Moment functions

1 Moments of a random variable

**Definition 1.1.** Let $X$ be a rrv on probability space $(\Omega, \mathcal{A}, P)$. For a given $r \in \mathbb{N}$, $\mathbb{E}[X^r]$, if it exists, is called the $r$-th moment of $X$. In particular,

- If $X$ is a discrete rrv with pmf $p_X$ and $X(\Omega)$ is the set of all values that $X$ can take on, then
  \[ \mathbb{E}[X^r] = \sum_{x \in X(\Omega)} x^r p_X(x) \]  

- If $X$ is a continuous rrv with pdf $f_X$, then
  \[ \mathbb{E}[X^r] = \int_{-\infty}^{\infty} x^r f_X(x) \, dx \]  

**Remarks:**
- The zero-th moment is 1.
- The first moment is the expected value of $X$.

**Definition 1.2.** Let $X$ be a rrv on probability space $(\Omega, \mathcal{A}, P)$. For a given $r \in \mathbb{N}$, $\mathbb{E}[(X - \mathbb{E}[X])^r]$, if it exists, is called the $r$-th moment about the mean or $r$-th central moment of $X$. In particular,

- If $X$ is a discrete rrv with pmf $p_X$ and $X(\Omega)$ is the set of all values that $X$ can take on, then
  \[ \mathbb{E}[(X - \mathbb{E}[X])^r] = \sum_{x \in X(\Omega)} (x - \mathbb{E}[X])^r p_X(x) \]  

- If $X$ is a continuous rrv with pdf $f_X$, then
  \[ \mathbb{E}[(X - \mathbb{E}[X])^r] = \int_{-\infty}^{\infty} (x - \mathbb{E}[X])^r f_X(x) \, dx \]  

**Remarks:**
- The zero-th central moment is 1.
- The first central moment is 0.
- The second central moment is the variance of $X$. 
2  Moment generating functions

**Definition 2.1.** Let $X$ be a rrv on probability space $(\Omega, \mathcal{A}, P)$. For a given $t \in \mathbb{R}$, the **moment generating function** (m.g.f.) of $X$, denoted $M_X(t)$, is defined as follows

$$M_X(t) = E\left[e^{tX}\right]$$

where there is a positive number $h$ such that the above summation exists for $-h < t < h$. In particular,

- If $X$ is a discrete rrv with pmf $p_X$ and $X(\Omega)$ is the set of all values that $X$ can take on, then

$$M_X(t) = \sum_{x \in X(\Omega)} e^{tx}p_X(x)$$

- If $X$ is a continuous rrv with pdf $f_X$, then

$$M_X(t) = \int_{-\infty}^{\infty} e^{tx}f_X(x) \, dx$$

**Remarks:**

- The m.g.f. $M_X(0)$ always exists and is equal to 1.

**Theorem 2.1.** A moment generating function completely determines the distribution of a real-valued random variable.

**Proposition 2.1.** Let $X$ be a rrv on probability space $(\Omega, \mathcal{A}, P)$. The $r$-th moment of $X$ can be found by evaluating the $r$-th derivative of the m.g.f. of $X$ at $t = 0$.

$$M_X^{(r)}(0) = E[X^r]$$

**Proof.** Using the expansion of the exponential function as a series, we have that:

$$e^{tX} = \sum_{n=0}^{\infty} \frac{(tX)^n}{n!}$$

Hence,

$$M_X(t) = E[e^{tX}]$$

$$= E \left[ \sum_{n=0}^{\infty} \frac{(tX)^n}{n!} \right]$$

$$= \sum_{n=0}^{\infty} \frac{t^n}{n!} E[X^n]$$

$$= 1 + t E[X] + \frac{t^2}{2} E[X^2] + \frac{t^3}{6} E[X^3] + \ldots$$
Differentiating \( M_X \) \( r \) times, the first \( r - 1 \) terms vanish. We now have the following :

\[
\frac{d^r M_X(t)}{dt^r} = \frac{r(r-1)\ldots 1}{r!} E[X^r] + \frac{(r+1)r\ldots 2}{(r+1)!} t E[X^{r+1}] + \frac{(r+2)(r+1)\ldots 3}{(r+2)!} t^2 E[X^{r+2}] + \ldots
\]

\[
= E[X^r] + \sum_{k=1}^{\infty} \frac{t^k}{k!} E[X^{r+k}]
\]

It now becomes clear that evaluating \( \frac{d^r M_X(t)}{dt^r} \) at \( t = 0 \) gives the result. \( \Box \)

**Lemma 2.2.** Let \( X \) be a rrv on probability space \( (\Omega, \mathcal{A}, P) \).

1. The expected value of \( X \), if it exists, can be found by evaluating the first derivative of the moment generating function at \( t = 0 \):

\[
E[X] = M_X'(0) \tag{9}
\]

2. The variance of \( X \), if it exists, can be found by evaluating the first and second derivatives of the moment generating function at \( t = 0 \):

\[
Var(X) = M_X''(0) - (M_X'(0))^2 \tag{10}
\]

## 3 Moment generating functions of Common Discrete Distributions

### 3.1 Discrete Uniform Distribution

**Definition 3.1** (Discrete Uniform Distribution). Let \( (\Omega, \mathcal{A}, P) \) be a probability space and let \( X \) be a random variable that can take \( n \in \mathbb{N}^* \) values on \( X(\Omega) = \{1, \ldots, n\} \). \( X \) is said to have a **discrete uniform distribution** \( U_n \), if its probability mass function is given by :

\[
p_X(i) = P(X = i) = \frac{1}{n}, \text{ for } i = 1, \ldots, n \tag{11}
\]
**Computation of the mgf.** Let $X$ be a random variable that follows a discrete uniform distribution $\mathcal{U}_n$. The mgf of $X$ is given by:

$$M_X(t) = \mathbb{E}[e^{tX}]$$

$$= \sum_{x=1}^{n} e^{tx} p_X(x)$$

$$= \frac{1}{n} \sum_{x=1}^{n} (e^t)^x$$

(sum of a geometric sequence)

$$= \frac{1}{n} \frac{1 - e^{nt}}{1 - e^t}$$

$$= \frac{e^t - e^{(n+1)t}}{n(1 - e^t)}$$

The above equality is true for all $t \neq 0$. If $t = 0$, we have:

$$M_X(0) = \frac{1}{n} \sum_{x=1}^{n} e^{0x}$$

$$= \frac{1}{n} \sum_{x=1}^{n} 1$$

$$= 1$$

In a nutshell, the mgf of $X$ is given by:

$$M_X(t) = \begin{cases} 
\frac{e^t - e^{(n+1)t}}{n(1 - e^t)} & \text{if } t \neq 0 \\
1 & \text{if } t = 0
\end{cases}$$

### 3.2 Binomial Distribution

**Definition 3.2 (Bernoulli process).** A **Bernoulli or binomial process** has the following features:

1. We repeat $n \in \mathbb{N}^*$ identical trials

2. A trial can result in only two possible outcomes, that is, a certain event $E$, called **success**, occurs with probability $p$, thus event $E^c$, called **failure**, occurs with probability $1 - p$

3. The probability of success $p$ remains constant trial after trial. In this case, the process is said to be **stationary**.

4. The trials are mutually independent.
Definition 3.3 (Binomial Distribution). Let \((\Omega, \mathcal{A}, \mathbb{P})\) be a probability space. Let \(E \in \mathcal{A}\) be an event labeled as success, that occurs with probability \(p\). If \(n \in \mathbb{N}^*\) trials are performed according to a Bernoulli process, then the random variable \(X\) defined as the number of successes among the \(n\) trials, is said to have a binomial distribution \(\text{Bin}(n, p)\) and its probability mass function is given by:

\[
p_X(x) = \mathbb{P}(X = x) = \binom{n}{x} p^x (1 - p)^{n-x} \quad \text{for} \quad x = 0, \ldots, n \tag{12}
\]

Computation of the mgf. Let \(X\) be a random variable that follows a binomial distribution \(\text{Bin}(n, p)\). The mgf of \(X\) is given by:

\[
M_X(t) = \sum_{x=0}^{n} e^{tx} p_X(x)
= \sum_{x=0}^{n} e^{tx} \binom{n}{x} p^x (1 - p)^{n-x}
= \sum_{x=0}^{n} \binom{n}{x} (pe^t)^x (1 - p)^{n-x}
= (pe^t + 1 - p)^n
\]

where the last equality comes from the binomial theorem.

3.3 Poisson Distribution

Definition 3.4 (Poisson Distribution). Let \((\Omega, \mathcal{A}, \mathbb{P})\) be a probability space. A random variable \(X\) is said to have a Poisson distribution \(\mathcal{P}(\lambda)\), with \(\lambda > 0\) if its probability mass function is given by:

\[
p_X(x) = \mathbb{P}(X = x) = e^{-\lambda} \frac{\lambda^x}{x!}, \quad \text{for} \quad x \in \mathbb{N} \tag{13}
\]

Computation of the mgf. Let \(X\) be a random variable that follows a Poisson distribution \(\mathcal{P}(\lambda)\). The mgf of \(X\) is given by:

\[
M_X(t) = \sum_{x=0}^{\infty} e^{tx} p_X(x)
= \sum_{x=0}^{\infty} e^{tx} e^{-\lambda} \frac{\lambda^x}{x!}
= e^{-\lambda} \sum_{x=0}^{\infty} \frac{(\lambda e^t)^x}{x!}
= e^{-\lambda} \exp(\lambda e^t)
= \exp(\lambda(e^t - 1))
\]
3.4 Geometric Distribution

**Definition 3.5** (Geometric Distribution). Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space. Let $E \in \mathcal{A}$ be an event labeled as success, that occurs with probability $p$. If all the assumptions of a Bernoulli process are satisfied, except that the number of trials is not preset, then the random variable $X$ defined as the number of trials until the first success is said to have a geometric distribution $G(p)$ and its probability mass function is given by:

$$p_X(x) = \mathbb{P}(X = x) = (1 - p)^{x-1}p, \text{ for } x \in \mathbb{N}^*$$

(14)

**Computation of the mgf.** Left as an exercise (Homework 3)!

<table>
<thead>
<tr>
<th>Distribution</th>
<th>Probability Mass Function $p_X(x) = \mathbb{P}(X = x)$</th>
<th>$M_X(t)$</th>
<th>$\mathbb{E}[X]$</th>
<th>$\text{Var}(X)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Uniform $U_n$</td>
<td>$\frac{1}{n}$</td>
<td>$\begin{cases} e^{t-\frac{e^{(n+1)t}}{n(1-e^t)}} \text{ if } t \neq 0 \ 1 \text{ if } t = 0 \end{cases}$</td>
<td>$n + 1$</td>
<td>$\frac{n^2 - 1}{12}$</td>
</tr>
<tr>
<td>$n \in \mathbb{N}^*$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Binomial $Bin(n, p)$</td>
<td>$\binom{n}{x} p^x (1 - p)^{n-x} \geq 0$ \text{ for } x = 0, \ldots, n</td>
<td>$(pe^t + 1 - p)^n$</td>
<td>$np$</td>
<td>$np(1 - p)$</td>
</tr>
<tr>
<td>$n \in \mathbb{N}^*, p \in (0, 1)$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Poisson $P(\lambda)$</td>
<td>$e^{-\lambda} \frac{\lambda^x}{x!}$ \text{ for } x \in \mathbb{N}$</td>
<td>$\exp(\lambda(e^t - 1))$</td>
<td>$\lambda$</td>
<td>$\lambda$</td>
</tr>
<tr>
<td>$\lambda &gt; 0$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Geometric $G(p)$</td>
<td>$(1 - p)^{x-1}p$</td>
<td>$\frac{pe^t}{1 - (1 - p)e^t}$ \text{ for } t &lt; -\ln(1 - p)</td>
<td>$\frac{1}{p}$</td>
<td>$\frac{1 - p}{p^2}$</td>
</tr>
<tr>
<td>$p \in (0, 1)$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
4 Moment generating functions of Common Continuous Distributions

4.1 Continuous Uniform Distribution

Definition 4.1. A continuous r.v. is said to follow a uniform distribution $U(a, b)$ on a segment $[a, b]$, with $a < b$, if its pdf is

$$f_X(x) = \frac{1}{b - a} \mathbb{1}_{[a,b]}(x)$$

(15)

Computation of the mgf. Let $X$ be a continuous random variable that follows a uniform distribution $U(a, b)$. The mgf of $X$ is given by:

$$M_X(t) = \int_{-\infty}^{\infty} e^{tx} f_X(x) \, dx$$

$$= \frac{1}{b - a} \int_a^b e^{tx} \, dx$$

$$= \frac{1}{b - a} \frac{e^{tx}}{t} \bigg|_a^b$$

$$= \frac{e^{tb} - e^{ta}}{t(b - a)}$$

The above equality holds for $t \neq 0$. We notice that $M_X(0) = 1$.

4.2 Normal Distribution

Definition 4.2. A continuous random variable is said to follow a normal (or Gaussian) distribution $N(\mu, \sigma^2)$ with parameters, mean $\mu$ and variance $\sigma^2$ if its pdf $f_X$ is given by:

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

(16)
Computation of the mgf. We start by computing the mgf a standard normal random variable \( Z \sim \mathcal{N}(0, 1) \). The mgf of \( Z \) is given by:

\[
M_Z(t) = \int_{-\infty}^{\infty} e^{tx} f_Z(x) \, dx
\]

\[
= \int_{-\infty}^{\infty} e^{tx} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \, dx
\]

\[
= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x^2 - 2tx)} \, dx
\]

\[
= e^{t^2/2} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x-t)^2} \, dx
\]

\[
= e^{t^2/2}
\]

We know (Property 4.5 in Chapter 5) that \( X = \mu + \sigma Z \) follows a normal distribution \( \mathcal{N}(\mu, \sigma^2) \). Therefore the mgf of a normal distribution \( \mathcal{N}(\mu, \sigma^2) \) is given by:

\[
M_{\mu+\sigma Z}(t) = E \left[ e^{t(\mu+\sigma Z)} \right]
\]

\[
= e^{t\mu} E \left[ e^{t\sigma Z} \right]
\]

\[
= e^{t\mu} M_Z(t\sigma)
\]

\[
= e^{t\mu} e^{t^2\sigma^2/2}
\]

\[
= \exp \left\{ \mu t + \frac{\sigma^2 t^2}{2} \right\}
\]

4.3 Exponential Distribution

Definition 4.3. A continuous random variable is said to follow an exponential distribution \( \mathcal{E}(\lambda) \) with \( \lambda > 0 \) if its pdf \( f_X \) is given by:

\[
f_X(x) = \lambda e^{-\lambda x} 1_{\mathbb{R}^+}(x)
\]

(17)
Computation of the mgf. Let $X$ be a continuous random variable that follows an exponential distribution $\mathcal{E}(\lambda)$. The mgf of $X$ is given by:

$$M_X(t) = \int_{-\infty}^{\infty} e^{tx} f_X(x) \, dx$$

$$= \int_{0}^{\infty} e^{tx} \lambda e^{-\lambda x} \, dx$$

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

$$= \lambda \left( \frac{1}{(\lambda-t)} \right)_{x=0}^{\infty}$$

$$= \frac{\lambda}{\lambda-t} [0 - 1]$$

$$= \frac{\lambda}{\lambda-t}$$

When integrating the exponential, we must be aware that $\lim_{x \to \infty} e^{-x(\lambda-t)} = 0$ if and only if $\lambda - t > 0$. Therefore the derived formula holds if and only if $t < \lambda$.

<table>
<thead>
<tr>
<th>Distribution</th>
<th>Probability Density Function $f_X(x)$</th>
<th>$M_X(t)$</th>
<th>$\mathbb{E}[X]$</th>
<th>$\text{Var}(X)$</th>
</tr>
</thead>
</table>
| Uniform $\mathcal{U}(a,b)$ $a,b \in \mathbb{R}$ with $a < b$ | $\frac{1}{b-a} \mathbb{1}_{[a,b]}(x)$ | \begin{cases} 
\frac{e^{tb} - e^{ta}}{b-a} & \text{if } t \neq 0 \\
1 & \text{if } t = 0 
\end{cases} | \frac{a+b}{2} | \frac{(b-a)^2}{12} |
| Normal $\mathcal{N}(\mu,\sigma^2)$ $\mu \in \mathbb{R}, \sigma > 0$ | $\frac{1}{\sigma \sqrt{2\pi}} \exp \left\{ -\frac{1}{2} \left( \frac{x - \mu}{\sigma} \right)^2 \right\}$ for $x \in \mathbb{R}$ | $\exp \left\{ \mu t + \frac{\sigma^2 t^2}{2} \right\}$ | $\mu$ | $\sigma^2$ |
| Exponential $\mathcal{E}(\lambda)$ $\lambda > 0$ | $\lambda e^{-\lambda x} \mathbb{1}_{\mathbb{R}_+}(x)$ | $\frac{\lambda}{\lambda-t}$ for $t < \lambda$ | $\frac{1}{\lambda}$ | $\frac{1}{\lambda^2}$ |