

## 2. 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

A [random variable](#)  $Z$  has the **standard normal distribution** if it has the [probability density function](#)  $\phi$  given by

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

1. Show that  $\phi$  really is a probability density function by showing that

$$\int_{-\infty}^{\infty} e^{-\frac{1}{2}z^2} dz = \sqrt{2\pi}$$

*Hint:* Let  $C$  denote the integral. Express  $C^2$  as a double integral over  $\mathbb{R}^2$  and then convert to polar coordinates.

2. Use basic calculus techniques to draw a careful sketch of the standard normal density function. In particular, show that

- $\phi$  is symmetric about  $z = 0$ .
- $\phi$  is increasing on  $(-\infty, 0)$  and decreasing on  $(0, \infty)$ .
- The mode occurs at  $z = 0$ .
- $\phi$  is concave upward on  $(-\infty, -1)$  and  $(1, \infty)$  and is concave downward on  $(-1, 1)$ .
- The inflection points of  $\phi$  occur at  $z = \pm 1$ .
- $\phi(z) \rightarrow 0$  as  $z \rightarrow \infty$  and as  $z \rightarrow -\infty$ .

3. In the [random variable experiment](#), 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, updating every 10 runs, and note the apparent convergence of the empirical density function to the true density function.

The standard normal [distribution function](#)  $\Phi$ , given by

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

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 [table of the standard normal distribution](#), the [quantile](#)

applet, and from most mathematics and statistics software.

4. Use symmetry to show that

- $\Phi(-z) = 1 - \Phi(z)$ ,  $z \in \mathbb{R}$
- $\Phi^{-1}(p) = -\Phi^{-1}(1 - p)$ ,  $p \in (0, 1)$
- The median is 0.

5. In the [quantile applet](#), select the standard normal distribution.

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

6. Use the [quantile applet](#) to find the quantiles of the following orders for the standard normal distribution:

- $p = 0.001$ ,  $p = 0.999$
- $p = 0.05$ ,  $p = 0.95$
- $p = 0.1$ ,  $p = 0.9$

### The General Normal Distribution

The general **normal distribution** is the **location-scale family** associated with the standard normal distribution. Thus, the basic properties of the density function and distribution function follow easily from general results for location scale families.

7. Show that the normal distribution with location parameter  $\mu \in \mathbb{R}$  and scale parameter  $\sigma > 0$  has **probability density function**  $f$  given by

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

8. Draw a careful sketch of the normal density function with location parameter  $\mu$  and scale parameter  $\sigma$ . In particular, show that

- $f$  is symmetric about  $x = \mu$
- $f$  is increasing on  $(-\infty, \mu)$  and is decreasing on  $(\mu, \infty)$
- The mode occurs at  $x = \mu$ .
- $f$  is concave upward on  $(-\infty, \mu - \sigma)$  and on  $(\mu + \sigma, \infty)$  and is concave downward on  $(\mu - \sigma, \mu + \sigma)$ .
- The inflection points of  $f$  occur at  $x = \mu \pm \sigma$ .
- $f(x) \rightarrow 0$  as  $x \rightarrow \infty$  and as  $x \rightarrow -\infty$

9. In the [random variable experiment](#), select the normal distribution. Vary the parameters and note the shape and location of the density function. With your choice of parameter settings, run the simulation 1000 times, updating every 10 runs and note the apparent convergence of the empirical density function to the true density function.

Let  $F$  denote the distribution function for the normal distribution with location parameter  $\mu$  and scale parameter  $\sigma$ , and as above, let  $\Phi$  denote the standard normal distribution function.

10. Show that

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

11. In the [quantile applet](#), select the normal distribution. Vary the parameters and note the shape of the density function and the distribution function.

### Moments

The important properties of the normal distribution are most easily obtained using the [moment generating function](#).

12. Suppose that  $Z$  has the standard normal distribution. Show that the moment generating function of  $Z$  is given by

$$\mathbb{E}\left(e^{tZ}\right) = e^{\frac{1}{2}t^2}, \quad t \in \mathbb{R}$$

*Hint:* In the integral for  $\mathbb{E}\left(e^{tZ}\right)$ , complete the square in  $z$  and look for a normal probability density function.

13. Suppose that  $X$  has the normal distribution with location parameter  $\mu$  and scale parameter  $\sigma$ . Use the result of the previous exercise to show that the moment generating function of  $X$  is given by

$$\mathbb{E}\left(e^{tX}\right) = \exp\left(\mu t + \frac{\sigma^2}{2} t^2\right), \quad t \in \mathbb{R}$$

As the notation suggests, the location and scale parameters are also the [mean](#) and [standard deviation](#), respectively.

14. Suppose that  $X$  has the normal distribution with location parameter  $\mu$  and scale parameter  $\sigma$ . Show that

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

More generally, we can compute all of the *central moments* of  $X$ :

15. Suppose that  $X$  has the normal distribution with mean  $\mu$  and standard deviation  $\sigma$ . Show that for  $n \in \mathbb{N}$ ,

$$\mathbb{E}\left((X - \mu)^{2n}\right) = \frac{(2n)!}{n! 2^n} \sigma^{2n}, \quad \mathbb{E}\left((X - \mu)^{2n+1}\right) = 0$$

16. In the simulation of the [random variable experiment](#), 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, updating every 10 runs and note the apparent convergence of the empirical moments to the true moments.

The following exercise gives the [skewness and kurtosis](#) of the normal distribution.

17. Suppose that  $X$  has the normal distribution with mean  $\mu$  and standard deviation  $\sigma$ . Show that
- $\text{skew}(X) = 0$
  - $\text{kurt}(X) = 3$

### Transformations

The normal family of distributions satisfies two very important properties: invariance under linear transformations 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. The proofs are easy using the moment generating function.

18. Suppose that  $X$  is normally distributed with mean  $\mu$  and variance  $\sigma^2$ . If  $a$  and  $b$  are constants and  $a$  is nonzero, show that  $aX + b$  is normally distributed with mean  $a\mu + b$  and variance  $a^2\sigma^2$ .

19. Prove the following special cases of the result in the previous exercise:

- If  $X$  has the normal distribution with mean  $\mu$  and standard deviation  $\sigma$  then  $Z = \frac{X - \mu}{\sigma}$  has the standard normal distribution.
- If  $Z$  has the standard normal distribution and if  $\mu$  and  $\sigma > 0$  are constants, then  $X = \mu + \sigma Z$  has the normal distribution with mean  $\mu$  and standard deviation  $\sigma$ .

20. Suppose that  $X_i$  is normally distributed with mean  $\mu_i$  and variance  $\sigma_i^2$  for  $i \in \{1, 2\}$ . Suppose also that  $X_1$  and  $X_2$  are independent. Show that  $X_1 + X_2$  is normally distributed with
- $\mathbb{E}(X_1 + X_2) = \mu_1 + \mu_2$
  - $\text{var}(X_1 + X_2) = \sigma_1^2 + \sigma_2^2$

The result of the previous exercise 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.

21. Suppose that  $X$  has the normal distribution with mean  $\mu$  and variance  $\sigma^2$ . Show that 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  $(X, X^2)$ .

### Computational Exercises

22. 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.
- Find the probability that a bottle will contain at least 0.48 liter.
  - Find the volume that corresponds to the 95th percentile



23. 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.



24. 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.

