확률과 통계

Class 3

Chapter 4
Continuous Distributions

4.1 Probability density

Probability density

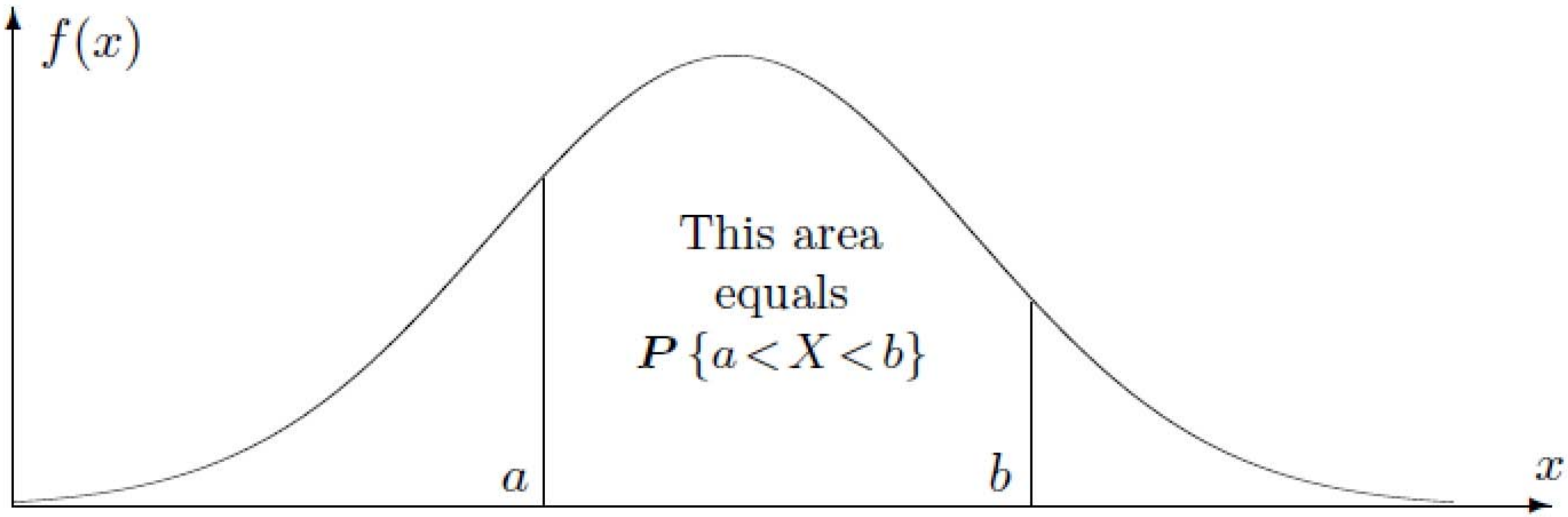


FIGURE 4.1: *Probabilities are areas under the density curve.*

DEFINITION 4.1

Probability density function (pdf, density) is the derivative of the cdf, $f(x) = F'(x)$. The distribution is called **continuous** if it has a density.

$$\int_a^b f(x)dx = F(b) - F(a) = \boldsymbol{P}\{a < X < b\}$$

Probability density function

$$f(x) = F'(x)$$

$$\boldsymbol{P}\{a < X < b\} = \int_a^b f(x)dx$$

$$\int_{-\infty}^{b} f(x)dx = \boldsymbol{P}\{-\infty < X < b\} = F(b)$$

$$\int_{-\infty}^{+\infty} f(x)dx = \boldsymbol{P}\{-\infty < X < +\infty\} = 1.$$

$$P(x) = \boldsymbol{P}\{x \leq X \leq x\} = \int_{x}^{x} f = 0.$$

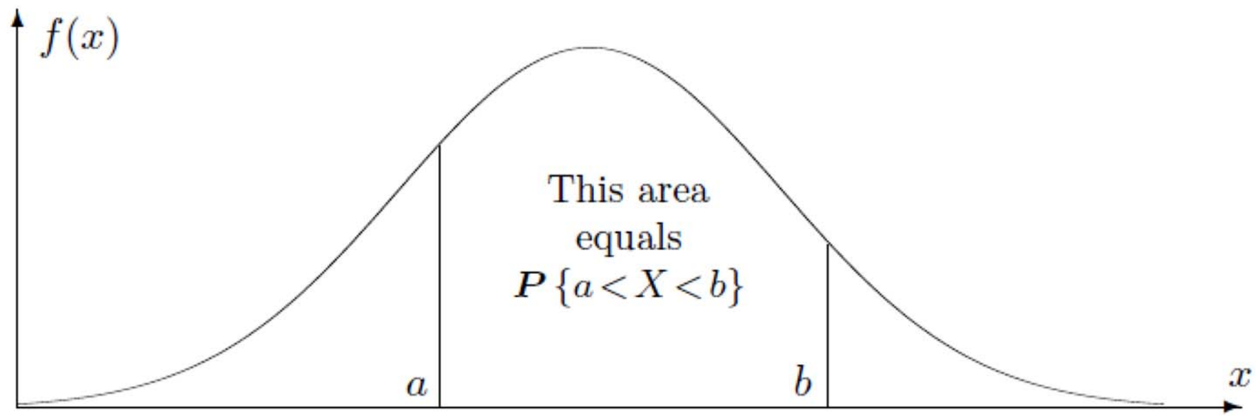


FIGURE 4.1: *Probabilities are areas under the density curve.*

Example 4.1. The lifetime, in years, of some electronic component is a continuous random variable with the density

$$f(x) = \begin{cases} \dfrac{k}{x^3} & \text{for} \quad x \geq 1 \\ 0 & \text{for} \quad x < 1. \end{cases}$$

Find k, draw a graph of the cdf $F(x)$, and compute the probability for the lifetime to exceed 5 years.

Solution. Find k from the condition $\int f(x)dx = 1$:

$$\int_{-\infty}^{+\infty} f(x)dx = \int_{1}^{+\infty} \frac{k}{x^3}dx = -\left.\frac{k}{2x^2}\right|_{x=1}^{+\infty} = \frac{k}{2} = 1.$$

$$\frac{d}{dx}x^c = cx^{c-1}$$

Hence, $k = 2$. Integrating the density, we get the cdf,

$$F(x) = \int_{-\infty}^{x} f(y)dy = \int_{1}^{x} \frac{2}{y^3}dy = -\left.\frac{1}{y^2}\right|_{y=1}^{x} = 1 - \frac{1}{x^2}$$

for $x > 1$. Its graph is shown in Figure 4.2.

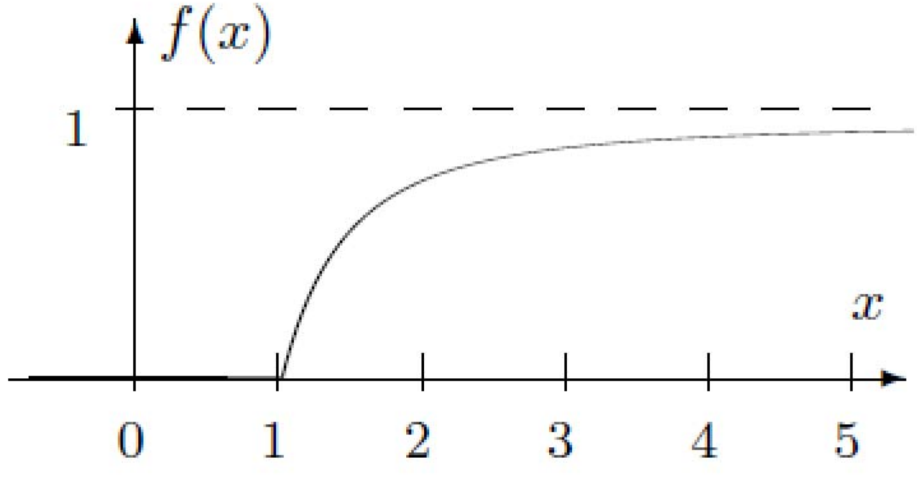


FIGURE 4.2: *Cdf for Example 4.1.*

$$\boldsymbol{P}\{X > 5\} = \int_5^{+\infty} f(x)dx = \int_5^{+\infty} \frac{2}{x^3}dx = -\left.\frac{1}{x^2}\right|_{x=5}^{+\infty} = \frac{1}{25} = 0.04.$$

Distribution	**Discrete**	**Continuous**
Definition	$P(x) = P\{X = x\}$ (pmf)	$f(x) = F'(x)$ (pdf)
Computing probabilities	$\boldsymbol{P}\{X \in A\} = \sum_{x \in A} P(x)$	$\boldsymbol{P}\{X \in A\} = \int_A f(x)dx$
Cumulative distribution function	$F(x) = \boldsymbol{P}\{X \le x\} = \sum_{y \le x} P(y)$	$F(x) = \boldsymbol{P}\{X \le x\} = \int_{-\infty}^{x} f(y)dy$
Total probability	$\sum_x P(x) = 1$	$\int_{-\infty}^{\infty} f(x)dx = 1$

TABLE 4.1: *Pmf* $P(x)$ *versus pdf* $f(x)$.

Joint and marginal densities

DEFINITION 4.2

For a vector of random variables, the **joint cumulative distribution function** is defined as

$$F_{(X,Y)}(x, y) = \boldsymbol{P}\{X \leq x \cap Y \leq y\}.$$

The **joint density** is the *mixed derivative* of the joint cdf,

$$f_{(X,Y)}(x, y) = \frac{\partial^2}{\partial x \partial y} F_{(X,Y)}(x, y).$$

Distribution	**Discrete**	**Continuous**
Marginal distributions	$P(x) = \sum_y P(x,y)$ $P(y) = \sum_x P(x,y)$	$f(x) = \int f(x,y)dy$ $f(y) = \int f(x,y)dx$
Independence	$P(x,y) = P(x)P(y)$	$f(x,y) = f(x)f(y)$
Computing probabilities	$\boldsymbol{P}\{(X,Y) \in A\}$ $= \sum_{(x,y)\in A}\sum P(x,y)$	$\boldsymbol{P}\{(X,Y) \in A\}$ $= \iint_{(x,y)\in A} f(x,y)\,dx\,dy$

TABLE 4.2: *Joint and marginal distributions in discrete and continuous cases.*

Expectation and variance

$$\mu = \mathbf{E}(X) = \int x f(x) dx$$

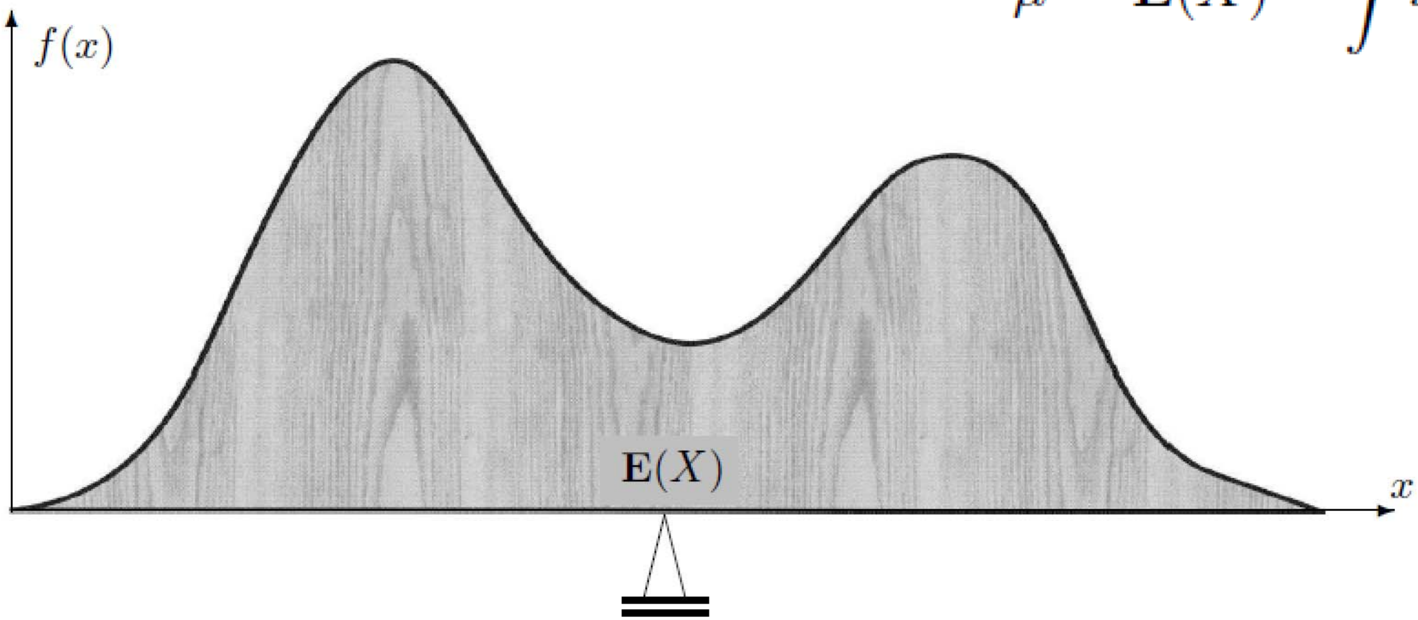


FIGURE 4.3: *Expectation of a continuous variable as a center of gravity.*

Example 4.2. A random variable X in Example 4.1 has density

$$f(x) = 2x^{-3} \text{ for } x \geq 1.$$

Its expectation equals

$$\mu = \mathbf{E}(X) = \int x\, f(x) dx = \int_1^\infty 2x^{-2} dx = -2x^{-1}\Big|_1^\infty = 2.$$

$$\sigma^2 = \mathrm{Var}(X) = \int x^2 f(x) dx - \mu^2 = \int_1^\infty 2x^{-1} dx - 4 = 2 \ln x|_1^\infty - 4 = +\infty.$$

Discrete	**Continuous**
$\mathbf{E}(X) = \sum_x xP(x)$	$\mathbf{E}(X) = \int xf(x)dx$
$\mathrm{Var}(X) = \mathbf{E}(X-\mu)^2$ $= \sum_x (x-\mu)^2 P(x)$ $= \sum_x x^2 P(x) - \mu^2$	$\mathrm{Var}(X) = \mathbf{E}(X-\mu)^2$ $= \int (x-\mu)^2 f(x)dx$ $= \int x^2 f(x)dx - \mu^2$
$\mathrm{Cov}(X,Y) = \mathbf{E}(X-\mu_X)(Y-\mu_Y)$ $= \sum_x \sum_y (x-\mu_X)(y-\mu_Y)P(x,y)$ $= \sum_x \sum_y (xy)P(x,y) - \mu_x\mu_y$	$\mathrm{Cov}(X,Y) = \mathbf{E}(X-\mu_X)(Y-\mu_Y)$ $= \iint (x-\mu_X)(y-\mu_Y)f(x,y)\,dx\,dy$ $= \iint (xy)f(x,y)\,dx\,dy - \mu_x\mu_y$

TABLE 4.3: *Moments for discrete and continuous distributions.*

4.2 Families of continuous distributions

4.2.1 Uniform distribution

$$f(x) = \frac{1}{b-a}, \quad a < x < b,$$

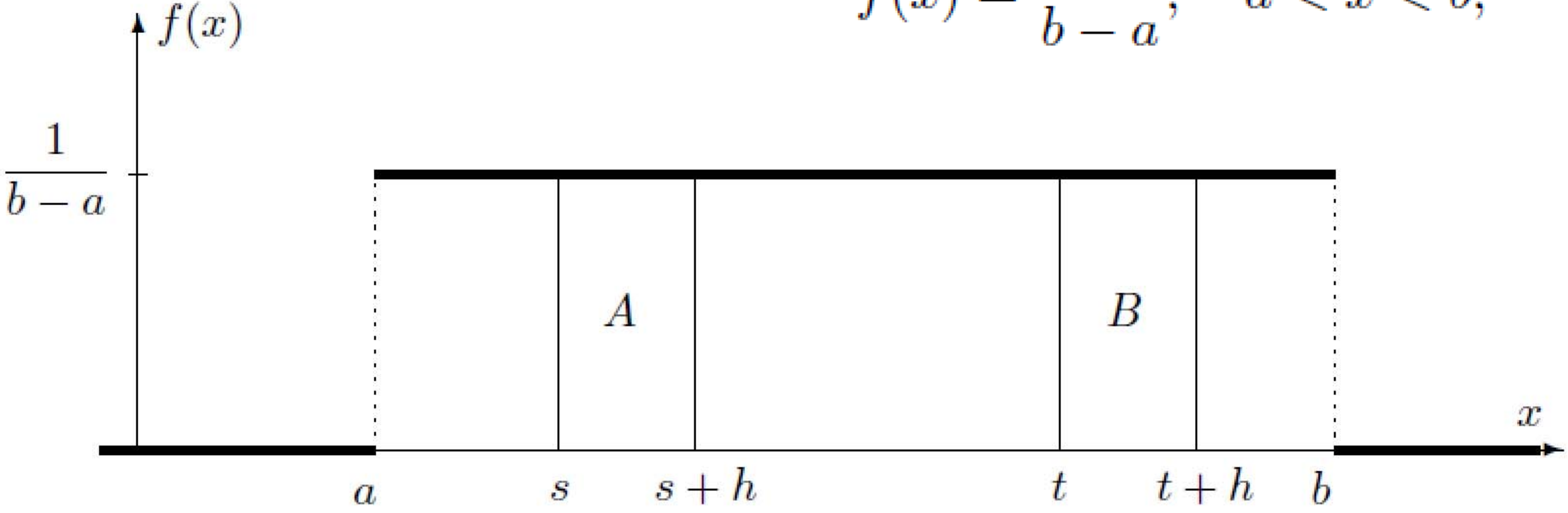


FIGURE 4.4: *The Uniform density and the Uniform property.*

The uniform property

For any $h > 0$ and $t \in [a, b-h]$, the probability

$$\boldsymbol{P}\{\, t < X < t+h \,\} = \int_t^{t+h} \frac{1}{b-a}\, dx = \frac{h}{b-a}$$

Example 4.3. In Figure 4.4, rectangles A and B have the same area, showing that $\boldsymbol{P}\{s < X < s+h\} = \boldsymbol{P}\{t < X < t+h\}$. $\Diamond$

Example 4.4. If a flight scheduled to arrive at 5 pm actually arrives at a Uniformly distributed time between 4:50 and 5:10, then it is equally likely to arrive before 5 pm and after 5 pm, equally likely before 4:55 and after 5:05, etc. $\Diamond$

Standard uniform distribution

- b=1, a=0

$$Y = \frac{X - a}{b - a} \qquad\qquad X = a + (b - a)Y$$

Expectation and variance

$$\mathbf{E}(Y) = \int y f(y) dy = \int_0^1 y dy = \frac{1}{2}$$

$$\mathrm{Var}(Y) = \mathbf{E}(Y^2) - \mathbf{E}^2(Y) = \int_0^1 y^2 dy - \left(\frac{1}{2}\right)^2 = \frac{1}{3} - \frac{1}{4} = \frac{1}{12}.$$

$$X = a + (b - a)Y$$

$$\mathbf{E}(X) = \mathbf{E}\{a + (b-a)Y\} = a + (b-a)\,\mathbf{E}(Y) = a + \frac{b-a}{2} = \frac{a+b}{2}$$

$$\mathrm{Var}(X) = \mathrm{Var}\{a + (b-a)Y\} = (b-a)^2\,\mathrm{Var}(Y) = \frac{(b-a)^2}{12}.$$

Uniform distribution

$$
\begin{aligned}
(a, b) &= \text{range of values} \\
f(x) &= \frac{1}{b-a}, \quad a < x < b \\
\mathbf{E}(X) &= \frac{a+b}{2} \\
\mathrm{Var}(X) &= \frac{(b-a)^2}{12}
\end{aligned}
$$

4.2.2 Exponential distribution

- Exponential distribution is often used to model time: waiting time, interarrival time, hardware lifetime, failure time, time between telephone calls, etc.
- As we shall see below, in a sequence of rare events, when the number of events is **Poisson**, the time between events is **Exponential**.

- Exponential distribution has density

$$f(x) = \lambda e^{-\lambda x} \text{ for } x > 0. \tag{4.1}$$

- With this density, we compute the Exponential cdf, mean, and variance as

$$\begin{aligned}
F(x) &= \int_0^x f(t)dt = \int_0^x \lambda e^{-\lambda t} dt = 1 - e^{-\lambda x} \quad (x > 0), && (4.2)\\
\mathbf{E}(X) &= \int t f(t)dt = \int_0^\infty t\lambda e^{-\lambda t} dt = \frac{1}{\lambda} \quad \left(\begin{array}{c} \textit{integrating} \\ \textit{by parts} \end{array} \right), && (4.3)\\
\mathrm{Var}(X) &= \int t^2 f(t)dt - \mathbf{E}^2(X) \\
&= \int_0^\infty t^2 \lambda e^{-\lambda t} dt - \left(\frac{1}{\lambda}\right)^2 \quad (\textit{by parts twice}) && (4.4)\\
&= \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}.
\end{aligned}$$

Partial integration

$$\int_a^b u(x)v'(x)\,dx = [u(x)v(x)]_a^b - \int_a^b u'(x)v(x)dx$$

$$= u(b)v(b) - u(a)v(a) - \int_a^b u'(x)v(x)\,dx$$

Event "the time T until the next event is greater than t" can be rephrased as "zero events occur by the time t," and further, as "$X = 0$," where X is the number of events during the time interval $[0, t]$. This X has Poisson distribution with parameter λt. It equals 0 with probability

$$P_X(0) = e^{-\lambda t}\frac{(\lambda t)^0}{0!} = e^{-\lambda t}.$$

Then we can compute the cdf of T as

$$F_T(t) = 1 - \boldsymbol{P}\{T > t\} = 1 - \boldsymbol{P}\{X = 0\} = 1 - e^{-\lambda t}, \tag{4.5}$$

Poisson distribution

$$\boxed{\begin{array}{lcl} \lambda & = & \text{frequency, average number of events} \\ P(x) & = & e^{-\lambda}\dfrac{\lambda^x}{x!},\ x = 0, 1, 2, \ldots \\ \mathbf{E}(X) & = & \lambda \\ \mathrm{Var}(X) & = & \lambda \end{array}}$$

Example 4.5. Jobs are sent to a printer at an average rate of 3 jobs per hour.
(a) What is the expected time between jobs?
(b) What is the probability that the next job is sent within 5 minutes?

Solution. Job arrivals represent rare events, thus the time T between them is Exponential with the given parameter $\lambda = 3$ hrs^{-1} (jobs per hour).
(a) $\mathbf{E}(T) = 1/\lambda = 1/3$ hours or 20 minutes between jobs;
(b) Convert to the same measurement unit: 5 min $= (1/12)$ hrs. Then,

$$\boldsymbol{P}\{T < 1/12 \text{ hrs}\} = F(1/12) = 1 - e^{-\lambda(1/12)} = 1 - e^{-1/4} = \underline{0.2212}.$$

$\Diamond$

Memoryless property

$$\boldsymbol{P}\{T > t + x \mid T > t\} = \boldsymbol{P}\{T > x\} \qquad \text{for } t, x > 0.$$

PROOF: From (4.2), $\boldsymbol{P}\{T > x\} = e^{-\lambda x}$. Also, by the formula (2.7) for conditional probability,

$$\boldsymbol{P}\{T > t + x \mid T > t\} = \frac{\boldsymbol{P}\{T > t + x \cap T > t\}}{\boldsymbol{P}\{T > t\}} = \frac{\boldsymbol{P}\{T > t + x\}}{\boldsymbol{P}\{T > t\}} = \frac{e^{-\lambda(t+x)}}{e^{-\lambda t}} = e^{-\lambda x}.$$

$\square$

Exponential distribution

$$\begin{aligned}
\lambda &= \text{frequency parameter, the number of events per time unit} \\
f(x) &= \lambda e^{-\lambda x}, \quad x > 0 \\
\mathbf{E}(X) &= \frac{1}{\lambda} \\
\mathrm{Var}(X) &= \frac{1}{\lambda^2}
\end{aligned}$$

4.2.3 Gamma distribution

- When a certain procedure consists of α independent steps, and each step takes Exponential(λ) amount of time, then the total time has **Gamma distribution** with parameters α and λ.
- Thus, Gamma distribution can be widely used for the total time of a multistage scheme, for example, related to downloading or installing a number of files.
- In a process of rare events, with Exponential times between any two consecutive events, the time of the α-th event has Gamma distribution because it consists of α independent Exponential times.

Example 4.6 (INTERNET PROMOTIONS). Users visit a certain internet site at the average rate of 12 hits per minute. Every sixth visitor receives some promotion that comes in a form of a flashing banner. Then the time between consecutive promotions has Gamma distribution with parameters $\alpha = 6$ and $\lambda = 12$. $\Diamond$

$$f(x) = \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}, \quad x > 0. \tag{4.7}$$

$$\boxed{\begin{array}{lcl} \text{Gamma}(1, \lambda) & = & \text{Exponential}(\lambda) \\ \text{Gamma}(\alpha, 1/2) & = & \text{Chi-square}(2\alpha) \end{array}}$$

$$\Gamma(z) = \int_0^\infty \frac{t^{z-1}\,dt}{\exp t} \qquad (\operatorname{Re} z > 0)$$

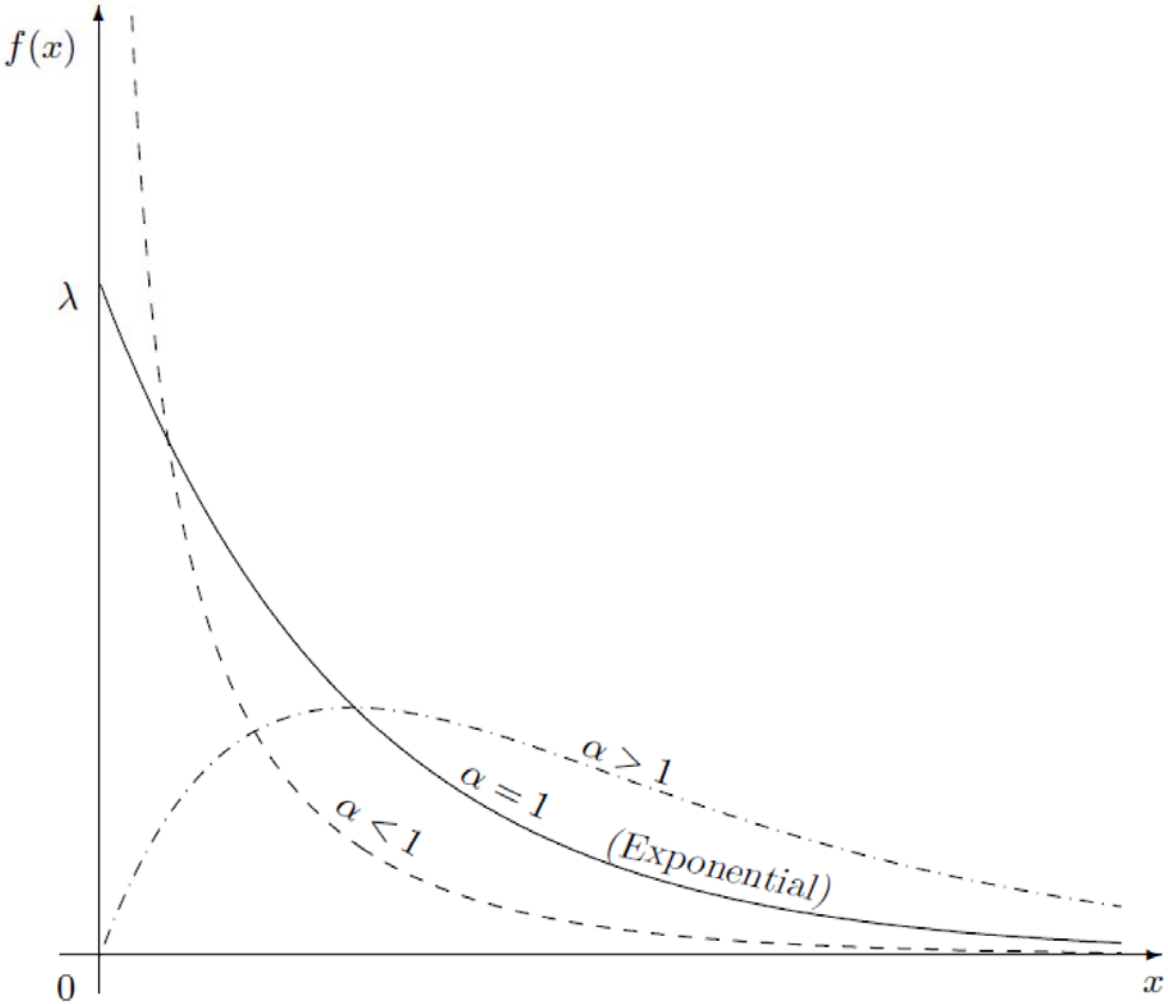


FIGURE 4.5: *Gamma densities with different shape parameters* α.

Gamma cdf has the form

$$F(t) = \int_0^t f(x)dx = \frac{\lambda^\alpha}{\Gamma(\alpha)} \int_0^t x^{\alpha-1} e^{-\lambda x} dx.$$

First, let us notice that $\int_0^\infty f(x)dx = 1$ for Gamma and all the other densities. Then, integrating (4.7) from 0 to ∞, we obtain that

$$\boxed{\int_0^\infty x^{\alpha-1} e^{-\lambda x} dx = \frac{\Gamma(\alpha)}{\lambda^\alpha} \qquad \text{for any } \alpha > 0 \text{ and } \lambda > 0} \tag{4.9}$$

Substituting $\alpha+1$ and $\alpha+2$ in place of α, we get for a Gamma variable X,

$$\mathbf{E}(X)=\int xf(x)dx=\frac{\lambda^\alpha}{\Gamma(\alpha)}\int_0^\infty x^\alpha e^{-\lambda x}dx=\frac{\lambda^\alpha}{\Gamma(\alpha)}\cdot\frac{\Gamma(\alpha+1)}{\lambda^{\alpha+1}}=\frac{\alpha}{\lambda}$$

(using the equality $\Gamma(t+1)=t\Gamma(t)$ that holds for all $t>0$),

$$\mathbf{E}(X^2)=\int x^2 f(x)dx=\frac{\lambda^\alpha}{\Gamma(\alpha)}\int_0^\infty x^{\alpha+1} e^{-\lambda x}dx=\frac{\lambda^\alpha}{\Gamma(\alpha)}\cdot\frac{\Gamma(\alpha+2)}{\lambda^{\alpha+2}}=\frac{(\alpha+1)\alpha}{\lambda^2},$$

and therefore,

$$\mathrm{Var}(X)=\mathbf{E}(X^2)-\mathbf{E}^2(X)=\frac{(\alpha+1)\alpha-\alpha^2}{\lambda^2}=\frac{\alpha}{\lambda^2}. \tag{4.11}$$

Gamma distribution

$$
\begin{aligned}
\alpha &= \text{shape parameter} \\
\lambda &= \text{frequency parameter} \\
f(x) &= \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}, \quad x > 0 \\
\mathbf{E}(X) &= \frac{\alpha}{\lambda} \\
\mathrm{Var}(X) &= \frac{\alpha}{\lambda^2}
\end{aligned}
$$

Example 4.7 (Total compilation time)

- Compilation of a computer program consists of 3 blocks that are processed sequentially, one after another.
- Each block takes Exponential time with the mean of 5 minutes, independently of other blocks.
  - (a) Compute the expectation and variance of the total compilation time.
  - (b) Compute the probability for the entire program to be compiled in less than 12 minutes.

- The total time T is a sum of three independent Exponential times, therefore, it has Gamma distribution with $\alpha = 3$.
- The frequency parameter λ equals (1/5) min^{-1} because the Exponential compilation time of each block has expectation $1/\lambda = 5$ min.
- (a) For a Gamma random variable T with $\alpha = 3$ and $\lambda = 1/5$,

$$\mathbf{E}(T) = \frac{3}{1/5} = 15 \text{ (min)} \quad \text{and} \quad \text{Var}(T) = \frac{3}{(1/5)^2} = 75 \text{ (min}^2\text{)}$$

(b) A direct solution involves two rounds of integration by parts,

$$
\begin{aligned}
\boldsymbol{P}\{T < 12\} &= \int_0^{12} f(t)dt = \frac{(1/5)^3}{\Gamma(3)} \int_0^{12} t^2 e^{-t/5} dt \\
&= \frac{(1/5)^3}{2!} \left(-5t^2 e^{-t/5} \Big|_{t=0}^{t=12} + \int_0^{12} 10te^{-t/5} dt \right) \\
&= \frac{1/125}{2} \left(-5t^2 e^{-t/5} - 50te^{-t/5} \Big|_{t=0}^{t=12} + \int_0^{12} 50e^{-t/5} dt \right) \\
&= \frac{1}{250} \left(-5t^2 e^{-t/5} - 50te^{-t/5} - 250e^{-t/5} \right) \Big|_{t=0}^{t=12} \\
&= 1 - e^{-2.4} - 2.4e^{-2.4} - 2.88e^{-2.4} = \underline{0.4303}.
\end{aligned} \tag{4.13}
$$

Gamma-Poisson formula

Indeed, let T be a Gamma variable with an integer parameter α and some positive λ. This is a distribution of the time of the α-th rare event. Then, the event $\{T > t\}$ means that the α-th rare event occurs after the moment t, and therefore, *fewer than α rare events occur before the time t*. We see that

$$\{T > t\} = \{X < \alpha\},$$

where X is the number of events that occur before the time t. This number of rare events X has Poisson distribution with parameter (λt); therefore, the probability

$$\boldsymbol{P}\{T > t\} = \boldsymbol{P}\{X < \alpha\}$$

and the probability of a complement

$$\boldsymbol{P}\{T \leq t\} = \boldsymbol{P}\{X \geq \alpha\}$$

Gamma-Poisson formula

For a Gamma(α, λ) variable T and a Poisson(λt) variable X,

$$\boldsymbol{P}\{T > t\} = \boldsymbol{P}\{X < \alpha\}$$

$$\boldsymbol{P}\{T \leq t\} = \boldsymbol{P}\{X \geq \alpha\}$$

Example 4.8 (Total compilation time, continued). Here is an alternative solution to Example 4.7(b). According to the Gamma-Poisson formula with $\alpha = 3$, $\lambda = 1/5$, and $t = 12$,

$$\boldsymbol{P}\{T < 12\} = \boldsymbol{P}\{X \geq 3\} = 1 - F(2) = 1 - 0.5697 = \underline{0.430}$$

from Table A3 for the Poisson distribution of X with parameter $\lambda t = 2.4$.

Furthermore, we notice that the four-term mathematical expression that we obtained in (4.13) after integrating by parts represents precisely

$$\boldsymbol{P}\{X \geq 3\} = 1 - P(0) - P(1) - P(2).$$

$\Diamond$

Example 4.9. Lifetimes of computer memory chips have Gamma distribution with expectation $\mu = 12$ years and standard deviation $\sigma = 4$ years. What is the probability that such a chip has a lifetime between 8 and 10 years?

<u>Solution</u>.

STEP 1, PARAMETERS. From the given data, compute parameters of this Gamma distribution. Using (4.12), obtain a system of two equations and solve them for α and λ,

$$\left\{ \begin{array}{lll} \mu & = & \alpha/\lambda \\ \sigma^2 & = & \alpha/\lambda^2 \end{array} \right. \Rightarrow \left\{ \begin{array}{lllllll} \alpha & = & \mu^2/\sigma^2 & = & (12/4)^2 & = & 9, \\ \lambda & = & \mu/\sigma^2 & = & 12/4^2 & = & 0.75. \end{array} \right.$$

STEP 2, PROBABILITY. We can now compute the probability,

$$\boldsymbol{P}\{8 < T < 10\} = F_T(10) - F_T(8). \tag{4.15}$$

For each term in (4.15), we use the Gamma-Poisson formula with $\alpha = 9$, $\lambda = 0.75$, and $t = 8, 10$,

$$F_T(10) = \boldsymbol{P}\{T \leq 10\} = \boldsymbol{P}\{X \geq 9\} = 1 - F_X(8) = 1 - 0.662 = 0.338$$

from Table A3 for a Poisson variable X with parameter $\lambda t = (0.75)(10) = 7.5$;

$$F_T(8) = \boldsymbol{P}\{T \leq 8\} = \boldsymbol{P}\{X \geq 9\} = 1 - F_X(8) = 1 - 0.847 = 0.153$$

from the same table, this time with parameter $\lambda t = (0.75)(8) = 6$. Then,

$$\boldsymbol{P}\{8 < T < 10\} = 0.338 - 0.153 = \underline{0.185}.$$

- Enter a value in BOTH of the first two text boxes.
- Click the **Calculate** button.
- The Calculator will compute the Poisson and Cumulative Probabilities.

Poisson random variable (x)	8
Average rate of success	7.5
Poisson Probability: $P(X = 8)$	0.137328592653878
Cumulative Probability: $P(X < 8)$	0.524638526487606
Cumulative Probability: $P(X \leq 8)$	0.661967119141484
Cumulative Probability: $P(X > 8)$	0.338032880858516
Cumulative Probability: $P(X \geq 8)$	0.475361473512394

- Enter a value in BOTH of the first two text boxes.
- Click the **Calculate** button.
- The Calculator will compute the Poisson and Cumulative Probabilities.

Poisson random variable (x)	8
Average rate of success	6
Poisson Probability: $P(X = 8)$	0.103257733530844
Cumulative Probability: $P(X < 8)$	0.743979760453717
Cumulative Probability: $P(X \leq 8)$	0.847237493984561
Cumulative Probability: $P(X > 8)$	0.152762506015439
Cumulative Probability: $P(X \geq 8)$	0.256020239546283

4.2.4 Normal distribution

Normal distribution	$$\begin{aligned} \mu &= \text{expectation, location parameter} \\ \sigma &= \text{standard deviation, scale parameter} \\ f(x) &= \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{\frac{-(x-\mu)^2}{2\sigma^2}\right\}, \quad -\infty < x < \infty \\ \mathrm{E}(X) &= \mu \\ \mathrm{Var}(X) &= \sigma^2 \end{aligned}$$

DEFINITION 4.3

Normal distribution with "standard parameters" $\mu = 0$ and $\sigma = 1$ is called **Standard Normal distribution.**

NOTATION

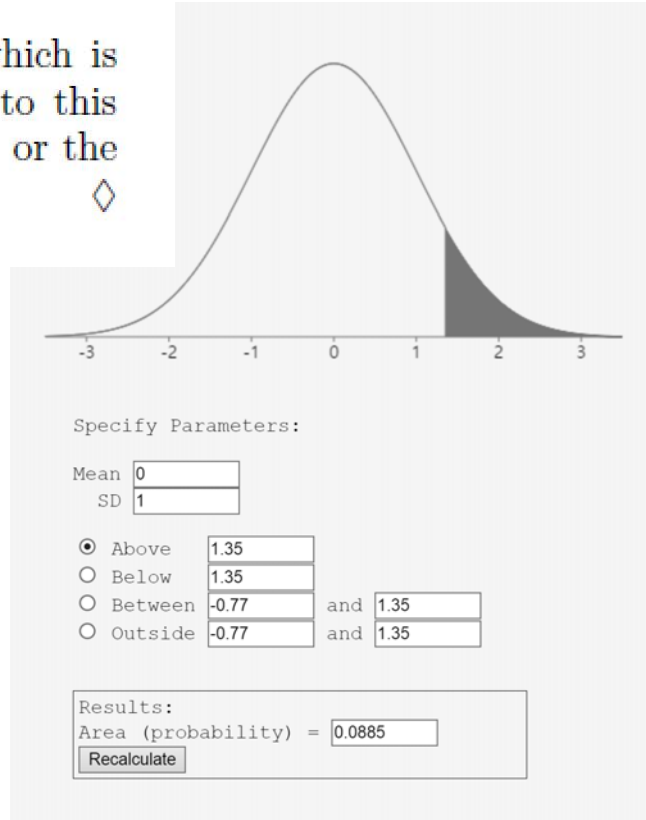
$$
\begin{aligned}
Z &= \text{Standard Normal random variable} \\
\phi(x) &= \frac{1}{\sqrt{2\pi}} e^{-x^2/2}, \text{ Standard Normal pdf} \\
\Phi(x) &= \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz, \text{ Standard Normal cdf}
\end{aligned}
$$

$$Z = \frac{X - \mu}{\sigma}. \qquad X = \mu + \sigma Z.$$

Example 4.10 (Computing Standard Normal probabilities). For a Standard Normal random variable Z,

$$\begin{aligned}
P\{Z < 1.35\} &= \Phi(1.35) = 0.9115 \\
P\{Z > 1.35\} &= 1 - \Phi(1.35) = 0.0885 \\
P\{-0.77 < Z < 1.35\} &= \Phi(1.35) - \Phi(-0.77) = 0.9115 - 0.2206 = 0.6909.
\end{aligned}$$

according to Table A4. Notice that $P\{Z < -1.35\} = 0.0885 = P\{Z > 1.35\}$, which is explained by the symmetry of the Standard Normal density in Figure 4.6. Due to this symmetry, "the left tail," or the area to the left of (-1.35) equals "the right tail," or the area to the right of 1.35. ◊



Example 4.11 (Computing non-standard Normal probabilities). Suppose that the average household income in some country is 900 coins, and the standard deviation is 200 coins. Assuming the Normal distribution of incomes, compute the proportion of "the middle class," whose income is between 600 and 1200 coins.

<u>Solution</u>. Standardize and use Table A4. For a Normal($\mu = 900$, $\sigma = 200$) variable X,

$$
\begin{aligned}
\boldsymbol{P}\{600 < X < 1200\} &= \boldsymbol{P}\left\{\frac{600-\mu}{\sigma} < \frac{X-\mu}{\sigma} < \frac{1200-\mu}{\sigma}\right\} \\
&= \boldsymbol{P}\left\{\frac{600-900}{200} < Z < \frac{1200-900}{200}\right\} = \boldsymbol{P}\{-1.5 < Z < 1.5\} \\
&= \Phi(1.5) - \Phi(-1.5) = 0.9332 - 0.0668 = \underline{0.8664}.
\end{aligned}
$$

$\Diamond$

Example 4.12 (Inverse problem). The government of the country in Example 4.11 decides to issue food stamps to the poorest 3% of households. Below what income will families receive food stamps?

Solution. We need to find such income x that $\boldsymbol{P}\{X < x\} = 3\% = 0.03$. This is an equation that can be solved in terms of x. Again, we standardize first, then use the table:

$$\boldsymbol{P}\{X < x\} = \boldsymbol{P}\left\{Z < \frac{x-\mu}{\sigma}\right\} = \Phi\left(\frac{x-\mu}{\sigma}\right) = 0.03,$$

from where

$$x = \mu + \sigma\Phi^{-1}(0.03).$$

In Table A4, we have to find the probability, the *table entry* of 0.03. We see that $\Phi(-1.88) \approx 0.03$. Therefore, $\Phi^{-1}(0.03) = -1.88$, and

$$x = \mu + \sigma(-1.88) = 900 + (200)(-1.88) = \underline{524} \text{ (coins)}$$

is the answer. In the literature, the value $\Phi^{-1}(\alpha)$ is often denoted by $z_{1-\alpha}$. ◇

4.3 Central Limit Theorem

Central limit theorem

Theorem 1 (CENTRAL LIMIT THEOREM) *Let $X_1, X_2, \ldots$ be independent random variables with the same expectation $\mu = \mathbf{E}(X_i)$ and the same standard deviation $\sigma = \mathrm{Std}(X_i)$, and let*

$$S_n = \sum_{i=1}^{n} X_i = X_1 + \ldots + X_n.$$

As $n \to \infty$, the standardized sum

$$Z_n = \frac{S_n - \mathbf{E}(S_n)}{\mathrm{Std}(S_n)} = \frac{S_n - n\mu}{\sigma\sqrt{n}}$$

converges in distribution to a Standard Normal random variable, that is,

$$F_{Z_n}(z) = \boldsymbol{P}\left\{\frac{S_n - n\mu}{\sigma\sqrt{n}} \le z\right\} \to \Phi(z) \tag{4.18}$$

for all z.

Example 4.13 (ALLOCATION OF DISK SPACE). A disk has free space of 330 megabytes. Is it likely to be sufficient for 300 independent images, if each image has expected size of 1 megabyte with a standard deviation of 0.5 megabytes?

Solution. We have $n = 300$, $\mu = 1$, $\sigma = 0.5$. The number of images n is large, so the Central Limit Theorem applies to their total size S_n. Then,

$$\begin{aligned} \boldsymbol{P}\{\text{sufficient space}\} &= \boldsymbol{P}\{S_n \le 330\} = \boldsymbol{P}\left\{\frac{S_n - n\mu}{\sigma\sqrt{n}} \le \frac{330 - (300)(1)}{0.5\sqrt{300}}\right\} \\ &\approx \Phi(3.46) = 0.9997. \end{aligned}$$

This probability is very high, hence, the available disk space is very likely to be sufficient. $\Diamond$

Example 4.14 (ELEVATOR). You wait for an elevator, whose capacity is 2000 pounds. The elevator comes with ten adult passengers. Suppose your own weight is 150 lbs, and you heard that human weights are normally distributed with the mean of 165 lbs and the standard deviation of 20 lbs. Would you board this elevator or wait for the next one?

<u>Solution</u>. In other words, is overload likely? The probability of an overload equals

$$\begin{aligned} \boldsymbol{P}\{S_{10}+150>2000\} &= \boldsymbol{P}\left\{\frac{S_{10}-(10)(165)}{20\sqrt{10}} > \frac{2000-150-(10)(165)}{20\sqrt{10}}\right\} \\ &= 1-\Phi(3.16) = 0.0008. \end{aligned}$$

Among the random variables discussed in Chapters 3 and 4, at least three have a form of S_n:

$$\begin{array}{lcl} \text{Binomial variable} & = & \text{sum of independent Bernoulli variables} \\ \text{Negative Binomial variable} & = & \text{sum of independent Geometric variables} \\ \text{Gamma variable} & = & \text{sum of independent Exponential variables} \end{array}$$

Hence, the Central Limit Theorem applies to all these distributions with sufficiently large n in the case of Binomial, k for Negative Binomial, and α for Gamma variables.

Normal approximation to Binomial distribution

Binomial variables represent a special case of $S_n = X_1 + \ldots + X_n$, where all X_i have Bernoulli distribution with some parameter p. We know from Section 3.4.5 that small p allows to approximate Binomial distribution with Poisson, and large p allows such an approximation for the number of failures. For the moderate values of p (say, $0.05 \leq p \leq 0.95$) and for large n, we can use Theorem 1:

$$\text{Binomial}(n, p) \approx Normal\left(\mu = np, \sigma = \sqrt{np(1-p)}\right) \tag{4.19}$$

Continuity correction

$$P_X(x) = \boldsymbol{P}\{X = x\} = \boldsymbol{P}\{x - 0.5 < X < x + 0.5\}$$

Example 4.15. A new computer virus attacks a folder consisting of 200 files. Each file gets damaged with probability 0.2 independently of other files. What is the probability that fewer than 50 files get damaged?

Solution. The number X of damaged files has Binomial distribution with $n = 200$, $p = 0.2$, $\mu = np = 40$, and $\sigma = \sqrt{np(1-p)} = 5.657$. Applying the Central Limit Theorem with the continuity correction,

$$\begin{aligned} \boldsymbol{P}\{X < 50\} &= \boldsymbol{P}\{X < 49.5\} = \boldsymbol{P}\left\{\frac{X-40}{5.657} < \frac{49.5-40}{5.657}\right\} \\ &= \Phi(1.68) = \underline{0.9535}. \end{aligned}$$

Notice that the properly applied continuity correction replaces 50 with 49.5, not 50.5. Indeed, we are interested in the event that X is *strictly* less than 50. This includes all values up to 49 and corresponds to the interval $[0, 49]$ that we *expand* to $[0, 49.5]$. In other words, events $\{X < 50\}$ and $\{X < 49.5\}$ are the same; they include the same possible values of X. Events $\{X < 50\}$ and $\{X < 50.5\}$ are different because the former includes $X = 50$, and the latter does not. Replacing $\{X < 50\}$ with $\{X < 50.5\}$ would have changed its probability and would have given a wrong answer. $\Diamond$

Q&A