Probability and Stochastic Processes
Lecture 11: Expectation

Prof. Washim Uddin Mondal
Department of Electrical Engineering
Indian Institute of Technology Kanpur

Introduction

Motivation

- Let X be a discrete random variable with range $\{x_1, \cdots, x_K\}$.
- One way to characterize X is by specifying its PMF: $\{\mathbf{P}(X = x_i)\}_{i=1}^{K}$. Another incomplete but cheap way is by specifying its average value.
- Assume that n independent samples are drawn from the same distribution as that of X, and n_i denotes the number of times the event $\{X = x_i\}$ occurs.
- Frequentist interpretation says that $\mathbf{P}(X = x_i) = \lim_{n\to\infty} n_i/n$. In the limiting regime $n \to \infty$, the average value obtained is

$$\lim_{n\to\infty} \frac{\sum_{i=1}^{K} n_i x_i}{n} = \sum_{i=1}^{K} x_i \mathbf{P}(X = x_i)$$

Definitions

Expectation for Discrete Random Variables

Let X be a discrete random variable defined on the probability space $(\Omega, \mathcal{F}, \mathbf{P})$ with range $\{x_i\}_{i \in I}$. The expectation of X is defined as follows.

$$\mathbf{E}[X] = \sum_{i \in I} x_i \mathbf{P}(X = x_i)$$

Expectation for Continuous Random Variables

Let X be a continuous random variable with PDF f_X. The expectation of X is:

$$\mathbf{E}[X] = \int_{\mathbb{R}} x f_X(x) \mathrm{d}x$$

Illustrative Example

Example

Let T denote the inter-emission time of a radioactive material. It is observed that $T \sim \exp(\lambda)$. Find the expected inter-emission time of the material.

Solution

- Note that $f_T(t) = \lambda e^{-\lambda t}\mathbf{1}(t \geq 0)$, where $\mathbf{1}$ is the indicator function.
- The expected inter-emission time is:

$$\mathbf{E}[T] = \int_{\mathbb{R}} t f_T(t)\mathrm{d}t = \int_0^{\infty} \lambda t e^{-\lambda t}\mathrm{d}t = \frac{1}{\lambda}\int_0^{\infty} y e^{-y}\mathrm{d}y = \frac{\Gamma(2)}{\lambda} = \frac{1}{\lambda}$$

Issues with Expectation

Notes

- For a random variable X defined on a probability space $(\Omega, \mathcal{F}, \mathbf{P})$, the expectation is defined using the *Lebesgue integration*, which is outside the scope of this course.

$$\mathbf{E}[X] = \int_{\Omega} X(\omega) \mathrm{d}\mathbf{P}(\omega)$$

- The above definition reduces to our familiar definition of expectation in the special cases of discrete and continuous random variables.
- In general, $\mathbf{E}[X]$ may not always be well-defined. In this course, however, we assume that $\mathbf{E}[X]$ is well-defined unless mentioned otherwise.

Expectation of a Function of a Random Variable

Proposition 11.1

Let X be a random variable associated with a probability space $(\Omega, \mathcal{F}, \mathbf{P})$, and $g : \mathbb{R} \to \mathbb{R}$ be a Borel-measurable function.

(a) If X is a discrete random variable with range $\{x_i\}_{i \in I}$, then

$$\mathbf{E}[g(X)] = \sum_{i \in I} g(x_i)\mathbf{P}(X = x_i)$$

(b) If X is a continuous random variable with PDF f_X, then

$$\mathbf{E}[g(X)] = \int_{\mathbb{R}} g(x) f_X(x) \mathrm{d}x$$

Expectation of a Function of a Random Variable

Proof of Proposition 11.1(a)

- Let $\{y_j\}_{j\in J}$ be the range of $Y = g(X)$. Note that $\mathbf{P}(Y = y_j) = \sum_{i\in I:g(x_i)=y_j} \mathbf{P}(X = x_i)$.

$$\mathbf{E}[Y] = \sum_{j\in J} y_j \mathbf{P}(Y = y_j) = \sum_{j\in J} y_j \sum_{i\in I:g(x_i)=y_j} \mathbf{P}(X = x_i) = \sum_{i\in I} g(x_i)\mathbf{P}(X = x_i)$$

Notes

- Proposition 11.1 is called the Law of the Unconscious Statistician (LOTUS) since people often use the formulae of $\mathbf{E}[g(X)]$ stated in this proposition as the definition of $\mathbf{E}[g(X)]$ rather than a derived result.

Linearity of Expectation

Proposition 11.2

Let X be a random variable associated with some probability space $(\Omega, \mathcal{F}, \mathbf{P})$. If a, b are arbitrary real numbers, then $\mathbf{E}[aX + b] = a\mathbf{E}[X] + b$.

Proof of Proposition 11.2 (Discrete Case)

- Assume that X is a discrete random variable with range $\{x_i\}_{i \in I}$. Clearly,

$$\begin{aligned}\mathbf{E}[aX + b] &= \sum_{i \in I} (ax_i + b)\mathbf{P}(X = x_i) \\ &= a \sum_{i \in I} x_i \mathbf{P}(X = x_i) + b \sum_{i \in I} \mathbf{P}(X = x_i) = a\mathbf{E}[X] + b\end{aligned}$$

Linearity of Expectation

Proof of Proposition 11.2 (Continuous Case)

- Let X be a continuous random variable with PDF f_X. We have the following.

$$\begin{aligned}\mathbf{E}[aX+b] &= \int_{\mathbb{R}} (ax+b) f_X(x)\mathrm{d}x \\ &= a\int_{\mathbb{R}} x f_X(x)\mathrm{d}x + b\int_{\mathbb{R}} f_X(x)\mathrm{d}x = a\mathbf{E}[X]+b\end{aligned}$$

Note

- If $a_0, a_1, \cdots, a_m$ are arbitrary real numbers, then

$$\mathbf{E}[a_0 + a_1 X + \cdots + a_m X^m] = a_0 + a_1\mathbf{E}[X] + \cdots + a^m \mathbf{E}[X^m]$$

Moments of a Random Variable

Definition

- The m-th raw moment of a random variable X is $\mathbf{E}[X^m]$.
- The m-th central moment of a random variable X is $\mathbf{E}[(X - \mathbf{E}[X])^m]$.
- The second central moment of X is called its variance: $\mathrm{Var}(X) = \mathbf{E}[(X - \mathbf{E}[X])^2]$.
- $\sqrt{\mathrm{Var}(X)}$ is called the standard deviation of X.

Note

- Central moments measure the tendency to deviate from the mean. The larger these moments, the more likely it is that the values of X will spread over a wide range.

Moments of a Random Variable

Notes

- The first central moment is $\mathbf{E}[X - \mathbf{E}[X]] = 0$ for any random variable X.
- The variance of X can be expanded as follows.

$$\begin{aligned}\mathrm{Var}(X) &= \mathbf{E}[(X - \mathbf{E}[X])^2] \\ &= \mathbf{E}[X^2 - 2X\mathbf{E}[X] + (\mathbf{E}[X])^2] \\ &\overset{(a)}{=} \mathbf{E}[X^2] - 2(\mathbf{E}[X])^2 + (\mathbf{E}[X])^2 = \mathbf{E}[X^2] - (\mathbf{E}[X])^2\end{aligned}$$

  where (a) uses the linearity of the expectation operation.
- Since $\mathrm{Var}(X) \geq 0$, we get $\mathbf{E}[X^2] \geq (\mathbf{E}[X])^2$.

Moments of a Random Variable

Proposition 11.3

If $Y = aX + b$ for some $a, b \in \mathbb{R}$, then

- $\mathbf{E}[(Y - \mathbf{E}[Y])^m] = a^m \mathbf{E}[(X - \mathbf{E}[X])^m]$, $\forall m \in \mathbb{N}$.
- In particular, $\mathrm{Var}(Y) = a^2 \mathrm{Var}(X)$.

Proof of Proposition 11.3

- Recall that $\mathbf{E}[Y] = a\mathbf{E}[X] + b$. Note that

$$\mathbf{E}[(Y - \mathbf{E}[Y])^m] = \mathbf{E}[a^m (X - \mathbf{E}[X])^m] \overset{(a)}{=} a^m \mathbf{E}[(X - \mathbf{E}[X])^m]$$

  where (a) follows from the linearity of $\mathbf{E}[\cdot]$.

Moment Generating Function (MGF)

Definition

The moment generating function (MGF) of a random variable X is a function of the form $M_X : \mathbb{R} \to [0, \infty]$ such that $M_X(t) = \mathbf{E}[e^{tX}], \ \forall t \in \mathbb{R}$.

Notes

- Since e^{tX} is a positive random variable, $M_X(t) \in [0, \infty], \forall t$.
- The set $\{t \in \mathbb{R} \mid M_X(t) < \infty\}$ is called the region of convergence (RoC) of M_X.
- If $M_X(t) = M_Y(t) < \infty, \forall t \in (-\delta, \delta)$ for some $\delta > 0$, then $F_X(x) = F_Y(x), \forall x \in \mathbb{R}$. This is known as the Uniqueness Theorem for MGFs.

Moment Generating Function (MGF)

Notes

- Expanding e^{tX}, we obtain the following.

$$M_X(t) = \mathbf{E}\left[\sum_{n=0}^{\infty} \frac{t^n X^n}{n!}\right] = \sum_{n=0}^{\infty} \frac{t^n}{n!}\mathbf{E}[X^n]$$

- This is why M_X is called the moment generating function.
- If $M_X^{(n)}(t)$ denotes the n-th derivative of M_X, evaluated at t, then

$$M_X^{(n)}(0) = \mathbf{E}[X^n]$$

- If $Y = aX + b$, then $M_Y(t) = \mathbf{E}[e^{(aX+b)t}] = e^{bt}M_X(at)$.

Bernoulli Random Variable

Moments

- If $X \sim \text{Bernoulli}(p)$, then $\mathbf{E}[X^m] = p$, $\forall m \in \mathbb{N}$. In particular,
  (a) $\mathbf{E}[X] = p$, and $\mathbf{E}[X^2] = p$.
  (b) $\text{Var}(X) = \mathbf{E}[X^2] - (\mathbf{E}[X])^2 = p(1-p)$.

Moment Generating Function (MGF)

- The MGF is $M_X(t) = 1 - p + pe^t$, $\forall t \in \mathbb{R}$.
- The region of convergence (RoC) is $\mathbb{R}$.

Geometric Random Variable

Moment Generating Function (MGF)

- If $X \sim \text{Geom}(p)$, then its MGF is given as follows.

$$M_X(t) = \mathbf{E}[e^{tX}] = \sum_{x=1}^{\infty} e^{tx}(1-p)^{x-1}p$$

$$= pe^t \sum_{x=1}^{\infty} \{e^t(1-p)\}^{x-1} = \begin{cases} \dfrac{pe^t}{1-e^t(1-p)} & \text{if } t < -\ln(1-p) \\ \infty & \text{elsewhere} \end{cases}$$

Geometric Random Variable

Moments

- The mean can be computed as follows.

$$\mathbf{E}[X] = M_X^{(1)}(0) = \left\{ \frac{\mathrm{d}}{\mathrm{d}t} \left[\frac{p}{e^{-t} - (1-p)} \right] \right\}_{t=0} = \left\{ \frac{pe^{-t}}{(e^{-t} - (1-p))^2} \right\}_{t=0} = \frac{1}{p}$$

- We can obtain the same result using the negative binomial series.

$$\mathbf{E}[X] = \sum_{x=1}^{\infty} xp(1-p)^{x-1} = p \sum_{y=2}^{\infty} \binom{y-1}{2-1} (1-p)^{y-2} = \frac{p}{(1-(1-p))^2} = \frac{1}{p}$$

Geometric Random Variable

Moments

- The second raw moment is computed as:

$$\begin{aligned}
\mathbf{E}[X^2] = M_X^{(2)}(0) &= \left\{ \frac{\mathrm{d}}{\mathrm{d}t} \left[\frac{pe^{-t}}{(e^{-t} - (1-p))^2} \right] \right\}_{t=0} \\
&= \left\{ \frac{-(e^{-t} - (1-p))^2 pe^{-t} + 2(e^{-t} - (1-p))pe^{-2t}}{(e^{-t} - (1-p))^4} \right\}_{t=0} \\
&= \frac{-p^3 + 2p^2}{p^4} = \frac{2-p}{p^2}
\end{aligned}$$

- The variance is given as $\mathrm{Var}(X) = \mathrm{E}[X^2] - (\mathbf{E}[X])^2 = (1-p)/p^2$.

Geometric Random Variable

Moments

- Alternatively, we can obtain the same result using the negative binomial series.

$$\mathbf{E}[X(X-1)] = \sum_{x=2}^{\infty} x(x-1)p(1-p)^{x-1}$$
$$= 2p(1-p)\sum_{y=3}^{\infty}\binom{y-1}{3-1}(1-p)^{y-3} = \frac{2p(1-p)}{(1-(1-p))^3} = \frac{2(1-p)}{p^2}$$

- Hence, $\mathbf{E}[X^2] = \mathbf{E}[X(X-1)] + \mathbf{E}[X] = (2-p)/p^2$.

Poisson Random Variable

Moment Generating Function (MGF)

- If $X \sim \text{Poisson}(\lambda)$, then its MGF is given as follows.

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

Moments

- The mean can be computed as:

$$\mathbf{E}[X] = \sum_{x=0}^{\infty} x e^{-\lambda} \frac{\lambda^x}{x!} = \lambda e^{-\lambda} \sum_{x=1}^{\infty} \frac{\lambda^{x-1}}{(x-1)!} = \lambda$$

Poisson Random Variable

Moments

- Alternatively, the mean can also be computed from the MGF.

$$\mathbf{E}[X] = M_X^{(1)}(0) = \left\{ \frac{\mathrm{d}}{\mathrm{d}t} \left[\exp\left(\lambda(e^t - 1)\right) \right] \right\}_{t=0} = \left\{ \lambda e^t \exp(\lambda(e^t - 1)) \right\}_{t=0} = \lambda$$

- The second raw moment can be obtained as:

$$\begin{aligned} \mathbf{E}[X^2] = M_X^{(2)}(0) &= \left\{ \frac{\mathrm{d}}{\mathrm{d}t} \left[\lambda \exp(\lambda(e^t - 1) + t) \right] \right\}_{t=0} \\ &= \left\{ \lambda(\lambda e^t + 1) \exp(\lambda(e^t - 1) + t) \right\}_{t=0} = \lambda(\lambda + 1) \end{aligned}$$

Poisson Random Variable

Moments

- Alternative, we can first compute $\mathbf{E}[X(X-1)]$ as shown below.

$$\mathbf{E}[X(X-1)] = \sum_{x=0}^{\infty} x(x-1)e^{-\lambda}\frac{\lambda^x}{x!} = \lambda^2 e^{-\lambda}\sum_{x=2}^{\infty}\frac{\lambda^{x-2}}{(x-2)!} = \lambda^2$$

- This leads to $\mathbf{E}[X^2] = \mathbf{E}[X(X-1)] + \mathbf{E}[X] = \lambda(\lambda+1)$.
- Therefore, the variance is $\mathrm{Var}(X) = \mathbf{E}[X^2] - (\mathbf{E}[X])^2 = \lambda$.

Binomial Random Variable

Moment Generating Function (MGF)

- If $X \sim \text{Binomial}(n, p)$, then $\forall t \in \mathbb{R}$, we get the following.

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

Moments

- The mean can be computed as:

$$\mathbf{E}[X] = M_X^{(1)}(0) = \left\{npe^t(pe^t + 1 - p)^{n-1}\right\}_{t=0} = np$$

Binomial Random Variable

Moments

- Alternatively, we can obtain the mean as follows.

$$
\begin{aligned}
\mathbf{E}[X] &= \sum_{x=0}^{n} x \binom{n}{x} p^x (1-p)^{n-x} \\
&= \sum_{x=1}^{n} \frac{n!}{(x-1)!(n-x)!} p^x (1-p)^{n-x} \\
&= np \sum_{y=0}^{n-1} \binom{n-1}{y} p^y (1-p)^{n-1-y} = np(p+1-p)^{n-1} = np
\end{aligned}
$$

Binomial Random Variable

Moments

- To obtain the second raw moment, we can first compute $\mathbf{E}[X(X-1)]$ as:

$$\mathbf{E}[X(X-1)] = \sum_{x=1}^{n} x(x-1)\binom{n}{x} p^x (1-p)^{n-x}$$

$$= \sum_{x=2}^{n} \frac{n!}{(x-2)!(n-x)!} p^x (1-p)^{n-x}$$

$$= n(n-1)p^2 \sum_{y=0}^{n-2} \binom{n-2}{y} p^y (1-p)^{n-2-y} = n(n-1)p^2$$

Binomial Random Variable

Moments

- This leads to $\mathbf{E}[X^2] = \mathbf{E}[X(X-1)] + \mathbf{E}[X] = n(n-1)p^2 + np$.
- We can alternatively obtain $\mathbf{E}[X^2]$ using the MGF.

$$
\begin{aligned}
\mathbf{E}[X^2] = M_X^{(2)}(0) &= \left\{ \frac{\mathrm{d}}{\mathrm{d}t} \left[npe^t (pe^t + 1 - p)^{n-1} \right] \right\}_{t=0} \\
&= \{npe^t (pe^t + 1 - p)^{n-1} + n(n-1)p^2 e^{2t} (pe^t + 1 - p)^{n-2}\}_{t=0} \\
&= np + n(n-1)p^2
\end{aligned}
$$

- The variance becomes $\mathrm{Var}(X) = \mathbf{E}[X^2] - (\mathbf{E}[X])^2 = np(1-p)$.

Negative Binomial Random Variable

Moment Generating Function (MGF)

- If $X \sim \mathrm{NB}(r, p)$, then we can compute its MGF as follows.

$$M_X(t) = \sum_{x=r}^{\infty} e^{tx} \binom{x-1}{r-1} p^r (1-p)^{x-r}$$

$$= (pe^t)^r \sum_{x=r}^{\infty} \binom{x-1}{r-1} \left\{e^t(1-p)\right\}^{x-r}$$

$$= \begin{cases} \left(\dfrac{pe^t}{1-e^t(1-p)}\right)^r & \text{if } t < -\ln(1-p) \\ \infty & \text{elsewhere} \end{cases}$$

Negative Binomial Random Variable

Moments

- Note that $M_X(t) = [M_Z(t)]^r$, where $Z \sim \text{Geom}(p)$. We have the following.

$$M_X^{(1)}(t) = \frac{\mathrm{d}}{\mathrm{d}t}\left[(M_Z(t))^r\right] = r(M_Z(t))^{r-1}M_Z^{(1)}(t)$$

$$M_X^{(2)}(t) = r(r-1)(M_Z(t))^{r-2}(M_Z^{(1)}(t))^2 + r(M_Z(t))^{r-1}M_Z^{(2)}(t)$$

- Recall that $M_Z^{(1)}(0) = \mathbf{E}[Z] = 1/p$, $M_Z^{(2)}(0) = \mathbf{E}[Z^2] = (2-p)/p^2$.
- Moreover, $M_Z(0) = 1$. Therefore, the mean can be obtained as:

$$\mathbf{E}[X] = M_X^{(1)}(0) = \frac{r}{p}$$

Negative Binomial Random Variable

Moments

- Alternatively, we can compute the mean as:

$$\mathbf{E}[X] = \sum_{x=r}^{\infty} x \binom{x-1}{r-1} p^r (1-p)^{x-r}$$

$$= \sum_{x=r}^{\infty} \frac{x!}{(r-1)!(x-r)!} p^r (1-p)^{x-r}$$

$$= rp^r \sum_{y=r+1}^{\infty} \binom{y-1}{r} (1-p)^{y-r-1} = \frac{rp^r}{(1-(1-p))^{r+1}} = \frac{r}{p}$$

Negative Binomial Random Variable

Moments

- To obtain the second raw moment, we need to first obtain $\mathbf{E}[X(X+1)]$.

$$\begin{aligned}
\mathbf{E}[X(X+1)] &= \sum_{x=r}^{\infty} x(x+1)\binom{x-1}{r-1}p^r(1-p)^{x-r} \\
&= \sum_{x=r}^{\infty} \frac{(x+1)!}{(r-1)!(x-r)!}p^r(1-p)^{x-r} \\
&= r(r+1)p^r \sum_{y=r+2}^{\infty} \binom{y-1}{r+1}(1-p)^{y-r-2} = \frac{r(r+1)}{p^2}
\end{aligned}$$

Negative Binomial Random Variable

Moments

- The second raw moment, thus, is given as follows.

$$\mathbf{E}[X^2] = \mathbf{E}[X(X+1)] - \mathbf{E}[X] = \frac{r(r+1)}{p^2} - \frac{r}{p} = \frac{r^2}{p^2} + \frac{r(1-p)}{p^2}$$

- The same result can also be obtained using the MGF.

$$\mathbf{E}[X^2] = M_X^{(2)}(0) = \frac{r(r-1)}{p^2} + \frac{r(2-p)}{p^2} = \frac{r^2}{p^2} + \frac{r(1-p)}{p^2}$$

- The variance is $\mathrm{Var}(X) = \mathbf{E}[X^2] - (\mathbf{E}[X])^2 = r(1-p)/p^2$.

Summary

Moments and MGFs of Discrete Random Variables

Distribution	Mean	Variance	MGF (within the RoC)
Bernoulli(p)	p	$p(1-p)$	$1-p+pe^t$
Geom(p)	$1/p$	$(1-p)/p^2$	$pe^t/(1-e^t(1-p))$
Poisson(λ)	λ	λ	$\exp(\lambda(e^t-1))$
Binomial(n,p)	np	$np(1-p)$	$(pe^t+1-p)^n$
NB(r,p)	r/p	$r(1-p)/p^2$	$(pe^t/(1-e^t(1-p)))^r$

Uniform Random Variable

Moment Generating Function (MGF)

- If $X \sim \text{Unif}(a, b)$, then its MGF is:

$$M_X(t) = \mathbf{E}[e^{tX}] = \frac{1}{b-a}\int_a^b e^{tx}\mathrm{d}x = \begin{cases} \dfrac{e^{tb} - e^{ta}}{t(b-a)} & \text{if } t \neq 0 \\ 1 & \text{if } t = 0 \end{cases}$$

Moments

$$\mathbf{E}[X^m] = \frac{1}{b-a}\int_a^b x^m \mathrm{d}x = \frac{b^{m+1} - a^{m+1}}{(m+1)(b-a)}, \ \forall m \in \mathbb{N}$$

Uniform Random Variable

Moments

- In particular, the first and second raw moments are given as follows.

$$\mathbf{E}[X] = \frac{b^2 - a^2}{2(b-a)} = \frac{a+b}{2}, \quad \mathbf{E}[X^2] = \frac{b^3 - a^3}{3(b-a)} = \frac{a^2 + ab + b^2}{3}$$

- Therefore, the variance can be computed as follows.

$$\begin{aligned}\mathrm{Var}(X) &= \mathbf{E}[X^2] - (E[X])^2 \\ &= \frac{a^2 + ab + b^2}{3} - \left(\frac{a+b}{2}\right)^2 = \frac{(b-a)^2}{12}\end{aligned}$$

Uniform Random Variable

Example

A stick of length 1 is split at a point U that is uniformly distributed in $(0,1)$. Find the expected length of the portion containing the point $p \in (0,1)$.

Solution

- If $L_p(U)$ denotes the length of the portion containing p, then

$$L_p(U) = \begin{cases} U & \text{if } p < U \leq 1 \\ 1-U & \text{if } 0 < U \leq p \end{cases}$$

- Hence, $\mathbf{E}[L_p(U)] = \int_0^p (1-u)\mathrm{d}u + \int_p^1 u\mathrm{d}u = \frac{1}{2} + p(1-p)$.

Exponential Random Variable

Moment Generating Function (MGF)

- If $X \sim \exp(\lambda)$, then its MGF is given as follows.

$$M_X(t) = \int_0^\infty e^{tx}\lambda e^{-\lambda x}\mathrm{d}x = \begin{cases} \dfrac{\lambda}{\lambda - t} & \text{if } t < \lambda \\ \infty & \text{elsewhere} \end{cases}$$

Moments

$$\mathbf{E}[X^m] = \int_0^\infty x^m \lambda e^{-\lambda x}\mathrm{d}x = \frac{1}{\lambda^m}\int_0^\infty y^m e^{-y}\mathrm{d}y = \frac{\Gamma(m+1)}{\lambda^m} = \frac{m!}{\lambda^m}$$

- In particular, $\mathbf{E}[X] = \lambda^{-1}$ and $\mathbf{E}[X^2] = 2\lambda^{-2}$, which implies that $\mathrm{Var}(X) = \lambda^{-2}$.

Standard Normal Random Variable

Moment Generating Function (MGF)

- If $Z \sim \mathcal{N}(0,1)$, then its MGF can be obtained as follows $\forall t \in \mathbb{R}$.

$$
\begin{aligned}
M_Z(t) = \mathbf{E}[e^{tZ}] &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left(tz - \frac{z^2}{2}\right) \mathrm{d}z \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left(\frac{t^2}{2} - \frac{1}{2}(z-t)^2\right) \mathrm{d}z \\
&\overset{(a)}{=} \exp\left(\frac{t^2}{2}\right) \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{y^2}{2}\right) \mathrm{d}y = \exp\left(\frac{t^2}{2}\right)
\end{aligned}
$$

where (a) follows from the substitution $y = z - t$.

Standard Normal Random Variable

Moments

- Expanding the MGF as an infinite series, we obtain:

$$M_Z(t) = \sum_{k=0}^{\infty} \frac{(t^2/2)^k}{k!} = \sum_{k=0}^{\infty} \frac{(2k)!}{2^k k!} \frac{t^{2k}}{(2k)!}$$

- The m-th raw moment can be expressed as:

$$\mathbf{E}[Z^m] = \begin{cases} \frac{(2k)!}{2^k k!} & \text{if } m = 2k, k \in \mathbb{N} \\ 0 & \text{if } m = 2k-1, k \in \mathbb{N} \end{cases}$$

- In particular, $\mathbf{E}[Z] = 0$, and $\mathbf{E}[Z^2] = 1$, implying that $\mathrm{Var}(Z) = 1$.

Normal Random Variable

Moment Generating Function (MGF)

- If $X \sim \mathcal{N}(\mu, \sigma^2)$, then $X = \mu + \sigma Z$, where $Z \sim \mathcal{N}(0, 1)$.

$$M_X(t) = \exp(\mu t) M_Z(\sigma t) = \exp\left(\mu t + \frac{\sigma^2 t^2}{2}\right), \ \forall t \in \mathbb{R}$$

Moments

- The mean can be computed as $\mathbf{E}[X] = \mu + \sigma \mathbf{E}[Z] = \mu$.
- Moreover, the variance can be computed as $\mathrm{Var}(X) = \sigma^2 \mathrm{Var}(Z) = \sigma^2$.

Gamma Random Variable

Moment Generating Function (MGF)

- If $X \sim \text{Gamma}(\alpha, \lambda)$, then its MGF can be obtained as follows.

$$M_X(t) = \int_0^\infty \frac{e^{tx}\lambda^\alpha x^{\alpha-1}e^{-\lambda x}}{\Gamma(\alpha)}\mathrm{d}x \overset{(a)}{=} \frac{1}{\Gamma(\alpha)}\int_0^\infty \exp\left(-\left(1-\frac{t}{\lambda}\right)y\right)y^{\alpha-1}\mathrm{d}y$$

  where (a) uses $y = \lambda x$. If $t < \lambda$, then using $z = (1 - t/\lambda)y$, we obtain:

$$M_X(t) = \frac{(1-t/\lambda)^{-\alpha}}{\Gamma(\alpha)}\int_0^\infty e^{-z}z^{\alpha-1}\mathrm{d}z = \left(1-\frac{t}{\lambda}\right)^{-\alpha}$$

- If $t \geq \lambda$, we obtain $M_X(t) = \infty$.

Gamma Random Variable

Moments

- The m-th raw moment can be expressed as follows.

$$\begin{aligned}\mathbf{E}[X^m] &= \frac{1}{\Gamma(\alpha)}\int_0^\infty x^m \lambda^\alpha x^{\alpha-1} e^{-\lambda x}\mathrm{d}x\\ &= \frac{1}{\lambda^m\Gamma(\alpha)}\int_0^\infty y^{m+\alpha-1}e^{-y}\mathrm{d}y = \frac{\Gamma(m+\alpha)}{\lambda^m\Gamma(\alpha)}\end{aligned}$$

- In particular, the first and second raw moments are as follows.

$$\mathbf{E}[X] = \frac{\Gamma(\alpha+1)}{\lambda\Gamma(\alpha)} = \frac{\alpha}{\lambda}, \quad \text{and} \quad \mathbf{E}[X^2] = \frac{\Gamma(\alpha+2)}{\lambda^2\Gamma(\alpha)} = \frac{\alpha(\alpha+1)}{\lambda^2}$$

- Therefore, the variance is $\mathrm{Var}(X) = \mathbf{E}[X^2] - (\mathbf{E}[X])^2 = \alpha/\lambda^2$.

Beta Random Variable

Moments

- The m-th raw moment of $X \sim \mathrm{Beta}(a, b)$ can be obtained as:

$$\mathbf{E}[X^m] = \frac{1}{B(a,b)} \int_0^1 x^m x^{a-1}(1-x)^{b-1} \mathrm{d}x = \frac{B(m+a,b)}{B(a,b)}$$

- In particular, the mean can be obtained as:

$$\mathbf{E}[X] = \frac{B(a+1,b)}{B(a,b)} = \frac{\Gamma(a+1)\Gamma(b)}{\Gamma(a+b+1)} \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} = \frac{a}{a+b}$$

Beta Random Variable

Moments

- The second raw moment can be obtained as:

$$\mathbf{E}[X^2] = \frac{B(a+2,b)}{B(a,b)} = \frac{\Gamma(a+2)\Gamma(b)}{\Gamma(a+b+2)} \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} = \frac{a(a+1)}{(a+b)(a+b+1)}$$

- The variance can be obtained as:

$$\mathrm{Var}(X) = \frac{ab}{(a+b)^2(a+b+1)}$$

- No simplified expression for the MGF.

$$M_X(t) = \sum_{m=0}^{\infty} \frac{B(m+a,b)}{B(a,b)} \frac{t^m}{m!}$$

Summary

Moments and MGFs of Continuous Random Variables

Distribution	Mean	Variance	MGF (within the RoC)
$\mathrm{Unif}(a,b)$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$	$\frac{e^{bt}-e^{at}}{t(b-a)}$, if $t \neq 0$; 1 if $t = 0$
$\exp(\lambda)$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$	$(1-\frac{t}{\lambda})^{-1}$
$\mathcal{N}(\mu,\sigma^2)$	μ	σ^2	$\exp(\mu t + \frac{\sigma^2 t^2}{2})$
$\mathrm{Gamma}(\alpha,\lambda)$	$\frac{\alpha}{\lambda}$	$\frac{\alpha}{\lambda^2}$	$(1-\frac{t}{\lambda})^{-\alpha}$
$\mathrm{Beta}(a,b)$	$\frac{a}{a+b}$	$\frac{ab}{(a+b)^2(a+b+1)}$	$\sum_{m=0}^{\infty} \frac{B(m+a,b)}{B(a,b)} \frac{t^m}{m!}$

An Interesting Case

Moments of a Pareto Random Variable

- If $X \sim \text{Pareto}(a, \lambda)$, then

$$\mathbf{E}[X^m] = \int_a^{\infty} \lambda x^m a^{\lambda} x^{-(\lambda+1)} \mathrm{d}x = \begin{cases} \dfrac{\lambda a^m}{\lambda - m} & \text{if } m < \lambda \\ \infty & \text{otherwise} \end{cases}$$

- For any $\lambda > 0$, only finite number of raw moments are finite.
- $M_X(t) = \sum_{n=0}^{\infty} \frac{t^n}{n!} \mathbf{E}[X^n] = \infty$ for any $t > 0$.