

# **MAT 2377**

## **Probability and Statistics for Engineers**

### **Chapter 2**

### **Discrete Distributions**

P. Boily (uOttawa)

Winter 2021

## Contents

### 2.1 – Random Variables and Distributions (p.3)

- Notation (p.4)
- Properties (p.5)

### 2.2 – Expectation of a Discrete Random Variable (p.13)

- Mean and Variance (p.16)
- Standard Deviation (p.18)
- Properties (p.22)

### 2.3 – Binomial Distributions (p.23)

- Binomial Experiments (p.24)
- Probability Mass Function (p.25)
- Expectation and Variance (p.26)

2.4 – Geometric Distributions (p.32)

2.5 – Negative Binomial Distribution (p.34)

2.6 – Poisson Distributions (p.36)

Appendix – Summary (p.44)

## 2.1 – Random Variables and Distributions

Recall that, for any random “experiment,” the set of all possible outcomes is denoted by  $\mathcal{S}$ .

A **random variable** (r.v.) is a function  $X : \mathcal{S} \rightarrow \mathbb{R}$ , i.e. it is a rule that associates a (real) number to every outcome of the experiment.

$\mathcal{S}$  is the **domain** of the r.v.  $X$ ;  $X(\mathcal{S}) \subseteq \mathbb{R}$  is its **range**.

A **probability distribution function** (p.d.f.) is a function  $f : \mathbb{R} \rightarrow \mathbb{R}$  which specifies the probabilities of the values in  $X(\mathcal{S})$ .

When  $\mathcal{S}$  is discrete, we say that  $X$  is a **discrete r.v.** and the p.d.f. is called a **probability mass function** (p.m.f.).

## Notation for R.V.

We use the following notation throughout:

- capital roman letters ( $X, Y$ , etc.) to denote r.v.
- corresponding lower case roman letters ( $x, y$ , etc.) to denote *generic values taken* by the r.v.

A discrete r.v. can be used to define events: if  $X$  takes values  $X(\mathcal{S}) = \{x_i\}$ , then we can define events  $A_i = \{s \in \mathcal{S} : X(s) = x_i\}$ .

- The p.m.f. of  $X$  is  $f(x) = P(\{s \in \mathcal{S} : X(s) = x\}) := P(X = x)$ .
- The **cumulative distribution function** (c.d.f.) of  $X$  is  $F(x) = P(X \leq x)$ .

## Properties of the P.M.F. and the C.D.F.

If  $X$  is a discrete random variable with p.m.f.  $f(x)$  and c.d.f.  $F(x)$ , then

- $0 < f(x) \leq 1$  for all  $x \in X(\mathcal{S})$ ;
- $\sum_{s \in \mathcal{S}} f(X(s)) = \sum_{x \in X(\mathcal{S})} f(x) = 1$ ;
- for any event  $A \subseteq \mathcal{S}$ ,  $P(X \in A) = \sum_{x \in A} f(x)$ ;
- for any  $a, b \in \mathbb{R}$ ,

$$P(a < X) = 1 - P(X \leq a) = 1 - F(a)$$

$$P(X < b) = P(X \leq b) - P(X = b) = F(b) - f(b)$$

- for any  $a, b \in \mathbb{R}$ ,

$$\begin{aligned}P(a \leq X) &= 1 - P(X < a) = 1 - (P(X \leq a) - P(X = a)) \\&= 1 - F(a) + f(a)\end{aligned}$$

We can use these results to compute the probability of a **discrete** r.v.  $X$  falling in various intervals:

$$P(a < X \leq b) = P(X \leq b) - P(X \leq a) = F(b) - F(a)$$

$$P(a \leq X \leq b) = P(a < X \leq b) + P(X = a) = F(b) - F(a) + f(a)$$

$$P(a < X < b) = P(a < X \leq b) - P(X = b) = F(b) - F(a) - f(b)$$

$$P(a \leq X < b) = P(a \leq X \leq b) - P(X = b) = F(b) - F(a) + f(a) - f(b)$$

**Examples:**

1. Flip a fair coin. The outcome space is  $\mathcal{S} = \{\text{Head}, \text{Tail}\}$ . Let  $X : \mathcal{S} \rightarrow \mathbb{R}$  be defined by  $X(\text{Head}) = 1$  and  $X(\text{Tail}) = 0$ . Then  $X$  is a discrete random variable (as a convenience, we write  $X = 1$  and  $X = 0$ ).

If the coin is fair, the p.m.f. of  $X$  is  $f : \mathbb{R} \rightarrow \mathbb{R}$ , where

$$f(0) = P(X = 0) = 1/2, \quad f(1) = P(X = 1) = 1/2, \quad f(x) = 0 \text{ for all other } x.$$

2. Roll a fair die. The outcome space is  $\mathcal{S} = \{1, \dots, 6\}$ . Let  $X : \mathcal{S} \rightarrow \mathbb{R}$  be defined by  $X(i) = i$  for  $i = 1, \dots, 6$ . Then  $X$  is a discrete r.v.

If the die is fair, the p.m.f. of  $X$  is  $f : \mathbb{R} \rightarrow \mathbb{R}$ , where

$$f(i) = P(X = i) = 1/6, \text{ for } i = 1, \dots, 6, \quad f(x) = 0 \text{ for all other } x.$$



3. For the random variable  $X$  from the previous example, the c.d.f. is  $F : \mathbb{R} \rightarrow \mathbb{R}$ , where

$$F(x) = P(X \leq x) = \begin{cases} 0 & \text{if } x < 1 \\ i/6 & \text{if } i \leq x < i + 1, \text{ for } i = 1, \dots, 6 \\ 1 & \text{if } x \geq 6 \end{cases}$$

4. For the same random variable, we can compute  $P(3 \leq X \leq 5)$  directly:

$$P(3 \leq X \leq 5) = P(X = 3) + P(X = 4) + P(X = 5) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{1}{2},$$

or we can use the c.d.f.

$$P(3 \leq X \leq 5) = F(5) - F(3) + f(3) = \frac{5}{6} - \frac{3}{6} + \frac{1}{6}$$

5. The number of calls received over a specific time period,  $X$ , is a discrete random variable, with potential values  $0, 1, 2, \dots$
6. Consider a 5-card poker hand consisting of cards selected at random from a 52-card deck. Find the probability distribution of  $X$ , where  $X$  indicates the number of red cards ( $\diamond$  and  $\heartsuit$ ) in the hand.





**Solution:** in all there are  $\binom{52}{5}$  ways to select a 5-card poker hand from a 52-card deck.

By construction,  $X$  can take on values  $x = 0, 1, 2, 3, 4, 5$ .

If  $X = 0$ , then none of the 5 cards in the hands are  or , and all of the 5 cards in the hands are  or .

There are thus  $\binom{26}{0} \cdot \binom{26}{5}$  5-card hands that only contain black cards, and

$$P(X = 0) = \frac{\binom{26}{0} \cdot \binom{26}{5}}{\binom{52}{5}}.$$

In general, if  $X = x$ ,  $x = 0, 1, 2, 3, 4, 5$ , there are  $\binom{26}{x}$  ways of having  $x$   or  in the hand, and  $\binom{26}{5-x}$  ways of having  $5 - x$   and  in the hand, so that

$$f(x) = P(X = x) = \frac{\binom{26}{x} \cdot \binom{26}{5-x}}{\binom{52}{5}}, \quad x = 0, 1, 2, 3, 4, 5; \quad f(x) = 0 \text{ otherwise.}$$

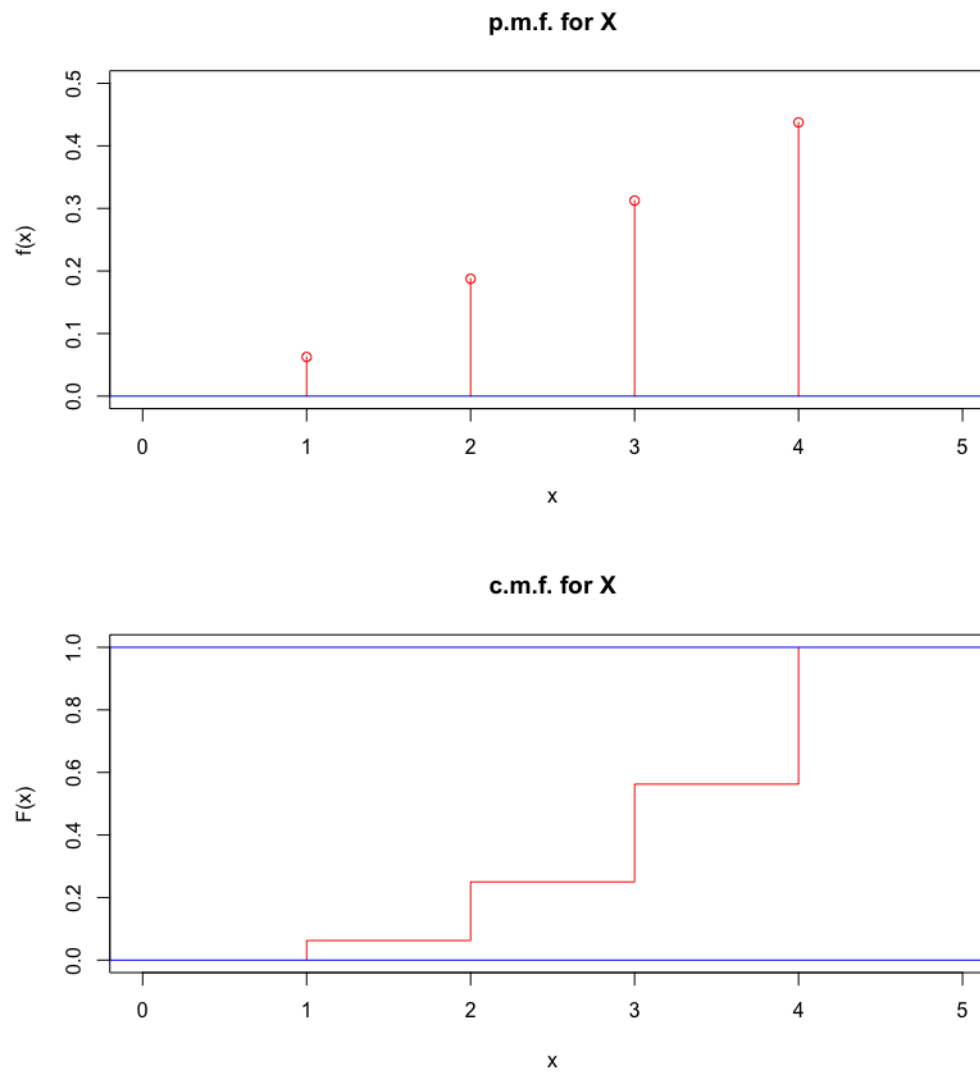
7. Find the c.d.f.  $F(x)$  of a discrete random variable  $X$  with p.m.f.  $f(x)$  defined by  $f(x) = 0.1x$  if  $x = 1, 2, 3, 4$  and  $f(x) = 0$  otherwise.

**Solution:**  $f(x)$  is indeed a p.m.f. because  $0 < f(x) \leq 1$  for all  $x$  and

$$\sum_{x=1}^4 0.1x = 0.1(1 + 2 + 3 + 4) = 0.1 \frac{4(5)}{2} = 1.$$

We have

$$F(x) = \begin{cases} 0 & \text{if } x < 1 \\ 0.1 & \text{if } 1 \leq x < 2 \\ 0.3 & \text{if } 2 \leq x < 3 \\ 0.6 & \text{if } 3 \leq x < 4 \\ 1 & \text{if } x \geq 4 \end{cases}$$



## 2.2 – Expectation of a Discrete R.V.

The **expectation** of a discrete random variable  $X$  is defined as

$$\mathbf{E}[X] = \sum_x x \cdot P(X = x) = \sum_x x f(x),$$

where the sum extends over all values of  $x$  taken by  $X$ .

The definition can be extended to a general function of  $X$ :

$$\mathbf{E}[u(X)] = \sum_x u(x)P(X = x) = \sum_x u(x)f(x).$$

As an important example,  $\mathbf{E}[X^2] = \sum_x x^2 P(X = x) = \sum_x x^2 f(x)$ .

## Examples:

1. What is the expectation on the roll  $Z$  of fair 6–sided die?

**Solution:** 
$$E[Z] = \sum_{z=1}^6 z \cdot P(Z = z) = \frac{1}{6} \sum_{z=1}^6 z = \frac{1}{6} \cdot \frac{6(7)}{2} = 3.5.$$

2. For each 1\$ bet in a gambling game, a player can win 3\$ with probability  $\frac{1}{3}$  and lose 1\$ with probability  $\frac{2}{3}$ . Let  $X$  be the net gain/loss from the game. Find the expected value of the game.

**Solution:**  $X$  can take on the value 2\$ (for a win) and  $-2$ \$ for a loss (outcome – bet). The expected value of  $X$  is thus

$$E[X] = 2 \cdot \frac{1}{3} + (-2) \cdot \frac{2}{3} = -\frac{2}{3}.$$

3. If  $Z$  is the number showing on a roll of a fair 6–sided die, find  $E[Z^2]$  and  $E[(Z - 3.5)^2]$ .

**Solution:**

$$E[Z^2] = \sum_z z^2 P(Z = z) = \frac{1}{6} \sum_{z=1}^6 z^2 = \frac{1}{6}(1^2 + \cdots + 6^2) = \frac{91}{6}$$

$$\begin{aligned} E[(Z - 3.5)^2] &= \sum_{z=1}^6 (z - 3.5)^2 P(Z = z) = \frac{1}{6} \sum_{z=1}^6 (z - 3.5)^2 \\ &= \frac{(1 - 3.5)^2 + \cdots + (6 - 3.5)^2}{6} = \frac{35}{12} \end{aligned}$$



## Mean and Variance of a Discrete R.V.

We can interpret the expectation as the average or the **mean** of  $X$ , which we often denote by  $\mu = \mu_X$ .

For instance, in the example of the fair die,  $\mu_Z = E[Z] = 3.5$ .

Note that in the final example, we could have written

$$E[(Z - 3.5)^2] = E[(Z - E[Z])^2].$$

This is an important quantity associated to a random variable  $X$ , its **variance**  $\text{Var}[X]$ .

The variance of a discrete random variable  $X$  is the **expected squared difference from the mean**:

$$\begin{aligned}\text{Var}(X) &= \text{E}[(X - \mu_X)^2] = \sum_x (x - \mu_X)^2 P(X = x) \\ &= \sum_x (x^2 - 2x\mu_X + \mu_X^2) f(x) \\ &= \sum_x x^2 f(x) - 2\mu_X \sum_x x f(x) + \mu_X^2 \sum_x f(x) \\ &= \text{E}[X^2] - 2\mu_X \mu_X + \mu_X^2 \cdot 1 = \text{E}[X^2] - \mu_X^2.\end{aligned}$$

This is also sometimes written as  $\text{Var}[X] = \text{E}[X^2] - \text{E}^2[X]$ .

## Standard Deviation

The **standard deviation** of a discrete random variable  $X$  is defined directly from the variance:

$$\text{SD}[X] = \sqrt{\text{Var}[X]}.$$

The mean gives some idea as to where the **bulk** of a distribution is located  
 $\Rightarrow$  measure of **centrality** (more on this later).

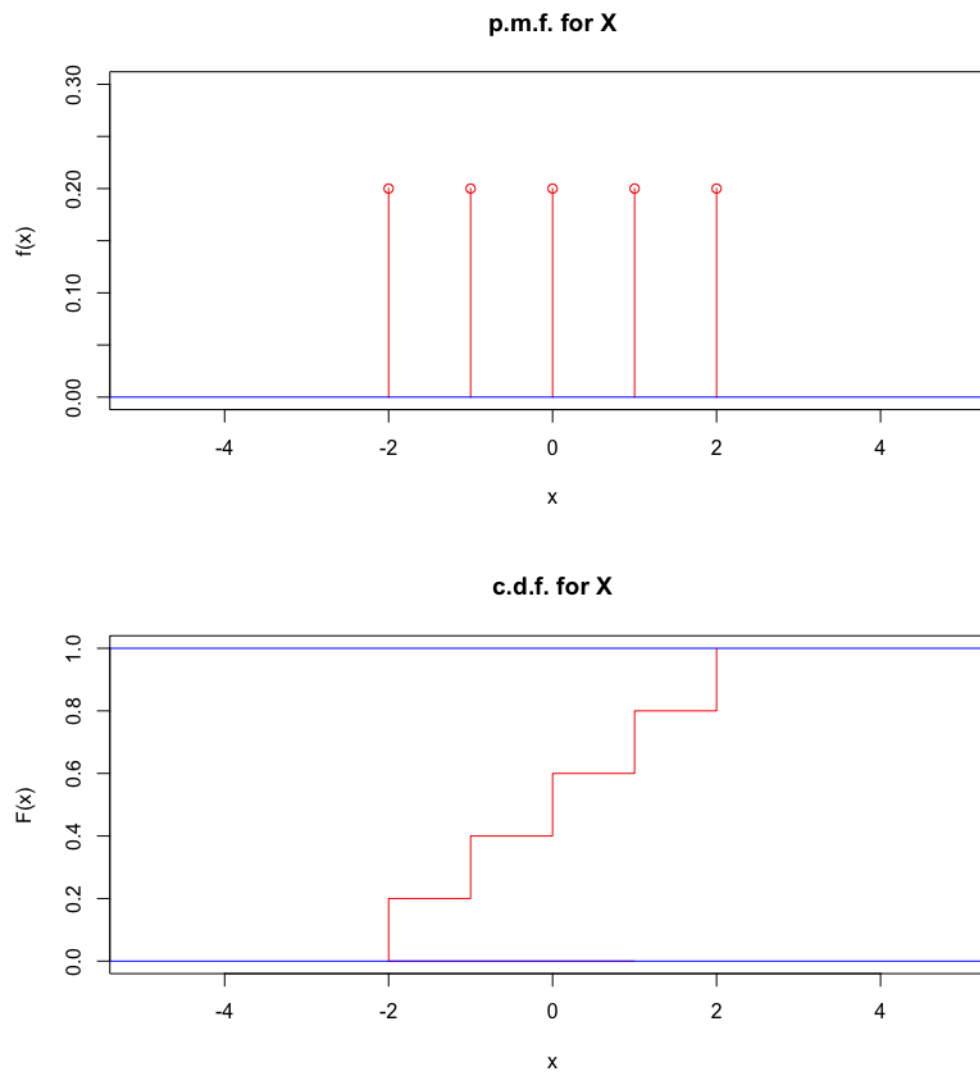
The variance and standard deviation provide information about the **spread**; distributions with higher variance/SD are **more spread out about the average**.

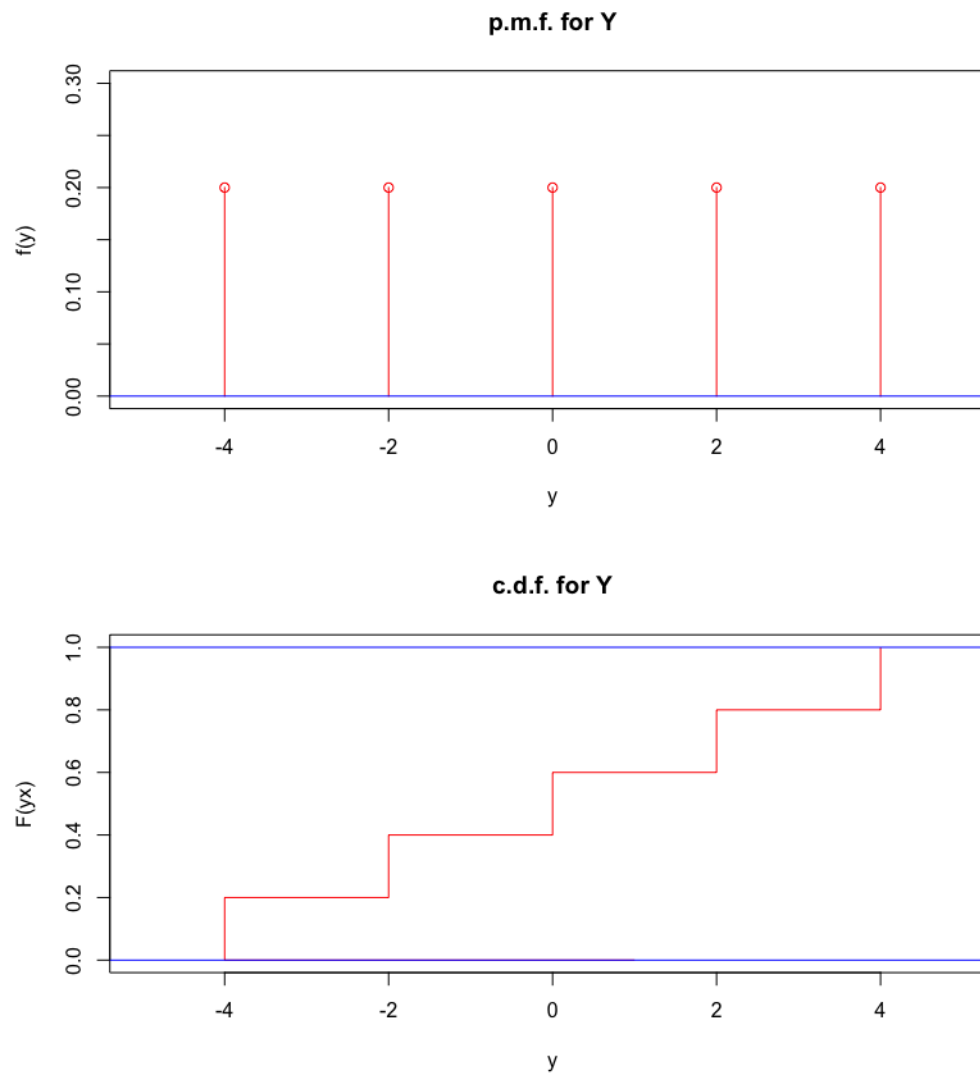
**Examples:** let  $X$  and  $Y$  be random variables with the following p.d.f.:

$x$	$P(X = x)$	$y$	$P(Y = y)$
-2	1/5	-4	1/5
-1	1/5	-2	1/5
0	1/5	0	1/5
1	1/5	2	1/5
2	1/5	4	1/5

Compute the expected values and compare the variances.

**Solution:** We have  $E[X] = E[Y] = 0$  and  $2 = \text{Var}[X] < \text{Var}[Y] = 8$ , meaning that we would expect both distributions to be centered at 0, but  $Y$  should be more spread-out than  $X$ .





## Properties of Expectations

For all  $a \in \mathbb{R}$ :

- $E[aX] = aE[X]$ ;
- $E[X + a] = E[X] + a$ ;
- $E[X + Y] = E[X] + E[Y]$ ;
- In general,  $E[XY] \neq E[X]E[Y]$ ;
- $\text{Var}[aX] = a^2\text{Var}[X]$ ,  $\text{SD}[aX] = |a|\text{SD}[X]$ ;
- $\text{Var}[X + a] = \text{Var}[X]$ ,  $\text{SD}[X + a] = \text{SD}[X]$ .

## 2.3 – Binomial Distributions

Recall the number of unordered samples of size  $r$  from a set of size  $n$  is:

$${}_nC_r = \binom{n}{r} = \frac{n!}{(n-r)!r!}.$$

- $2! \times 4! = (1 \times 2) \times (1 \times 2 \times 3 \times 4) = 48$ , but  $(2 \times 4)! = 8! = 40320$ .
- $\binom{5}{1} = \frac{5!}{1! \times 4!} = \frac{1 \times 2 \times 3 \times 4 \times 5}{1 \times (1 \times 2 \times 3 \times 4)} = \frac{5}{1} = 5$ ; In general:  $\binom{n}{1} = n$ . Also  $\binom{n}{0} = 1$
- $\binom{6}{2} = \frac{6!}{2! \times 4!} = \frac{4! \times 5 \times 6}{2! \times 4!} = \frac{5 \times 6}{2} = 15$
- $\binom{27}{22} = \frac{27!}{22! \times 5!} = \frac{22! \times 23 \times 24 \times 25 \times 26 \times 27}{5! \times 22!} = \frac{23 \times 24 \times 25 \times 26 \times 27}{120}$



## Binomial Experiments

A **Bernoulli trial** is a random experiment with two possible outcomes, “success” and “failure”. Let  $p$  denote the probability of a success.

A **binomial experiment** consists of  $n$  repeated independent Bernoulli trials, each with the same probability of success,  $p$ .

### Examples:

- female/male births;
- satisfactory/defective items on a production line;
- sampling with replacement with two types of item, etc.

## Probability Mass Function

In a binomial experiment of  $n$  independent events, each with probability of success  $p$ , the number of successes  $X$  is a discrete random variable that follows a **binomial distribution** with parameters  $(n, p)$ :

$$f(x) = P(X = x) = \binom{n}{x} p^x (1 - p)^{n-x}, \quad \text{for } x = 0, 1, 2, \dots, n.$$

This is often abbreviated to “ $X \sim \mathcal{B}(n, p)$ ”.

If  $X \sim B(1, p)$  then  $P(X = 0) = 1 - p$ ,  $P(X = 1) = p$ , and so

$$E[X] = (1 - p) \cdot 0 + p \cdot 1 = p.$$

## Expectation and Variance for $\mathcal{B}(n, p)$

If  $X \sim \mathcal{B}(n, p)$ , then  $P(X = x)$  is as on the previous slide and it can be shown that

$$E[X] = \sum_{x=0}^n xP(X = x) = np,$$

and

$$\text{Var}[X] = E[(X - np)^2] = \sum_{x=0}^n (x - np)^2 P(X = x) = np(1 - p).$$

Later we will see an easier way to derive these by interpreting  $X$  as a sum of other discrete random variables.

**Examples:**

1. Suppose that each water sample taken in some well-defined region has a 10% probability of being polluted.

If 12 samples are selected independently, then it is reasonable to model the number  $X$  of polluted samples as  $\mathcal{B}(12, 0.1)$ . Find

- a)  $E[X]$  and  $\text{Var}[X]$ ;
- b)  $P(X = 3)$ ;
- c)  $P(X \leq 3)$ .

**Solution:**

- a) If  $X \sim \mathcal{B}(n, p)$  then  $E[X] = np$  and  $\text{Var}[X] = np(1 - p)$ , so

$$E[X] = 12 \times 0.1 = 1.2; \quad \text{Var}[X] = 12 \times 0.1 \times 0.9 = 1.08.$$

b) By definition,  $P(X = 3) = \binom{12}{3}(0.1)^3(0.9)^9 \approx 0.0852$ .

c) By definition,

$$P(X \leq 3) = P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3) = \dots$$

However, for  $X \sim \mathcal{B}(12, 0.1)$ ,  $P(X \leq 3)$  is tabulated on p.430 of text (Table A.1), and is  $\approx 0.9744$ .

The table can also be used to compute

$$P(X = 3) = P(X \leq 3) - P(X \leq 2) = 0.9744 - 0.8891 \approx 0.0853.$$

Note the rounding error.

12	0	0.2824	0.0687	0.0138	0.0022	0.0002	0.0000			
	1	0.6590	0.2749	0.0850	0.0196	0.0032	0.0003	0.0000		
	2	0.8891	0.5583	0.2528	0.0834	0.0193	0.0028	0.0002		
	3	0.9744	0.7946	0.4925	0.2253	0.0730	0.0153	0.0017	0.0000	
	4	0.9957	0.9274	0.7237	0.4382	0.1938	0.0573	0.0095	0.0006	
	5	0.9995	0.9806	0.8822	0.6652	0.3872	0.1582	0.0386	0.0039	0.0000
	6	0.9999	0.9961	0.9614	0.8418	0.6128	0.3348	0.1178	0.0194	0.0005
	7	1.0000	0.9994	0.9905	0.9427	0.8062	0.5618	0.2763	0.0726	0.0043
	8		0.9999	0.9983	0.9847	0.9270	0.7747	0.5075	0.2054	0.0256
	9		1.0000	0.9998	0.9972	0.9807	0.9166	0.7472	0.4417	0.1109
	10			1.0000	0.9997	0.9968	0.9804	0.9150	0.7251	0.3410
	11				1.0000	0.9998	0.9978	0.9862	0.9313	0.7176
	12					1.0000	1.0000	1.0000	1.0000	1.0000

Table of c.d.f.  $F(x) = P(X \leq x)$  for  $X \sim \mathcal{B}(12, p)$ ,  $p = 0.1, \dots, 0.9$ .

2. An airline sells 101 tickets for a flight with 100 seats. Each passenger with a ticket is known to have a  $p = 0.97$  probability of showing up for their flight. What is the probability of 101 passengers showing up (and the airline being caught overbooking)? Make appropriate assumptions. What if the airline sells 125 tickets?

**Solution:** let  $X$  be the number of passengers that show up. We want to compute  $P(X > 100)$ .

If all passengers show up independently of one another (no families or late bus?), we can model  $X \sim \mathcal{B}(101, 0.97)$  and

$$P(X > 100) = P(X = 101) = \binom{101}{101} (0.97)^{101} (0.03)^0 \approx 0.046$$

If the airline sells  $n = 125$  tickets, we can model  $X \sim \mathcal{B}(125, 0.97)$  and

$$P(X > 100) = 1 - P(X \leq 100) = 1 - \sum_{x=0}^{100} \binom{125}{x} (0.97)^x (0.03)^{125-x}.$$

This is harder to compute directly, but is very nearly 1 (try it in R).



## 2.4 – Geometric Distributions

Consider a sequence of Bernoulli trials, with probability  $p$  of success at each step.

Let the **geometric** random variable  $X$  denote the number of steps before the first success occurs. The probability distribution is given by

$$f(x) = P(X = x) = (1 - p)^{x-1}p, \quad x = 1, \dots$$

We will write  $X \sim \text{Geo}(p)$ . For this random variable, we have

$$E[X] = \frac{1}{p} \quad \text{and} \quad \text{Var}[X] = \frac{1 - p}{p^2}.$$

## Examples:

- A fair 6–sided die is thrown until it shows a 6. What is the probability that 5 throws are required?

**Solution:** If 5 throws are required, we have to compute  $P(X = 5)$ , where  $X$  is geometric  $\text{Geo}(1/6)$ :

$$P(X = 5) = (1 - p)^{5-1}p = (5/6)^4(1/6) \approx 0.0804.$$

- In the example above, how many throws would you expect to need?

**Solution:**  $E[X] = \frac{1}{1/6} = 6.$

## 2.5 – Negative Binomial Distributions

Consider a sequence of Bernoulli trials, with probability  $p$  of success at each step.

Let the **negative binomial** random variable  $X$  denote the number of steps before the  $r$ th success occurs. The probability distribution is given by

$$f(x) = P(X = x) = \binom{x-1}{r-1} (1-p)^{x-r} p^r, \quad x = r, \dots$$

We will write  $X \sim \text{NegBin}(p, r)$ . For this random variable, we have

$$E[X] = \frac{r}{p} \quad \text{and} \quad \text{Var}[X] = \frac{r(1-p)}{p^2}.$$

## Examples:

- A fair 6–sided die is thrown until it three 6's are rolled. What is the probability that 5 throws are required?

**Solution:** If 5 throws are required, we have to compute  $P(X = 5)$ , where  $X$  is geometric  $\text{NegBin}(1/6, 3)$ :

$$P(X = 5) = \binom{5-1}{3-1} (1-p)^{5-3} p^3 = \binom{4}{2} (5/6)^2 (1/6)^3 \approx 0.0193.$$

- In the example above, how many throws would you expect to need?

**Solution:**  $E[X] = \frac{3}{1/6} = 18.$

## 2.6 – Poisson Distributions

We count the number of “changes” that occur in a continuous interval of time or space (such as # of defects on a production line over a 1 hr period, # of customers that arrive at a teller over a 15 min interval, etc.).

We have a **Poisson process** with rate  $\lambda$ , denoted by  $\mathcal{P}(\lambda)$ , if:

- a) the number of changes occurring in non-overlapping intervals are independent;
- b) the probability of exactly one change in a short interval of length  $h$  is approximately  $\lambda h$ , and
- c) The probability of 2+ changes in a sufficiently short interval is essentially 0.

Assume that an experiment satisfies the above properties. Let  $X$  be the number of changes in a **unit interval** (this could be 1 day, or 15 minutes, or 10 years, etc.).

What is  $P(X = x)$ , for  $x = 0, 1, \dots$ ?

Partition the unit interval into  $n$  disjoint sub-intervals of length  $1/n$ .

1. By condition b), the probability of one change occurring in one of the sub-intervals is approximately  $\lambda/n$ .
2. By condition c), the probability of 2+ changes is  $\approx 0$ .
3. By condition a), we have a sequence of  $n$  Bernoulli trials with probability  $p = \lambda/n$ .

Therefore,

$$\begin{aligned}
 f(x) = P(X = x) &\approx \frac{n!}{x!(n-x)!} \left(\frac{\lambda}{n}\right)^x \left(1 - \frac{\lambda}{n}\right)^{n-x} \\
 &= \frac{\lambda^x}{x!} \cdot \underbrace{\frac{n!}{(n-x)!} \cdot \frac{1}{n^x}}_{\text{term 1}} \cdot \underbrace{\left(1 - \frac{\lambda}{n}\right)^n}_{\text{term 2}} \cdot \underbrace{\left(1 - \frac{\lambda}{n}\right)^{-x}}_{\text{term 3}}.
 \end{aligned}$$

Letting  $n \rightarrow \infty$ , we get

$$\begin{aligned}
 P(X = x) &= \lim_{n \rightarrow \infty} \frac{\lambda^x}{x!} \cdot \underbrace{\frac{n!}{(n-x)!} \cdot \frac{1}{n^x}}_{\text{term 1}} \cdot \underbrace{\left(1 - \frac{\lambda}{n}\right)^n}_{\text{term 2}} \cdot \underbrace{\left(1 - \frac{\lambda}{n}\right)^{-x}}_{\text{term 3}} \\
 &= \frac{\lambda^x}{x!} \cdot 1 \cdot \exp(-\lambda) \cdot 1 = \frac{\lambda^x e^{-\lambda}}{x!}, \quad x = 0, 1, \dots
 \end{aligned}$$

Let  $X \sim \mathcal{P}(\lambda)$ . Then it can be shown that

$$E[X] = \lambda \quad \text{and} \quad \text{Var}[X] = \lambda,$$

that is, the mean and the variance of a Poisson random variable are identical!

### Examples:

1. A traffic flow is typically modeled by a Poisson distribution. It is known that the traffic flowing through an intersection is 6 cars/minute, on average. What is the probability of no cars entering the intersection in a 30 second period?

**Solution:** 6 cars/min = 3 cars/30 sec. Thus  $\lambda = 3$ , and we need to compute

$$P(X = 0) = \frac{3^0 e^{-3}}{0!} = \frac{e^{-3}}{1} \approx 0.0498.$$



## 2. A hospital needs to schedule night shifts in the maternity ward.

It is known that there are 3000 deliveries per year; if these happened randomly round the clock (is this a reasonable assumption?), we would expect 1000 deliveries between the hours of midnight and 8.00 a.m., a time when much of the staff is off-duty.

It is thus important to ensure that the night shift is sufficiently staffed to allow the maternity ward to cope with the workload on any particular night, or at least, on a high proportion of nights.

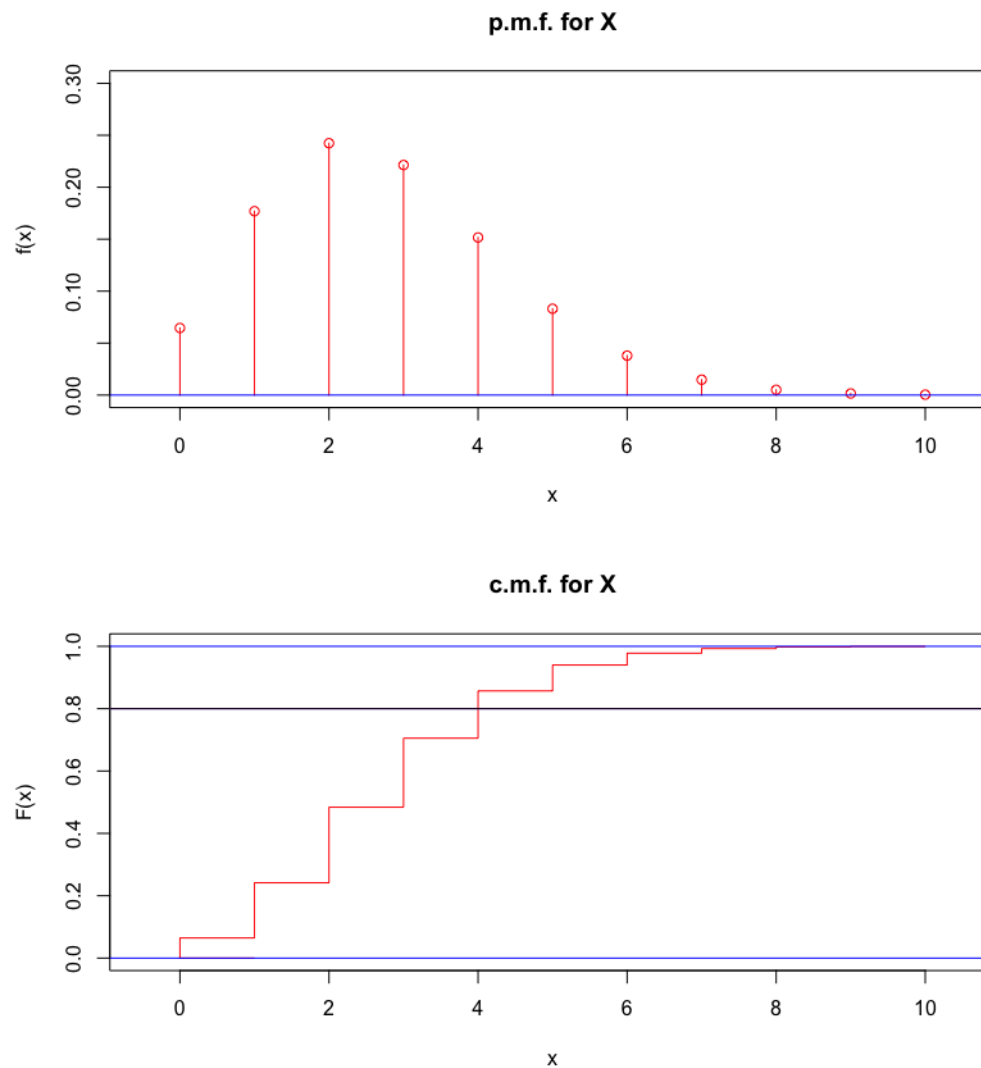
The average number of deliveries per night is  $\lambda = 1000/365.25 \approx 2.74$ . If the daily number  $X$  of night deliveries follows a Poisson process  $\mathcal{P}(\lambda)$ , we can compute the probability of delivering  $x = 0, 1, 2, \dots$  babies on each night.

Some of the probabilities are:

$P(X = x)$	$\lambda^x \cdot \exp(-\lambda)/x!$
$P(X = 0)$	$2.74^0 \cdot \exp(-2.74)/0! = 0.065$
$P(X = 1)$	$2.74^1 \cdot \exp(-2.74)/1! = 0.177$
$P(X = 2)$	$2.74^2 \cdot \exp(-2.74)/2! = 0.242$
...	...

3. If the maternity ward wants to prepare for the greatest possible traffic on 80% of the nights, how many deliveries should be expected?

**Solution:** we seek an  $x$  for which  $P(X \leq x - 1) \leq 0.80 \leq P(X \leq x)$ : since  $\text{ppois}(3, 2.74) = .705$  and  $\text{ppois}(4, 2.74) = .857$ , if they prepare for 4 deliveries a night, they will be ready for the worst on at least 80% of the nights (closer to 85.7%, actually). Note that this is different than asking how many deliveries are expected nightly (namely,  $E[X] = 2.74$ ).



4. On how many nights in the year would 5 or more deliveries be expected?

**Solution:** we need to evaluate

$$\begin{aligned} 365.25 \cdot P(X \geq 5) &= 365.25(1 - P(X \leq 4)) \\ &= 365.25 * (1 - \text{ppois}(4, 2.74)) \approx 52.27. \end{aligned}$$

5. Over the course of one year, what is the greatest number of deliveries expected on any night?

**Solution:** look for largest value of  $x$  for which  $365.25 \cdot P(X = x) \geq 1$ .

```
> nights=c() # initializing vector  
> for(j in 0:10){nights[j+1]=365.25*dpois(j,2.74)}; # p.m.f.  
> max(which(nights>1))-1 # identify largest index  $\Rightarrow x = 8$ .
```

## Appendix – Summary

$X$	Description	$P(X = x)$	Domain	$E[X]$	$\text{Var}[X]$
<b>Uniform (Discrete)</b>	Equally likely outcomes	$\frac{1}{b-a+1}$	$a, \dots, b$	$\frac{a+b}{2}$	$\frac{(b-a+2)(b-a)}{12}$
<b>Binomial</b>	Number of successes in $n$ trials	$\binom{n}{x} p^x (1-p)^{n-x}$	$0, \dots, n$	$np$	$np(1-p)$
<b>Poisson</b>	Number of arrivals in a fixed period of time	$\frac{\lambda^x \exp(-\lambda)}{x!}$	$0, 1, \dots$	$\lambda$	$\lambda$

## Summary

$X$	Description	$P(X = x)$	Domain	$E[X]$	$\text{Var}[X]$
<b>Geometric</b>	Number of trials until $1^{st}$ success	$(1 - p)^{x-1}p$	$1, 2, \dots$	$\frac{1}{p}$	$\frac{1-p}{p^2}$
<b>Negative Binomial</b>	Number of trials until $k^{th}$ successes	$\binom{x-1}{k-1} (1 - p)^{x-k} p^k$	$k, k+1, \dots$	$\frac{k}{p}$	$\frac{k(1-p)}{p^2}$