for continuous random variables. Furthermore, the definition and properties of the variance are the same. So to evaluate the variance of a random variable $X$, as before,

$$\operatorname{Var}(X) = \mathbb{E}\,[(X - \mathbb{E}\,X)^2] = \mathbb{E}\,[X^2] - (\mathbb{E}\,X)^2.$$

**Example:** Let $X$ be the random variable having pdf

$$f_X(x) = \begin{cases} 3/4(1 - x^2), & \text{if } x \in [-1, 1] \\ 0, & \text{if } x \notin [-1, 1]. \end{cases}$$

What is the expected value and variance of $X$?

$$\mathbb{E}\,X = \int_{-\infty}^{\infty} x\, f_X(x)\, dx = \int_{-1}^{1} x \cdot \frac{3}{4}(1-x^2)\, dx = \frac{3}{4}\left(\frac{1}{2}x^2 - \frac{1}{4}x^4\right)\Big|_{-1}^{1} = 0$$

$$\mathbb{E}\,X^2 = \int_{-\infty}^{\infty} x^2\, f_X(x)\, dx = \int_{-1}^{1} x^2 \cdot \frac{3}{4}(1-x^2)\, dx$$

$$= \frac{3}{4}\left(\frac{1}{3}x^3 - \frac{1}{5}x^5\right)\Big|_{-1}^{1} = \frac{3}{4}\left(\frac{2}{3} - \frac{2}{5}\right) = \frac{1}{5}$$

$$\operatorname{Var}(X) = \mathbb{E}\,X^2 - (\mathbb{E}\,X)^2 = 1/5 - 0 = 1/5.$$

## Common continuous distributions

### Uniform Distribution

Consider a random variable $X$ taking values in the interval $[a, b]$. If all sub-intervals of equal length have the same probability of containing $X$, the we say $X$ has a *continuous* **uniform distribution** on $[a, b]$.

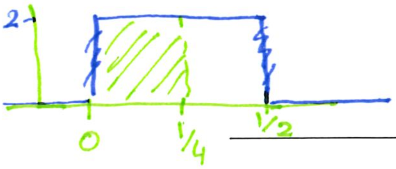
**Example:** If $X \sim \mathsf{Uniform}[0, 1/2]$ then the pdf of $X$ is

$$f_X(x) = \begin{cases} c, \; 2. & x \in (0, 1/2), \\ 0, & x \notin (0, 1/2), \end{cases}$$

where $c$ is some constant. What is $c$?

$$1 = \int_{-\infty}^{\infty} f_X(x)\, dx = \int_0^{1/2} c\, dx = \frac{c}{2}$$

giving $c = 2$.




$$f_X(x) = \begin{cases} 2, & x \in (0, 1/2) \\ 0, & x \notin (0, 1/2) \end{cases}$$



What is $\mathbb{P}(X \leqslant 1/4)$?

$$\mathbb{P}(X \leqslant 1/4) = \int_{-\infty}^{1/4} f_X(x)dx = \int_0^{1/4} 2\, dx = 1/2.$$

In general, the probability density function of the uniform distribution on $[a, b]$ is

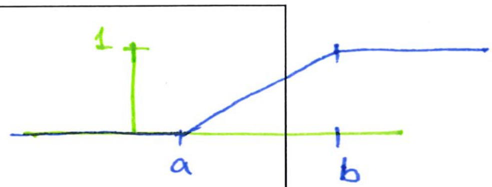
$$f_X(x) = \begin{cases} \frac{1}{b-a}, & x \in [a, b], \\ 0, & x \notin [a, b], \end{cases}$$

and the cumulative distribution function of the uniform distribution on $[a, b]$ is

$$F_X(x) = \int_{-\infty}^{x} f_X(u)du$$

$$x \in (a,b) \quad F_X(x) = \int_a^x \frac{1}{b-a} du$$

$$F_X(x) = \begin{cases} 0, & x \in (-\infty, a] \\ \frac{x-a}{b-a}, & x \in [a, b], \\ 1, & x \in [b, \infty), \end{cases}$$




Since $\mathbb{P}(X = x) = 0$ for a continuous distribution, the uniform distributions on $[a, b]$, $[a, b)$, $(a, b]$ and $(a, b)$ are all essentially the same.

The expected value and variance of the uniform distribution on $[a, b]$ are

$$\mathbb{E}\, X = \int_{-\infty}^{\infty} x\, f_X(x)\, dx = \int_a^b x \cdot \frac{1}{b-a} = \frac{1}{(b-a)} \cdot \frac{x^2}{2}\Big|_a^b = \frac{b^2-a^2}{2(b-a)} = \frac{(b+a)\cancel{(b-a)}}{2\cancel{(b-a)}}$$

$$\mathbb{E}\, X^2 = \int_{-\infty}^{\infty} x^2\, f_X(x)\, dx = \int_a^b \frac{1}{(b-a)} x^2 dx = \frac{1}{(b-a)} \cdot \frac{x^3}{3}\Big|_a^b = \frac{b^3-a^3}{3(b-a)}$$

$$= \frac{(b-a)(b^2+ab+a^2)}{3(b-a)} = \frac{1}{3}(b^2+ab+a^2)$$

$$\mathrm{Var}(X) = \mathbb{E}\, X^2 - (\mathbb{E}\, X)^2 = \frac{1}{3}(b^2+ab+a^2) - \left(\frac{b+a}{2}\right)^2$$

$$= \frac{b^2+ab+a^2}{3} - \frac{b^2+2ab+a^2}{4} = \frac{(b-a)^2}{12}$$

**Question.** Let $n$ be a positive integer. If $X$ has a continuous uniform distribution on $(0, 1)$, what is the distribution of $Y = \lceil nX \rceil$? (*Note: $\lceil x \rceil$ means rounding $x$ up to the nearest integer.*)

First note that $Y$ is a discrete random variable taking values in $\{1, 2, \ldots, n\}$. We now want to determine the probability mass function of $Y$. For $y \in \{1, 2, \ldots, n\}$,

$y \in \{1, 2, \ldots, n\}$

pmf
$$\begin{aligned} f_Y(y) = \mathbb{P}(Y = y) &= \mathbb{P}(\lceil nX \rceil = y) = \mathbb{P}(y-1 < nX \leqslant y) \\ &= \mathbb{P}\left(\frac{y-1}{n} < X \leqslant y/n\right) \\ &= \int_{(y-1)/n}^{y/n} 1 \cdot dx \\ &= \frac{y}{n} - \frac{(y-1)}{n} = 1/n \end{aligned}$$

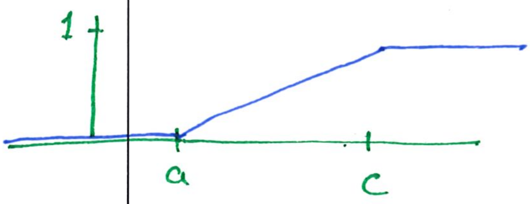
So $Y$ has a Discrete Uniform on $\{1, 2, \ldots, n\}$.

**Question.** Suppose $X$ has a continuous uniform distribution on $(a, b)$ and let $c \in (a, b)$. What is the distribution of $X$ conditioned on the event $\{X \leqslant c\}$?

For $x \in (a, c)$,

$\{X \leqslant x\} \subseteq \{X \leqslant c\}$

$$\begin{aligned} \mathbb{P}(X \leqslant x \mid X \leqslant c) &= \frac{\mathbb{P}(X \leqslant x, X \leqslant c)}{\mathbb{P}(X \leqslant c)} \\ &= \frac{\mathbb{P}(X \leqslant x)}{\mathbb{P}(X \leqslant c)} \\ &= \left(\int_a^x \frac{1}{b-a}\, du\right) \Big/ \left(\int_a^c \frac{1}{b-a}\, du\right) \\ &= \left(\frac{x-a}{b-a}\right) \Big/ \left(\frac{c-a}{b-a}\right) = \frac{x-a}{c-a} \end{aligned}$$




So conditional on the event $\{X \leqslant c\}$, $X$ has a cont. Uniform on $(a, c)$.

## Exponential Distribution

**Example:** Let $X$ be the random variable describing the time until the first bug is reported in a computer program. Suppose the cdf of $X$ is given by

$$F_X(x) = \mathbb{P}(X \leqslant x) = \begin{cases} 0, & \text{if } x < 0 \\ 1 - e^{-\lambda x}, & \text{if } x \geqslant 0, \end{cases}$$

where $\lambda$ is a positive constant that may depend on the complexity of the program. What is the pdf of $X$?

Recall: If $F_X$ is differentiable at $x$, then the pdf at $x$ is $F_X'(x)$.

If $x < 0$, then $f_X(x) = F_X'(x) = 0$.

If $x > 0$, then $f_X(x) = F_X'(x) = \lambda e^{-\lambda x}$.

However, $F_X'(0)$ does not exist. This doesn't matter; $f_X(0)$ can be set to anything non-negative, say $f_X(0) = \lambda$. In this case we get

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

Note that the "0 if $x < 0$" part of the function is essential.

We call the distribution with this pdf the **Exponential distribution** with rate parameter $\lambda$. We denote this distribution by $\mathsf{Exp}(\lambda)$. It is one of the most important distribution in *Applied Probability* due to its connection to continuous time Markov chains and the Poisson distribution (which we will see later) and its many useful properties. Like the discrete geometric distribution, the exponential distribution has the *memoryless* property.

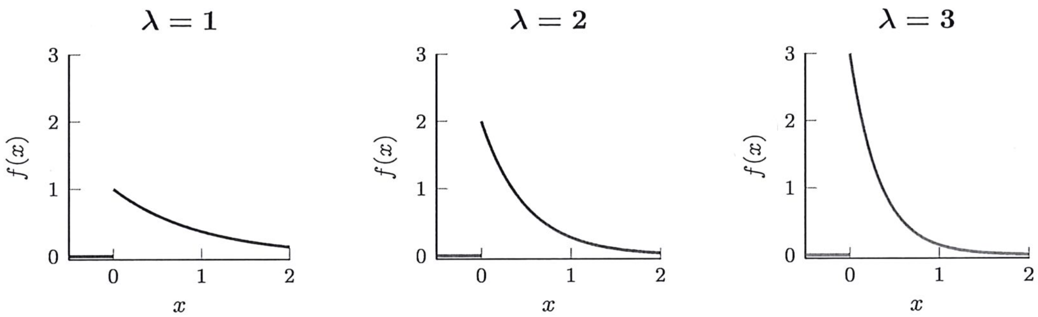



Figure 5.1: Probability density function for $X \sim \mathsf{Exp}(\lambda)$ for $\lambda = 1$, 2, and 3. Compare this to Figure 4.4.

**Question.** What is the probability that the time until the first bug is reported is greater than time $x$?

Assuming $x > 0$,

$$\begin{aligned} \mathbb{P}(X > x) &= 1 - \mathbb{P}(X \leq x) \\ &= 1 - F_X(x) \\ &= 1 - (1 - e^{-\lambda x}) = e^{-\lambda x}. \end{aligned}$$

The function $\mathbb{P}(X > x)$ is often called the "survivor function" or "reliability function" and is related to cumulative distribution function by $\mathbb{P}(X > x) = 1 - F_X(x)$.

**Question.** Given that no bug is found after $y$ time units, what is the probability that it takes longer than time $x + y$ for the first bug to be found?

$$\begin{aligned} \mathbb{P}(X > x + y \mid X > y) &= \mathbb{P}(X > x+y, X > y) / \mathbb{P}(X > y) \\ &= \mathbb{P}(X > x+y) / \mathbb{P}(X > y) \\ &= e^{-\lambda(x+y)} / e^{-\lambda y} \\ &= e^{-\lambda x} \end{aligned}$$

$\{X > x+y\} \subseteq \{X > y\}$

So $\mathbb{P}(X > x+y \mid X > y) = \mathbb{P}(X > x)$

"This is the memory-less property".

**Question.** If $X \sim \mathsf{Exp}(\lambda)$, what is the distribution of $Y = \lceil X \rceil$? The support of the distribution of $Y$ is $\{1, 2, \dots\}$. The probability mass function of $Y$ is given by

$$\begin{aligned} f_Y(y) = \mathbb{P}(Y = y) &= \mathbb{P}(y-1 < X \leq y) \\ &= \int_{y-1}^{y} \lambda e^{-\lambda x}\, dx \\ &= -e^{-\lambda x}\Big|_{y-1}^{y} \\ &= e^{-\lambda(y-1)} - e^{-\lambda y} = (e^{-\lambda})^{y-1}(1-e^{-\lambda}) \end{aligned}$$

Geo$(p)$
pmf $p(1-p)^{k-1}$

where $y \in \{1, 2, \dots\}$. We can recognise this as the probability mass function of a $\boxed{\text{Geometric }(1-e^{-\lambda})}$ distribution.

What is the expected value and variance of the $\mathsf{Exp}(\lambda)$ distribution? It will be useful to recall that for two functions $u$ and $v$ we may use *integration by parts* as follows

$$\int u(x)v'(x)\mathrm{d}x = u(x)v(x) - \int v(x)u'(x)\mathrm{d}x\,.$$

So

$$\begin{aligned} \mathbb{E}[X] &= \int_{-\infty}^{\infty} x \cdot f_X(x)\,dx = \int_0^{\infty} x \cdot \lambda e^{-\lambda x}\, dx \\ &= x\cdot(-e^{-\lambda x})\Big|_0^{\infty} + \int_0^{\infty} 1 \cdot e^{-\lambda x}\,dx \quad [u(x) = x,\ v(x) = -e^{-\lambda x}] \\ &= 0 + \left(\frac{-1}{\lambda} e^{-\lambda x}\right)\Big|_0^{\infty} = \frac{1}{\lambda} \\ &= \left(\frac{-1}{\lambda} e^{-\lambda\cdot\infty} - \frac{-1}{\lambda} e^{-\lambda\cdot 0}\right) \end{aligned}$$

$$\begin{aligned} \mathbb{E}[X^2] &= \int_0^{\infty} x^2 f_X(x)\,\mathrm{d}x = \int_0^{\infty} x^2 \underbrace{\lambda e^{-\lambda x}}\,\mathrm{d}x \quad [u(x) = x^2,\ v(x) = -e^{-\lambda x}] \quad v'(x) = \lambda e^{-\lambda x} \\ &= x^2 \cdot (-e^{-\lambda x})\Big|_0^{\infty} + \int_0^{\infty} 2x e^{-\lambda x}\,dx \\ &= 0 + \frac{2}{\lambda}\int_0^{\infty} x \cdot \lambda e^{-\lambda x}\,dx = \frac{2}{\lambda}\cdot\frac{1}{\lambda} = \frac{2}{\lambda^2} \\ &= \end{aligned}$$

$$\mathrm{Var}(X) = \mathbb{E}[X^2] - (\mathbb{E}\,X)^2 = \frac{2}{\lambda^2} - \left(\frac{1}{\lambda}\right)^2 = \frac{1}{\lambda^2}\,.$$

## Standard Normal Distribution

The normal (or *Gaussian*) distribution plays a central role in probability and statistics. Many observed quantities appear to follow a normal distribution. Later we will see that the central limit theorem states that, under mild conditions, the average of independent