IES302 2011/1 Part I.4 Dr.Prapun

7 Random variables

In performing a chance experiment, one is often not interested in the particular outcome that occurs but in a specific numerical value associated with that outcome. In fact, for most applications, measurements and observations are expressed as numerical quantities.

7.1. The advantage of working with numerical quantities is that we can perform mathematical operations on them.

In order to exploit the axioms and properties of probability that we studied earlier, we technically define random variables as functions on an underlying sample space.

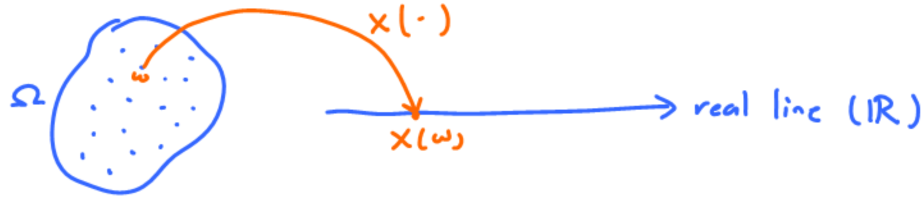
Fortunately, once some basic results are derived, we can think of random variables in the traditional manner, and not worry about, or even mention the underlying sample space.

Any function that assigns a real number to each outcome in the sample space of the experiment is called a random variable. Intuitively, a random variable is a variable that takes on its values by chance.

Definition 7.2. A real-valued function $X(\omega)$ defined for all points ω in a sample space Ω is called a ***random variable*** (r.v. or RV) [12].

[12] The term "random variable" is a misnomer. Technically, if you look at the definition carefully, a random variable is a deterministic function; that is, it is not random and it is not a variable. [Toby Berger][23, p 254]

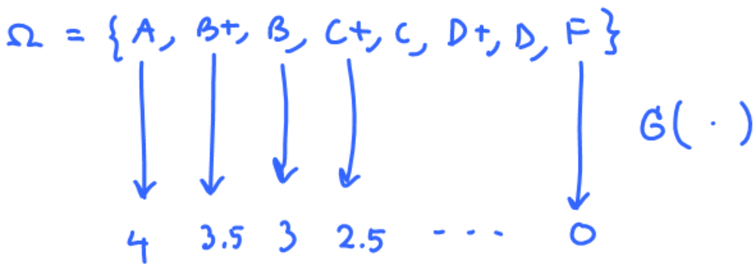
- So, a random variable is a rule that assigns a numerical value to each possible outcome of a chance experiment.



- Random variables are important because they provide a compact way of referring to events via their numerical attributes.

- The convention is to use capital letters such as X, Y, Z to denote random variables.

Example 7.3. Take this course and observe your grades.



7.4. Remark: Unlike probability models defined on arbitrary sample space, random variables allow us to compute averages.

In the mathematics of probability, averages are called expectations or expected values.

- As a function, it is simply a rule that maps points/outcomes ω in Ω to real numbers.
- It is also a deterministic function; nothing is random about the mapping/assignment. The randomness in the observed values is due to the underlying randomness of the argument of the function X, namely the experiment outcomes ω.
- In other words, the randomness in the observed value of X is induced by the underlying random experiment, and hence we should be able to compute the probabilities of the observed values in terms of the probabilities of the underlying outcomes.

Example 7.5. Roll a fair dice: $\Omega = \{1, 2, 3, 4, 5, 6\}$.

$\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6\} = \{\omega_i : \omega_i = i, 1 \le i \le 6\}$

1 2 3 4 5 6

$X(\omega) = \omega$

$Y(\omega) = (\omega - 3)^2$

$Z(\omega) = \sqrt{Y(\omega)}$

$U(\omega) = \begin{cases} 1, & \omega \geq 3 \\ 0, & \omega < 3 \end{cases}$

$V(\omega) = \sqrt{Y^2(\omega) + Z^2(\omega)}$

Note that more than one r.v. can be defined on a sample space.

Example 7.6. Counting the number (N) of heads in a sequence of three (fair) coin tosses.

$$\Omega = \{TTT, TTH, THT, THH, HTT, HTH, HHT, HHH\}$$

$P[N=0] = \frac{1}{8}$

$P[N=1] = \frac{3}{8}$

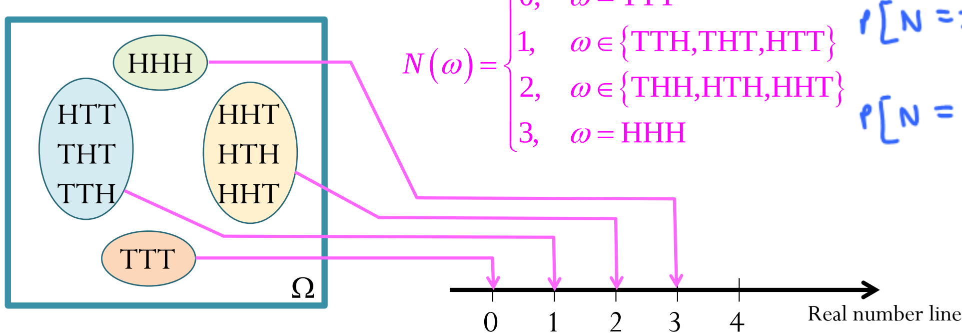
$P[N=2] = \frac{3}{8}$

$P[N=3] = \frac{1}{8}$

$\{0,1,2,3\}$ is a support of the r.v. N.

$\{0,1,2,3,\sqrt{5}\}$ is another support of the r.v. N.

$$N(\omega) = \begin{cases} 0, & \omega = TTT \\ 1, & \omega \in \{TTH, THT, HTT\} \\ 2, & \omega \in \{THH, HTH, HHT\} \\ 3, & \omega = HHH \end{cases}$$



Example 7.7. If X is the sum of the dots when rolling one fair dice twice, the random variable X assigns the numerical value $i+j$ to the outcome (i, j) of the chance experiment.

Example 7.8. Continue from Example 7.5,

(a) What is the probability that $X = 4$?

$P[X=4] = \frac{1}{6}$

$X = 4 \Leftrightarrow X(\omega) = 4 \Leftrightarrow \omega = 4$

$[X=4] = \{\omega \in \Omega : X(\omega) = 4\}$

$= \{4\}$

$P([X=4])$

(b) What is the probability that $Y = 4$?

$[Y=4] = \{\omega : Y(\omega) = 4\}$

$= \{1, 5\}$

$P[Y=4] = P(\{1,5\}) = P(\{1\}) + P(\{5\}) = \frac{1}{6} + \frac{1}{6} = \frac{1}{3}$

$Y = 4 \Leftrightarrow Y(\omega) = 4 \Leftrightarrow (\omega - 3)^2 = 4 \Leftrightarrow \omega = 1, 5$

$P([Y=4])$

Definition 7.9. Shorthand Notation:

- $[X \in B] = \{\omega \in \Omega : X(\omega) \in B\}$
- $[a \le X < b] = [X \in [a,b)] = \{\omega \in \Omega : a \le X(\omega) < b\}$
- $[X > a] = \{\omega \in \Omega : X(\omega) > a\}$
- $[X = x] = \{\omega \in \Omega : X(\omega) = x\}$

To evaluate the probability, $P([a \le X < b])$

- ◦ We usually use the corresponding lowercase letter to denote
  (a) a possible value (realization) of the random variable
  (b) the value that the random variable takes on
  (c) the running values for the random variable

All of the above items are sets of outcomes. They are all events!

To avoid double use of brackets (round brackets over square brackets), we write $P[X \in B]$ when we means $P([X \in B])$. Hence,

$$P[X \in B] = P([X \in B]) = P(\{\omega \in \Omega : X(\omega) \in B\}).$$

Similarly,

$$P[X < x] = P([X < x]) = P(\{\omega \in \Omega : X(\omega) < x\}).$$

Example 7.10. Continue from Examples 7.5 and 7.8,

(a) $[X = 4] = \{\omega : X(\omega) = 4\}$

(b) $[Y = 4] = \{\omega : Y(\omega) = 4\} = \left\{\omega : (\omega - 3)^2 = 4\right\}$

Example 7.11. In Example 7.6, if the coins is fair, then

$$P[N < 2] = P([N < 2]) = P(\{\omega : N(\omega) < 2\})$$
$$= P(\{TTT, TTH, THT, HTT\})$$
$$= \frac{1}{8} + \frac{1}{8} + \frac{1}{8} + \frac{1}{8} = \frac{1}{2}$$

7.12. At a certain point in most probability courses, the sample space is rarely mentioned anymore and we work directly with random variables. The sample space often "disappears" along with the "(ω)" of $X(\omega)$ but they are really there in the background.

Definition 7.13. A set S is called a ***support*** of a random variable X if $P[X \in S] = 1$.

- To emphasize that S is a support of a particular variable X, we denote a support of X by S_X.
- Recall that a support of a probability measure P is any set $A \subset \Omega$ such that $P(A) = 1$.

7.14. There are three types of of random variables. The first type, which will be discussed in Section 8, is called ***discrete random variable***. To tell whether a random variable is discrete, one simple way is to consider the possible values of the random variable. If it is limited to only a finite or countably infinite number of possibilities, then it is discrete. We will later discuss ***continuous random variables*** whose possible values can be anywhere in some intervals of real numbers.

7.15. In some cases, the random variable X is actually discrete but, because the range of possible values is so large, it might be more convenient to analyze X as a continuous random variable. For example, suppose that current measurements are read from a digital instrument that displays the current to the nearest one-hundredth of a milliampere. Because the possible measurements are limited, the random variable is discrete. However, it might be a more convenient, simple approximation to assume that the current measurements are values of a continuous random variable.

8 Discrete Random Variables

Definition 8.1. A random variable X is said to be a ***discrete random variable*** if there exists countable distinct real numbers x_k such that

$x_1, \dots, x_n$

$$\sum_k P[X = x_k] = 1. \tag{10}$$

$x_1, x_2, x_3, \dots$

Example 8.2. Voice Lines: A voice communication system for a business contains 48 external lines. At a particular time, the system is observed, and some of the lines are being used. Let the random variable X denote the number of lines in use. Then, X can assume any of the integer values 0 through 48. [11, Ex 3-1]

Definition 8.3. An ***integer-valued random variable*** is a discrete random variable whose x_k in (10) above are all integers.

8.4. The ***probability distribution*** of a random variable X is a description of the probabilities associated with X. For a discrete random variable, the distribution is often characterized by just a list of the possible values $(x_1, x_2, x_3, \dots)$ along with the probability of each $(P[X = x_1], P[X = x_2], P[X = x_3], \dots,$ respectively).

In some cases, it is convenient to express the probability in terms of a formula.

8.1 PMF: Probability Mass Function

Definition 8.5. When X is a discrete random variable satisfying (10), we define its ***probability mass function*** (pmf) by[13]

$$p(x) = p_X(x) = P[X = x].$$

subscript indicates the name of the RV.

- Sometimes, when we only deal with one random variable or when it is clear which random variable the pmf is associated with, we write $p(x)$ or p_x instead of $p_X(x)$.

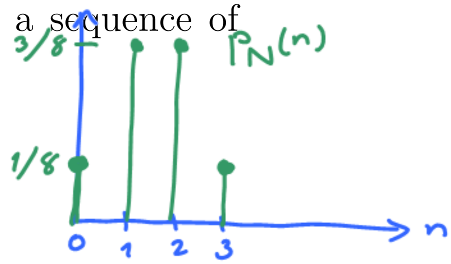
[13] Many references (including [11] and MATLAB) use $f_X(x)$ for pmf instead of $p_X(x)$. We will *NOT* use $f_X(x)$ for pmf. Later, we will define $f_X(x)$ as a probability density function which will be used primarily for another type of random variable (continuous r.v.)

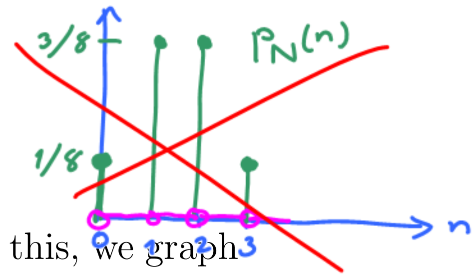
- The argument (x) of a pmf ranges over all real numbers. Hence, the pmf is defined for x that is not among the x_k in (10). In such case, the pmf is simply 0. This is usually expressed as "$p_X(x) = 0$, otherwise" when we specify a pmf for a particular r.v.

Example 8.6. Let ~~X~~ N be the number of heads in a sequence of three coin tosses.

Four possible values: 0, 1, 2, 3

$$\left.\begin{array}{l} P[N=0] = \frac{1}{8} \\ P[N=1] = \frac{3}{8} \\ P[N=2] = \frac{3}{8} \\ P[N=3] = \frac{1}{8} \end{array}\right\} \quad \text{pmf} \quad p_N(n) = \begin{cases} 1/8, & n = 0, 3 \\ 3/8, & n = 1, 2 \\ 0, & \text{otherwise} \end{cases}$$





8.7. We can use ***stem plot*** to visualize p_X. To do this, we graph a pmf by marking on the horizontal axis each value with nonzero probability and drawing a vertical bar with length proportional to the probability.

8.8. Any pmf $p(\cdot)$ satisfies two properties:

(a) $p(\cdot) \geq 0$

$\Sigma = 1$

(b) there exists numbers $x_1, x_2, x_3, \ldots$ such that $\sum_k p(x_k) = 1$ and $p(x) = 0$ for other x.

When you are asked to verify that a function is a pmf, check these two properties.

8.9. Finding probability from pmf: for any subset B of $\mathbb{R}$, we can find

$$P[X > 1] = \sum_{x : x > 1} p_X(x)$$

$$P[X \in B] = \sum_{x_k \in B} P[X = x_k] = \sum_{x_k \in B} p_X(x_k).$$

In particular, for integer-valued random variables,

$$P[X \in B] = \sum_{k \in B} P[X = k] = \sum_{k \in B} p_X(k).$$

Example 8.10. Suppose a random variable X has pmf

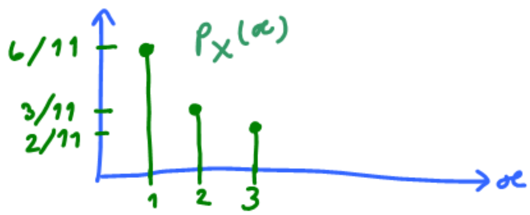
$$p_X(x) = \begin{cases} c/x, & x = 1, 2, 3, \\ 0, & \text{otherwise.} \end{cases} = \begin{cases} \frac{6}{11x}, & x = 1, 2, 3, \\ 0, & \text{otherwise} \end{cases}$$

(a) The value of the constant c is

To be a pmf, "$\sum = 1$" $\sum_x p_X(x) = p_X(1) + p_X(2) + p_X(3) = 1$

$$\frac{c}{1} + \frac{c}{2} + \frac{c}{3} = 1 \Rightarrow c = \frac{6}{11}$$

(b) Sketch of pmf



(c) $P[X = 1] = p_X(1) = \frac{6}{11}$

(d) $P[X \geq 2] = p_X(2) + p_X(3) = \frac{3}{11} + \frac{2}{11} = \frac{5}{11}$

$= 1 - P[X < 2] = 1 - p_X(1) = 1 - \frac{6}{11} = \frac{5}{11}$

(e) $P[X > 3] = 0$

8.11. Any function $p(\cdot)$ on $\mathbb{R}$ which satisfies

(a) $p(\cdot) \geq 0$, and

(b) there exists numbers $x_1, x_2, x_3, \ldots$ such that $\sum_k p(x_k) = 1$ and $p(x) = 0$ for other x

is a pmf of some discrete random variable.

8.2 CDF: Cumulative Distribution Function

Definition 8.12. The ***(cumulative) distribution function (cdf)*** of a random variable X is the function $F_X(x)$ defined by

$p_X(x) = P[X = x]$

$$F_X(x) = P[X \leq x].$$

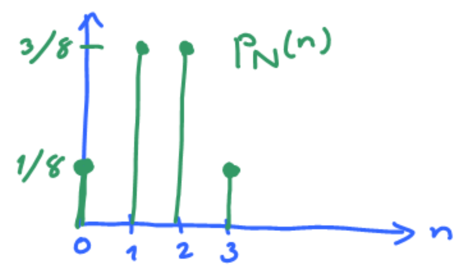
- The argument (x) of a cdf ranges over all real numbers.
- From its definition, we know that $0 \leq F_X \leq 1$.
- Think of it as a function that collects the "probability mass" from $-\infty$ up to the point x.

8.13. In general, for any discrete random variable with possible values $x_1, x_2, \ldots$, the cdf of X is given by

$$F_X(x) = P[X \leq x] = \sum_{x_k < x} p_X(x_i).$$

Example 8.14. Continue from Example 8.14 where N is defined as the number of heads in a sequence of three coin tosses. We have

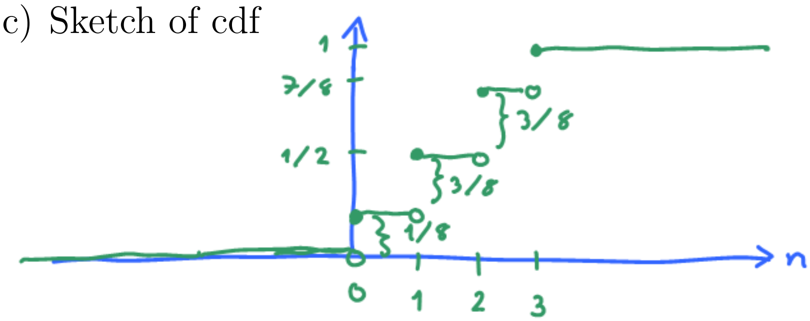
$$p_N(0) = p_N(3) = \frac{1}{8} \text{ and } p_N(1) = p_N(2) = \frac{3}{8}.$$



(a) $F_N(0) = P[N \leq 0] = \frac{1}{8}$

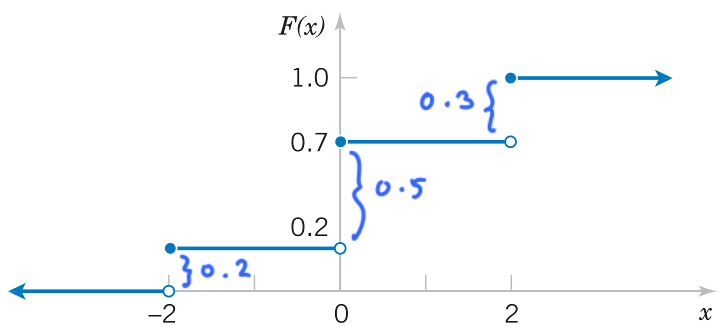
(b) $F_N(1.5) = P[N \leq 1.5] = p_N(0) + p_N(1) = \frac{1}{8} + \frac{3}{8} = \frac{4}{8} = \frac{1}{2} = 0.5.$

(c) Sketch of cdf



8.15. For any discrete r.v. X,
F_X is a right-continuous, ***staircase*** function of x with jumps at a countable set of points x_k. When you are given the cdf of a discrete r.v., you can derive its pmf by looking at the locations/sizes of the jumps. If a jump happens at $x = c$, then $P[X = c]$ is the same as the amount of jump at c. At the location x where there is no jump, $P[X = x] = 0$.

Example 8.16. Consider a discrete random variable X whose cdf $F_X(x)$ is shown below.



Determine the pmf $p_X(x)$.

$$p_X(x) = \begin{cases} 0.2, & x = -2, \\ 0.5, & x = 0, \\ 0.3, & x = 2, \\ 0, & \text{otherwise.} \end{cases}$$

8.17. Characterizing properties of cdf: These properties hold for all types of RVs.

CDF1 F_X is non-decreasing (monotone increasing)

if $a < b$, then $F_X(a) \leq F_X(b)$

CDF2 F_X is right continuous (continuous from the right)

CDF3 $\lim_{x \to -\infty} F_X(x) = 0$ and $\lim_{x \to \infty} F_X(x) = 1$.

useful and popular

8.3 Families of Discrete Random Variables

Many physical systems can be modeled by the same or similar random experiments and random variables. In this subsection, we present the analysis of several discrete random variables that frequently arise in applications.[14]

Remember the pmf of these RVs $\begin{cases} \text{uniform} \\ \text{Bernoulli} \\ \text{binomial} \\ \text{geometric} \\ \text{Poisson} \end{cases}$

Definition 8.18. X is ***uniformly distributed*** on a finite set S if

$$p_X(x) = P[X = x] = \frac{1}{|S|}, \quad \forall\, x \in S.$$

- We write $X \sim \mathcal{U}(S)$ or $X \sim \text{Uniform}(S)$.
- Read "X is uniform on S" or "X is a uniform random variable on set S".
- The pmf is usually referred to as the uniform discrete distribution.

Example 8.19. X is uniformly distributed on $1, 2, \ldots, n$ if

$$p_X(x) = \begin{cases} 1/n, & x = 1, 2, \ldots, n, \\ 0, & \text{otherwise.} \end{cases}$$

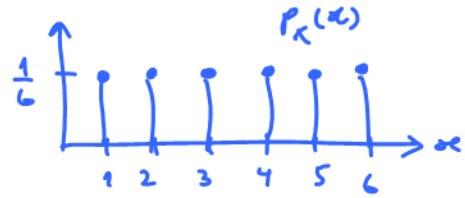
Example 8.20. Uniform pmf is used when the random variable can take finite number of "equally likely" or "totally random" values.

- Classical game of chance / classical probability drawing at random
- Fair gaming devices (well-balanced coins and dice, well shuffled decks of cards)

[14] As mention in 7.12, we often omit a discussion of the underlying sample space of the random experiment and directly describe the distribution of a particular random variable.

Example 8.21. Roll a fair dice. Let X be the outcome.

$$p_X(x) = \begin{cases} 1/6, & x = 1, 2, \dots, 6, \\ 0, & \text{otherwise.} \end{cases}$$



Definition 8.22. X is a ***Bernoulli*** random variable if

$$p_X(x) = \begin{cases} 1-p, & x = 0, \\ p, & x = 1, \\ 0, & \text{otherwise,} \end{cases} \qquad p \in (0,1)$$

- Write $X \sim \mathcal{B}(1,p)$ or $X \sim \text{Bernoulli}(p)$
- Some references denote $1-p$ by q for brevity.
- $p_0 = q = 1-p$, $p_1 = p$
- Bernoulli random variable is usually denoted by I. (Think about indicator function; it also has only two possible values, 0 and 1.)

Definition 8.23. X is a ***binary*** random variable if

$$p_X(x) = \begin{cases} 1-p, & x = a, \\ p, & x = b, \\ 0, & \text{otherwise,} \end{cases} \qquad p \in (0,1), \quad b > a.$$

- X takes only two values: a and b

Definition 8.24. X is a ***Binomial*** random variable with size $n \in \mathbb{N}$ and parameter $p \in (0,1)$ if

$$p_X(x) = \begin{cases} \binom{n}{x} p^x (1-p)^{n-x}, & x \in \{0, 1, 2, \dots, n\} \\ 0, & \text{otherwise} \end{cases} \tag{11}$$

Let's check that this is a pmf.

1) ≥ 0 ✓

2) "$\sum = 1$"

$$\sum_{x=0}^{n} \binom{n}{x} p^x (1-p)^{n-x} = (p + (1-p))^n = 1$$

$$\sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k} = (x+y)^n$$

- Write $X \sim \mathcal{B}(n,p)$ or $X \sim \text{Binomial}(n,p)$.
  - Observe that $\mathcal{B}(1,p)$ is Bernoulli with parameter p.

I_1 I_2 I_3 ... I_n $X = I_1 + I_2 + \cdots + I_n$

"Binomial RV is a sum of independent Bernoulli RVs"

n times

Bernoulli RV's with success probability p

- Use `binopdf(x,n,p)` in MATLAB. $P[X=k] = \binom{n}{k} p^k (1-p)^{n-k}$
- Interpretation: X is the number of successes in n independent Bernoulli trials. p = probability of success for 1 trial.

Example 8.25. Daily Airlines flies from Amsterdam to London every day. The price of a ticket for this extremely popular flight route is \$75. The aircraft has a passenger capacity of 150. The airline management has made it a policy to sell 160 tickets for this flight in order to protect themselves against no-show passengers. Experience has shown that the probability of a passenger being a no-show is equal to 0.1. The booked passengers act independently of each other. Given this overbooking strategy, what is the probability that some passengers will have to be bumped from the flight?

Solution: This problem can be treated as 160 independent trials of a Bernoulli experiment with a success rate of $p = 9/10$, where a passenger who shows up for the flight is counted as a success. Use the random variable X to denote number of passengers that show up for a given flight. The random variable X is binomial distributed with the parameters $n = 160$ and $p = 9/10$. The probability in question is given by

$$P[X > 150] = 1 - P[X \le 150] = 1 - F_X(150).$$

In `MATLAB`, we can enter `1-binocdf(150,160,9/10)` to get 0.0359. Thus, the probability that some passengers will be bumped from any given flight is roughly 3.6%. [19, Ex 4.1]

Definition 8.26. A geometric random variable X is defined by the fact that for some $\beta \in (0,1)$,

$$p_X(k+1) = \beta \times p_X(k)$$

for all $k \in S$ where S can be either $\mathbb{N}$ or $\mathbb{N} \cup \{0\}$.

There are two types of geometric RVs:

a) • When its support is $\mathbb{N} = \{1, 2, \ldots\}$,

$$p_X(x) = \begin{cases} (1-\beta)\,\beta^{x-1}, & x \in \mathbb{N} \\ 0, & \text{otherwise.} \end{cases}$$

X = # trials until the next success

$P[X = 3] = P[FFS] = (1-p)(1-p)p = p(1-p)^2$

$P[X = 10] = P[FFFFFFFFFS] = p(1-p)^9$

$P[X = k]$

$= p(1-p)^{k-1}$

- Write $X \sim \mathcal{G}_1(\beta)$ or geometric$_1(\beta)$.
- In MATLAB, use `geopdf(k-1,1-`β`)`.
- Interpretation: X is the number of trials required in a Bernoulli trials to achieve the ~~first~~ next success.

  In particular, in a series of Bernoulli trials (independent trials with constant probability p of a success), let the random variable X denote the number of trials until the first success. Then X is a geometric random variable with parameter $\beta = 1 - p$ and

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

b) • When its support is $\mathbb{N} \cup \{0\} = \{0, 1, 2, 3, \ldots\}$

$$p_X(x) = \begin{cases} (1-\beta)\,\beta^{x}, & \forall x \in \mathbb{N} \cup \{0\}, \\ 0, & \text{otherwise} \end{cases} = \begin{cases} p(1-p)^x, & x = 0, 1, 2, \ldots \\ 0, & \text{otherwise} \end{cases}$$

- Write $X \sim \mathcal{G}_0(\beta)$ or geometric$_0(\beta)$.
- In MATLAB, use `geopdf(k,1-`β`)`.
- Interpretation: X is the number of failures in a Bernoulli trials before the ~~first~~ next success occurs.

8.27. In 1837, the famous French mathematician Poisson introduced a probability distribution that would later come to be known as the Poisson distribution, and this would develop into one of the most important distributions in probability theory. As is often remarked, Poisson did not recognize the huge practical importance of the distribution that would later be named after him. In his book, he dedicates just one page to this distribution. It was Bortkiewicz in 1898, who first discerned and explained the importance of the Poisson distribution in his book Das Gesetz der Kleinen Zahlen (*The Law of Small Numbers*).

Definition 8.28. X is a ***Poisson*** random variable with ***parameter*** $\alpha > 0$ if

$$p_X(k) = \begin{cases} e^{-\alpha}\frac{\alpha^k}{k!}, & k \in \{0,1,2,\dots\} \\ 0, & \text{otherwise} \end{cases}$$

- In `MATLAB`, use `poisspdf(k,alpha)`.
- Write $X \sim \mathcal{P}(\alpha)$ or Poisson(α).
- We will see later in Example 8.49 that α is the "average" or expected value of X.
- Instead of X, Poisson random variable is usually denoted by Λ. The parameter α is often replaced by $\lambda\tau$ where λ is referred to as the ***intensity/rate parameter*** of the distribution

Example 8.29. The first use of the Poisson model is said to have been by a Prussian (German) physician, Bortkiewicz, who found that the annual number of late-19th-century Prussian (German) soldiers kicked to death by horses fitted a Poisson distribution [4, p 150],[2, Ex 2.23][15].

[15] I. J. Good and others have argued that the Poisson distribution should be called the Bortkiewicz distribution, but then it would be very difficult to say or write.

$$p_N(n) = e^{-\alpha}\frac{\alpha^n}{n!}$$

Example 8.30. The number of hits to a popular website during a 1-minute interval is given by $N \sim \mathcal{P}(\alpha)$ where $\alpha = 2$.

(a) Find the probability that there is at least one hit between 3:00AM and 3:01AM.

$P[N \geq 1] = p_N(1) + p_N(2) + p_N(3) + \cdots$ ← difficult

$$= 1 - P[N<1] = 1 - \left(p_N(0)\right) = 1 - e^{-\alpha}\frac{\cancel{\alpha^0}^{\,1}}{\cancel{0!}_{\,1}} = 1 - e^{-\alpha} = 1 - e^{-2} \approx .865$$

(b) Find the probability that there are at least 2 hits during the time interval above.

$P[N \geq 2] = p_N(2) + p_N(3) + p_N(4) + \cdots$

$$= 1 - P[N<2] = 1 - \left(p_N(0) + p_N(1)\right) = 1 - \left(e^{-\alpha} + e^{-\alpha}\frac{\alpha^1}{1!}\right) = 1 - e^{-\alpha}(1+\alpha) = 1 - 3e^{-2} \approx .594$$

8.31. One of the reasons why Poisson distribution is important is because many natural phenomenons can be modeled by ***Poisson processes***.

Definition 8.32. A ***Poisson process*** (PP) is a random arrangement of "marks" (denoted by "×" below) on the time line.

The "marks" may indicate the arrival times or occurrences of event/phenomenon of interest.

Example 8.33. Examples of processes that can be modeled by ***Poisson process*** include

(a) the sequence of times at which lightning strikes occur or mail carriers get bitten within some region

(b) the emission of particles from a radioactive source

(c) the arrival of

- telephone calls at a switchboard or at an automatic phone-switching system
- urgent calls to an emergency center
- (filed) claims at an insurance company
- incoming spikes (action potential) to a neuron in human brain

(d) the occurrence of

- serious earthquakes
- traffic accidents
- power outages

in a certain area.

(e) page view requests to a website

8.34. It is convenient to consider the Poisson process in terms of customers arriving at a facility.

We focus on a type of Poisson process that is called *homogeneous Poisson process.*

Definition 8.35. For ***homogeneous Poisson process***, there is only one parameter that describes the whole process. This number is call the ***rate*** and usually denoted by λ.

Example 8.36. If you think about modeling customer arrival as a Poisson process with rate $\lambda = 5$ customers/hour, then it means that during any fixed time interval of duration 1 hour (say, from noon to 1PM), you expect to have about 5 customers arriving in that interval. If you consider a time interval of duration two hours (say, from 1PM to 3PM), you expect to have about $2 \times 5 = 10$ customers arriving in that time interval.

8.37. More generally, For a homogeneous Poisson process of rate λ, during a time interval of length τ, the average number of arrivals will be $\lambda \times \tau$.

One important fact which we will show later is that, when we consider a fixed time interval, the ***number of arrivals*** for a Poisson process is a Poisson random variable. So, now we know that the "average" or expected value of this random variable must be λT.

Summary: **For a homogeneous Poisson process, the number of arrivals during a time interval of duration T is a Poisson random variable with parameter $\alpha = \lambda T$.**

Example 8.38. Examples of Poisson *random variables*:

- #photons emitted by a light source of intensity λ [photons/second] in time τ
- #atoms of radioactive material undergoing decay in time τ
- #clicks in a Geiger counter in τ seconds when the average number of click in 1 second is λ.
- #dopant atoms deposited to make a small device such as an FET
- #customers arriving in a queue or workstations requesting service from a file server in time τ
- Counts of demands for telephone connections in time τ
- Counts of defects in a semiconductor chip.

Example 8.39. Thongchai produces a new hit song every 7 months on average. Assume that songs are produced according to a Poisson process. Find the probability that Thongchai produces more than two $\underbrace{\text{hit songs}}_{N}$ in $\underbrace{\text{1 year}}_{\tau}$.

$N \sim \mathcal{P}(\alpha)$

$$\lambda = \frac{1 \text{ song}}{7 \text{ months}}$$

$$\alpha = \lambda \times \tau = \frac{1 \text{ song}}{7 \text{ months}} \times \frac{12 \text{ months}}{\cancel{1 \text{ year}}} = \frac{12}{7}$$

$$P[N > 2] = 1 - P[N \le 2] = 1 - \left(P_N(0) + P_N(1) + P_N(2)\right) = 1 - \left(e^{-\alpha}\frac{\alpha^0}{0!} + e^{-\alpha}\frac{\alpha^1}{1!} + e^{-\alpha}\frac{\alpha^2}{2!}\right)$$

$$= 1 - e^{-\alpha}\left(1 + \alpha + \frac{\alpha^2}{2}\right) = 0.2466$$

$\alpha = \frac{12}{7}$

8.40. ***Poisson approximation*** of Binomial distribution: When p is small and n is large, $\mathcal{B}(n,p)$ can be approximated by $\mathcal{P}(np)$

(a) In a large number of independent repetitions of a Bernoulli trial having a small probability of success, the total number of successes is approximately Poisson distributed with parameter $\alpha = np$, where n = the number of trials and p = the probability of success. [19, p 109]

(b) More specifically, suppose $X_n \sim \mathcal{B}(n, p_n)$. If $p_n \to 0$ and $np_n \to \alpha$ as $n \to \infty$, then

binomial   apply calculus   Poisson

$$P\left[X_n = k\right] = \binom{n}{k} p_n^k (1-p_n)^{n-k} \to e^{-\alpha}\frac{\alpha^k}{k!}.$$

Example 8.41. Recall that Bortkiewicz applied the Poisson model to the number of Prussian cavalry deaths attributed to fatal horse kicks. Here, indeed, one encounters a very large number of trials (the Prussian cavalrymen), each with a very small probability of "success" (fatal horse kick).

8.42. Summary:

$X \sim$	Support S_X	$p_X(x)$
Uniform $\mathcal{U}_n$	$\{1, 2, \ldots, n\}$	$\frac{1}{n}$
$\mathcal{U}_{\{0,1,\ldots,n-1\}}$	$\{0, 1, \ldots, n-1\}$	$\frac{1}{n}$
Bernoulli $\mathcal{B}(1,p)$	$\{0, 1\}$	$\begin{cases} 1-p, & x = 0 \\ p, & x = 1 \end{cases}$
Binomial $\mathcal{B}(n,p)$	$\{0, 1, \ldots, n\}$	$\binom{n}{x} p^x (1-p)^{n-x}$
Geometric $\mathcal{G}_0(\beta)$	$\mathbb{N} \cup \{0\}$	$(1-\beta)\beta^x$
Geometric $\mathcal{G}_1(\beta)$	$\mathbb{N}$	$(1-\beta)\beta^{x-1}$
Poisson $\mathcal{P}(\alpha)$	$\mathbb{N} \cup \{0\}$	$e^{-\alpha}\frac{\alpha^x}{x!}$

Table 3: Examples of probability mass functions. Here, $p, \beta \in (0,1)$. $\alpha > 0$. $n \in \mathbb{N}$