Note 1.191. If X is discrete with p.m.f. f_X such that $\mathbb{E}X$ exists, then $\mathbb{E}|X| = \sum_{x \in S_X} |x| f_X(x) < \infty$. Similarly, if X is continuous with p.d.f. f_X such that $\mathbb{E}X$ exists, then $\mathbb{E}|X| = \int_{-\infty}^{\infty} |x| f_X(x)\, dx < \infty$. Therefore $\mathbb{E}X$ exists if and only if $\mathbb{E}|X| < \infty$. In other words, $\mathbb{E}X$ is finite if and only if $\mathbb{E}|X|$ is finite.

Note 1.192. Fix $a, b \in \mathbb{R}$ with $a \neq 0$. Let X be a discrete/continuous RV with p.m.f./p.d.f. f_X such that $\mathbb{E}X$ exists. Then $Y = aX + b$ is also a discrete/continuous RV. If X is discrete, then

$$\sum_{x \in S_X} |ax + b| f_X(x) \leq |a| \sum_{x \in S_X} |x| f_X(x) + |b| \sum_{x \in S_X} f_X(x) = |a| \mathbb{E}|X| + |b| < \infty$$

and hence $\mathbb{E}(aX + b)$ exists and equals

$$\mathbb{E}(aX + b) = \sum_{x \in S_X} (ax + b) f_X(x) = a \sum_{x \in S_X} x f_X(x) + b \sum_{x \in S_X} f_X(x) = a\, \mathbb{E}X + b.$$

If X is continuous, a similar argument shows $\mathbb{E}(aX + b) = a\, \mathbb{E}X + b$.

Using arguments similar to the above observations, we obtain the next result. We skip the details for brevity.

Proposition 1.193. *Let X be a discrete/continuous RV with p.m.f./p.d.f. f_X.*

(a) *Let $h_i : \mathbb{R} \to \mathbb{R}$ be functions and let $a_i \in \mathbb{R}$ for $i = 1, 2, \cdots, n$. Then*

$$\mathbb{E}\left(\sum_{i=1}^{n} a_i h_i(X)\right) = \sum_{i=1}^{n} a_i\, \mathbb{E}h_i(X),$$

provided all the expectations above exist.

(b) *Let $h_1, h_2 : \mathbb{R} \to \mathbb{R}$ be functions such that $h_1(x) \leq h_2(x), \forall x \in S_X$, where S_X denotes the support of X. Then,*

$$\mathbb{E}h_1(X) \leq \mathbb{E}h_2(X),$$

provided all the expectations above exist.

(c) *Take $h_1(x) := -|x|, h_2(x) := x, h_3(x) := |x|, \forall x \in \mathbb{R}$. If $\mathbb{E}X$ exists, then*

$$-\mathbb{E}|X| \leq \mathbb{E}X \leq \mathbb{E}|X|,$$

i.e. $|\mathbb{E}X| \leq \mathbb{E}|X|$.

(d) *If* $\mathbb{P}(a \leq X \leq b) = 1$ *for some* $a, b \in \mathbb{R}$*, then* $\mathbb{E}X$ *exists and* $a \leq \mathbb{E}X \leq b$.

Note 1.194. Given an RV X, by choosing different functions $h : \mathbb{R} \to \mathbb{R}$, we obtain several quantities of interest of the form $\mathbb{E}h(X)$.

Definition 1.195 (Moments)**.** The quantity $\mu_r' := \mathbb{E}[X^r]$, if it exists, is called the r-th moment of RV X for $r > 0$.

Definition 1.196 (Absolute Moments)**.** The quantity $\mathbb{E}[|X|^r]$, if it exists, is called the r-th absolute moment of RV X for $r > 0$.

Definition 1.197 (Moments about a point)**.** Let $c \in \mathbb{R}$. The quantity $\mathbb{E}[(X - c)^r]$, if it exists, is called the r-th moment of RV X about c for $r > 0$.

Definition 1.198 (Absolute Moments about a point)**.** Let $c \in \mathbb{R}$. The quantity $\mathbb{E}[|X - c|^r]$, if it exists, is called the r-th absolute moment of RV X about c for $r > 0$.

Note 1.199. It is clear from the definitions above that the usual moments and absolute moments are moments and absolute moments about origin, respectively.

Proposition 1.200. *Let X be a discrete/continuous RV such that $\mathbb{E}|X|^r < \infty$ for some $r > 0$. Then $\mathbb{E}|X|^s < \infty$ for all $0 < s < r$.*

Proof. Observe that for all $x \in \mathbb{R}$, we have $|x|^s \leq \max\{|x|^r, 1\} \leq |x|^r + 1$ and hence

$$\mathbb{E}|X|^s \leq \mathbb{E}|X|^r + 1 < \infty.$$

$\square$

Remark 1.201. Suppose that the m-th moment $\mathbb{E}X^m$ of X exists for some positive integer m. Then we have $\mathbb{E}|X|^m < \infty$ (see Note 1.191). By Proposition 1.200, we have $\mathbb{E}|X|^n < \infty$ for all positive integers $n \leq m$ and hence the n-th moment $\mathbb{E}X^n$ exists for X. In particular, the existence of the second moment $\mathbb{E}X^2$ implies the existence of the first moment $\mathbb{E}X$, which is the expectation of X.

Definition 1.202 (Central Moments)**.** Let X be an RV such that $\mu_1' = \mathbb{E}X$ exists. The quantity $\mu_r := \mathbb{E}[(X-\mu_1')^r]$, if it exists, is called the r-th moment of RV X about the mean or r-th central moment of X for $r > 0$.

Definition 1.203 (Variance)**.** The second central moment μ_2 of an RV X, if it exists, is called the variance of X and denoted by $Var(X)$. Note that $Var(X) = \mu_2 = \mathbb{E}\left[(X-\mu_1')^2\right]$.

Remark 1.204. The following are some simple observations about the variance of an RV X.

(a) We have

$$Var(X) = \mathbb{E}\left[(X-\mu_1')^2\right] = \mathbb{E}[X^2+(\mu_1')^2-2\mu_1'X] = \mu_2'-2(\mu_1')^2+(\mu_1')^2 = \mu_2'-(\mu_1')^2 = \mathbb{E}X^2-(\mathbb{E}X)^2.$$

(b) Since the RV $(X-\mu_1')^2$ takes non-negative values, we have $Var(X) = \mathbb{E}(X-\mu_1')^2 \geq 0$.

(c) We have $(\mathbb{E}X)^2 \leq \mathbb{E}X^2$.

(d) $Var(X) = 0$ if and only if $\mathbb{P}(X = \mu_1') = 1$. (see problem set 5).

(e) For any $a, b \in \mathbb{R}$, we have $Var(aX+b) = a^2 Var(X)$.

(f) Let $Var(X) > 0$. Then $Y := \frac{X-\mathbb{E}X}{\sqrt{Var(X)}}$ has the property that $\mathbb{E}Y = 0$ and $Var(Y) = 1$.

Definition 1.205 (Standard Deviation)**.** The quantity $\sigma(X) = \sqrt{Var(X)}$ is defined to be the standard deviation of X.

Example 1.206. Let X be a discrete RV with p.m.f.

$$f_X(x) := \begin{cases} \frac{1}{6}, \forall x \in \{1,2,3,4,5,6\} \\ 0, \text{ otherwise.} \end{cases}$$

Here, existence of $\mu_1' = \mathbb{E}X$ and $\mu_2' = \mathbb{E}X^2$ can be established by standard calculations. Moreover,

$$\mathbb{E}X = \sum_{x \in S_X} x f_X(x) = \frac{1}{6}(1+2+3+4+5+6) = \frac{7}{2}$$

and

$$\mathbb{E}X^2 = \sum_{x \in S_X} x^2 f_X(x) = \frac{1}{6}(1^2+2^2+3^2+4^2+5^2+6^2) = \frac{91}{6}.$$

Variance can now be computed using the relation $Var(X) = \mathbb{E}X^2 - (\mathbb{E}X)^2$.

Example 1.207. In Example 1.184, we had shown $\mathbb{E}X = \frac{1}{2}$, where X is a continuous RV with the p.d.f.

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

Now, $\mathbb{E}X^2 = \int_{-\infty}^{\infty} x^2 f_X(x)\,dx = \int_0^1 x^2\,dx = \frac{1}{3}$. Then $Var(X) = \mathbb{E}X^2 - (\mathbb{E}X)^2 = \frac{1}{3} - \frac{1}{4} = \frac{1}{12}$.

Note 1.208. We are familiar with the Laplace transform of a given real-valued function defined on $\mathbb{R}$. We also know that under certain conditions, the Laplace transform of a function determines the function almost uniquely. In probability theory, the Laplace transform of a p.m.f./p.d.f. of a random variable X plays an important role.

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 1.209 (Moment Generating Function (MGF))**.** We say that the moment generating function (MGF) of X exists, denoted by M_X and equals $M_X(t) := \mathbb{E}e^{tX}$, provided $\mathbb{E}e^{tX}$ exists for all $t \in (-h, h)$, for some $h > 0$.

Note 1.210. Observe that $e^x > 0, \forall x \in \mathbb{R}$.

Note 1.211. 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

$$M_X(t) = \mathbb{E}e^{tX} = \begin{cases} \sum\limits_{x \in S_X} e^{tx} f_X(x), & \text{if } \sum\limits_{x \in S_X} e^{tx} f_X(x) < \infty \text{ for discrete } X, \forall t \in (-h,h) \text{ for some } h > 0 \\ \int_{-\infty}^{\infty} e^{tx} f_X(x)\,dx, & \text{if } \int_{-\infty}^{\infty} e^{tx} f_X(x)\,dx < \infty \text{ for continuous } X, t \in (-h,h) \text{ for some } h > 0. \end{cases}$$

In this case, we shall say that the MGF M_X exists on $(-h, h)$.

Remark 1.212.

(a) $M_X(0) = 1$ and hence $A := \left\{t \in \mathbb{R} : \mathbb{E}[e^{tX}] \text{ is finite}\right\} \neq \emptyset$.

(b) $M_X(t) > 0\ \forall t \in A$, with A as above.

(c) For $c \in \mathbb{R}$, consider the constant/degenerate RV X given by the p.m.f. (see Example 1.179)

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

Here, the support is $S_X = \{c\}$ and $M_X(t) = \mathbb{E}e^{tX} = \sum_{x \in S_X} e^{tx} f_X(x) = e^{tc}$ exists for all $t \in \mathbb{R}$.

(d) Suppose the MGF M_X exists on $(-h, h)$. 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,

$$M_Y(t) = \mathbb{E}e^{t(cX+d)} = e^{td} M_X(ct)$$

exists for all $t \in (-\frac{h}{|c|}, \frac{h}{|c|})$.

Note 1.213. The MGF can be used to compute the moments of an RV and this is the motivation behind the term 'Moment Generating Function'. This result is stated below. We skip the proof for brevity.

Theorem 1.214. *Let X be an RV with MGF M_X which exists on $(-h, h)$ for some $h > 0$. Then, we have the following results.*

(a) $\mu_r' = \mathbb{E}[X^r]$ is finite for each $r \in \{1, 2, \ldots\}$.

(b) $\mu_r' = \mathbb{E}[X^r] = M_X^{(r)}(0)$, 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 $r \in \{1, 2, \ldots\}$.

(c) M_X has the following Maclaurin's series expansion around $t = 0$ of the following form $M_X(t) = \sum_{r=0}^{\infty} \frac{t^r}{r!} \mu_r'$ with $t \in (-h, h)$.

Proposition 1.215. *Continue with the notations and assumptions of Theorem 1.214 and define $\psi_X : (-h, h) \to \mathbb{R}$ by $\psi_X(t) := \ln M_X(t), t \in (-h, h)$. Then*

$$\mu_1' = \mathbb{E}[X] = \psi_X^{(1)}(0) \quad \text{and} \quad \mu_2 = Var(X) = \psi_X^{(2)}(0),$$

where $\psi_X^{(r)}$ denotes the r-th $(r \in \{1, 2\})$ derivative of ψ_X.

Proof. We have, for $t \in (-h, h)$

$$\psi_X^{(1)}(t) = \frac{M_X^{(1)}(t)}{M_X(t)} \quad \text{and} \quad \psi_X^{(2)}(t) = \frac{M_X(t)M_X^{(2)}(t) - \left(M_X^{(1)}(t)\right)^2}{(M_X(t))^2}.$$

Evaluating the above equalities at $t = 0$ give the required results. □

Example 1.216. Let X be a discrete RV with p.m.f.

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

where $\lambda > 0$. We have

$$M_X(t) = \mathbb{E}\left[e^{tX}\right] = \sum_{x=0}^{\infty} e^{tx}\frac{e^{-\lambda}\lambda^x}{x!} = e^{-\lambda}\sum_{x=0}^{\infty}\frac{(\lambda e^t)^x}{x!} = e^{-\lambda}e^{\lambda e^t} = e^{\lambda\left(e^t-1\right)} \ \forall\ t \in \mathbb{R}$$

since $A = \left\{t \in \mathbb{R} : \mathbb{E}\left(e^{tX}\right) < \infty\right\} = \mathbb{R}$. Now,

$$M_X^{(1)}(t) = \lambda e^t e^{\lambda\left(e^t-1\right)} \quad \text{and} \quad M_X^{(2)}(t) = \lambda e^t e^{\lambda\left(e^t-1\right)}\left(1 + \lambda e^t\right) \ \forall\ t \in \mathbb{R}.$$

Then,

$$\mu_1' = \mathbb{E}(X) = M_X^{(1)}(0) = \lambda,\ \mu_2' = \mathbb{E}(X^2) = M_X^{(2)}(0) = \lambda(1+\lambda),\ Var(X) = \mu_2 = \mu_2' - (\mu_1')^2 = \lambda.$$

Again, for $t \in \mathbb{R}$, $\psi_X(t) = \ln(M_X(t)) = \lambda\left(e^t - 1\right)$, which yields $\psi_X^{(1)}(t) = \psi_X^{(2)}(t) = \lambda e^t, \forall t \in \mathbb{R}$. Then, $\mu_1' = \mathbb{E}(X) = \lambda, \mu_2 = Var(X) = \lambda$. Higher order moments can be calculated by looking at higher order derivatives of M_X.

Example 1.217. Let X be a continuous RV with p.d.f.

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

We have

$$M_X(t) = \mathbb{E}\left(e^{tX}\right) = \int_0^{\infty} e^{tx}e^{-x}\ \mathrm{d}x = \int_0^{\infty} e^{-(1-t)x}\ \mathrm{d}x = (1-t)^{-1} < \infty,\ \text{if } t < 1.$$

In particular, M_X exists on $(-1,1)$ and $A = \left\{t \in \mathbb{R} : \mathbb{E}\left(e^{tX}\right) < \infty\right\} = (-\infty, 1) \supset (-1,1)$. Now,

$$M_X^{(1)}(t) = (1-t)^{-2} \quad \text{and} \quad M_X^{(2)}(t) = 2(1-t)^{-3}, t < 1.$$

Then,

$$\mu_1' = \mathbb{E}(X) = M_X^{(1)}(0) = 1,\ \mu_2' = \mathbb{E}(X^2) = M_X^{(2)}(0) = 2,\ Var(X) = \mu_2 = \mu_2' - (\mu_1')^2 = 1.$$

Again, for $t < 1$, $\psi_X(t) = \ln\left(M_X(t)\right) = -\ln(1-t)$, which yields $\psi_X^{(1)}(t) = \frac{1}{1-t}, \psi_X^{(2)}(t) = \frac{1}{(1-t)^2}, \forall t < 1$. Then, $\mu_1' = \mathbb{E}(X) = 1, \mu_2 = Var(X) = 1$.

Now, consider the Maclaurin's series expansion for M_X around $t = 0$. We have

$$M_X(t) = (1-t)^{-1} = \sum_{r=0}^{\infty} t^r, \forall t \in (-1,1)$$

and hence $\mu_r' = r!$, which is the coefficient of $\frac{t^r}{r!}$ in the above power series.

Example 1.218. Let X be a continuous RV with p.d.f.

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

As observed earlier in Example 1.186, $\mathbb{E}X$ does not exist. Since the existence of moments is a necessary condition for the existence of MGF, we conclude that the MGF does not exist for this RV X.

Remark 1.219 (Identically distributed RVs). Let X and Y be two RVs, possibly defined on different probability spaces.

(a) Recall from Remark 1.112 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 1.113 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 1.127, 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 1.138 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 1.141, 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 1.220. We write $X \stackrel{d}{=} Y$ to denote that X and Y are identically distributed.

Example 1.221. If Y is an RV with the MGF $M_Y(t) = (1-t)^{-1}, \forall t \in (-1, 1)$, then by Example 1.217, 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 1.222. 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 1.223. We may refer to expectations of the form $\mathbb{E}e^{tX}$ as exponential moments of the RV X.

Definition 1.224 (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 1.225. *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.