MAT 2377
Probability and Statistics for Engineers

Chapter 3
Continuous Distributions

P. Boily (uOttawa)

Winter 2021

Contents

3.1 – Continuous Random Variables

How do we approach probabilities where there there are **uncountably infinitely many outcomes**, such as one might encounter if X represents the height of an individual in the population, for instance (e.g., the outcomes reside in a continuous interval on the real line)?

What's the probability that a randomly selected person is 6 feet tall?

In the discrete case, the probability mass function $f_X(x) = P(X = x)$ was the main object of interest. In the continuous case, the analogous role is played by the **probability density function** (p.d.f.), still denoted by $f_X(x)$, but

$$f_X(x) \neq P(X = x).$$

The **(cumulative) distribution function** (c.d.f.) of any such random variable X is still defined by

$$F_X(x) = P(X \leq x)\,,$$

viewed as a function of a real variable x; but $P(X \leq x)$ is not simply computed by adding a few terms of the form $P(X = x_i)$. Note that

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

We can describe the **distribution** of the random variable X *via* the following relationship between $f_X(x)$ and $F_X(x)$:

$$f_X(x) = \frac{d}{dx}F_X(x).$$

Area Under a Curve

For any $a < b$, we have

$$\{X \leq b\} = \{X \leq a\} \cup \{a < X \leq b\},$$

so that

$$\begin{aligned} P(X \leq a) + P(a < X \leq b) &= P(X \leq b) \\ P(a < X \leq b) &= P(X \leq b) - P(X \leq a) \\ &= F_X(b) - F_X(a) = \int_a^b f_X(x)\,dx \end{aligned}$$

Probability Density Functions

The **probability density function** (p.d.f.) of a continuous random variable X is an **integrable** function $f_X : X(\mathcal{S}) \to \mathbb{R}$ such that

- $f_X(x) > 0$ for all $x \in X(\mathcal{S})$ and $\lim_{x\to\pm\infty} f_X(x) = 0$;
- $\int_{\mathcal{S}} f_X(x)\,dx = 1$;
- for any event $A = (a,b) = \{X|a < X < b\}$,

$$P(A) = P((a,b)) = \int_a^b f_X(x)\,dx;$$

- for any x,

$$F_X(x) = P(X \leq x) = \int_{-\infty}^{x} f_X(t)\,dt;$$

- for any x,

$$P(x > X) = 1 - P(X \leq x) = 1 - F_X(x) = 1 - \int_{-\infty}^{x} f_X(t)\,dt;$$

- for any a, b,

$$P(a < X < b) = P(a \leq X < b) = P(a < X \leq b) = P(a \leq X \leq b)$$
$$= F_X(b) - F_X(a) = \int_a^b f(x)\,dx.$$

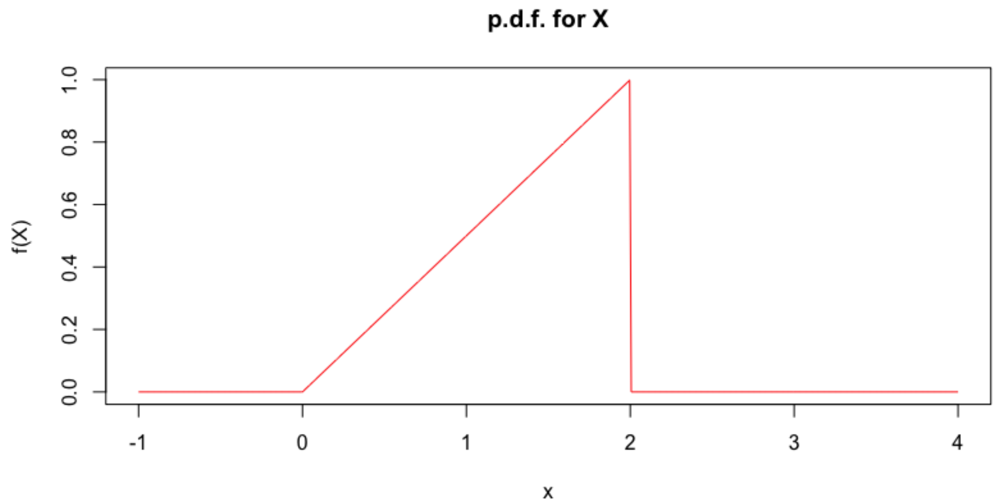
Examples:

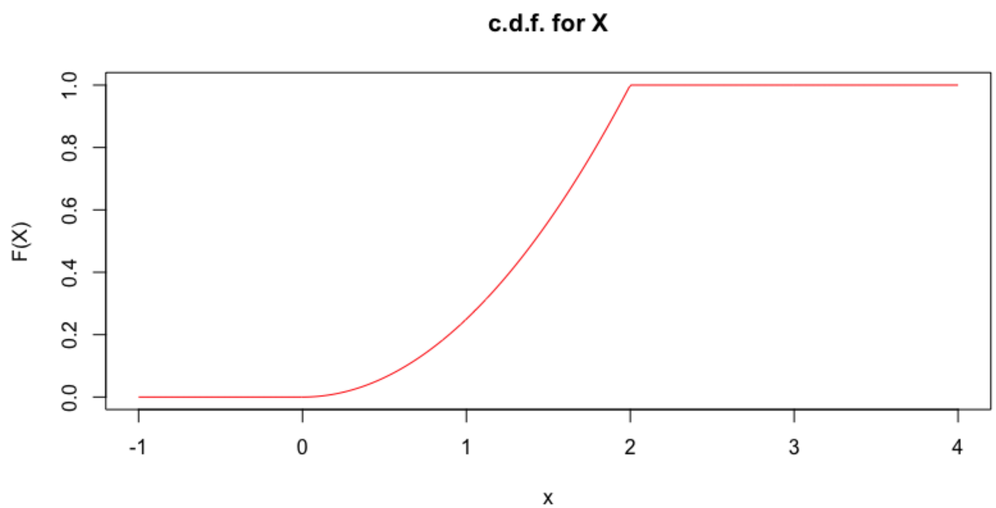
1. Assume that X has the following p.d.f.

$$f_X(x) = \begin{cases} 0 & \text{if } x < 0 \\ x/2 & \text{if } 0 \le x \le 2 \\ 0 & \text{if } x > 2 \end{cases} \qquad \text{(note that } \int_0^2 f(x)\,dx = 1\text{).}$$

The corresponding c.d.f. is given by:

$$\begin{aligned} F_X(x) = P(X \le x) &= \int_{-\infty}^{x} f_X(t)\,dt \\ &= \begin{cases} 0 & \text{if } x < 0 \\ 1/2 \cdot \int_0^x t\,dt = 1/2 \cdot [t^2/2]_0^x = x^2/4 & \text{if } 0 < x < 2 \\ 1 & \text{if } x \ge 2 \end{cases} \end{aligned}$$


p.d.f. for X
f(X)
0.0
0.2
0.4
0.6
0.8
1.0
-1
0
1
2
3
4
x


c.d.f. for X
F(X)
0.0
0.2
0.4
0.6
0.8
1.0
-1
0
1
2
3
4
x

2. What is the probability of the event $A = \{X|0.5 < X < 1.5\}$?

   Solution: we need to evaluate

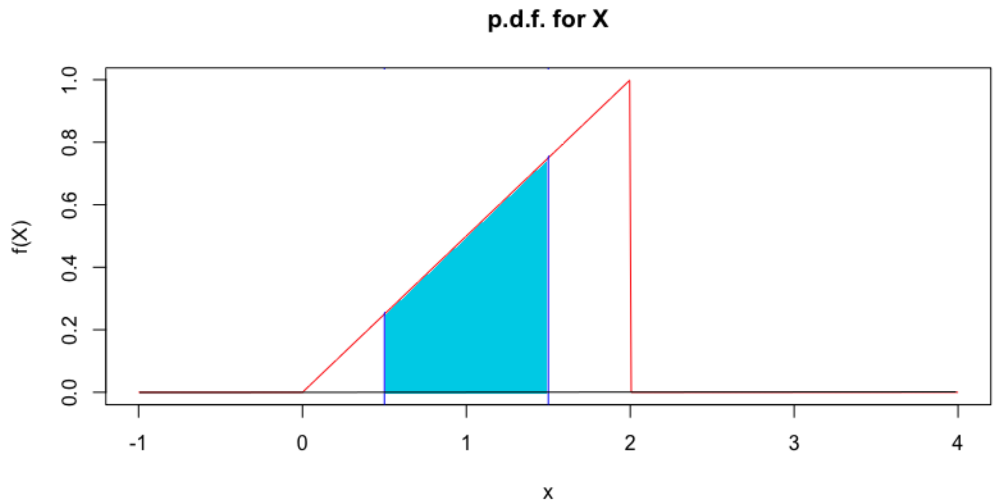
$$P(A) = P(0.5 < X < 1.5) = F_X(1.5) - F_X(0.5) = \frac{(1.5)^2}{4} - \frac{(0.5)^2}{4} = \frac{1}{2}.$$

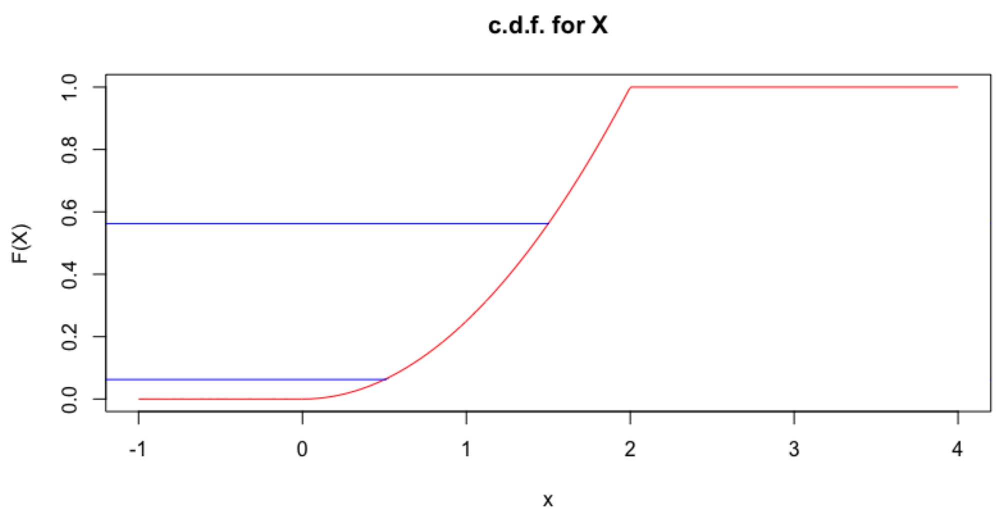
3. What is the probability of the event $B = \{X|X = 1\}$?

   Solution: we need to evaluate

$$P(B) = P(X = 1) = P(1 \leq X \leq 1) = F_X(1) - F_X(1) = 0.$$

   This is unexpected: even though $f_X(1) = 0.5 \neq 0$, $P(X = 1) = 0$! The probability that a continuous random variable X take on any particular single value is nil.


p.d.f. for X
f(X)
0.0
0.2
0.4
0.6
0.8
1.0
-1
0
1
2
3
4
x


c.d.f. for X
F(X)
0.0
0.2
0.4
0.6
0.8
1.0
-1
0
1
2
3
4
x

4. Assume that, for some $\lambda > 0$, X has the following p.d.f.:

$$f_X(x) = \begin{cases} \lambda \exp(-\lambda x) & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases} \qquad \left(\text{is } \int_{-\infty}^{\infty} f(x)\,dx = 1?\right)$$

What is the probability that $X > 10.2$?

Solution: the corresponding c.d.f. is given by:

$$F_X(x;\lambda) = P_\lambda(X \leq x) = \int_{-\infty}^{x} f_X(t)\,dt = \begin{cases} 0 & \text{if } x < 0 \\ \lambda \int_0^x \exp(-\lambda t)\,dt & \text{if } x \geq 0 \end{cases}$$

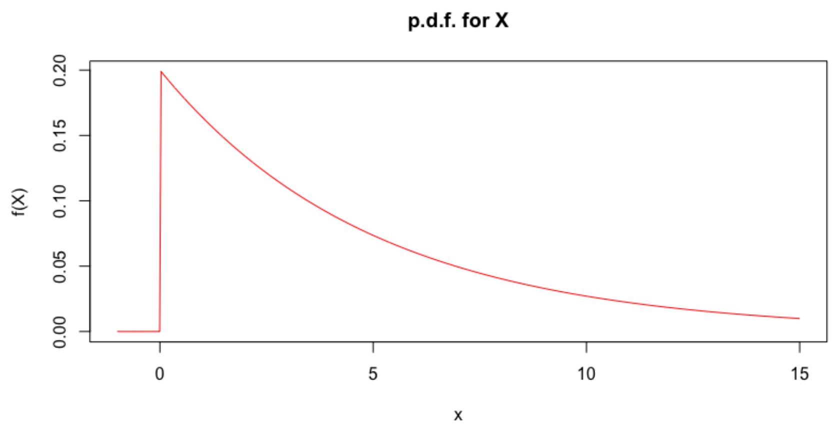
$$= \begin{cases} 0 & \text{if } x < 0 \\ [-\exp(-\lambda t)]_0^x = 1 - \exp(-\lambda x) & \text{if } x \geq 0 \end{cases}$$

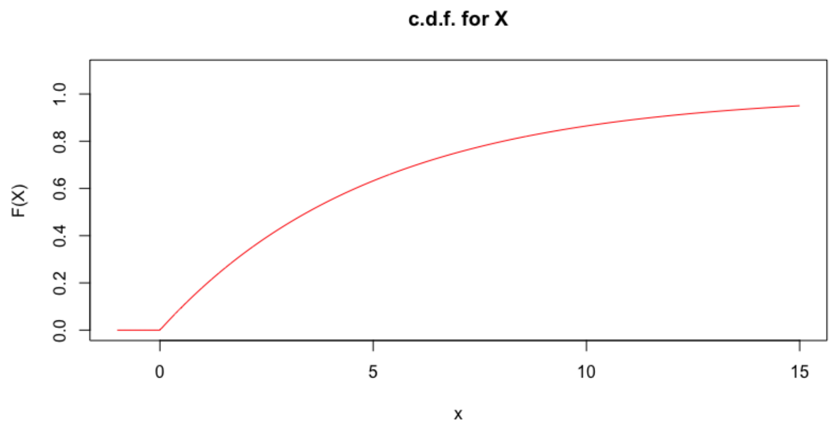
Then

$$P_\lambda(X > 10.2) = 1 - F_X(10.2; \lambda) = 1 - [1 - \exp(-10.2\lambda)] = \exp(-10.2\lambda)$$

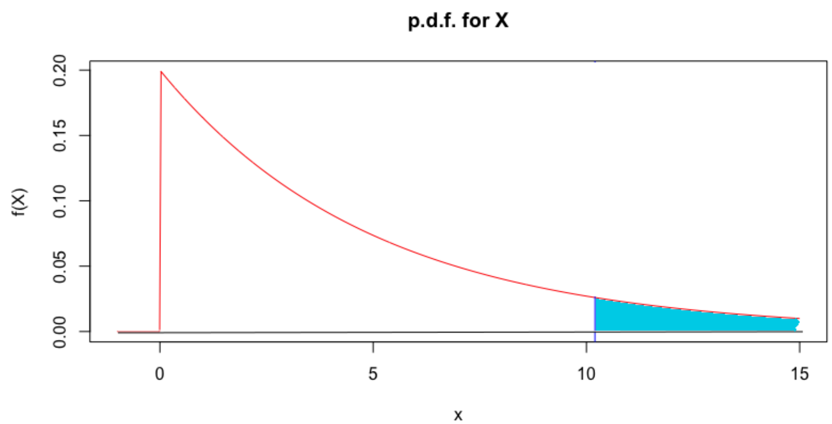
is a function of the **distribution parameter** λ itself:

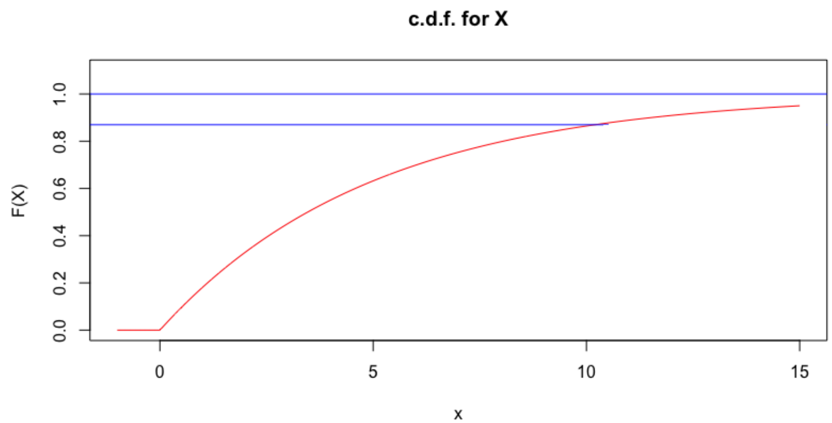
λ	$P_\lambda(X > 10.2)$
0.002	0.9798
0.02	0.8155
0.2	0.1300
2	1.38×10^{-9}
20	2.54×10^{-89}
200	0 (for all intents and purposes)


p.d.f. for X
f(X)
0.00
0.05
0.10
0.15
0.20
0
5
10
15
x

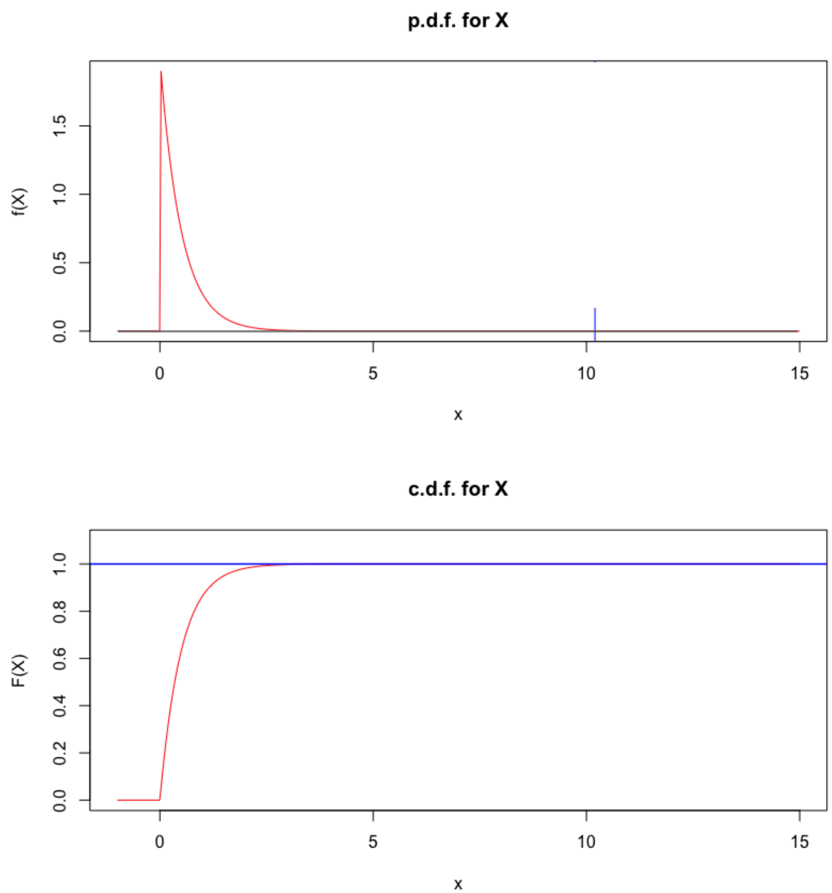

c.d.f. for X
F(X)
0.0
0.2
0.4
0.6
0.8
1.0
0
5
10
15
x

$$\lambda = 0.2$$





$$\lambda = 0.2,\ P_{0.2}(X > 10.2) \approx 0.1300$$



$$\lambda = 2,\ P_2(X > 10.2) \approx 1.38 \times 10^{-9}$$

3.2 – Expectation of a Continuous Random Variable

For a continuous random variable X with p.d.f. $f_X(x)$, the **expectation** of X is defined as

$$\mathrm{E}[X] = \int_{-\infty}^{\infty} x f_X(x)\,dx\,.$$

In a similar way to the discrete case, for any function $h(X)$, we have

$$\mathrm{E}\left[h(X)\right] = \int_{-\infty}^{\infty} h(x) f_X(x)\,dx\,.$$

Note that the **expectation need not exist**!

Examples:

1. Find the expected value of X in the example 1, above.

   Solution: we need to evaluate

$$\mathrm{E}[X] = \int_{-\infty}^{\infty} x f_X(X)\,dx = \int_0^2 x f_X(x)\,dx = \int_0^2 x \cdot x/2\,dx$$
$$= \int_0^2 \frac{x^2}{2}\,dx = \left[\frac{x^3}{6}\right]_{x=0}^{x=2} = \frac{4}{3}.$$

2. What about X^2?

   Solution: we have $\mathrm{E}[X^2] = \int_0^2 \frac{x^3}{2}\,dx = 2.$

3. Compute the expectation of the random variable X with p.d.f.

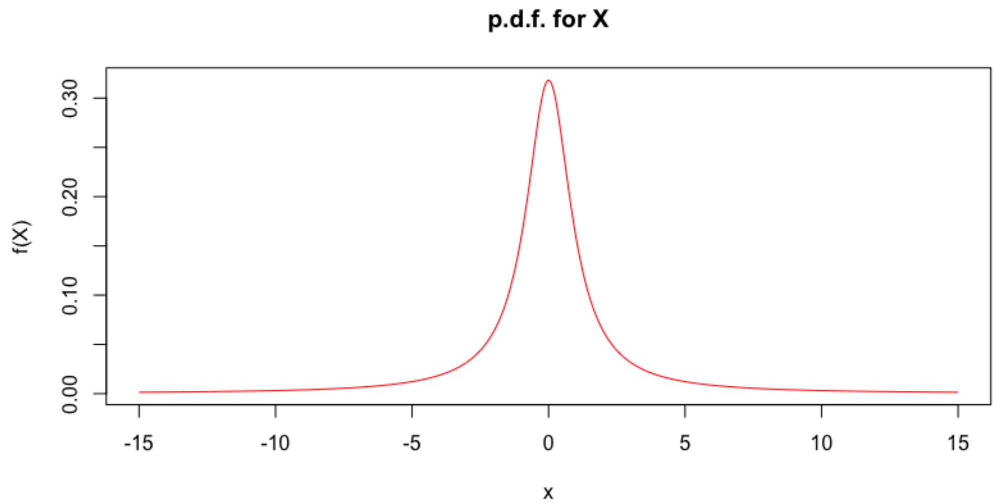
$$f_X(x) = \frac{1}{\pi(1+x^2)}, \quad -\infty < x < \infty.$$

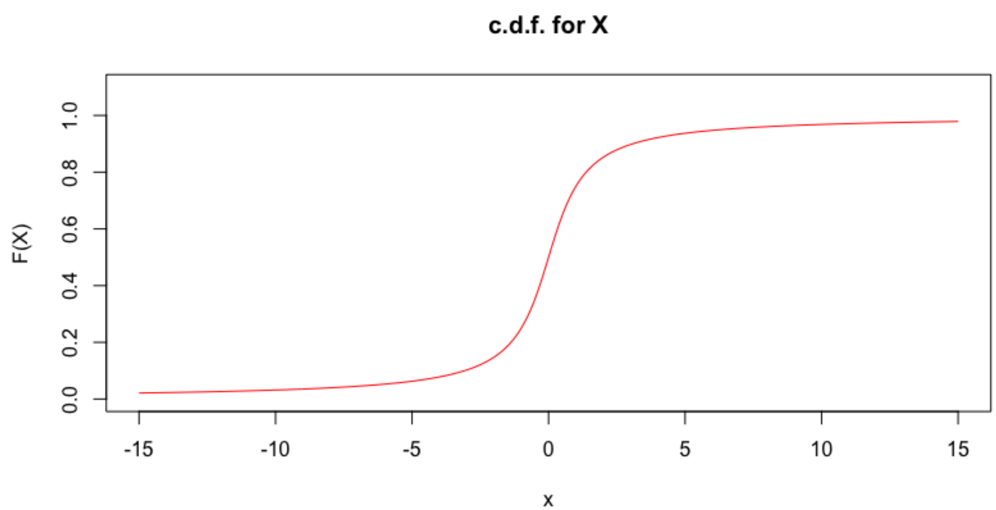
Solution: let's verify that $f_X(x)$ is indeed a p.d.f.:

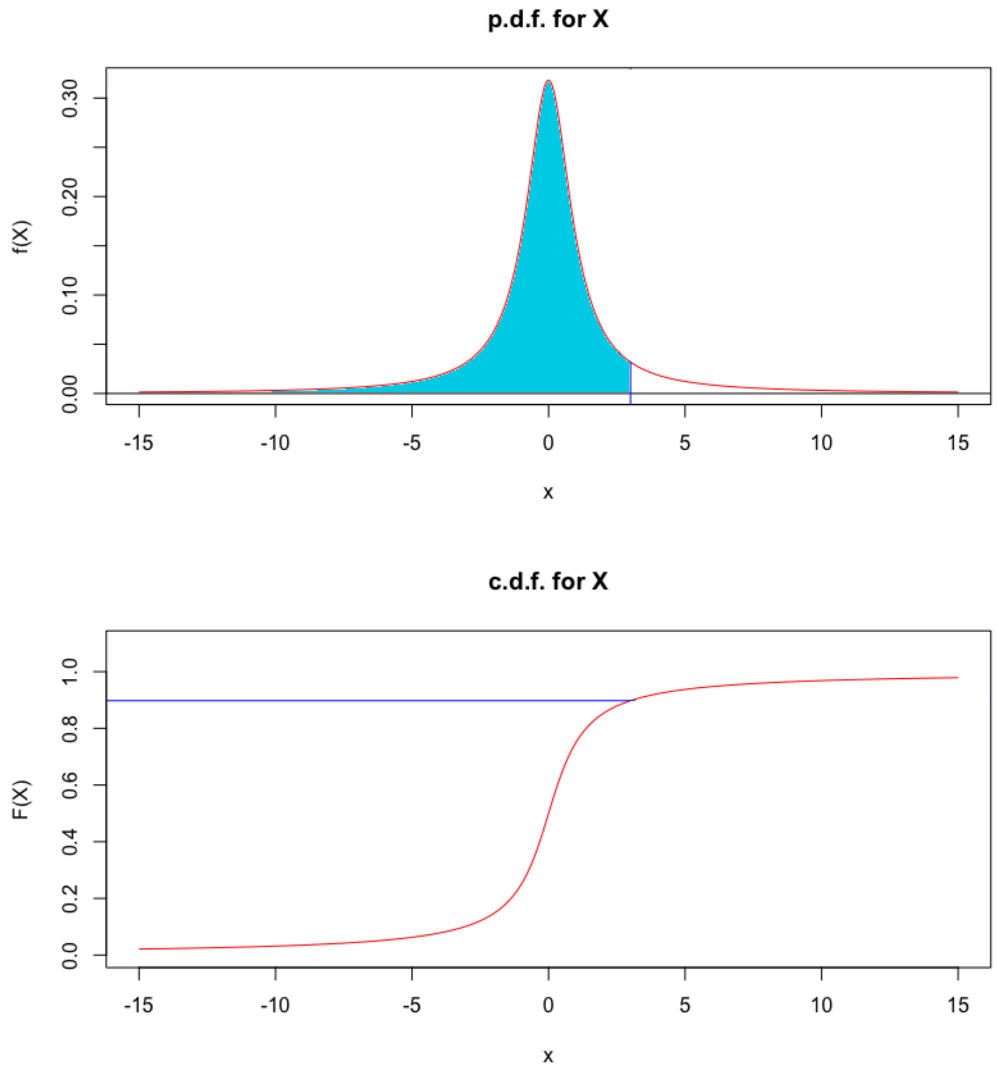
$$\int_{-\infty}^{\infty} f_X(x)\,dx = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{1}{1+x^2}\,dx = \frac{1}{\pi}[\arctan(x)]_{-\infty}^{\infty} = \frac{1}{\pi}\left[\frac{\pi}{2} - \frac{-\pi}{2}\right] = 1.$$

We can also easily see that

$$F_X(x) = P(X \leq x) = \int_{-\infty}^{x} f_X(t)\,dt = \frac{1}{\pi}\int_{-\infty}^{x} \frac{1}{1+t^2}\,dt = \frac{1}{\pi}\arctan(x) + \frac{1}{2}.$$


p.d.f. for X
f(X)
0.00
0.10
0.20
0.30
-15
-10
-5
0
5
10
15
x


c.d.f. for X
F(X)
0.0
0.2
0.4
0.6
0.8
1.0
-15
-10
-5
0
5
10
15
x


p.d.f. for X
f(X)
0.00
0.10
0.20
0.30
-15
-10
-5
0
5
10
15
x
c.d.f. for X
F(X)
0.0
0.2
0.4
0.6
0.8
1.0
-15
-10
-5
0
5
10
15
x

In particular, $P(X \leq 3) = F_X(3) = \frac{1}{\pi}\arctan(3) + \frac{1}{2} \approx 0.8976$, say.

The expectation of X is

$$\mathrm{E}[X] = \int_{-\infty}^{\infty} x f_X(x)\,dx = \int_{-\infty}^{\infty} \frac{x}{\pi(1+x^2)}\,dx.$$

If this improper integral exists, then it needs to be equal, among other things, **both** to

$$\underbrace{\int_{-\infty}^{0} \frac{x}{\pi(1+x^2)}\,dx + \int_{0}^{\infty} \frac{x}{\pi(1+x^2)}\,dx}_{\text{candidate 1}} \quad \text{and to} \quad \underbrace{\lim_{a\to\infty} \int_{-a}^{a} \frac{x}{\pi(1+x^2)}\,dx}_{\text{candidate 2}}.$$

It is straightforward to find an antiderivative of $\frac{x}{\pi(1+x^2)}$.

Set $u = 1 + x^2$. Then $du = 2xdx$ and $xdx = \frac{du}{2}$, and we obtain

$$\int \frac{x}{\pi(1+x^2)}\,dx = \frac{1}{2\pi}\int u\,du = \frac{1}{2\pi}\ln|u| = \frac{1}{2\pi}\ln(1+x^2).$$

The candidate 2 integral reduces to

$$\lim_{a\to\infty}\left[\frac{\ln(1+x^2)}{2\pi}\right]_{-a}^{a} = \lim_{a\to\infty}\left[\frac{\ln(1+a^2)}{2\pi} - \frac{\ln(1+(-a)^2)}{2\pi}\right] = \lim_{a\to\infty} 0 = 0;$$

while the candidate 1 integral reduces to

$$\left[\frac{\ln(1+x^2)}{2\pi}\right]_{-\infty}^{0} + \left[\frac{\ln(1+x^2)}{2\pi}\right]_{0}^{\infty} = 0 - (\infty) + \infty - 0 = \infty - \infty$$

which is **undefined**. Thus $\mathrm{E}[X]$ does not exist (or is undefined).

Mean and Variance of a Continuous Random Variable

In a similar way to the discrete case, the **mean** of X is defined to be $\mathrm{E}[X]$, and the **variance** and **standard deviation** of X are, as before,

$$\mathrm{Var}[X] \overset{\text{def}}{=} \mathrm{E}\left[(X - \mathrm{E}(X))^2\right] \overset{\text{comp. formula}}{=} \mathrm{E}[X^2] - \mathrm{E}^2[X]\,,$$
$$\mathrm{SD}[X] = \sqrt{\mathrm{Var}[X]}\,.$$

As in the discrete case, if X, Y are continuous random variables, and $a, b \in \mathbb{R}$,

$$\mathrm{E}[aY + bX] = a\mathrm{E}[Y] + b\mathrm{E}[X]$$
$$\mathrm{Var}[a + bX] = b^2\mathrm{Var}[X]$$
$$\mathrm{SD}[a + bX] = |b|\mathrm{SD}[X]$$

3.3 – Normal Distributions

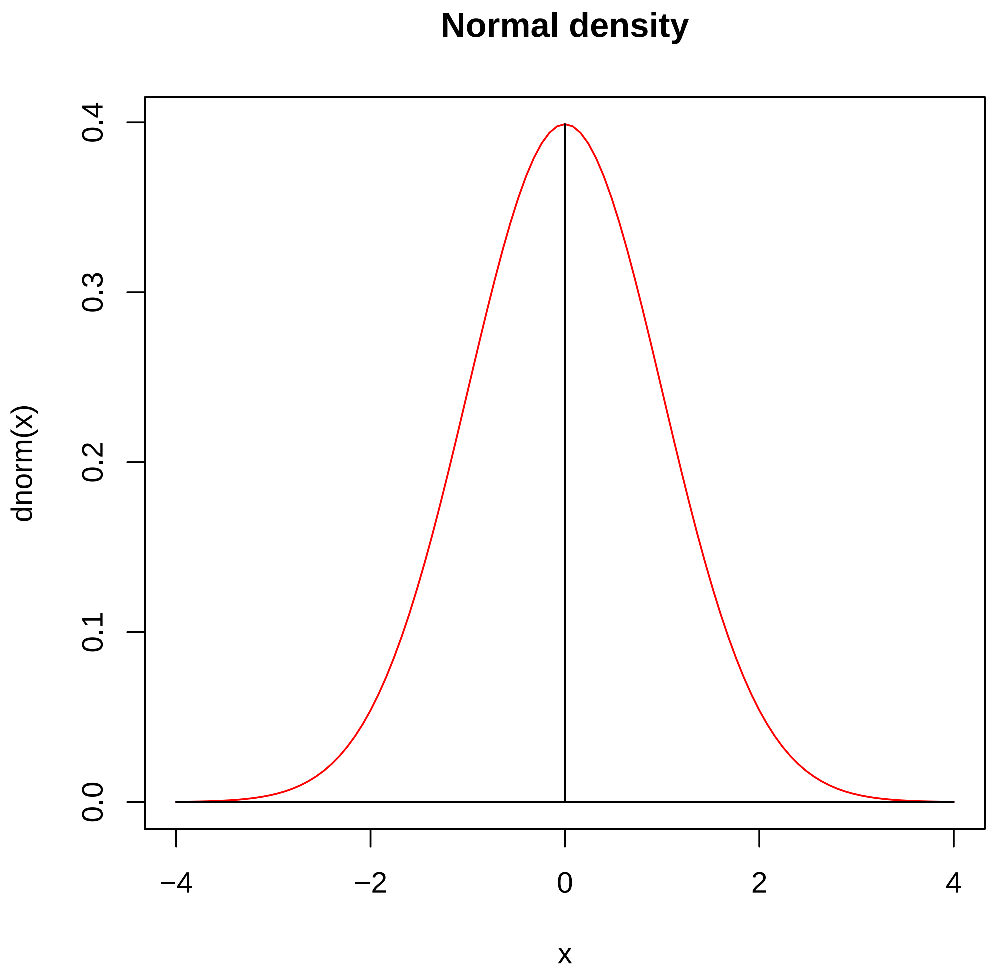
A **very** important example of continuous distributions is that of the special probability distribution function

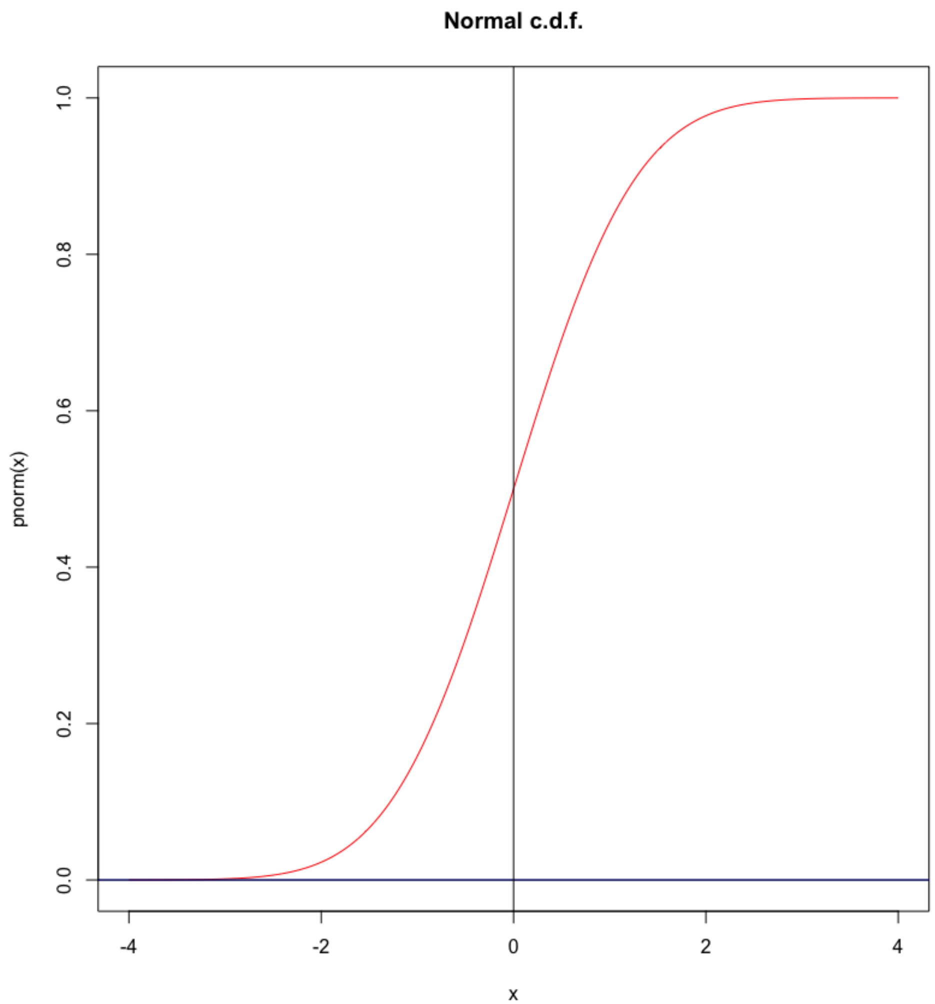
$$\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2} .$$

The corresponding cumulative distribution function is denoted by

$$\Phi(z) = P(Z \leq z) = \int_{-\infty}^{z} \phi(t)\, dt .$$

A random variable Z with this c.d.f. is said to have a **standard normal distribution**, and we write $Z \sim \mathcal{N}(0,1)$.


Normal density
dnorm(x)
0.0
0.1
0.2
0.3
0.4
−4
−2
0
2
4
x


Normal c.d.f.
pnorm(x)
0.0
0.2
0.4
0.6
0.8
1.0
-4
-2
0
2
4
x

Standard Normal Random Variables

The expectation and variance of $Z \sim \mathcal{N}(0,1)$ are

$$\mathrm{E}[Z] = \int_{-\infty}^{\infty} z\,\phi(z)\,dz = \int_{-\infty}^{\infty} z\,\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}z^2}\,dz = 0,$$

$$\mathrm{Var}[Z] = \int_{-\infty}^{\infty} z^2\,\phi(z)\,dz = 1, \quad \mathrm{SD}[Z] = \sqrt{\mathrm{Var}[Z]} = \sqrt{1} = 1.$$

Other quantities of interest include:

$$\Phi(0) = P(Z \leq 0) = \frac{1}{2},\ \Phi(-\infty) = 0,\ \Phi(\infty) = 1,$$

$$\Phi(1) = P(Z \leq 1) = \texttt{pnorm(1)} \approx 0.8413, etc.$$

General Normal Random Variables

Let $\sigma > 0$ and $\mu \in \mathbb{R}$. If $Z \sim \mathcal{N}(0,1)$ and $X = \mu + \sigma Z$, then

$$\frac{X-\mu}{\sigma} = Z \sim \mathcal{N}(0,1).$$

However, the c.d.f. of X is given by

$$\begin{aligned} F_X(x) &= P(X \leq x) = P(\mu + \sigma Z \leq x) = P\left(Z \leq \frac{x-\mu}{\sigma}\right) \\ &= \Phi\left(\frac{x-\mu}{\sigma}\right). \end{aligned}$$

The p.d.f. of X is then

$$f_X(x) = \frac{d}{dx}F_X(x) = \frac{d}{dx}\Phi\left(\frac{x-\mu}{\sigma}\right) = \frac{1}{\sigma}\phi\left(\frac{x-\mu}{\sigma}\right)$$

Any random variable X with this c.d.f./p.d.f. must satisfy

$$\mathrm{E}[X] = \mu + \sigma\mathrm{E}[Z] = \mu, \quad \mathrm{Var}[X] = \sigma^2\mathrm{Var}[Z] = \sigma^2 \ \Rightarrow \ \mathrm{SD}[X] = \sigma$$

and is said to be **normal with mean μ and variance σ^2**, denoted by $X \sim \mathcal{N}(\mu, \sigma^2)$.

Every general normal X can be obtained by a linear transformation of the standard normal Z!

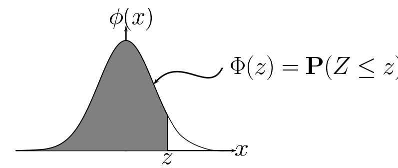


Table 1. Normal Distribution Function
Lower tail of the standard normal distribution is tabulated

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.0
0.00	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5
0.10	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5
0.20	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6
0.30	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6
0.40	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6
0.50	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7
0.60	0.7258	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7
0.70	0.7580	0.7612	0.7642	0.7673	0.7703	0.7734	0.7764	0.7793	0.7823	0.7
0.80	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8079	0.8106	0.8
0.90	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8
1.00	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8
1.10	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8
1.20	0.8849	0.8869	0.8888	0.8906	0.8925	0.8943	0.8962	0.8980	0.8997	0.9
1.30	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9
1.40	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9
1.50	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9430	0.9
1.60	0.9452	0.9463	0.9474	0.9485	0.9495	0.9505	0.9515	0.9525	0.9535	0.9
1.70	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9
1.80	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9700	0.9
1.90	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9
2.00	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9
2.10	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9

Examples:

1. Assume that Z represents the standard normal random variable. Evaluate the following probabilities:

   a) $P(Z \leq 0.5) = 0.6915$
   b) $P(Z < -0.3) = 0.3821$
   c) $P(Z > 0.5) = 1 - P(Z \leq 0.5) = 1 - 0.6915 = 0.3085$,
   d) $P(0.1 < Z < 0.3) = P(Z < 0.3) - P(Z < 0.1) = 0.6179 - 0.5398 = 0.0781$,
   e) $P(-1.2 < Z < 0.3) = P(Z < 0.3) - P(Z < -1.2) = 0.5028$.

2. Suppose that the waiting time (in minutes) for a coffee at 9am is normally distributed with mean 5 and standard deviation 0.5. What is the probability that one such waiting time is at most 6 minutes?

   Solution: let X denote the waiting time; then $X \sim \mathcal{N}(5, 0.5^2)$ and the **standardised random variable** is a standard normal:

$$Z = \frac{X-5}{0.5} \sim \mathcal{N}(0,1)\,.$$

   The desired probability is

$$P\left(X \leq 6\right) = P\left(\frac{X-5}{0.5} \leq \frac{6-5}{0.5}\right) = P\left(Z \leq \frac{6-5}{0.5}\right) = \Phi\left(\frac{6-5}{0.5}\right)$$
$$= \Phi(2) = P(Z \leq 2) \approx 0.9772 \text{ (reading from the table).}$$

3. Suppose that bottles of beer are filled in such a way that the actual volume of the liquid in them (in mL) varies randomly according to a normal distribution with $\mu = 376.1$ and $\sigma = 0.4$. What is the probability that the volume in any randomly selected bottle is less than 375mL?

   Solution: let X denote the volume of the liquid in the bottle; then

$$X \sim \mathcal{N}(376.1, 0.4^2) \quad \text{and so} \quad Z = \frac{X - 376.1}{0.4} \sim \mathcal{N}(0,1)\,.$$

   The desired probability is

$$P\left(X < 375\right) = P\left(\frac{X - 376.1}{0.4} < \frac{375 - 376.1}{0.4}\right) = P\left(Z < \frac{-1.1}{0.4}\right)$$
$$= P(Z \leq -2.75) = \Phi\left(-2.75\right) \approx 0.003\,.$$

4. If $Z \sim \mathcal{N}(0,1)$, for which values a, b and c do we have

   a) $P(Z \leq a) = 0.95$;
   b) $P(|Z| \leq b) = P(-b \leq Z \leq b) = 0.99$;
   c) $P(|Z| \geq c) = 0.01$.

   Solution:

   a) From the table we see that

   $$P(Z \leq 1.64) \approx 0.9495 \quad \text{and} \quad P(Z \leq 1.65) \approx 0.9505\,.$$

   Clearly we must have $1.64 < a < 1.65$; a linear interpolation provides a decent guess at $a \approx 1.645$, although this level of precision is usually not necessary. It is often sufficient to simply present the initial interval estimate.

b) Note that

$$P(-b \leq Z \leq b) = P(Z \leq b) - P(Z < -b)$$

However the p.d.f. $\phi(z)$ is symmetric about $z = 0$, which means that

$$P(Z < -b) = P(Z > b) = 1 - P(Z \leq b),$$

and so that

$$P(-b \leq Z \leq b) = P(Z \leq b) - [1 - P(Z \leq b)] = 2P(Z \leq b) - 1$$

In the question, $P(-b \leq Z \leq b) = 0.99$, so that

$$2P(Z \leq b) - 1 = 0.99 \;\Rightarrow\; P(Z \leq b) = \frac{1 + 0.99}{2} = 0.995\,;$$

Consulting the table we see that

$$P(Z \leq 2.57) \approx 0.9949 \quad \text{and} \quad P(Z \leq 2.58) \approx 0.9951;$$

linear interpolation suggests taking $b \approx 2.575$.

c) Note that $\{|Z| \geq c\} = \{|Z| < c\}^c$, so we need to find c such that

$$P(|Z| < c) = 1 - P(|Z| \geq c) = 0.99.$$

But this is equivalent to

$$P(-c < Z < c) = P(-c \leq Z \leq c) = 0.99$$

since $|x| < y \Leftrightarrow -y < x < y$, and $P(Z = c) = 0$ for all c. This problem was solved in the preceding example; take $c \approx 2.575$.

3.4 – Exponential Distributions

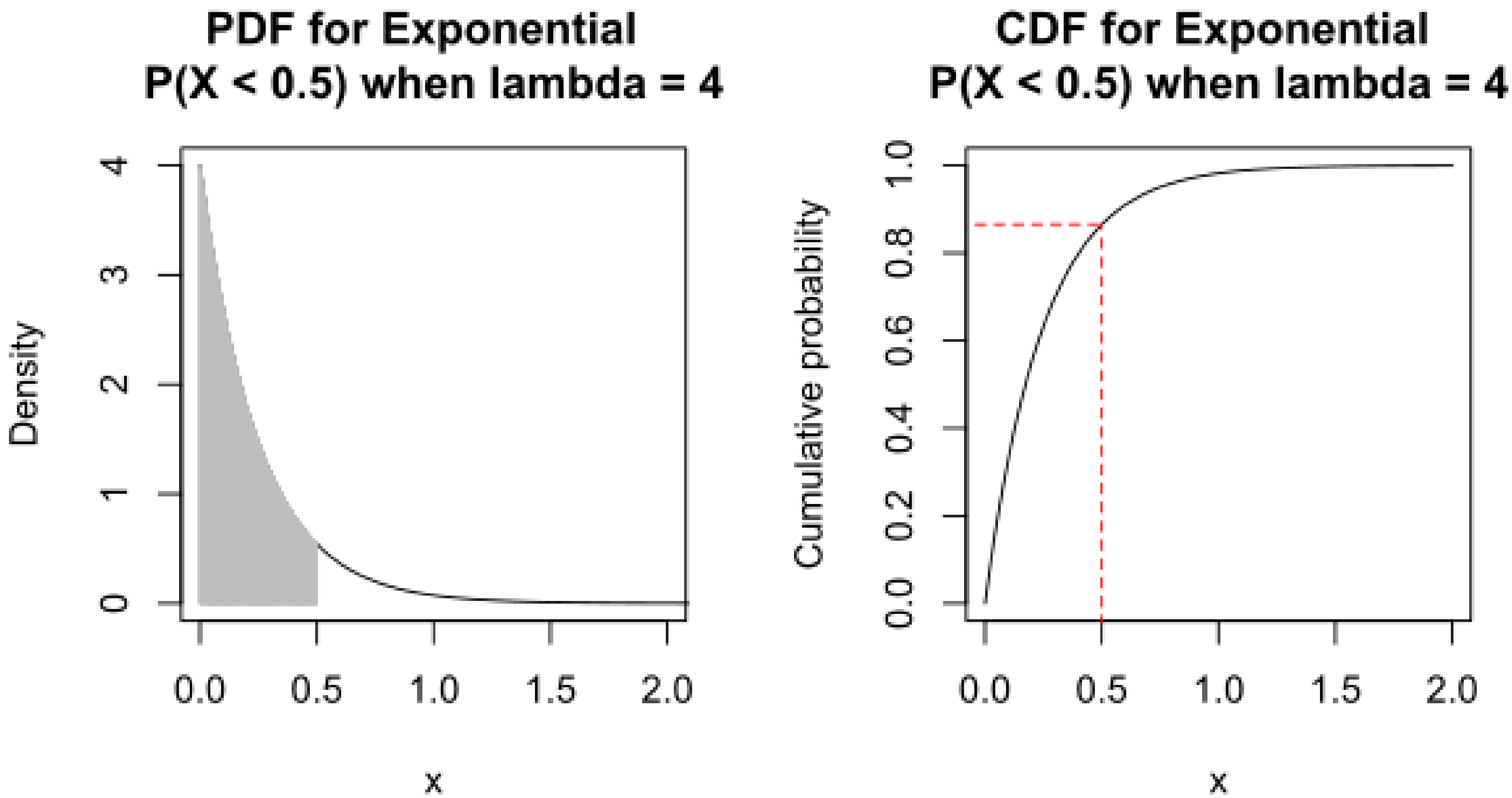
Assume that cars arrive according to a **Poisson process with rate** λ, i.e. the number of cars arriving within a fixed unit time period is a Poisson random variable with parameter λ.

Over a period of time x, we would expect the number of arrivals N to follow a Poisson process with parameter λx. Let X be the wait time to the first car arrival. Then

$$P(X > x) = 1 - P(X \leq x) = P(N = 0) = \exp(-\lambda x).$$

We say that X follows a **exponential distribution** $\text{Exp}(\lambda)$, and

$$F_X(x) = \begin{cases} 0 & \text{for } x < 0 \\ 1 - e^{-\lambda x} & \text{for } 0 \leq x \end{cases} \quad \text{and} \quad f_X(x) = \begin{cases} 0 & \text{for } x \leq 0 \\ \lambda e^{-\lambda x} & \text{for } 0 \leq x \end{cases}$$



If $X \sim \text{Exp}(4)$, then $P(X < 0.5) = F_X(0.5) = 1 - e^{-4(0.5)} \approx 0.865.$

Properties of Exponential Random Variables

- $\mu = \mathrm{E}[X] = 1/\lambda$;
- $\sigma^2 = \mathrm{Var}[X] = 1/\lambda^2$;
- **Memory-Less Property**:

$$P(X > s + t \mid X > t) = P(X > s),$$

- Exp(λ) is the continuous analogue to the **geometric** distribution Geo(p).

Example: the lifetime of a certain type of light bulb has an exponential distribution with mean 100 hours (i.e. $\lambda = 1/100$).

1. What is the probability that a light bulb will last at least 100 hours?

   Solution: $X \sim \text{Exp}(1/100)$, so

   $$P(X > 100) = 1 - P(X \leq 100) = \exp(-100/100) = e^{-1} \approx 0.3679.$$

2. Given that a light bulb has already been burning for 100 hours, what is the probability that it will last at least 100 hours more?

   Solution: we are interested in evaluating $P(X > 200|X > 100)$. By the memory-less property,

   $$P(X > 200|X > 100) = P(X > 200 - 100) = P(X > 100) \approx 0.3679.$$

3. The manufacturer wants to guarantee that their light bulbs will last at least t hours. What should t be in order to ensure that 90% of the light bulbs will last longer than t hours?

   Solution: we need to find t such that $P(X > t) = 0.9$. In other words, we are looking for t such that

$$0.9 = P(X > t) = 1 - P(X \leq t) = 1 - F_X(t) = e^{-0.01t},$$

   that is

$$\ln 0.9 = -0.01t \quad \Longrightarrow \quad t = -100 \ln 0.9 \approx 10.53605 \text{ hours.}$$

3.5 – Gamma Distributions

Assume that cars arrive according to a Poisson process with rate λ. Recall that if X is the time to the first car arrival, then $X \sim \text{Exp}(\lambda)$.

If Y is the wait time to the rth arrival, then Y follows a **Gamma distribution** with parameters λ and r, $Y \sim \Gamma(\lambda, r)$, for which the p.d.f. is

$$f_Y(y) = \begin{cases} 0 & \text{for } y < 0 \\ \frac{y^{r-1}}{(r-1)!}\lambda^r e^{-\lambda y} & \text{for } 0 \leq y \end{cases}$$

$F_Y(y)$ cannot be expressed with elementary functions. We also have

$$\mu = \mathrm{E}[Y] = \frac{r}{\lambda} \quad \text{and} \quad \sigma^2 = \mathrm{Var}[Y] = \frac{r}{\lambda^2}.$$

Examples:

1. Suppose that an average of 30 customers per hour arrive at a shop in accordance with a Poisson process, that is to say, $\lambda = 1/2$ customers arrive on average every minute. What is the probability that the shopkeeper will wait more than 5 minutes before both of the first two customers arrive?

   Solution: let Y denote the wait time in minutes until the second customer arrives. Then $Y \sim \Gamma(1/2, 2)$ and

$$P(Y > 5) = \int_5^\infty \frac{y^{2-1}}{(2-1)!}(1/2)^2 e^{-y/2}\,dy = \int_5^\infty \frac{ye^{-y/2}}{4}\,dy$$
$$= \frac{1}{4}\left[-2ye^{-y/2} - 4e^{-y/2}\right]_5^\infty = \frac{7}{2}e^{-5/2} \approx 0.287.$$

2. Telephone calls arrive at a switchboard at a mean rate of $\lambda = 2$ per minute, according to a Poisson process. Let Y be the waiting time until the 5th call arrives. What is the p.d.f., the mean, and the variance of Y?

   Solution: we have

$$f_Y(y) = \frac{2^5 y^4}{4!} e^{-2y}, \text{ for } 0 \leq y < \infty, \quad \mathrm{E}[Y] = \frac{5}{2}, \quad \mathrm{Var}[Y] = \frac{5}{4}.$$

The Gamma distribution can be extended to cases where $r > 0$ is not an integer by replacing $(r-1)!$ by $\Gamma(r) = \int_0^\infty t^{r-1} e^{-t}\, dt$.

The exponential and the χ^2 distributions (we will discuss that one later) are special cases of $\Gamma(\lambda, r)$: Exp$(\lambda) = \Gamma(\lambda, 1)$ and $\chi^2(r) = \Gamma(1/2, r)$.

3.6 – Joint Distributions

Let X, Y be two continuous random variables. The **joint probability distribution function** (joint p.d.f.) of X, Y is a function $f(x, y)$ satisfying

1. $f(x, y) \geq 0$, for all x, y;
2. $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y)\, dxdy = 1$, and
3. $P(A) = \iint_A f(x, y)\, dxdy$, where $A \subseteq \mathbb{R}^2$.

Properties for discrete r.v.: replace integrals by sums, cap $f(x, y) \leq 1$.

Property 3 implies that $P(A)$ is the volume of the solid over the region A in the xy plane bounded by the surface $z = f(x, y)$.

Examples:

1. Roll a pair of unbiased dice. For each of the 36 possible outcomes, let X denote the smaller roll, and Y the larger roll.

   a) How many outcomes correspond to the event $A = \{(X = 2, Y = 3)\}$?
   Solution: the rolls $(3, 2)$ and $(2, 3)$ both give rise to event A.

   b) What is $P(A)$?
   Solution: there are 36 possible outcomes, so $P(A) = \frac{2}{36} \approx 0.0556$.

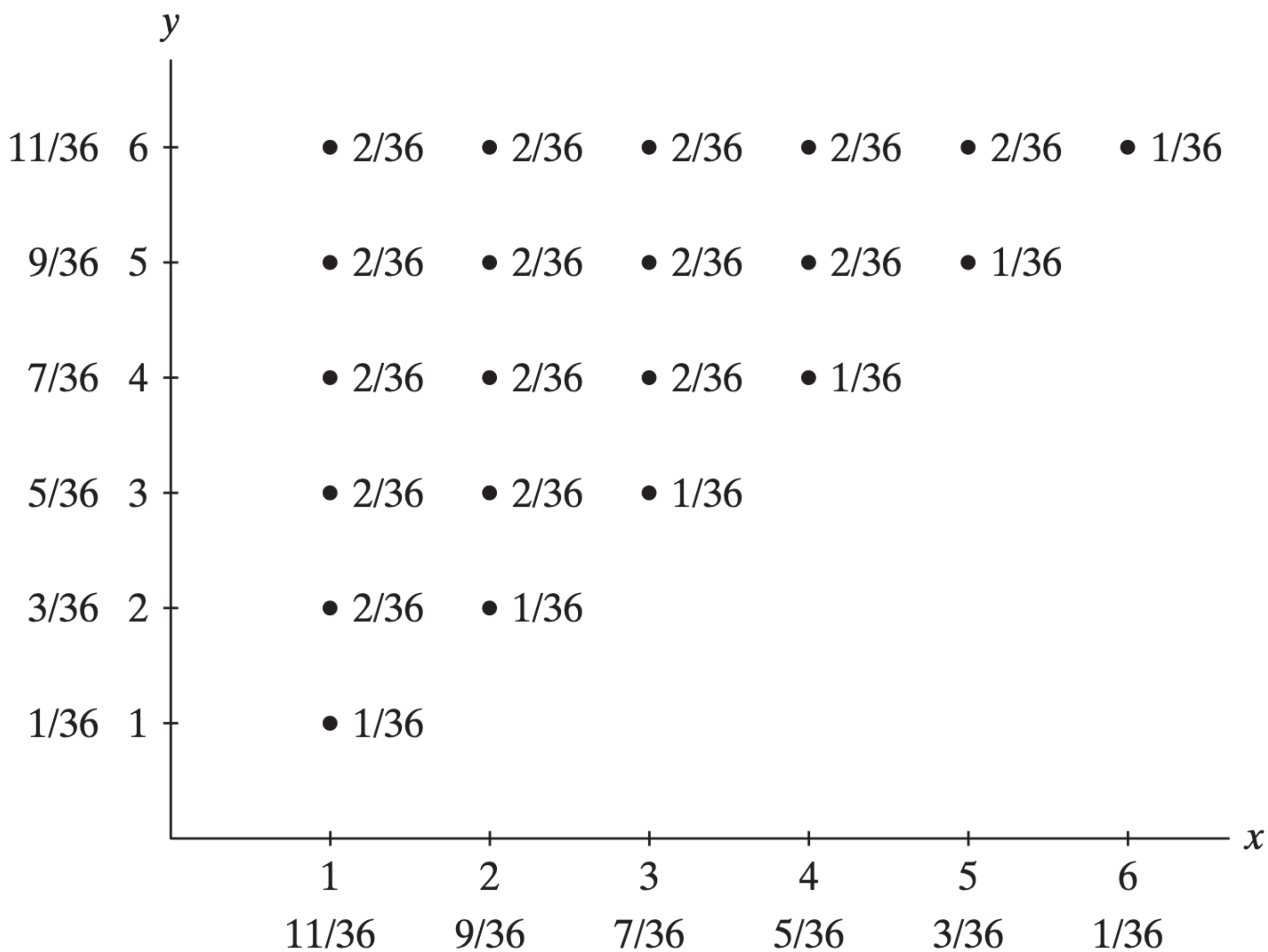
   c) What is the joint p.m.f. of X, Y?
   Solution: there is only one outcome $(X = a, Y = a)$ that gives rise to $\{X = Y = a\}$. For every other event $\{X \neq Y\}$, two outcomes do

the trick: (X, Y) and (Y, X). The joint p.m.f. is thus

$$f(x,y)=\begin{cases}1/36 & 1\leq x=y\leq 6\\ 2/36 & 1\leq x<y\leq 6\end{cases}$$

The first property is automatically satisfied, as is the third (by construction). There are only 6 outcomes for which $X = Y$, all the remaining outcomes (of which there are 15) have $X < Y$. Thus,

$$\sum_{x=1}^{6}\sum_{y=x}^{6}f(x,y)=6\cdot\frac{1}{36}+15\cdot\frac{2}{36}=1.$$


y
x
11/36 6
9/36 5
7/36 4
5/36 3
3/36 2
1/36 1
• 2/36 • 2/36 • 2/36 • 2/36 • 2/36 • 1/36
• 2/36 • 2/36 • 2/36 • 2/36 • 1/36
• 2/36 • 2/36 • 2/36 • 1/36
• 2/36 • 2/36 • 1/36
• 2/36 • 1/36
• 1/36
1 2 3 4 5 6
11/36 9/36 7/36 5/36 3/36 1/36

d) Compute $P(X=a)$ and $P(Y=b)$, for $a,b=1,\ldots,6$.
Solution: for every $a=1,\ldots,6$, the event $\{X=a\}$ corresponds to the following union of events:

$$\{X=a,Y=a\}\cup\{X=a,Y=a+1\}\cup\cdots\cup\{X=a,Y=6\}.$$

These events are mutually exclusive, so that

$$P(X=a)=\sum_{y=a}^{6}P(\{X=a,Y=y\})=\frac{1}{36}+\sum_{y=a+1}^{6}\frac{2}{36}$$
$$=\frac{1}{36}+\frac{2(6-a)}{36},\quad a=1,\ldots,6.$$

Similarly, we get $P(Y=b)=\frac{1}{36}+\frac{2(b-6)}{36}$, $b=1,\ldots,6$. These **marginal probabilities** can be found in the margins of the p.m.f.

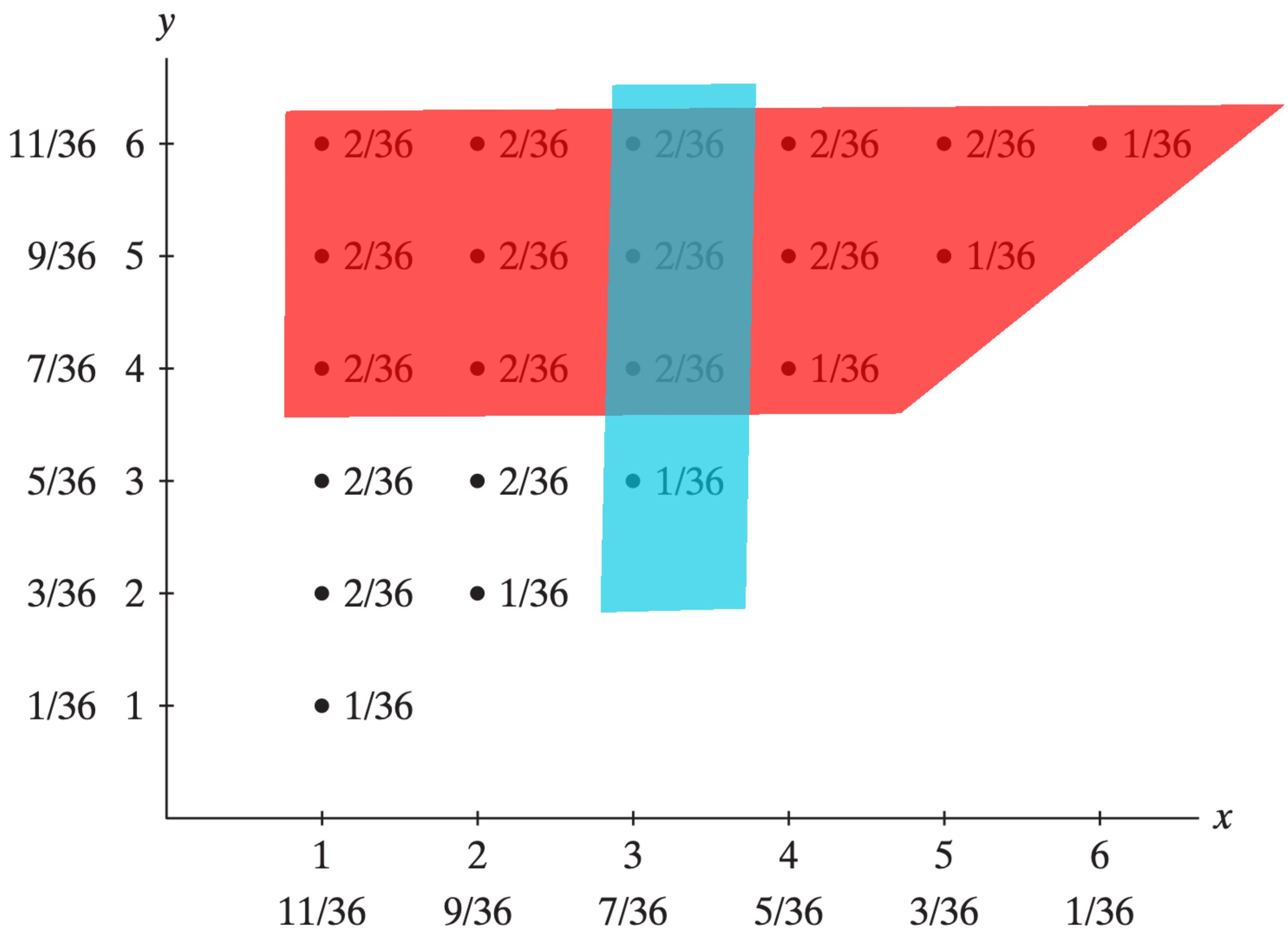
e) Compute $P(X = 3|Y > 3)$ and $P(Y \leq 3|X \geq 4)$.
Solution: the notation suggests how to compute these **conditional probabilities**:

$$P(X = 3|Y > 3) = \frac{P(X = 3 \cap Y > 3)}{P(Y > 3)}$$

The region corresponding to $P(Y > 3) = \frac{27}{36}$ is shaded in red (see next slide); the region corresponding to $P(X = 3) = \frac{7}{36}$ is shaded in blue.

The region corresponding to $P(X = 3 \cap Y > 3) = \frac{6}{36}$ is the intersection of the blue and the red regions, so

$$P(X = 3|Y > 3) = \frac{6/36}{27/36} = \frac{6}{27} \approx 0.2222.$$


y
11/36 6
9/36 5
7/36 4
5/36 3
3/36 2
1/36 1
2/36 2/36 2/36 2/36 2/36 1/36
2/36 2/36 2/36 2/36 1/36
2/36 2/36 2/36 1/36
2/36 2/36 1/36
2/36 1/36
1/36
x
1 2 3 4 5 6
11/36 9/36 7/36 5/36 3/36 1/36

Since $P(Y \leq 3 \cap X \geq 4) = 0$, $P(Y \leq 3 | X \geq 4) = 0$.

f) Are X and Y independent?
Solution: why don't we simply use the multiplicative rule to compute $P(X = 3 \cap Y > 3) = P(X = 3)P(Y > 3)$?

Well, we don't yet know if X and Y are **independent**, that is, we don't know if

$$P(X = x, Y = y) = P(X = x)P(Y = y) \quad \text{for all allowable } x, y.$$

As it is, $P(X = 1, Y = 1) = \frac{1}{36}$, but $P(X = 1)P(Y = 1) = \frac{11}{36} \cdot \frac{1}{36}$, so X and Y are **dependent** (this is often the case when the domain of the joint p.d.f./p.m.f. is not rectangular).

2. There are 8 similar chips in a bowl: three marked $(0,0)$, two marked $(1,0)$, two marked $(0,1)$ and one marked $(1,1)$. A player selects a chip at random and is given the sum of the two coordinates in dollars.

   a) What is the joint probability mass function of X_1, and X_2?
   Solution: let X_1 and X_2 represent the coordinates; we have

$$f(x_1,x_2)=\frac{3-x_1-x_2}{8},\quad x_1,x_2=0,1.$$

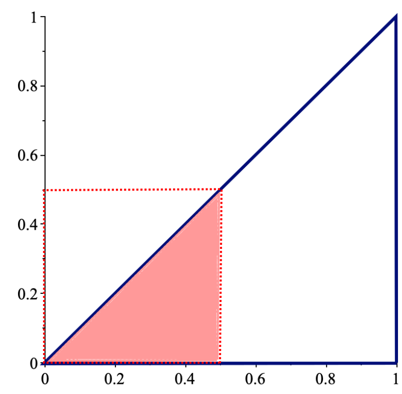
   a) What is the expected pay-off for this game?
   Solution: the pay-off is simply X_1+X_2. The expected pay-off is thus

$$\mathrm{E}[X_1+X_2]=\sum_{x_1=0}^{1}\sum_{x_2=1}^{0}(x_1+x_2)f(x_1,x_2)=0\cdot\frac{3}{8}+1\cdot\frac{2}{8}+1\cdot\frac{2}{8}+2\cdot\frac{1}{8}=0.75.$$

3. Let X and Y have joint p.d.f.

$$f(x,y) = 2, \quad 0 \leq y \leq x \leq 1.$$

a) What is the support of $f(x,y)$?
Solution: the support is the set $S = \{(x,y) : 0 \leq y \leq x \leq 1\}$, a triangle in the xy plane bounded by the $x-$axis, the line $y = 1$, and the line $y = x$. The support is the blue triangle shown below.



b) What is $P(0 \leq X \leq 0.5, 0 \leq Y \leq 0.5)$?
Solution: we need to evaluate the integral over the shaded area:

$$\begin{aligned} P(0 \leq X \leq 0.5, 0 \leq Y \leq 0.5) &= P(0 \leq X \leq 0.5, 0 \leq Y \leq X) \\ &= \int_0^{0.5} \int_0^x 2\, dydx = \int_0^{0.5} [2y]_{y=0}^{y=x}\, dx \\ &= \int_0^{0.5} 2x\, dx = 1/4. \end{aligned}$$

c) What are the marginal probabilities $P(X = x)$ and $P(Y = y)$?
Solution: we get

$$P(X = x) = \int_{-\infty}^{\infty} f(x, y)\, dy = \int_{y=0}^{y=x} 2\, dy = [2y]_{y=0}^{y=x} = 2x, \quad 0 \leq x \leq 1$$

and

$$P(Y=y)=\int_{-\infty}^{\infty} f(x,y)\,dx=\int_{x=y}^{x=1} 2\,dx$$
$$=[2x]_{x=y}^{x=1}=2-2y,\quad 0\leq y\leq 1.$$

d) Compute $\mathrm{E}[X]$, $\mathrm{E}[Y]$, and $\mathrm{E}[Y^2]$
Solution: we have

$$\mathrm{E}[X]=\iint_S xf(x,y)\,dA=\int_0^1\int_0^x 2x\,dydx=\int_0^1 [2xy]_{y=0}^{y=x}\,dx$$
$$=\int_0^1 2x^2\,dx=\left[\frac{2}{3}x^3\right]_0^1=\frac{2}{3};$$

$$
\begin{aligned}
\mathrm{E}[Y] &= \iint_S y f(x,y)\, dA = \int_0^1 \int_y^1 2y \, dx dy = \int_0^1 \left[2xy\right]_{x=y}^{x=1} \, dy \\
&= \int_0^1 (2y - 2y^2)\, dy = \left[y^2 - \frac{2}{3}y^3\right]_0^1 = \frac{1}{3}; \\
\mathrm{E}[Y^2] &= \iint_S y^2 f(x,y)\, dA = \int_0^1 \int_y^1 2y^2 \, dx dy = \int_0^1 \left[2xy^2\right]_{x=y}^{x=1} \, dy \\
&= \int_0^1 (2y - 2y^3)\, dy = \left[\frac{2}{3}y^3 - \frac{1}{2}y^4\right]_0^1 = \frac{1}{6}
\end{aligned}
$$

e) Are X and Y independent?
Solution: they are not independent as the support of the joint p.d.f. is not rectangular.

3.7 – Normal Approximation of the Binomial Distribution

If $X \sim \mathcal{B}(n,p)$ then we may interpret X as a sum of **independent and identically distributed** random variables

$$X = I_1 + I_2 + \cdots + I_n \text{ where each } I_i \sim \mathcal{B}(1,p)\,.$$

Thus, according to the **Central Limit Theorem** (more on this later), for large n, we have

$$\frac{X - np}{\sqrt{np(1-p)}} \overset{\text{approx}}{\sim} \mathcal{N}(0,1)\,,$$

i.e. for large n if $X \overset{\text{exact}}{\sim} \mathcal{B}(n,p)$ then $X \overset{\text{approx}}{\sim} \mathcal{N}(np, np(1-p))$.

Normal Approximation with Continuity Correction

Let $X \sim \mathcal{B}(n,p)$. Recall that $\mathrm{E}[X] = np$ and $\mathrm{Var}[X] = np(1-p)$.

If n is large, we may approximate X by a normal random variable in the following way:

$$P(X \leq x) = P(X < x + 0.5) = P\left(Z < \frac{x - np + 0.5}{\sqrt{np(1-p)}}\right)$$

and

$$P(X \geq x) = P(X > x - 0.5) = P\left(Z > \frac{x - np - 0.5}{\sqrt{np(1-p)}}\right).$$

Example: suppose $X \sim \mathcal{B}(36, 0.5)$. Provide a normal approximation to the probability $P(X \leq 12)$. *Note:* For $n = 36$ the binomial probabilities are not available in the textbook tables.

Solution: since $\mathrm{E}[X] = 36 \times 0.5 = 18$ and $\mathrm{Var}[X] = 36 \times 0.5 \times 0.5 = 9$,

$$P(X \leq 12) = P\left(\frac{X-18}{3} \leq \frac{12-18+0.5}{3}\right)$$

$$\overset{\text{norm.approx'n}}{\approx} \Phi(-1.83) \overset{\text{table}}{\approx} 0.033\,.$$

Compare this to the R value of $\texttt{pbinom}(12, 36, 0.5) = 0.0326$.

Computing Binomial Probabilities

We thus have at least 3 ways to compute (or approximate) binomial probabilities:

- Use the exact formula: if $X \sim \mathcal{B}(n,p)$ then for each $x = 0, 1, \ldots, n$, $P(X=x) = \binom{n}{x} p^x (1-p)^{n-x}$;
- Use tables: if $n \leq 15$ and p is one of 0.1, 0.2,. . . , 0.9, then the CDF is in the textbook (must express desired probability in terms of CDF, i.e. in form $P(X \leq x)$ first), i.e.

$$P(X<3) = P(X \leq 2); \qquad P(X=7) = P(X \leq 7) - P(X \leq 6)\,;$$
$$P(X>7) = 1 - P(X \leq 7); \qquad P(X \geq 5) = 1 - P(X \leq 4) \text{ etc.}$$

- Use normal approximation: the suggested "rule of thumb " in the binomial case is: if np and $n(1-p)$ are both ≥ 5, the normal approximation $X \sim \mathcal{N}(np, np(1-p))$

$$P(X \leq x) \approx \Phi\left(\frac{x + 0.5 - np}{\sqrt{np(1-p)}}\right)$$

$$P(X \geq x) \approx 1 - \Phi\left(\frac{x - 0.5 - np}{\sqrt{np(1-p)}}\right)$$

for $x = 0, 1, \ldots, n$ should provide a decent approximation.

Appendix – Summary

X	**Example**	$f(x)$	**Domain**	$\mathrm{E}[X]$	$\mathrm{Var}[X]$
Uniform	Select a point at random from $[a,b]$	$\frac{1}{b-a}$	$a \leq x \leq b$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
Normal	Meas. errors; children heights; breaking strengths, etc.	$\frac{\exp(-(x-\mu)^2/2\sigma^2)}{\sigma\sqrt{2\pi}}$	$-\infty < x < \infty$	μ	σ^2

Summary

X	**Example**	$f(x)$	**Domain**	$\mathrm{E}[X]$	$\mathrm{Var}[X]$
Exponential	Waiting time to first arrival in a Poisson process with rate λ	$\lambda e^{-\lambda x}$	$0 \leq x < \infty$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
Gamma	Waiting time to rth arrival in a Poisson process with rate λ	$\frac{x^{r-1}}{(r-1)!}\lambda^r e^{-\lambda x}$	$0 \leq x < \infty$	$\frac{r}{\lambda}$	$\frac{r}{\lambda^2}$