**ECS 232**

## 3. การแจกแจงความน่าจะเป็นและค่าที่คาดว่าควรจะเป็นของตัวแปรสุ่ม

## (Probability Distribution and Expected value of Random Variables)

ในทางสถิติโดยปกติจะแบ่งชนิดของตัวแปรสุ่มได้เป็น 2 ชนิด ได้แก่ ตัวแปรสุ่มไม่ต่อเนื่องกัน (discrete random variable): ตัวแปร x จะเป็นตัวแปรชนิดสุ่มชนิดไม่ต่อเนื่องกัน ถ้า range ของฟังก์ชันของตัวแปรสุ่มเป็นเซ็ตของเลขจำนวนนับทั้งหมด และตัวแปรสุ่มชนิดต่อเนื่อง (continuous random variable): ตัวแปรสุ่ม x จะเป็นตัวแปรชนิดสุ่มชนิดต่อเนื่องกัน ถ้า range ของฟังก์ชันของตัวแปรสุ่มเป็นเซ็ตของเลขจำนวนจริงใด ๆทั้งหมด

**Definition 1** *A random variable X is said to be* **discrete** *if it can assume only a finite or accountably infinite number of distinct values.*

**Example 1**

In the tossing of three fair coins, let the random variable $X$ be defined as $X$ = number of tails. Then $X$ canassume values 0, 1, 2, and 3. We can associate these values with probabilities in the following way:

$P(X = 0) = P(\{H, H, H\}) = 1/8$

$P(X = 1) = P(\{H, H, T\} \cup \{H, T, H\} \cup \{T, H, H\}) = 3/8$

$P(X = 2) = P(\{T, T, H\} \cup \{T, H, T\} \cup \{H, T, T\}) = 3/8$

$P(X = 3) = P(\{T, T, T\}) = 1/8.$

เราสามารถเขียนตารางการแจกแจงความถี่ของความน่าจะเป็นได้ดังนี้

| x | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| *p(x)* | 1/8 | 3/8 | 3/8 | 1/8 |

Let $X$ be a discrete random variable assuming values $x1, x2, x3, \ldots$ เราสามารถเขียนได้ว่า

**Definition 2** *The* **discrete probability mass function (pmf )** *of a discrete random variable X is the function*

$$p(xi) = P(X = xi),\ i = 1, 2, 3, \ldots$$

*A probability mass function (pmf) is more simply called a probability function (pf).*

*The* **cumulative distribution function (cdf )** *F of the random variable X is defined by*

$$F(x) = P(X \le x)$$

$$= \sum_{\text{all } y \le x} p(y)$$

*For* $-\infty < x < \infty$.

*A cumulative distribution function is also called a* **probability distribution function** *or simply the* **distribution function***.*

The probability function *p(x)* is nonnegative. In addition, because *X* must take on one of the values in {*x*1*, x*2*, x*3 . . .}, we have $\sum_{i=1}^{\infty} p(xi) = 1.$

**Example 2**

ในการโยนเหรียญเที่ยงตรง 1เหรียญสองครั้งมี sample space *S* = {*HH, HT, TH, TT* }ให้ *X* เป็นจำนวนการออกหัว

**(a)** Find the probability function for *X*.

**(b)** Find the cumulative distribution function of *X*.

**Definition 3** *Let X be a random variable. Suppose that there exists a nonnegative real-valued function: f* : *R* → [0,∞*) such that for any interval* [*a, b*]*,*

$$P(X \in [a, b]) = \int_a^b f(t)dt.$$

*Then X is called a* **continuous random variable***. The function f is called the* **probability density function (pdf)** *of X.*

*The* **cumulative distribution function (cdf)** *is given by*

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)dt.$$

For a given function $f$ to be a pdf, it needs to satisfy the following two conditions: $f(x) \geq 0$ for all values of $x$, and $\int_{-\infty}^{\infty} f(x)dx = 1$.

Also, if $f$ is continuous, then $\frac{dF(x)}{dx} = f(x)$, where $F(x)$ is the cdf. This follows from the fundamental theorem of calculus. If $f$ is the pdf of a random variable $X$, then

$$P(a \leq X \leq b) = \int_a^b f(x)dx.$$

Figure 2.5 represents $P(a \leq X \leq b)$.

As a result, for any real number $a$, $P(X = a) = 0$. Also,

$$P(a \leq X \leq b) = P(a < X \leq b) = P(a \leq X < b) = P(a < X < b).$$

If we have cdf $F(x)$, then we have

$$P(a \leq X \leq b) = F(b) - F(a).$$

**SOME PROPERTIES OF DISTRIBUTION FUNCTION**

1. $0 \leq F(x) \leq 1$.
2. $\lim_{x \to -\infty} F(x) = 0$, and $\lim_{x \to \infty} F(x) = 1$.
3. $F$ is a nondecreasing function, and right continuous.

**Example 3**

Let the function

$$f(x) = \begin{cases} \lambda x e^{-x}, & x > 0 \\ 0, & \text{otherwise.} \end{cases}$$

(a) For what value of $\lambda$ is $f$ a pdf?

(b) Find $F(x)$.

***Solution***

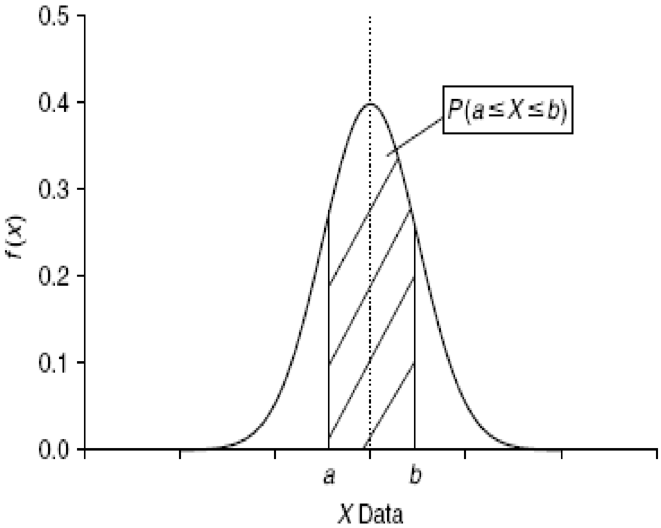
(a) *First note that* $f(x) \geq 0$. *Now, for* $f(x)$ *to be a pdf, we need* $\int_{-\infty}^{\infty} f(x)dx = 1$. *Because* $f(x) = 0$ *for* $x \leq 0$,

*Therefore* $\lambda = 1$. *See Figure 2.6.*

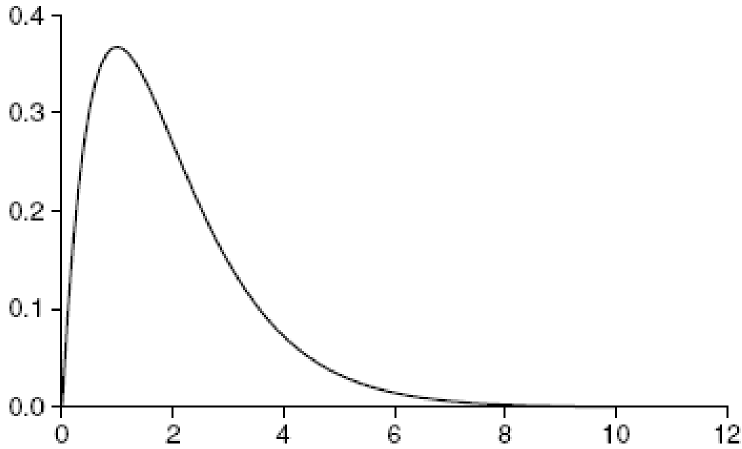
$$1 = \int_{-\infty}^{\infty} f(x)dx = \int_{0}^{\infty} \lambda x e^{-x} dx$$

$$= \lambda \int_{0}^{\infty} x e^{-x} dx = \lambda \left[ -x e^{-x} \Big|_{0}^{\infty} + \int_{0}^{\infty} e^{-x} dx \right] \text{(using integration by parts)}$$

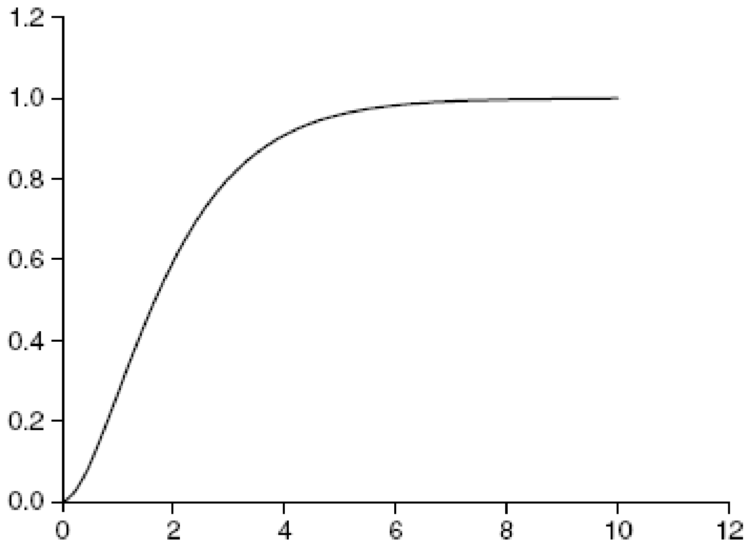
$$= \lambda \left[ 0 - e^{-x} \Big|_{0}^{\infty} \right] = \lambda.$$




■ **FIGURE 2.6** Probability as an area under a curve.




■ FIGURE 2.7 Graph of $f(x) = xe^{-x}$.




■ FIGURE 2.8 Graph of $F(x)$, $x \geq 0$.

**(b)** *The cumulative distribution function is*

$$F(x) = \int_{-\infty}^{x} f(t)dt = \begin{cases} 0, & x < 0 \\ \int_{0}^{x} te^{-t}dt = 1 - (x+1)\,e^{-x}, & x \geq 0. \end{cases}$$

*Figure 2.8 represents the cumulative distribution.*

**Example 4**

สมมติให้ห้างค้าปลีกแห่งหนึ่งมีพื้นที่สำหรับวางเครื่องดื่มน้ำผลไม้ 150 ขวดและมาส่งเป็นประจำทุกวัน ยอดขายรายสัปดาห์เพิ่มขึ้นอย่างค่อยเป็นค่อยไป จาก 100 ขวด เป็น 100-150 ขวด สมมติให้ Y เป็นความต้องการรายสัปดาห์ของน้ำผลไม้ (หน่วย 100 ขวด)

และให้ pdf. ของ Y สามารถประมาณค่าได้จาก

$$f(y) = \begin{cases} y, & 0 \le y \le 1 \\ 1, & 1 < y \le 1.5 \\ 0, & \text{elsewhere.} \end{cases}$$

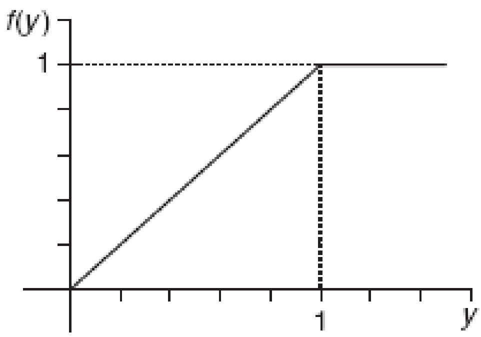
(a) Find $F(y)$,
(b) Find $P(0 \le Y \le 0.5)$,
(c) Find $P(0.5 \le Y \le 1.2)$.

***Solution***

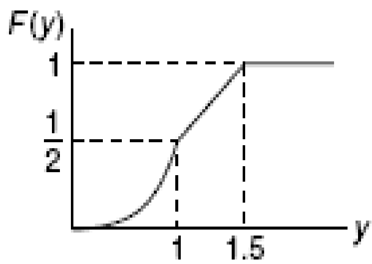
(a) *The graph of the density function* $f(y)$ *is shown in Figure 2.9.*
*From the definition of cdf, we have (Figure 2.10)*

$$F(y) = \int_{-\infty}^{y} f(t)dt = \begin{cases} 0, & y < 0 \\ \int_0^y t\,dt, & 0 \le y < 1 \\ \int_0^1 t\,dt + \int_1^y dt, & 1 \le y < 1.5 \\ \int_0^1 t\,dt + \int_1^{1.5} dt, & y \ge 1.5 \end{cases}$$

$$= \begin{cases} 0, & y < 0 \\ y^2/2, & 0 \le y < 1 \\ y - 1/2, & 1 \le y < 1.5 \\ 1, & y \ge 1.5. \end{cases}$$



■ **FIGURE 2.9** Graph of $f(y)$.



■ **FIGURE 2.10** Graph of cdf.

**(b)** *The probability,*

$$P(0 \leq Y \leq 0.5) = F(0.5) - F(0)$$
$$= (0.5)^2/2 = 1/8 = 0.125.$$

**(c)**

$$P(0.5 \leq Y \leq 1.2) = F(1.2) - F(0.5)$$
$$= (1.2 - 1/2) - 0.125 = 0.575.$$

**Moments and Moment- Generating Functions**

หนึ่งในเรื่องที่ใช้ประโยชน์มากที่สุดในทฤษฎีของความน่าจะเป็นคือเรื่องค่าที่คาดว่าควรจะเป็นของตัวแปรสุ่ม (Expected Value of a Random Variable) โดยทั่วไปค่าซึ่งถือว่าเป็นตัวแทนที่ดีที่สุดของการแจกแจงหนึ่ง คือค่าศูนย์กลางของการแจกแจงนั้น หรือค่าเฉลี่ย ค่านี้วัดได้ด้วยค่าที่คาดว่าควรจะเป็น (Expected Value)

**Definition 4** ถ้า x เป็นตัวแปรสุ่มชนิดไม่ต่อเนื่อง และมีการแจกแจงความน่าจะเป็น pf p(x) แล้วค่าที่คาดว่าจะเป็นของตัวแปร สามารถประมาณค่าได้จาก

$$\mu = E(X) = \sum_{\text{all } x} x p(x), \text{ provided } \sum_{\text{all } x} |x|\, p(x) < \infty.$$

**Definition 5** ถ้า x เป็นตัวแปรสุ่มชนิดต่อเนื่อง และมีการแจกแจงความน่าจะเป็น pdf p(x) แล้วค่าที่คาดว่าจะเป็นของตัวแปร สามารถประมาณค่าได้จาก

$$\mu = E(X) = \int_{-\infty}^{\infty} x f(x) dx, \text{ provided } \int_{-\infty}^{\infty} |x|\, f(x) dx < \infty.$$

The expected value of $X$ is also called the *expectation* or mathematical expectation of $X$. We denote the expected value of $X$ by $\mu$.

**Example 5**

Let

$$X = \begin{cases} 1, & \text{with a probability } 1/2 \\ 0, & \text{with a probability } 1/2. \end{cases}$$

Then $E(X) = 1(1/2) + 0(1/2) = 1/2.$

**Example 6**

Let $X$ be a discrete random variable whose probability density function is given in the following table:

| $x$ | $-1$ | $0$ | $1$ | $2$ | $3$ | $4$ | $5$ |
|---|---|---|---|---|---|---|---|
| $p(x)$ | $\frac{1}{7}$ | $\frac{1}{7}$ | $\frac{1}{14}$ | $\frac{2}{7}$ | $\frac{1}{14}$ | $\frac{1}{7}$ | $\frac{1}{7}$ |

Find $E(X)$.

**Example 7**

Let $X \geq 0$ be an integer-valued random variable such that $P(X = n) = p_n$. Show that $EX = \sum_{n=1}^{\infty} P(X \geq n)$.

## Solution

*Using the definition of expectation, and the fact that* $(0)(p_0) = 0$, *we have*

$$\begin{aligned} EX &= \sum_{n=1}^{\infty} np_n = 1p_1 + 2p_2 + 3p_3 + \cdots \\ &= p_1 + p_2 + p_3 + \cdots \\ &\quad + p_2 + p_3 + p_4 + \cdots \\ &\quad + p_3 + p_4 + \cdots \\ &= P(X \geq 1) + P(X \geq 2) + \cdots \\ &= \sum_{n=1}^{\infty} P(X \geq n). \end{aligned}$$

**Definition 6** ความแปรปรวนของตัวแปรสุ่ม x ถูกกำหนดสามารถหาได้จาก

$$\sigma^2 = Var(X) = E(X - \mu)^2$$

*The square root of variance, denoted by* $\sigma$, *is called the* **standard deviation**.

The variance is a measure of spread or variability of values of a random variable around the mean.

**EXPECTATION OF FUNCTION OF A RANDOM VARIABLE**

Theorem 2.6.1 *Let* $g(X)$ *be a function of* $X$, *then the expected value of* $g(X)$ *is*

$$E[g(X)] = \begin{cases} \sum_x g(x)p(x), & \text{if X is discrete} \\ \int_{-\infty}^{\infty} g(x)f(x)dx, & \text{if X is continuous} \end{cases}$$

*provided the sum or the integral exists.*

**SOME PROPERTIES OF EXPECTED VALUE AND VARIANCE**

Theorem 2.6.2 *Let $c$ be a constant and let $g(X)$, $g_1(X), \ldots, g_n(X)$ be functions of a random variable $X$ such that $E(g(X))$ and $E(g_i(X))$ for $i = 1, 2, \ldots, n$ exist. Then the following results hold:*

(a) $E(c) = c$.

(b) $E[cg(X)] = cE[g(X)]$.

(c) $E[\sum_i g_i(X)] = \sum_i E[g_i(X)]$.

(d) $Var(aX + b) = a^2 Var(X)$. *In particular,* $Var(aX) = a^2 Var(X)$.

(e) $Var(X) = E(X^2) - \mu^2$.

*Proof.* Proof of (a) through (d) will be given as an exercise. We will prove (e).

$$\begin{aligned} Var(X) &= E(X - \mu)^2 \\ &= E\left(X^2 - 2X\mu + \mu^2\right) \\ &= E\left(X^2\right) - 2\mu E(X) + \mu^2 \\ &= E\left(X^2\right) - 2\mu^2 + \mu^2 \\ &= E\left(X^2\right) - \mu^2. \end{aligned}$$

**Example 8**

A discrete random variable $X$ is said to be *uniformly distributed* over the numbers $1, 2, 3, \ldots, n$, if

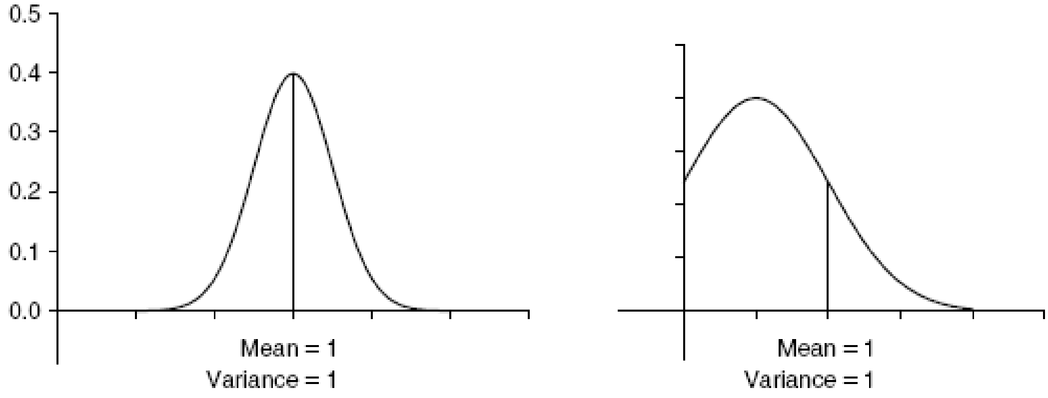
$$P(X = i) = \frac{1}{n}, \quad i = 1, 2, \ldots, n.$$

Find $EX$ and $Var(X)$.

**Note**

Even though the mean $\mu$ and the standard deviation $\sigma$ are significant descriptive measures that locate the center and describe the spread or dispersion of probability density function $f(x)$, they do not provide a unique characterization of the distribution. Two distributions may have the same mean and variance and yet could be very different, as in Figure 2.12.

To better approximate the probability distribution of a random variable, we may need higher moments.




■ **FIGURE 2.12** Same mean and variance.

**Definition 2.6.4** *The* **kth moment about the origin** *of a random variable X is defined as* $EX^k$ *and denoted by* $\mu'_k$*, whenever it exists. The* **kth moment about its mean** *(also called* **central kth moment***) of a random variable X is defined as* $E\left[(X-\mu)^k\right]$ *and denoted by* $\mu_k, k = 2, 3, 4, \ldots,$ *whenever it exists.*

In particular, we have $E(X) = \mu'_1 = \mu$, and $\sigma^2 = \mu_2$. We have seen earlier that the second moment about mean (variance, $\sigma^2$) is used as a measure of dispersion about the mean.

**Definition 2.6.5** *The standardized third moment about mean*

$$\alpha_3 = \frac{E(X-\mu)^3}{\sigma^3} = \frac{\mu_3}{\mu_2^{3/2}}$$

*is called the* **skewness** *of the distribution of X. The standardized fourth moment about mean*

$$\alpha_4 = \frac{E(X-\mu)^4}{\sigma^4}$$

*is called the* **kurtosis** *of the distribution.*

Let $X$ be a random variable with pf

$$p(x) = \binom{n}{x} p^x (1-p)^{n-x}, \quad x = 0, 1, 2, \ldots, n.$$

(This random variable is called a binomial random variable, and the pf is called a binomial distribution.) Show that $M_X(t) = \left[(1-p) + pe^t\right]^n$, for all real values of $t$. Also obtain mean and variance of the random variable $X$.

**Definition 2.6.6** *For a random variable X, suppose that there is a positive number h such that for* $-h < t < h$ *the mathematical expectation* $E\left(e^{tX}\right)$ *exists. The* **moment-generating function (mgf)** *of the random variable X is defined by*

$$M_X(t) = E\left(e^{tX}\right) = \begin{cases} \sum e^{tx} p(x), & \text{if discrete} \\ \int e^{tx} f(x)dx, & \text{if continuous} \end{cases}.$$

An advantage of the moment generating function is its ability to give the moments. Recall that the Maclaurin series of the function $e^{tx}$ is

$$e^{tx} = 1 + tx + \frac{(tx)^2}{2!} + \frac{(tx)^3}{3!} + \cdots + \frac{(tx)^n}{n!} + \cdots.$$

By using the fact that the expected value of the sum equals the sum of the expected values, the moment-generating function can be written as

$$M_X(t) = E\left[e^{tX}\right] = E\left[1 + tX + \frac{(tX)^2}{2!} + \frac{(tX)^3}{3!} + \cdots + \frac{(tX)^n}{n!} + \cdots\right]$$

$$= 1 + tE[X] + \frac{t^2}{2!}E\left[X^2\right] + \frac{t^3}{3!}E\left[X^3\right] + \cdots + \frac{t^n}{n!}E\left[X^n\right] + \cdots$$

Taking the derivative of $M_X(t)$ with respect to $t$, we obtain

$$\frac{dM_X(t)}{dt} = M'_X(t) = E[X] + tE[X] + \frac{t^2}{2!}E\left[X^2\right]$$

$$+ \frac{t^3}{3!}E\left[X^3\right] + \cdots + \frac{t^{(n-1)}}{(n-1)!}E\left[X^n\right] + \cdots$$

Evaluating this derivative at $t = 0$, all terms except $E[X]$ become zero. We have

$$M'_X(0) = E[X].$$

Similarly, taking the second derivative of $M_X(t)$, we obtain

$$M''_X(0) = E\left[X^2\right].$$

Continuing in this manner, from the $n$th derivative $M_X^{(n)}(t)$ with respect to $t$, we obtain all the moments to be

$$M_X^{(n)}(0) = E\left[X^n\right], \quad n = 1, 2, 3, \ldots.$$

We summarize these calculations in the following theorem.

Theorem 2.6.3 *If $M_X(t)$ exists, then for any positive integer $k$,*

$$\left.\frac{d^k M_X(t)}{dt^k}\right|_{t=0} = M_X^{(k)}(0) = \mu_k'.$$

The usefulness of the foregoing theorem lies in the fact that, if the mgf can be found, the often difficult process of integration or summation involved in calculating different moments can be replaced by the much easier process of differentiation. The following examples illustrate this fact.

**Example 9**

Let $X$ be a random variable with pf

$$p(x) = \binom{n}{x} p^x (1-p)^{n-x}, \quad x = 0, 1, 2, \ldots, n.$$

(This random variable is called a binomial random variable, and the pf is called a binomial distribution.) Show that $M_X(t) = \left[(1-p) + pe^t\right]^n$, for all real values of $t$. Also obtain mean and variance of the random variable $X$.

***Solution***

*The moment-generating function of $X$ is*

$$M_X(t) = E\left(e^{tX}\right) = \sum_{x=0}^{n} e^{tx} \binom{n}{x} p^x (1-p)^{n-x}$$

$$= \sum_{x=0}^{n} \binom{n}{x} (pe^t)^x (1-p)^{n-x}.$$

*Using the binomial formula, we have*

$$M_X(t) = \left[pe^t + (1-p)\right]^n, \quad -\infty < t < \infty.$$

*The first two derivatives of $M_X(t)$ are*

$$M_X'(t) = n\left[(1-p) + pe^t\right]^{(n-1)} \left(pe^t\right)$$

*and*

$$M_X''(t) = n(n-1)\left[(1-p) + pe^t\right]^{(n-2)} \left(pe^t\right)^2 + n\left[(1-p) + pe^t\right]^{(n-1)} \left(pe^t\right).$$

*Thus,*

$$\mu = E(X) = M_X'(0) = np$$

*and*

$$\sigma^2 = E\left(X^2\right) - \mu^2 = M''(0) - (np)^2$$
$$= n(n-1)\,p^2 + np - (np)^2 = np(1-p)\,.$$

**Example 10**

Let $X$ be a random variable with pdf given by

$$f(x) = \begin{cases} \frac{1}{\beta} e^{-x/\beta}, & x > 0 \\ 0, & \text{otherwise.} \end{cases}$$

Find mgf $M_X(t)$.

***Solution***

*By definition of mgf,*

$$M_x(t) = \int_{\infty}^{-\infty} e^{tx} f(x) dx$$
$$= \int_0^{\infty} e^{tx} \frac{1}{\beta} e^{-x/\beta} dx$$
$$= \frac{1}{\beta} \int_0^{\infty} e^{-\left(\frac{1}{\beta}-t\right)x} dx, \left(t < \frac{1}{\beta}\right)$$
$$= \frac{1}{\beta}\left[-\frac{1}{((1/\beta)-t)} e^{-\left(\frac{1}{\beta}-t\right)x}\Big|_{x=0}^{\infty}\right]$$
$$= \frac{1}{\beta}\frac{\beta}{1-\beta t} = \frac{1}{1-\beta t}, \quad t < \frac{1}{\beta}.$$

**PROPERTIES OF THE MOMENT-GENERATING FUNCTION**

1. The moment-generating function of $X$ is unique in the sense that, if two random variables $X$ and $Y$ have the same mgf ($M_X(t) = M_Y(t)$, for $t$ in an interval containing 0), then $X$ and $Y$ have the same distribution.
2. If $X$ and $Y$ are independent, then

$$M_{X+Y}(t) = M_X(t)M_Y(t).$$

   That is, the mgf of the sum of two independent random variables is the product of the mgfs of the individual random variables. The result can be extended to '$n$' random variables.
3. Let $Y = aX + b$. Then

$$M_Y(t) = e^{bt} M_X(at).$$

## แบบฝึกหัดที่ 3

1.

The probability function of a random variable $Y$ is given by $p(i) = \frac{c\lambda^i}{i!}$, $i = 0, 1, 2, \ldots$, where $\lambda$ is a known positive value and $c$ is a constant.

(a) Find $c$.
(b) Find $P(Y = 0)$.
(c) Find $P(Y > 2)$.

2.

The probability density function of a random variable $X$ is given by

$$f(x) = \begin{cases} cx, & 0 < x < 4 \\ 0, & \text{otherwise.} \end{cases}$$

(a) Find $c$.
(b) Find the distribution function $F(x)$.
(c) Compute $P(1 < x < 3)$.

3.

The grades from a statistics class for the first test are given by

| $x_i$ | 96 | 87 | 65 | 49 | 77 | 74 | 99 | 68 | 56 | 84 |
|---|---|---|---|---|---|---|---|---|---|---|
| $p(x_i)$ | 3/15 | 2/15 | 1/15 | 1/15 | 2/15 | 1/15 | 1/15 | 1/15 | 1/15 | 2/15 |

(a) Find mean $\mu$ and variance $\sigma^2$.
(b) Find the mgf.

4.

The probability density function of the random variable $X$ is given by

$$f(x) = \begin{cases} \frac{x^2}{2}, & 0 < x \le 1, \\ \frac{6x-2x^2-3}{2}, & 1 < x \le 2, \\ \frac{(x-3)^2}{2}, & 2 < x \le 3, \\ 0, & \text{otherwise.} \end{cases}$$

Find the expected value of the random variable $X$.

5.

Let the random variable $X$ be normally distributed with mean 0 and variance $\sigma^2$. Show that $E\left(X^{2k+1}\right)=0$, where $k = 0, 1, 2, \ldots$.