7. Week 7

Remark 7.1 (Identically distributed RVs). Let X and Y be two RVs, possibly defined on different probability spaces.

(a) Recall from Remark 3.19 that their law/distribution may be the same and in this case, we have $F_X = F_Y$, i.e. $F_X(x) = F_Y(x), \forall x \in \mathbb{R}$. The statement '$X$ and Y are equal in law/distribution' is equivalent to 'X and Y are identically distributed'.

(b) Recall from Remark 3.20 that the DF uniquely identifies the law/distribution, i.e. if $F_X = F_Y$, then X and Y are identically distributed.

(c) Suppose X and Y are discrete RVs. Recall from Remark 4.13, the p.m.f. is uniquely determined by the DF and vice versa. In the case of discrete RVs, X and Y are identically distributed if and only if the p.m.f.s are equal (i.e., $f_X = f_Y$).

(d) Suppose X and Y are continuous RVs. Recall from Note 4.24 that the p.d.f.s in this case are uniquely identified upto sets of 'length 0'. We may refer to such an almost equal p.d.f. as a 'version of a p.d.f.'. Recall from Note 4.27, the p.d.f. is uniquely determined by the DF and vice versa. In the case of continuous RVs, X and Y are identically distributed if and only if the p.d.f.s are versions of each other. In other words, X and Y are identically distributed if and only if there exist versions f_X and f_Y of the p.d.f.s such that $f_X = f_Y$, i.e. $f_X(x) = f_Y(x), \forall x \in \mathbb{R}$.

(e) Suppose X and Y are identically distributed and let $h : \mathbb{R} \to \mathbb{R}$ be a function. Then we have that the RVs $h(X)$ and $h(Y)$ are identically distributed. In particular, $\mathbb{E}h(X) = \mathbb{E}h(Y)$, provided one of the expectations exists.

(f) Suppose X and Y are identically distributed. By (e), X^2 and Y^2 are identically distributed and $\mathbb{E}X^2 = \mathbb{E}Y^2$, provided one of the expectations exists. More generally, the n-th moments $\mathbb{E}X^n$ and $\mathbb{E}Y^n$ of X and Y are the same, provided they exist.

(g) There are examples where $\mathbb{E}X^n = \mathbb{E}Y^n, \forall n = 1, 2, \cdots$, but X and Y are not identically distributed. We may discuss such an example later in this course. Consequently, the moments do not uniquely identify the distribution. Under certain sufficient conditions on the moments, such as the Carleman's condition, it is however possible to uniquely identify the distribution. This is beyond the scope of this course.

(h) Suppose X and Y are identically distributed and suppose that the MGF M_X exists on $(-h, h)$ for some $h > 0$. By the above observation (e), the MGF M_Y exists and $M_X = M_Y$, i.e. $M_X(t) = M_Y(t), \forall t \in (-h, h)$.

(i) We now state a result without proof. Suppose the MGFs M_X and M_Y exist. If $M_X(t) = M_Y(t), \forall t \in (-h, h)$, then X and Y are identically distributed. Therefore, the MGF uniquely identifies the distribution.

Notation 7.2. We write $X \overset{d}{=} Y$ to denote that X and Y are identically distributed.

Example 7.3. If Y is an RV with the MGF $M_Y(t) = (1-t)^{-1}, \forall t \in (-1, 1)$, then by Example 6.47, we conclude that Y is a continuous RV with p.d.f.

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

Example 7.4. If X is a discrete RV with support S_X and p.m.f. f_X, then the MGF M_X is of the form

$$M_X(t) = \sum_{x \in S_X} e^{tx} f_X(x).$$

We can also make a converse statement. Since the MGF uniquely identifies a distribution, if an MGF is given by a sum of the above form, we can immediately identify the corresponding discrete RV with its support and p.m.f.. For example, if $M_X(t) = \frac{1}{2} + \frac{1}{3}e^t + \frac{1}{6}e^{-t}$, then X is discrete with the p.m.f.

$$f_X(x) := \begin{cases} \frac{1}{2}, & \text{if } x = 0, \\ \frac{1}{3}, & \text{if } x = 1, \\ \frac{1}{6}, & \text{if } x = -1, \\ 0, & \text{otherwise.} \end{cases}$$

Notation 7.5. We may refer to expectations of the form $\mathbb{E}e^{tX}$ as exponential moments of the RV X.

We now discuss a function similar to the MGF, arising through the Fourier transform. We first recall some notations from Complex Analysis.

Remark 7.6. Recall that $\exp(iy) = \cos(y) + i\sin(y), \forall y \in \mathbb{R}$. Moreover, $|\exp(iy)| = 1, \forall y \in \mathbb{R}$.

Let X be a discrete/continuous RV defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with DF F_X, p.m.f./p.d.f. f_X and support S_X.

Definition 7.7 (Characteristic Function)**.** The Characteristic function of X, denoted by $\Phi_X : \mathbb{R} \to \mathbb{C}$ is defined by $\Phi_X(t) := \mathbb{E}e^{itX} = \mathbb{E}\cos(tX) + i\,\mathbb{E}\sin(tX), \forall t \in \mathbb{R}$.

Remark 7.8. Since cos and sin are bounded continuous functions on $\mathbb{R}$, the random variables $\cos(tX)$ and $\sin(tX)$ are bounded for all $t \in \mathbb{R}$ and all RVs X. Using Proposition 6.22(d), we conclude that $\mathbb{E}\cos(tX)$ and $\mathbb{E}\sin(tX)$ exist for all $t \in \mathbb{R}$ and all RVs X. As such, the Characteristic function exists for all RVs X.

Note 7.9. If X is discrete/continuous with p.m.f./p.d.f. f_X, then following the definition of an expectation of an RV, we write

$$\begin{aligned}
\Phi_X(t) &= \mathbb{E}e^{itX} \\
&= \begin{cases} \sum\limits_{x \in S_X} e^{itx} f_X(x), & \text{for discrete } X, \\ \int_{-\infty}^{\infty} e^{itx} f_X(x)\,dx, & \text{for continuous } X. \end{cases} \\
&= \mathbb{E}\cos(tX) + i\,\mathbb{E}\sin(tX) \\
&= \begin{cases} \sum\limits_{x \in S_X} \cos(tx) f_X(x) + i \sum\limits_{x \in S_X} \sin(tx) f_X(x), & \text{for discrete } X, \\ \int_{-\infty}^{\infty} \cos(tx) f_X(x)\,dx + i \int_{-\infty}^{\infty} \sin(tx) f_X(x)\,dx, & \text{for continuous } X. \end{cases}
\end{aligned}$$

Note 7.10. For any RV X, $\Phi_X(0) = 1$.

Example 7.11. For $c \in \mathbb{R}$, consider the constant/degenerate RV X given by the p.m.f. (see Example 6.8)

$$f_X(x) = \begin{cases} 1, & \text{if } x = c \\ 0, & \text{otherwise.} \end{cases}$$

Here, the support is $S_X = \{c\}$ and $\Phi_X(t) = \mathbb{E}e^{itX} = \sum_{x \in S_X} e^{itx} f_X(x) = e^{itc} = \cos(tc) + i\sin(tc), \forall t \in \mathbb{R}$.

Note 7.12. Take constants $c, d \in \mathbb{R}$ with $c \neq 0$. Then, the RV $Y = cX + d$ is discrete/continuous, according to X being discrete/continuous and moreover,

$$\Phi_Y(t) = \mathbb{E}e^{it(cX+d)} = e^{itd}\Phi_X(ct)$$

exists for all $t \in \mathbb{R}$.

Note 7.13. We have $\Phi_{-X} = \bar{\Phi}_X$.

We state the next result without proof.

Theorem 7.14. *If $\mathbb{E}|X|^k < \infty$ for some positive integer k, then*

(a) $\Phi_X^{(r)}(u) = i^r \mathbb{E}\left(X^r e^{iuX}\right)$, *where* $M_X^{(r)}(0) = \left[\frac{d^r}{dt^r} M_X(t)\right]_{t=0}$ *is the r-th derivative of $M_X(t)$ at the point 0 for each $r \in \{1, 2, \cdots, k\}$.*

(b) In particular, $\mathbb{E}(X^r) = (-i)^r \Phi_X^{(r)}(0)$ *for each $r \in \{1, 2, \cdots, k\}$.*

(c) $\Phi_X^{(k)}(u) = \sum_{j=0}^{k} \frac{i^j}{j!}(\mathbb{E}X^j)u^j + o(|u|^k)$, *as $u \to 0$.*

Remark 7.15. Suppose X and Y are two RVs, possibly defined on different probability spaces. If $\Phi_X(t) = \Phi_Y(t), \forall t \in \mathbb{R}$, then X and Y are identically distributed. Therefore, the Characteristic function uniquely identifies the distribution.

Definition 7.16 (Symmetric Distribution)**.** An RV X is said to have a symmetric distribution about a point $\mu \in \mathbb{R}$ if $X - \mu \stackrel{d}{=} \mu - X$.

Proposition 7.17. *Let X be an RV which is symmetric about 0.*

(a) If X is discrete, then the p.m.f. f_X has the property that $f_X(x) = f_X(-x), \forall x \in \mathbb{R}$. Further, $\mathbb{E}X^n = 0, \forall n = 1, 3, 5, \cdots$, provided the moments exist.

(b) If X is continuous, then the p.d.f. f_X has the property that $f_X(x) = f_X(-x), \forall x \in \mathbb{R}$. Further, $\mathbb{E}X^n = 0, \forall n = 1, 3, 5, \cdots$, provided the moments exist.

Proof. We prove the statement when X is a continuous RV. The proof for the case when X is discrete is similar.

If X is symmetric about 0, then $X \stackrel{d}{=} -X$ and hence for any $x \in \mathbb{R}$, $F_X(x) = F_{-X}(x)$ and hence $F_X(x) = \mathbb{P}(X \leq x) = \mathbb{P}(-X \leq x) = \mathbb{P}(X \geq -x) = 1 - F_X(-x)$. This implies $f_X(x) = f_X(-x), \forall x \in \mathbb{R}$.

Assume that the moments in question exist. Then $\mathbb{E}X^n = \int_{-\infty}^{\infty} x^n f_X(x)\, dx = 0$, since the function $x \mapsto x^n f_X(x)$ is odd. $\square$

Remark 7.18. Let X be a continuous RV, which is symmetric about $\mu \in \mathbb{R}$. As argued in the above proposition, we have for all $x \in \mathbb{R}$

$$F_X(\mu + x) = \mathbb{P}(X \leq \mu + x) = \mathbb{P}(X - \mu \leq x) = \mathbb{P}(\mu - X \leq x) = \mathbb{P}(X - \mu \geq -x) = 1 - F_X(\mu - x)$$

and hence $f_X(\mu + x) = f_X(\mu - x), \forall x$. Conversely, given a continuous RV X such that $f_X(\mu + x) = f_X(\mu - x), \forall x$ for some $\mu \in \mathbb{R}$, we have $F_{X-\mu} = F_{\mu - X}$ and hence X is symmetric about μ.

We now look at some special examples of discrete RVs.

Example 7.19 (Degenerate RV)**.** We have already mentioned this example earlier in Example 6.8. Fix $c \in \mathbb{R}$. Say that X is degenerate at c if its distribution is given by the p.m.f.

$$f_X(x) = \mathbb{P}(X = x) = \begin{cases} 1, & \text{if } x = c \\ 0, & \text{otherwise.} \end{cases}$$

This is a discrete RV with support $S_X = \{c\}$. As computed earlier, $\mathbb{E}X = c$. We also have $\mathbb{E}X^n = c^n, \forall n \geq 1, M_X(t) = e^{tc}, \forall t \in \mathbb{R}$ and $\Phi_X(t) = e^{itc}, \forall t \in \mathbb{R}$. Note that $Var(X) = \mathbb{E}X^2 - (\mathbb{E}X)^2 = 0$.

Remark 7.20 (Bernoulli Trial). Suppose that a random experiment has exactly two outcomes, identified as a 'success' and a 'failure'. For example, while tossing a coin, we may think of obtaining a head as a success and a tail as a failure. Here, the sample space is $\Omega = \{Success, Failure\}$. A single trial of such an experiment is referred to as a Bernoulli trial. In this case, $Probability(\{Success\}) = 1 - Probability(\{Failure\})$. If we define an RV $X : \Omega \to \mathbb{R}$ by $X(Success) = 1$ and $X(Failure) = 0$, then X is a discrete RV with p.m.f.

$$f_X(x) = \mathbb{P}(X = x) = \begin{cases} 1 - Probability(\{Success\}), & \text{if } x = 0, \\ Probability(\{Success\}), & \text{if } x = 1, \\ 0, & \text{otherwise.} \end{cases}$$

If $Probability(\{Success\}) = 0$ or $Probability(\{Failure\}) = 0$, then X is degenerate at 0 or 1, respectively. The case when $Probability(\{Success\}) \in (0, 1)$ is therefore of interest.

Example 7.21 (Bernoulli(p) RV)**.** Let $p \in (0, 1)$. An RV X is said to follow Bernoulli(p) distribution or equivalently, X is a Bernoulli(p) RV if its distribution is given by the p.m.f.

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

In relation with the Bernoulli trial described above, p may be treated as the probability of success. Here, $\mathbb{E}X = p$, $\mathbb{E}X^2 = p$, $Var(X) = \mathbb{E}X^2 - (\mathbb{E}X)^2 = p - p^2 = p(1-p)$, $M_X(t) = 1 - p + pe^t, t \in \mathbb{R}$ and $\Phi_X(t) = 1 - p + pe^{it}, t \in \mathbb{R}$. By standard arguments, we can establish the existence of these moments.

Notation 7.22. We may write $X \sim Bernoulli(p)$ to mean that X is a Bernoulli(p) RV. Similar notations shall be used for other RVs and their distributions.

Example 7.23 (Binomial(n, p) RV)**.** Fix a positive integer n and let $p \in (0, 1)$. By the Binomial theorem, we have

$$1 = [p + (1-p)]^n = \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k}$$

and hence the function $f : \mathbb{R} \to [0,1]$ given by

$$f(x) = \begin{cases} \binom{n}{x} p^x (1-p)^{n-x}, & \text{if } x \in \{0, 1, \cdots, n\}, \\ 0, & \text{otherwise.} \end{cases}$$

is a p.m.f.. An RV X is said to follow Binomial(n,p) distribution or equivalently, X is a Binomial(n,p) RV if its distribution is given by the above p.m.f.. Here,

$$\mathbb{E}X = \sum_{k=0}^{n} k \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} = np \sum_{k=1}^{n} \binom{n-1}{k-1} p^{k-1} (1-p)^{n-k} = np\, [p + (1-p)]^{n-1} = np,$$

and

$$\mathbb{E}X(X-1) = \sum_{k=0}^{n} k(k-1) \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} = n(n-1)p^2 \sum_{k=2}^{n} \binom{n-2}{k-2} p^{k-2} (1-p)^{n-k} = n(n-1)p^2.$$

Then $\mathbb{E}X^2 = \mathbb{E}X(X-1) + \mathbb{E}X = n(n-1)p^2 + np$ and $Var(X) = \mathbb{E}X^2 - (\mathbb{E}X)^2 = n(n-1)p^2 + np - n^2p^2 = np(1-p)$. Also

$$M_X(t) = \mathbb{E}e^{tX} = \sum_{k=0}^{n} e^{tk} \binom{n}{k} p^k (1-p)^{n-k} = (1 - p + pe^t)^n, \forall t \in \mathbb{R}.$$

By standard arguments, we can establish the existence of these moments. We also have $\Phi_X(t) = \mathbb{E}e^{itX} = \sum_{k=0}^{n} e^{itk} \binom{n}{k} p^k (1-p)^{n-k} = (1 - p + pe^{it})^n, \forall t \in \mathbb{R}$.

Note 7.24. Observe that Binomial$(1,p)$ distribution is the same as Bernoulli(p) distribution. We shall explore the connection between Binomial and Bernoulli distributions later in the course.

Remark 7.25 (Factorial moments). In the computation for $\mathbb{E}X^2$ for $X \sim Binomial(n,p)$, we first computed $\mathbb{E}X(X-1)$, which is easy to compute. It turns out that expectations of the form $\mathbb{E}X(X-1)$, $\mathbb{E}X(X-1)(X-2)$ etc. are often easy to compute for integer valued RVs X. We refer to such expectations as factorial moments of X.

Remark 7.26 (Symmetry of Binomial$(n, \frac{1}{2})$ distribution). Let $X \sim Binomial(n,p)$ and let $Y := n - X$. Since $M_X(t) = (1 - p + pe^t)^n, \forall t \in \mathbb{R}$, we have

$$M_Y(t) = \mathbb{E}e^{tY} = \mathbb{E}e^{t(n-X)} = e^{-nt} M_X(-t) = e^{nt}(1 - p + pe^{-t})^n = (p + (1-p)e^t)^n.$$

Since MGFs determine the distribution, we conclude that $Y \sim Binomial(n, 1-p)$. In particular, if $p = \frac{1}{2}$, then $Y = n - X \stackrel{d}{=} X \sim Binomial(n, \frac{1}{2})$. Rewriting the relation, we get $\frac{n}{2} - X \stackrel{d}{=} X - \frac{n}{2}$. Therefore, $X \sim Binomial(n, \frac{1}{2})$ is symmetric about $\frac{n}{2}$.

We now look at more examples of discrete RVs. Later in the course, we shall discuss their motivation through various random experiments.

Example 7.27 (Uniform RVs with support on a finite set)**.** Consider a discrete RV X with support $S_X = \{x_1, x_2, \cdots, x_n\}$ and p.m.f. $f_X : \mathbb{R} \to [0,1]$ given by

$$f_X(x) = \begin{cases} \frac{1}{n}, & \text{if } x \in S_X, \\ 0, & \text{otherwise.} \end{cases}$$

We had considered the case $S_X = \{1, 2, \cdots, 6\}$ in Example 6.9 and computed the expectation. In the general setting, we have

$$\mathbb{E}X = \frac{1}{n} \sum_{x \in S_X} x, \quad \mathbb{E}X^2 = \frac{1}{n} \sum_{x \in S_X} x^2, \quad M_X(t) = \mathbb{E}e^{tX} = \frac{1}{n} \sum_{x \in S_X} e^{tx}, \forall t \in \mathbb{R}$$

and hence $Var(X)$ can be computed by the formula $\mathbb{E}X^2 - (\mathbb{E}X)^2$. By standard arguments, we can establish the existence of these moments.

Example 7.28 (Poisson (λ) RV)**.** Fix $\lambda > 0$. Note that $e^\lambda = \sum_{k=0}^\infty \frac{\lambda^k}{k!}$ and hence the function $f : \mathbb{R} \to [0,1]$ given by

$$f(x) = \begin{cases} e^{-\lambda} \frac{\lambda^x}{x!}, & \text{if } x \in \{0, 1, 2, \cdots\}, \\ 0, & \text{otherwise.} \end{cases}$$

is a p.m.f.. An RV X is said to follow Poisson(λ) distribution or equivalently, X is a Poisson(λ) RV if its distribution is given by the above p.m.f.. Recall that we have already computed the following $\mathbb{E}X = \lambda, Var(X) = \lambda$ and $M_X(t) = e^{\lambda\left(e^t - 1\right)}, \forall t \in \mathbb{R}$ in Example 6.46. As done for the case of Binomial(n, p) RVs, we can compute factorial moments. For example,

$$\mathbb{E}X(X-1) = \sum_{k=0}^\infty k(k-1) e^{-\lambda} \frac{\lambda^k}{k!} = \lambda^2 \sum_{k=2}^\infty e^{-\lambda} \frac{\lambda^{k-2}}{(k-2)!} = \lambda^2.$$

In fact, $\mathbb{E}X(X-1)\cdots(X-(n-1)) = \lambda^n$ for all $n \geq 1$. Here, the Characteristic function is $\Phi_X(t) = e^{\lambda\left(e^{it}-1\right)}$, $\forall t \in \mathbb{R}$.

Example 7.29 (Geometric (p) RV)**.** Fix $p \in (0,1)$. Note that $\sum_{k=0}^\infty p(1-p)^k = 1$ and hence the function $f : \mathbb{R} \to [0,1]$ given by

$$f(x) = \begin{cases} p(1-p)^x, & \text{if } x \in \{0,1,2,\cdots\}, \\ 0, & \text{otherwise.} \end{cases}$$

is a p.m.f.. An RV X is said to follow Geometric(p) distribution or equivalently, X is a Geometric(p) RV if its distribution is given by the above p.m.f.. Let us compute the MGF. Here,

$$M_X(t) = \mathbb{E}e^{tX} = \sum_{k=0}^\infty e^{tk} p(1-p)^k = \frac{p}{1-(1-p)e^t},$$

for all t such that $0 < (1-p)e^t < 1$ or equivalently, $t < \ln\left(\frac{1}{1-p}\right)$. Looking at the derivatives of M_X and evaluating at $t = 0$, we have $\mathbb{E}X = \frac{1-p}{p}$ and $Var(X) = \frac{1-p}{p^2}$. The Characteristic function is given by $\Phi_X(t) = \mathbb{E}e^{itX} = \frac{p}{1-(1-p)e^{it}}, t \in \mathbb{R}$.

We now look at special examples of continuous RVs.

Example 7.30 (Uniform(a,b) RV)**.** Fix $a, b \in \mathbb{R}$ with $a < b$. An RV X is said to follow Uniform(a,b) distribution or equivalently, X is a Uniform(a,b) RV if its distribution is given by the p.d.f.

$$f_X(x) = \begin{cases} \frac{1}{b-a}, & \text{if } x \in (a,b), \\ 0, & \text{otherwise.} \end{cases}$$

We had considered the case $a = 0, b = 1$ in Example 6.13 and computed the expectation. In the general setting, we have

$$\mathbb{E}X = \int_a^b \frac{x}{b-a}\,dx = \frac{b^2-a^2}{2(b-a)} = \frac{a+b}{2}, \quad \mathbb{E}X^2 = \int_a^b \frac{x^2}{b-a}\,dx = \frac{b^3-a^3}{3(b-a)} = \frac{a^2+ab+b^2}{3}$$

and hence $Var(X)$ can be computed by the formula $\mathbb{E}X^2 - (\mathbb{E}X)^2$. The MGF is given by

$$\mathbb{E}e^{tX} = \int_a^b \frac{e^{tx}}{b-a}\, dx = \begin{cases} \frac{e^{tb}-e^{ta}}{t(b-a)}, \text{ if } t \neq 0, \\ 1, \text{ if } t = 0. \end{cases}$$

By standard arguments, we can establish the existence of these moments. Further, observe that $f_X(\frac{a+b}{2} - x) = f_X(\frac{a+b}{2} + x), \forall x \in \mathbb{R}$. Using Remark 7.18, we conclude that X is symmetric about its mean.

Example 7.31 (Cauchy(μ, θ) RV)**.** Let $\theta > 0$ and $\mu \in \mathbb{R}$. An RV X is said to follow Cauchy(μ, θ) distribution if its distribution is given by the p.d.f.

$$f_X(x) = \frac{\theta}{\pi} \frac{1}{\theta^2 + (x-\mu)^2}, \forall x \in \mathbb{R}.$$

The fact that f_X is a p.d.f. is easy to check. Set $y = \frac{x-\mu}{\theta}$ and observe that

$$\int_{-\infty}^{\infty} f_X(x)\, dx = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{1}{1+y^2}\, dy = \frac{2}{\pi}\int_0^{\infty} \frac{1}{1+y^2}\, dy = \frac{2}{\pi}\tan^{-1}(y)\mid_0^{\infty} = 1.$$

We have already considered the case $\mu = 0, \theta = 1$ in Example 6.15 and Example 6.48, where we have seen that $\mathbb{E}X$ and the MGF do not exist for this distribution. In the general setting, note that $\frac{X-\mu}{\theta} \sim Cauchy(0,1)$ and by a similar argument, we can show that $\mathbb{E}X$ and MGF do not exist. Moreover, $f_X(\mu + x) = f_X(\mu - x), \forall x \in \mathbb{R}$ and using Remark 7.18, we conclude that X is symmetric about μ. The Characteristic function for X is given by

$$\Phi_X(u) = \exp(iu\mu - \theta|u|), \forall u \in \mathbb{R}.$$

Example 7.32 (Exponential(λ) RV)**.** Let $\lambda > 0$. Note that $\int_0^\infty \exp(-\frac{x}{\lambda})\, dx = \lambda$ and hence the function $f : \mathbb{R} \to [0, \infty)$ given by

$$f(x) = \begin{cases} \frac{1}{\lambda}\exp(-\frac{x}{\lambda}), \text{ if } x > 0, \\ 0, \text{ otherwise.} \end{cases}$$

is a p.d.f.. An RV X is said to follow Exponential(λ) distribution or equivalently, X is an Exponential(λ) RV if its distribution is given by the above p.d.f.. We have already considered the case $\lambda = 1$ in Example 6.47, where we computed the moments and the MGF. Following similar arguments, in the general setting we have

$$\mathbb{E}X^n = \lambda^n n!, \quad Var(X) = \lambda^2, \quad M_X(t) = (1 - \lambda t)^{-1}, \forall t < \frac{1}{\lambda}.$$

By standard arguments, we can establish the existence of these moments.

Definition 7.33 (Gamma function)**.** Recall that the integral $\int_0^\infty x^{\alpha-1}e^{-x}\,dx$ exists if and only if $\alpha > 0$. On $(0,\infty)$, consider the function $\alpha \mapsto \int_0^\infty x^{\alpha-1}e^{-x}\,dx$. It is called the Gamma function and the value at any $\alpha > 0$ is denoted by $\Gamma(\alpha)$.

Remark 7.34. We recall some important properties of the Gamma function.

(a) For $\alpha > 0$, we have $\Gamma(\alpha) > 0$.
(b) $\Gamma(\alpha) = (\alpha - 1)\Gamma(\alpha - 1)$, if $\alpha > 1$.
(c) $\Gamma(1) = \int_0^\infty e^{-x}\,dx = 1$ and hence using (b), $\Gamma(n) = (n-1)!$ for all positive integers n.
(d) $\Gamma(\frac{1}{2}) = \int_0^\infty \frac{1}{\sqrt{x}}e^{-x}\,dx = \sqrt{\pi}$. Putting $x = \frac{y^2}{2}$, this relation may be rewritten as

$$\Gamma\left(\frac{1}{2}\right) = \sqrt{2}\int_0^\infty \exp\left(-\frac{y^2}{2}\right)\,dy = \sqrt{\pi}.$$

(e) Fix $\beta > 0$. Putting $x = \frac{y}{\beta}$, in the integral for $\Gamma(\alpha) = \int_0^\infty x^{\alpha-1}e^{-x}\,dx$, we get $\Gamma(\alpha) = \int_0^\infty y^{\alpha-1}\beta^{-\alpha}\exp(-\frac{y}{\beta})\,dy$.

Example 7.35 (Gamma(α,β) RV)**.** Fix $\alpha > 0, \beta > 0$. By the properties of the Gamma function described above, the function $f : \mathbb{R} \to [0,\infty)$ defined by

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

is a p.d.f.. An RV X is said to follow Gamma(α,β) distribution or equivalently, X is a Gamma(α,β) RV if its distribution is given by the above p.d.f.. Note that for $\alpha = 1$, we get back the p.d.f. for an Exponential(β) RV (see Example 7.32), i.e. Gamma$(1,\beta)$ distribution is the same as

Exponential(β) distribution. For general $\alpha > 0, \beta > 0$, we have

$$\mathbb{E}X = \alpha\beta, \quad Var(X) = \alpha\beta^2, \quad M_X(t) = (1-\beta t)^{-\alpha}, \forall t < \frac{1}{\beta}.$$

By standard arguments, we can establish the existence of these moments.

Example 7.36 (Normal(μ, σ^2) RV)**.** Fix $\mu \in \mathbb{R}, \sigma > 0$. Note that $\Gamma(\frac{1}{2}) = \int_0^\infty \frac{1}{\sqrt{t}} e^{-t}\,dt = \sqrt{\pi}$ (see Remark 7.34). Putting $t = \frac{y^2}{2}$ and after suitable manipulation, we have $\frac{1}{\sqrt{2\pi}} \int_{-\infty}^\infty \exp\left(-\frac{y^2}{2}\right)\,dy = 1$. Putting $y = \frac{1}{\sigma}(x-\mu)$ (equivalently, $x = \sigma y + \mu$), we have

$$\frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^\infty \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)\,dx = 1.$$

Therefore, the function $f : \mathbb{R} \to [0,\infty)$ defined by

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right), \forall x \in \mathbb{R}$$

is a p.d.f.. An RV X is said to follow Normal(μ, σ^2) distribution or equivalently, X is a Normal(μ, σ^2) RV, denoted by $X \sim N(\mu, \sigma^2)$ if its distribution is given by the above p.d.f.. If $X \sim N(\mu, \sigma^2)$, from our above discussion we conclude that $Y = \frac{X-\mu}{\sigma} \sim N(0,1)$. Now,

$$\begin{aligned} M_Y(t) = \mathbb{E}e^{tY} &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^\infty e^{ty} \exp\left(-\frac{y^2}{2}\right)\,dy \\ &= \exp\left(\frac{t^2}{2}\right) \frac{1}{\sqrt{2\pi}} \int_{-\infty}^\infty \exp\left(-\frac{(y-t)^2}{2}\right)\,dy \\ &= \exp\left(\frac{t^2}{2}\right), \forall t \in \mathbb{R}. \end{aligned}$$

In particular, $\psi_Y(t) = \ln M_Y(t) = \frac{t^2}{2}, \forall t \in \mathbb{R}$ with $\psi'(t) = t, \psi''(t) = 1, \forall t \in \mathbb{R}$. Evaluating at $t = 0$, by Proposition 6.45 we conclude that $\mathbb{E}Y = 0$ and $Var(Y) = 1$. But $X = \sigma Y + \mu$ and hence $\mathbb{E}X = \mu, Var(X) = \sigma^2$. This yields the interpretation of the parameters μ and σ in the distribution of X. Further, $M_X(t) = \mathbb{E}e^{tX} = \mathbb{E}e^{t(\sigma Y + \mu)} = e^{\mu t} M_Y(\sigma t) = \exp(\mu t + \frac{1}{2}\sigma^2 t^2), \forall t \in \mathbb{R}$. The Characteristic function for X is given by

$$\Phi_X(u) = \exp(iu\mu - \frac{1}{2}\sigma^2 u^2), \forall u \in \mathbb{R}.$$

Definition 7.37 (Standard Normal RV)**.** We say X is a Standard Normal RV if $X \sim N(0, 1)$, i.e. $\mathbb{E}X = 0$ and $Var(X) = 1$.

Notation 7.38. Normal RVs are also referred to as Gaussian RVs and Normal distribution as Gaussian distribution.

Remark 7.39 (Symmetry of Gaussian Distribution)*.* If $X \sim N(\mu, \sigma^2)$, note that $f_X(\mu + x) = f_X(\mu - x), \forall x \in \mathbb{R}$ and using Remark 7.18, we conclude that X is symmetric about its mean μ.