*Weiss* §*5.4*)

**Definition 4.15.** Hypergeometric Random Variable

Assume that, in a population of $N$ individuals, a proportion $p$ of them have a certain attribute.

From this population, $n$ individuals are sampled *without replacement*, and we let

$X$ : number of individuals in the sample having the attribute.

Then, the PMF of $X$ is

$$P_X(x) = \begin{cases} \dfrac{\binom{Np}{x}\binom{N(1-p)}{n-x}}{\binom{N}{n}} & \text{for } x = 0, 1, \ldots, n, \\ 0 & \text{otherwise,} \end{cases}$$

with maximal range $R_X = \{0, 1, \ldots, n\}$ that is finite.

In addition, $X$ is said to have the **Hypergeometric distribution** with parameters $N$, $n$ and $p$ and we write $X \sim H(N, n, p)$.

► Notes:

- Sampling without replacement implies that individuals cannot be selected more than once to be included in the sample.
- If the sampling was instead done with replacement, we would have that that $X \sim B(n, p)$ for any value of $N$.
- The actual range of $X$ is all integers such that
  - $x \geq 0$,
  - $Np - x \geq 0$,
  - $n - x \geq 0$,
  - $N(1-p) - (n-x) \geq 0$,

  or,
  $$\max\big[0, n - N(1-p)\big] \leq x \leq \min\big[n, Np\big],$$
  that is, for $x$ to cover the whole range of values $0, 1, \ldots, n$ we need enough individuals with/without the attribute.

**Example 4.8.** (cont'd)

When playing the Lotto 6/49, let

$X$ : number of selected balls, out of 6, that are drawn on a given draw.

Find the range and probability distribution of $X$. Also, using this, recalculate the probability of winning a prize on any given draw.

**Example 4.11.** Drawing balls from two urns

There are two urns, each containing 15 balls. Specifically, we have

| | | | |
|---|---|---|---|
| Urn 1 : | 2 red balls | Urn 2 : | 5 red balls |
| | 13 blue balls | | 10 blue balls |

An urn is randomly selected by flipping a fair coin, and then, three balls are drawn without replacement from the selected urn.

Let $X$ denote the number of red balls drawn. Find the PMF of $X$.

**Example 4.12.** Politics

The new leader of a political party is to be chosen. Assume that this party has 30 000 members and that 65% of them support the leading candidate.

Ten randomly selected members are to be surveyed. Let

$X$ : number of surveyed members that support the candidate.

Compare the PMF of $X$ in the cases where the sampling is done with and without replacement.

**Proposition 4.16.** Binomial Approximation to the Hypergeometric

Assume that $X \sim H(N, n, p)$ where the sample size $n$ is small relative to the population size $N$ (usually, it is required that $n/N \leq 0.05$).

Then,

$$P_X(x) = \frac{\binom{Np}{x}\binom{N(1-p)}{n-x}}{\binom{N}{n}} \simeq \binom{n}{x} p^x (1-p)^{n-x} \qquad \text{for } x = 0, 1, \ldots, n.$$

In other words, the Hypergeometric distribution $H(N, n, p)$ can be approximated by the $B(n, p)$ distribution,

or, sampling with/without replacement makes almost no difference.

► Note: hypergeometric probabilities satisfy the basic properties of PMF's!

## 4.4 Poisson RV's (*Weiss* §5.5)

**Definition 4.17.** Poisson Random Variable

A random variable X such that

$$P_X(x) = \begin{cases} \dfrac{e^{-\lambda}\lambda^x}{x!} & \text{for } x = 0, 1, \ldots \\ 0 & \text{otherwise,} \end{cases}$$

i.e. having range $R_X = \{0, 1, \ldots\} = \mathbb{N} \cup \{0\}$ that is infinite countable, is said to have a **Poisson distribution** with parameter $\lambda$.

We write $X \sim P(\lambda)$.

► Notes:

- The Poisson distribution is typically used for

  X : number of occurrences of an event over a time-interval of given length (or area in space),

  when events can be assumed to occur independently of each other. In this case, $\lambda$ represents the average rate of occurrence of these events.

- Poisson probabilities satisfy the basic properties of PMF's!

**Example 4.13.** Passengers at a bus stop

Passengers arrive at a downtown bus stop at an average rate of 1 per minute.

Calculate the probability that no more than 2 passengers arrive at the bus stop

- in the next minute,
- in the next 3 minutes.

**Example 4.14.** Knots and planks

A certain type of planks from a lumberyard is known to have, on average, 0.5 knots per running foot of wood. What is the probability that a randomly selected 8-foot plank has no more that one knot?

Out of 10 such planks, what is the probability that none are free of knots?

**Proposition 4.18.** Poisson Approximation to the Binomial

Assume that $X \sim B(n, p)$ where the sample size $n$ is large and $p$ is small (usually, it is required that $np^2 \leq 0.05$).

Then,

$$P_X(x) = \binom{n}{x} p^x (1-p)^{n-x} \simeq \frac{e^{-np}(np)^x}{x!} \qquad \text{for } x = 0, 1, \ldots, n.$$

In other words, the binomial distribution $B(n, p)$ can be approximated by the $P(\lambda)$ distribution, where $\lambda = np$.

**Example 4.8.** (cont'd)

When playing Lotto 6/49, let

$X$ : number of prizes won over one year.

It should be clear that $X \sim B(104, 0.01864)$.

Compare these binomial probabilities to their Poisson approximation.

**Example 4.15.** Insecticide effectiveness

A large number of insects are expected to be attracted to the rose plants located in a garden. Suppose that 10 000 insects infest that garden and that a commercial insecticide has been applied prior to the infestation. Suppose also that the insecticide is 99.5% effective in killing insects that come in contact with it and that all insects entering the garden will eventually come in contact with the insecticide. Finally, let X denote the number of surviving insects.

Calculate the probability that at least 52 and no more than 55 of the 10 000 insects survive entering the garden.

## 4.5 Functions of a Discrete RV (*Weiss §5.8*)

**Example 4.16.** Profit

Assume that the weekly demand at a certain store for a product is uniformly distributed on $\{0, 1, \ldots, N\}$. Further assume that

- each unit of the product sold leads to a profit of \$10,
- each unsold unit still in stock at the end of the week leads to a loss of \$5,
- any demand that occurs while the store is out-of-stock is lost,
- at the beginning of the week, the stock is replenished to its maximum level of $s$, with $0 < s \leq N$.

Given a stock level of $s$, what is the probability the store makes a profit on this product over a week?

**Proposition 4.19.** PMF of a Function of a Random Variable

Let $X$ be a discrete random variable with PMF $P_X$ (and rande $R_X$) and $g$ be a real-valued function of $R_X \to \mathbb{R}$. Then, the PMF of the random variable $Y = g(X)$ is given by

$$P_Y(y) = \begin{cases} \displaystyle\sum_{\{x \in R_X \,:\, g(x)=y\}} P_X(x) & \text{for } y \in R_Y, \\ 0 & \text{otherwise,} \end{cases}$$

where the rande of $Y$ is $R_Y = \{y : y = g(x) \text{ for some } x \in R_X\}$.

**Example 4.17.** Quadratic transformation

A random variable $X$ has the following PMF

| $x$ | $-2$ | $-1$ | $0$ | $1$ | $2$ | $3$ |
|---|---|---|---|---|---|---|
| $P_X(x)$ | 0.10 | 0.20 | 0.20 | 0.15 | 0.15 | 0.20 |

Find the PMF of $Y = X^2$ and use it to calculate the probability that $X^2 \geq 2$.