IC105: Probability and Statistics **2021-22-M**

Lecture 13: Some Special Continuous Distributions

Instructor: Dr. Kuldeep Kumar Kataria *Scribe:*

13.0.2. Discrete Uniform Distribution

Let N be a given positive integer and $x_1 < x_2 < \cdots < x_N$ be given real numbers. A r.v. X is said to follow a discrete uniform distribution on the set $\{x_1, x_2, \ldots, x_N\}$ (written as $X \sim U(\{x_1, x_2, \ldots, x_N\})$) if its p.m.f. is given by

$$f_X(x) = P(X = x) = \begin{cases} \frac{1}{N}, & x \in \{x_1, x_2, \ldots, x_N\}, \\ 0, & \text{otherwise.} \end{cases}$$

Suppose that $X \sim U(\{x_1, x_2, \ldots, x_N\})$. Then,

$$\mu'_r = E(X^r) = \frac{1}{N}\sum_{i=1}^{N} x_i^r,$$

$$\text{Mean} = \mu'_1 = \frac{1}{N}\sum_{i=1}^{N} x_i,$$

$$\text{Var}(X) = \sigma^2 = E((X - \mu'_1)^2) = \frac{1}{N}\sum_{i=1}^{N}(x_i - \mu'_1)^2,$$

$$\text{m.g.f. } M_X(t) = E(e^{tX}) = \frac{1}{N}\sum_{i=1}^{N} e^{tx_i}.$$

Suppose that $Y \sim U(\{1, 2, \ldots, N\})$. Then,

$$\mu'_1 = \text{Mean} = E(Y) = \frac{1}{N}\sum_{i=1}^{N} i = \frac{N+1}{2},$$

$$\mu'_2 = E(Y^2) = \frac{1}{N}\sum_{i=1}^{N} i^2 = \frac{(N+1)(2N+1)}{6},$$

$$\mu'_3 = E(Y^3) = \frac{1}{N}\sum_{i=1}^{N} i^3 = \frac{N(N+1)^2}{4},$$

$$\mu'_4 = E(Y^4) = \frac{1}{N}\sum_{i=1}^{N} i^4 = \frac{(N+1)(2N+1)(3N^2+3N-1)}{30},$$

$$\mu_2 = E((Y - \mu'_1)^2) = \frac{N^2-1}{12},$$

$$\mu_3 = E((Y - \mu'_1)^3) = 0,$$

$$\mu_4 = E((Y - \mu'_1)^4) = \frac{(3N^2-7)(N^2-1)}{240},$$

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

$$\text{Kurtosis} = \nu_1 = \frac{\mu_4}{\mu_2^2} = \frac{3}{5}\frac{(3N^2-7)}{N^2-1},$$

$$\text{m.g.f. } M_Y(t) = E(e^{tY}) = \frac{1}{N}\sum_{j=1}^{N} e^{tj} = \begin{cases} \frac{e^t(e^{Nt}-1)}{N(e^t-1)}, & t \neq 0, \\ 1, & t = 0. \end{cases}$$

Example 13.5. *A person has to open a lock whose key is lost among a set of N keys. Assume that out of these N keys only one can open the lock. To open the lock the person tries keys one by one by choosing at each attempt one of the keys at random from the unattempted keys. The unsuuccessful keys are not considered for future attempts. Let Y denote the number of attempts the person will have to make to open the lock. Show that $Y \sim U(\{1, 2, \ldots, N\})$ and hence find the mean and variance of the r.v. Y.*

Solution: For $r \neq \{1, 2, \ldots, N\}$, we have $P(Y = r) = 0$. For $r \in \{1, 2, \ldots, N\}$, we have

$$P(Y = r) = \frac{N-1}{N} \cdot \frac{N-2}{N-1} \cdots \frac{N-(r-1)}{N-(r-2)} \cdot \frac{1}{N-(r-1)} = \frac{1}{N} \implies Y \sim U(\{1, 2, \ldots, N\}).$$

This implies $E(Y) = \dfrac{N+1}{2}$ and $\mathrm{Var}(Y) = \dfrac{N^2-1}{12}$.

13.1. Some Special Continuous Distributions

13.1.1. Uniform or Rectangular Distribution

Let $-\infty < \alpha < \beta < \infty$. An absolutely continuous type r.v. X is said to have a uniform (or rectangular) distribution over the interval (α, β) (written as $X \sim U(\alpha, \beta)$) if its p.d.f. is given by

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

$\{U(\alpha, \beta) : -\infty < \alpha < \beta < \infty\}$ is a family of 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

$$\mu_r' = E(X^r) = \int_\alpha^\beta \frac{x^r}{\beta - \alpha} \mathrm{d}x = \frac{\beta^{r+1} - \alpha^{r+1}}{(r+1)(\beta - \alpha)} = \frac{\beta^r}{r+1}\left[1 + \frac{\alpha}{\beta} + \left(\frac{\alpha}{\beta}\right)^2 + \cdots + \left(\frac{\alpha}{\beta}\right)^r\right],$$

$$E(X) = \frac{\alpha + \beta}{2} = \mu_1',$$

$$\mu_r = E(X - \mu_1')^r = \int_\alpha^\beta \left(x - \frac{\alpha+\beta}{2}\right)^r \frac{1}{\beta - \alpha} \mathrm{d}x = \int_{-\frac{\beta-\alpha}{2}}^{\frac{\beta-\alpha}{2}} \frac{t^r}{\beta - \alpha} \mathrm{d}t = \begin{cases} 0, & r = 1, 3, 5, \ldots, \\ \dfrac{(\beta-\alpha)^r}{2^r(r+1)}, & r = 2, 4, 6, \ldots. \end{cases}$$

Also,

$$f_X\left(x - \frac{\alpha+\beta}{2}\right) = f_X\left(\frac{\alpha+\beta}{2} - x\right) = \begin{cases} \dfrac{1}{\beta - \alpha}, & -\frac{\beta-\alpha}{2} < x < \frac{\beta-\alpha}{2}, \\ 0, & \text{otherwise} \end{cases} \implies X - \frac{\alpha+\beta}{2} \stackrel{d}{=} \frac{\alpha+\beta}{2} - X.$$

This implies distribution of X is symmetric about its mean $\mu_1' = \frac{\alpha+\beta}{2}$.

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

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

$$\text{Kurtosis} = \nu_1 = \frac{\mu_4}{\mu_2^2} = \frac{9}{5} = 1.8,$$

$$\text{The d.f. of } X \sim U(\alpha,\beta) \text{ is given by } F(x) = \begin{cases} 0, & x < \alpha, \\ \dfrac{x-\alpha}{\beta-\alpha}, & \alpha \le x < \beta, \\ 1, & x \ge \beta. \end{cases}$$

Theorem 13.6. *Let $-\infty < \alpha < \beta < \infty$ and let X be a r.v. of continuous type with $P(\alpha \le X \le \beta) = 1$. Then $X \sim U(\alpha,\beta) \iff P(X \in I) = P(X \in J)$, for any pairs of intervals $I, J \subseteq (\alpha,\beta)$ having the same length.*

Proof. Suppose that $X \sim U(\alpha,\beta)$. Then, for $\alpha \le a < b \le \beta$, we have

$$\begin{aligned} P(X \in (a,b)) &= P(X \in [a,b)) = P(X \in (a,b]) = P(X \in [a,b]) \\ &= F(b|\alpha,\beta) - F(a|\alpha,\beta) \\ &= \frac{b-a}{\beta-\alpha} \to \text{dependes only on length } b-a \text{ of the interval } (a,b)/[a,b)/(a,b]/[a,b]. \end{aligned}$$

Conversely, suppose that $P(X \in I) = P(X \in J)$, for all pairs of intervals $I, J \subseteq (\alpha,\beta)$ having the same length. For $0 < s \le 1$, let $G(s) = P(\alpha < X \le \alpha + (\beta-\alpha)s) = F(\alpha + (\beta-\alpha)s|\alpha,\beta)$. Then for $0 < s_1, s_2 \le 1$, $0 < s_1 + s_2 \le 1$,

$$\begin{aligned} G(s_1+s_2) &= P(\alpha < X \le \alpha + (\beta-\alpha)(s_1+s_2)) \\ &= P(\alpha < X \le \alpha + (\beta-\alpha)s_1) + P\left(\underbrace{\alpha + (\beta-\alpha)s_1 < X \le \alpha + (\beta-\alpha)(s_1+s_2)}_{\text{Depends only on the length } (\beta-\alpha)s_2 \text{ of } (\alpha+(\beta-\alpha)s_1,\, \alpha+(\beta-\alpha)(s_1+s_2))}\right) \\ &= G(s_1) + P(\alpha < X \le \alpha + (\beta-\alpha)s_2) = G(s_1) + G(s_2). \end{aligned}$$

By induction, for $0 < s_i \le 1$, $i = 1, 2, \ldots, n$, $0 < \sum_{i=1}^n s_i \le 1$, we have $G(s_1 + s_2 + \cdots + s_n) = G(s_1) + G(s_2) + \cdots + G(s_n)$. This implies that

$$G(ms) = mG(s),\ \forall\, 0 < s \le \frac{1}{m}, \tag{13.1}$$

$$G(s) = G\left(\underbrace{\frac{s}{n} + \frac{s}{n} + \cdots + \frac{s}{n}}_{n \text{ times}}\right) = nG\left(\frac{s}{n}\right). \tag{13.2}$$

For $m, n \in \{1, 2, \dots\}$, $m < n$, we get

$$\begin{aligned}
G\left(\frac{m}{n}\right) &= G\left(\underbrace{\frac{1}{n}+\frac{1}{n}+\cdots+\frac{1}{n}}_{m \text{ times}}\right) \\
&= mG\left(\frac{1}{n}\right), \quad \text{using (13.1)} \\
&= \frac{m}{n}G(1), \quad \text{using (13.2)} \\
&= \frac{m}{n}F(\beta|\alpha,\beta) = \frac{m}{n} \implies G(r) = r, \ \forall\, r \in IQ \cap (0,1),
\end{aligned}$$

where IQ denotes the set of rational numbers. Now let $x \in (0,1)$. Then there exists a sequence $\{r_n\}_{n\geq 1}$ in $IQ \cap (0,1)$ such that $r_n \downarrow x$ (rationals are dense in $(0,1)$). Then, since G is continuous, we have

$$G(x) = \lim_{n\to\infty} G(r_n) = \lim_{n\to\infty} r_n = x, \ \ \forall\, x \in (0,1).$$

This implies

$$\begin{aligned}
& F(\alpha + (\beta-\alpha)x|\alpha,\beta) = x, \ \ \forall\, x \in (0,1) \\
\implies & F(x|\alpha,\beta) = \frac{x-\alpha}{\beta-\alpha}, \ \ x \in (\alpha,\beta) \\
\implies & F(x|\alpha,\beta) = \begin{cases} 0, & x < \alpha, \\ \dfrac{x-\alpha}{\beta-\alpha}, & \alpha \le x < \beta, \\ 1, & x \ge \beta, \end{cases} \implies X \sim U(\alpha,\beta).
\end{aligned}$$

This completes the proof. □

$$\text{M.g.f. } M_X(t) = E(e^{tX}) = \int_\alpha^\beta \frac{e^{tx}}{\beta-\alpha}\mathrm{d}x = \begin{cases} \dfrac{e^{t\beta}-e^{t\alpha}}{t(\beta-\alpha)}, & t \neq 0, \\ 1, & t = 0. \end{cases}$$

Theorem 13.7. *Let $X \sim U(\alpha,\beta)$, $-\infty < \alpha < \beta < \infty$. Then,*

(i) for $a > 0$ and $b \in \mathbb{R}$, $Y = aX + b \sim U(a\alpha + b, a\beta + b)$,

(ii) for $a < 0$ and $b \in \mathbb{R}$, $Y = aX + b \sim U(a\beta + b, a\alpha + b)$,

(iii) $Z = \dfrac{X-\alpha}{\beta-\alpha} \sim U(0,1)$.

Proof. Exercise. □

Recall that quantile function is defined by $Q_X(p) = \inf\{s \in \mathbb{R} : F_X(s) \geq p\}$, $0 < p < 1$.

Theorem 13.8. *Let X be a r.v. with d.f. F and quantile function $Q(\cdot)$. Then*

*(i) (**Probability Integral Transform**)*

$$X \text{ is of continuous type} \implies F(X) \sim U(0,1)$$

(ii) $U \sim U(0,1) \implies Q(U) \stackrel{d}{=} X$.

Proof. (i) Let G be the d.f. of $Y = F(X)$. Then $G(y) = P(F(X) \le y)$, $y \in \mathbb{R}$. Clearly, for $y < 0$, $G(y) = 0$ and for $y \ge 1$, $G(y) = 1$. For $y \in [0, 1)$,

$$\begin{aligned}
&\{s \in \mathbb{R} : F(s) \ge y\} = \{s \in \mathbb{R} : s \ge Q(y)\} \\
\implies &P(F(X) \ge y) = P(X \ge Q(y)) \\
\implies &P(F(X) < y) = P(X < Q(y)) \\
\implies &P(F(X) < y) = P(X \le Q(y)) = F(Q(y)) = y, \quad \text{since } X \text{ is of continuous type.}
\end{aligned}$$

Since X is of continuous type $P(F(X) = y) = P(x_1 \le X \le x_2) = 0$ for some x_1, x_2 with $F(x_1) = F(x_2)$. Thus,

$$\begin{aligned}
&P(F(X) \le y) = y, \ \forall\, y \in (0,1), \\
\implies G(y) = &\begin{cases} 0, & \text{if } y < 0, \\ y, & \text{if } 0 \le y < 1, \\ 1, & \text{if } y \ge 1, \end{cases} \implies Y \sim U(0,1).
\end{aligned}$$

(ii) Let $U \sim U(0,1)$ and let $Z = Q(U)$. Then the d.f. of Z is

$$H(z) = P(Z \le z) = P(Q(U) \le z) = P(Q(U) \le z, 0 < U < 1).$$

Note that for $z \in (0,1)$, $\{p \in \mathbb{R} : Q(p) \le z\} = \{p \in \mathbb{R} : F(z) \ge p\}$. Thus, for $z \in (0,1)$

$$H(z) = P(F(Z) \ge U, 0 < U < 1) = P(U \le F(z)) = F(z) \implies Z = Q(U) \stackrel{d}{=} X.$$

This completes the proof. □

Remark 13.9. *The above theorem provides a method to generate observations from any arbitrary distributions using $U(0,1)$ observations. Suppose that we require an observation X from a distribution having d.f. F and quantile functions Q. To do so, the above theorem suggests that generate an observation U from $U(0,1)$ distribution and take $X = Q(U)$.*

13.1.2. Gamma and Related Distributions

Gamma Function: $\Gamma : (0,\infty) \to (0,\infty)$ is defined as

$$\Gamma(\alpha) = \int_0^\infty e^{-t} t^{\alpha-1} \mathrm{d}t, \quad \alpha > 0.$$

It converges for any $\alpha > 0$. Integration by parts yields $\Gamma(\alpha+1) = \alpha\Gamma(\alpha)$, $\alpha > 0$ and $\Gamma(1) = 1$. For any $n \in \mathbb{N}$,

$$\Gamma(n) = (n-1)!, \quad \Gamma\left(\frac{1}{2}\right) = \int_0^\infty e^{-t} t^{-1/2} \mathrm{d}t = 2\int_0^\infty e^{-x^2} \mathrm{d}x$$

This implies

$$\begin{aligned}
\left(\Gamma\left(\frac{1}{2}\right)\right)^2 &= 4\int_0^\infty \int_0^\infty e^{-(x^2+y^2)} \mathrm{d}x\mathrm{d}y \\
&= 4\int_0^\infty \int_0^{\pi/2} re^{-r^2} \mathrm{d}\theta\mathrm{d}r, \quad (x = r\cos\theta,\ y = r\sin\theta) \\
&= \pi \implies \Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}.
\end{aligned}$$

Also,

$$\Gamma\left(\frac{3}{2}\right)=\frac{1}{2}\Gamma\left(\frac{1}{2}\right)=\frac{\sqrt{\pi}}{2},\ \ \Gamma\left(\frac{5}{2}\right)=\frac{3}{2}\cdot\frac{1}{2}\Gamma\left(\frac{1}{2}\right)=\frac{1\cdot 3\sqrt{\pi}}{2^2},$$
$$\Gamma\left(\frac{2n+1}{2}\right)=\frac{1\cdot 3\cdot 5\cdots(2n-1)\sqrt{\pi}}{2^n}=\frac{(2n)!}{n!4^n}\sqrt{\pi},\ \ n\in\mathbb{N}.$$

Clearly,

$$\int_0^\infty e^{-x/\theta}x^{\alpha-1}\mathrm{d}x=\theta^\alpha\Gamma(\alpha),\ \ \alpha>0,\ \theta>0.$$

Definition 13.10. *A r.v. X is said to have a gamma distribution with shape parameter $\alpha>0$ and scale parameter $\theta>0$ (written as $X\sim GAM(\alpha,\theta)$ if its p.d.f. is given by*

$$f(x|\alpha,\theta)=\begin{cases}\dfrac{e^{-x/\theta}x^{\alpha-1}}{\theta^\alpha\Gamma(\alpha)}, & x>0,\\ 0, \ \textit{otherwise}.\end{cases}\rightarrow \textit{family of distributions}\ \ \{GAM(\alpha,\theta),\alpha>0,\theta>0\}.$$

Let $X\sim GAM(\alpha,\theta)\implies \dfrac{X}{\theta}\sim GAM(\alpha,1)$ (θ is called scale parameter since the distribution of $\dfrac{X}{\theta}$ does not depend on θ). The p.d.f. of $Z\sim GAM(\alpha,1)$ is $f(z)=\begin{cases}\dfrac{e^{-z}z^{\alpha-1}}{\Gamma(\alpha)}, & z>0,\\ 0, & \text{otherwise.}\end{cases}$

Also,

$$\begin{aligned}E(Z^r)=\frac{1}{\Gamma(\alpha)}\int_0^\infty z^{\alpha+r-1}e^{-z}\mathrm{d}z&=\frac{\Gamma(\alpha+r)}{\Gamma(\alpha)},\ \ r>-\alpha,\ \alpha>0,\\ &=\alpha(\alpha+1)\cdots(\alpha+r-1),\ \ \text{if}\ \ r\in\mathbb{N}.\end{aligned}$$

$$\begin{aligned}&\text{Mean}=\mu_1'=E(X)=\alpha\theta,\ \ \mu_2'=E(X^2)=\alpha(\alpha+1)\theta^2,\ \ \mu_2=\sigma^2=\text{Var}(X)=\alpha\theta^2,\\ &\mu_3=E((X-\mu_1')^3)=\mu_3'-3\mu_1'\mu_2'+2(\mu_1')^3=2\alpha\theta^3,\\ &\mu_4=E((X-\mu_1')^4)=\mu_4'-4\mu_1'\mu_3'+6(\mu_1')^2\mu_2'-3(\mu_1')^4=3\alpha(\alpha+2)\theta^4,\\ &\text{Coefficient of skewness}=\beta_1=\frac{\mu_3}{\mu_2^{\frac{3}{2}}}=\frac{2}{\sqrt{\alpha}},\ \ \text{Kurtosis}=\nu_1=\frac{\mu_4}{\mu_2^2}=3+\frac{6}{\alpha}.\end{aligned}$$

For $0<\alpha\le 1$, $f(x|\alpha,\theta)\downarrow$ and for $\alpha>1$, $f(x|\alpha,\theta)\uparrow$ in $(0,(\alpha-1)\theta)$ and $\downarrow$ in $((\alpha-1)\theta,\infty)$.

$$\begin{aligned}\text{m.g.f. } M_X(t)&=E(e^{tX})=E(e^{t\theta Z}),\ \ (Z=X/\theta)\\ &=\frac{1}{\Gamma(\alpha)}\int_0^\infty e^{t\theta z}e^{-z}z^{\alpha-1}\mathrm{d}z=\frac{1}{\Gamma(\alpha)}\int_0^\infty e^{-(1-t\theta)z}z^{\alpha-1}\mathrm{d}z=(1-t\theta)^{-\alpha},\ \ t<\frac{1}{\theta}.\end{aligned}$$

Theorem 13.11. *Let $X_1,X_2,\ldots,X_k$ be independent r.v.'s such that $X_i\sim GAM(\alpha_i,\theta)$, for some $\alpha_i>0$, $\theta>0$, $i=1,2,\ldots,k$. Then $Y=\sum_{i=1}^k X_i\sim GAM(\sum_{i=1}^k\alpha_i,\theta)$.*

Proof. Note that

$$M_Y(t)=\prod_{i=1}^k M_{X_i}(t)=\prod_{i=1}^k(1-t\theta)^{-\alpha_i}=(1-t\theta)^{-\sum_{i=1}^k\alpha_i},\ \ t<\frac{1}{\theta}=\text{m.g.f. of } GAM(\sum_{i=1}^k\alpha_i,\theta).$$

This completes the proof. $\square$

Theorem 13.12 (Relationship beween Gamma and Poisson distribution)**.** *For $n \in \mathbb{N}$, $\theta > 0$ and $t > 0$, let $X \sim GAM(n, \theta)$ and $Y \sim Po(t/\theta)$. Then $P(X > t) = P(Y \leq n-1)$, i.e.*

$$\frac{1}{(n-1)!\theta^n} \int_t^{\infty} e^{-x/\theta} x^{n-1} \mathrm{d}x = \sum_{j=0}^{n-1} \frac{e^{-t/\theta}(t/\theta)^j}{j!}.$$

Proof. Use integration by parts. □

Remark 13.13. *For $n \in \mathbb{N}$ and $\theta > 0$, let $X \sim GAM(n, \theta)$. Then*

$$\sum_{j=n}^{\infty} \frac{e^{-X/\theta}(X/\theta)^j}{j!} \sim U(0,1) \ \textit{and} \ \sum_{j=0}^{n-1} \frac{e^{-X/\theta}(X/\theta)^j}{j!} \sim U(0,1) \quad (U \sim U(0,1) \implies 1-U \sim U(0,1)).$$

Definition 13.14. *For a $\theta > 0$, a $GAM(1, \theta)$ distribution is called exponential distribution with scale parameter θ (denoted by $Exp(\theta)$).*

The p.d.f. of $T \sim Exp(\theta)$ is given by

$$f_T(t) = \begin{cases} \frac{1}{\theta} e^{-t/\theta}, & t > 0, \\ 0, & \text{otherwise.} \end{cases}$$

and its d.f. is given by

$$F_T(t) = P(T \leq t) = 1 - P(T > t) = \begin{cases} 0, & t \leq 0, \\ 1 - e^{-t/\theta}, & t > 0. \end{cases}$$

Mean= $E(T) = \theta$, variance= θ^2, $\mu_r' = E(T^r) = r!\theta^r$, $r \in \mathbb{N}$, coefficient of skewness=$\beta_1 = 2$, Kurtosis=$\nu_1 = 9$. M.g.f.=$M_T(t) = (1 - t\theta)^{-1}$, $t < 1/\theta$ and

$$P(T > t) = \begin{cases} 1, & t \leq 0, \\ e^{-t/\theta}, & t > 0. \end{cases}$$

For $s > 0, t > 0$

$$P(T > s+t | T > s) = \frac{P(T > s+t)}{P(T > s)} = e^{-t/\theta} = P(T > t)$$
$$\implies P(T > s+t) = P(T > s)P(T > t), \ \forall \ s, t > 0 \rightarrow \text{Lack of Memory Property.}$$

Let T denote the lifetime of a system. Given that the system has survived $s(> 0)$ units of time the probability that it will survive t additional units of time is the same as the probability that a fresh system (of age 0) will survive t units of time. In other words, the system has no memory of its current age or it is not ageing with time.

13.1.3. Beta Distribution

For $\alpha > 0$ and $\beta > 0$, we have

$$\begin{aligned}
\Gamma(\alpha)\Gamma(\beta) &= \int_0^\infty \int_0^\infty e^{-(s+t)} s^{\alpha-1} t^{\beta-1} \mathrm{d}s\mathrm{d}t \\
&= \int_0^1 \int_0^\infty e^{-v} (uv)^{\alpha-1} ((1-u)v)^{\beta-1} |v| \mathrm{d}v\mathrm{d}u, \\
&\qquad\qquad \text{making transformation: } s = uv,\ t = (1-u)v,\ \text{Jacobian} : J = v \\
&= \Gamma(\alpha+\beta) \int_0^1 u^{\alpha-1}(1-u)^{\beta-1} \mathrm{d}u \\
\implies \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)} &= \int_0^1 u^{\alpha-1}(1-u)^{\beta-1}\mathrm{d}u \to \text{Beta function (function of } (\alpha,\beta),\ \alpha>0,\ \beta>0).
\end{aligned}$$

Note that $B(\alpha, \beta) = B(\beta, \alpha), \forall\ \alpha,\ \beta > 0$.

Definition 13.15. *For given $\alpha > 0$ and $\beta > 0$, a r.v. X is said to have the beta distribution with parameter (α, β) (written as $X \sim Be(\alpha, \beta)$) if its p.d.f. is given by*

$$f(x|\alpha,\beta) = \begin{cases} \dfrac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha,\beta)}, & 0 < x < 1, \\ 0, & \textit{otherwise}. \end{cases}$$

Suppose that $X \sim Be(\alpha, \beta)$), for some $\alpha > 0$ and $\beta > 0$. Then

$$\begin{aligned}
& E(X^r) = \frac{B(\alpha+r,\beta)}{B(\alpha,\beta)} = \frac{\Gamma(\alpha+r)\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\alpha+\beta+r)},\ \ r > -\alpha, \\
& \text{Mean} = \mu_1' = E(X) = \frac{\alpha}{\alpha+\beta},\ \ \mu_2' = E(X^2) = \frac{\alpha(\alpha+1)}{(\alpha+\beta)(\alpha+\beta+1)}, \\
& \mu_2 = \sigma^2 = \mathrm{Var}(X) = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}, \\
& \text{Mode} = M_0 = \frac{\alpha-1}{\alpha+\beta-2},\ \ \text{if } \alpha > 1 \text{ and } \alpha+\beta > 2, \\
& \text{Skewness} = \beta_1 = \frac{\mu_3^2}{\mu_2^3} = \frac{2(\beta-\alpha)\sqrt{\alpha+\beta+1}}{\sqrt{\alpha\beta}(\alpha+\beta+2)}, \\
& \text{Kurtosis} = \nu_1 = \frac{\mu_4}{\mu_2^2} = \frac{6[(\alpha-\beta)^2(\alpha+\beta+1) - \alpha\beta(\alpha+\beta+2)]}{\alpha\beta(\alpha+\beta+2)(\alpha+\beta+3)} + 3 \\
& \qquad\qquad\qquad\quad = \frac{6[\alpha^3 + \alpha^2(1-2\beta) + \beta^2(1+\beta) - 2\alpha\beta(2+\beta)]}{\alpha\beta(\alpha+\beta+2)(\alpha+\beta+3)}.
\end{aligned}$$

Let $X \sim Be(\alpha, \alpha)$, $\alpha > 0$. Then

$$f(x|\alpha,\beta) = \begin{cases} \dfrac{x^{\alpha-1}(1-x)^{\alpha-1}}{B(\alpha,\alpha)}, & 0 < x < 1, \\ 0, & \text{otherwise}. \end{cases}$$

Clearly, $X \stackrel{d}{=} 1 - X \implies X - \frac{1}{2} \stackrel{d}{=} \frac{1}{2} - X$. Thus, if $X \sim Be(\alpha, \alpha)$. Then the distribuion of X is symmetric about $1/2$.

Theorem 13.16 (Relationship between Beta and Binomial Distribution)**.** *For $m, n \in \mathbb{N}$ and $x \in (0,1)$, let $X \sim Be(m,n)$ and $Y \sim Bin(m+n-1, x)$. Then $P(X \leq x) = P(Y \geq m)$, i.e.*

$$\frac{1}{B(m,n)} \int_0^x t^{m-1}(1-t)^{n-1}\mathrm{d}t = \sum_{j=m}^{m+n-1} \binom{m+n-1}{j} x^j (1-x)^{m+n-1-j}.$$

Proof. Fix $m, n \in \mathbb{N}$ and $x \in (0,1)$. Let

$$\begin{aligned}
I_{m,n} = LHS &= \frac{(m+n-1)!}{(m-1)!(n-1)!} \int_0^x t^{m-1}(1-t)^{n-1}\mathrm{d}t \\
&= \binom{m+n-1}{m} x^m (1-x)^{n-1} + \frac{(m+n-1)!}{m!(n-2)!} \int_0^x t^m (1-t)^{n-2}\mathrm{d}t \\
&= \binom{m+n-1}{m} x^m (1-x)^{n-1} + I_{m+1,n-1}.
\end{aligned}$$

Proceeding recursively give the result. $\square$

$$\begin{aligned}
\text{m.g.f. } M_X(t) = E(e^{tX}) &= \frac{1}{B(\alpha,\beta)} \int_0^1 e^{tx} x^{\alpha-1}(1-x)^{\beta-1}\mathrm{d}x \\
&= \frac{1}{B(\alpha,\beta)} \int_0^1 \left(\sum_{j=0}^{\infty} \frac{t^j x^j}{j!} \right) x^{\alpha-1}(1-x)^{\beta-1}\mathrm{d}x \\
&= \frac{1}{B(\alpha,\beta)} \sum_{j=0}^{\infty} \frac{t^j}{j!} \int_0^1 x^{j+\alpha-1}(1-x)^{\beta-1}\mathrm{d}x = \frac{1}{B(\alpha,\beta)} \sum_{j=0}^{\infty} \frac{B(\alpha+j,\beta)t^j}{j!}, \quad t \in \mathbb{R}.
\end{aligned}$$

Example 13.17. *Time (in hours) to finish a job follows beta distribution with mean $\frac{1}{3}$ hrs. and variance $\frac{2}{63}$ hrs. Find the probability that the job will be finished in 30 minutes.*

Solution: Define $X =$ time to finish job (in hours) $\sim Be(\alpha,\beta)$, say.

$E(X) = \frac{1}{3} \implies \frac{\alpha}{\alpha+\beta} = \frac{1}{3}$, $\mathrm{Var}(X) = \frac{2}{63} \implies \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)} = \frac{2}{63}$. This implies $\alpha = 2$ and $\beta = 4$. Thus, $X \sim Be(2,4)$. Required probability

$$P(X < \frac{1}{2}) = \frac{1}{B(2,4)} \int_0^{1/2} x(1-x)^3\mathrm{d}x = \frac{13}{16}.$$