# Actuary Probability I

## Random Variables II

Irving Gómez Méndez

August-December, 2021

UNIVERSIDAD IBEROAMERICANA
CIUDAD DE MÉXICO ®

# Introduction

# Introduction

We considered discrete random variable–that is, random variables whose set of possible values is either finite or countably infinite. However, there also exist random variables whose set of possible values is uncountable. Two examples are the time that a train arrives at a specified stop and the lifetime of a transistor. Let X be such a random variable.

In general, we say that a random variable $X$ is continuous if its distribution function is continuous. Since $\mathbb{P}(X = x) = F(x) - F(x^-)$, being $F$continuous. Then $X$ is continuous if $\mathbb{P}(X = x) = 0$ for all $x \in \mathbb{R}$.

# Example

Consider the experiment of choosing a random point in a disc $D$ of radius $R$ with center in the origin. We interpret the expression "at random" as "if $A$ and $B$ are subsets of the disc with equal area and $\omega$ is a point chosen at random then $\mathbb{P}(\omega \in A) = \mathbb{P}(\omega \in B)$". We conclude that, the probability of a point chosen in a subset $A$ of the disc is proportional to the area of $A$:

$$\mathbb{P}(\omega \in A) = C|A|,$$

where $C$ is a constant and $|A|$ is the area of $A$. Because

$$\mathbb{P}(\omega \in C) = 1 = C|D|$$

we have that

$$C = \frac{1}{|D|} \quad \text{and} \quad \mathbb{P}(\omega \in A) = \frac{|A|}{|D|}$$

In particular, consider the case where $A$ is a disc, and define over the space $D$ the variable $X$ as the distance from the chosen point to the origin. If $0 \leq x \leq R$, the event

$$\{\omega, \text{ s.t. } X(\omega) \leq x\}$$

is the disc center in the origin of radius $x$. Its area is $\pi x^2$. Hence,

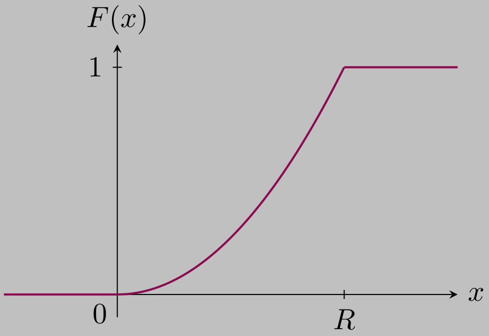
$$F(x) = \mathbb{P}(X \leq x) = \frac{\pi x^2}{\pi R^2} = \frac{x^2}{R^2}, \quad 0 \leq x \leq R.$$

Moreover, if $x < 0$ then $\mathbb{P}(X \leq x) = 0$ and if $x > R$ then $\mathbb{P}(X \leq x) = 1$.

Thus,

$$F(x) = \begin{cases} 0 & \text{if } x < 0, \\ \frac{x^2}{R^2} & \text{if } 0 \leq x \leq R, \\ 1 & \text{if } x > R, \end{cases}$$

which is a continuous function, so $X$ is a continuous random variable. The graph of $F$ is



# Density Function

# Density Function

Let $X$ be a random variable with distribution function $F$. We say that $F$ has density (or is absolutely continuous), if exists a non-negative function $f$ such that

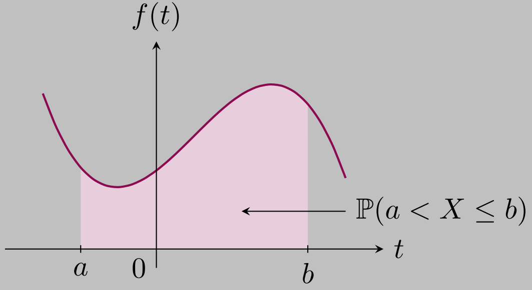
$$F(x) = \int_{-\infty}^{x} f(t)dt \quad \text{for all } x \in \mathbb{R}.$$

The function $f$ is called the density of the distribution function or of the distribution or of the random variable.
Because $\lim_{x\to\infty} F(x) = 1$, we have that $\int_{-\infty}^{\infty} f(t)dt = 1$.

We also have that

$$\mathbb{P}(a < X \leq b) = \mathbb{P}(X \leq b) - \mathbb{P}(X \leq a) = F(b) - F(a) = \int_a^b f(t)dt.$$

Thus, the probability that the random variable $X$ belongs to the interval $(a, b]$ is the area between the graph of the function $f$, the x-axis and the verticals $a$ and $b$.




In general, if $B$ is any set of real numbers,

$$\mathbb{P}(X \in B) = \int_B f(t)dt.$$

Remember that we defined the density function $f$ of the random variable $X$ with distribution function $F$, as the non-negative function such that

$$F(x) = \int_{-\infty}^{x} f(t)dt.$$

Differentiating both sides of the equation yields

$$\frac{d}{dx}F(x) = f(x),$$

that is, the density is the derivative of the cumulative distribution function.

A some-what more intuitive interpretation of the density function may be obtained from as follows:

$$\mathbb{P}\left[a-\frac{\varepsilon}{2}<X\leq a+\frac{\varepsilon}{2}\right]=\int_{a-\frac{\varepsilon}{2}}^{a+\frac{\varepsilon}{2}} f(t)dt \approx \varepsilon f(a)$$

if $\varepsilon$ is small and $f$ is continuous at $a$.

In other words, the probability that $X$ will be contained in an interval of length $\varepsilon$ around the point $a$ is approximately $\varepsilon f(a)$. From this result we see that $f(a)$ is a measure of how likely it is that the random variable will be near $a$.

# Examples

1. Suppose that $X$ is a continuous random variable whose probability density function is given by

$$f(x) = \begin{cases} C(4x - 2x^2) & \text{for } 0 < x < 2, \\ 0 & \text{otherwise.} \end{cases}$$

   (a) What is the value of $C$?
   (b) Find $\mathbb{P}(X > 1)$.

(a). Since $f$ is a probability density function, we must have

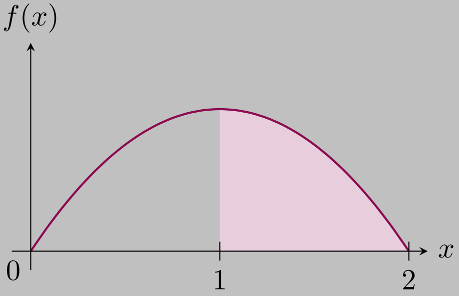
$$C \int_0^2 (4x - 2x^2) dx = 1$$

$$\Leftrightarrow C \left[ 2x^2 - \frac{2x^3}{3} \right]_{x=0}^{x=2} = 1$$

$$\Leftrightarrow C = \frac{3}{8}$$

(b).

$$\mathbb{P}(X > 1) = \int_1^\infty f(x)dx = \frac{3}{8}\int_1^2 (4x - 2x^2) = \frac{1}{2}$$



2. The amount of time in hours that a computer functions before breaking down is a continuous random variable with probability density function given by

$$f(x) = \begin{cases} \lambda e^{-x|100} & \text{for } x \geq 0, \\ 0 & \text{otherwise.} \end{cases}$$

What is the probability that

(a) a computer will function between 50 and 150 hours before breaking down?

(b) it will function for fewer than 100 hours?

(a). Since

$$1 = \int_0^\infty \lambda e^{-x/100} dx = \lambda \int_0^\infty e^{-x/100} dx$$

we obtain

$$1 = -\lambda(100)e^{x/\lambda}|_0^\infty = 100\lambda \Leftrightarrow \lambda = \frac{1}{100}.$$

Hence, the probability that a computer will function between 50 and 150 hours before breaking down is given by

$$\begin{aligned}\mathbb{P}(50 < X < 150) &= \int_{50}^{150} \frac{1}{100} e^{-x/100} dx = -e^{-x/100}|_{50}^{150} \\ &= e^{-1/2} - e^{-3/2} \approx 0.384\end{aligned}$$

(b). Similarly,

$$\mathbb{P}(X < 100) = \int_0^{100} \frac{1}{100} e^{-x/100} dx = -e^{-x/100}|_0^{100} = 1 - e^{-1} \approx 0.633$$

3. The lifetime in hours of a certain kind of radio tube is a random variable having a probability density function given by

$$f(x) = \begin{cases} 0 & x \leq 100, \\ \frac{100}{x^2} & x > 100. \end{cases}$$

What is the probability that exactly 2 of 5 such tubes in a radio set will have to be replaced within the first 150 hours of operation? Assume that the events $E_i$, $i = 1, 2, 3, 4, 5$, that the $i$th such tube will have to be replaced within this time are independent.

From the statement of the problem, we have

$$\mathbb{P}(E_i) = \int_0^{150} f(x)dx = 100\int_{100}^{150} \frac{1}{x^2}dx = \frac{1}{3}.$$

Hence, from the independence of the events $E_i$ , it follows that the desired probability is

$$\binom{5}{2}\left(\frac{1}{3}\right)^2\left(\frac{2}{3}\right)^3 = \frac{80}{243}.$$

# Uniform, Normal, Exponential and Laplace Distributions

# Uniform Distribution

A random variable $X$ has uniform distribution in the interval $[a, b]$ if for any interval $I$ contained in $[a, b]$, $\mathbb{P}(X \in I)$ is proportional to the length of $I$. This is denoted as $X \sim \mathcal{U}[a, b]$. We can calculate the distribution function of $X$,

$$F(x) = \mathbb{P}(X \in [a, x]) = K(x - a)$$

where $K$ is the constant of proportionality. Because $F(b) = \mathbb{P}(X \in [a, b]) = 1$ we have

$$K(b - a) = 1 \Leftrightarrow K = \frac{1}{b - a}.$$

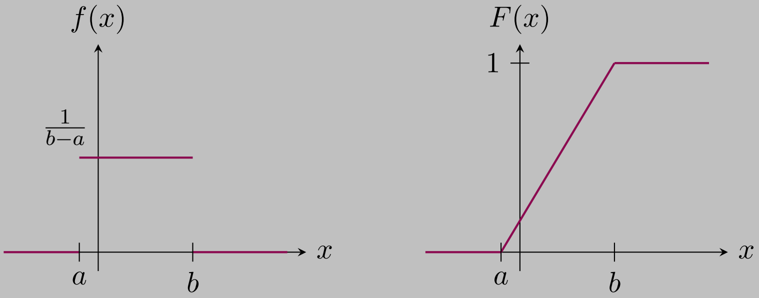
Therefore, the distribution function of $X$ is

$$F(x) = \frac{x - a}{b - a}\mathbb{1}_{a \leq x \leq b} + 1\mathbb{1}_{x > b}$$

The density of this function is

$$f(x) = \frac{1}{b-a}\mathbb{1}_{a \leq x \leq b}$$

since we can verify that $F(x) = \int_{-\infty}^{x} f(t)dt$ for all $x \in \mathbb{R}$.



# Example

Between 7 a.m. and 8 a.m. the trains departure from a certain station every 10 minutes beginning at 7:03. Find the probability that a person that arrives to the station has to wait less than 2 minutes for the train if the arrival of the person follows a uniform distribution in the interval:

(a) from 7 a.m. to 8 a.m.

(b) from 7:15 a.m. to 7:30 a.m.

Note that to wait less than 2 minutes, a person must arrive to the station in an interval of the form $(t-2, t)$ where $t$ is one of the moments where the train departures.

(a). In the first case, the intervals of interest are

$$(7:01, 7:03), \quad (7:11, 7:13), \quad (7:21, 7:23),$$
$$(7:31, 7:33), \quad (7:41, 7:43), \quad (7:51, 7:53).$$

Let $B$ be the union of this intervals. We know that the distribution of the arrival is uniform in $[7:00, 8:00]$ and we want to find the probability of $\{X \in B\}$. Because the total length of $B$ is 12 minutes we have

$$\mathbb{P}(X \in B) = \frac{\text{length of } B}{60} = \frac{12}{60} = \frac{1}{5}.$$

(b). In the second case, we have that $B = (7.21, : 23)$ so

$$\mathbb{P}(X \in B) = \frac{2}{15}.$$

# Normal Distribution

A random variable $Y$ is said to have or follow a normal distribution with parameters $\mu$ and $\sigma^2$ (whose meaning we are going to see later), if its density function is given by

$$f(y) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(y-\mu)^2}{2\sigma^2}\right\} \mathbb{1}_{y\in\mathbb{R}}.$$

Since this is a non-negative function, to see that it is effectively a density of probability we have to prove that

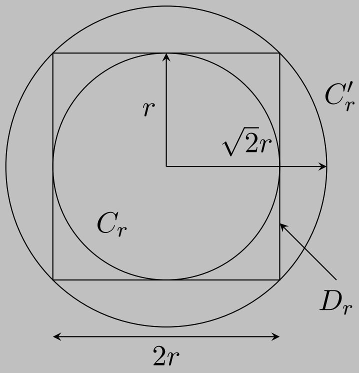
$$\int_{-\infty}^{\infty} f(y)dy = 1,$$

considering the change of variable $z = (y-\mu)/2\sigma$ we have

$$\begin{aligned}\int_{-\infty}^{\infty} f(y)dy &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-z^2} \sigma\sqrt{2}dz \\ &= \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} e^{-z^2} dz\end{aligned}$$

One way to calculate this last integral is as follows.

Let $C_r$ be the disc with center in the origin and radius $r$ and $C'_r$ the disc with the same center and radius $\sqrt{2}r$. Let $D_r$ be the square with center in the origin and side $2r$.



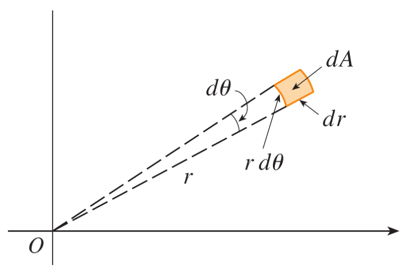
Given that the common integrand of the following integrals is non-negative, we have

$$\int\int_{C_r} e^{-(x^2+y^2)}dxdy \leq \int\int_{D_r} e^{-(x^2+y^2)}dxdy \leq \int\int_{C'_r} e^{-(x^2+y^2)}dxdy,$$

also

$$\int\int_{D_r} e^{-(x^2+y^2)}dxdy = \int_{-r}^{r} e^{-x^2}dx \int_{-r}^{r} e^{-y^2}dy = \left(\int_{-r}^{r} e^{-x^2}dx\right)^2.$$

Consider now the first integral of the inequalities. Changing to polar coordinates $\rho, \theta$ through the transformation $x = \rho\cos\theta$ and $y = \rho\cos\theta$, we have




De forma análoga, cambiando r por 2r resulta

$$
\begin{aligned}
\int\int_{C_r} e^{-(x^2+y^2)} dxdy &= \int_0^{2\pi} d\theta \int_0^r e^{-\rho^2} \rho d\rho \\
&= 2\pi \left[ -\frac{1}{2} e^{-\rho^2} \right]_0^r \\
&= 2\pi \left[ \frac{1}{2} \left( 1 - e^{-r^2} \right) \right] \\
&= \pi \left( 1 - e^{-r^2} \right).
\end{aligned}
$$

Analogously, changing $r$ by $\sqrt{2}r$, we have

$$\int\int_{C'_r} e^{-(x^2+y^2)}dxdy = \pi\left(1 - e^{-2r^2}\right).$$

Replacing these quantities in the previous inequalities,

$$\pi\left(1 - e^{-r^2}\right) \leq \left(\int_{-r}^{r} e^{-x^2}dx\right)^2 \leq \pi\left(1 - e^{-2r^2}\right),$$

letting $r \to \infty$

$$\pi \leq \left(\int_{-\infty}^{\infty} e^{-x^2}dx\right)^2 \leq \pi,$$

thus

$$\int_{-\infty}^{\infty} e^{-x^2}dx = \sqrt{\pi}.$$

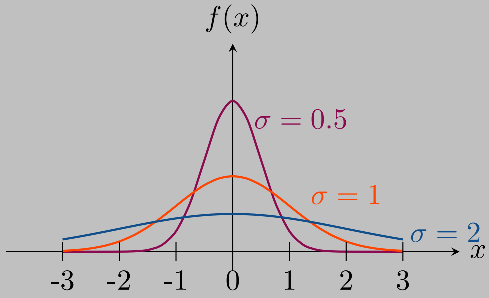
Finally, we have

$$\int_{-\infty}^{\infty} f(y)dy = \frac{1}{\sqrt{\pi}}\int_{-\infty}^{\infty} e^{-z^2}dz = 1.$$

If $Y$ has a normal distribution with parameters $\mu$ and $\sigma^2$, denoted as $Y \sim \mathcal{N}(\mu, \sigma^2)$. If $X \sim \mathcal{N}(0, 1)$ it is said that $X$ has a standard normal distribution, whose density is usually denoted as $\phi$, so

$$\phi(x) = \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{x^2}{2}\right\} \mathbb{1}_{x \in \mathbb{R}}.$$

The distribution function corresponding to the density $\phi$ is usually denoted as $\Phi$. This function does not have an explicit formula and must be calculated numerically.



Because $\phi$ is symmetric with respect to the origin, we have that

$$\begin{aligned}\Phi(-x) &= \int_{-\infty}^{-x} \phi(x)dx = \int_{x}^{\infty} \phi(x)dx \\ &= \int_{-\infty}^{\infty} \phi(x)dx - \int_{-\infty}^{x} \phi(x)dx \\ &= 1 - \Phi(x),\end{aligned}$$

so, for any value of $x$ we have that $\Phi(-x) = 1 - \Phi(x)$, this formula allows us to find the value of $\Phi(-x)$ from the value of $\Phi(x)$. Hence, it is enough to know the values of $\Phi(x)$ for $x \geq 0$.

Let $X \sim \mathcal{N}(0,1)$ and consider $Y = \mu + \sigma X$ with $\sigma > 0$. Let $F_Y$ denote the distribution function of $Y$, then

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

By differentiation, the density function of $Y$ is then

$$
\begin{aligned}
f_Y(x) &= \frac{1}{\sigma}\phi\left(\frac{x-\mu}{\sigma}\right) \\
&= \frac{1}{\sqrt{2\pi}\sigma}\exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}\mathbb{1}_{x\in\mathbb{R}}.
\end{aligned}
$$

which shows that $Y \sim \mathcal{N}(\mu, \sigma)$

# Examples

1. If $X$ is a normal random variable with parameters $\mu = 3$ and $\sigma^2 = 9$, find (a) $\mathbb{P}(2 < X < 5)$, (b) $\mathbb{P}(X > 0)$ and (c) $\mathbb{P}(|X - 3| > 6)$.

   (a).

$$\begin{aligned}\mathbb{P}(2 < X < 5) &= \mathbb{P}\left(\frac{2-3}{3} < \frac{X-3}{3} < \frac{5-3}{3}\right)\\ &= \mathbb{P}\left(-\frac{1}{3} < Z < \frac{2}{3}\right)\\ &= \Phi\left(\frac{2}{3}\right) - \Phi\left(-\frac{1}{3}\right)\\ &= \Phi\left(\frac{2}{3}\right) - \left[1 - \Phi\left(\frac{1}{3}\right)\right] \approx 0.378\end{aligned}$$

(b).

$$\begin{aligned}\mathbb{P}(X>0) &= \mathbb{P}\left(\frac{X-3}{3} > \frac{0-3}{3}\right)\\ &= \mathbb{P}(Z>-1)\\ &= 1-\Phi(-1)\\ &= \Phi(1) \approx 0.841\end{aligned}$$

(c).

$$\begin{aligned}\mathbb{P}(|X-3|>6) &= \mathbb{P}(X>9)+\mathbb{P}(X<-3)\\ &= \mathbb{P}\left(\frac{X-3}{3} > \frac{9-3}{3}\right)+\mathbb{P}\left(\frac{X-3}{3} < \frac{-3-3}{3}\right)\\ &= \mathbb{P}(Z>2)+\mathbb{P}(Z<-2)\\ &= 1-\Phi(2)+\Phi(-2)\\ &= 2[1-\Phi(2)] \approx 0.046\end{aligned}$$

# Normal Approximation to Binomial Distribution

## De Moivre - Laplace Limit Theorem

If $S_n \sim \text{Bin}(np)$ for $n \geq 1$. Define $q = 1 - p$ and

$$T_n = \frac{S_n - np}{\sqrt{npq}}.$$

Then for all $x \in \mathbb{R}$,

$$\mathbb{P}(T_n \leq x) \to \Phi(x)$$

This result was proved originally for the special case of $p = 12$ by A. de Moivre in 1733 and was then extended to general $p$ by P. S. Laplace in 1812. It represents the first version of the central limit theorem presented by P. S. Laplace and proved rigorously by A. M. Lyapunov in the period 1901–1902.

# Examples

1. Let $X$ be the number of times that a fair coin that is flipped 40 times lands on heads. Find the probability that $X = 20$. Use the normal approximation and compare it with the exact solution.

   To employ the normal approximation, note that because the binomial is a discrete integer-valued random variable, whereas the normal is a continuous random variable, it is best to write $\mathbb{P}(X = i)$ as $\mathbb{P}(i - 1/2 < X < i + 1/2)$ before applying the normal approximation (this is called the continuity correction).

Therefore

$$\begin{aligned}\mathbb{P}(X = 20) &= \mathbb{P}(19.5 < X < 20.5) \\ &= \mathbb{P}\left(\frac{19.5 - 20}{\sqrt{10}} < \frac{X - 20}{\sqrt{10}} < \frac{20.5 - 20}{\sqrt{10}}\right) \\ &\approx \mathbb{P}(-0.158 < Z < 0.158) \\ &\approx \Phi(0.158) - \Phi(-0.158) \approx 0.1255.\end{aligned}$$

The exact result is

$$\mathbb{P}(X = 20) = \binom{40}{20}\left(\frac{1}{2}\right)^4 0 \approx 0.1254.$$

2. To determine the effectiveness of a certain diet in reducing the amount of cholesterol in the bloodstream, 100 people are put on the diet. After they have been on the diet for a sufficient length of time, their cholesterol count will be taken. The nutritionist running this experiment has decided to endorse the diet if at least 65 percent of the people have a lower cholesterol count after going on the diet. What is the probability that the nutritionist endorses the new diet if, in fact, it has no effect on the cholesterol level?

Let us assume that if the diet has no effect on the cholesterol count, then, strictly by chance, each person's count will be lower than it was before the diet with probability 1/2 . Hence, if $X$ is the number of people whose count is lowered, then the probability that the nutritionist will endorse the diet when it actually has no effect on the cholesterol count is

$$\sum_{i=65}^{1} 00 \binom{100}{i} \left(\frac{1}{2}\right)^{100} = \mathbb{P}(X \geq 65)$$

$$= \mathbb{P}\left(\frac{X-50}{\sqrt{25}} \geq \frac{65-50}{\sqrt{25}}\right)$$

$$\approx \mathbb{P}(Z \geq 3)$$

$$1 - \Phi(3) \approx 0.0014$$

3. Fifty-two percent of the residents of New York City are in favor of outlawing cigarette smoking in publicly owned areas. Approximate the probability that more than 50 percent of a random sample of n people from New York are in favor of this prohibition when (a) $n = 11$, (b) $n = 101$, (c) $n = 1001$. How large would n have to be to make this probability exceed 0.95?

Let $N$ denote the number of residents of New York City. To answer the preceding question, we must first understand that a random sample of size $n$ is a sample such that the $n$ people were chosen in such a manner that each of the $\binom{N}{n}$ subsets of $n$ people had the same chance of being the chosen subset. Consequently, $S_n$ , the number of people in the sample who are in favor of the smoking prohibition, is a hypergeometric random variable.

But because $N$ and $0.52N$ are both large in comparison with the sample size $n$, it follows from the binomial approximation to the hypergeometric that the distribution of $S_n$ is closely approximated by a binomial distribution with parameters $n$ and $p = 0.52$. The normal approximation to the binomial distribution then shows that

$$\begin{aligned}\mathbb{P}(S_n > 0.5n) &= \mathbb{P}\left(\frac{S_n - 0.52n}{\sqrt{n(0.52)(0.48)}} > \frac{0.5n - 0.52n}{\sqrt{n(0.52)(0.48)}}\right)\\ &= \mathbb{P}(Z > -0.04\sqrt{n})\\ &\approx \Phi(0.04\sqrt{n}).\end{aligned}$$

Thus,

$$\mathbb{P}(S_n > 0.5n) \approx \begin{cases} \Phi(0.1327) \approx 0.5528 & \text{if } n = 11,\\ \Phi(0.4020) \approx 0.6562 & \text{if } n = 101,\\ \Phi(1.2655) \approx 0.8972 & \text{if } n = 1001.\end{cases}$$

In order for this probability to be at least 0.95, we would need $\Phi(0.04\sqrt{n}) > 0.95$. Because $\Phi(x)$ is an increasing function and $\Phi((1.645)) = 0.95$, this means that

$$0.04n > 1.645 \Leftrightarrow n > 1691.266.$$

That is, the sample size would have to be at least 1692.

# Exponential Distribution

We say that a random variable $X$ has exponential distribution if

$$\mathbb{P}(X > x) = e^{-\lambda x} \quad (x \geq 0)$$

where $\lambda > 0$. Thus, its distribution function is

$$F(x) = \mathbb{P}(X \leq x) = 1 - \mathbb{P}(X > x) = \begin{cases} 0 & \text{if } x < 0, \\ 1 - e^{-\lambda x} & \text{if } x \geq 0. \end{cases}$$

and the density function of this distribution is

$$f(x) = \begin{cases} 0 & \text{if } x < 0 \\ \lambda e^{-\lambda x} & \text{if } x \geq 0. \end{cases}$$

We use the notation $X \sim \text{Exp}(\lambda)$.

An important property of the exponential distribution is that for $a$ and $b$ non-negative

$$\mathbb{P}(X > a + b) = \mathbb{P}(X > a)\mathbb{P}(X > b).$$

We can verify this property from the definition of the distribution. An equivalent way to write this property is

$$\mathbb{P}(X > a + b | X > a) = \mathbb{P}(X > b), \quad a \geq 0, \quad b \geq b,$$

which is known as the "memoryless" property of the exponential function or the property of a lifetime product that never gets old.

# Examples

1. if we assume that the decay rate of a mass $m$ of a radioactive material is proportional to the quantity of the material at time $t$, then $m$ satisfies the equation

   $$\frac{dm}{dt} = -\lambda m$$

   where $\lambda$ is a constant that depends on the material. The solution to this equation is

   $$m = m_0 e^{-\lambda t},$$

   where $m_0$ is the quantity of the material at time $t = 0$.

The proportion of material that has decay in $t$ units of time is given by $(m_0 - m)/m_0$, which can be interpreted as the probability that an atom selected at random from the original material decays in a period of time $t$. If $X$ represents the lifetime of this atom, then

$$F(x) = \mathbb{P}(X \leq t) = \frac{m_0 - m}{m_0} = 1 - e^{-\lambda t},$$

so $X \sim \text{Exp}(\lambda)$.

# Laplace Distribution

A random variable $X$ that can take either positive or negative and whose absolute value is exponentially distributed with parameter $\lambda$, $\lambda > 0$ is said to have a Laplace distribution, its density and distribution function are given by

$$f(x) = \frac{1}{2}e^{-\lambda|x|}\mathbb{1}_{x\in\mathbb{R}}$$

and

$$F(x) = \frac{1}{2}e^{-\lambda x}\mathbb{1}_{x<0} + \left[1 - \frac{1}{2}e^{-\lambda x}\right]\mathbb{1}_{x\geq 0},$$

respectively.

# Survivor Function

# Introduction

In the lifetime analysis we consider situations in which the time to the occurrence of some event is of interest for individuals in some population. Sometimes the events are actual deaths of individuals and “lifetime” is the length of life measured from some particular starting point. In other instances “lifetime” and the words “death” or “failure,” which denote the event of interest, are used in a figurative sense.

# Survivor Function for Continuous Models

Let $T$ be a non-negative continuous random variable representing the lifetime of individuals in some population. Let $f(t)$denote the probability density function of $T$ The probability of an individual surviving to time $t$ is given by the survivor function

$$S(t) = \mathbb{P}(T \geq t) = \int_t^\infty f(x)dx.$$

In some contexts, $S(t)$ is referred to as the reliability function.

- $S(t)$ is a monotone decreasing function continuous.
- $S(0) = 1$ and $\lim_{t\to\infty} S(t) = 0$.

# Hazard Function for Continuous Models

The hazard function $h(t)$ is defined as

$$\begin{aligned} h(t) &= \lim_{\Delta t \to 0} \frac{\mathbb{P}(t \leq T < t + \Delta t | T \geq t)}{\Delta t} \\ &= \frac{f(t)}{S(t)}. \end{aligned}$$

The hazard function specifies the instantaneous rate of death or failure at time $t$, given that the individual survives up to $t$; $h(t)\Delta t$ is the approximate probability of death in $[t, t + \Delta t)$ given survival up to $t$. The hazard function is sometimes called hazard rate or force of mortality.

# Relation between Functions

- $f(t) = -\frac{d}{dt}S(t)$.
- $h(t) = -\frac{d}{dt}\log S(t)$.
- $S(t) = \exp\left\{-\int_0^t h(x)dx\right\}$.

We define the cumulative hazard function

$$H(t) = \int_0^t h(x)dx.$$

- $S(t) = \exp[-H(t)]$.
- $\lim_{t\to\infty} H(t) = \infty$.
- $f(t) = h(t)S(t) = h(t)\exp\left\{-\int_0^t h(x)dx\right\}$.

# Survivor Function for Discrete Models

Sometimes, for example, when lifetime are grouped or measured as a number of cycles or some sort, $T$ may be treated as a discrete random variable. Suppose $T$ can take the values $t_1, t_2, \ldots$ with $0 \leq t_1 < t_2 < \cdots$, and let the probability function be

$$f(t_j) = \mathbb{P}(T = t_j), \quad j = 1, 2, \ldots.$$

The survivor function is then

$$S(t) = \mathbb{P}(T \geq t) = \sum_{j : t_j \geq t} f(t_j).$$

When considered as a function for all $t \geq 0$, $S(t)$ is a left-continuous, non.increasing step function with $S(0) = 1$ and $\lim_{t \to \infty} S(t) = 0$.

# Hazard Function for Discrete Models

The discrete hazard function is defined as

$$\begin{aligned} h(t_j) &= \mathbb{P}(T = t_j | T \geq t_j) \\ &= \frac{f(t_j)}{S(t_j)}, \quad j = 1, 2, \ldots \end{aligned}$$

Since $f(t_j) = S(t_j) - S(t_{j+1})$,

$$h(t_j) = 1 - \frac{S(t_{j+1})}{S(t_j)}, \quad j = 1, 2, \ldots$$

and

$$S(t) = \prod_{j:t_j<t} [1 - h(t_j)].$$

The analog of the continuous $H(t)$ is defined as

$$H(t) = \sum_{j:t_j<t} h(t_j).$$

# Example

One often hears that the death rate of a person who smokes is, at each age, twice that of a nonsmoker. What does this mean? Does it mean that a nonsmoker has twice the probability of surviving a given number of years as does a smoker of the same age?

If $h(t)$ denotes the hazard rate of a smoker of age $t$ and $h_s(t)$ that of a nonsmoker of age $t$, then the statement at issue is equivalent to the statement that $h_s(t) = 2h_n(t)$. The probability that an $A$-year-old nonsmoker will survive until age $B$, $A < B$, is

$$
\begin{aligned}
&\mathbb{P}(A\text{-year-old nonsmoker reaches age } B) \\
&\quad = \mathbb{P}(\text{nonsmoker's lifetime } > B | \text{nonsmoker's lifetime } > A) \\
&\quad = \frac{S_{\text{non}}(B)}{S_{\text{non}}(A)} \\
&\quad = \frac{\exp\left\{-\int_0^B h_n(t)dt\right\}}{\exp\left\{-\int_0^A h_n(t)dt\right\}} \\
&\quad = \exp\left\{-\int_A^B h_n(t)dt\right\},
\end{aligned}
$$

whereas the corresponding probability for a smoker is, by the same reasoning,

$$
\begin{aligned}
&\mathbb{P}(A\text{-year-old smoker reaches age } B)\\
&= \exp\left\{-\int_A^B h_s(t)dt\right\}\\
&= \exp\left\{-2\int_A^B h_n(t)dt\right\}\\
&= \left[\exp\left\{-\int_A^B h_n(t)dt\right\}\right]^2.
\end{aligned}
$$

In other words, for two people of the same age, one of whom is a smoker and the other a nonsmoker, the probability that the smoker survives to any given age is the square (not one-half) of the corresponding probability for a nonsmoker. For instance, if $h_n(t) = 1/30$, $50 \leq t \leq 60$, then the probability that a 50-year-old nonsmoker reaches age 60 is $e^{-1/3} \approx 0.7165$, whereas the corresponding probability for a smoker is $e^{-2/3} \approx 0.5134$.

# Exponential, Rayleigh, Gamma, Erlang, $\chi^2$, Weibull, Log-normal, Cauchy and Beta Distributions

# Exponential and Rayleigh Distributions

The exponential distribution is characterized by a constant hazard function

$$h(t) = \lambda, \quad t \geq 0,$$

where $\lambda > 0$. The survivor function is

$$S(t) = e^{-\lambda t}.$$

If a random variable has a linear hazard rate function–that is, if

$$h(t) = a + bt, \quad t \geq 0$$

then its survivor and distribution functions are given by

$$S(t) = e^{-at-bt^2/2}, \quad F(t) = 1 - e^{-at-bt^2/2}$$

and differentiation yields its density, namely,

$$f(t) = (a + bt)e^{-at-bt^2/2}\mathbb{1}_{t \geq} 0.$$

When $a = 0$, the preceding equation is known as the Rayleigh density function.

# Gamma Distribution

A random variable is said to have a gamma distribution with parameters $\alpha, \beta$, $\alpha > 0, \beta > 0$, if its density function is given by

$$f(t) = \frac{\beta^\alpha}{\Gamma(\alpha) t^{\alpha-1}} e^{-\beta t} \mathbb{1}_{t \geq 0},$$

where $\Gamma(\alpha)$, called the gamma function, is defined as

$$\Gamma(\alpha) = \int_0^\alpha e^{-y} y^{\alpha-1} dy,$$

$\beta$ is called the rate parameter and $\alpha$ the shape parameter. It is often to see the density parametrized in terms of $\alpha$ and $\theta = 1/\beta$, $\theta$ is called the scale parameter of the distribution. This distribution includes the exponential as a special case ($\alpha = 1$).

# Erlang Distribution

The gamma distribution arises in some situations involving the exponential distribution, because of the well-known result that sums of independent and identically distributed (i.i.d.) exponential random variables have a gamma distribution. Specifically, if $T_1, \ldots, T_n$, are independent, each with exponential distribution $\lambda$, then $T_1 + \cdots + T_n$ has a gamma distribution with parameters $\alpha = n$ and $\beta = \lambda$. This distribution is often referred to in the literature as the $n$-Erlang distribution.

# $\chi^2$ Distribution

The gamma distribution with $\alpha = n/2$ and $\beta = 1/2$, $n$ a positive integer, is called the $\chi^2$ distribution with $n$ degrees of freedom. The chi-squared distribution often arises in practice as the distribution of the error involved in attempting to hit a target in $n$-dimensional space when each coordinate error is normally distributed.

# Weibull Dsitribution

The Weibull distribution is perhaps the most widely used lifetime distribution model. Application to the lifetimes or durability of manufactured items is common, and it is used as a model with diverse types of items, such as ball bearings, automobile components, and electrical insulation. It is also used in biological and medical applications, for example, in studies on the time to the occurrence of tumors in human populations or in laboratory animals.

The Weibull distribution has a hazard function of the form

$$h(t) = \lambda\beta(\lambda t)^{\beta-1},$$

where $\lambda > 0$ is called the rate parameter ($\theta = 1/\lambda$ is called the scale parameter) and $\beta > 0$ is called the shape parameter. It includes the exponential distribution as the special case where $\beta = 1$.

The distribution, survivor and density functions are given by

$$F(t) = 1 - \exp\left\{-(\lambda t)^{\beta}\right\}, \quad S(t) = \exp\left\{-(\lambda t)^{\beta}\right\}$$

and

$$f(t) = \lambda\beta(\lambda t)^{\beta-1}\exp\left\{-(\lambda t)^{\beta}\right\}\mathbb{1}_{t\geq 0}.$$

The Weibull hazard function is monotone increasing if $\beta > 1$, decreasing if $\beta < 1$, and constant for $\beta = 1$. The model is fairly flexible and has been found to provide a good description of many types of lifetime data.

# Log-normal Distribution

The log-normal distribution has been used as a model in diverse applications in engineering, medicine, and other areas. A random variable $T$ is said to be log-normally distributed if $X = \log T$ is normally distributed, say with mean $\mu$, variance $\sigma^2$.

The density, distribution and survivor functions are given by

$$f(t) = \frac{1}{\sqrt{2\pi}\sigma t} \exp\left\{-\frac{1}{2}\left(\frac{\log t - \mu}{\sigma}\right)^2\right\} \mathbb{1}_{t>0}$$

$$F(t) = \Phi\left(\frac{\log t - \mu}{\sigma}\right), \quad S(t) = 1 - \Phi\left(\frac{\log t - \mu}{\sigma}\right)$$

The hazard function can be shown to have the value 0 at $t = 0$, increase to a maximum, and then decrease, approaching 0 as $t \to \infty$. This shape arises in many situations, for example, when a population consists of a mixture of individuals who tend to have short and long lifetimes, respectively. Examples include survival after treatment for some forms of cancer, where persons who are cured become long-term survivors, and the duration of marriages, where after a certain number of years the risk of marriage dissolution due to divorce tends to decrease.

# Cauchy Distribution

A random variable is said to have a Cauchy distribution with parameters $\mu \in \mathbb{R}$ and $\sigma > 0$, if its density is given by

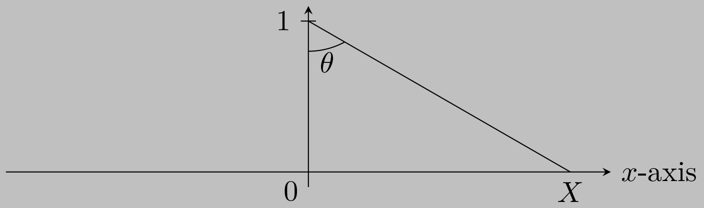
$$f(x) = \frac{1}{\pi\sigma} \frac{1}{1 + \left(\frac{x-\mu}{\sigma}\right)^2} \mathbb{1}_{x \in \mathbb{R}},$$

$\mu$ is called the location parameter and $\sigma$ the scale parameter. And its distribution function is given by

$$\frac{1}{2} + \frac{1}{\pi} \arctan\left(\frac{x - \mu}{\sigma}\right).$$

# Example

Suppose that a narrow-beam flashlight is spun around its center, which is located a unit distance from the $x$-axis. Consider the point $X$ at which the beam intersects the $x$-axis when the flashlight has stopped spinning. (If the beam is not pointing toward the $x$-axis, repeat the experiment.)



The point $X$ is determined by the angle $\theta$ between the flashlight and the $y$-axis, which, from the physical situation, appears to be uniformly distributed between $-\pi/2$ and $\pi/2$. The distribution function of $X$ is thus given by

$$\begin{aligned} F(x) &= \mathbb{P}(X \leq x) \\ &= \mathbb{P}(\tan\theta \leq x) \\ &= \mathbb{P}(\theta \leq \arctan x) \\ &= \frac{\arctan x + \pi/2}{\pi/2 + \pi/2} \\ &= \frac{1}{2} + \frac{\arctan x}{\pi} \end{aligned}$$

Hence, the density function of $X$ is given by

$$f(x) = \frac{d}{dx}F(x) = \frac{1}{\pi(1 + x^2)}$$

and we see that $X$ has the Cauchy distribution.

# Beta Distribution

A random variable $X$ is said to have a beta distribution if its density is given by

$$f(x) = \frac{1}{B(a,b)} x^{a-1}(1-x)^{b-1}\mathbb{1}_{x\in(0,1)},$$

where

$$B(a,b) = \int_0^1 x^{a-1}(1-x)^{b-1}dx = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$$

and $a, b > 0$.

When $a = b$, the beta density is symmetric about $1/2$ , giving more and more weight to regions about $1/2$ as the common value a increases. When $b > a$, the density is skewed to the left (in the sense that smaller values become more likely); and it is skewed to the right when $a > b$.

The uniform distribution in $(0,1)$ is a special case of the beta distribution where $a = b = 1$.

# Mixed Random Variables

# Mixed Random Variables

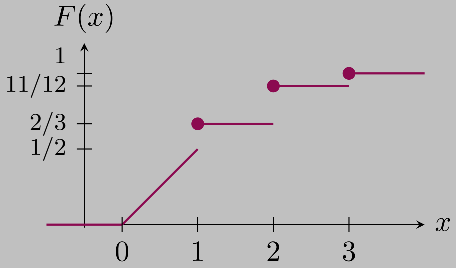
These are random variables that are neither discrete nor continuous, but are a mixture of both. In particular, a mixed random variable has a continuous part and a discrete part.

## Examples

1. The distribution function of the random variable $X$ is given by

$$F(x) = \begin{cases} 0 & \text{if } x < 0, \\ \frac{x}{2} & \text{if } 0 \leq x < 1, \\ \frac{2}{3} & \text{if } 1 \leq x < 2, \\ \frac{11}{12} & \text{if } 2 \leq x < 3, \\ 1 & \text{if } 3 \leq x. \end{cases}$$

A graph of $F(x)$ is




(a) $\mathbb{P}(X < 3) = \lim_{x\uparrow 3} F(x) = \frac{11}{12}$.

(b) $\mathbb{P}(X = 1) = \mathbb{P}(X \leq 1) - \mathbb{P}(X < 1) = F(1) - \lim_{x\uparrow 1} F(x) = \frac{2}{3} - \frac{1}{2} = \frac{1}{6}$.

(c) $\mathbb{P}(X > 1/2) = 1 - \mathbb{P}(X \leq 1/2) = 1 - F(1/2) = \frac{3}{4}$.

(d) $\mathbb{P}(2 < X \leq 4) = F(4) - F(2) = \frac{1}{12}$.

2. Let $X$ be a continuous random variable with the following density: $f(x) = 2x\mathbb{1}_{x\in[0,1]}$. Let also

$$Y = g(X) = \begin{cases} X & \text{if } 0 \leq X \leq \frac{1}{2}, \\ \frac{1}{2} & \text{if } X > \frac{1}{2}. \end{cases}$$

Find the distribution function of $Y$ and calculate the probabilities: (a) $\mathbb{P}(\frac{1}{4} \leq Y \leq \frac{3}{8})$, (b) $\mathbb{P}(Y \geq \frac{1}{4})$.

First note that the support of $X$ is $[0, 1]$For $x \in [0, 1]$, $0 \leq g(x) \leq \frac{1}{2}$. Thus, the support of $Y$ is $[0, \frac{1}{2}]$, and therefore

$$F_Y(y) = 0, \quad \text{for } y < 0,$$
$$F_Y(y) = 1, \quad \text{for } y > \tfrac{1}{2}.$$

Now note that

$$\mathbb{P}(Y = \frac{1}{2}) = \mathbb{P}(X > \frac{1}{2})$$
$$= \int_{\frac{1}{2}}^{1} 2x dx = \frac{3}{4}.$$

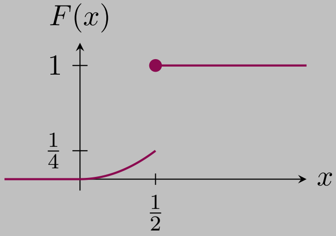
Also, for $0 < y < \frac{1}{2}$,

$$\begin{aligned} F_Y(y) &= \mathbb{P}(Y \le y) \\ &= \mathbb{P}(X \le y) \\ &= \int_0^y 2x dx \\ &= y^2. \end{aligned}$$

Thus, the distribution function of $Y$ is given by

$$F_Y(y) = y^2 \mathbb{1}_{y \in (0, \frac{1}{2})} + \mathbb{1}_{y \ge \frac{1}{2}},$$

whose graph is



(a)

$$\begin{aligned}\mathbb{P}\left(\frac{1}{4} \le Y \le \frac{3}{8}\right) &= \mathbb{P}\left(\frac{1}{4} < Y \le \frac{3}{8}\right) + \mathbb{P}\left(Y = \frac{1}{4}\right) \\ &= F_Y\left(\frac{3}{8}\right) - F_Y\left(\frac{1}{4}\right) + \mathbb{P}\left(Y = \frac{1}{4}\right) \\ &= \left(\frac{3}{8}\right)^2 - \left(\frac{1}{4}\right)^2 + 0 = \frac{5}{64}.\end{aligned}$$

(b)

$$\begin{aligned}\mathbb{P}\left(Y \ge \frac{1}{4}\right) &= 1 - F_Y\left(\frac{1}{4}\right) + \mathbb{P}\left(Y = \frac{1}{4}\right) \\ &= 1 - \left(\frac{1}{4}\right)^2 + 0 = \frac{15}{16}.\end{aligned}$$