Econ 325

Notes on Uniform Distribution and Normal Distribution[1]

By Hiro Kasahara

Continuous Random Variables

Many types of data, such as thickness of an item, height, and weight, can take any value in some interval. A **continuous random variable** is a random variable that can take any values in some interval. Define the **cumulative distribution function** of a continuous random variable X by the probability that a random variable X takes less than or equal to some value x.

Definition 1 *The* ***cumulative distribution function (cdf)*** *of a continuous random variable X is defined by*

$$F_X(x) = P(X \leq x).$$

The properties of the cumulative distribution function are:

1. $0 \leq F_X(x) \leq 1$ for all $x \in \mathbb{R}$.
2. $F_X(x)$ is non-decreasing in x.
3. $\lim_{x\to-\infty} F_X(x) = 0$ and $\lim_{x\to\infty} F_X(x) = 1$.

When X is a continuous random variable, then $F_X(x)$ is also continuous everywhere.[2]

Let S_X be the space of possible values of X. Analogous to the probability mass function (pmf) of a discrete random variable, we define the **probability density function (pdf)** of a continuous random variable X as follows.

Definition 2 *The* ***probability density function (pdf)*** *of a continuous random variable X, denoted by $f(x)$, is a function that satisfies the following properties:*

1. *$f_X(x) \geq 0$ for any $x \in S_X$.*
2. *$\int_{x\in S_X} f_X(x)dx = 1$.*
3. *For any a and b such that $(a, b) \subset S_X$, the probability of the event $\{a < X < b\}$ is*

$$P(a < X < b) = \int_a^b f_X(x)dx.$$

For any $x \in \mathbb{R}$, we have $P(X = x) = 0$ if X is a continuous random variable.

Given the third property of the probability density function, we can express the cumulative distribution function as the integral of the probability density function:

$$F_X(x) = P(X < x) = \int_{-\infty}^{x} f_X(t)dt.$$

This has an intuitive meaning that we may obtain the value of cumulative distribution function as the area under the curve defined by the probability density function up to the value x. On the

[2]When X is a discrete random variable, then $F_X(x)$ is "right continuous," i.e., $F_X(x) = \lim_{h>0,h\to0} F_X(x + h)$ everywhere but $F_X(x)$ is not left continuous at points with positive probability mass.

other hand, the probability density function can be obtained from differentiating the cumulative density function. In fact,

$$\frac{dF_X(x)}{dx} = \frac{d\int_{-\infty}^{x} f_X(t)dt}{dx} = f_X(x)$$

for any value of x for which the derivative of $F_X(x)$ exists. Therefore, the value of probability density function can be obtained from the slope of the cumulative distribution function.

Definition 3 *If $f(x)$ is the pdf of a random variable X, then the mathematical expectation, or the expected value, of X is defined by*

$$E[X] = \int_{x\in S_X} x f_X(x)dx.$$

We often denote the expected value of X using the Greek letter μ.

Definition 4 *If $f(x)$ is the pdf of a random variable X, then the variance σ^2 and the standard deviation σ of X are defined by*

$$\sigma^2 = \int_{x\in S_X} (x-\mu)^2 f(x)dx \quad \textit{and} \quad \sigma = \sqrt{\int_{x\in S_X} (x-\mu)^2 f(x)dx},$$

respectively.

Uniform Distribution

Consider a random variable X of which outcome is a point selected at random from an interval $[a, b]$ for $-\infty < a < b < \infty$. The cumulative distribution function (cdf) and the probability density function (pdf) of X are given by

$$F(x) = \begin{cases} 0, & x < a, \\ \frac{x-a}{b-a}, & a \le x < b, \\ 1, & b \le x, \end{cases}$$

and

$$f(x) = \frac{dF(x)}{dx} = \begin{cases} \frac{1}{b-a}, & a \le x < b, \\ 0, & \text{otherwise}, \end{cases}$$

respectively. The random variable X is said to be uniformly distributed on the interval $[a, b]$ and write

$$X \sim U[a, b].$$

The uniform distribution is an example of continuous probability distribution because the support of random variable X is continuous.

The expected value of X is given by

$$\begin{aligned}\mu = E(X) = \int xf(x)dx = \int_a^b \frac{x}{b-a}dx = \frac{1}{b-a}\int_a^b xdx = \frac{1}{b-a}[(1/2)x^2]_a^b \\ = \frac{1}{2}\frac{b^2-a^2}{b-a} = \frac{1}{2}\frac{(b-a)(b+a)}{b-a} = \frac{a+b}{2}.\end{aligned}$$

The variance of X is

$$\begin{aligned}
\mathrm{Var}(X) &= \int_a^b (x-\mu)^2 \frac{1}{b-a} dx = \int_a^b (x^2 - 2\mu x + \mu^2) \frac{1}{b-a} dx = \frac{1}{b-a}[(1/3)x^3 - \mu x^2 + \mu^2 x]_a^b \\
&= \frac{1}{b-a}[(1/3)(b^3 - a^3) - \mu(b^2 - a^2) + \mu^2(b-a)] \\
&= \frac{1}{b-a}[(1/3)(b-a)(b^2 + a^2 + ab) - \mu(b-a)(b+a) + \mu^2(b-a)] \\
&= (1/3)(b^2 + a^2 + ab) - (a+b)^2/2 + (a+b)^2/4 = \frac{b^2 + a^2 - 2ab}{12} = \frac{(b-a)^2}{12},
\end{aligned}$$

where $\mu = \frac{a+b}{2}$.

Example 1 *Suppose that $X \sim U[2,6]$. Then $f(x) = \frac{1}{6-2} = \frac{1}{4}$ for $2 \leq x \leq 6$ and $= 0$ otherwise. We may compute $E[X] = \frac{2+6}{2} = 4$ and $Var[X] = \frac{(6-2)^2}{12} = 4/3$. What is the probability of $P(3 < X < 5)$? Because we can compute $P(a < X < b) = \int_a^b f(x)dx$, we have*

$$P(3 < X < 5) = \int_3^5 \frac{1}{4} dx = \left[\frac{x}{4}\right]_3^5 = \frac{5-3}{4} = \frac{1}{2}.$$

Normal Distribution

The normal distribution plays a very important role in statistics. First, in empirical applications, many variables have a "bell-shaped" frequency distribution that is approximately symmetric and has higher frequency around the mean than at the tail parts. As a result, the normal distribution approximates the probability distributions of a wide range of random variables. Second, distributions of sample means approach a normal distribution as the sample size gets large.

The probability density function and the cumulative distribution function

The probability density function (pdf) of the normal random variable X is given by

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

and its cumulative distribution function (cdf) is

$$F_x(x) = \Pr(X \leq x) = \int_{-\infty}^{x} f_x(t)dt = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(t-\mu)^2}{2\sigma^2}\right) dt. \tag{2}$$

When X is normally distributed with mean μ and variance σ^2, we write

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

It is possible to show that $\int \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) dx = 1$ (See Chapter 3.3 of Hogg, Tanis, and Zimmerman). Also, the mean and the variance of normal random variable are given as

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

Standard normal distribution and change of variables

Consider a standard normal random variable $Z \sim N(0,1)$, of which pdf is given by

$$\phi(z) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right), \quad -\infty < z < \infty$$

and its cdf is given by

$$\Phi(z) = \Pr(Z \leq z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right) dt.$$

There is no analytical expression for $\Phi(z)$ but the value of $\Phi(z)$ across different values of z can be computed in any statistical software and any textbook on econometrics or statistics report the table for $\Phi(z)$ (for example, Table 1 of Newbold, Carlson, and Thorne).

The shape of $\phi(z)$ can be analyzed by taking derivatives. Note that $\phi'(z) := \frac{d\phi(z)}{dz} = \frac{-z}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right) = -z\phi(z)$. Because $\phi'(z) = 0$ if and only if $z = 0$, the standard normal density function is flat at $z = 0$, taking the maximum value of $\phi(0) = \frac{1}{\sqrt{2\pi}}$. We may also examine how the slope of $\phi(z)$ changes by taking the second order derivatives. $\phi''(z) := \frac{d^2\phi(z)}{dz^2} = \frac{d}{dz}\phi'(z) = \frac{d}{dz}(-z\phi(z)) = -z\phi'(z) - \phi(z) = (z^2-1)\phi(z)$. Therefore, $\phi''(z) = 0$ only if $z = \pm 1$. This means that the points of inflection of the graph of the pdf of Z occurs at $z = \pm 1$; i.e., $\phi(z)$ is a concave function for $|z| < 1$ while $\phi(z)$ is a convex function for $|z| > 1$.

For $X \sim N(\mu, \sigma^2)$, we may compute $F_x(x) = \Pr(X \leq x)$ from the value of $\Phi((x-\mu)/\sigma)$, i.e., we may show that

$$F_x(x) = \Phi\left(\frac{x-\mu}{\sigma}\right) \tag{3}$$

as follows:

$$\begin{aligned} F_x(x) &= \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(t-\mu)^2}{2\sigma^2}\right) dt \\ &= \int_{-\infty}^{\frac{x-\mu}{\sigma}} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{z^2}{2}\right) \sigma dz \qquad \text{(using change of variables } z = \frac{t-\mu}{\sigma} \text{ and } dt = \sigma dz) \\ &= \Phi\left(\frac{x-\mu}{\sigma}\right), \end{aligned}$$

where the value x in the domain of X is changed into the value $\frac{x-\mu}{\sigma}$ in the domain of z in the second line. Similarly, we may show that $\Pr(a \leq X \leq b) = F_x(b) - F_x(a) = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)$.

Next example shows that the linear transformation of a normal random variable is also a random variable.

Example 2 *If $X \sim N(\mu, \sigma^2)$, then what is the distribution of $Y = a + bX$ for $b \neq 0$? Answer: Y is $N(a+b\mu, (b\sigma)^2)$. In view of equation (3), this can be shown by showing that the cdf of Y is given by $P(Y \leq y) = \Phi\left(\frac{y-\mu_y}{\sigma_y}\right)$, where $\mu_y = a + b\mu$ and $\sigma_y = b\sigma$ as follows:*

$$P(Y \leq y) = P(a + bX \leq y) = P\left(X \leq \frac{y-a}{b}\right) = \Phi\left(\frac{\frac{y-a}{b} - \mu}{\sigma}\right) = \Phi\left(\frac{y-(a+b\mu)}{b\sigma}\right).$$

The standardized value of normally distributed random variable X plays an important role in computing the probability. Given $X \sim N(\mu, \sigma^2)$, define the standardized normal random variable

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

This is the special case of Example 2 above where $a = -\mu$ and $b = \frac{1}{\sigma}$. Therefore, the standardized normal random variable has a normal distribution.

Example 3 *If $X \sim N(\mu, \sigma^2)$, then $Z := \frac{X-\mu}{\sigma} \sim N(0,1)$ so that*

$$P\left(\frac{X-\mu}{\sigma} \leq z\right) = \Phi(z).$$

.

Example 4 *Suppose $X \sim N(8, (5)^2)$. What is $P(X < 8.6)$? To answer this question, we may transform X by subtracting its mean and dividing it by its standard deviation to $\frac{X-8}{5}$ as follows.*

$$P(X < 8.6) = P\left(\frac{X-8}{5} < \frac{8.6-8}{5}\right) = P\left(Z < 0.12\right), \quad \text{where } Z \sim N(0,1).$$

Now, looking at the standardized normal probability table in the textbook (Table 1 in Appendix), we find that $P(Z < 0.12) = 0.5478$. What is $P(X > 8.6)$? We follow the similar computation to above to conclude that $P(X > 8.6) = P(Z > 0.12)$. Because $P(Z < 0.12) = 0.5478$, we have $P(X > 8.6) = P(Z > 0.12) = 1 - P(Z < 0.12) = 1 - 0.5478 = 0.4522$.

Example 5 *Suppose $X \sim N(8, (5)^2)$. What is the value of x such that $P(X < x) = 0.2$? To answer this question, we will transform X to $Z = \frac{X-8}{5}$ as*

$$P(X < x) = P\left(\frac{X-8}{5} < \frac{x-8}{5}\right) = P\left(Z < \frac{x-8}{5}\right) = 0.2, \quad \text{where } Z \sim N(0,1).$$

Looking at the standardized normal probability table in the textbook, we find that $P(Z < -0.84) = 0.2$. Therefore, the value of x that satisfies the above equation must satisfy

$$\frac{x-8}{5} = 0.84.$$

Solving this for x, we have $x = 8 + (-0.84) \times 5 = 3.80$.

Linear combination of random variables

Consider three random variables X, Y, and Z. Then,

$$E[X+Y+Z] = E[X] + E[Y] + E[Z]$$

and

$$Var[X+Y+Z] = Var[X] + Var[Y] + Var[Z] + 2Cov[X,Y] + 2Cov[X,Z] + 2Cov[Y,Z].$$

To see why the latter equation holds, define $\tilde{X} = X - E[X]$, $\tilde{Y} = Y - E[Y]$, and $\tilde{Z} = Z - E[Z]$. Then, by definition of variance,

$$\begin{aligned}
Var[X+Y+Z] &= E\left[(X+Y+Z-E[X+Y+Z])^2\right] \\
&= E\left[\{(X-E[X]) + (Y-E[Y]) + (Z-E[Z])\}^2\right] \\
&= E\left[\{\tilde{X} + \tilde{Y} + \tilde{Z}\}^2\right] \\
&= E\left[\tilde{X}^2\tilde{Y}^2 + \tilde{Z}^2 + 2\tilde{X}\tilde{Y} + 2\tilde{X}\tilde{Z} + 2\tilde{Y}\tilde{Z}\right] \\
&= E[\tilde{X}^2] + E[\tilde{Y}^2] + E[\tilde{Z}^2] + 2E[\tilde{X}\tilde{Y}] + 2E[\tilde{X}\tilde{Z}] + 2E[\tilde{Y}\tilde{Z}] \\
&= Var[X] + Var[Y] + Var[Z] + 2Cov[X,Y] + 2Cov[X,Z] + 2Cov[Y,Z].
\end{aligned}$$

We can generalize this result to a sequence of k random variables as follows.

Proposition 1 *Let X_1, X_2, ..., X_k be k random variables. Then,*

$$E[X_1 + X_2 + ... + X_k] = E[X_1] + E[X_2] + ... + E[X_k]$$

and

$$Var[X_1+X_2+...+X_k] = Var[X_1]+Var[X_2]+...+Var[X_k]+2Cov[X_1,X_2]+...+2Cov[X_{k-1},X_k]$$

where the summation for covariance terms is over all possible pairs of X_i and X_j for $i,j = 1,...,k$ with $i \neq j$.

Linear combination of normal random variables

Suppose that X and Y are two normal random variables with means μ_x and μ_y, variances σ_x^2 and σ_y^2, and covariance of X and Y given by σ_{xy}. Then, the linear function of X and Y has the normal distribution. Specifically, let a and b are some constant, then $aX + bY$ has the normal distribution so that

$$aX + bY \ \sim \ N\left(a\mu_x + b\mu_y, a^2\sigma_x^2 + b^2\sigma_y^2 + 2ab\sigma_{xy}\right).$$

Note that, *for any two random variables*, even when X and Y are not normally distributed, we have $E[aX + bY] = a\mu_x + b\mu_y$ and $Var(aX + bY) = a^2\sigma_x^2 + b^2\sigma_y^2 + 2ab\sigma_{xy}$. The unique property of normal random variables here is that the distribution of the linear combination of two normal random variables remains the normal distribution. For the random variables that are not normally distributed, the distribution of the linear combination of two random variables is generally neither the same as the original distribution nor the same as the normal distribution. For example, the linear combination two independent Bernouilli random variables neither has Bernouilli distribution nor has the normal distribution.

We can also consider the linear combination of n *independent* normal random variables, X_1, X_2, ..., X_n, with means μ_1, μ_2, ..., μ_n and variances σ_1^2, σ_2^2, ..., σ_n^2. Let c_1, c_2, ..., c_n be some constant. Then

$$\sum_{i=1}^n c_i X_i \ \sim \ N\left(\sum_{i=1}^n \mu_i, c_i^2 \sum_{i=1}^n \sigma_i^2\right) \tag{4}$$

so that the linear combination of normal random variables is normally distributed. Note that, because we assume that X_1, X_2, ..., X_n are independent to each other, the variance of $\sum_{i=1}^n c_i X_i$ does not contain the covariance terms, $2Cov(X_i, X_j)$'s.

The special case of the above result is the sample average of n independent normally distributed random variables. Suppose that X_i is independently drawn from $N(\mu, \sigma^2)$ for $i = 1, ..., n$, and define the sample average by $\bar{X} = (1/n)\sum_{i=1}^n X_i$. Then, by letting $c_i = 1/n$ and setting $\mu = \mu_i$ and $\sigma^2 = \sigma_i^2$ for $i = 1, ..., n$ in (5), we have

$$\bar{X}_i \ \sim \ N\left(\mu, \sigma^2/n\right).$$

Note that this result holds even when n is small (and hence this result does not use the Central Limit Theorem).

Proofs

We have proved many formulas when X and Y are discrete random variables in "Notes on Mathematical Expectation, Variance, and Covariance." We may apply the same line of proofs when X and Y are continuously distributed. The proofs are essentially the same by replacing the summation operator with the integral operator, where the *probability mass functions* (i.e., probability weights) are replaced with the *probability density functions*. For example, consider a continuous random variable X with the *probability density function* given by $f_X(x)$. Then,

$$
\begin{aligned}
E(a+bX) &= \int (a+bx) f_X(x) dx \\
&= \int a f_X(x) dx + \int bx f_X(x) dx \\
&= a \int f_X(x) dx + b \int x f_X(x) dx \\
&= a \times 1 + b \times E[X] \\
&= a + bE[X].
\end{aligned}
\tag{5}
$$

Let $f_{xy}(x,y)$ be the joint probability density function of X and Y, where the support of X and Y is $(-\infty,\infty)$. The marginal cumulative distribution of x can be obtained by integrating y from $f_{xy}(x,y)$ as

$$f_X(x) = \int_{-\infty}^{\infty} f_{xy}(x,y')dy'.$$

Similarly, $f_Y(y) = \int_{-\infty}^{\infty} f_{xy}(x',y)dx'$.

For any x with $f_X(x) > 0$, the probability density function of a random variable Y conditional on $X = x$ is given by

$$f_{y|x}(y|x) = \frac{f_{xy}(x,y)}{f_X(x)}.$$

When X and Y are continuously distributed, X and Y are independent if and only if

$$f_{xy}(x,y) = f_X(x) f_Y(y).$$

More examples.

1. Suppose that X and Y are continuous random variables that are independent to each other with the density function $f_X(x)$ and $f_Y(y)$. In this case, $Cov\,(X,Y) = 0$. Note that the independence means that the joint density function of X and Y, $f_{xy}(x,y)$, is equal to the product of $f_X(x)$ and $f_Y(y)$.

$$
\begin{aligned}
Cov\,(X,Y) &= E[(X-E(X))(Y-E(Y))] \\
&= \int\int (x-E(X))(y-E(Y)) f_{xy}(x,y) dxdy \\
&= \int\int (x-E(X))(y-E(Y)) f_X(x) f_Y(y) dxdy \quad \text{(by indepedence)} \\
&= \int (x-E(X)) f_X(x) dx \int (y-E(Y)) f_Y(y) dy \\
&= \left(\int x f_X(x) dx - E(X)\right)\left(\int y f_Y(y) dy - E(Y)\right) \\
&= 0 \times 0 = 0.
\end{aligned}
$$

2. Let X be continuously distributed with the density function given by $f_X(x)$. Define $Z = \frac{X-E(X)}{\sqrt{\mathrm{Var}(X)}}$. Then $E[Z] = 0$ and $\mathrm{Var}(X) = 1$. To prove this, we use $E(a+bX) = a + bE(X)$

which we prove in (5) for the case of continuous random variable.

$$
\begin{aligned}
E[Z] &= E\left(\frac{X - E(X)}{\sqrt{\text{Var}(X)}}\right) \\
&= \frac{1}{\sqrt{\text{Var}(X)}} E\left(X - E(X)\right) \quad (\text{by } E(bX) = bE(X) \text{ with } b = 1/\sqrt{\text{Var}(X)}) \\
&= \frac{1}{\sqrt{\text{Var}(X)}} \left(E(X) - E(X)\right) \quad (\text{by } E(X - a) = E(X) - a \text{ with } a = E(X)) \\
&= 0
\end{aligned}
$$

and

$$
\begin{aligned}
Var[Z] &= Var\left(\frac{X - E(X)}{\sqrt{\text{Var}(X)}}\right) \\
&= E\left(\frac{X - E(X)}{\sqrt{\text{Var}(X)}}\right)^2 \quad (\text{by } E[Z] = 0 \text{ and the definition of variance}) \\
&= E\left(\left(X - E(X)\right)^2 / \text{Var}(X)\right) \\
&= \frac{1}{\text{Var}(X)} E\left(\left(X - E(X)\right)^2\right) \quad (\text{by } E(bX) = bE(X) \text{ with } b = 1/\text{Var}(X)) \\
&= \frac{1}{\text{Var}(X)} \text{Var}(X) \quad (\text{by the definition of variance}) \\
&= 1.
\end{aligned}
$$

3. Law of Iterated Expectation: $E_Y[Y] = E_X[E_Y[Y|X]]$. Note that

$$
E_Y[Y|X = x] = \int_{-\infty}^{\infty} y' f_{y|x}(y'|x) dy' = \int_{-\infty}^{\infty} y' \frac{f_{xy}(x, y')}{f_X(x)} dy'.
$$

Therefore,

$$
\begin{aligned}
E_X[E_Y[Y|X]] &= \int_{-\infty}^{\infty} E_Y[Y|X = x] f_X(x) dx \\
&= \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} y' \frac{f_{xy}(x, y')}{f_X(x)} dy'\right) f_X(x) dx \\
&= \int_{-\infty}^{\infty} y' \left(\int_{-\infty}^{\infty} f_{xy}(x, y') dx\right) dy' \\
&= \int_{-\infty}^{\infty} y' f_Y(y') dy' \\
&= E_Y[Y].
\end{aligned}
$$