ECE 302: Probability and Applications[1]

Week 6 Topics

- Gaussian Random Variable
- Functions of a Random Variable
- Transform Methods
  - Characteristic Function
  - Probability Generating Function
  - Laplace Transform

1 The Gaussian (Normal) Random Variable

Many man-made and natural phenomena involve a random variable X that consists of the sum of a large number of "small" random variables.

- While the exact description of the pdf of X in terms of the component random variables can become quite complex and unwieldy, one finds that under very general conditions, as the number of components becomes large, the cdf of X approaches that of the Gaussian (normal) random variable.
- This result is called the **central limit theorem**.
- This random variable appears so often in problems involving randomness that it has come to be known as the "normal" random variable.

The **pdf** for the **Gaussian random variable** X is given by

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-m)^2/2\sigma^2} \qquad -\infty < x < \infty$$

where m and $\sigma > 0$ are real numbers, and which we will show are the mean and standard deviation of X.

- Figure 1 shows that the Gaussian pdf is a "bell-shaped" curve centered and symmetric about m and whose "width" increases with σ.

The cdf of the Gaussian random variable is given by

$$P[X \leq x] = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{x} e^{-(x'-m)^2/2\sigma^2} dx'$$

The change of variable $t = (x' - m)/\sigma$ results in

$$\begin{aligned} F_X(x) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{(x-m)/\sigma} e^{-t^2/2} dt \\ &= \Phi\left(\frac{x-m}{\sigma}\right) \end{aligned}$$

where $\Phi(x)$ is the cdf of a Gaussian random variable with $m = 0$ and $\sigma = 1$:

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

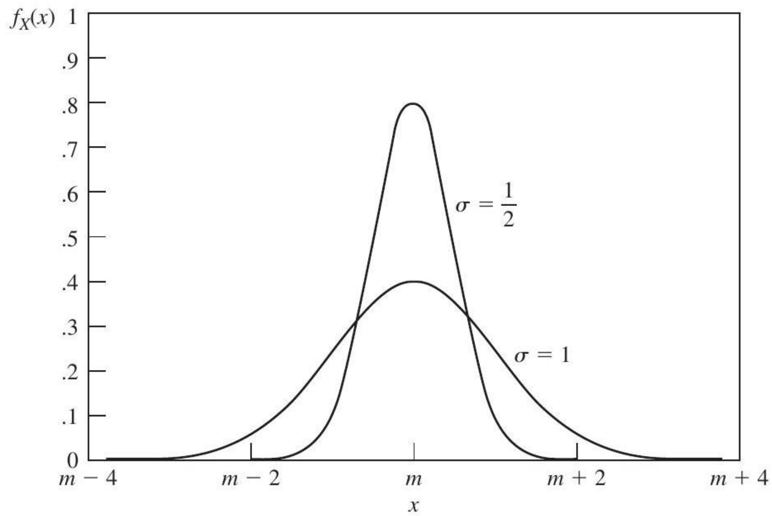

Figure 1: pdf of Gaussian random variable

Therefore any probability involving an arbitrary Gaussian random variable can be expressed in terms of $\Phi(x)$.

Example 1: Integrating the Gaussian pdf

Show that the Gaussian pdf integrates to one. Consider the square of the integral of the pdf:

$$\begin{aligned}\left[\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-x^2/2}dx\right]^2 &= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-x^2/2}dx \int_{-\infty}^{\infty} e^{-y^2/2}dy \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-\left(x^2+y^2\right)/2}dxdy\end{aligned}$$

Let $x = r\cos\theta$ and $y = r\sin\theta$. Changing from Cartesian to polar coordinates, we obtain:

$$\begin{aligned}\frac{1}{2\pi}\int_{0}^{\infty}\int_{0}^{2\pi} e^{-r^2/2} r dr d\theta &= \int_{0}^{\infty} r e^{-r^2/2} dr \\ &= \left[-e^{-r^2/2}\right]_0^{\infty} \\ &= 1\end{aligned}$$

Example 2: Mean and Variance of the Gaussian random variable

First we note that the pdf is symmetric about m, that is,

$$f_X(m-x) = f_X(m+x) \quad \text{for all } x$$

In addition the term $(m-t)$ has odd symmetry about m, so

$$0 = \int_{-\infty}^{+\infty} (m-t) f_X(t)dt = m - \int_{-\infty}^{+\infty} t f_X(t)dt$$

We then have that $E[X] = m$.

To find the variance of a Gaussian random variable, we first multiply the integral of the pdf of X by $\sqrt{2\pi}\sigma$ to obtain

$$\int_{-\infty}^{\infty} e^{-(x-m)^2/2\sigma^2} dx = \sqrt{2\pi}\sigma$$

We then differentiate both sides with respect to σ

$$\int_{-\infty}^{\infty} \left(\frac{(x-m)^2}{\sigma^3} \right) e^{-(x-m)^2/2\sigma^2} dx = \sqrt{2\pi}$$

By rearranging the above equation, we obtain

$$\text{VAR}[X] = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} (x-m)^2 e^{-(x-m)^2/2\sigma^2} dx = \sigma^2$$

It is sometimes customary to work with the Q-function, which is defined by

$$\begin{aligned} Q(x) &= 1 - \Phi(x) \\ &= \frac{1}{\sqrt{2\pi}} \int_x^{\infty} e^{-t^2/2} dt \end{aligned}$$

$Q(x)$ is simply the probability of the "tail" of the pdf. The symmetry of the pdf implies that

$$Q(0) = 1/2 \quad \text{and} \quad Q(-x) = 1 - Q(x)$$

- The integral in Eq. (1) does not have a closed-form expression.

- Traditionally the integrals have been evaluated by looking up tables that list $Q(x)$ or by using approximations that require numerical evaluation.
- The following expression has been found to give good accuracy for $Q(x)$ over the entire range $0 < x < \infty$:

$$Q(x) \simeq \left[\frac{1}{(1-a)x + a\sqrt{x^2+b}} \right] \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \tag{2}$$

  where $a = 1/\pi$ and $b = 2\pi$.
- Table 1 shows $Q(x)$ and the value given by the above approximation.
- In some problems, we are interested in finding the value of x for which $Q(x) = 10^{-k}$.
- Table 2 gives these values for $k = 1, \ldots, 10$.

The Gaussian random variable plays a very important role in communication systems, where transmission signals are corrupted by noise resulting from the thermal motion of electrons. It can be shown from physical principles that these voltages will have a Gaussian pdf.

x	$Q(x)$	Approximation	x	$Q(x)$	Approximation
0	5.00E − 01	5.00E − 01	2.7	3.47E − 03	3.46E − 03
0.1	4.60E − 01	4.58E − 01	2.8	2.56E − 03	2.55E − 03
0.2	4.21E − 01	4.17E − 01	2.9	1.87E − 03	1.86E − 03
0.3	3.82E − 01	3.78E − 01	3.0	1.35E − 03	1.35E − 03
0.4	3.45E − 01	3.41E − 01	3.1	9.68E − 04	9.66E − 04
0.5	3.09E − 01	3.05E − 01	3.2	6.87E − 04	6.86E − 04
0.6	2.74E − 01	2.71E − 01	3.3	4.83E − 04	4.83E − 04
0.7	2.42E − 01	2.39E − 01	3.4	3.37E − 04	3.36E − 04
0.8	2.12E − 01	2.09E − 01	3.5	2.33E − 04	2.32E − 04
0.9	1.84E − 01	1.82E − 01	3.6	1.59E − 04	1.59E − 04
1.0	1.59E − 01	1.57E − 01	3.7	1.08E − 04	1.08E − 04
1.1	1.36E − 01	1.34E − 01	3.8	7.24E − 05	7.23E − 05
1.2	1.15E − 01	1.14E − 01	3.9	4.81E − 05	4.81E − 05
1.3	9.68E − 02	9.60E − 02	4.0	3.17E − 05	3.16E − 05
1.4	8.08E − 02	8.01E − 02	4.5	3.40E − 06	3.40E − 06
1.5	6.68E − 02	6.63E − 02	5.0	2.87E − 07	2.87E − 07
1.6	5.48E − 02	5.44E − 02	5.5	1.90E − 08	1.90E − 08
1.7	4.46E − 02	4.43E − 02	6.0	9.87E − 10	9.86E − 10
1.8	3.59E − 02	3.57E − 02	6.5	4.02E − 11	4.02E − 11
1.9	2.87E − 02	2.86E − 02	7.0	1.28E − 12	1.28E − 12
2.0	2.28E − 02	2.26E − 02	7.5	3.19E − 14	3.19E − 14
2.1	1.79E − 02	1.78E − 02	8.0	6.22E − 16	6.22E − 16
2.2	1.39E − 02	1.39E − 02	8.5	9.48E − 18	9.48E − 18
2.3	1.07E − 02	1.07E − 02	9.0	1.13E − 19	1.13E − 19
2.4	8.20E − 03	8.17E − 03	9.5	1.05E − 21	1.05E − 21
2.5	6.21E − 03	6.19E − 03	10.0	7.62E − 24	7.62E − 24
2.6	4.66E − 03	4.65E − 03			

Table 1: Comparison of $Q(x)$ and approximation given by Eq. (2)

Example 3: Binary Communications

A communication system accepts a positive voltage V as input and outputs a voltage $Y = \alpha V + N$, where $\alpha = 10^{-2}$ and N is a Gaussian random variable with parameters $m = 0$ and $\sigma = 2$. Find the value of V that gives $P[Y < 0] = 10^{-6}$.

The probability $P[Y < 0]$ is written in terms of N as follows:

$$\begin{aligned} P[Y < 0] &= P[\alpha V + N < 0] \\ &= P[N < -\alpha V] = \Phi\left(\frac{-\alpha V}{\sigma}\right) = Q\left(\frac{\alpha V}{\sigma}\right) = 10^{-6} \end{aligned}$$

From Table 2 we see that the argument of the Q-function should be $\alpha V/\sigma = 4.753$. Thus $V = (4.753)\sigma/\alpha = 950.6$.

k	$x = Q^{-1}\left(10^{-k}\right)$
1	1.2815
2	2.3263
3	3.0902
4	3.7190
5	4.2649
6	4.7535
7	5.1993
8	5.6120
9	5.9978
10	6.3613

Table 2: $Q(x) = 10^{-k}$

2 Functions of a Random Variable

Let X be a random variable and let $g(x)$ be a real-valued function defined on the real line. Define $Y = g(X)$, that is, Y is determined by evaluating the function $g(x)$ at the value assumed by the random variable X. Then Y is also a random variable.

- The probabilities with which Y takes on various values depend on the function $g(x)$ as well as the cumulative distribution function of X.

- In this section we consider the problem of finding the cdf and pdf of Y.

Example 4

Let the function $h(x) = (x)^{+}$be defined as follows:

$$(x)^{+} = \begin{cases} 0 & \text{if } x < 0 \\ x & \text{if } x \geq 0 \end{cases}$$

For example, let X be the number of working devices in a group of N devices, and let Y be the number of working in excess of M, then $Y = (X - M)^{+}$. In machine learning, let X be an input to a rectifier linear unit (RELU) activation function then $Y = (X)^{+}$is the output.

Example 5

Let the function $q(x)$ be defined as shown in Fig. 4.8(a), where the set of points on the real line are mapped into the nearest representation point from the set $S_Y = \{-3.5d, -2.5d, -1.5d, -0.5d,\ 0.5d, 1.5d, 2.5d, 3.5d\}$. Thus, for example, all the points in the interval $(0, d)$ are mapped into the point $d/2$. The function $q(x)$ represents an eight-level uniform quantizer.

Example 6

Consider the linear function $c(x) = ax + b$, where a and b are constants. This function arises in many situations. For example, $c(x)$ could be the cost associated with the quantity x, with the constant a being the cost per unit of x, and b being a fixed cost component. In a signal processing context, $c(x) = ax$ could be the amplified version (if $a > 1$) or attenuated version (if $a < 1$) of the voltage x.

T*he probability of an event C involving Y is equal to the probability of the equivalent event B of values of X such that $g(X)$ is in C*:

$$P[Y \text{ in } C] = P[g(X) \text{ in } C] = P[X \text{ in } B].$$

Three types of equivalent events are useful in determining the cdf and pdf of $Y = g(X)$:

- The event $\{g(X) = y_k\}$ is used to determine the magnitude of the jump at a point y_k where the cdf of Y is known to have a *discontinuity*;
- The event $\{g(X) \leq y\}$ is used to *find the cd*f of Y directly; and
- The event $\{y < g(X) \leq y + h\}$ is useful in *determining the pdf* of Y.
- We will demonstrate the use of these three methods in a series of examples.

The next two examples demonstrate how the pmf is computed in cases where $Y = g(X)$ is discrete. In the first example, X is discrete. In the second example, X is continuous.

Example 7

Let X be the number of working devices in a group of N independent devices. Let p be the probability that a device is working. X has a binomial distribution with parameters N and p. Suppose that a system can use up to M devices at a time, and that when X exceeds $M, X - M$ is the number of extra devices available. Let Y be the number of extra devices, then

$$Y = (X - M)^+$$

Y takes on values from the set $S_Y = \{0, 1, \ldots, N - M\}$. Y will equal zero whenever X is less than or equal to M, and Y will equal $k > 0$ when X is equal to $M + k$. Therefore

$$P[Y = 0] = P[X \text{ in } \{0, 1, \ldots, M\}] = \sum_{j=0}^{M} p_j$$

and

$$P[Y = k] = P[X = M + k] = p_{M+k} \quad 0 < k \leq N - M$$

where p_j is the pmf of X.

Example 8

Let X be a sample voltage of a speech waveform, and suppose that X has a uniform distribution in the interval $[-4d, 4d]$. Let $Y = q(X)$, where the quantizer input-output characteristic is as shown in Fig. 4.10. Find the pmf for Y.

The event $\{Y = q\}$ for q in S_Y is equivalent to the event $\{X \text{ in } I_q\}$, where I_q is an interval of points mapped into the representation point q. The pmf of Y is therefore found by evaluating

$$P[Y = q] = \int_{I_q} f_X(t)dt$$

It is easy to see that the representation point has an interval of length d mapped into it. Thus the eight possible outputs are equiprobable, that is, $P[Y = q] = 1/8$ for q in S_Y.

In Example 8, each constant section of the function $q(X)$ produces a delta function in the pdf of Y. In general, if the function $g(X)$ is constant during certain intervals and if the pdf of X is nonzero in these intervals, then the pdf of Y will contain delta functions. Y will then be either discrete or of mixed type.

The cdf of Y is defined as the probability of the event $\{Y \le y\}$. *In principle, it can always be obtained by finding the probability of the equivalent event* $\{g(X) \le y\}$ as shown in the next examples.

Example 9: A Linear Function

Let the random variable Y be defined by

$$Y = aX + b$$

where a is a nonzero constant. Suppose that X has cdf $F_X(x)$, then find $F_Y(y)$.

The event $\{Y \le y\}$ occurs when $A = \{aX + b \le y\}$ occurs. If $a > 0$, then $A = \{X \le (y - b)/a\}$ (see Fig. 2), and thus

$$F_Y(y) = P\left[X \le \frac{y-b}{a}\right] = F_X\left(\frac{y-b}{a}\right) \quad a > 0.$$

On the other hand, if $a < 0$, then $A = \{X \ge (y - b)/a\}$, and

$$F_Y(y) = P\left[X \ge \frac{y-b}{a}\right] = 1 - F_X\left(\frac{y-b}{a}\right) \quad a < 0$$

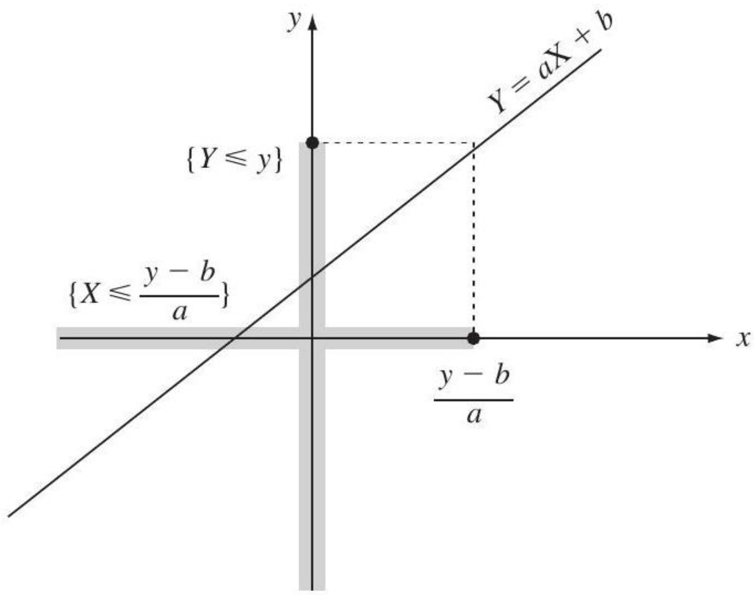

Figure 2: The equivalent event for $\{Y \leq y\}$ is the event $\{x \leq (y-b)/a\}$, if $a > 0$

We obtain the pdf of Y by differentiating with respect to y. To do this we need to use the chain rule for derivatives:

$$\frac{dF}{dy} = \frac{dF}{du}\frac{du}{dy}$$

where u is the argument of F. For $u = (y-b)/a$, and we then obtain

$$f_Y(y) = \frac{1}{a} f_X\left(\frac{y-b}{a}\right) \qquad a > 0$$

and

$$f_Y(y) = \frac{1}{-a} f_X\left(\frac{y-b}{a}\right) \qquad a < 0$$

The above two results can be written compactly as

$$f_Y(y) = \frac{1}{|a|} f_X\left(\frac{y-b}{a}\right) \tag{3}$$

Example 10: A Linear Function of a Gaussian Random Variable

Let X be a random variable with a Gaussian pdf with mean m and standard deviation σ:

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-m)^2/2\sigma^2} \qquad -\infty < x < \infty \tag{4}$$

Let $Y = aX + b$, then find the pdf of Y.

Substitution of Eq. (4) into Eq. (3) yields

$$f_Y(y) = \frac{1}{\sqrt{2\pi}|a\sigma|} e^{-(y-b-am)^2/2(a\sigma)^2}$$

Note that Y also has a Gaussian distribution with mean $b + am$ and standard deviation $|a|\sigma$. T*herefore a linear function of a Gaussian random variable is also a Gaussian random variable.*

Example 11: Square Function

Let the random variable Y be defined by

$$Y = X^2$$

where X is a continuous random variable. Find the cdf and pdf of Y.

The event $\{Y \leq y\}$ occurs when $\{X^2 \leq y\}$ or equivalently when $\{-\sqrt{y} \leq X \leq \sqrt{y}\}$ for y nonnegative; see Fig. 3. The event is null when y is negative. Thus

$$F_Y(y) = \begin{cases} 0 & y < 0 \\ F_X(\sqrt{y}) - F_X(-\sqrt{y}) & y > 0 \end{cases}$$

and differentiating with respect to y,

$$\begin{aligned} f_Y(y) &= \frac{f_X(\sqrt{y})}{2\sqrt{y}} - \frac{f_X(-\sqrt{y})}{-2\sqrt{y}} \qquad y > 0 \\ &= \frac{f_X(\sqrt{y})}{2\sqrt{y}} + \frac{f_X(-\sqrt{y})}{2\sqrt{y}} \end{aligned} \tag{5}$$

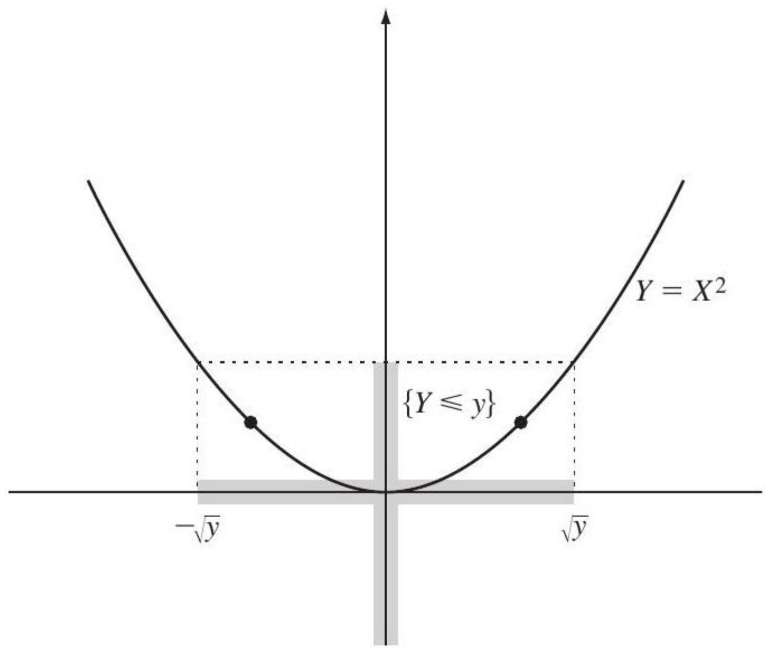

Figure 3: The equivalent event for $\{Y \leq y\}$ is the event $\{-\sqrt{y} \leq X \leq \sqrt{y}\}$, if $y \geq 0$

Example 12: A Chi-Square Random Variable

Let X be a Gaussian random variable with mean $m = 0$ and standard deviation $\sigma = 1$. X is then said to be a standard normal random variable. Let $Y = X^2$. Find the pdf of Y.

Substitution of Eq. (4) into Eq. (5) yields

$$f_Y(y) = \frac{e^{-y/2}}{\sqrt{2y\pi}} \quad y \geq 0$$

From the table of continuous random variables presented in Week 5, we see that $f_Y(y)$ is the pdf of a chi-square random variable with one degree of freedom.

The result in Example 11 suggests that if the equation $y_0 = g(x)$ has n solutions, $x_0, x_1, \ldots, x_n$, then $f_Y(y_0)$ will be equal to n terms of the type on the right-hand side of Eq. (5). We now show that this is generally true

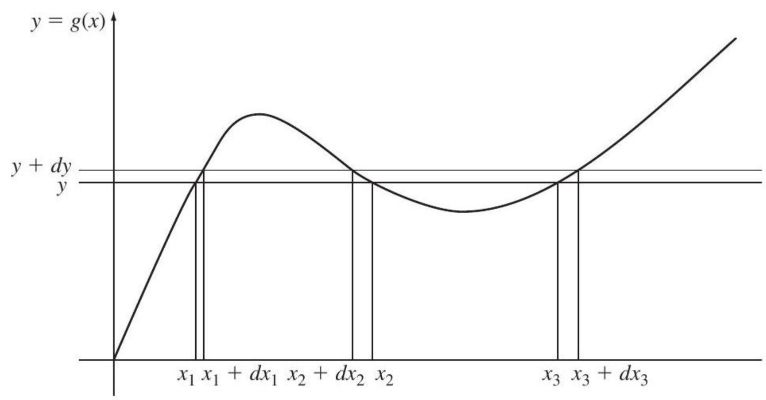

Figure 4: The equivalent event of $\{y < Y < y+dy\}$ is $\{x_1 < X < x_1 + dx_1\} \cup \{x_2 + dx_2 < X < x_2\} \cup \{x_3 < X < x_3 + dx_3\}$.

by using a method for directly obtaining the pdf of Y in terms of the pdf of X.

Consider a nonlinear function $Y = g(X)$ such as the one shown in Fig. 4. Consider the event $C_y = \{y < Y < y + dy\}$ and let B_y be its equivalent event. For y indicated in the figure, the equation $g(x) = y$ has three solutions x_1, x_2, and x_3, and the equivalent event B_y has a segment corresponding to each solution:

$$\begin{aligned} B_y = \{x_1 < X < x_1 + dx_1\} &\cup \{x_2 + dx_2 < X < x_2\} \\ &\cup \{x_3 < X < x_3 + dx_3\} \end{aligned}$$

The probability of the event C_y is approximately

$$P[C_y] = f_Y(y)|dy| \tag{6}$$

where $|dy|$ is the length of the interval $y < Y \leq y + dy$. Similarly, the probability of the event B_y is approximately

$$P[B_y] = f_X(x_1)|dx_1| + f_X(x_2)|dx_2| + f_X(x_3)|dx_3| \tag{7}$$

Since C_y and B_y are equivalent events, their probabilities must be equal. By equating Eqs. (6) and (7) we obtain

$$f_Y(y) = \sum_k \left. \frac{f_X(x)}{|dy/dx|} \right|_{x=x_k} \tag{8}$$

$$= \sum_k f_X(x) \left| \frac{dx}{dy} \right| \Bigg|_{x=x_k} \tag{9}$$

It is clear that if the equation $g(x) = y$ has n solutions, the expression for the pdf of Y at that point is given by Eqs. (8) and (9), and contains n terms.

Example 13

Let $Y = X^2$ as in Example 12. For $y \geq 0$, the equation $y = x^2$ has two solutions, $x_0 = \sqrt{y}$ and $x_1 = -\sqrt{y}$, so Eq. (8) has two terms. Since $dy/dx = 2x$, Eq. (8) yields

$$f_Y(y) = \frac{f_X(\sqrt{y})}{2\sqrt{y}} + \frac{f_X(-\sqrt{y})}{2\sqrt{y}}$$

This result is in agreement with Eq. (5). To use Eq. (9), we note that

$$\frac{dx}{dy} = \frac{d}{dy} \pm \sqrt{y} = \pm \frac{1}{2\sqrt{y}}$$

which when substituted into Eq. (9) then yields Eq. (5) again.

Example 14: Amplitude Samples of a Sinusoidal Waveform

Let $Y = \cos(X)$, where X is uniformly distributed in the interval $(0, 2\pi]$. Y can be viewed as the sample of a sinusoidal waveform at a random instant of time that is uniformly distributed over the period of the sinusoid. Find the pdf of Y.

It can be seen in Fig. 5 that for $-1 < y < 1$ the equation $y = \cos(x)$ has two solutions in the interval of interest, $x_0 = \cos^{-1}(y)$ and $x_1 = 2\pi - x_0$. Since (see an introductory calculus textbook)

$$\frac{dy}{dx}\bigg|_{x_0} = -\sin(x_0) = -\sin\left(\cos^{-1}(y)\right) = -\sqrt{1 - y^2}$$

and since $f_X(x) = 1/2\pi$ in the interval of interest, Eq. (8) yields

$$f_Y(y) = \frac{1}{2\pi\sqrt{1 - y^2}} + \frac{1}{2\pi\sqrt{1 - y^2}}$$

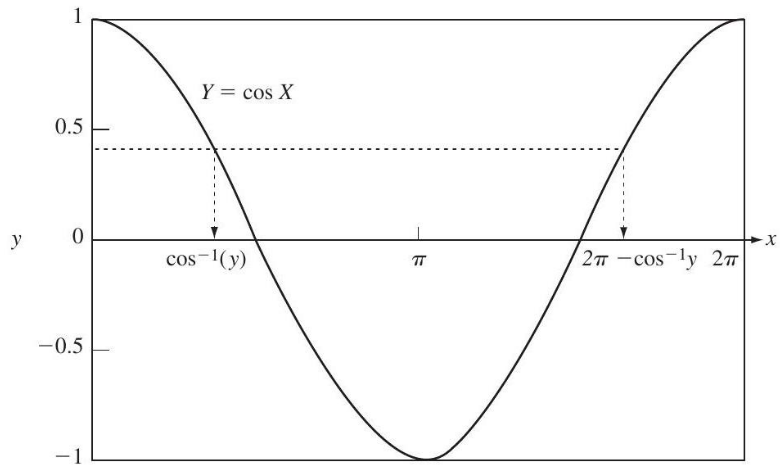

Figure 5: $y = \cos x$ has two roots in the interval $(0, 2\pi)$

$$= \frac{1}{\pi\sqrt{1-y^2}} \quad \text{for } -1 < y < 1$$

The cdf of Y is found by integrating the above:

$$F_Y(y) = \begin{cases} 0 & y < -1 \\ \frac{1}{2} + \frac{\sin^{-1} y}{\pi} & -1 \le y \le 1 \\ 1 & y > 1 \end{cases}$$

Y is said to have the **arcsine distribution**.

3 Transform Methods

We have seen that the pdf of a random variable X includes all the information about the behavior of the random variable. We will now introduce **transforms** that map a pdf $f_X(x)$ into another equivalent function $\mathscr{F}_X(\omega)$, that also includes all the information about the random variable X. We will present the **moment theorem** which shows that the moments of X can be obtained from $\mathscr{F}_X(\omega)$.

3.1 The Characteristic Function

The characteristic function of a random variable X is defined by

$$\Phi_X(\omega) = E\left[e^{j\omega X}\right]$$

$$= \int_{-\infty}^{\infty} f_X(x) e^{j\omega x} dx$$

where $j = \sqrt{-1}$ is the imaginary unit number. The two expressions on the right-hand side motivate two interpretations of the characteristic function.

- In the first expression, $\Phi_X(\omega)$ can be viewed as the expected value of a function of $X, e^{j\omega X}$, in which the parameter ω is left unspecified.
- In the second expression, $\Phi_X(\omega)$ is simply the Fourier transform of the pdf $f_X(x)$ (with a reversal in the sign of the exponent).
- Both of these interpretations prove useful in different contexts.

If we view $\Phi_X(\omega)$ as a **Fourier transform**, then we have from the Fourier transform inversion formula that the pdf of X is given by

$$f_X(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \Phi_X(\omega) e^{-j\omega x} d\omega \tag{10}$$

It then follows that every pdf and its characteristic function form a unique Fourier transform pair. The list of popular continuous random variables provides the characteristic function of some continuous random variables.

Example 15: Exponential Random Variable

The characteristic function for an exponentially distributed random variable with parameter λ is given by

$$\begin{aligned}\Phi_X(\omega) &= \int_0^{\infty} \lambda e^{-\lambda x} e^{j\omega x} dx = \int_0^{\infty} \lambda e^{-(\lambda - j\omega)x} dx \\ &= \frac{\lambda}{\lambda - j\omega}\end{aligned}$$

If X is a *discrete random variable*, substitution of Eq. 15 (in Week 4 Notes) into the definition of $\Phi_X(\omega)$ gives

$$\Phi_X(\omega) = \sum_k p_X(x_k) e^{j\omega x_k} \quad \text{discrete random variables.}$$

Most of the time we deal with discrete random variables that are integer-valued. The characteristic function then becomes

$$\Phi_X(\omega) = \sum_{k=-\infty}^{\infty} p_X(k) e^{j\omega k} \quad \text{integer-valued random variables.} \tag{11}$$

Equation (11) is the **Fourier transform of the sequence** $p_X(k)$. Note that the Fourier transform in Eq. (11) is a periodic function of ω with period 2π, since $e^{j(\omega+2\pi)k} = e^{j\omega k} e^{jk2\pi}$ and $e^{jk2\pi} = 1$. Therefore the characteristic function of integer-valued random variables is a periodic function of ω. The following inversion formula allows us to recover the probabilities $p_X(k)$ from $\Phi_X(\omega)$:

$$p_X(k) = \frac{1}{2\pi} \int_0^{2\pi} \Phi_X(\omega) e^{-j\omega k} d\omega \quad k = 0, \pm 1, \pm 2, \ldots \tag{12}$$

Indeed, a comparison of the above two equations shows that the $p_X(k)$ are simply the coefficients of the Fourier series of the periodic function $\Phi_X(\omega)$.

Example 16: Geometric Random Variable

The characteristic function for a geometric random variable is given by

$$\begin{aligned}\Phi_X(\omega) &= \sum_{k=0}^{\infty} p q^k e^{j\omega k} = p \sum_{k=0}^{\infty} \left(q e^{j\omega}\right)^k \\ &= \frac{p}{1 - q e^{j\omega}}.\end{aligned}$$

Since $f_X(x)$ and $\Phi_X(\omega)$ form a transform pair, we would expect to be able to obtain the moments of X from $\Phi_X(\omega)$. The **moment theorem** states that the moments of X are given by

$$E\left[X^n\right] = \frac{1}{j^n} \frac{d^n}{d\omega^n} \Phi_X(\omega) \bigg|_{\omega=0} \tag{13}$$

To show this, first expand $e^{j\omega x}$ in a power series in the definition of $\Phi_X(\omega)$:

$$\Phi_X(\omega) = \int_{-\infty}^{\infty} f_X(x) \left\{ 1 + j\omega X + \frac{(j\omega X)^2}{2!} + \cdots \right\} dx$$

Assuming that all the moments of X are finite and that the series can be integrated term by term, we obtain

$$\Phi_X(\omega) = 1 + j\omega E[X] + \frac{(j\omega)^2 E\left[X^2\right]}{2!} + \cdots + \frac{(j\omega)^n E\left[X^n\right]}{n!} + \cdots$$

If we differentiate the above expression once and evaluate the result at $\omega = 0$ we obtain

$$\left.\frac{d}{d\omega}\Phi_X(\omega)\right|_{\omega=0} = jE[X]$$

If we differentiate n times and evaluate at $\omega = 0$, we finally obtain

$$\left.\frac{d^n}{d\omega^n}\Phi_X(\omega)\right|_{\omega=0} = j^n E\left[X^n\right]$$

which yields Eq. (13).

Note that when the above power series converges, the characteristic function and hence the pdf by Eq. (11) are completely determined by the moments of X.

Example 17

To find the mean of an exponentially distributed random variable, we differentiate $\Phi_X(\omega) = \lambda(\lambda - j\omega)^{-1}$ once, and obtain

$$\Phi'_X(\omega) = \frac{\lambda j}{(\lambda - j\omega)^2}$$

The moment theorem then implies that $E[X] = \Phi'_X(0)/j = 1/\lambda$. If we take two derivatives, we obtain

$$\Phi''_X(\omega) = \frac{-2\lambda}{(\lambda - j\omega)^3}$$

so the second moment is then $E\left[X^2\right] = \Phi''_X(0)/j^2 = 2/\lambda^2$. The variance of X is then given by

$$\mathrm{VAR}[X] = E\left[X^2\right] - E[X]^2 = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}$$

3.2 The Probability Generating Function

In problems where random variables are nonnegative, it is usually more convenient to use the z-transform or the Laplace transform. The **probability generating function** $G_N(z)$ of a nonnegative integer-valued random variable N is defined by

$$\begin{aligned} G_N(z) &= E\left[z^N\right] \\ &= \sum_{k=0}^{\infty} p_N(k) z^k \end{aligned}$$

- The first expression is the expected value of the function of N, z^N.
- The second expression is the z-transform of the pmf (with a sign change in the exponent).
- The list of continuous random variables from Week 5 shows the probability generating function for some discrete random variables.
- Note that the characteristic function of N is given by $\Phi_N(\omega) = G_N\left(e^{j\omega}\right)$.

Using a derivation similar to that used in the moment theorem, it is easy to show that *the pmf of N is given* by

$$p_N(k) = \frac{1}{k!} \frac{d^k}{dz^k} G_N(z) \Big|_{z=0}$$

This is why $G_N(z)$ is called the probability generating function.

By taking the first two derivatives of $G_N(z)$ and evaluating the result at $z = 1$, it is possible to find the first two moments of X:

$$\frac{d}{dz} G_N(z) \Big|_{z=1} = \sum_{k=0}^{\infty} p_N(k) k z^{k-1} \Big|_{z=1} = \sum_{k=0}^{\infty} k p_N(k) = E[N]$$

and

$$\begin{aligned} \frac{d^2}{dz^2} G_N(z) \Big|_{z=1} &= \sum_{k=0}^{\infty} p_N(k) k(k-1) z^{k-2} \Big|_{z=1} \\ &= \sum_{k=0}^{\infty} k(k-1) p_N(k) = E[N(N-1)] = E\left[N^2\right] - E[N] \end{aligned}$$

Thus the **mean and variance of** X are given by

$$E[N] = G_N'(1)$$

and

$$\mathrm{VAR}[N] = G_N''(1) + G_N'(1) - \left(G_N'(1)\right)^2$$

Example 18: Poisson Random Variable
The probability generating function for the Poisson random variable with parameter α is given by

$$\begin{aligned} G_N(z) &= \sum_{k=0}^{\infty} \frac{\alpha^k}{k!} e^{-\alpha} z^k = e^{-\alpha} \sum_{k=0}^{\infty} \frac{(\alpha z)^k}{k!} \\ &= e^{-\alpha} e^{\alpha z} = e^{\alpha(z-1)} \end{aligned}$$

The first two derivatives of $G_N(z)$ are given by

$$G_N'(z) = \alpha e^{\alpha(z-1)}$$

and

$$G_N''(z) = \alpha^2 e^{\alpha(z-1)}$$

Therefore the mean and variance of the Poisson are

$$\begin{aligned} E[N] &= \alpha \\ \mathrm{VAR}[N] &= \alpha^2 + \alpha - \alpha^2 = \alpha \end{aligned}$$

3.3 The Laplace Transform of the pdf

In waiting line problems one deals with service times, waiting times, and delays. All of these are nonnegative continuous random variables. It is therefore customary to work with the *Laplace transform of the pdf*,

$$X^*(s) = \int_0^{\infty} f_X(x) e^{-sx} dx = E\left[e^{-sX}\right]$$

Note that $X^*(s)$ can be interpreted as a Laplace transform of the pdf or as an expected value of a function of X, e^{-sX}.

The **moment theorem** also holds for $X^*(s)$:

$$E\left[X^n\right] = (-1)^n \frac{d^n}{ds^n} X^*(s)\bigg|_{s=0}$$

Example 19: Gamma Random Variable

The Laplace transform of the gamma pdf is given by

$$\begin{aligned} X^*(s) &= \int_0^\infty \frac{\lambda^\alpha x^{\alpha-1} e^{-\lambda x} e^{-sx}}{\Gamma(\alpha)} dx = \frac{\lambda^\alpha}{\Gamma(\alpha)} \int_0^\infty x^{\alpha-1} e^{-(\lambda+s)x} dx \\ &= \frac{\lambda^\alpha}{\Gamma(\alpha)} \frac{1}{(\lambda+s)^\alpha} \int_0^\infty y^{\alpha-1} e^{-y} dy = \frac{\lambda^\alpha}{(\lambda+s)^\alpha}, \end{aligned}$$

where we used the change of variable $y = (\lambda + s)x$. We can then obtain the first two moments of X as follows:

$$E[X] = -\frac{d}{ds} \frac{\lambda^\alpha}{(\lambda+s)^\alpha}\bigg|_{s=0} = \frac{\alpha\lambda^\alpha}{(\lambda+s)^{\alpha+1}}\bigg|_{s=0} = \frac{\alpha}{\lambda}$$

and

$$E\left[X^2\right] = \frac{d^2}{ds^2} \frac{\lambda^\alpha}{(\lambda+s)^\alpha}\bigg|_{s=0} = \frac{\alpha(\alpha+1)\lambda^\alpha}{(\lambda+s)^{\alpha+2}}\bigg|_{s=0} = \frac{\alpha(\alpha+1)}{\lambda^2}.$$

Thus the variance of X is

$$\text{VAR}(X) = E\left[X^2\right] - E[X]^2 = \frac{\alpha}{\lambda^2}$$