Chapter 6: Some Continuous Probability Distributions

Juhyung Lee

Department of Statistics
University of Florida

Continuous Uniform Distribution

Definition (Uniform Distribution)

The pdf of the continuous uniform random variable X on the interval $[a, b]$ is

$$f(x) = \begin{cases} \dfrac{1}{b-a}, & a \leq x \leq b, \\ 0, & \text{elsewhere.} \end{cases}$$

We write $X \sim Unif(a, b)$.

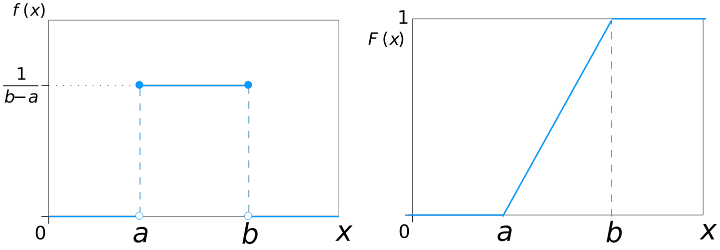

Figure: Uniform pdf and cdf.

Continuous Uniform Distribution

Theorem (Theorem 6.1 of WMMY)

The mean and variance of the uniform distribution Unif(a, b) are

$$\mu = \frac{a+b}{2} \text{ and } \sigma^2 = \frac{(b-a)^2}{12}.$$

Example (Example 6.1 of WMMY)

Suppose that a large conference room at a certain company can be reserved for no more than 4 hours. Both long and short conferences occur quite often. In fact, it can be assumed that the length X of a conference has a uniform distribution on the interval $[0, 4]$.

a. What is the pdf of X?
b. What is the probability that any given conference lasts at least 3 hours?
c. What are the mean and variance of X?

Normal Distribution

- The **normal distribution** is the most important continuous probability distribution in statistics as many variables are well described by the normal distribution (e.g., people's heights, IQs, test scores, etc.).
- The normal distribution has a **bell-shaped** pdf called the **normal curve**.
- The normal distribution is determined by two parameters μ and σ, which are in fact the mean and standard deviation of the distribution.
- The normal distribution is symmetric about the mean μ and it is also the median.

Definition (Normal Distribution)

The pdf of the normal random variable X, with mean μ and variance σ^2, is

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right], \quad -\infty < x < \infty.$$

We write $X \sim N(\mu, \sigma^2)$.

Normal Distribution

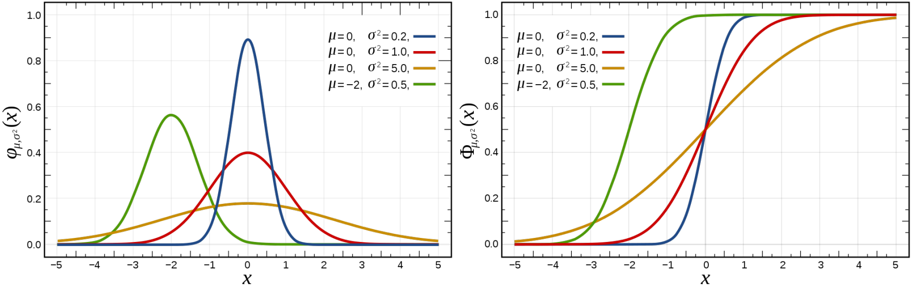

Figure: Normal pdfs and cdfs.

See also Figures 6.3, 6.4, and 6.5 of WMMY.

Areas under the Normal Curve

Definition (Standard Normal Distribution)

The normal distribution with mean 0 and variance 1 is called the standard normal distribution.

Theorem

If $X \sim N(\mu, \sigma^2)$*, then*

$$Z \triangleq \frac{X - \mu}{\sigma} \sim N(0, 1).$$

Areas under the Normal Curve

Example (Normal Probability)

Suppose $X \sim N(\mu, \sigma^2)$. Then for $-\infty < x_1 < x_2 < \infty$,

$$\begin{aligned}
P(x_1 < X < x_2) &= P\left(\frac{x_1 - \mu}{\sigma} < \frac{X - \mu}{\sigma} < \frac{x_2 - \mu}{\sigma}\right) \\
&= P\left(\frac{x_1 - \mu}{\sigma} < Z < \frac{x_2 - \mu}{\sigma}\right) \\
&= P\left(Z < \frac{x_2 - \mu}{\sigma}\right) - P\left(Z < \frac{x_1 - \mu}{\sigma}\right),
\end{aligned}$$

where $Z \sim N(0, 1)$. Therefore, we need only a standard normal table (e.g., Table A.3 of WMMY) to find a normal probability. See also Figure 6.8 of WMMY.

Areas under the Normal Curve

Normal Distribution in R

In R,

- the function `dnorm()` computes the normal pdf;
- the function `pnorm()` computes the normal cdf;
- the function `qnorm()` computes the normal quantiles.

Areas under the Normal Curve

Example (Example 6.4 of WMMY)

Consider $X \sim N(50, 10^2)$.

a. Find $P(45 < X < 62)$.
b. Find the lower and upper 0.025 quantiles of X.

Areas under the Normal Curve

Example (Example 6.4 of WMMY)

```
> # a
> pnorm(62, mean = 50, sd = 10) - pnorm(45, mean = 50, sd = 10)
[1] 0.5763928
> # alternatively,
> pnorm(1.2) - pnorm(-0.5)
[1] 0.5763928
>
> # b
> qnorm(c(0.025, 0.975), 50, 10)
[1] 30.40036 69.59964
> # alternatively,
> 50 + qnorm(0.025) * 10 # same as 50 - qnorm(0.975) * 10
[1] 30.40036
> 50 + qnorm(0.975) * 10 # same as 50 - qnorm(0.025) * 10
[1] 69.59964
```

Applications of the Normal Distribution

Example (Example 6.7 of WMMY)

A certain type of storage battery lasts, on average, 3.0 years with a standard deviation of 0.5 year. Assuming that battery life is normally distributed, find the probability that a given battery will last less than 2.3 years.

```
> pnorm(2.3, mean = 3, sd = 0.5)
[1] 0.08075666
> # alternatively,
> pnorm((2.3 - 3) / 0.5)
[1] 0.08075666
```

Applications of the Normal Distribution

Example (Example 6.10 of WMMY)

Gauges are used to reject all components for which a certain dimension is not within the specification $1.50 \pm d$. It is known that this measurement is normally distributed with mean 1.50 and standard deviation 0.2. Determine the value d such that the specifications "cover" 95% of the measurements.

```
> # 1.5 - d and 1.5 + d can be found as
> qnorm(c(0.025, 0.975), mean = 1.5, sd = 0.2)
[1] 1.108007 1.891993
> # alternatively,
> 1.5 + c(-1, 1) * qnorm(0.975) * 0.2
[1] 1.108007 1.891993
```

Chi-Squared Distribution

Definition (Chi-Squared Distribution)

The continuous random variable X has a **chi-squared distribution**, with ν (*pronounced nu*) **degrees of freedom (df)**, if its density function is given by

$$f(x) = \begin{cases} \frac{1}{2^{\nu/2}\Gamma(\nu/2)} x^{\nu/2-1} e^{-x/2}, & x > 0, \\ 0, & \text{elsewhere,} \end{cases}$$

where $\Gamma(\cdot)$ is the gamma function. We write $X \sim \chi^2_\nu$ (χ *is pronounced kai*).

Theorem (Theorem 6.5 of WMMY)

The mean and variance of the chi-squared distribution are

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

Chi-Squared Distribution

Some Properties of the Chi-Squared Distribution

- If $Z_i, i = 1, \ldots, k$, are independent standard normal random variables (i.e., $Z_i \overset{ind}{\sim} N(0,1)$), then

$$\sum_{i=1}^{k} Z_i^2 \sim \chi_k^2$$

  (e.g., $Z_1^2 \sim \chi_1^2$, $Z_1^2 + Z_2^2 \sim \chi_2^2$, $Z_2^2 + Z_4^2 + Z_6^2 \sim \chi_3^2$, etc.).

- If $X_i, i = 1, \ldots, k$, are independent chi-squared random variables with df ν_i (i.e., $X_i \overset{ind}{\sim} \chi_{\nu_i}^2$), then

$$\sum_{i=1}^{k} X_i \sim \chi^2_{\sum_{i=1}^{k} \nu_i}$$

  (i.e., the degrees of freedom add).

Chi-Squared Distribution

Chi-Squared Distribution in R

In R,

- the function `dchisq()` computes the chi-squared pdf;
- the function `pchisq()` computes the chi-squared cdf;
- the function `qchisq()` computes the chi-squared quantiles.

Example

```
> qnorm(0.975)
[1] 1.959964
> pchisq(1.96 ^ 2, df = 1)
[1] 0.9500042
>
> qchisq(0.95, df = 1)
[1] 3.841459
> pnorm(sqrt(3.841))
[1] 0.9749932
```