

## 5.6: The Normal Distribution

The normal distribution holds an honored role in probability and statistics, mostly because of the central limit theorem, one of the fundamental theorems that forms a bridge between the two subjects. In addition, as we will see, the normal distribution has many nice mathematical properties. The normal distribution is also called the *Gaussian distribution*, in honor of Carl Friedrich Gauss, who was among the first to use the distribution.

### The Standard Normal Distribution

#### Distribution Functions

The *standard normal distribution* is a continuous distribution on  $\mathbb{R}$  with probability density function  $\phi$  given by

$$\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}, \quad z \in \mathbb{R} \quad (5.6.1)$$

Proof that  $\phi$  is a probability density function

Let  $c = \int_{-\infty}^{\infty} e^{-z^2/2} dz$ . We need to show that  $c = \sqrt{2\pi}$ . That is,  $\sqrt{2\pi}$  is the normalizing constant for the function  $z \mapsto e^{-z^2/2}$ . The proof uses a nice trick:

$$c^2 = \int_{-\infty}^{\infty} e^{-x^2/2} dx \int_{-\infty}^{\infty} e^{-y^2/2} dy = \int_{\mathbb{R}^2} e^{-(x^2+y^2)/2} d(x, y) \quad (5.6.2)$$

We now convert the double integral to polar coordinates:  $x = r \cos \theta$ ,  $y = r \sin \theta$  where  $r \in [0, \infty)$  and  $\theta \in [0, 2\pi)$ . So,  $x^2 + y^2 = r^2$  and  $d(x, y) = r d(r, \theta)$ . Thus, converting back to iterated integrals,

$$c^2 = \int_0^{2\pi} \int_0^{\infty} r e^{-r^2/2} dr d\theta \quad (5.6.3)$$

Substituting  $u = r^2/2$  in the inner integral gives  $\int_0^{\infty} e^{-u} du = 1$  and then the outer integral is  $\int_0^{2\pi} 1 d\theta = 2\pi$ . Thus,  $c^2 = 2\pi$  and so  $c = \sqrt{2\pi}$ .

The standard normal probability density function has the famous “bell shape” that is known to just about everyone.

The standard normal density function  $\phi$  satisfies the following properties:

1.  $\phi$  is symmetric about  $z = 0$ .
2.  $\phi$  increases and then decreases, with mode  $z = 0$ .
3.  $\phi$  is concave upward and then downward and then upward again, with inflection points at  $z = \pm 1$ .
4.  $\phi(z) \rightarrow 0$  as  $z \rightarrow \infty$  and as  $z \rightarrow -\infty$ .

Proof

These results follow from standard calculus. Note that  $\phi'(z) = -z\phi(z)$  (which gives (b)) and hence also  $\phi''(z) = (z^2 - 1)\phi(z)$  (which gives (c)).

In the Special Distribution Simulator, select the normal distribution and keep the default settings. Note the shape and location of the standard normal density function. Run the simulation 1000 times, and compare the empirical density function to the probability density function.

The standard normal distribution function  $\Phi$ , given by

$$\Phi(z) = \int_{-\infty}^z \phi(t) dt = \int_{-\infty}^z \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt \quad (5.6.4)$$

and its inverse, the quantile function  $\Phi^{-1}$ , cannot be expressed in closed form in terms of elementary functions. However approximate values of these functions can be obtained from the special distribution calculator, and from most mathematics and statistics software. Indeed these functions are so important that they are considered *special functions* of mathematics.

The standard normal distribution function  $\Phi$  satisfies the following properties:

1.  $\Phi(-z) = 1 - \Phi(z)$  for  $z \in \mathbb{R}$
2.  $\Phi^{-1}(p) = -\Phi^{-1}(1-p)$  for  $p \in (0, 1)$
3.  $\Phi(0) = \frac{1}{2}$ , so the median is 0.

Proof

Part (a) follows from the symmetry of  $\phi$ . Part (b) follows from part (a). Part (c) follows from part (a) with  $z = 0$ .

In the special distribution calculator, select the normal distribution and keep the default settings.

1. Note the shape of the density function and the distribution function.
2. Find the first and third quartiles.
3. Compute the interquartile range.

In the special distribution calculator, select the normal distribution and keep the default settings. Find the quantiles of the following orders for the standard normal distribution:

1.  $p = 0.001, p = 0.999$
2.  $p = 0.05, p = 0.95$
3.  $p = 0.1, p = 0.9$

## Moments

Suppose that random variable  $Z$  has the standard normal distribution.

The mean and variance of  $Z$  are

1.  $\mathbb{E}(Z) = 0$
2.  $\text{var}(Z) = 1$

Proof

1. Of course, by symmetry, if  $Z$  has a mean, the mean must be 0, but we have to argue that the mean exists. Actually it's not hard to compute the mean directly. Note that

$$\mathbb{E}(Z) = \int_{-\infty}^{\infty} z \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz = \int_{-\infty}^0 z \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz + \int_0^{\infty} z \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz \quad (5.6.5)$$

The integrals on the right can be evaluated explicitly using the simple substitution  $u = z^2/2$ . The result is  $\mathbb{E}(Z) = -1/\sqrt{2\pi} + 1/\sqrt{2\pi} = 0$ .

2. Note that

$$\text{var}(Z) = \mathbb{E}(Z^2) = \int_{-\infty}^{\infty} z^2 \phi(z) dz \quad (5.6.6)$$

Integrate by parts, using the parts  $u = z$  and  $dv = z\phi(z) dz$ . Thus  $du = dz$  and  $v = -\phi(z)$ . Note that  $z\phi(z) \rightarrow 0$  as  $z \rightarrow \infty$  and as  $z \rightarrow -\infty$ . Thus, the integration by parts formula gives  $\text{var}(Z) = \int_{-\infty}^{\infty} \phi(z) dz = 1$ .

- 3.

In the Special Distribution Simulator, select the normal distribution and keep the default settings. Note the shape and size of the mean  $\pm$  standard deviation bar. Run the simulation 1000 times, and compare the empirical mean and standard deviation to the true mean and standard deviation.

More generally, we can compute all of the moments. The key is the following recursion formula.

For  $n \in \mathbb{N}_+$ ,  $\mathbb{E}(Z^{n+1}) = n\mathbb{E}(Z^{n-1})$

Proof

First we use the differential equation in the proof of the [PDF properties](#) above, namely  $\phi'(z) = -z\phi(z)$ .

$$\mathbb{E}(Z^{n+1}) = \int_{-\infty}^{\infty} z^{n+1} \phi(z) dz = \int_{-\infty}^{\infty} z^n z \phi(z) dz = - \int_{-\infty}^{\infty} z^n \phi'(z) dz \quad (5.6.7)$$

Now we integrate by parts, with  $u = z^n$  and  $dv = \phi'(z) dz$  to get

$$\mathbb{E}(Z^{n+1}) = -z^n \phi(z) \Big|_{-\infty}^{\infty} + \int_{-\infty}^{\infty} n z^{n-1} \phi(z) dz = 0 + n \mathbb{E}(Z^{n-1}) \quad (5.6.8)$$

The moments of the standard normal distribution are now easy to compute.

For  $n \in \mathbb{N}$ ,

1.  $\mathbb{E}(Z^{2n}) = 1 \cdot 3 \cdots (2n-1) = (2n)! / (n! 2^n)$
2.  $\mathbb{E}(Z^{2n+1}) = 0$

Proof

The result follows from the [mean and variance](#) and [recursion relation](#) above.

1. Since  $\mathbb{E}(Z) = 0$  it follows that  $\mathbb{E}(Z^n) = 0$  for every odd  $n \in \mathbb{N}$ .
2. Since  $\mathbb{E}(Z^2) = 1$ , it follows that  $\mathbb{E}(Z^4) = 1 \cdot 3$  and then  $\mathbb{E}(Z^6) = 1 \cdot 3 \cdot 5$ , and so forth. You can use induction, if you like, for a more formal proof.

Of course, the fact that the odd-order moments are 0 also follows from the symmetry of the distribution. The following theorem gives the skewness and kurtosis of the standard normal distribution.

The skewness and kurtosis of  $Z$  are

1.  $\text{skew}(Z) = 0$
2.  $\text{kurt}(Z) = 3$

Proof

1. This follows immediately from the symmetry of the distribution. Directly, since  $Z$  has mean 0 and variance 1,  $\text{skew}(Z) = \mathbb{E}(Z^3) = 0$ .
2. Since  $\mathbb{E}(Z) = 0$  and  $\text{var}(Z) = 1$ ,  $\text{kurt}(Z) = \mathbb{E}(Z^4) = 3$ .

Because of the last result, (and the use of the standard normal distribution literally as a *standard*), the *excess kurtosis* of a random variable is defined to be the ordinary kurtosis minus 3. Thus, the excess kurtosis of the normal distribution is 0.

Many other important properties of the normal distribution are most easily obtained using the moment generating function or the characteristic function.

The moment generating function  $m$  and characteristic function  $\chi$  of  $Z$  are given by

1.  $m(t) = e^{t^2/2}$  for  $t \in \mathbb{R}$ .
2.  $\chi(t) = e^{-t^2/2}$  for  $t \in \mathbb{R}$ .

Proof

1. Note that

$$m(t) = \mathbb{E}(e^{tZ}) = \int_{-\infty}^{\infty} e^{tz} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz = \int_{-\infty}^{\infty} \frac{1}{2\pi} \exp\left(-\frac{1}{2}z^2 + tz\right) dz \quad (5.6.9)$$

We complete the square in  $z$  to get  $-\frac{1}{2}z^2 + tz = -\frac{1}{2}(z-t)^2 + \frac{1}{2}t^2$ . Thus we have

$$\mathbb{E}(e^{tZ}) = e^{\frac{1}{2}t^2} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}(z-t)^2\right] dz \quad (5.6.10)$$

In the integral, if we use the simple substitution  $u = z - t$  then the integral becomes  $\int_{-\infty}^{\infty} \phi(u) du = 1$ . Hence

$$\mathbb{E}(e^{tZ}) = e^{\frac{1}{2}t^2},$$

2. This follows from (a) since  $\chi(t) = m(it)$ .

Thus, the standard normal distribution has the curious property that the characteristic function is a multiple of the probability density function:

$$\chi = \sqrt{2\pi}\phi \quad (5.6.11)$$

The moment generating function can be used to give another derivation of the moments of  $Z$ , since we know that  $\mathbb{E}(Z^n) = m^{(n)}(0)$ .

## The General Normal Distribution

The general normal distribution is the location-scale family associated with the standard normal distribution.

Suppose that  $\mu \in \mathbb{R}$  and  $\sigma \in (0, \infty)$  and that  $Z$  has the standard normal distribution. Then  $X = \mu + \sigma Z$  has the *normal distribution* with location parameter  $\mu$  and scale parameter  $\sigma$ .

### Distribution Functions

Suppose that  $X$  has the normal distribution with location parameter  $\mu \in \mathbb{R}$  and scale parameter  $\sigma \in (0, \infty)$ . The basic properties of the density function and distribution function of  $X$  follow from general results for location scale families.

The probability density function  $f$  of  $X$  is given by

$$f(x) = \frac{1}{\sigma} \phi\left(\frac{x-\mu}{\sigma}\right) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right], \quad x \in \mathbb{R} \quad (5.6.12)$$

Proof

This follows from the change of variables formula corresponding to the transformation  $x = \mu + \sigma z$ .

The probability density function  $f$  satisfies the following properties:

1.  $f$  is symmetric about  $x = \mu$ .
2.  $f$  increases and then decreases with mode  $x = \mu$ .
3.  $f$  is concave upward then downward then upward again, with inflection points at  $x = \mu \pm \sigma$ .
4.  $f(x) \rightarrow 0$  as  $x \rightarrow \infty$  and as  $x \rightarrow -\infty$ .

Proof

These properties follow from the corresponding [properties of  \$\phi\$](#) .

In the special distribution simulator, select the normal distribution. Vary the parameters and note the shape and location of the probability density function. With your choice of parameter settings, run the simulation 1000 times and compare the empirical density function to the true probability density function.

Let  $F$  denote the distribution function of  $X$ , and as above, let  $\Phi$  denote the standard normal distribution function.

The distribution function  $F$  and quantile function  $F^{-1}$  satisfy the following properties:

1.  $F(x) = \Phi\left(\frac{x-\mu}{\sigma}\right)$  for  $x \in \mathbb{R}$ .
2.  $F^{-1}(p) = \mu + \sigma \Phi^{-1}(p)$  for  $p \in (0, 1)$ .
3.  $F(\mu) = \frac{1}{2}$  so the median occurs at  $x = \mu$ .

Proof

Part (a) follows since  $X = \mu + \sigma Z$ . Parts (b) and (c) follow from (a).

In the special distribution calculator, select the normal distribution. Vary the parameters and note the shape of the density function and the distribution function.

## Moments

Suppose again that  $X$  has the normal distribution with location parameter  $\mu \in \mathbb{R}$  and scale parameter  $\sigma \in (0, \infty)$ . As the notation suggests, the location and scale parameters are also the mean and standard deviation, respectively.

The mean and variance of  $X$  are

1.  $\mathbb{E}(X) = \mu$
2.  $\text{var}(X) = \sigma^2$

Proof

This follows from the representation  $X = \mu + \sigma Z$  and basic properties of expected value and variance.

So the parameters of the normal distribution are usually referred to as the *mean* and *standard deviation* rather than *location* and *scale*. The central moments of  $X$  can be computed easily from the moments of the standard normal distribution. The ordinary (raw) moments of  $X$  can be computed from the central moments, but the formulas are a bit messy.

For  $n \in \mathbb{N}$ ,

1.  $\mathbb{E}[(X - \mu)^{2n}] = 1 \cdot 3 \cdots (2n - 1)\sigma^{2n} = (2n)!\sigma^{2n} / (n!2^n)$
2.  $\mathbb{E}[(X - \mu)^{2n+1}] = 0$

All of the odd central moments of  $X$  are 0, a fact that also follows from the symmetry of the probability density function.

In the special distribution simulator select the normal distribution. Vary the mean and standard deviation and note the size and location of the mean/standard deviation bar. With your choice of parameter settings, run the simulation 1000 times and compare the empirical mean and standard deviation to the true mean and standard deviation.

The following exercise gives the skewness and kurtosis.

The skewness and kurtosis of  $X$  are

1.  $\text{skew}(X) = 0$
2.  $\text{kurt}(X) = 3$

Proof

The skewness and kurtosis of a variable are defined in terms of the standard score, so these results follows from the [corresponding result for  \$Z\$](#) .

The moment generating function  $M$  and characteristic function  $\chi$  of  $X$  are given by

1.  $M(t) = \exp(\mu t + \frac{1}{2}\sigma^2 t^2)$  for  $t \in \mathbb{R}$ .
2.  $\chi(t) = \exp(i\mu t - \frac{1}{2}\sigma^2 t^2)$  for  $t \in \mathbb{R}$

Proof

1. This follows from the representation  $X = \mu + \sigma Z$ , basic properties of expected value, and the MGF of  $Z$  in (12):

$$\mathbb{E}(e^{tX}) = \mathbb{E}(e^{t\mu + t\sigma Z}) = e^{t\mu} \mathbb{E}(e^{t\sigma Z}) = e^{t\mu} e^{\frac{1}{2}t^2\sigma^2} = e^{t\mu + \frac{1}{2}\sigma^2 t^2} \quad (5.6.13)$$

2. This follows from (a) since  $\chi(t) = M(it)$ .

## Relations

The normal family of distributions satisfies two very important properties: invariance under linear transformations of the variable and invariance with respect to sums of independent variables. The first property is essentially a restatement of the fact that the normal distribution is a location-scale family.

Suppose that  $X$  is normally distributed with mean  $\mu$  and variance  $\sigma^2$ . If  $a \in \mathbb{R}$  and  $b \in \mathbb{R} \setminus \{0\}$ , then  $a + bX$  is normally distributed with mean  $a + b\mu$  and variance  $b^2\sigma^2$ .

Proof

The MGF of  $a + bX$  is

$$\mathbb{E} \left[ e^{t(a+bX)} \right] = e^{ta} \mathbb{E} \left[ e^{(tb)X} \right] = e^{ta} e^{\mu(tb) + \sigma^2(tb)^2/2} = e^{(a+b\mu)t + b^2\sigma^2 t^2/2} \quad (5.6.14)$$

which we recognize as the MGF of the normal distribution with mean  $a + b\mu$  and variance  $b^2\sigma^2$ .

Recall that in general, if  $X$  is a random variable with mean  $\mu$  and standard deviation  $\sigma > 0$ , then  $Z = (X - \mu)/\sigma$  is the *standard score* of  $X$ . A corollary of the last result is that if  $X$  has a normal distribution then the standard score  $Z$  has a standard normal distribution. Conversely, any normally distributed variable can be constructed from a standard normal variable.

Standard score.

1. If  $X$  has the normal distribution with mean  $\mu$  and standard deviation  $\sigma$  then  $Z = \frac{X - \mu}{\sigma}$  has the standard normal distribution.
2. If  $Z$  has the standard normal distribution and if  $\mu \in \mathbb{R}$  and  $\sigma \in (0, \infty)$ , then  $X = \mu + \sigma Z$  has the normal distribution with mean  $\mu$  and standard deviation  $\sigma$ .

Suppose that  $X_1$  and  $X_2$  are independent random variables, and that  $X_i$  is normally distributed with mean  $\mu_i$  and variance  $\sigma_i^2$  for  $i \in \{1, 2\}$ . Then  $X_1 + X_2$  is normally distributed with

1.  $\mathbb{E}(X_1 + X_2) = \mu_1 + \mu_2$
2.  $\text{var}(X_1 + X_2) = \sigma_1^2 + \sigma_2^2$

Proof

The MGF of  $X_1 + X_2$  is the product of the MGFs, so

$$\mathbb{E}(\exp[t(X_1 + X_2)]) = \exp(\mu_1 t + \sigma_1^2 t^2/2) \exp(\mu_2 t + \sigma_2^2 t^2/2) = \exp[(\mu_1 + \mu_2)t + (\sigma_1^2 + \sigma_2^2)t^2/2] \quad (5.6.15)$$

which we recognize as the MGF of the normal distribution with mean  $\mu_1 + \mu_2$  and variance  $\sigma_1^2 + \sigma_2^2$ .

This theorem generalizes to a sum of  $n$  independent, normal variables. The important part is that the sum is still normal; the expressions for the mean and variance are standard results that hold for the sum of independent variables generally. As a consequence of this result and the one for [linear transformations](#), it follows that the normal distribution is stable.

The normal distribution is stable. Specifically, suppose that  $X$  has the normal distribution with mean  $\mu \in \mathbb{R}$  and variance  $\sigma^2 \in (0, \infty)$ . If  $(X_1, X_2, \dots, X_n)$  are independent copies of  $X$ , then  $X_1 + X_2 + \dots + X_n$  has the same distribution as  $(n - \sqrt{n})\mu + \sqrt{n}X$ , namely normal with mean  $n\mu$  and variance  $n\sigma^2$ .

Proof

As a consequence of the result for [sums](#)  $X_1 + X_2 + \dots + X_n$  has the normal distribution with mean  $n\mu$  and variance  $n\sigma^2$ . As a consequence of the result for [linear transformations](#),  $(n - \sqrt{n})\mu + \sqrt{n}X$  has the normal distribution with mean  $(n - \sqrt{n})\mu + \sqrt{n}\mu = n\mu$  and variance  $(\sqrt{n})^2\sigma^2 = n\sigma^2$ .

All stable distributions are infinitely divisible, so the normal distribution belongs to this family as well. For completeness, here is the explicit statement:

The normal distribution is infinitely divisible. Specifically, if  $X$  has the normal distribution with mean  $\mu \in \mathbb{R}$  and variance  $\sigma^2 \in (0, \infty)$ , then for  $n \in \mathbb{N}_+$ ,  $X$  has the same distribution as  $X_1 + X_2 + \dots + X_n$  where  $(X_1, X_2, \dots, X_n)$  are independent, and each has the normal distribution with mean  $\mu/n$  and variance  $\sigma^2/n$ .

Finally, the normal distribution belongs to the family of general exponential distributions.

Suppose that  $X$  has the normal distribution with mean  $\mu$  and variance  $\sigma^2$ . The distribution is a two-parameter exponential family with natural parameters  $\left(\frac{\mu}{\sigma^2}, -\frac{1}{2\sigma^2}\right)$ , and natural statistics  $(X, X^2)$ .

Proof

Expanding the square, the normal PDF can be written in the form

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{\mu^2}{2\sigma^2}\right) \exp\left(\frac{\mu}{\sigma^2}x - \frac{1}{2\sigma^2}x^2\right), \quad x \in \mathbb{R} \quad (5.6.16)$$

so the result follows from the definition of the general exponential family.

A number of other special distributions studied in this chapter are constructed from normally distributed variables. These include

- The lognormal distribution
- The folded normal distribution, which includes the half normal distribution as a special case
- The Rayleigh distribution
- The Maxwell distribution
- The Lévy distribution

Also, as mentioned at the beginning of this section, the importance of the normal distribution stems in large part from the central limit theorem, one of the fundamental theorems of probability. By virtue of this theorem, the normal distribution is connected to many other distributions, by means of limits and approximations, including the special distributions in the following list. Details are given in the individual sections.

- The binomial distribution
- The negative binomial distribution
- The Poisson distribution
- The gamma distribution
- The chi-square distribution
- The student  $t$  distribution
- The Irwin-Hall distribution

### Computational Exercises

Suppose that the volume of beer in a bottle of a certain brand is normally distributed with mean 0.5 liter and standard deviation 0.01 liter.

1. Find the probability that a bottle will contain at least 0.48 liter.
2. Find the volume that corresponds to the 95th percentile

Answer

Let  $X$  denote the volume of beer in liters

1.  $\mathbb{P}(X > 0.48) = 0.9772$
2.  $x_{0.95} = 0.51645$

A metal rod is designed to fit into a circular hole on a certain assembly. The radius of the rod is normally distributed with mean 1 cm and standard deviation 0.002 cm. The radius of the hole is normally distributed with mean 1.01 cm and standard deviation 0.003 cm. The machining processes that produce the rod and the hole are independent. Find the probability that the rod is too big for the hole.

Answer

Let  $X$  denote the radius of the rod and  $Y$  the radius of the hole.  $\mathbb{P}(Y - X < 0) = 0.0028$

The weight of a peach from a certain orchard is normally distributed with mean 8 ounces and standard deviation 1 ounce. Find the probability that the combined weight of 5 peaches exceeds 45 ounces.

Answer

Let  $X$  denote the combined weight of the 5 peaches, in ounces.  $\mathbb{P}(X > 45) = 0.0127$

### A Further Generalization

In some settings, it's convenient to consider a constant as having a normal distribution (with mean being the constant and variance 0, of course). This convention simplifies the statements of theorems and definitions in these settings. Of course, the formulas for

the [probability density function](#) and the [distribution function](#) do not hold for a constant, but the other results involving the [moment generating function](#), [linear transformations](#), and [sums](#) are still valid. Moreover, the result for [linear transformations](#) would hold for all  $a$  and  $b$ .

---

This page titled [5.6: The Normal Distribution](#) is shared under a [CC BY 2.0](#) license and was authored, remixed, and/or curated by [Kyle Siegrist](#) ([Random Services](#)) via [source content](#) that was edited to the style and standards of the LibreTexts platform.