Core course for *App.AI*, *Bioinfo.*, *Data Sci.&Eng.*, *Dec.Analytics*, *Q.Fin.*, *Risk Mgmt.* and *Stat.* Majors:

STAT2601A Probability and Statistics I (2023-2024 First Semester)

Chapter 5: Continuous Distributions

5.1 Distribution of the Continuous Type

Definition 5.1.
A random variable X is said to be of the (*absolute*) *continuous* type if its distribution function $F(x) = \Pr(X \le x)$ has the form

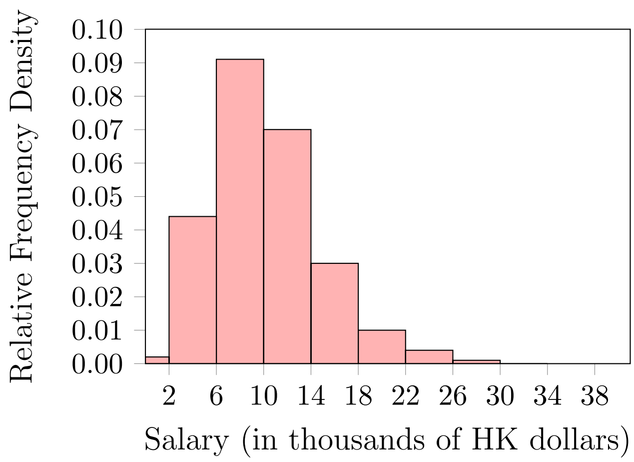
$$F(x) = \int_{-\infty}^{x} f(t)\mathrm{d}t, \qquad -\infty < x < \infty,$$

for some function $f:\ \mathbb{R} \to [0, \infty)$.
If a random variable X is of continuous type, then its probabilistic behaviour is no longer described by pmf defined by $p(x) = \Pr(X = x)$. Instead, the function f, called the *probability density function* (*pdf*), will be used. ◀

5.1.1 Probability Density Function of Continuous Random Variable

A continuous variable of a group of individuals can be represented by a histogram as shown below.

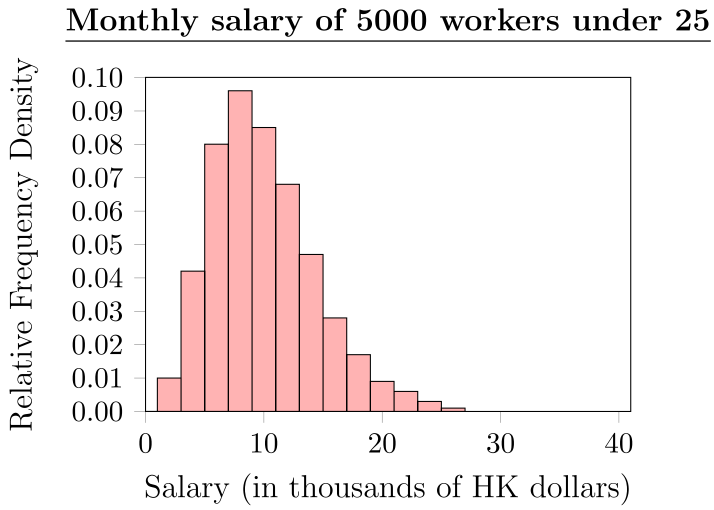
Monthly salary of 5000 workers under 25

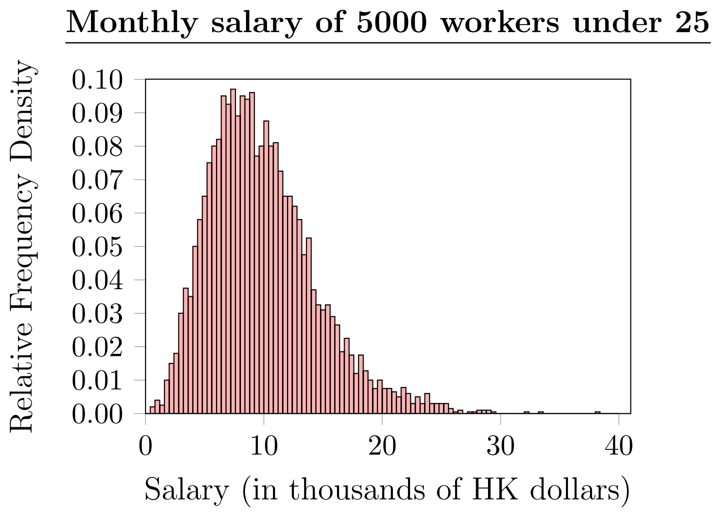


In a histogram, the area of each rectangular block should be directly proportional to the frequency of the corresponding class. Usually, the height of each block is set to the following relative frequency density so that the total area of all the blocks will be equal to 1.

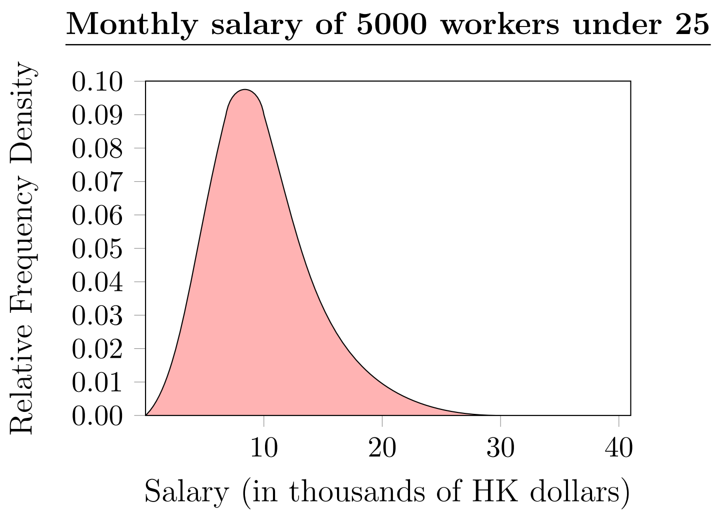
$$\text{relative frequency density} = \frac{\text{relative frequency}}{\text{class width}}.$$

For a large data set, one can use smaller class width to produce finer histogram.





Therefore for an infinite population, it is reasonable to model the distribution of a continuous variable by a smooth curve, that is, the probability density function.



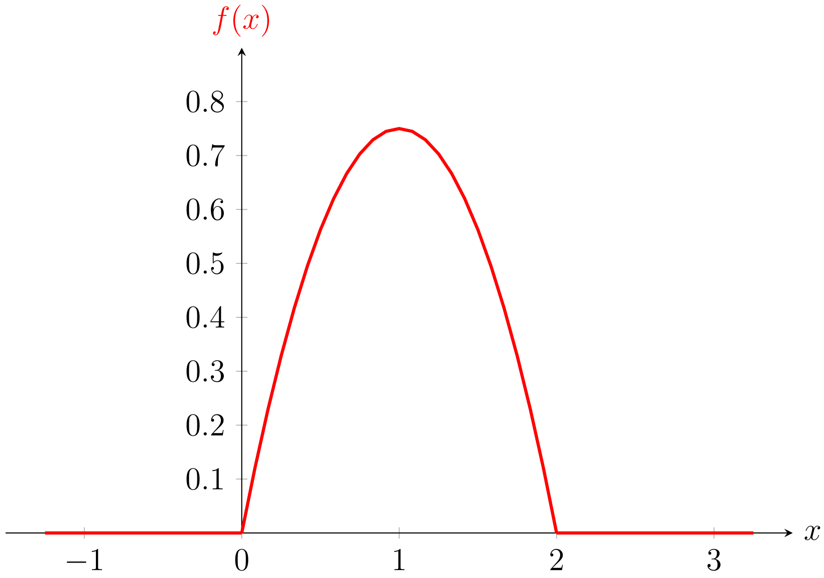
Properties of a pdf

1. If F is differentiable, then
$$f(x) = \lim_{t \to 0} \frac{\Pr(X \leq x+t) - \Pr(X \leq x)}{t} = \frac{\mathrm{d}}{\mathrm{d}x} F(x).$$
2. $f(x) \geq 0$ for all x.
3. $\int_{-\infty}^{\infty} f(x)\mathrm{d}x = F(\infty) = 1$.
4. $\Pr(X \in A) = \int_A f(x)\mathrm{d}x$ where A is any subset of $\mathbb{R}$.
5. If X is of continuous type, then
   (a) $\Pr(X = a) = \int_a^a f(x)\mathrm{d}x = 0$ for all a; and hence
   (b) $\Pr(a \leq X \leq b) = \Pr(a < X \leq b) = \Pr(a \leq X < b) = \Pr(a < X < b)$
   $= \int_a^b f(x)\mathrm{d}x = F(b) - F(a)$.
   (c) The distribution function F is continuous everywhere.

Example 5.1.

Consider the following probability density function.
$$f(x) = \begin{cases} c(2x - x^2), & \text{for } 0 < x < 2; \\ 0, & \text{otherwise.} \end{cases}$$

The constant c can be solved by
$$\int_0^2 c(2x - x^2)\mathrm{d}x = 1 \qquad \Longrightarrow c\left[x^2 - \frac{x^3}{3}\right]_0^2 = 1 \qquad \Longrightarrow \frac{4c}{3} = 1 \qquad \Longrightarrow c = \frac{3}{4}.$$



For any $a, b \in (0, 2)$ where $a < b$,

$$\Pr(a \leq X \leq b) = \int_a^b \frac{3}{4}(2x - x^2)\mathrm{d}x = \frac{1}{4}\left[3(b^2 - a^2) - (b^3 - a^3)\right].$$

Distribution function:

$$F(x) = \int_0^x \frac{3}{4}(2t - t^2)\mathrm{d}t = \frac{1}{4}(3x^2 - x^3), \qquad \text{for } 0 < x < 2.$$

$$F(x) = 0 \qquad \text{for } x \leq 0; \qquad F(x) = 1 \qquad \text{for } x \geq 2.$$

$$\Pr(0.5 \leq X \leq 1) = F(1) - F(0.5) = 0.5 - 0.15625 = 0.34375.$$

★

Remarks

Note that $f(x)$ is **not** a probability. For a very small $\varepsilon > 0$,

$$\Pr\left(a - \frac{\varepsilon}{2} \leq X \leq a + \frac{\varepsilon}{2}\right) = \int_{a-\frac{\varepsilon}{2}}^{a+\frac{\varepsilon}{2}} f(x)\mathrm{d}x \approx \varepsilon f(a).$$

Hence, $f(a)$ is a measure of how likely it is that the random variable will be near a.

5.1.2 Mean, Variance and Moment Generating Function

Definition 5.2.
If $f(x)$ is the pdf of a continuous random variable, then

$$\mathrm{E}[g(X)] = \int_{-\infty}^{\infty} g(x)f(x)\mathrm{d}x$$

is the *mathematical expectation* (*expected value*) of $g(X)$ if it exists.

◀

The properties of the mathematical expectation of a discrete random variable also apply to the mathematical expectation of a continuous random variable.

Properties

1. If c is a constant, then $\mathrm{E}(c) = c$.
2. If c is a constant, then $\mathrm{E}[cg(X)] = c\mathrm{E}[g(X)]$.
3. If $c_1, c_2, \ldots, c_n$ are constants, then $\mathrm{E}\left[\sum_{i=1}^{n} c_i g_i(X)\right] = \sum_{i=1}^{n} c_i \mathrm{E}\left[g_i(X)\right]$.
4. $X(\omega) \geq Y(\omega)$ for all $\omega \in \Omega \Longrightarrow \mathrm{E}(X) \geq \mathrm{E}(Y)$.
5. $\mathrm{E}(|X|) \geq |\mathrm{E}(X)|$.

Mean:

$$\mu = \mathrm{E}(X) = \int_{-\infty}^{\infty} x f(x) \mathrm{d}x.$$

Variance:

$$\sigma^2 = \mathrm{Var}(X) = \mathrm{E}\left[(X-\mu)^2\right] = \int_{-\infty}^{\infty} (x-\mu)^2 f(x) \mathrm{d}x = \mathrm{E}(X^2) - \mu^2.$$

Moment generating function:

$$M_X(t) = \mathrm{E}(e^{tX}) = \int_{-\infty}^{\infty} e^{tx} f(x) \mathrm{d}x,$$

$$\mathrm{E}(X^r) = M_X^{(r)}(0).$$

Cumulant generating function:

$$R_X(t) = \ln M_X(t),$$

$$\mu = R_X'(0), \quad \sigma^2 = R_X''(0).$$

Example 5.2.

Consider the following probability density function.

$$f(x) = \begin{cases} xe^{-x}, & \text{for } 0 < x < \infty; \\ 0, & \text{otherwise.} \end{cases}$$

The mgf is

$$M_X(t) = \int_0^{\infty} e^{tx} x e^{-x} \mathrm{d}x = \int_0^{\infty} x e^{-(1-t)x} \mathrm{d}x = \frac{1}{(1-t)^2}, \qquad t < 1.$$

The cgf is

$$R_X(t) = \ln M_X(t) = -2\ln(1-t).$$

And the mean and the variance are

$$\mu = R'(0) = 2 \qquad \text{and} \qquad \sigma^2 = R''(0) = 2.$$

★

5.2 Common Continuous Distributions

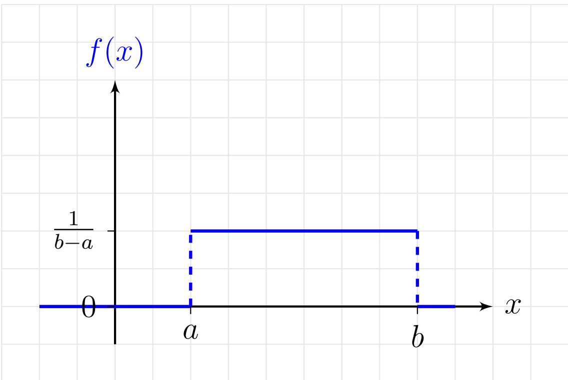
5.2.1 Uniform Distribution

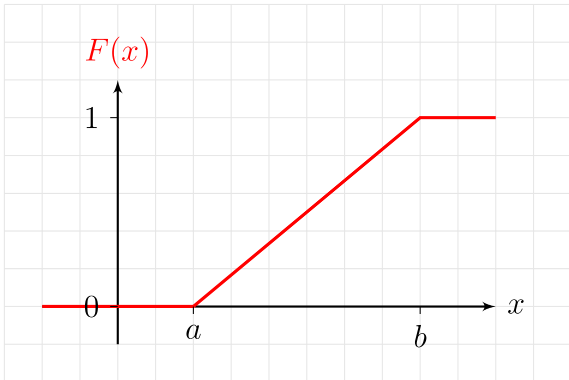
Definition 5.3.
For an interval (a, b), let X be the point randomly drawn from this interval. If the pdf of X is a constant function on (a, b), i.e.,

$$f(x) = \begin{cases} \dfrac{1}{b-a}, & \text{for } a < x < b; \\ 0, & \text{otherwise,} \end{cases}$$

then X is said to have a *uniform distribution* and is denoted as $X \sim \mathrm{U}(a, b)$.

◀





Roughly speaking, it is a point randomly selected in such a way that it has no preference to be located near any particular region.

Distribution function:

$$F(x) = \begin{cases} 0, & \text{for } x \le a; \\ \dfrac{x-a}{b-a}, & \text{for } a < x < b; \\ 1, & \text{for } x \ge b. \end{cases}$$

Mean:

$$\mu = \mathrm{E}(X) = \frac{a+b}{2} \quad \text{(midpoint of the interval).}$$

Variance:

$$\sigma^2 = \mathrm{Var}(X) = \mathrm{E}(X^2) - \mu^2 = \frac{(b-a)^2}{12}.$$

Example 5.3.

A straight rod drops freely onto a horizontal plane. Let X be the angle between the rod and North direction: $0 \le X \le 2\pi$. Then, $X \sim \mathrm{U}(0, 2\pi)$.

$$\mu = \frac{0+2\pi}{2} = \pi, \qquad \sigma^2 = \frac{(2\pi - 0)^2}{12} = \frac{\pi^2}{3}.$$

$$F(x) = \frac{x-0}{2\pi - 0} = \frac{x}{2\pi} \qquad \text{for } 0 \le x < 2\pi.$$

$$\begin{aligned}
\Pr(\text{pointing towards direction between NE and E}) &= \Pr\left(\frac{\pi}{4} \le X \le \frac{\pi}{2}\right) \\
&= F\left(\frac{\pi}{2}\right) - F\left(\frac{\pi}{4}\right) \\
&= \frac{1}{2\pi}\left(\frac{\pi}{2} - \frac{\pi}{4}\right) \\
&= \frac{1}{8}.
\end{aligned}$$

★

Property
Let $X \sim \mathrm{U}(0,1)$ and $Y = cX + d$, then

$$Y \sim \mathrm{U}(d, c+d), \qquad \text{if } c \text{ is positive;}$$

$$Y \sim \mathrm{U}(c+d, d), \qquad \text{if } c \text{ is negative.}$$

Proof. Consider

$$F_X(x) = \Pr(X \le x) = \frac{x-0}{1-0} = x \qquad \text{for } 0 \le x \le 1.$$

If $c > 0$, then the cdf of Y is given by

$$\begin{aligned}
F_Y(y) &= \Pr(Y \le y) = \Pr(cX + d \le y) \\
&= \Pr\left(X \le \frac{y-d}{c}\right) = \frac{y-d}{c} = \frac{y-d}{(c+d)-d} \qquad \text{for } d \le y \le c+d.
\end{aligned}$$

Compare with the cdf of a uniform random variable, we have

$$Y \sim \mathrm{U}(d, c+d).$$

Similarly for $c < 0$, then

$$Y \sim \mathrm{U}(c+d, d).$$

□

Example 5.4.

If $X \sim \mathrm{U}(0,1)$, then $2X \sim \mathrm{U}(0,2)$ and $2X - 1 \sim \mathrm{U}(-1,1)$.

If $X \sim \mathrm{U}(-2,6)$, then $\dfrac{X+2}{8} \sim \mathrm{U}(0,1)$. Also, $6 - X \sim \mathrm{U}(0,8)$ and $\dfrac{6-X}{8} \sim \mathrm{U}(0,1)$.

★

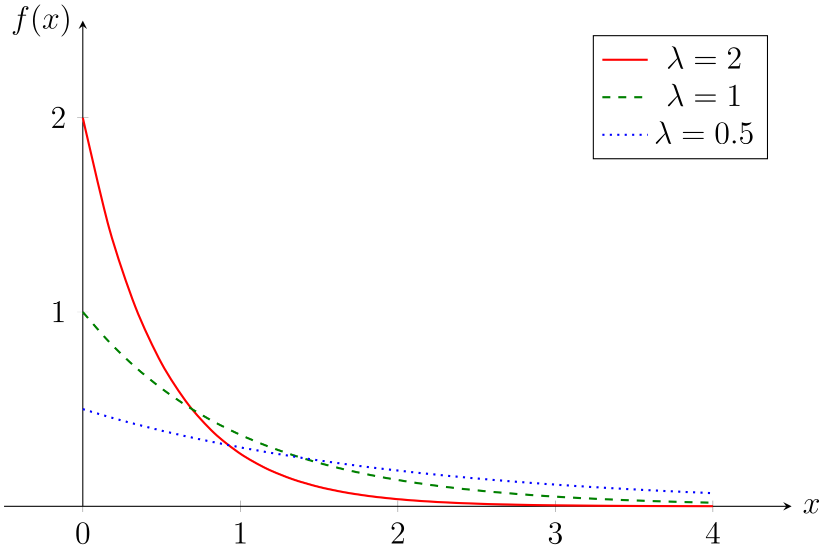
5.2.2 Exponential Distribution

Definition 5.4.
Let X be a positive random variable with pdf

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & \text{for } x > 0; \\ 0, & \text{for } x \le 0, \end{cases}$$

then X is said to have an *exponential distribution* and is denoted as $X \sim \text{Exp}(\lambda)$. ◄

Probability Density Functions of Exp(λ)



Recall in Poisson process,

$$\text{Random variable } N(t) = \text{number of occurrences in a time interval } [0, t]$$

$$p(y) = \Pr(N(t) = y) = \frac{e^{-\lambda t}(\lambda t)^y}{y!}, \qquad y = 0, 1, 2, \ldots.$$

Define random variable X be the waiting time until the first occurrence in a Poisson process with rate λ. Then the distribution function of X can be derived as follows:

For $x \le 0$,

$$F(x) = \Pr(X \le x) = 0.$$

For $x > 0$,

$$\begin{aligned} F(x) &= \Pr(X \le x) = 1 - \Pr(X > x) \\ &= 1 - \Pr(\text{no occurrence in time interval } [0, x]) \\ &= 1 - \Pr(N(x) = 0) = 1 - e^{-\lambda x} \qquad (\text{Set } y = 0 \text{ in } p(y).) \end{aligned}$$

Therefore, the cdf of X is

$$F(x) = \begin{cases} 1 - e^{-\lambda x}, & \text{for } x > 0; \\ 0, & \text{for } x \le 0. \end{cases}$$

Hence, the pdf of X is given by

$$f(x) = F'(x) = \lambda e^{-\lambda x}, \qquad \text{for } x > 0.$$

Therefore X is distributed as exponential with parameter λ. An exponential random variable therefore can describe the random time elapsing between unpredictable events (e.g., telephone calls, earthquakes, arrivals of buses or customers, etc.)

The moment generating function of $\text{Exp}(\lambda)$ is given by

$$M_X(t) = \int_0^\infty e^{tx}\lambda e^{-\lambda x}\mathrm{d}x = \int_0^\infty \lambda e^{-(\lambda - t)x}\mathrm{d}x = \left[-\frac{\lambda e^{-(\lambda-t)x}}{\lambda - t}\right]_0^\infty = \frac{\lambda}{\lambda - t}, \qquad t < \lambda.$$

The cumulant generating function is then

$$R_X(t) = \ln M_X(t) = \ln\lambda - \ln(\lambda - t).$$

$$R'(t) = \frac{1}{\lambda - t}, \qquad R''(t) = \frac{1}{(\lambda - t)^2},$$

$$\mu = R'(0) = \frac{1}{\lambda}, \qquad \sigma^2 = R''(0) = \frac{1}{\lambda^2}.$$

In summary, we have for $X \sim \text{Exp}(\lambda)$,

Distribution function:

$$F(x) = \begin{cases} 1 - e^{-\lambda x}, & \text{for } x > 0; \\ 0, & \text{for } x \le 0. \end{cases}$$

Moment generating function:

$$M_X(t) = \frac{\lambda}{\lambda - t}, \qquad t < \lambda.$$

Mean and variance:

$$\mu = \frac{1}{\lambda} \qquad \text{and} \qquad \sigma^2 = \frac{1}{\lambda^2}.$$

Example 5.5.

Let X be the failure time of a machine. Assume that X follows an exponential distribution with rate $\lambda = 2$ failures per month. Then the average failure time is $\mathrm{E}(X) = \dfrac{1}{2}$ month. The probability that the machine can run over 4 months without any failure is given by

$$\Pr(X > 4) = 1 - F(4) = 1 - \left(1 - e^{-2(4)}\right) = e^{-8} = 0.000335.$$

★

Remarks

For any $a > 0$ and $b > 0$,

$$\begin{aligned}\Pr(X > a+b) &= 1 - F(a+b) \\ &= e^{-\lambda(a+b)} \\ &= e^{-\lambda a}e^{-\lambda b} \\ &= (1-F(a))\,(1-F(b)) \\ &= \Pr(X > a)\Pr(X > b).\end{aligned}$$

This implies

$$\Pr(X > a+b | X > a) = \Pr(X > b).$$

That is, knowing that event has not occurred in the past a units of time does not alter the distribution of the arrival time in the future, i.e., we may assume the process starts *afresh* at any point of observation.

Among all continuous random variable with support $(0, \infty)$, the exponential distribution is the only distribution that has the *memoryless property*.

Example 5.6.

In the previous example (**Example 5.5.**), what is the probability that the machine can run 4 months more given that it has run for 1 year already?

Solution:

$\Pr(X > 16|X > 12) = \Pr(X > 4) = e^{-8} = 0.000335.$

Remarks

Another commonly used parametrization of exponential distribution is to define the pdf as

$$f(x) = \begin{cases} \dfrac{1}{\beta}e^{-\frac{x}{\beta}}, & \text{for } x > 0; \\ 0, & \text{for } x \le 0, \end{cases}$$

where $\beta > 0$ is mean, standard deviation, and scale parameter of the distribution, the reciprocal of the rate parameter, λ, defined earlier.

In this parametrization, β is a *survival parameter* that if a random variable X is considered the survival time of some systems or lives, the expected duration of survival is β units of time. The previous parametrization involving the *rate parameter* λ arises in the context of events arriving at a rate λ, when the time between events (which might be modeled using an exponential distribution) has a mean of $\beta = \lambda^{-1}$. The alternative specification with β is sometimes more convenient than the one with λ, and some authors will use it as a standard definition.

5.2.3 Gamma and Chi-Squared Distribution

Definition 5.5.
The *gamma function* is defined by

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1}e^{-x}\mathrm{d}x, \qquad \text{for } \alpha > 0.$$

It is also called the *Euler integral of the second kind.*

◀

Properties of the gamma function

1. $\Gamma(1) = 1$.
2. For $\alpha > 1$, $\Gamma(\alpha) = (\alpha - 1)\Gamma(\alpha - 1)$.
3. For any integer $n \geq 1$, $\Gamma(n) = (n-1)!$.
4. $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$.

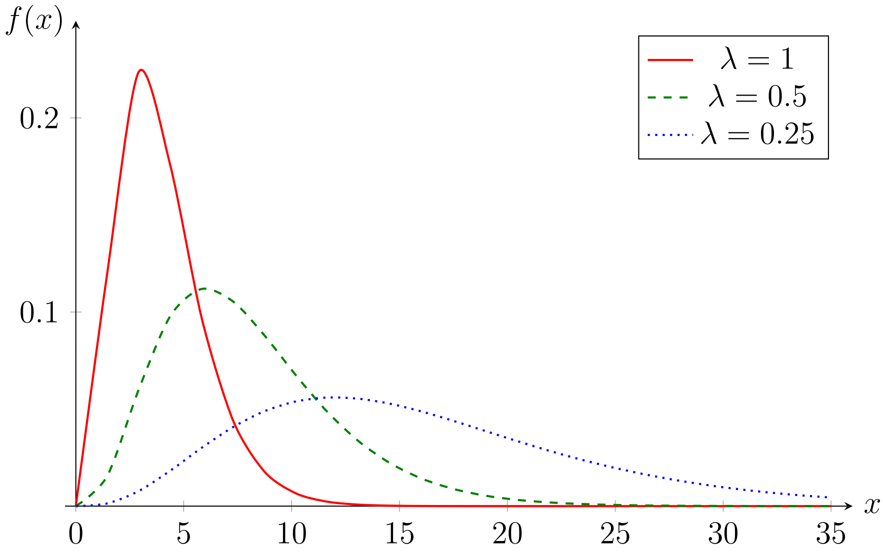
Definition 5.6.
Let X be a positive random variable with pdf

$$f(x) = \begin{cases} \dfrac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1}e^{-\lambda x}, & \text{for } x > 0; \\ 0, & \text{for } x \leq 0, \end{cases}$$

where $\alpha > 0$ and $\lambda > 0$, then X is said to have a *gamma distribution* and is denoted as $X \sim \Gamma(\alpha, \lambda)$, or $X \sim \mathrm{Gam}(\alpha, \lambda)$.

◀

Probability Density Functions of $\Gamma(\alpha = 4, \lambda)$



Gamma random variable as a waiting time

Let T_n be the waiting time until the n-th occurrence of an event according to a Poisson process with rate λ. Then the distribution function of T_n can be derived as

$$\begin{aligned} F(t) &= \Pr(T_n \le t) = 1 - \Pr(T_n > t) \\ &= 1 - \Pr(N(t) < n) \\ &= 1 - \sum_{k=0}^{n-1} \frac{(\lambda t)^k e^{-\lambda t}}{k!}, \qquad t > 0. \end{aligned}$$

Hence, the pdf of T_n is

$$f(t) = F'(t) = \frac{\lambda(\lambda t)^{n-1} e^{-\lambda t}}{(n-1)!} = \frac{\lambda^n}{\Gamma(n)} t^{n-1} e^{-\lambda t}, \qquad t > 0.$$

Therefore T_n is distributed as gamma with parameters $\alpha = n$ and rate λ. A gamma random variable therefore can describe the random time elapsing until the accumulation of a specific number of unpredictable events (e.g., telephone calls, earthquakes, arrivals of buses or customers, etc.)

The moment generating function of $X \sim \Gamma(\alpha, \lambda)$ is given by

$$\begin{aligned} M_X(t) &= \int_0^\infty e^{tx} \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x} \mathrm{d}x = \frac{\lambda^\alpha}{\Gamma(\alpha)} \int_0^\infty x^{\alpha-1} e^{-(\lambda - t)x} \mathrm{d}x \\ &= \frac{\lambda^\alpha}{(\lambda - t)^\alpha} \int_0^\infty \frac{(\lambda - t)^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-(\lambda - t)x} \mathrm{d}x \\ &= \left(\frac{\lambda}{\lambda - t} \right)^\alpha, \qquad t < \lambda. \end{aligned}$$

From this mgf, one can easily derive the mean and variance of X.

In summary, we have for $X \sim \Gamma(\alpha, \lambda)$,

Distribution function:

$$F(x) = \int_0^x \frac{\lambda^\alpha}{\Gamma(\alpha)} t^{\alpha-1} e^{-\lambda t} \mathrm{d}t, \qquad \text{for } x > 0.$$

Moment generating function:

$$M_X(t) = \left(\frac{\lambda}{\lambda - t}\right)^\alpha, \qquad t < \lambda.$$

Mean and variance:

$$\mu = \frac{\alpha}{\lambda} \qquad \text{and} \qquad \sigma^2 = \frac{\alpha}{\lambda^2}.$$

For a Poisson process, the mean waiting time until the n-th occurrence is $\mathrm{E}(T_n) = \dfrac{n}{\lambda}$.

Example 5.7.

Assume that the number of phone calls received by a customer service representative follows a Poisson process with rate 20 calls per hour.

Let T_8 be the waiting time (in hours) until the 8th phone call. Then $T_8 \sim \Gamma(8, 20)$.

Mean waiting time for 8 phone calls is $\mathrm{E}(T_8) = \dfrac{8}{20} = \dfrac{2}{5}$ hour.

Suppose that a customer service representative only needs to serve 8 calls before taking a break. Then the probability that he/she has to work for more than 48 minutes (i.e., 0.8 hour) before taking a rest is

$$\begin{aligned} \Pr(T_8 > 0.8) &= 1 - \Pr(T_8 \le 0.8) \\ &= 1 - \int_0^{0.8} \frac{20^8}{\Gamma(8)} x^7 e^{-20x} \mathrm{d}x = 1 - \left(1 - \sum_{k=0}^{7} \frac{16^k e^{-16}}{k!}\right) = 0.01. \end{aligned}$$

★

Properties of gamma distribution

1. The exponential distribution is a special case of the gamma distribution with $\alpha = 1$, i.e., $\Gamma(1, \lambda) \equiv \mathrm{Exp}(\lambda)$.

2. Suppose $X \sim \Gamma(\alpha, \lambda)$. Let $Y = bX$ where b is a positive number. Then,

$$Y \sim \Gamma\left(\alpha, \frac{\lambda}{b}\right).$$

Proof. The moment generating function of Y is given by

$$M_Y(t) = \mathrm{E}(e^{tY}) = \mathrm{E}(e^{tbX}) = M_X(bt) = \left(\frac{\lambda}{\lambda - bt}\right)^\alpha = \left(\frac{\frac{\lambda}{b}}{\frac{\lambda}{b} - t}\right)^\alpha, \qquad t < \frac{\lambda}{b},$$

which is the moment generating function of $\Gamma\left(\alpha, \dfrac{\lambda}{b}\right)$. □

3. The parameter α is called the *shape parameter* while λ is called the *rate parameter*. According to property 2, we may tabulate the distribution function for the gamma distribution with a standardized scale parameter.

4. Similar to exponential distribution, there is another parametrization of gamma distribution. Consider a *shape parameter* of $k > 0$ and a *scale parameter* of $\theta > 0$, the corresponding pdf is

$$f(x) = \begin{cases} \dfrac{1}{\Gamma(k)\theta^k} x^{k-1} e^{-\frac{x}{\theta}}, & \text{for } x > 0; \\ 0, & \text{for } x \leq 0. \end{cases}$$

Essentially, $k = \alpha$ and $\theta = \lambda^{-1}$.

Definition 5.7.
Let r be a positive integer. If $X \sim \Gamma(\alpha, \lambda)$ with $\alpha = \dfrac{r}{2}$, $\lambda = \dfrac{1}{2}$, then we say that X has a *chi-squared distribution* with degrees of freedom r and is denoted by $X \sim \chi^2(r)$ or $X \sim \chi^2_r$.
Probability density function:

$$f(x) = \frac{1}{\Gamma\left(\frac{r}{2}\right) 2^{\frac{r}{2}}} x^{\frac{r}{2}-1} e^{-\frac{x}{2}}, \qquad x > 0.$$

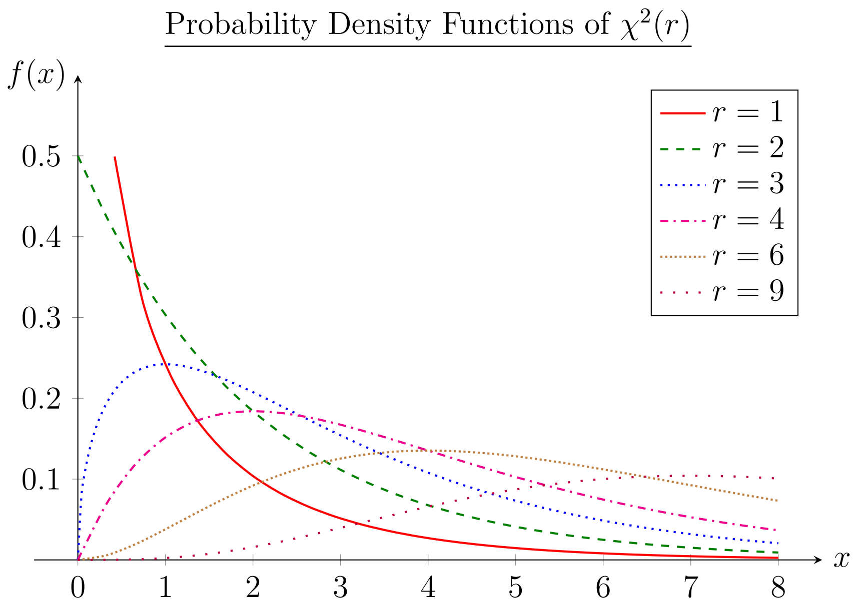
Moment generating function:

$$M_X(t) = \frac{1}{(1-2t)^{\frac{r}{2}}}, \qquad t < \frac{1}{2}.$$

Mean and variance:

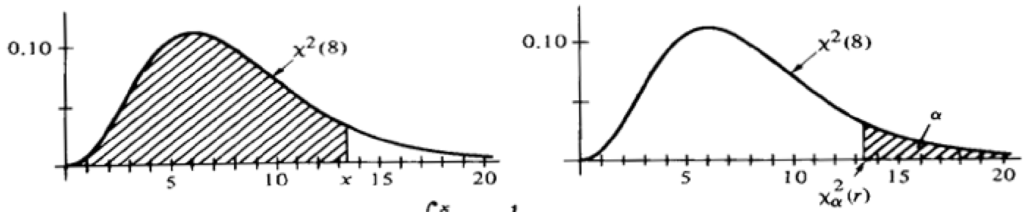
$$\mu = r \qquad \text{and} \qquad \sigma^2 = 2r.$$

◀



The values of the chi-squared distribution function are tabulated.

Table V The Chi-Square Distribution



$$P(X \le x) = \int_0^x \frac{1}{\Gamma(r/2)2^{r/2}} w^{r/2-1} e^{-w/2}\, dw$$

	$P(X \le x)$							
	0.010	0.025	0.050	0.100	0.900	0.950	0.975	0.990
r	$\chi^2_{0.99}(r)$	$\chi^2_{0.975}(r)$	$\chi^2_{0.95}(r)$	$\chi^2_{0.90}(r)$	$\chi^2_{0.10}(r)$	$\chi^2_{0.05}(r)$	$\chi^2_{0.025}(r)$	$\chi^2_{0.01}(r)$
1	0.000	0.001	0.004	0.016	2.706	3.841	5.024	6.635
2	0.020	0.051	0.103	0.211	4.605	5.991	7.378	9.210
3	0.115	0.216	0.352	0.584	6.251	7.815	9.348	11.34
4	0.297	0.484	0.711	1.064	7.779	9.488	11.14	13.28
5	0.554	0.831	1.145	1.610	9.236	11.07	12.83	15.09
6	0.872	1.237	1.635	2.204	10.64	12.59	14.45	16.81
7	1.239	1.690	2.167	2.833	12.02	14.07	16.01	18.48
8	1.646	2.180	2.733	3.490	13.36	15.51	17.54	20.09
9	2.088	2.700	3.325	4.168	14.68	16.92	19.02	21.67
10	2.558	3.247	3.940	4.865	15.99	18.31	20.48	23.21
11	3.053	3.816	4.575	5.578	17.28	19.68	21.92	24.72
12	3.571	4.404	5.226	6.304	18.55	21.03	23.34	26.22
13	4.107	5.009	5.892	7.042	19.81	22.36	24.74	27.69
14	4.660	5.629	6.571	7.790	21.06	23.68	26.12	29.14
15	5.229	6.262	7.261	8.547	22.31	25.00	27.49	30.58
16	5.812	6.908	7.962	9.312	23.54	26.30	28.84	32.00
17	6.408	7.564	8.672	10.08	24.77	27.59	30.19	33.41
18	7.015	8.231	9.390	10.86	25.99	28.87	31.53	34.80
19	7.633	8.907	10.12	11.65	27.20	30.14	32.85	36.19
20	8.260	9.591	10.85	12.44	28.41	31.41	34.17	37.57
21	8.897	10.28	11.59	13.24	29.62	32.67	35.48	38.93
22	9.542	10.98	12.34	14.04	30.81	33.92	36.78	40.29
23	10.20	11.69	13.09	14.85	32.01	35.17	38.08	41.64
24	10.86	12.40	13.85	15.66	33.20	36.42	39.36	42.98
25	11.52	13.12	14.61	16.47	34.38	37.65	40.65	44.31
26	12.20	13.84	15.38	17.29	35.56	38.88	41.92	45.64
27	12.88	14.57	16.15	18.11	36.74	40.11	43.19	46.96
28	13.56	15.31	16.93	18.94	37.92	41.34	44.46	48.28
29	14.26	16.05	17.71	19.77	39.09	42.56	45.72	49.59
30	14.95	16.79	18.49	20.60	40.26	43.77	46.98	50.89
40	22.16	24.43	26.51	29.05	51.80	55.76	59.34	63.69
50	29.71	32.36	34.76	37.69	63.17	67.50	71.42	76.15
60	37.48	40.48	43.19	46.46	74.40	79.08	83.30	88.38
70	45.44	48.76	51.74	55.33	85.53	90.53	95.02	100.4
80	53.34	57.15	60.39	64.28	96.58	101.9	106.6	112.3

This table is abridged and adapted from Table III in *Biometrika Tables for Statisticians*, edited by E. S. Pearson and H. O. Hartley. It is published here with the kind permission of the *Biometrika* Trustees.

Example 5.8.

For **Example 5.7.**, $T_8 \sim \Gamma(8, 20)$. Consider $Y = 40T_8 \sim \Gamma\left(8, \frac{1}{2}\right) \equiv \chi^2(16)$.

From the chi-squared distribution table,

$$\Pr(T_8 > 0.8) = 1 - \Pr(T_8 \le 0.8) = 1 - \Pr(Y \le 32) = 1 - 0.99 = 0.01.$$

★

5.2.4 Normal Distribution

Many naturally occurring variables have distributions that are well-approximated by a "bell-shaped curve", or a *normal distribution*. These variables have histograms which are approximately symmetric, have a single mode at the center, and tail away to both sides. Two parameters, the mean μ and the standard deviation σ describe a normal distribution completely and allow one to approximate the proportions of observations in any interval by finding corresponding areas under the appropriate normal curve.

Definition 5.8.
A random variable X is said to have a *normal distribution* (*Gaussian distribution*) if its pdf is defined by

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \qquad -\infty < x < \infty,$$

where $-\infty < \mu < \infty$ and $\sigma^2 > 0$ are the *location parameter* and *scale parameter* respectively. It is denoted as $X \sim \mathrm{N}(\mu, \sigma^2)$.
Distribution function:

$$F(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(t-\mu)^2}{2\sigma^2}} \mathrm{d}t, \qquad -\infty < x < \infty.$$

Moment generating function:

$$M_X(t) = \exp\left(\mu t + \frac{1}{2}\sigma^2 t^2\right), \qquad t \in \mathbb{R}.$$

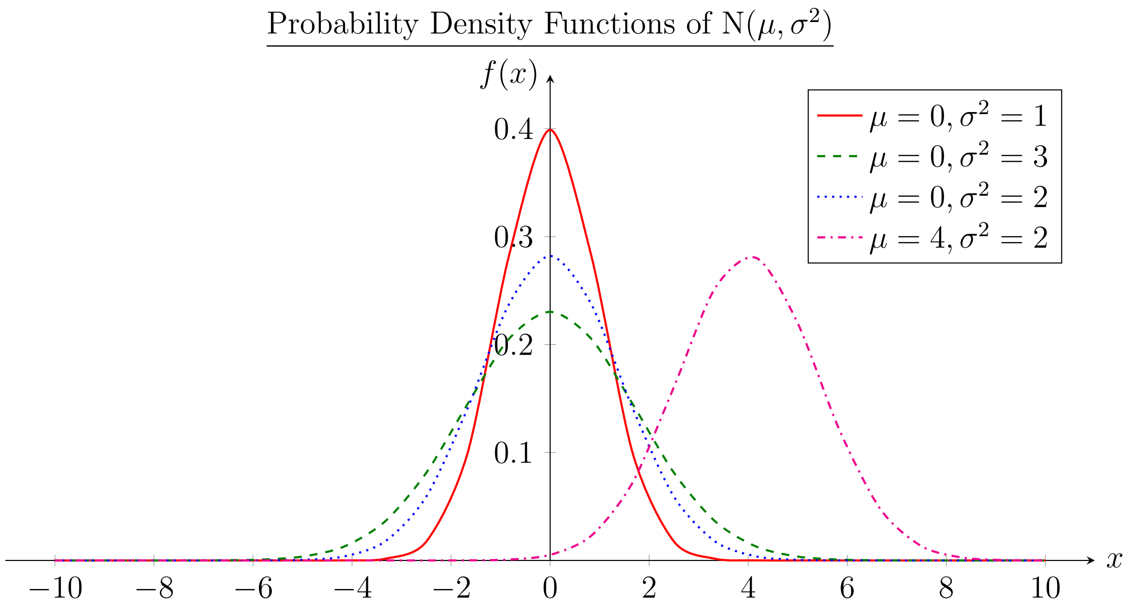
Mean and variance:

$$\mathrm{E}(X) = \mu \qquad \text{and} \qquad \mathrm{Var}(X) = \sigma^2.$$

◀

Characteristics of all normal curves

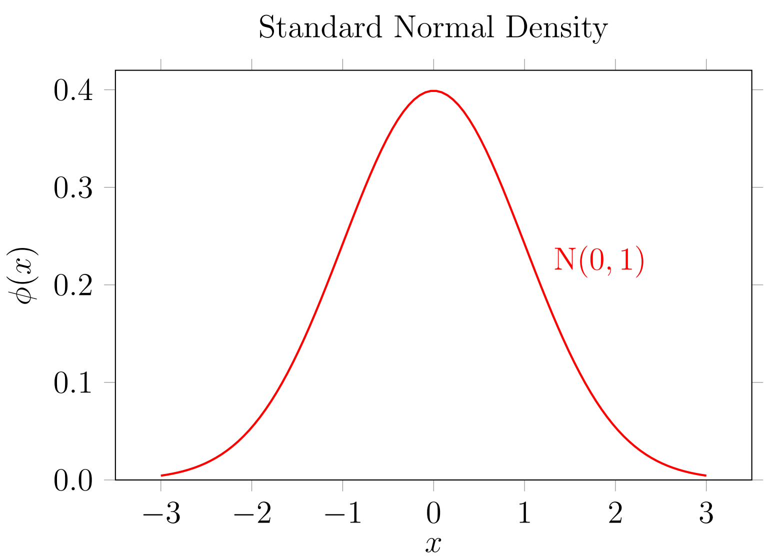
1. Each bell-shaped normal curve is symmetric and centered at its mean μ.
2. About 68% of the area is within one standard deviation of the mean, about 95% of the area is within two standard deviations of the mean, and almost all (99.7%) of the area is within three standard deviations of the mean.
3. The places where the normal curve is the steepest are one standard deviation below and above the mean ($\mu - \sigma$ and $\mu + \sigma$).



The standard normal distribution

Areas under all normal curves are related. For example, the area to the right of 1.76 standard deviations above the mean is identical for all normal curves. Because of this, we can find an area over an interval for any normal curve by finding the corresponding area under a *standard normal curve* which has mean $\mu = 0$ and standard deviation $\sigma = 1$.

If $\mu = 0$ and $\sigma^2 = 1$, then $Z \sim \mathrm{N}(0, 1)$ is said to have the *standard normal distribution*.



Usually the probability density function and distribution function of the standard normal distribution are denoted as

$$\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}, \qquad -\infty < z < \infty,$$
$$\Phi(z) = \Pr(Z \le z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} \mathrm{d}t, \qquad -\infty < z < \infty.$$

The values of the standard normal distribution function are tabulated.

The standard normal distribution table

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998

The standard normal table tells the area under a standard normal curve to the left of a number z, i.e., the value of $\Phi(z)$. The number z is rounded to two decimal places in the table. The ones place and the tenths place are listed at the left side of the table. The hundredths place is listed across the top. Areas are listed in the table entries. The standard normal table here does not have information for negative z values. Because of symmetry, this is not necessary, one may consider $\Phi(-z) = 1 - \Phi(z)$.

Properties

1. Normal distribution is symmetric about its mean. That is, if $X \sim \mathrm{N}(\mu, \sigma^2)$, then
$$\Pr(X \le \mu - x) = \Pr(X \ge \mu + x).$$
In particular, $\Phi(x) = 1 - \Phi(-x)$.

2. If $X \sim \mathrm{N}(\mu, \sigma^2)$, then $aX + b \sim \mathrm{N}(a\mu + b, a^2\sigma^2)$. In particular, the normal score $Z = \dfrac{X-\mu}{\sigma}$ is distributed as standard normal.

3. If $X \sim \mathrm{N}(0, 1)$, then $X^2 \sim \chi^2(1)$.

Standardization

In working with normal curves, the first step in a calculation is invariably to standardize.
$$Z = \frac{X-\mu}{\sigma} \sim \mathrm{N}(0,1).$$
This normal score (standard score, z-score) tells how many standard deviations an observation X is from its mean. Positive z-score indicates that X is greater than the mean; negative z-score indicates that X is below the mean.

If the z-score is known and the value of X is needed, solving the previous equation of X gives
$$X = \mu + Z \times \sigma.$$
In words, this states that X is Z standard deviations above the mean.

Example 5.9.

Studies show that gasoline use for compact cars sold in the United States is normally distributed, with a mean use of 30.5 miles per gallon (mpg) and a standard deviation of 4.5 mpg. What percentage of compact cars use between 25 and 40 mpg?

Solution:

Let X mpg be the gasoline used for a compact car. Then, $X \sim \mathrm{N}(30.5, 4.5^2)$.
$$\begin{aligned}
\Pr(25 \le X \le 40) &= \Pr\left(\frac{25-30.5}{4.5} \le \frac{X-30.5}{4.5} \le \frac{40-30.5}{4.5}\right) \\
&\approx \Pr(-1.22 \le Z \le 2.11), \qquad Z \sim \mathrm{N}(0,1) \\
&= \Phi(2.11) - [1 - \Phi(1.22)] \\
&= 0.9826 - (1 - 0.8888) \\
&= 0.8714 \\
&= 87.14\%.
\end{aligned}$$

★

In general, if $X \sim \mathrm{N}(\mu, \sigma^2)$, then $\Pr(a \le X \le b) = \Phi\left(\dfrac{b-\mu}{\sigma}\right) - \Phi\left(\dfrac{a-\mu}{\sigma}\right)$.

Example 5.10.

In general, for a normal random variable $X \sim \mathrm{N}(\mu, \sigma^2)$, the probability of that its value is within one standard deviation from the mean is

$$\Pr(\mu - \sigma \leq X \leq \mu + \sigma) = \Phi(1) - \Phi(-1) = 0.682 = 68.2\%.$$

Similarly, the probability of that its value is within two standard deviations from the mean is

$$\Pr(\mu - 2\sigma \leq X \leq \mu + 2\sigma) = \Phi(2) - \Phi(-2) = 0.954 = 95.4\%,$$

and the probability of that its value is within three standard deviations from the mean is

$$\Pr(\mu - 3\sigma \leq X \leq \mu + 3\sigma) = \Phi(3) - \Phi(-3) = 0.997 = 99.7\%.$$

★

Example 5.11.

A manufacturer needs washers between 0.1180 and 0.1220 inches thick; any thickness outside this range is unusable. One machine shop sells washers as \$30 per 100. Their thickness is normally distributed with a mean of 0.1200 inch and a standard deviation of 0.0010 inch.

A second machine shop sells washers at \$26 per 100. Their thickness is normally distributed with a mean of 0.1200 inch and a standard deviation of 0.0015 inch.

Which shop offers the better deal?

Solution:

Washers purchased from shop 1: $\quad X \sim \mathrm{N}(0.12, 0.001^2)$

$$\begin{aligned}
\Pr(0.118 \leq X \leq 0.122) &= \Phi\left(\frac{0.122 - 0.12}{0.001}\right) - \Phi\left(\frac{0.118 - 0.12}{0.001}\right) \\
&= \Phi(2) - \Phi(-2) \\
&= 0.9772 - (1 - 0.9772) \\
&= 0.9544.
\end{aligned}$$

Hence on average, 95.44 of 100 washers from shop 1 are usable. Average cost of each effective washer is $\$30/95.44 = \$0.3143$.

Washers purchased from shop 2: $\quad Y \sim \mathrm{N}(0.12, 0.0015^2)$

$$\begin{aligned}
\Pr(0.118 \leq Y \leq 0.122) &= \Phi\left(\frac{0.122 - 0.12}{0.0015}\right) - \Phi\left(\frac{0.118 - 0.12}{0.0015}\right) \\
&\approx \Phi(1.33) - \Phi(-1.33) \\
&= 0.9082 - (1 - 0.9082) \\
&= 0.8164.
\end{aligned}$$

Hence on average, 81.64 of 100 washers from shop 2 are usable. Average cost of each effective washer is $\$26/81.64 = \$0.3185$.

We may conclude that washers from shop 1 are slightly better.

★

Example 5.12.

A brass polish manufacturer wishes to set his filling equipment so that in the long run only five cans in 1,000 will contain less than a desired minimum net fill of 800 gm. It is known from experience that the filled weights are approximately normally distributed with a standard deviation of 6 gm. At what level will the mean fill have to be set in order to meet this requirement?

Solution:

Let X gm be the net fill of a particular can. Then $X \sim \mathrm{N}(\mu, 36)$. The requirement can be expressed as the following probability statement:

$$
\begin{aligned}
\Pr(X < 800) = 0.005 \quad &\Longrightarrow \quad \Pr\left(\frac{X-\mu}{\sigma} < \frac{800-\mu}{6}\right) = 0.005 \\
&\Longrightarrow \quad \Phi\left(\frac{800-\mu}{6}\right) = 0.005 \\
&\Longrightarrow \quad \Phi\left(-\frac{800-\mu}{6}\right) = 1 - 0.005 = 0.995 \\
&\Longrightarrow \quad -\frac{800-\mu}{6} = 2.575 \\
&\Longrightarrow \quad \mu = 815.45.
\end{aligned}
$$

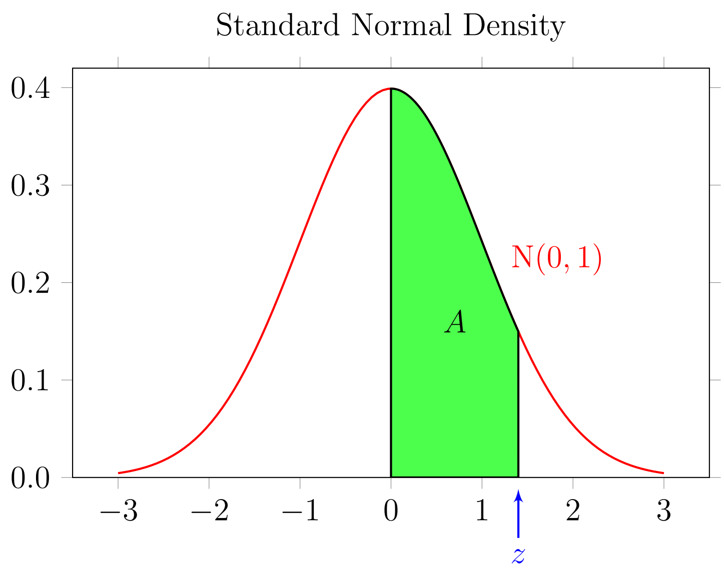
★

Remarks

Another version of the standard normal distribution table may take the follow form:

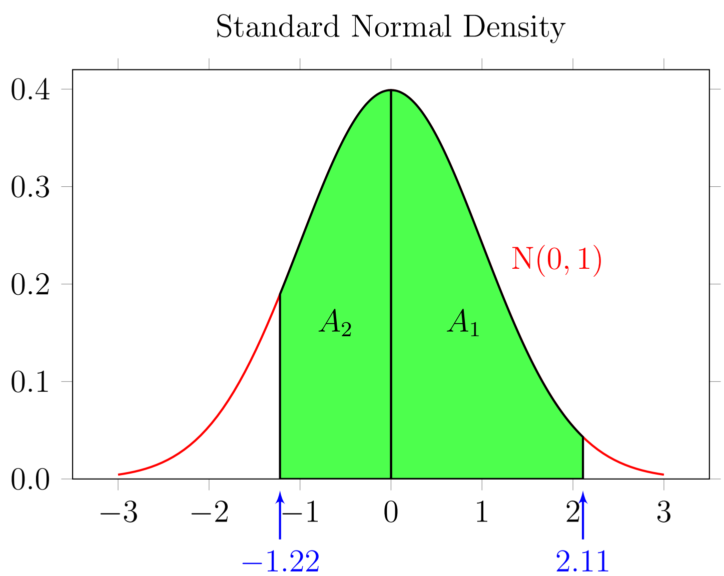
z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.0000	.0040	.0080	.0120	.0160	.0199	.0239	.0279	.0319	.0359
0.1	.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0753
0.2	.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
0.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
0.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
0.5	.1915	.1950	.1985	.2019	.2054	.2088	.2123	.2157	.2190	.2224
0.6	.2257	.2291	.2324	.2357	.2389	.2422	.2454	.2486	.2517	.2549
0.7	.2580	.2611	.2642	.2673	.2704	.2734	.2764	.2794	.2823	.2852
0.8	.2881	.2910	.2939	.2967	.2995	.3023	.3051	.3078	.3106	.3133
0.9	.3159	.3186	.3212	.3238	.3264	.3289	.3315	.3340	.3365	.3389
1.0	.3413	.3438	.3461	.3485	.3508	.3531	.3554	.3577	.3599	.3621
1.1	.3643	.3665	.3686	.3708	.3729	.3749	.3770	.3790	.3810	.3830
1.2	.3849	.3869	.3888	.3907	.3925	.3944	.3962	.3980	.3997	.4015
1.3	.4032	.4049	.4066	.4082	.4099	.4115	.4131	.4147	.4162	.4177
1.4	.4192	.4207	.4222	.4236	.4251	.4265	.4279	.4292	.4306	.4319
1.5	.4332	.4345	.4357	.4370	.4382	.4394	.4406	.4418	.4429	.4441
1.6	.4452	.4463	.4474	.4484	.4495	.4505	.4515	.4525	.4535	.4545
1.7	.4554	.4564	.4573	.4582	.4591	.4599	.4608	.4616	.4625	.4633
1.8	.4641	.4649	.4656	.4664	.4671	.4678	.4686	.4693	.4699	.4706
1.9	.4713	.4719	.4726	.4732	.4738	.4744	.4750	.4756	.4761	.4767
2.0	.4772	.4778	.4783	.4788	.4793	.4798	.4803	.4808	.4812	.4817
2.1	.4821	.4826	.4830	.4834	.4838	.4842	.4846	.4850	.4854	.4857
2.2	.4861	.4864	.4868	.4871	.4875	.4878	.4881	.4884	.4887	.4890
2.3	.4893	.4896	.4898	.4901	.4904	.4906	.4909	.4911	.4913	.4916
2.4	.4918	.4920	.4922	.4925	.4927	.4929	.4931	.4932	.4934	.4936
2.5	.4938	.4940	.4941	.4943	.4945	.4946	.4948	.4949	.4951	.4952
2.6	.4953	.4955	.4956	.4957	.4959	.4960	.4961	.4962	.4963	.4964
2.7	.4965	.4966	.4967	.4968	.4969	.4970	.4971	.4972	.4973	.4974
2.8	.4974	.4975	.4976	.4977	.4977	.4978	.4979	.4979	.4980	.4981
2.9	.4981	.4982	.4982	.4983	.4984	.4984	.4985	.4985	.4986	.4986
3.0	.4987	.4987	.4987	.4988	.4988	.4989	.4989	.4989	.4990	.4990
3.1	.4990	.4991	.4991	.4991	.4992	.4992	.4992	.4992	.4993	.4993
3.2	.4993	.4993	.4994	.4994	.4994	.4994	.4994	.4995	.4995	.4995
3.3	.4995	.4995	.4995	.4996	.4996	.4996	.4996	.4996	.4996	.4997
3.4	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4998
3.5	.4998	.4998	.4998	.4998	.4998	.4998	.4998	.4998	.4998	.4998

The table entries give the areas between zero and the value of z, as shown in the following graph:



Example 5.13.

For the normal probability calculation in **Example 5.9.**, to use the above table to find the area between -1.22 and 2.11, we can add up the two areas shown in the following graph:



From the table, the area A_1 is obtained as 0.4826. Based on the symmetric property of normal distribution, the area A_2 is equal to the area between zero and 1.22, which can be obtained directly from the table as 0.3888. Therefore, the total area between -1.22 and 2.11 is $0.4826 + 0.3888 = 87.14\%$.

Example 5.14.

Refer to **Example 5.9.**. In times of scarce energy resources, a competitive advantage is given to an automobile manufacturer who can produce a car obtaining substantially better fuel economy than the competitors' cars. If a manufacturer wishes to develop a compact car that outperforms 95% of the current compacts in fuel economy, what must the gasoline use rate for the new car be?

The gasoline use X (mpg) is normally distributed: $X \sim \mathrm{N}(30.5, 4.5^2)$. According to the manufacturer's requirement, we want to find the value c such that $\Pr(X \leq c) = 0.95$. (In fact, c is said to be the 95th percentile of the distribution of X).

Using the standard normal distribution table (of the version with areas between zero and z), we need to find the table entry as close as possible to $0.95 - 0.5 = 0.45$. We observe that the table entry is 0.4495 for $z = 1.64$ and is 0.4505 for $z = 1.65$. Therefore, the required value of the z-score is halfway between 1.64 and 1.65, i.e., $z = 1.645$.

The required value of c is thus given by $c = 30.5 + 1.645 \times 4.5 = 37.9$ mpg.

The manufacturer's new compact car must therefore obtain a fuel economy of 37.9 mpg to outperform 95% of the compact cars currently available on the U.S. market.

★

5.2.4.1 Normal Approximation to Binomial Probability

One useful application of normal distribution is to approximate the probabilities of other distributions, in particular, binomial distribution. The theory behind is indeed an application of the *central limit theorem* (*CLT*) which is introduced in Chapter 11 and is a topic in STAT2602 Probability and Statistics II. Here we only focus on the application of normal approximation to binomial probability, which is common in practice. (Recall from Chapter 4 that binomial distribution can also be approximated by Poisson distribution.)

For a random variable X following a binomial distribution $\mathrm{B}(n, p)$, it is known that

$$\mathrm{E}(X) = np \qquad \text{and} \qquad \mathrm{Var}(X) = np(1-p).$$

If n is *sufficiently large* (a classical general guideline is to require $np \geq 5$ and $n(1-p) \geq 5$ while some stricter sources request 9 rather than 5), then X can be approximated by the normal distribution $\mathrm{N}(np, np(1-p))$.

The approximation works especially well if n is large and p is close to 0.5, because in this case the binomial distribution is more symmetric like a bell-shaped normal distribution.

Continuity Correction

Since binomial distribution is a discrete distribution while normal distribution is a continuous distribution, the approximation should be made more accurate by making an adjustment of values. This is called *continuity correction*.

For example, let X be a discrete binomial random variable and Y be a continuous normal random variable to approximate X. Suppose $\Pr(X = a)$ is to be calculated where a is an integer, it is clear that using $\Pr(Y = a)$ for approximation does not work because it must be zero. Therefore, it is more appropriate to consider

$$\Pr(X = a) = \Pr(a - 0.5 < X < a + 0.5) \approx \Pr(a - 0.5 < Y < a + 0.5).$$

Example 5.15.

Recall **Example 4.22.**, $X \sim \mathrm{B}\left(8000, \frac{1}{1000}\right)$ can be approximated by Po(8) in calculating $\Pr(X < 8)$. Here we consider a normal approximation by first checking that

$$8000 \times \frac{1}{1000} = 8 \geq 5 \qquad \text{and} \qquad 8000 \times \frac{999}{1000} = 7992 \geq 5.$$

Note that

$$\begin{aligned} \mathrm{E}(X) &= 8000 \times \frac{1}{1000} = 8, \\ \mathrm{Var}(X) &= 8000 \times \frac{1}{1000} \times \frac{999}{1000} = 7.992. \end{aligned}$$

The binomial random variable X can be approximated by $Y \sim \mathrm{N}(8, 7.992)$.

Therefore,

$$\begin{aligned} \Pr(X < 8) &= \Pr(X \leq 7) \\ &\approx \Pr(Y \leq 7.5) \\ &= \Pr\left(Z \leq \frac{7.5 - 8}{\sqrt{7.992}}\right) \qquad \text{where } Z \sim \mathrm{N}(0,1) \\ &= \Pr(Z \leq -0.17686515) \\ &\approx 0.429807157. \end{aligned}$$

Note that the previous Poisson approximation gives 0.452960809 while the accurate binomial probability is 0.452890977.

We see that in this case normal approximation does not perform as accurate as Poisson approximation because $p = \frac{1}{1000}$ is quite far from 0.5.

★

5.2.5 Beta Distribution

Definition 5.9.
Let X be a positive random variable with pdf

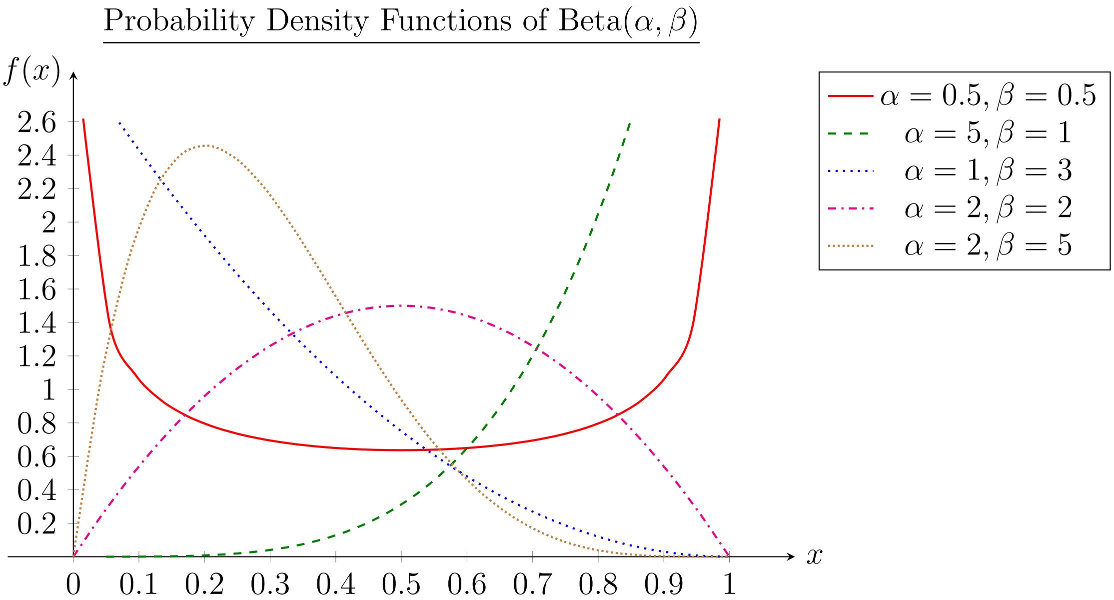
$$f(x) = \begin{cases} \dfrac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}, & \text{for } 0 < x < 1; \\ 0, & \text{otherwise,} \end{cases}$$

where α and β are positive numbers, then X is said to have a *beta distribution* and is denoted as $X \sim \mathrm{Beta}(\alpha, \beta)$.
Mean and variance:

$$\mu = \frac{\alpha}{\alpha+\beta} \qquad \text{and} \qquad \sigma^2 = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}.$$

◀

It is commonly used for modelling a random quantity distributed within a finite interval.



Remarks

1. The function
$$\mathrm{B}(\alpha, \beta) = \int_0^1 x^{\alpha-1}(1-x)^{\beta-1}\mathrm{d}x = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$$
is called the *beta function*, it is also the *Euler integral of the first kind*. Therefore the pdf of a beta distribution is sometimes expressed as
$$f(x) = \begin{cases} \dfrac{1}{\mathrm{B}(\alpha,\beta)} x^{\alpha-1}(1-x)^{\beta-1}, & \text{for } 0 < x < 1; \\ 0, & \text{otherwise.} \end{cases}$$

2. If $\alpha = \beta = 1$, the beta distribution becomes the uniform distribution $\mathrm{U}(0, 1)$.

Example 5.16.

After assessing the current political, social, economical and financial factors, a financial analyst believes that the proportion of the stocks that will increase in value tomorrow is a beta random variable with $\alpha = 5$ and $\beta = 3$. What is the expected value of this proportion? How likely is that the values of at least 70% of the stocks will move up?

Solution:

Let P be the proportion. Then, $P \sim \text{Beta}(5, 3)$ and $\mathrm{E}(P) = \dfrac{5}{5+3} = \dfrac{5}{8} = 0.625$.

$$\Pr(P \geq 0.7) = \int_{0.7}^1 \frac{\Gamma(8)}{\Gamma(5)\Gamma(3)} x^{5-1}(1-x)^{3-1}\mathrm{d}x = \frac{7!}{4!2!}\int_{0.7}^1 (x^4 - 2x^5 + x^6)\mathrm{d}x = 0.3529.$$

★

∼ End of Chapter 5 ∼