MODULE 5

SPECIAL ABSOLUTELY CONTINUOUS DISTRIBUTIONS AND THEIR PROPERTIES

LECTURES 20-24

Topics

MODULE 5

SPECIAL ABSOLUTELY CONTINUOUS DISTRIBUTIONS AND THEIR PROPERTIES

LECTURE 20

Topics

5.1 UNIFORM OR RECTANGULAR DISTRIBUTION

5.1.1 Quantile function and uniform distribution

Recall that a random variable (r.v.) X is said to be of absolutely continuous type if there exists a function $f_X: \mathbb{R} \to [0, \infty)$ such that the distribution function (d.f.) of X is given by

$$F_X(x) = \int_{-\infty}^{x} f_X(t)dt \,, \quad x \in \mathbb{R} \,.$$

The function $f_X(\cdot)$ is called a probability density function (p.d.f) of r.v. X and the set $S_X = \{x \in \mathbb{R} : f_X(x) > 0\}$ is called the support of the p.d.f. $f_X(\cdot)$ (or of r.v. X).

We have seen that the probability distribution of an absolutely continuous type r.v. is completely determined by its p.d.f (or its d.f.). Recall that a function $g: \mathbb{R} \rightarrow \mathbb{R}$ is a p.d.f of some r.v. if, and only if, $g(x) \geq 0, \forall x \in \mathbb{R}$ and

$$\int_{-\infty}^{\infty} g(x)\, dx = 1.$$

5.1 UNIFORM OR RECTANGULAR DISTRIBUTION

Let α and β be real numbers such that $-\infty < \alpha < \beta < \infty$. An absolutely continuous type r.v. X is said to have uniform (or rectangular) distribution over the interval (α, β) (written as $X \sim \mathrm{U}(\alpha, \beta)$) if the p.d.f. of X is given by

$$f_X(x) = \begin{cases} \dfrac{1}{\beta - \alpha}, & \text{if } \alpha < x < \beta \\ 0\ , & \text{otherwise} \end{cases}.$$

Clearly $f_X(x) > 0,\ \forall x \in S_X = (\alpha, \beta)$ and

$$\int_{-\infty}^{\infty} f_X(x)\, dx = 1,$$

i.e., $f_X(\cdot)$ is a proper p.d.f. with support $S_X = (\alpha, \beta)$.

We have a family $\{U(\alpha, \beta): -\infty < \alpha < \beta < \infty\}$ of uniform distributions corresponding to different choices of α and β $(-\infty < \alpha < \beta < \infty)$.

Suppose that $X \sim U(\alpha, \beta)$, for some $-\infty < \alpha < \beta < \infty$. Then, for $r \in \{1,2, \dots\}$,

$$\begin{aligned} \mu_r' &= E(X^r) \\ &= \int_{-\infty}^{\infty} x^r f_X(x) dx \\ &= \int_{\alpha}^{\beta} \frac{x^r}{\beta - \alpha} dx \\ &= \frac{\beta^{r+1} - \alpha^{r+1}}{(r+1)(\beta - \alpha)} \\ &= \frac{\beta^r}{r+1}\left[1 + \frac{\alpha}{\beta} + \left(\frac{\alpha}{\beta}\right)^2 + \cdots + \left(\frac{\alpha}{\beta}\right)^r\right] \end{aligned}$$

and

$$\mu_r = E\left(\left(X-\mu_1'\right)^r\right)$$

$$= E\left(\left(X-\frac{\alpha+\beta}{2}\right)^r\right)$$

$$= \int_\alpha^\beta \left(\left(x-\frac{\alpha+\beta}{2}\right)^r\right)\frac{1}{\beta-\alpha}dx$$

$$= \int_{-\frac{\beta-\alpha}{2}}^{\frac{\beta-\alpha}{2}} \frac{t^r}{\beta-\alpha}dt$$

$$= \begin{cases} 0, & \text{if } r = 1,3,5,\dots \\ \dfrac{(\beta-\alpha)^r}{2^r(r+1)}, & \text{if } r = 2,4,6,\dots \end{cases}.$$

Thus we have

$$\boxed{\mu_r' = E(X^r) = \frac{\beta^r}{r+1}\left[1+\frac{\alpha}{\beta}+\left(\frac{\alpha}{\beta}\right)^2+\cdots+\left(\frac{\alpha}{\beta}\right)^r\right], r = 1,2,\dots}$$

and

$$\boxed{\mu_r = E\left(\left(X-\mu_1'\right)^r\right) = \begin{cases} 0, & \text{if } r = 1,3,5,\dots \\ \dfrac{(\beta-\alpha)^r}{2^r(r+1)}, & \text{if } r = 2,4,6,\dots \end{cases}.}$$

Consequently,

$$\boxed{\text{Mean} = E(X) = \mu_1' = \frac{\beta+\alpha}{2};\ \ \text{Var}(X) = \mu_2 = \frac{(\beta-\alpha)^2}{12},}$$

$$\mu_3 = 0,\ \ \mu_4 = \frac{(\beta-\alpha)^4}{80},$$

$$\boxed{\text{Coefficient of Skewness } = \beta_1 = \frac{\mu_3}{\mu_2^{\frac{3}{2}}} = 0,}$$

and

$$\boxed{\text{Kurtosis } = \gamma_1 = \frac{\mu_4}{\mu_2^2} = \frac{9}{5} = 1.8.}$$

Thus an uniform distribution is highly platykurtic (i.e., in comparison with normal distribution having mean $(\alpha+\beta)/2$, p.d.f. of $U(\alpha,\beta)$ distribution has a flatter peak around its mean). The flatness of p.d.f. around mean is due to distribution being less concentrated around its mean. Moreover the value of coefficient of skewness $\beta_1 = 0$ suggests that the distribution of X may be symmetric about the mean μ_1'. Clearly

$$f_X(\mu_1' - x) = f_X(x - \mu_1'), \quad \forall x \in \mathbb{R},$$

and, therefore, the distribution of $X \sim U(\alpha,\beta)$ is symmetric about $\mu_1' = (\alpha+\beta)/2$.

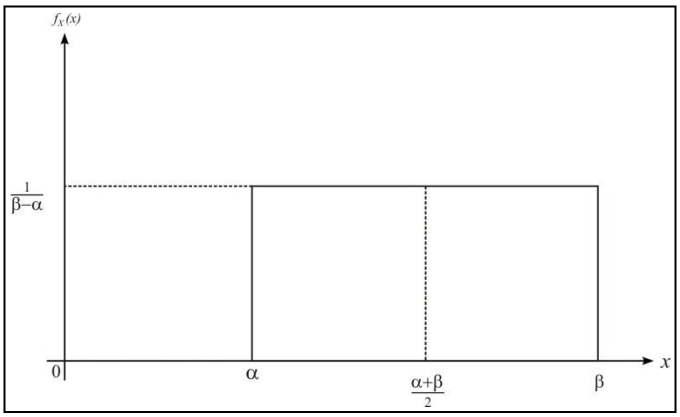

Figure 1.1. Plot of p.d.f. of $U(\alpha,\beta)$ distribution.

Since the distribution of $X \sim U(\alpha,\beta)$ is symmetric about $(\alpha+\beta)/2$, we have

$$\boxed{X - \frac{\alpha+\beta}{2} \overset{d}{=} \frac{\alpha+\beta}{2} - X,}$$

or equivalently

$$\boxed{X \overset{d}{=} \alpha + \beta - X.}$$

The distribution function of $X \sim U(\alpha,\beta)$ is given by

$$F_X(x) = \int_{-\infty}^{x} f_X(t)\,dt$$

$$\boxed{\text{i.e.,} \qquad F_X(x) = \begin{cases} 0, & \text{if } x < \alpha \\ \dfrac{x-\alpha}{\beta-\alpha}, & \text{if } \alpha \le x < \beta \\ 1, & \text{if } x \ge \beta \end{cases}.}$$

Since the distribution of $X \sim U(\alpha,\beta)$ is symmetric about $\mu_1' = (\alpha+\beta)/2$, we have

$$\boxed{\text{Median} = \text{Mean} = \frac{\alpha+\beta}{2}.}$$

One may directly check that

$$F_X\left(\frac{\alpha+\beta}{2}\right) = \frac{1}{2},$$

implying that $(\alpha+\beta)/2$ is the median of $X \sim U(\alpha,\beta)$.

The lower quartile q_1 and the upper quartile q_3 of $X \sim U(\alpha,\beta)$ are given by

$$F_X(q_1) = \frac{1}{4} \text{ and } F_X(q_3) = \frac{3}{4}$$

$$\boxed{\Rightarrow q_1 = \frac{\beta+3\alpha}{4} \text{ and } q_3 = \frac{3\beta+\alpha}{4}.}$$

Also,

$$\boxed{\text{Quartile deviation (QD)} = \frac{q_3-q_1}{2} = \frac{\beta-\alpha}{4}.}$$

The moment generating function of $X \sim U(\alpha,\beta)$ is given by

$$\begin{aligned} M_X(t) &= E(e^{tX}) \\ &= \int_{-\infty}^{\infty} e^{tx} f_X(x)dx \\ &= \int_{\alpha}^{\beta} \frac{e^{tx}}{\beta-\alpha} dt \end{aligned}$$

$$\boxed{\text{i.e.,} \quad M_X(t) = \begin{cases} \dfrac{e^{t\beta}-e^{t\alpha}}{(\beta-\alpha)t}, & \text{if } t \neq 0 \\ 1, & \text{if } t = 0 \end{cases}.}$$

The following theorem provides a characterization of $X \sim U(\alpha,\beta)$ in terms of the property that, for any interval $I \subseteq [\alpha,\beta]$, $P(\{X \in I\})$ depends only on the length of the interval I and not on the location of I on $[\alpha,\beta]$.

Theorem 1.1

Let α and β be real constants such that $-\infty < \alpha < \beta < \infty$ and let X be a random variable of absolutely continuous type with $P(\{\alpha \le X \le \beta\}) = 1$. Then $X \sim U(\alpha,\beta)$, if, and only if, $P(\{X \in I\}) = P(\{X \in J\})$, for any pair of intervals $I, J \subseteq [\alpha, \beta]$ having the same lengths.

Proof. First suppose that $X \sim U(\alpha, \beta)$ and $\alpha \leq a < b \leq \beta$. Then

$$P(\{X \in (a,b)\}) = P(\{X \in [a,b)\}) = P(\{X \in (a,b]\}) = P(\{X \in [a,b]\})$$

$$= F_X(b) - F_X(a)$$

$$= \frac{b-a}{\beta-\alpha},$$

depends only on the length $(= b - a)$ of the interval $(a,b)/[a,b)/(a,b]/[a,b]$.

Conversely suppose that $P(\{X \in I\}) = P(\{X \in J\})$ for any pair of intervals $I, J \subseteq [\alpha, \beta]$ having the same lengths. For $0 < s \leq 1$, let

$$G(s) = P(\{\alpha < X \leq \alpha + (\beta - \alpha)s\}) = F_X(\alpha + (\beta - \alpha)s),$$

where $F_X(\cdot)$ is the d.f. of X. Then, for $0 < s_1 \leq 1$, $0 < s_2 \leq 1$, $0 < s_1 + s_2 \leq 1$,

$$P(\{\alpha + (\beta - \alpha)s_1 < X \leq \alpha + (\beta - \alpha)(s_1 + s_2)\}) = P(\{\alpha < X \leq \alpha + (\beta - \alpha)s_2\}),$$

and therefore

$$G(s_1 + s_2) = P(\{\alpha < X \leq \alpha + (\beta - \alpha)(s_1 + s_2)\})$$

$$= P(\{\alpha < X \leq \alpha + (\beta - \alpha)s_1\}) + P(\{\alpha + (\beta - \alpha)s_1 < X \leq \alpha + (\beta - \alpha)(s_1 + s_2)\})$$

$$= P(\{\alpha < X \leq \alpha + (\beta - \alpha)s_1\}) + P(\{\alpha < X \leq \alpha + (\beta - \alpha)s_2\})$$

$$= G(s_1) + G(s_2).$$

By induction, for $0 < s_i \leq 1$, $i = 1, \ldots, n$ and $0 < \sum_{i=1}^{n} s_i \leq 1$, we have

$$G(s_1 + s_2 + \cdots + s_n) = G(s_1) + G(s_2) + \cdots + G(s_n).$$

Consequently

$$G(ms) = mG(s), \qquad \forall\ 0 < s \leq \frac{1}{m} \tag{1.1}$$

and $$G(s) = G\Bigg(\underbrace{\frac{s}{n} + \cdots + \frac{s}{n}}_{\text{n times}}\Bigg) = nG\left(\frac{s}{n}\right),\ \ 0 < s \leq 1. \tag{1.2}$$

Also, for $m, n \in \{1,2 \ldots\ldots\}$, $m < n$,

$$G\left(\frac{m}{n}\right) = G\left(\underbrace{\frac{1}{n} + \cdots + \frac{1}{n}}_{\text{m times}}\right)$$

$$= mG\left(\frac{1}{n}\right) \qquad \text{(using(1.1))}$$

$$= \frac{m}{n} G(1) \qquad \text{(using (1.2))}$$

$$= \frac{m}{n} F_X(\beta)$$

$$= \frac{m}{n}$$

$$\Longrightarrow G(r) = r, \qquad \forall r \in \mathbb{Q} \cap (0,1), \qquad (1.3)$$

where $\mathbb{Q}$ denotes the set of rational numbers. Now let $x \in (0,1)$. Choose a sequence $\{r_n : n = 1, 2, \dots\}$ in $\mathbb{Q} \cap (0,1)$ such that $r_n \downarrow x$ (existence of such a sequence is guaranteed). Then

$$G(x) = \lim_{n\to\infty} G(r_n) \qquad \text{(since } G(x) = F_X(\alpha + (\beta - \alpha)x) \text{ is right continuous)}$$

$$= \lim_{n\to\infty} r_n \qquad \text{(using (1.3))}$$

$$= x.$$

It follows that

$$F_X(\alpha + (\beta - \alpha)x) = x, \qquad \forall x \in (0,1)$$

$$\Longrightarrow F_X(x) = \frac{x - \alpha}{\beta - \alpha}, \quad \forall x \in (\alpha, \beta).$$

Also, since F_X is continuous on $\mathbb{R}$ and $P(\{\alpha \le X \le \beta\}) = 1$, we have

$$F_X(x) = \begin{cases} 0, & \text{if } x < \alpha \\ \dfrac{x-\alpha}{\beta-\alpha}, & \text{if } \alpha \le x < \beta \\ 1, & \text{if } x \ge \beta \end{cases}$$

$$\Longrightarrow X \sim U(\alpha, \beta). \blacksquare$$

Theorem 1.2

Suppose that $X \sim U(\alpha, \beta)$, for some real constants α and β such that $-\infty < \alpha < \beta < \infty$. Then $Y = \frac{X-\alpha}{\beta-\alpha} \sim U(0,1)$.

Proof. The p.d.f of X is

$$f_X(x) = \begin{cases} \frac{1}{\beta-\alpha}, & \text{if } \alpha < x < \beta \\ 0, & \text{otherwise} \end{cases}.$$

Let $Y = \frac{X-\alpha}{\beta-\alpha} = h(X)$, say. Clearly $h(x) = \frac{x-\alpha}{\beta-\alpha},\ x \in S_X = (\alpha, \beta)$ is strictly increasing on S_X. Therefore the r.v. $Y = \frac{X-\alpha}{\beta-\alpha}$ is of absolutely continuous type with support $S_Y = h(S_X) = (0,1)$ and p.d.f.

$$f_Y(y) = f_X\left(h^{-1}(y)\right)\left|\frac{d}{dy}h^{-1}(y)\right| I_{h(S_X)}(y).$$

We have $h(S_X) = (0,1)$ and $h^{-1}(y) = \alpha + (\beta - \alpha)y, y \in h(S_X) = (0, 1)$. Therefore

$$f_Y(y) = f_X(\alpha + (\beta - \alpha)y)|\beta - \alpha| I_{(0,1)}(y)$$

$$= \begin{cases} 1, & \text{if } 0 < y < 1 \\ 0, & \text{otherwise} \end{cases}.$$

$$\Rightarrow Y = \frac{X-\alpha}{\beta-\alpha} \sim U(0,1). \blacksquare$$

Example 1.1

Let $a > 0$ be a real constant. A point X is chosen at random on the interval $(0, a)$ (i.e., $X \sim U(0, a)$).

(i) If Y denotes the area of equilateral triangle having sides of length X, find the mean and variance of Y.

(ii) If the point X divides the interval $(0, a)$ into subintervals $I_1 = (0, X)$ and $I_2 = [X, a)$, find the probability that the larger of these two subintervals is at least the double the size of the smaller subinterval.

Solution.

(i) In the equilateral triangle ABC

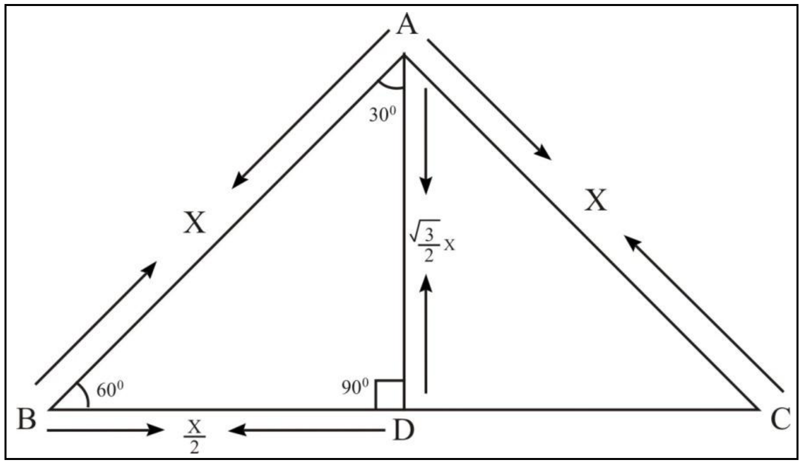

Figure 1.2

$$AB = BC = X, \qquad BD = \frac{X}{2} \text{ and } AD = \frac{\sqrt{3}}{2}X.$$

Therefore

$$Y = \frac{1}{2} \times X \times \frac{\sqrt{3}}{2}X = \frac{\sqrt{3}}{4}X^2$$

$$E(Y) = \frac{\sqrt{3}}{4}E(X^2) = \frac{\sqrt{3}}{12}a^2,$$

$$E(Y^2) = \frac{3}{16}E(X^4) = \frac{3}{80}a^4$$

and $\quad \mathrm{Var}(Y) = E(Y^2) - \left(E(Y)\right)^2 = \frac{a^4}{60}.$

(ii) The required probability is

$$\begin{aligned} p &= P(\{\max\{X, a - X\} > 2\min\{X, a - X\}\}) \\ &= P\left(\left\{a - X > 2X, X \le \frac{a}{2}\right\}\right) + P\left(\left\{X > 2(a - X), X > \frac{a}{2}\right\}\right) \\ &= P\left(\left\{X \le \frac{a}{3}\right\}\right) + P\left(\left\{X > \frac{2}{3}a\right\}\right) \\ &= F_X\left(\frac{a}{3}\right) + 1 - F_X\left(\frac{2}{3}a\right) \\ &= \frac{1}{3} + 1 - \frac{2}{3} = \frac{2}{3}. \blacksquare \end{aligned}$$

Remark 1.1

Uniform distribution is applicable in situations where the outcome of random experiment is a number X chosen at random from an interval $[\alpha, \beta]$ in the sense that if $I \subseteq [\alpha, \beta]$ is any interval then $P(\{X \in I\})$ depends only on the length of I and not on its location in $[\alpha, \beta]$. ■

5.1.1 *Quantile function and uniform distribution*

We begin this section with the definition of quantile function.

Definition 1.1

Let X be a random variable (not necessarily of absolutely continuous type) with distribution function $F_X(\cdot)$.

(i) The function $Q_X: (0,1) \to \mathbb{R}$, defined by, $Q_X(p) = \inf\{s \in \mathbb{R}: F_X(s) \geq p\}, 0 < p < 1$, is called the *quantile function* (q.f.) of the random variable X (or of distribution function $F_X(\cdot)$).

(ii) For a fixed $p \in (0,1)$, the quantity $Q_X(p) = \inf\{s \in \mathbb{R}: F_X(s) \geq p\}$ is called the *p-th quantile* of X (or of $F_X(\cdot)$). ■

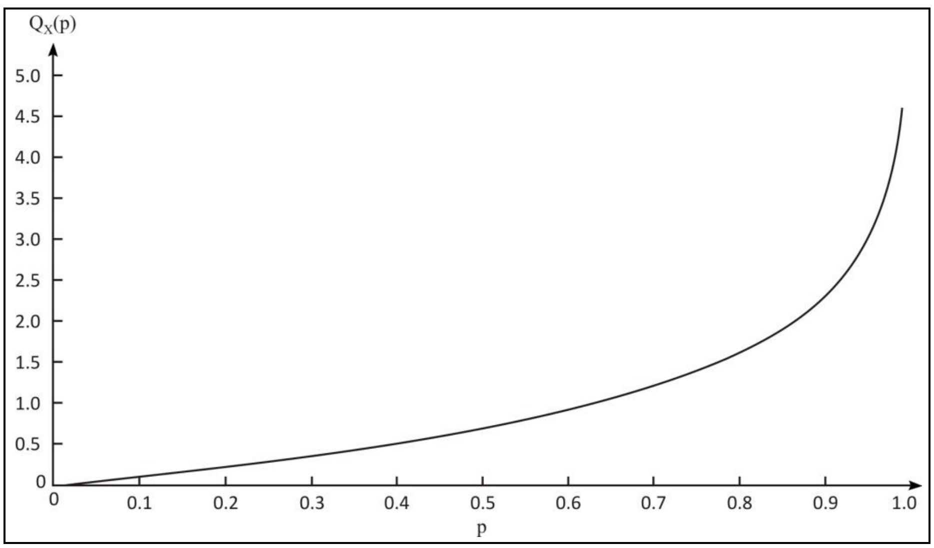

Figure 1.3. Plot of quantile function.

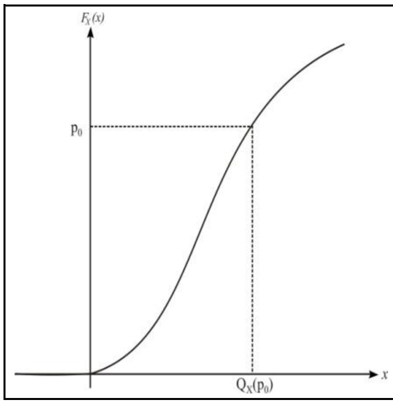

Figure 1.4. (a)

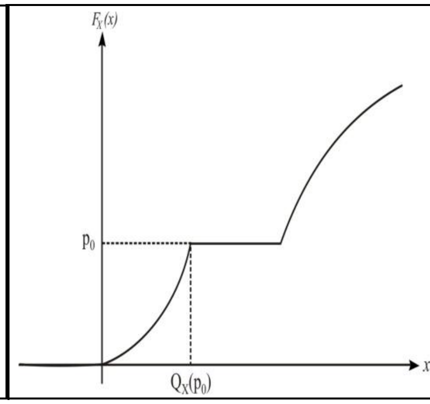

Figure 1.4. (b)

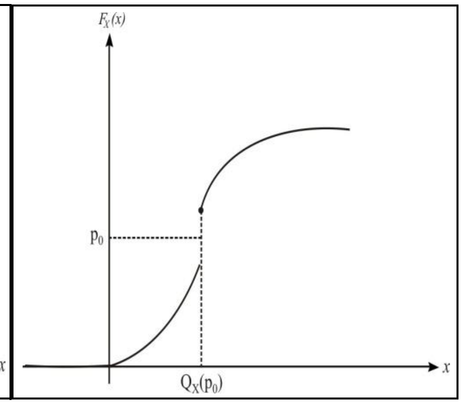

Figure 1.4. (c)

Remark 1.2

If the distribution function $F_X(\cdot)$ is continuous and strictly increasing on $\mathbb{R}$ then $Q_X(p) = F_X^{-1}(p),\ 0 < p < 1.$ ■