Moment Generating Function, Expectation and Variance of Ubiquitous Distributions with Applications in Decision Sciences: A Review

Kim-Hung Pho
Fractional Calculus, Optimization and Algebra Research Group,
Faculty of Mathematics and Statistics, Ton Duc Thang University,
Ho Chi Minh City, Vietnam

Thi Diem-Chinh Ho
Faculty of Mathematics and Statistics, University of Natural Sciences,
Ho Chi Minh City, Vietnam
and
General Faculty, Binh Duong Economics & Technology University,
Binh Duong City, Vietnam

Tuan-Kiet Tran
College of Science, Can Tho University, Viet Nam

Wing-Keung Wong**
Department of Finance, Fintech Center, and Big Data Research Center,
Asia University, Taiwan
and
Department of Medical Research, China Medical University Hospital, Taiwan
and
Department of Economics and Finance, the Hang Seng University of Hong Kong,
Hong Kong, China

Revised: June 2019

* The authors with to acknowledge a reviewer for helpful comments and suggestions.
** Corresponding author: wong@asia.edu.tw
Abstract

Statistics have been widely used in many disciplines including science, social science, business, engineering, and many others. One of the most important areas in statistics is to study the properties of distribution functions. To bridge the gap in the literature, this paper presents the theory of some important distribution functions and their moment generating functions. We introduce two approaches to derive the expectations and variances for all the distribution functions being studied in our paper and discuss the advantages and disadvantages of each approach in our paper. In addition, we display the diagrams of the probability mass function, probability density function, and cumulative distribution function for each distribution function being investigated in this paper. Furthermore, we review the applications of the theory discussed and developed in this paper to decision sciences.

Keywords: Moment Generating Function, Expectation, Variance, Distribution Functions

JEL: A12, G05, G35, O34
1 Introduction

The main task of Statistics, a branch of Mathematics, is to collect data, conduct analysis, make interpretation, evaluate and present the results, etc. Statistics have been widely used in many disciplines including science, social science, business, engineering, economics, finance, education, and many others. Hence, Statistics are very useful to research. Two ubiquitous statistical methods are utilized in data analysis consisting of descriptive statistics and statistical inference. The primary objective of descriptive statistics is to summarize data from a sample including the expectation and variance. Meanwhile, the main purpose of statistical inference is to draw conclusion from data analysis that is subject to random variation such as observational errors and sampling variation.

When studying statistics, we are often interested in knowing some properties of each distribution function, including probability density function (PDF), probability mass function (PMF), and cumulative distribution function (CDF). If one knows about the specific formulas of the distribution function, then one can explain these problems related to them easily. One of the main objectives in statistics is to know the distribution functions, and when working the functions, one usually cares about the expectation and variance.

It well known that there are two most important approaches to obtain the expectation and variance of distribution functions: one is based on the moment generating function, and another one is to compute the expectation and variance based on the definition of expectation and variance. Until today there have been several articles and books presented to this issue (see e.g. Tallis (1961), Cressie et al. (1981), Cain (1994), Ghosh et al. (2018), Wang et al. (2017), Yamamoto et al. (2018)). Nevertheless, most of books/papers only introduce to the result of the MGF, expectation, and variance of some distribution functions and they provide only a few ubiquitous distribution functions.

The distribution functions are mainly classify into two categories: discrete and continuous distributions. Discrete distributions include Bernoulli, binomial, negative binomial, Poisson, geometric, discrete uniform distributions, etc. Continuous distributions include Normal, log-normal, gamma, beta, uniform continuous distributions, etc. The distribution functions have been widely used in many disciplines. Readers may read Bakouch et al. (2014), Hajmohammadi et al. (2013), Jazi et al. (2010), Kibzun et al. (2013), Paisley et al. (2012), Cowpertwait (2010), Griggs et al. (2012), Ranodolph et al. (2012), Stickel et al. (2012), Zhang et al. (2015), and Zhao et al. (2017).
Therefore, it is important to have a paper presenting the detail about distribution functions and their moment generating function, expectation, and variance. To bridge the gap in the literature, this paper presents the theory of some important distribution functions and their moment generating functions. We introduce two approaches to derive the expectations and variances for all the distribution functions being studied in our paper. The first approach is to use the first and second derivatives of the moment generating function to calculate the expectation and variance of the corresponding distribution while the second approach is to use direct calculation. We discuss the advantages and disadvantages of each approach in our paper.

In addition, we display the diagrams of the probability mass function, probability density function, and cumulative distribution function for each distribution function being investigated in this paper. For each distribution, we show how to construct the corresponding regression models. We also discuss the difficulty when the outcome of the variables have much more zeros than expected and how to overcome the difficulty. In addition, we review the applications of the theory discussed and developed in this paper to decision sciences. Moreover, we have checked many books and papers. So far, we cannot find any book or paper present the detail of the theory discussed in our paper. Thus, we strongly believe though some or even all theories developed in our paper are well-known, our paper is the first paper discussing the details of the theory for some important distribution functions with applications, and thus, our paper could still have important some contributions to the literature.

The rest of the paper is structured as follows. We provide the definitions and discuss some basic properties of the MGF, expectation, and variance of a random variable in Section 2. We present some distribution functions and their MGF, expectation, and variance of some ubiquitous distributions in Sections 3 and 4, and discuss some properties of the distribution functions being studied in our paper in Section 5. We then review the applications of theories discussed in our paper in Decision Sciences in Section 6. The last section concludes.

2 Definitions and Basic Properties

In this section, we briefly discuss some of the most basic and important definitions and properties in statistics related to moment generating functions. We first state some definitions in the next subsection.
Definition 1. The collection $\mathcal{S}$ of subsets of $\Omega$ is called a $\sigma$-algebra if it satisfies the following properties:

(i) $\Omega \in \mathcal{S}$,

(ii) $E \in \mathcal{S} \Rightarrow E^c \in \mathcal{S}$, (closure under complementation)

where $E^c$ refers to the complement of $E$ with respect to $\Omega$.

(iii) $E_j \in \mathcal{S}, \ j = 1, 2, ... \Rightarrow \bigcup_{j=1}^{\infty} E_j \in \mathcal{S}$. (closure under countable union)

Definition 2. A probability measure, denoted by $P(\cdot)$, is a real-valued set function that is defined over a $\sigma$-algebra $\mathcal{S}$ and satisfies the following properties:

(i) $P(\Omega) = 1$;

(ii) $E \in \mathcal{S} \Rightarrow P(E) \geq 0$;

(iii) If $\{E_j\}$ is a countable collection of disjoint sets in $\mathcal{S}$, then $P\left(\bigcup_{j=1}^{\infty} E_j\right) = \sum_{j=1}^{\infty} P(E_j)$.

Definition 3. Given a sample space $\Omega$, a $\sigma$-algebra $\mathcal{S}$ associated with $\Omega$, and a probability measure $P(\cdot)$ defined over $\mathcal{S}$, we call the triplet $(\Omega, \mathcal{S}, P)$ a probability space.

Definition 4. A random variable on $(\Omega, \mathcal{S}, P)$ is a real-valued function defined over a sample space $\Omega$, denoted by $X(\omega)$ for $\omega \in \Omega$, such that for any real number $x$, $\{\omega | X(\omega) < x\} \in \mathcal{S}$. A random variable is always defined relative to some specific $\sigma$-algebra $\mathcal{S}$. It is discrete if its range forms a discrete(countable) set of real number. It is continuous if its range forms a continuous (uncountable) set of real numbers and the probability of $X$ equalling any single value in its range is zero.

We next state the definition of probability density function, probability mass function and cumulative distribution function as follows:

Definition 5. Let $X$ be a continuous random variable. The probability distribution function of $X$ is defined as $F_X(u) = \Pr(-\infty < X \leq u)$, with $F_X(\infty) = 1$. The probability density function (PDF) is $f(x) = \frac{dF(x)}{dx}$, with $f(x) \geq 0$, and $f(-\infty) = f(\infty) = 0$.

Definition 6. Suppose that $X : S \mapsto A$ for $A \subseteq \mathbb{R}$ is a discrete random variable defined on a sample space $S$. Then the probability mass function (PMF) $f_X : A \mapsto [0,1]$
for $X$ is defined as

$$f_X(x) = \Pr(X = x) = P\{s \in S : X(s) = x\}$$

**Definition 7.** The **cumulative distribution function** (CDF) of a real-valued random variable $X$, or just distribution function of $X$, evaluated at $x$, is the probability that $X$ will take a value less than or equal to $x$. The cumulative distribution function of a real-valued random variable $X$ is the function given by

$$F_X(x) = P(X \leq x)$$

Next, we state the definition of moment generating function as follows:

**Definition 8.** The **moment generating function** (MGF) of a random variable $X$ is a function $M_X : \mathbb{R} \to [0, \infty)$ defined as follows:

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

given that the expectation exists for $t$ in some neighborhoods of zero. The MGF of $X$ can be expressed as follows:

$$M_X(t) = \int_{-\infty}^{\infty} e^{tx} f_X(x) dx, \text{ if } X \text{ is continuous,} \tag{1}$$

$$M_X(t) = \sum_{x \in \chi} e^{tx} P(X = x), \text{ if } X \text{ is discrete.} \tag{2}$$

We turn to define the expectation and the variance of a random variable.

**Definition 9.** The **expectation** of a random variable $X$ is defined as follows:

$$E(X) = \int_{-\infty}^{\infty} x f_X(x) dx, \text{ if } X \text{ is continuous,} \tag{3}$$

$$E(X) = \sum_{x \in \chi} x P(X = x), \text{ if } X \text{ is discrete.} \tag{4}$$

**Definition 10.** The **variance** of a random variable $X$ is defined as follows:

$$\text{Var}(X) = \int_{-\infty}^{\infty} (x - E(X))^2 f_X(x) dx, \text{ if } X \text{ is continuous,} \tag{5}$$

$$\text{Var}(X) = \sum_{x \in \chi} (x - E(X))^2 P(X = x), \text{ if } X \text{ is discrete.} \tag{6}$$
In addition, for any \( X \), one can easily get

\[
\text{Var}(X) = E[(X - E(X))^2] = E(X^2) - [E(X)]^2.
\]

(7)

Using the above definitions, the following proposition can be obtained.

**Proposition 1.** If \( M_X(t) \) is the moment generating function of \( X \), then

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

where

\[
M_X^{(n)}(0) = \frac{d^n}{dt^n} M_X(t)|_{t=0}.
\]

From Proposition 1, one could easily obtain the following property

**Property 1.** The \( n \)-th moment will be equal to the \( n \)-th derivative of the MGF executed at \( t = 0 \), such that

\[
\frac{d^n}{dt^n} M_X(t)|_{t=0} = E(X^n e^{tX})|_{t=0} = E(X^n).
\]

(8)

In particular,

\[
\frac{d}{dt} M_X(t)|_{t=0} = E(Xe^{tX})|_{t=0} = E(X),
\]

(9)

\[
\frac{d^2}{dt^2} M_X(t)|_{t=0} = E(X^2 e^{tX})|_{t=0} = E(X^2),
\]

(10)

3 **Theory**

We discuss the distribution functions, moment generating functions, expectations, and variances of different discrete distributions in this section.

3.1 **Bernoulli distribution**

Before we state the probability mass function, moment generating function, expectation, and variance for the Bernoulli distribution, we first define Bernoulli random variable. Random variable \( X \) is Bernoulli random variable if \( X \) only takes two values, say 1 and 0 with probability \( p \) and \( q = 1 - p \), respectively, then we are ready to state to the probability mass function, moment generating function, expectation, and variance of the Bernoulli distribution:
The PMF and CDF of Bernoulli distribution are:

\[
    f(x; p) = p^x (1 - p)^{1-x}, \quad x \in \{0, 1\}
\]

\[
    F(x; p) = P(X \leq x) = \begin{cases} 
    0 & x < 0 \\
    1 - p & 0 \leq x < 1 \\
    1 & x \geq 1 
\end{cases}
\]

respectively. The diagram of PMF and CDF of Bernoulli distribution is described in Figure 1.
Figure 1: PDF and CDF of Bernoulli distribution
It has been seen that
\[ M_X(t) = E(e^{tX}) = \sum_{x=0,1} e^{tx}f(x;p) = \sum_{x=0,1} e^{tx}p^x(1-p)^{1-x} \]
\[ = e^0p^0(1-p)^{1-0} + e^1p^1(1-p)^{1-1} = (1-p) + pe^t \]
therefore
\[
\frac{d}{dt} M_X(t) = \frac{d}{dt}(1 - p + pe^t) = pe^t 
\]
and
\[
\frac{d^2}{dt^2} M_X(t) = \frac{d}{dt}(pe^t) = pe^t 
\]

\section*{Approach 1}

\[ E(X) = \frac{d}{dt} M_X(t) \Big|_{t=0} = (pe^t) \Big|_{t=0} = p \]
and
\[ E(X^2) = \frac{d^2}{dt^2} M_X(t) \Big|_{t=0} = (pe^t) \Big|_{t=0} = p \]

\section*{Approach 2}

\[ E(X) = \sum_{x=0,1} xf(x;p) = \sum_{x=0,1} xp^x(1-p)^{1-x} = 0 + p^1(1-p)^{1-1} = p \]
and
\[ E(X^2) = \sum_{x=0,1} x^2f(x;p) = \sum_{x=0,1} x^2p^x(1-p)^{1-x} = 0 + p^1(1-p)^{1-1} = p \]
Hence
\[ Var(X) = E(X^2) - [E(X)]^2 = p - p^2 = p(1-p) = pq \]

\section*{3.2 Binomial distribution}

The PMF and CDF of Binomial distribution can be written as follows
\[ f(x;p,n) = \binom{n}{x}p^x(1-p)^{n-x}, x \in \{0,1,...,n\} \]
\[ F(x;p,n) = \sum_{i=0}^{x} \binom{n}{i}p^i(1-p)^{n-i} \]
respectively, with \( n = 1 \) then binomial distribution becomes Bernoulli distribution. The diagram of PMF and CDF of binomial distribution is illustrated in Figure 2.
Figure 2: PDF and CDF of binomial distribution
It can be seen that

\[ M_X (t) = E (e^{tX}) = \sum_{x=0}^{n} e^{tx} f (x; p, n) = \sum_{x=0}^{n} e^{tx} \binom{n}{x} p^x (1-p)^{n-x} \]

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

thus

\[ \frac{d}{dt}M_X (t) = \frac{d}{dt}(1-p + pe^t)^n = n(1-p + pe^t)^{n-1} \frac{d}{dt}(1-p + pe^t) = npe^t(1-p + pe^t)^{n-1} \]

and

\[ \frac{d^2}{dt^2}M_X (t) = \frac{d}{dt} \left[ (npe^t)(1-p + pe^t)^{n-1} \right] = npe^t(1-p + pe^t)^{n-1} + npe^t \frac{d}{dt} \left[ (1-p + pe^t)^{n-1} \right] \]

\[ = npe^t(1-p + pe^t)^{n-1} + npe^t(n-1)(1-p + pe^t)^{n-2} \frac{d}{dt}(1-p + pe^t) \]

\[ = npe^t(1-p + pe^t)^{n-1} + npe^t(n-1)(1-p + pe^t)^{n-2}(pe^t) \]

**Approach 1**

\[ E (X) = \frac{d}{dt}M_X (t) \bigg|_{t=0} = \left[ npe^t(1-p + pe^t)^{n-1} \right] \bigg|_{t=0} = npe^0(1-p + pe^0)^{n-1} = np \]

and

\[ E(X^2) = \frac{d^2}{dt^2}M_X (t) \bigg|_{t=0} = npe^t(1-p + pe^t)^{n-1} + npe^t(n-1)(1-p + pe^t)^{n-2}(pe^t) \bigg|_{t=0} \]

\[ = npe^0(1-p + pe^0)^{n-1} + npe^0(n-1)(1-p + pe^0)^{n-2}(pe^0) \]

\[ = np + np(n-1)p = np + n(n-1)p^2 = np + (n^2 - n)p^2 \]

**Approach 2**

\[ E(X) = \sum_{x=0}^{n} xf (x; p) = \sum_{x=0}^{n} x \binom{n}{x} p^x (1-p)^{n-x} = \sum_{x=0}^{n} x \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x} \]

\[ = \sum_{x=1}^{n} \frac{n!}{(x-1)!(n-x)!} p^x (1-p)^{n-x} = np \sum_{x=1}^{n} \frac{(n-1)!}{(x-1)!(n-x)!} p^{x-1}(1-p)^{n-x} \]

For sake of simplicity, let \( y = x - 1 \) and \( m = n - 1 \) then

\[ E(X) = np \sum_{y=0}^{m} \frac{m!}{y!(m-y)!} p^y (1-p)^{m-y} = np(p + 1 - p)^m = np \]

and

\[ E[X(X-1)] = \sum_{x=0}^{n} x(x-1)f (x; p) = \sum_{x=0}^{n} x(x-1) \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x} \]

\[ = \sum_{x=2}^{n} \frac{n!}{(x-2)!(n-x)!} p^x (1-p)^{n-x} = n(n-1)p^2 \sum_{x=2}^{n} \frac{(n-2)!}{(x-2)!(n-x)!} p^{x-2}(1-p)^{n-x} \]
likewise, let \( t = x - 2 \) and \( k = n - 2 \) then

\[
E[X(X - 1)] = n(n - 1)p^2 \sum_{t=0}^{k} \frac{k!}{t!(k - t)!} p^t(1 - p)^{k-t} = n(n - 1)p^2(p + 1 - p)^m = n(n - 1)p^2
\]

and

\[
E(X^2) = E[X(X - 1)] + E(X) = n(n - 1)p^2 + np = np + (n^2 - n)p^2
\]

thus

\[
Var(X) = E(X^2) - [E(X)]^2 = np + (n^2 - n)p^2 - (np)^2
\]

\[
= np + n^2p^2 - np^2 - (np)^2 = np - np^2 = np(1 - p) = npq
\]

### 3.3 Negative Binomial distribution

\( Y \) = the number of failures before the \( r \)th success.

The PMF and CDF of Negative binomial distribution can be expressed as follows

\[
P(Y = y) = \binom{r + y - 1}{y} p^r(1 - p)^y
\]

\[
F(Y) = \sum_{k=0}^{y} \binom{r + k - 1}{k} p^r(1 - p)^k
\]

respectively, where \( y = 0, 1, ..., 0 < p < 1, r > 0 \) and \( r \) is an integer. The diagram of PMF and CDF of negative binomial distribution is provided in Figure 3.
Figure 3: PDF and CDF of negative binomial distribution
we have
\[ M_Y(t) = E(e^{tY}) = \sum_{y=0}^{\infty} e^{ty} \binom{r+y-1}{y} p^r(1-p)^y = p^r \sum_{y=0}^{\infty} \binom{r+y-1}{y} ((1-p)e^{t})^y \]
it will be know that
\[ \sum_{y=0}^{\infty} \binom{r+y-1}{y} x^y = (1-x)^{-r} \]
Therefore
\[ \sum_{y=0}^{\infty} \binom{r+y-1}{y} x^y(1-x)^r = 1 \]
Let \( x = (1-p)e^t \) then
\[ \sum_{y=0}^{\infty} \binom{r+y-1}{y} ((1-p)e^{t})^y (1-(1-p)e^{t})^r = 1 \]
for \((1-p)e^t < 1\).
thus
\[ \sum_{y=0}^{\infty} \binom{r+y-1}{y} ((1-p)e^{t})^x = (1-(1-p)e^{t})^{-r} = \frac{1}{(1-(1-p)e^{t})^r} \]
and
\[ M_Y(t) = E(e^{tY}) = p^r \frac{1}{(1-(1-p)e^{t})^r} = \left( \frac{p}{1-(1-p)e^{t}} \right)^r \]
for \( t < -\log (1-p) \). One has
\[ \frac{d}{dt} M_Y(t) = \left( \frac{p^r(1-(1-p)e^{t})}{(1-(1-p)e^{t})^r} \right)' = p^r.r.(1-p).e^{t}.(1-(1-p)e^{t})^{-r-1} = \frac{p^r.r.(1-p)e^{t}}{[1-(1-p)e^{t})^{2r}]^{r+1}} \]
and
\[ \frac{d^2}{dt^2} M_Y(t) = \left[ \frac{p^r.r.(1-p)e^{t}}{[1-(1-p)e^{t})^{r+1}]} \right]' = p^r.r.(1-p). \left[ \frac{e^{t}}{(1-(1-p)e^{t})^{r+1}} \right]' = p^r.r.(1-p).\left[ \frac{e^{t}}{(1-(1-p)e^{t})^{r+1}} \right] + e^{t}.(r+1).(1-(1-p)e^{t})^{r}.(1-p) \]
\[ = p^r.r.(1-p).\left[ \frac{e^{t}}{(1-(1-p)e^{t})^{r+1}} \right] + e^{t}.(r+1).(1-p) \]
\[ = p^r.r.(1-p).\left[ \frac{e^{t}}{(1-(1-p)e^{t})^{r+1}} \right] + e^{t}.(r+1).(1-p) \]
\[ = p^r.r.(1-p).\left[ \frac{e^{t}}{(1-(1-p)e^{t})^{r+1}} \right] + e^{t}.(r+1).(1-p) \]
\[ = p^r.r.(1-p).\left[ \frac{e^{t}}{(1-(1-p)e^{t})^{r+1}} \right] + e^{t}.(r+1).(1-p) \]
\[ = p^r.r.(1-p).\left[ \frac{e^{t}}{(1-(1-p)e^{t})^{r+1}} \right] + e^{t}.(r+1).(1-p) \]
\[ = p^r.r.(1-p).\left[ \frac{e^{t}}{(1-(1-p)e^{t})^{r+1}} \right] + e^{t}.(r+1).(1-p) \]
\[ = p^r.r.(1-p).\left[ \frac{e^{t}}{(1-(1-p)e^{t})^{r+1}} \right] + e^{t}.(r+1).(1-p) \]
Approach 1
\[ E(Y) = \left. \frac{d}{dt} M_Y(t) \right|_{t=0} = \left. \frac{p^r.r.(1-p)e^{t}}{[1-(1-p)e^{t})^{r+1}]^{r+1} \right|_{t=0} = \frac{r(1-p)}{p} \]
We have

\[ E(Y^2) = \frac{d^2}{dt^2} M_Y(t)|_{t=0} = p^r r(1-p) \left[ \frac{e^t}{(1-(1-p)e^t)^{r+1}} + \frac{e^t(r+1)(1-p)}{(1-(1-p)e^t)^{r+2}} \right]|_{t=0} \]

\[ = p^r r(1-p) \left[ \frac{1}{(1-(1-p)e^0)^{r+1}} + \frac{e^0(r+1)(1-p)}{(1-(1-p)e^0)^{r+2}} \right] \]

\[ = p^r r(1-p) \left[ \frac{1}{p^{r+1}} + \frac{(r+1)(1-p)}{p^{r+2}} \right] = \frac{r(1-p) + r(r+1)(1-p)^2}{p^2} \]

**Approach 2**

We have

\[ E(Y) = \sum_{y=0}^{\infty} y P(Y = y) = \sum_{y=0}^{\infty} y \left( \frac{r+y-1}{r} \right) p^r (1-p)^y = \sum_{y=1}^{\infty} y \frac{(y+r-1)!}{y!(r-1)!} p^r (1-p)^y \]

\[ = \sum_{y=1}^{\infty} \frac{(y+r-1)!}{y!(r-1)!} p^r (1-p)^y = \sum_{z=0}^{\infty} \frac{(z+r)!}{z!(r-1)!} p^r (1-p)^{z+1} \quad \text{(let } z = y - 1) \]

\[ = (1-p) \sum_{z=0}^{\infty} \frac{r(z+r)!}{z!(r-1)!} p^r (1-p)^z = r(1-p) \sum_{z=0}^{\infty} \frac{(z+r)!}{z!(r-1)!} p^r (1-p)^z \]

\[ = r \frac{1-p}{p} \sum_{y=0}^{\infty} \left( \frac{r+y-1}{y} \right) p^r (1-p)^y = \frac{r(1-p)}{p} \]

and

\[ E[Y(Y-1)] = \sum_{y=0}^{\infty} y(y-1) P(Y = y) = \sum_{y=0}^{\infty} y(y-1) \left( \frac{r+y-1}{y} \right) p^r (1-p)^y \]

\[ = \sum_{y=0}^{\infty} y(y-1) \frac{(y+r-1)!}{y!(r-1)!} p^r (1-p)^y = \sum_{y=2}^{\infty} \frac{(y+r-1)!}{y!(r-1)!} p^r (1-p)^y \]

\[ = \sum_{y=2}^{\infty} \frac{(z+r+1)!}{z!(r-1)!} p^r (1-p)^{z+2} \quad \text{(let } z = y - 2) \]

\[ = (1-p)^2 \sum_{z=0}^{\infty} \frac{r(r+1)(z+r+1)!}{z!(r+1)(r-1)!} p^r (1-p)^z \]

\[ = r(r+1)(1-p)^2 \sum_{y=0}^{\infty} \frac{(z+r+1)!}{z!(r+1)!} p^r (1-p)^z \]

\[ = r(r+1)(1-p)^2 \sum_{y=0}^{\infty} \left( \frac{(r+2)+z-1}{z} \right) p^{r+2}(1-p)^z = r(r+1) \frac{(1-p)^2}{p^2} \]

and

\[ E(Y^2) = E[Y(Y-1) + Y] = E[Y(Y-1)] + E(Y) = r(r+1) \frac{(1-p)^2}{p^2} + \frac{r(1-p)}{p} \]
\[ \text{Var}(Y) = E(Y^2) - [E(Y)]^2 = r(r + 1) \frac{(1-p)^2}{p^2} + r(1-p) \frac{(1-p)}{p} - \left[ r \frac{(1-p)}{p} \right]^2 \]
\[ = r(r + 1) \frac{(1-p)^2}{p^2} + r \frac{(1-p)p}{p^2} - r^2 \frac{(1-p)^2}{p^2} \]
\[ = r \frac{(1-p)}{p^2} \left[ (r + 1)(1-p) + p - r(1-p) \right] \]
\[ = r \frac{(1-p)}{p^2} \left( r - rp + 1 - p + p - r + rp \right) = r \frac{(1-p)}{p^2} \]

3.4 Poisson distribution

The PMF and CDF of Poisson distribution is given by
\[ P(X = x | \lambda) = \frac{e^{-\lambda} \lambda^x}{x!}, x = 0, 1, 2, ... \]
\[ F(x | \lambda) = \sum_{i=0}^{x} \frac{e^{-\lambda} \lambda^i}{i!} \]
respectively. The diagram of PMF and CDF of Poisson distribution is presented in Figure 4.
Figure 4: PDF and CDF of Poisson distribution
It has been seen that

\[ M_X(t) = E(e^{tX}) = \sum_{x=0}^{\infty} e^{tx} P(X = x | \lambda) = \sum_{x=0}^{\infty} e^{tx} \frac{e^{-\lambda} \lambda^x}{x!} = e^{-\lambda} \sum_{x=0}^{\infty} \frac{(\lambda e^t)^x}{x!} = e^{-\lambda} e^{\lambda e^t} = e^{\lambda(e^t-1)} \]

and

\[ \frac{d}{dt} M_X(t) = \frac{d}{dt} \left[ e^{\lambda(e^t-1)} \right] = e^{\lambda(e^t-1)} \frac{d}{dt} [\lambda (e^t - 1)] = \lambda e^t e^{\lambda(e^t-1)} \]

and

\[ \frac{d^2}{dt^2} M_X(t) = \frac{d}{dt} \left[ (\lambda e^t) e^{\lambda(e^t-1)} \right] = (\lambda e^t) e^{\lambda(e^t-1)} + (\lambda e^t) e^{\lambda(e^t-1)} \frac{d}{dt} [e^{\lambda(e^t-1)}] \\
= (\lambda e^t) e^{\lambda(e^t-1)} + (\lambda e^t) e^{\lambda(e^t-1)} (\lambda e^t) = \lambda e^t e^{\lambda(e^t-1)} (\lambda e^t + 1) \]

**Approach 1**

we have

\[ E(X) = \left. \frac{d}{dt} M_X(t) \right|_{t=0} = \left. (\lambda e^t e^{\lambda(e^t-1)}) \right|_{t=0} = \lambda e^0 e^{\lambda(0-1)} = \lambda \]

and

\[ E(X^2) = \left. \frac{d^2}{dt^2} M_X(t) \right|_{t=0} = \left. \left[ \lambda e^t e^{\lambda(e^t-1)} (\lambda e^t + 1) \right] \right|_{t=0} = (\lambda e^0) e^{\lambda(0-1)} (\lambda e^0 + 1) = \lambda + \lambda^2 \]

**Approach 2**

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

\[ = e^{-\lambda} \sum_{x=1}^{\infty} \frac{\lambda^x}{(x-1)!} = \lambda e^{-\lambda} \sum_{x=1}^{\infty} \frac{\lambda^{x-1}}{(x-1)!} = \lambda e^{-\lambda} e^{\lambda} = \lambda \]

Consider the expectation of random variable \(X(X - 1)\) we have

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

\[ = e^{-\lambda} \lambda^2 \sum_{x=2}^{\infty} \frac{\lambda^{x-2}}{(x-2)!} = \lambda^2 \]

Therefore \(E(X^2) = E[X(X - 1)] + E(X) = \lambda + \lambda^2\) and

\[ \text{Var}(X) = E(X^2) - [E(X)]^2 = \lambda + \lambda^2 - \lambda^2 = \lambda \]
3.5 Geometric distribution

The PMF and CDF of geometric distribution is written as follows

\[ P(X = x | p) = p(1 - p)^{x-1}, x = 1, 2, ... \]
\[ F(x | p) = 1 - (1 - p)^x \]

respectively. The diagram of PMF and CDF of geometric distribution is described in Figure 5.
Figure 5: PDF and CDF of geometric distribution
one has

\[ M_X(t) = E(e^{tX}) = \sum_{x=0}^{n} e^{tx} P(X = x | p) = \sum_{x=0}^{n} e^{tx} p(1-p)^{x-1} = \sum_{x=0}^{n} e^{tx} \frac{(1-p)^x}{(1-p)} \]

\[ = \frac{p}{1-p} \sum_{x=0}^{n} e^{tx}(1-p)^x = \frac{p}{1-p} \sum_{x=0}^{n} [(1-p) e^t]^x \]

\[ = \frac{p}{1-p} e^t (1-p) e^t = \frac{p}{1-p} e^t \]

Because \(|(1-p)e^t| < 1\).

\[ \frac{d}{dt} M_X(t) = \frac{d}{dt} \left[ \frac{p}{1-(1-p)e^t} \right] = p \frac{d}{dt} \left[ \frac{e^t}{1-(1-p)e^t} \right] \]

\[ = p \frac{e^t [1-(1-p)e^t] - e^t \frac{d}{dt} [1-(1-p)e^t]}{[1-(1-p)e^t]^2} \]

\[ = p \frac{e^t - (1-p)e^{2t} - e^t [- (1-p)e^t]}{[1-(1-p)e^t]^2} \]

\[ = p \frac{e^t - (1-p)e^{2t} + (1-p)e^{2t}}{[1-(1-p)e^t]^2} = p \frac{e^t}{[1-(1-p)e^t]^2} \]

and

\[ \frac{d^2}{dt^2} M_X(t) = \frac{d}{dt} \left[ \frac{e^t}{[1-(1-p)e^t]^2} \right] = \frac{e^t [1-(1-p)e^t]^2 - e^t \frac{d}{dt} [1-(1-p)e^t]^2}{[1-(1-p)e^t]^4} \]

\[ = p \frac{e^t [1-(1-p)e^t]^2 - e^t [1-(1-p)e^t] \frac{d}{dt} [1-(1-p)e^t]}{[1-(1-p)e^t]^4} \]

\[ = p \frac{e^t [1-(1-p)e^t]^2 - e^t [1-(1-p)e^t] [- (1-p)e^t]}{[1-(1-p)e^t]^4} \]

\[ = p \frac{e^t [1-(1-p)e^t] - e^t [1-(1-p)e^t]}{[1-(1-p)e^t]^4} \]

\[ = p \frac{e^t - (1-p)e^{2t} + 2(1-p)e^{2t}}{[1-(1-p)e^t]^4} = p \frac{e^t + (1-p)e^{2t}}{[1-(1-p)e^t]^3} \]

**Approach 1**

we have

\[ E(X) = \frac{d}{dt} M_X(t) \bigg|_{t=0} = \left( \frac{p}{1-(1-p)e^t} \right)^2 \bigg|_{t=0} = p \frac{e^0}{[1-(1-p)e^0]^2} = \frac{p}{1-p^2} = \frac{1}{p} \]

\[ E(X^2) = \frac{d^2}{dt^2} M_X(t) \bigg|_{t=0} = \left[ \frac{p e^t + (1-p)e^{2t}}{[1-(1-p)e^t]^3} \right] \bigg|_{t=0} = p \frac{e^0 + (1-p)e^0}{[1-(1-p)e^0]^3} \bigg|_{t=0} = p \frac{1 + (1-p)}{[1-(1-p)]^3} = \frac{2-p}{p^2} \]

**Approach 2**

\[ E(X) = \sum_{x=1}^{\infty} xP(X = x | p) = \sum_{x=1}^{\infty} x(1-p)^{x-1} = \frac{p}{1-p} \sum_{x=1}^{\infty} x(1-p)^x = \frac{p}{1-p} \frac{1-p}{p^2} = \frac{1}{p} \]
Fact, if we put \( a = (1 - p) \) then \(|a| < 1\) hence sequence \( \sum_{x=1}^{\infty} x(1 - p)^x = \sum_{x=1}^{\infty} xa^x = S \) and we have \( \frac{S}{a} = 1 + 2.a + 3.a^2 + \ldots \) to compute that sequence we need compute

\[
\int \frac{S}{a} \, da = a + a^2 + a^3 + \ldots = \frac{1}{1 - a}
\]
deduced

\[
\frac{S}{a} = \frac{1}{(1 - a)^2}
\]
so

\[
S = \frac{a}{(1 - a)^2}
\]

\[
E(X^2) = \sum_{x=1}^{\infty} x^2.P(X = x|p) = \sum_{x=1}^{\infty} x^2.p(1 - p)^{x-1} = \frac{p}{1 - p} \sum_{x=1}^{\infty} x^2.(1 - p)^x
\]
put as above we have \( \sum_{x=1}^{\infty} x^2.a^x = H \)
we will be calculate \( H = \frac{(2 - p)(1 - p)}{p^3} \)
because we have

\[
\frac{H}{a} = 1^2 + 2^2.a + 3^2.a^2 + 4^2.a^3 + \ldots
\]
deduced

\[
\int \frac{H}{a} \, da = \int (1^2 + 2^2.a + 3^2.a^2 + 4^2.a^3 + \ldots) \, da = 1.a + 2.a^2 + 3.a^3 + 4.a^4 + \ldots
\]
\[
= a.(1 + 2a + 3a^2 + 4a^3 + \ldots)
\]
we have

\[
\int (1 + 2a + 3a^2 + \ldots) \, da = a + a^2 + a^3 + \ldots = a \frac{1}{1 - a}
\]
therefore

\[
1 + 2a + 3a^2 + 4a^3 + \ldots = \left( \frac{a}{1 - a} \right)' = \frac{1}{(1 - a)^2}
\]
thus

\[
\int \frac{H}{a} \, da = \frac{a}{(1 - a)^2}
\]

23
and

\[
\frac{H}{a} = \frac{1 + a}{(1 - a)^3}
\]

we have

\[
H = \frac{a(1 + a)}{(1 - a)^3} = \frac{(1 - p)(2 - p)}{p^3}
\]

therefore

\[
E(X^2) = \frac{p}{1 - p} \left( \frac{(1 - p)(2 - p)}{p^3} \right) = \frac{2 - p}{p^2}
\]

hence

\[
Var(X) = E(X^2) - [E(X)]^2 = \frac{2 - p}{p^2} - \left( \frac{1}{p} \right)^2 = \frac{2 - p - 1}{p^2} = \frac{1 - p}{p^2}
\]

### 3.6 Discrete uniform distribution

Since X is a random variable with the general discrete uniform \((N_0, N_1)\) distribution

So PMF and CDF of X can be written as follows

\[
f(x) = \frac{1}{N_1 - N_0 + 1}, \quad x = N_0, N_0 + 1, ..., N_1
\]

\[
F(x) = P(X \leq x) = \frac{x - N_0 + 1}{N_1 - N_0 + 1}
\]

respectively. The diagram of PMF and CDF of discrete uniform distribution is illustrated in Figure 6.
Figure 6: PDF and CDF of discrete uniform distribution
one has
\[ M_X(t) = E(e^{tX}) = \sum_{x=N_0}^{N_1} e^{tx} f(x) = \frac{1}{N_1 - N_0 + 1} \sum_{x=N_0}^{N_1} e^{tx} \]
deduced
\[ \frac{d}{dt} M_X(t) = \left( \frac{1}{N_1 - N_0 + 1} \sum_{x=N_0}^{N_1} e^{tx} \right)' = \frac{1}{N_1 - N_0 + 1} \sum_{x=N_0}^{N_1} x.e^{tx} \]
and
\[ \frac{d^2}{dt^2} M_X(t) = \left( \frac{1}{N_1 - N_0 + 1} \sum_{x=N_0}^{N_1} x.e^{tx} \right)' \]
\[ = \frac{1}{N_1 - N_0 + 1} \left( N_1^2 e^{tN_0} + (N_0 + 1)^2 e^{t(N_0 + 1)} + \ldots + N_1^2 e^{tN_1} \right) \]
\[ = \frac{1}{N_1 - N_0 + 1} \sum_{x=N_0}^{N_1} x^2.e^{tx} \]

**Approach 1**

we have
\[ E(X) = \frac{d}{dt} M_X(t)|_{t=0} = \frac{1}{N_1 - N_0 + 1} \sum_{x=N_0}^{N_1} x.e^{tx}|_{t=0} = \frac{1}{N_1 - N_0 + 1} \sum_{x=N_0}^{N_1} x \]
\[ = \frac{1}{N_1 - N_0 + 1} (N_1 - N_0 + 1). \frac{N_1 + N_0}{2} = \frac{N_0 + N_1}{2} \]
and
\[ E(X^2) = \frac{d^2}{dt^2} M_X(t)|_{t=0} = \frac{1}{N_1 - N_0 + 1} \sum_{x=N_0}^{N_1} x^2.e^{tx}|_{t=0} = \frac{1}{N_1 - N_0 + 1} \sum_{x=N_0}^{N_1} x^2 \]
\[ = \frac{1}{N_1 - N_0 + 1} \left( \sum_{x=1}^{N_1} x^2 - \sum_{x=1}^{N_0-1} x^2 \right) \]
\[ = \frac{1}{N_1 - N_0 + 1} \left( \frac{N_1(N_1+1)(2N_1 + 1)}{6} - \frac{(N_0 - 1)N_0(2N_0 - 1)}{6} \right) \]
\[ = \frac{1}{N_1 - N_0 + 1} \left( \frac{N_1^2 + N_1}{2N_1 + 1} - \frac{N_0^2 - N_0}{2N_0 - 1} \right) \]
\[ = \frac{1}{N_1 - N_0 + 1} \left( \frac{2N_1^3 - 2N_1^3 + 3N_1^2 + 3N_0^2 - N_1 - N_0}{6} \right) \]
\[ = \frac{2(N_1 - N_0)(N_1^2 + N_1N_0 + N_0^2)}{6(N_1 - N_0 + 1)} + \frac{N_1 - N_0 + 3(N_1^2 + N_0^2)}{6(N_1 - N_0 + 1)} \]
\[ = \frac{(N_1 - N_0)[2(N_1^2 + N_1N_0 + N_0^2) + 1]}{6(N_1 - N_0 + 1)} + \frac{3(N_1^2 + N_0^2)}{6(N_1 - N_0 + 1)} \]
\[ = \frac{(N_1 - N_0)[4(N_1^2 + N_1N_0 + N_0^2) + 1]}{12(N_1 - N_0 + 1)} + \frac{6(N_1^2 + N_0^2)}{12(N_1 - N_0 + 1)} \]
Approach 2

we have

\[
E(X) = \sum_{x=N_0}^{x=N_1} \frac{1}{N_1 - N_0 + 1} \sum_{x=0}^{x=N_1} x = \frac{1}{N_1 - N_0 + 1} \sum_{x=1}^{x=N_1} x - \sum_{x=1}^{x=N_0-1} x
\]

\[
E(X) = \frac{1}{N_1 - N_0 + 1} \left( \frac{(N_1 + 1)}{2} \right) (N_1 + N_0 + 1 + N_1 + N_0) = \frac{(N_1 - N_0)}{2} (N_1 + N_0 + 1)
\]

\[
E(X^2) = \sum_{x=N_0}^{x=N_1} x^2 = \frac{1}{N_1 - N_0 + 1} \sum_{x=0}^{x=N_1} x^2 = \frac{1}{N_1 - N_0 + 1} \left( \sum_{x=1}^{x=N_1} x^2 - \sum_{x=1}^{x=N_0-1} x^2 \right)
\]

\[
E(X^2) = \frac{1}{N_1 - N_0 + 1} \left( \frac{(N_1 (N_1 + 1)(2N_1 + 1)}{6} - (N_0 - 1) N_0 (2N_0 - 1) \right)
\]

\[
E(X^2) = \frac{1}{N_1 - N_0 + 1} \left( \frac{(N_1^2 + N_1)(2N_1 + 1)}{6} - (N_1^2 - N_0) (2N_1 - 1) \right)
\]

\[
E(X^2) = \frac{2(N_1 - N_0)(N_1^2 + N_1 N_0 + N_0^2)}{6(N_1 - N_0 + 1)} + \frac{N_1 - N_0 + 3(N_1^2 + N_0^2)}{6(N_1 - N_0 + 1)}
\]

\[
E(X^2) = \frac{(N_1 - N_0)[2(N_1^2 + N_1 N_0 + N_0^2) + 1]}{6(N_1 - N_0 + 1)} + \frac{3(N_1^2 + N_0^2)}{6(N_1 - N_0 + 1)}
\]

\[
E(X^2) = \frac{(N_1 - N_0)[2(N_1 N_0 + N_0^2) + 1]}{6(N_1 - N_0 + 1)} + \frac{3(N_1^2 + N_0^2)}{6(N_1 - N_0 + 1)}
\]

\[
E(X^2) = \frac{(N_1 - N_0)[4(N_1 N_0 + N_0^2) + 2]}{12(N_1 - N_0 + 1)} + \frac{6(N_1^2 + N_0^2)}{12(N_1 - N_0 + 1)} - \frac{(N_1 + N_0)^2}{4}
\]

\[
\text{Var} = E(X^2) - [E(X)]^2 = \frac{(N_1 - N_0)[4(N_1 N_0 + N_0^2) + 2]}{12(N_1 - N_0 + 1)} + \frac{6(N_1^2 + N_0^2)}{12(N_1 - N_0 + 1)} - \frac{(N_1 + N_0)^2}{4}
\]

\[
\text{Var} = \frac{(N_1 - N_0)[4(N_1 N_0 + N_0^2) + 2]}{12(N_1 - N_0 + 1)} + \frac{6(N_1^2 + N_0^2)}{12(N_1 - N_0 + 1)} - \frac{(N_1 + N_0)^2}{4}
\]

\[
\text{Var} = \frac{(N_1 - N_0)[4(N_1 N_0 + N_0^2) + 2]}{12(N_1 - N_0 + 1)} + \frac{6(N_1^2 + N_0^2)}{12(N_1 - N_0 + 1)} - \frac{3(N_1 - N_0 + 1)(N_1^2 + N_0^2) + 2N_1 N_0}{12(N_1 - N_0 + 1)}
\]

\[
\text{Var} = \frac{(N_1 - N_0)(N_1 - N_0 + 1)(N_1 - N_0 + 2)}{12(N_1 - N_0 + 1)} = \frac{(N_1 - N_0)(N_1 - N_0 + 2)}{12}
\]
4 Moment generating function, expectation and variance of continuous distributions

4.1 Normal distribution

The PDF and CDF of normal distribution is given by

\[
f(x) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, -\infty < x < \infty
\]

\[
F(x) = \frac{1}{2} \left[ 1 + \text{erf} \left( \frac{x - \mu}{\sigma \sqrt{2}} \right) \right]
\]

respectively, where \( \text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt \). The diagram of PDF and CDF of normal distribution is provided in Figure 7.
Figure 7: PDF and CDF of the standard normal distribution
It has been seen that

\[ M_X(t) = E(e^{tX}) = \int_{-\infty}^{\infty} e^{tx} f(x) \, dx = \int_{-\infty}^{\infty} e^{tx} \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \, dx \]

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

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

one has

\[ x^2 - 2\mu x - 2\sigma^2 t x + \mu^2 = x^2 - 2(\mu + \sigma^2 t) x + (\mu + \sigma^2 t)^2 + \mu^2 - (\mu + \sigma^2 t)^2 \]

\[ = [x - (\mu + \sigma^2 t)]^2 + \mu^2 - (\mu^2 + 2\mu \sigma^2 t + (\sigma^2 t)^2) \]

\[ = [x - (\mu + \sigma^2 t)]^2 - (2\mu \sigma^2 t + \sigma^4 t^2) \]

\[ = [x - (\mu + \sigma^2 t)]^2 - 2\sigma^2 \left( \mu + \frac{\sigma^2 t^2}{2} \right) \]

thus

\[ M_X(t) = E(e^{tX}) = \frac{1}{\sqrt{2\pi\sigma}} \int_{-\infty}^{\infty} e^{-\frac{(x^2 - 2\mu x - 2\sigma^2 t x + \mu^2)}{2\sigma^2}} \, dx = \frac{1}{\sqrt{2\pi\sigma}} \int_{-\infty}^{\infty} \frac{[x - (\mu + \sigma^2 t)]^2 + 2\sigma^2 (\mu + \frac{\sigma^2 t^2}{2})}{2\sigma^2} \, dx \]

\[ = \frac{1}{\sqrt{2\pi\sigma}} \int_{-\infty}^{\infty} e^{-\frac{(x^2 - 2\mu x - 2\sigma^2 t x + \mu^2)}{2\sigma^2}} \, dx = \frac{1}{\sqrt{2\pi\sigma}} \int_{-\infty}^{\infty} e^{-\frac{(x^2 - 2\mu x + (\sigma^4 t^2 + \sigma^2 t^2))}{2\sigma^2}} \, dx = e^{(\mu + \frac{\sigma^2 t^2}{2})} \]

Because, let \( z = x - (\mu + \sigma^2 t) \) \( \Rightarrow dz = dx \), then

\[ \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(z^2 - 2\mu z + (\sigma^2 t^2))}{2\sigma^2}} \, dz = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{z^2}{2\sigma^2}} \, dz = 1 \]

we have

\[ \frac{d}{dt} M_X(t) = \frac{d}{dt} \left( \mu + \sigma^2 t \right) e^{\left( \mu + \frac{\sigma^2 t^2}{2} \right)} = (\mu + \sigma^2 t) e^{\left( \mu + \frac{\sigma^2 t^2}{2} \right)} \]

and

\[ \frac{d^2}{dt^2} M_X(t) = \frac{d}{dt} \left[ (\mu + \sigma^2 t) e^{\left( \mu + \frac{\sigma^2 t^2}{2} \right)} \right] = (\mu + \sigma^2 t) \frac{d}{dt} \left[ e^{\left( \mu + \frac{\sigma^2 t^2}{2} \right)} \right] + e^{\left( \mu + \frac{\sigma^2 t^2}{2} \right)} \frac{d}{dt} (\mu + \sigma^2 t) \]

\[ = (\mu + \sigma^2 t) (\mu + \sigma^2 t) e^{\left( \mu + \frac{\sigma^2 t^2}{2} \right)} + e^{\left( \mu + \frac{\sigma^2 t^2}{2} \right)} (\sigma^2) = e^{\left( \mu + \frac{\sigma^2 t^2}{2} \right)} \left[ (\mu + \sigma^2 t)^2 + \sigma^2 \right] \]

**Approach 1**

It can be seen that

\[ E(X) = \frac{d}{dt} M_X(t) |_{t=0} = \mu \]

\[ E(X^2) = \frac{d^2}{dt^2} M_X(t) |_{t=0} = \mu^2 + \sigma^2 \]
Approach 2

It can be observed that

\[ E(X) = \int_{-\infty}^{\infty} x.f(x)dx = \int_{-\infty}^{\infty} x. \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \frac{1}{\sigma \sqrt{2\pi}} \int_{-\infty}^{\infty} x.e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \]

if we put \( t = \frac{x - \mu}{\sigma} \) deduced \( dt = \frac{1}{\sigma} dx \) and we alway have an equation

\[ \int_{-\infty}^{\infty} f(x)dx = 1 \]

detail

\[ \int_{-\infty}^{\infty} \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} = 1 \]

or

\[ \int_{-\infty}^{\infty} \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{t^2}{2}} = 1 \]

from that the expectation became

\[
E(X) = \frac{1}{\sigma \sqrt{2\pi}} \int_{-\infty}^{\infty} (\sigma t + \mu)e^{-\frac{t^2}{2}} \sigma dt = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (\sigma t + \mu)e^{-\frac{t^2}{2}} dt \\
= \frac{\sigma}{\sqrt{2\pi}} \int_{-\infty}^{\infty} t.e^{-\frac{t^2}{2}} dt + \frac{\mu}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{t^2}{2}} dt = \frac{\sigma}{\sqrt{2\pi}} \int_{-\infty}^{\infty} t.e^{-\frac{t^2}{2}} dt + \mu.1 \\
= -\frac{\sigma}{\sqrt{2\pi}}. \int e^{-\frac{t^2}{2}} (-\frac{t^2}{2}) + \mu = -\frac{\sigma}{\sqrt{2\pi}}. e^{-\frac{t^2}{2}} |_{-\infty}^{\infty} + \mu = -\frac{\sigma}{\sqrt{2\pi}}.0 + \mu = \mu
\]

and

\[
E(X^2) = \int_{-\infty}^{\infty} x^2 f(x)dx = \frac{1}{\sigma \sqrt{2\pi}} \int_{-\infty}^{\infty} x^2.e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (\sigma t + \mu)^2e^{-\frac{t^2}{2}} dt \\
= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (\sigma^2.t^2.e^{-\frac{t^2}{2}} + 2\mu\sigma t.e^{-\frac{t^2}{2}} + \mu^2e^{-\frac{t^2}{2}}) dt \\
= \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} t^2.e^{-\frac{t^2}{2}} dt + \frac{\mu\sigma}{\sqrt{2\pi}} \int_{-\infty}^{\infty} t.e^{-\frac{t^2}{2}} dt + \mu^2. \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt
\]

Considering the following integral

\[ \int_{-\infty}^{\infty} t^2.e^{-\frac{t^2}{2}} \]

31
using partial integral and let
\[ u = t; \quad dv = t \cdot e^{-\frac{t^2}{2}} dt \]

then
\[ du = dt; \quad v = \int t \cdot e^{-\frac{t^2}{2}} dt = -e^{-\frac{t^2}{2}} \]

the integral need compute equal to
\[ \int_{-\infty}^{\infty} t^2 \cdot e^{-\frac{t^2}{2}} dt = -e^{-\frac{t^2}{2}} \big|_{-\infty}^{\infty} + \int_{-\infty}^{\infty} e^{-\frac{t^2}{2}} dt \]

In addition, one need to find the limit of \( t \cdot e^{-\frac{t^2}{2}} \)

by the L'Hospital rules we have the limit
\[ \lim_{t \to \infty} t \cdot e^{-\frac{t^2}{2}} = \lim_{t \to \infty} \frac{t}{e^{\frac{t^2}{2}}} = \lim_{t \to \infty} \frac{t'}{(e^{\frac{t^2}{2}})} = \lim_{t \to \infty} \frac{1}{t \cdot e^{\frac{t^2}{2}}} = 0 \]

therefore
\[ \lim_{t \to -\infty} t \cdot e^{-\frac{t^2}{2}} = 0 \]

and finally
\[ \int_{-\infty}^{\infty} e^{-\frac{t^2}{2}} dt = \sqrt{2\pi} \]

then
\[ E(X^2) = \frac{\sigma^2}{\sqrt{2\pi}} \cdot \sqrt{2\pi} + \frac{\mu \sigma}{\sqrt{2\pi}} \cdot 0 + \mu^2 \cdot \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt = \sigma^2 + \mu^2 \]

so
\[ Var(X) = E(X^2) - [E(X)]^2 = \sigma^2 + \mu^2 - \mu^2 = \sigma^2 \]

### 4.2 Log-normal distribution

The PDF and CDF of log-normal distribution can be expressed as follows
\[ f(x) = \frac{1}{\sqrt{2\pi} \sigma} x e^{-\frac{(\ln x - \mu)^2}{2\sigma^2}}, \quad 0 < x < \infty, -\infty < \mu < \infty, \sigma > 0 \]
\[ F(x) = \frac{1}{2} + \frac{1}{2} \left[ 1 + \text{erf} \left( \frac{\ln x - \mu}{\sigma \sqrt{2}} \right) \right] \]

respectively, where \( \text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt \). The diagram of PDF and CDF of log-normal distribution is described in Figure 8. If \( X \) follows log-normal distribution, then \( Y = \ln X \sim N(\mu, \sigma^2) \).
Figure 8: PDF and CDF of log-normal distribution
one has
\[ f(y) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(y-\mu)^2}{2\sigma^2}}, \]
\[ M_Y(t) = e^{\mu t + \frac{\sigma^2 t^2}{2}}, \quad -\infty < y < \infty, \quad -\infty < \mu < \infty, \quad \sigma > 0 \]

**Approach 1**
we have
\[ E(X) = E(e^{\ln(X)}) = E(e^Y) = M_Y(1) = \left( e^{\mu t + \frac{\sigma^2 t^2}{2}} \right) \Big|_{t=1} = e^{\mu + \frac{\sigma^2}{2}} \]
and
\[ E(X^2) = E(e^{2\ln(X)}) = E(e^{2Y}) = M_Y(2) = \left( e^{\mu t + \frac{\sigma^2 t^2}{2}} \right) \Big|_{t=2} = e^{2\mu + 2\sigma^2} \]

**Approach 2**
we have
\[ E(X) = \int_0^\infty x f(x) dx = \int_0^\infty x \cdot \frac{1}{\sigma \sqrt{2\pi}} \cdot \frac{1}{x} e^{-\frac{(\ln x - \mu)^2}{2\sigma^2}} dx = \int_0^\infty \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(\ln x - \mu)^2}{2\sigma^2}} dx \]
it will be known that
\[ \frac{1}{b \sqrt{2\pi}} \int_{-\infty}^\infty e^{-\frac{(x-b)^2}{2\sigma^2}} = 1 \]
put \( y = \ln x - \mu \) then \( dx = e^{y+\mu} dy \)
\[ E(X) = \int_{-\infty}^\infty \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{y^2}{2\sigma^2}} e^{y+\mu} dy = e^\mu \cdot \frac{1}{\sigma \sqrt{2\pi}} \int_{-\infty}^\infty e^{-\frac{y^2}{2\sigma^2} + y} dy \]
have
\[ -\frac{1}{2\sigma^2} y^2 + y = -\frac{1}{2\sigma^2} [(y - \sigma^2)^2 - \sigma^4] \]
deduced
\[ E(X) = e^\mu \cdot \frac{1}{\sigma \sqrt{2\pi}} \int_{-\infty}^\infty e^{-\frac{1}{2\sigma^2} (y-\sigma^2)^2} \cdot \frac{dy}{\sigma \sqrt{2\pi}} = e^\mu \cdot \frac{1}{\sigma \sqrt{2\pi}} \int_{-\infty}^\infty e^{-\frac{1}{2\sigma^2} (y-\sigma^2)^2} dy = e^{\mu + \frac{\sigma^2}{2}} \]
we have
\[ E(X^2) = \int_0^\infty x^2 f(x) dx = \int_0^\infty x^2 \cdot \frac{1}{\sigma \sqrt{2\pi}} \cdot \frac{1}{x} e^{-\frac{(\ln x - \mu)^2}{2\sigma^2}} dx = \frac{1}{\sigma \sqrt{2\pi}} \int_0^\infty x \cdot e^{-\frac{(\ln x - \mu)^2}{2\sigma^2}} dx \]
put \( y \) as above deduced
\[
E(X^2) = \frac{1}{\sigma \sqrt{2\pi}} \int_{-\infty}^{\infty} e^{y+\mu} e^{-\frac{y^2}{2\sigma^2}} e^{y+\mu} dy = \frac{e^{2\mu}}{\sigma \sqrt{2\pi}} \int_{-\infty}^{\infty} e^{2y} e^{-\frac{y^2}{2\sigma^2}} dy = \frac{e^{2\mu}}{\sigma \sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{2y^2}{2\sigma^2} + 2y} dy
\]
\[
\text{fact}
\]
\[
-\frac{y^2}{2\sigma^2} + 2y = -\frac{1}{2\sigma^2} (y - 2\sigma^2)^2 + 2\sigma^2
\]
then
\[
E(X^2) = e^{2\mu + 2\sigma^2} \left( \frac{1}{\sigma \sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2\sigma^2} (y - 2\sigma^2)^2} dy \right) = e^{2\mu + 2\sigma^2} \cdot 1 = e^{2\mu + 2\sigma^2}
\]
and
\[
\text{Var} (X) = E(X^2) - [E(X)]^2 = e^{2\mu + 2\sigma^2} - \left( e^{\mu + \frac{\sigma^2}{2}} \right)^2 = e^{2\mu + 2\sigma^2} - e^{2\mu + \sigma^2} = e^{2\mu + \sigma^2} \left( e^{\sigma^2} - 1 \right)
\]

### 4.3 Gamma distribution

The PDF and CDF of gamma distribution can be written as follows
\[
f(x; \alpha, \beta) = \frac{1}{\Gamma(\alpha) \beta^\alpha} x^{\alpha-1} e^{-x/\beta}, 0 < x < \infty; 0 < \alpha, \beta
\]
\[
F(x; \alpha, \beta) = \frac{1}{\Gamma(\alpha) \beta^\alpha} \int_0^x t^{\alpha-1} e^{-t/\beta} dt
\]
respectively, the diagram of PDF and CDF of gamma distribution is provided in Figure 9.
Figure 9: PDF and CDF of gamma distribution
It can be seen that

\[ M_X(t) = E(e^{tX}) = \int_{-\infty}^{\infty} e^{tx} f(x; \alpha, \beta) dx = \int_{0}^{\infty} \frac{1}{\Gamma(\alpha) \beta^\alpha} x^{\alpha-1} e^{-x/\beta} dx \]

\[ = \frac{1}{\Gamma(\alpha) \beta^\alpha} \int_{0}^{\infty} x^{\alpha-1} e^{-x/\beta} dx = \frac{1}{\Gamma(\alpha) \beta^\alpha} \int_{0}^{\infty} \int_{0}^{\alpha} \frac{1}{\beta^\alpha} \Gamma(\alpha)\beta^\alpha x^{\alpha-1} e^{-x/\beta} dx \]

\[ = \frac{1}{\Gamma(\alpha) \beta^\alpha} \int_{0}^{\infty} \[ \frac{1}{\beta^{\alpha}} \int_{0}^{\alpha} x^{\alpha-1} e^{-x/\beta} \] \mathcal{d}x = \frac{1}{\Gamma(\alpha) \beta^\alpha} \int_{0}^{\infty} \frac{1}{\beta^{\alpha}} \int_{0}^{\alpha} x^{\alpha-1} e^{-x/\beta} \mathcal{d}x \]

\[ = \frac{1}{\Gamma(\alpha) \beta^\alpha} \left( \frac{\beta}{1-t\beta} \right)^{\alpha} \int_{0}^{\infty} x^{\alpha-1} e^{-x/\beta} \mathcal{d}x \]

\[ = \frac{1}{\Gamma(\alpha) \beta^\alpha} \left( \frac{\beta}{1-t\beta} \right)^{\alpha} \int_{0}^{\infty} x^{\alpha-1} e^{-x/\beta} \mathcal{d}x = \frac{1}{\Gamma(\alpha) \beta^\alpha} \Gamma(\alpha) = \frac{1}{(1-t\beta)^\alpha} \]

hence

\[ \frac{d}{dt} M_X(t) = \frac{d}{dt} \left( \frac{1}{1-t\beta} \right)^\alpha = \frac{-\alpha (-\beta) (1-t\beta)^{\alpha-1}}{(1-t\beta)^{2\alpha}} = \frac{\alpha \beta}{(1-t\beta)^{\alpha+1}} \]

and

\[ \frac{d^2}{dt^2} M_X(t) = \frac{d}{dt} \left( \frac{\alpha \beta}{(1-t\beta)^{\alpha+1}} \right) = \alpha \beta \frac{d}{dt} \left( \frac{1}{(1-t\beta)^{\alpha+1}} \right) = \frac{-\alpha \beta (\alpha+1) (1-t\beta)^{\alpha} (-\beta)}{(1-t\beta)^{2\alpha+2}} = \frac{\alpha (\alpha+1) \beta^2}{(1-t\beta)^{\alpha+2}} \]

**Approach 1**

Therefore, one can obtain

\[ E(X) = \frac{d}{dt} M_X(t) \bigg|_{t=0} = \alpha \beta \]

\[ E(X^2) = \frac{d^2}{dt^2} M_X(t) \bigg|_{t=0} = \alpha (\alpha+1) \beta^2 \]

**Approach 2**

It has been seen that

\[ E(X) = \int_{0}^{\infty} x f(x; \alpha, \beta) dx = \int_{0}^{\infty} \frac{1}{\Gamma(\alpha) \beta^\alpha} x^{\alpha} e^{-x/\beta} dx \]

\[ = \frac{\beta \Gamma(\alpha+1)}{\Gamma(\alpha)} \left( \int_{0}^{\infty} \frac{x^{\alpha+1} e^{-x/\beta}}{\Gamma(\alpha+1) \beta^{\alpha+1}} dx \right) = \frac{\beta \Gamma(\alpha+1)}{\Gamma(\alpha)} \cdot 1 = \frac{\beta \Gamma(\alpha+1)}{\Gamma(\alpha)} = \beta \alpha \]
because fixed $\beta$ and increase $\alpha$ one unit with fact

$$\int_0^\infty f(x; \alpha, \beta)dx = 1$$

$$\int_0^\infty \frac{x^\alpha e^{-\frac{x}{\beta}}}{\Gamma(\alpha) \beta^\alpha}dx = 1$$

or

$$\int_0^\infty \frac{x^{\alpha+1} e^{-\frac{x}{\beta}}}{\Gamma(\alpha + 1) \beta^{\alpha+1}}dx = 1$$

and

$$E(X^2) = \int_0^\infty x^2 f(x; \alpha, \beta)dx = \int_0^\infty \frac{x^{\alpha+2} e^{-\frac{x}{\beta}}}{\Gamma(\alpha) \beta^\alpha}dx$$

respectively as above we fixed $\beta$ and increase $\alpha$ two units have

$$\int_0^\infty \frac{x^{\alpha+2} e^{-\frac{x}{\beta}}}{\Gamma(\alpha + 2) \beta^{\alpha+2}}dx = 1$$

so

$$E(X^2) = \frac{\Gamma(\alpha + 2) \beta^{\alpha+2}}{\Gamma(\alpha) \beta^\alpha} \left( \int_0^\infty \frac{x^{\alpha+2} e^{-\frac{x}{\beta}}}{\Gamma(\alpha + 2) \beta^{\alpha+2}}dx \right) = \frac{\Gamma(\alpha + 2)}{\Gamma(\alpha)} \beta^2 \cdot 1 = \beta^2 \cdot \alpha \cdot (\alpha + 1)$$

and

$$\text{Var}(X) = E(X^2) - [E(X)]^2 = \alpha (\alpha + 1) \beta^2 - (\alpha \beta)^2 = \alpha \beta^2 (\alpha + 1 - \alpha) = \alpha \beta^2$$

### 4.4 Beta distribution

The PDF and CDF of beta distribution is written by

$$f(x | \alpha, \beta) = \frac{1}{B(\alpha, \beta)} x^{\alpha-1} (1-x)^{\beta-1}, 0 < x < 1, \alpha > 0, \beta > 0$$

$$F(x; \alpha, \beta) = \frac{1}{B(\alpha, \beta)} \int_0^x x^{\alpha-1} (1-x)^{\beta-1}dx$$

respectively, where

$$B(\alpha, \beta) = \int_0^1 x^{\alpha-1} (1-x)^{\beta-1}dx, B(\alpha, \beta) = \frac{\Gamma(\alpha) \Gamma(\beta)}{\Gamma(\alpha + \beta)}, \Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x}dx$$

The diagram of PDF and CDF of beta distribution is presented in Figure 10.
Figure 10: PDF and CDF of beta distribution
It can be observed that

\[
M_X(t) = E(e^{tX}) = \int_0^1 e^x \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha, \beta)} dx = \frac{1}{B(\alpha, \beta)} \int_0^1 \left( \sum_{k=0}^{\infty} \frac{(tx)^k}{k!} \right) x^{\alpha-1}(1-x)^{\beta-1} dx
\]

\[
= \frac{1}{B(\alpha, \beta)} \sum_{k=0}^{\infty} \frac{t^k}{k!} \int_0^1 x^{\alpha+k-1}(1-x)^{\beta-1} dx = \frac{1}{B(\alpha, \beta)} \sum_{k=0}^{\infty} \frac{t^k}{k!} B(\alpha+k, \beta)
\]

\[
= \sum_{k=1}^{\infty} \frac{t^k}{k!} B(\alpha+k, \beta) + B(\alpha+0, \beta) \frac{t^0}{0!} = 1 + \sum_{k=1}^{\infty} \frac{\Gamma(\alpha+k) \cdot B(\alpha+\beta+k) \cdot B(\alpha+\beta)}{\Gamma(\alpha+\beta+k) \cdot \Gamma(\alpha) \cdot \Gamma(\alpha+\beta)} \cdot t^k \frac{1}{k!}
\]

\[
= 1 + \sum_{k=1}^{\infty} \left( \frac{\Gamma(\alpha+k) \cdot \Gamma(\alpha+\beta+k) \cdot \Gamma(\alpha)}{\Gamma(\alpha+\beta+k) \cdot \Gamma(\alpha+\beta)} \right) \frac{t^k}{k!}
\]

\[
= 1 + \sum_{k=1}^{\infty} \left( \prod_{r=0}^{k} \frac{\alpha+r}{\alpha+\beta+r} \right) \frac{t^k}{k!}
\]

hence

\[
\frac{d}{dt} M_X(t) = \left( 1 + \sum_{k=1}^{\infty} \frac{t^k}{k!} B(\alpha+k, \beta) \right) \frac{1}{B(\alpha, \beta)} = \left( \frac{1}{1!} \right) \cdot B(\alpha+1, \beta) + \left( \sum_{k=1}^{\infty} \frac{t^k}{k!} B(\alpha+k, \beta) \right) \frac{1}{B(\alpha, \beta)}
\]

\[
= \frac{\alpha}{\alpha+\beta} + \sum_{k=0}^{\infty} \frac{B(\alpha+k+1, \beta)}{B(\alpha, \beta)} \frac{t^k}{k!}
\]

and

\[
\frac{d^2}{dt^2} M_X(t) = \frac{d}{dt} \left( \frac{d}{dt} M_X(t) \right) = \left( \frac{\alpha}{\alpha+\beta} + \sum_{k=0}^{\infty} \frac{B(\alpha+k+1, \beta)}{B(\alpha, \beta)} \frac{t^k}{k!} \right) \frac{1}{B(\alpha, \beta)} = \sum_{k=0}^{\infty} \frac{t^k}{k!} \frac{B(\alpha+k+2, \beta)}{B(\alpha, \beta)}
\]

**Approach 1**

one has

\[
E(X) = \frac{d}{dt} M_X(t)|_{t=0} = \frac{\alpha}{\alpha+\beta} + \left( \sum_{k=0}^{\infty} \frac{t^k}{k!} \frac{B(\alpha+k+1, \beta)}{B(\alpha, \beta)} \right)|_{t=0} = \frac{\alpha}{\alpha+\beta}
\]

and

\[
E(X^2) = \frac{d^2}{dt^2} M_X(t)|_{t=0} = \sum_{k=0}^{\infty} \frac{t^k}{k!} \frac{B(\alpha+k+2, \beta)}{B(\alpha, \beta)}|_{t=0} = \frac{B(\alpha+2, \beta)}{B(\alpha, \beta)} + \sum_{k=1}^{\infty} \frac{t^k}{k!} \frac{B(\alpha+k+2, \beta)}{B(\alpha, \beta)}|_{t=0}
\]

\[
= \frac{B(\alpha+2, \beta)}{B(\alpha, \beta)} = \frac{\alpha(\alpha+1)}{(\alpha+\beta)(\alpha+\beta+1)}
\]
Approach 2

It can be seen that

\[ E(X^n) = \int_{-\infty}^{\infty} x^n f(x | \alpha, \beta) \, dx = \int_{0}^{1} x^n \frac{1}{B(\alpha, \beta)} x^{\alpha-1} (1 - x)^{\beta-1} \, dx \]

\[ = \frac{1}{B(\alpha, \beta)} \int_{0}^{1} x^{(\alpha+n)-1} (1 - x)^{\beta-1} \, dx = \frac{B(\alpha + n, \beta)}{B(\alpha, \beta)} \]

\[ = \frac{\Gamma(\alpha + n) \Gamma(\beta)}{\Gamma(\alpha + n + \beta)} \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha) \Gamma(\beta)} = \frac{\Gamma(\alpha + n) \Gamma(\alpha + \beta)}{\Gamma(\alpha + n + \beta)} \]

therefore

\[ E(X) = \frac{\Gamma(\alpha + 1) \Gamma(\alpha + \beta)}{\Gamma(\alpha) \Gamma(\alpha + 1 + \beta)} = \frac{\alpha! (\alpha + \beta - 1)!}{(\alpha - 1)! \Gamma(\alpha + \beta)!} = \frac{\alpha}{\alpha + \beta} \]

and

\[ E(X^2) = \frac{\Gamma(\alpha + 2) \Gamma(\alpha + \beta)}{\Gamma(\alpha) \Gamma(\alpha + 2 + \beta)} = \frac{(\alpha + 1)! (\alpha + \beta - 1)!}{(\alpha - 1)! \Gamma(\alpha + \beta + 1)!} = \frac{\alpha (\alpha + 1)}{(\alpha + \beta) (\alpha + \beta + 1)} \]

thus

\[ \text{Var}(X) = E(X^2) - [E(X)]^2 = \frac{\alpha (\alpha + 1)}{(\alpha + \beta) (\alpha + \beta + 1)} - \left( \frac{\alpha}{\alpha + \beta} \right)^2 \]

\[ = \frac{\alpha}{(\alpha + \beta) (\alpha + \beta + 1)} \left[ \frac{(\alpha + 1)}{(\alpha + \beta + 1)} - \frac{\alpha}{(\alpha + \beta)} \right] \]

\[ = \frac{\alpha}{(\alpha + \beta)} \frac{(\alpha + \beta + 1) (\alpha + \beta)}{(\alpha + 1) (\alpha + \beta + 1)} - \frac{\alpha}{(\alpha + \beta)} \frac{(\alpha + \beta + 1) (\alpha + \beta)}{(\alpha + 1) (\alpha + \beta + 1)} \]

\[ = \frac{\alpha}{(\alpha + \beta) (\alpha + \beta + 1) (\alpha + \beta)} = \frac{\alpha \beta}{(\alpha + \beta + 1) (\alpha + \beta)^2} \]

4.5 Continuous uniform distribution

The PDF and CDF of continuous uniform distribution can be expressed as follows

\[ f(x) = \begin{cases} 
\frac{1}{b-a} & , a \leq x \leq b \\
0 & , \text{otherwise}
\end{cases} \]

\[ F(x) = \begin{cases} 
0 & , x < a \\
\frac{x-a}{b-a} & , a \leq x < b \\
1 & , x \leq b
\end{cases} \]

respectively, the diagram of PDF and CDF of continuous uniform distribution is illustrated in Figure 11.
Figure 11: PDF and CDF of continuous uniform distribution
It has been seen that

\[ M_X(t) = E(e^{tX}) = \int_{-\infty}^{\infty} e^{tx} f(x) \, dx = \int_{a}^{b} e^{tx} \frac{1}{b-a} \, dx = \frac{1}{b-a} \int_{a}^{b} e^{tx} \, dx \]

so

\[ \lim_{t \to 0} \frac{1}{b-a} \left. e^{tx} \right|_{a}^{b} = \frac{1}{t(b-a)} (e^{tb} - e^{ta}) \]

thus

\[ \frac{d}{dt} M_X(t) = \frac{1}{b-a} (\frac{e^{tb} - e^{ta}}{t})' = \frac{1}{b-a} \frac{t(b.e^{tb} - a.e^{ta}) - e^{tb} + e^{ta}}{t^2} \]

and

\[ \frac{d^2}{dt^2} M_X(t) = \left[ \frac{1}{b-a} \frac{t(b.e^{tb} - a.e^{ta}) - e^{tb} + e^{ta}}{t^2} \right]' = \frac{1}{b-a} \left[ \left( \frac{b.e^{tb} - a.e^{ta}}{t} \right)' - \left( \frac{e^{tb} - e^{ta}}{t^2} \right)' \right] \]

\[ = \frac{1}{b-a} \left[ \frac{(b.e^{tb} - a.e^{ta})t - (b.e^{tb} - a.e^{ta})}{t^2} \right] - \frac{1}{b-a} \left[ \frac{(b.e^{tb} - a.e^{ta})t^2 - 2(e^{tb} - e^{ta})}{t^4} \right] \]

\[ = \frac{1}{b-a} \left[ \frac{b^2 e^{tb} - a^2 e^{ta}}{t} - \frac{2(b.e^{tb} - a.e^{ta})}{t^2} + \frac{2(e^{tb} - e^{ta})}{t^3} \right] \]

**Approach 1**

It can be observed that

\[ E(X) = \frac{d}{dt} M_X(t) \bigg|_{t=0} = \frac{1}{b-a} \lim_{t \to 0} \frac{t(b.e^{tb} - a.e^{ta}) - e^{tb} + e^{ta}}{t^2} \]

use the L'Hospital rules we have

\[ \lim_{t \to 0} \frac{t(b.e^{tb} - a.e^{ta}) - e^{tb} + e^{ta}}{t^2} = \lim_{t \to 0} \frac{b.e^{tb} - a.e^{ta} + t.(b^2 e^{tb} - a^2 e^{ta}) - (b.e^{tb} - a.e^{ta})}{2t} \]

\[ = \lim_{t \to 0} \frac{b^2 e^{tb} - a^2 e^{ta}}{2} = \frac{b^2 - a^2}{2} \]

so

\[ E(X) = \frac{1}{b-a} \frac{b^2 - a^2}{2} = \frac{b + a}{2} \]

\[ E(X^2) = \frac{d^2}{dt^2} M_X(t) \bigg|_{t=0} = \lim_{t \to 0} \frac{1}{b-a} \left[ \frac{b^2 e^{tb} - a^2 e^{ta}}{t} - \frac{2(b.e^{tb} - a.e^{ta})}{t^2} + \frac{2(e^{tb} - e^{ta})}{t^3} \right] \]

use L'Hospital rules we have

\[ \lim_{t \to 0} \left[ \frac{b^2 e^{tb} - a^2 e^{ta}}{t} - \frac{2(b.e^{tb} - a.e^{ta})}{t^2} + \frac{2(e^{tb} - e^{ta})}{t^3} \right] \]

\[ = \lim_{t \to 0} \left[ \frac{t^2(b^2 e^{tb} - a^2 e^{ta}) - 2t(b.e^{tb} - a.e^{ta}) + 2(e^{tb} - e^{ta})}{t^3} \right] \]

\[ = \lim_{t \to 0} \left[ \frac{2t(b^2 e^{tb} - a^2 e^{ta}) + t^2(b^3 e^{tb} - a^3 e^{ta})}{3t^2} \right] - \lim_{t \to 0} \left[ \frac{2(b.e^{tb} - a.e^{ta}) + 2t(b^2 e^{tb} - a^2 e^{ta})}{3t^2} \right] + \lim_{t \to 0} \left[ \frac{2(b.e^{tb} - a.e^{ta})}{3t^2} \right] \]

\[ = \lim_{t \to 0} \left[ \frac{b^2 e^{tb} - a^2 e^{ta}}{3t^2} \right] - \lim_{t \to 0} \left[ \frac{b^3 e^{tb} - a^3 e^{ta}}{3} \right] = \frac{b^2 - a^2}{3} - \frac{b^3 - a^3}{3} \]
deduced

\[ E(X^2) = \frac{1}{b-a} \cdot \frac{b^3 - a^3}{3} = \frac{a^2 + ab + b^2}{3} \]

**Approach 2**

we have

\[
E(X) = \int_a^b x.f(x)dx = \int_a^b \frac{1}{b-a} \cdot x^2 \left[ \frac{1}{b-a} \cdot \frac{b^3 - a^3}{3} \right]^{b-a} = \frac{1}{b-a} \cdot \frac{b^2 - a^2}{2} = \frac{a+b}{2}
\]

and

\[
E(X^2) = \int_a^b x^2.f(x)dx = \int_a^b \frac{1}{b-a} \cdot x^3 dx = \frac{1}{b-a} \cdot \frac{x^3}{3} \bigg|_a^b = \frac{b^3 - a^3}{3(b-a)} = \frac{a^2 + ab + b^2}{3}
\]

so

\[
Var(X) = E(X^2) - [E(X)]^2 = \frac{a^2 + ab + b^2}{3} - \left( \frac{a+b}{2} \right)^2 = \frac{4(a^2 + ab + b^2) - 3(a+b)^2}{12} = \frac{4}{12} \left( a^2 - 2ab + b^2 \right) = \frac{(b-a)^2}{12}
\]

**4.6 Exponential distribution**

The PDF and CDF of exponential distribution can be written as follows

\[
f(x) = \begin{cases} 
\lambda e^{-\lambda x}, & x \geq 0 \\
0, & \text{otherwise}
\end{cases}
\]

\[
F(x) = \begin{cases} 
1 - e^{-\lambda x}, & x \geq 0 \\
0, & \text{otherwise}
\end{cases}
\]

respectively, with parameter \( \lambda > 0 \). The diagram of PDF and CDF of exponential distribution is described in Figure 12.
Figure 12: PDF and CDF of exponential distribution
It can be seen that
\[ M_X(t) = E(e^{tx}) = \int_{-\infty}^{\infty} e^{tx} f(x) \, dx = \int_{0}^{\infty} e^{tx} \lambda e^{-\lambda x} \, dx = \lambda \int_{0}^{\infty} e^{(t-\lambda)x} \, dx \]
for \( t < \lambda \)

Therefore
\[
\frac{d}{dt} M_X(t) = \frac{d}{dt} \left( \frac{\lambda}{\lambda - t} \right) = \frac{\lambda}{(\lambda - t)^2} \\
\frac{d^2}{dt^2} M_X(t) = \frac{d}{dt} \left( \frac{\lambda}{(\lambda - t)^2} \right) = \frac{2\lambda}{(\lambda - t)^3}
\]

**Approach 1**

one has
\[
E(X) = \frac{d}{dt} M_X(t)|_{t=0} = \frac{\lambda}{(\lambda - t)^2} \bigg|_{t=0} = \frac{1}{\lambda}
\]
and
\[
E(X^2) = \frac{d^2}{dt^2} M_X(t)|_{t=0} = \frac{2\lambda}{(\lambda - t)^3} \bigg|_{t=0} = \frac{2}{\lambda^2}
\]

**Approach 2**

\[
E(X) = \int_{-\infty}^{\infty} x f(x) \, dx = \int_{0}^{\infty} x \lambda e^{-\lambda x} \, dx = x \lambda \frac{e^{-\lambda x}}{-\lambda} \bigg|_{x=0}^{x=\infty} - \int_{0}^{\infty} \lambda \frac{e^{-\lambda x}}{-\lambda} \, dx = \int_{0}^{\infty} e^{-\lambda x} \, dx = \frac{1}{\lambda}
\]
and
\[
E(X^2) = \int_{-\infty}^{\infty} x^2 f(x) dx = \int_{0}^{\infty} x^2 \lambda e^{-\lambda x} \, dx = x^2 \lambda \frac{e^{-\lambda x}}{-\lambda} \bigg|_{x=0}^{x=\infty} - \int_{0}^{\infty} 2x \lambda \frac{e^{-\lambda x}}{-\lambda} \, dx = \frac{2}{\lambda} \int_{0}^{\infty} x \lambda e^{-\lambda x} \, dx = \frac{2}{\lambda} E(X) = \frac{2}{\lambda^2}
\]
Hence
\[
Var(X) = E(X^2) - [E(X)]^2 = \frac{2}{\lambda^2} - \left( \frac{1}{\lambda} \right)^2 = \frac{1}{\lambda^2}
\]
4.7 Chi-square distribution

The PDF and CDF of Chi-square distribution is given by

\[ f_X(x) = \frac{1}{2^{k/2} \Gamma(k/2)} x^{k/2-1} e^{-x/2}, \quad x \in (0, +\infty) \]

\[ F_X(x) = \frac{1}{\Gamma(k/2)} \gamma \left( \frac{k}{2}, \frac{x}{2} \right) \]

respectively. The diagram of PDF and CDF of Chi-square distribution is provided in Figure 13.
Figure 13: PDF and CDF of chi-square distribution
It has been seen that
\[ M_X(t) = E(e^{tX}) = \int_{-\infty}^{\infty} e^{tx} f_X(x) \, dx = \frac{1}{2^{k/2} \Gamma(k/2)} \int_{0}^{\infty} e^{tx} x^{k/2-1} e^{-x/2} \, dx \]
\[ = \frac{1}{2^{k/2} \Gamma(k/2)} \int_{0}^{\infty} x^{k/2-1} e^{(t-1/2)x} \, dx \]

For the case where \( t < \frac{1}{2} \), let \( u = (1/2 - t)x \) we have
\[ M_X(t) = \frac{1}{2^{k/2} \Gamma(k/2)} \int_{0}^{\infty} x^{k/2-1} e^{(t-1/2)x} \, dx = \frac{1}{2^{k/2} \Gamma(k/2)} \left( \frac{1}{2} - t \right)^{-k/2} \int_{0}^{\infty} u^{k/2-1} e^{-u} \, du \]
\[ = (1 - 2t)^{-k/2} \frac{1}{\Gamma(k/2)} \int_{0}^{\infty} u^{k/2-1} e^{-u} \, du = (1 - 2t)^{-k/2} \]

therefore
\[ \frac{d}{dt} M_X(t) = \frac{d}{dt} ((1 - 2t)^{-k/2}) = k(1 - 2t)^{-k/2-1} \]
and
\[ \frac{d^2}{dt^2} M_X(t) = \frac{d}{dt} (k(1 - 2t)^{-k/2-1}) = k(k + 2)(1 - 2t)^{-k/2-2} \]

**Approach 1**

one has
\[ E(X) = \frac{d}{dt} M_X(t) \big|_{t=0} = (1 - 2t)^{-k/2-1} \big|_{t=0} = k \]

and
\[ E(X^2) = \frac{d^2}{dt^2} M_X(t) \big|_{t=0} = k(k + 2)(1 - 2t)^{-k/2-2} \big|_{t=0} = k(k + 2) \]

**Approach 2**

\[ E(X) = \int_{-\infty}^{\infty} x f_X(x) \, dx = \frac{1}{2^{k/2} \Gamma(k/2)} \int_{0}^{\infty} x^{k/2} e^{-x/2} \, dx \]

Let \( u = x/2 \) then
\[ E(X) = \frac{2}{\Gamma(k/2)} \int_{0}^{\infty} u^{k/2} e^{-u} \, du = \frac{2}{\Gamma(k/2)} \left[ \left. (-u^{k/2} e^{-u}) \right|_{u=0}^{u=\infty} + \frac{k}{2} \int_{0}^{\infty} u^{k/2-1} e^{-u} \, du \right] \]
\[ = \frac{2}{\Gamma(k/2)} \frac{k}{2} \Gamma(k/2) = k \]

and
\[ E(X^2) = \int_{-\infty}^{\infty} x^2 f_X(x) \, dx = \frac{1}{2^{k/2} \Gamma(k/2)} \int_{0}^{\infty} x^{k/2+1} e^{-x/2} \, dx \]
Let \( u = x/2 \) then

\[
E(X^2) = \frac{4}{\Gamma(k/2)} \int_0^\infty u^{k/2+1}e^{-u}du
\]

\[
= \frac{4}{\Gamma(k/2)} \left[ (-u^{k/2+1}e^{-u})\big|_{u=0}^{u=\infty} - \frac{k+2}{2} (u^{k/2}e^{-u})\big|_{u=0}^{u=\infty} + \frac{k(k+2)}{4} \int_0^\infty u^{k/2-1}e^{-u}du \right]
\]

\[
= \frac{4}{\Gamma(k/2)} \cdot \frac{k(k+2)}{4} \Gamma(k/2) = k(k+2)
\]

Hence

\[
Var(X) = E(X^2) - [E(X)]^2 = k(k+2) - k^2 = 2k
\]

### 4.8 Weibull distribution

The PDF and CDF of Weibull distribution can be expressed as follows

\[
f_X(x) = \begin{cases} 
\frac{k}{\lambda} \left(\frac{x}{\lambda}\right)^{k-1}e^{-\left(\frac{x}{\lambda}\right)^k}, & x \geq 0 \\
0, & \text{otherwise}
\end{cases}
\]

\[
F_X(x) = \begin{cases} 
1 - e^{-\left(\frac{x}{\lambda}\right)^k}, & x \geq 0 \\
0, & \text{otherwise}
\end{cases}
\]

respectively. The diagram of PDF and CDF of Weibull distribution is presented in Figure 14.
Figure 14: PDF and CDF of Weibull distribution
It can be observed that

\[ M_X (t) = E (e^{tX}) = \int_{-\infty}^{\infty} e^{tx} f_X (x) \, dx = \int_0^{\infty} e^{tx} \frac{k}{\lambda} x^{k-1} \, e^{-(x/\lambda)^k} \, dx \]

Let \( u = x/\lambda \), for \( \lambda > 0 \)

\[ M_X (t) = \int_0^{\infty} e^{tu} k u^{k-1} e^{-u^k} \, du \]

Let \( x = u^k \), for \( k > 0 \)

\[ M_X (t) = \int_0^{\infty} e^{\lambda x^{1/k}} e^{-x} \, dx = \int_0^{\infty} \sum_{n=0}^{\infty} \frac{(\lambda t)^n}{n!} x^{n/k} e^{-x} \, dx \]

\[ = \sum_{n=0}^{\infty} \frac{(\lambda t)^n}{n!} \int_0^{\infty} x^{n/k} e^{-x} \, dx = \sum_{n=0}^{\infty} \frac{(\lambda t)^n}{n!} \Gamma (n/k + 1) \]

thus

\[ \frac{d}{dt} M_X (t) = \frac{d}{dt} \left( \sum_{n=0}^{\infty} \frac{(\lambda t)^n}{n!} \Gamma (n/k + 1) \right) = \sum_{n=1}^{\infty} \frac{(\lambda t)^{n-1}}{(n-1)!} \Gamma (n/k + 1) \]

and

\[ \frac{d^2}{dt^2} M_X (t) = \frac{d}{dt} \left( \sum_{n=1}^{\infty} \frac{(\lambda t)^{n-1}}{(n-1)!} \Gamma (n/k + 1) \right) = \sum_{n=2}^{\infty} \frac{(\lambda t)^{n-2}}{(n-2)!} \Gamma (n/k + 1) \]

**Approach 1**

one has

\[ E (X) = \frac{d}{dt} M_X (t) \bigg|_{t=0} = \sum_{n=1}^{\infty} \frac{(\lambda t)^{n-1}}{(n-1)!} \Gamma (n/k + 1) \bigg|_{t=0} = \lambda \Gamma (1/k + 1) \]

and

\[ E (X^2) = \frac{d^2}{dt^2} M_X (t) \bigg|_{t=0} = \sum_{n=2}^{\infty} \frac{(\lambda t)^{n-2}}{(n-2)!} \Gamma (n/k + 1) \bigg|_{t=0} = \lambda^2 \Gamma (2/k + 1) \]

**Approach 2**

\[ E (X) = \int_{-\infty}^{\infty} x f_X (x) \, dx = \left( \frac{k}{\lambda} \right) \frac{1}{\lambda^{k-1}} \int_0^{\infty} x^{k} e^{-(x/\lambda)^k} \, dx \]

Let \( t = (x/\lambda)^k \), then we have \( x = \lambda t^{1/k} \) and \( dx = \frac{\lambda}{k} t^{1/k-1} \, dt \).

\[ E (X) = \left( \frac{k}{\lambda} \right) \frac{1}{\lambda^{k-1}} \int_0^{\infty} \lambda^k t e^{-t} \frac{\lambda}{k} t^{1/k-1} \, dt = \lambda \int_0^{\infty} e^{-t} t^{1/k-1} \, dt = \lambda \Gamma \left( 1 + \frac{1}{k} \right) \]

and

\[ E (X^2) = \int_0^{\infty} x^2 f_X (x) \, dx = \left( \frac{k}{\lambda} \right) \frac{1}{\lambda^{k-1}} \int_0^{\infty} x^{k+1} e^{-(x/\lambda)^k} \, dx \]
Let \( t = (x/\lambda)^k \), then we have \( x = \lambda t^{1/k} \) and \( dx = \frac{\lambda t^{1/k - 1}}{k} dt \).

\[
E(X^2) = \left( \frac{k}{\lambda} \right) \frac{1}{\lambda^{k-1}} \int_{0}^{\infty} \lambda^{k+1} t^{k+1} e^{-t^2} \frac{1}{k} t^{1/k-1} dt = \lambda^2 \int_{0}^{\infty} e^{-t^2} \frac{1}{k} t^{1/k-1} dt
\]

Hence

\[
Var(X) = E(X^2) - [E(X)]^2 = \lambda^2 \Gamma \left( 1 + \frac{2}{k} \right) - \left[ \lambda \Gamma \left( 1 + \frac{1}{k} \right) \right]^2
\]

### 4.9 Laplace distribution

The PDF and CDF of Laplace distribution is given by

\[
f(x|\mu, \sigma) = \frac{1}{2\sigma} \exp \left( -\frac{|x - \mu|}{\sigma} \right)
\]

\[
F(x|\mu, \sigma) = \begin{cases} 
\frac{1}{2} \exp \left( \frac{x - \mu}{\sigma} \right) & \text{if } x \leq \mu \\
1 - \frac{1}{2} \exp \left( -\frac{x - \mu}{\sigma} \right) & \text{if } x \geq \mu
\end{cases}
\]

respectively. The diagram of PDF and CDF of Laplace distribution is described in Figure 15.
Figure 15: PDF and CDF of Laplace distribution
It can be seen that

\[ M_X(t) = E(e^{tX}) = \int_{-\infty}^{\infty} e^{tx} f_X(x) \, dx = \frac{1}{2\sigma} \int_{-\infty}^{\infty} e^{tx} e^{-\frac{|x-\mu|}{\sigma}} \, dx \]

Let \( y = \frac{x-\mu}{\sigma} \), then \( x = y\sigma + \mu \)

\[ M_X(t) = \frac{1}{2\sigma} \int_{-\infty}^{\infty} e^{t(y\sigma+\mu)} e^{-|y|\sigma} \, dy = \frac{1}{2} e^{\mu t} \int_{-\infty}^{\infty} e^{ty\sigma} e^{-|y|\sigma} \, dy \]

\[ = \frac{1}{2} e^{\mu t} \left( \int_{-\infty}^{t\sigma} e^{ty\sigma} \, dy + \int_{t\sigma}^{\infty} e^{ty\sigma} \, dy \right) = \frac{1}{2} e^{\mu t} \left( \int_{-\infty}^{0} e^{y(t+\sigma)} \, dy + \int_{0}^{\infty} e^{-y(-t+\sigma)+1} \, dy \right) \]

\[ = \frac{1}{2} e^{\mu t} \left( \frac{1}{t\sigma+1} e^{y(t+\sigma+1)} \right)_{y=0}^{y=-\infty} + \frac{1}{t\sigma-1} e^{-y(-t+\sigma)} \right)_{y=0}^{y=\infty} \]

\[ = \frac{1}{2} e^{\mu t} \left( \frac{1}{t\sigma+1} e^{y(t+\sigma+1)} \right)_{y=0}^{y=-\infty} + \frac{1}{t\sigma-1} e^{-y(-t+\sigma)} \right)_{y=0}^{y=\infty} \]

thus

\[ \frac{d}{dt} M_X(t) = \frac{d}{dt} \left( \frac{e^{\mu t}}{1-t^2\sigma^2} \right) = \frac{e^{\mu t}(\mu - t^2\mu\sigma^2 + 2t\sigma^2)}{(1-t^2\sigma^2)^2} \]

and

\[ \frac{d^2}{dt^2} M_X(t) = \frac{d}{dt} \left( \frac{e^{\mu t}(\mu - t^2\mu\sigma^2 + 2t\sigma^2)}{(1-t^2\sigma^2)^2} \right) = \frac{e^{\mu t}(1-t^2\sigma^2)^3}{(1-t^2\sigma^2)^4} \left[ \mu^2 + 2\sigma^2 + 6t^2\sigma^2 + 2t\mu\sigma^2 (2 - t\mu - 2t^2\sigma^2) \right] \]

**Approach 1**

therefore

\[ E(X) = \left. \frac{d}{dt} M_X(t) \right|_{t=0} = \left. \frac{e^{\mu t}(\mu - t^2\mu\sigma^2 + 2t\sigma^2)}{(1-t^2\sigma^2)^2} \right|_{t=0} = \mu \]

and

\[ E(X^2) = \left. \frac{d^2}{dt^2} M_X(t) \right|_{t=0} = \left. \frac{e^{\mu t}(1-t^2\sigma^2)^3}{(1-t^2\sigma^2)^4} \left[ \mu^2 + 2\sigma^2 + 6t^2\sigma^2 + 2t\mu\sigma^2 (2 - t\mu - 2t^2\sigma^2) \right] \right|_{t=0} = \mu^2 + 2\sigma^2 \]

**Approach 2**

\[ E(X) = \int_{-\infty}^{\infty} x f_X(x) \, dx = \frac{1}{2\sigma} \int_{-\infty}^{\infty} x e^{-\frac{|x-\mu|}{\sigma}} \, dx \]

Let \( u = x - \mu \), then \( x = u + \mu \) and \( dx = du \).

By linearity one has \( E(X) = E(U + \mu) = E(U) + \mu \), hence

\[ E(X) = \frac{1}{2\sigma} \int_{-\infty}^{\infty} u e^{-\frac{|u|}{\sigma}} \, du + \mu = \frac{1}{2\sigma} \int_{\infty}^{0} u e^{\frac{u}{\sigma}} \, du + \frac{1}{2\sigma} \int_{0}^{\infty} u e^{-\frac{u}{\sigma}} \, du + \mu \]

\[ = -\frac{1}{2\sigma} \int_{0}^{\infty} u e^{-\frac{u}{\sigma}} \, du + \frac{1}{2\sigma} \int_{0}^{\infty} u e^{-\frac{u}{\sigma}} \, du + \mu = \mu \]
and

\[ E(X^2) = E[(U + \mu)^2] = E(U^2) + 2\mu E(U) + \mu^2 = E(U^2) + 2\mu E(X - \mu) + \mu^2 \]

\[ = \frac{1}{2\sigma} \int_{-\infty}^{\infty} u^2 e^{-\frac{|u|}{\sigma}} du + \mu^2 = \frac{1}{2\sigma} \int_{0}^{\infty} u^2 e^{\frac{u}{\sigma}} du + \frac{1}{2\sigma} \int_{0}^{\infty} u^2 e^{-\frac{u}{\sigma}} du + \mu^2 \]

\[ = \frac{1}{2\sigma} \int_{0}^{\infty} u^2 e^{-\frac{u}{\sigma}} du + \frac{1}{2\sigma} \int_{0}^{\infty} u^2 e^{-\frac{u}{\sigma}} du + \mu^2 = \frac{1}{\sigma} \int_{0}^{\infty} u^2 e^{-\frac{u}{\sigma}} du + \mu^2 \]

Let \( v = u/\sigma \), then

\[ E(X^2) = \sigma^2 \int_{0}^{\infty} v^2 e^{-v} dv + \mu^2 = 2\sigma^2 + \mu^2 \]

Hence

\[ Var(X) = E(X^2) - [E(X)]^2 = 2\sigma^2 + \mu^2 - \mu^2 = 2\sigma^2 \]

### 4.10 Gumbel distribution

The PDF and CDF of Gumbel distribution can be written as follows

\[ f(x) = \frac{1}{\beta} e^{-(x + e^{-\frac{x}{\beta}})} \]

\[ F(x) = e^{-(x - \mu)/\beta} \]

respectively, where \( z = \frac{x - \mu}{\beta} \). The diagram of PDF and CDF of Gumbel distribution is illustrated in Figure 16.
Figure 16: PDF and CDF of Gumbel distribution
It can be observed that
\[ M_X(t) = E(e^{tX}) = \int_{-\infty}^{\infty} e^{tx} f(x) \, dx = \int_{-\infty}^{\infty} e^{tx} \frac{1}{\beta} e^{-(z+e^{-z})} \, dx = \int_{-\infty}^{\infty} e^{(\beta z + \mu)} e^{-z} e^{-e^{-z}} \, dz \]

Let \( y = e^{-z} \), then \( z = -\ln y \) and \( dy = -e^{-z} \, dz \)
\[ M_X(t) = \int_{0}^{\infty} e^{t\mu} e^{-t\ln y} e^{-y} \, dy = \int_{0}^{\infty} e^{t\mu} y^{-t\beta} e^{-y} \, dy = e^{t\mu} \Gamma(1 - t\beta) \]

hence
\[ \frac{d}{dt} M_X(t) = \frac{d}{dt}(e^{t\mu} \Gamma(1 - t\beta)) = \mu e^{t\mu} \Gamma(1 - t\beta) - \beta e^{t\mu} \Gamma'(1 - t\beta) \]

and
\[ \frac{d^2}{dt^2} M_X(t) = \frac{d}{dt}(\mu e^{t\mu} \Gamma(1 - t\beta) - \beta e^{t\mu} \Gamma'(1 - t\beta)) = \mu^2 e^{t\mu} \Gamma(1 - t\beta) - 2\beta \mu e^{t\mu} \Gamma'(1 - t\beta) + \beta^2 e^{t\mu} \Gamma''(1 - t\beta) \]

**Approach 1**

One has
\[ E(X) = \frac{d}{dt} M_X(t)|_{t=0} = [\mu e^{t\mu} \Gamma(1 - t\beta) - \beta e^{t\mu} \Gamma'(1 - t\beta)]|_{t=0} = \mu \Gamma(1) - \beta \Gamma'(1) = \mu + \beta \gamma \]
with \( \gamma \) is the Euler-Mascheroni constant.

And
\[ E(X^2) = \frac{d^2}{dt^2} M_X(t)|_{t=0} = \left[ \mu^2 e^{t\mu} \Gamma(1 - t\beta) - 2\beta \mu e^{t\mu} \Gamma'(1 - t\beta) + \beta^2 e^{t\mu} \Gamma''(1 - t\beta) \right]|_{t=0} \]
\[ = \mu^2 + 2\beta \mu \gamma + \beta^2 \left( \gamma^2 + \frac{\pi^2}{6} \right) \]

because
\[ \Gamma''(1 - t\beta) = \int_{0}^{\infty} y^{-t\beta} e^{-y \ln^2 y} \, dy \]
\[ \Gamma''(1) = \int_{0}^{\infty} e^{-y \ln^2 y} \, dy = \gamma^2 + \frac{\pi^2}{6} \]

**Approach 2**

\[ E(X) = \int_{-\infty}^{\infty} x f(x) \, dx = \int_{-\infty}^{\infty} x \frac{1}{\beta} e^{-(z+e^{-z})} \, dx = \int_{-\infty}^{\infty} (\beta z + \mu) e^{-z} e^{-e^{-z}} \, dz \]
Let \( y = e^{-z} \), then \( z = -\ln y \) and \( dy = -e^{-z}dz \)

\[
E(X) = \int_{0}^{\infty} -\beta \ln y + \mu e^{-y}dy = \int_{-\infty}^{\infty} e^{y\beta}e^{-y}dy = -\mu e^{-y}\big|_{y=0}^{y=\infty} - \beta \int_{0}^{\infty} e^{-y} \ln ydy = \mu + \beta \gamma
\]

and

\[
E(X^2) = \int_{\infty}^{\infty} x^2 f(x)dx = \int_{-\infty}^{\infty} x^2 \frac{1}{\beta} e^{-(z+\epsilon^+)} dx = \int_{-\infty}^{\infty} (\beta z + \mu)^2 e^{-z}e^{-\epsilon^+} dz
\]

Let \( y = e^{-z} \), then \( z = -\ln y \) and \( dy = -e^{-z}dz \)

\[
E(X^2) = \int_{0}^{\infty} (-\beta \ln y + \mu)^2 e^{-y}dy = -\mu^2 e^{-y}\big|_{y=0}^{y=\infty} - 2\mu \beta \int_{0}^{\infty} e^{-y} \ln ydy + \beta^2 \int_{0}^{\infty} e^{-y}\ln^2 ydy
\]

\[
= \mu^2 + 2\mu \beta \gamma + \beta^2 \left(\gamma^2 + \frac{\pi^2}{6}\right)
\]

Hence

\[
Var(X) = E(X^2) - [E(X)]^2 = \mu^2 + 2\mu \beta \gamma + \beta^2 \left(\gamma^2 + \frac{\pi^2}{6}\right) - (\mu + \beta \gamma)^2 = \beta^2 \frac{\pi^2}{6}
\]

### 5 Discussion on the distribution functions

It can be seen that, to calculate the expectation and variance of distribution functions, one can utilize 2 different approaches, including based on the first and second derivatives of the moment generating function (Approach 1) or direct calculation (Approach 2). It can be observed that, if variable \( X \) is discrete, it is difficult to obtain the results of the expectation and variance of distribution functions by Approach 2 in some cases, for example, \( X \) follows the binomial, negative binomial, and geometric distributions because the distribution functions contain the expression of combination. Thus, to avoid the complication in calculations, one should utilize the first approach that based on the first and second derivatives of moment generating function.

On the other hand, if variable \( X \) is continuous, it is arduous to obtain the results of the expectation and variance of the distribution functions by using Approach 2 in some cases, for example, when \( X \) follows normal, log-normal, Chi-square, and Gumble distributions. Thus, Approach 1 should be used in this situation. In general, it can be seen that, Approach 1 that is based on the first and second derivatives of the moment generating function will be easier to implement in the calculation of the expectation and variance for the distribution functions. Nevertheless, some distribution functions have the simple formulas such as Bernoulli, Poisson, and continuous uniform distribution, one should utilize Approach 2, regardless whether the variable is discrete or continue.
Furthermore, it has been seen that, distribution functions play a very crucial role in the literature by using some specific distribution functions, including Bernoulli, binomial, negative binomial, and Poisson distributions. For example, one can develop models like logistic model by using Bernoulli, binomial, negative binomial, and Poisson distributions. Readers can refer in Cameron (1990), Chin et al. (2003), Hosmer et al. (2013), King et al. (2001), and Kupper et al. (1978) for more information.

If the outcome of the variables have much more zeros than expected, then it is not easy to handle the models. Some of models are proposed to improve this issue such as zero-inflated Bernoulli (ZIBer) model, zero-inflated binomial (ZIB) model, zero-inflated negative binomial (ZINB) model, and zero-inflated Poisson (ZIP) model. The study and discuss about zero-inflated models are diverse and abundant.

For instance, Diop et al. (2011) introduce the maximum likelihood estimation to estimate parameters in the zero-inflated Bernoulli (ZIBer) model. Diallo et al. (2017) develop some asymptotic properties of the maximum-likelihood estimator in zero-inflated binomial (ZIB) regression. Pho and Nguyen (2018) utilize the Newton-Raphson method and maxLik function in the statistical software R to compare the results of estimation parameter for the zero-inflated binomial (ZIB) regression model. Pho et al. (2019) mention to some of zero-inflated regression models. Lambert (1992) introduces the zero-inflated Poisson (ZIP) regression, with an application to measure defects in manufacturing, etc.

In addition, readers may refer in Hall (2000), Pho et al. (2019), Ridout et al. (2001), etc to see the applications of the other zero-inflated models. In addition, Bian, et al. (2011) develop a trinomial test for paired data when there are many ties. Matsumura, et al. (1990) develop an extended Multinomial-Dirichlet model for error bounds for dollar-unit sampling in which there are many zeros.

6 Applications in Decision Sciences

In this section, we review the applications of the theory discussed and developed in this paper to decision sciences. There are many applications of the theory discussed and developed in this paper to decision sciences. In this paper, we will mainly discuss the applications related to our work. The obvious application is estimation and testing, especially parametric estimation and testing because all parametric estimation and testing involve distribution. We first discuss to robust estimation.
6.1 Robust Estimation

Tiku and Wong (1998) develop a unit root test to take care of data follow an AR(1) model. Tiku, Wong and Bian, (1999) derive the MML (modified maximum likelihood) estimators of the parameters for AR(q) models with asymmetric innovations represented by gamma and generalized logistic distributions while Tiku, Wong, Vaughan, and Bian (2000) derive the MML estimators of the parameters for AR(q) models with non-normal innovations represented by Student’s t distribution. They show that the estimators are remarkably efficient and easy to compute.

On the other hand, Tiku, Wong and Bian, (1999a) derive the estimator for coefficients in a simple regression model with autocorrelated errors in which the underlying distribution is assumed to be symmetric, one of Student’s t family for illustration. Wong and Bian (2005) extend the theory by considering the underlying distribution is a generalized logistic distribution. They develop the MML estimators since the ML (maximum likelihood) estimators are intractable for the generalized logistic data. They then study the asymptotic properties of the proposed estimators and conduct simulation to the study.

6.2 Bayesian estimation

Bian and Wong (1997) develop the normal g-prior Bayesian estimator for regression coefficients using independent Cauchy and inverted gamma prior distributions. Their proposed estimator has a simple mathematical expression and it is an adaptive weighted average of the least square estimator (LSE) and the prior location with weights depending on residuals. There are many applications of the models to decision science.

For example, Wong and Bian (2000) introduce the robust Bayesian estimator to the estimation of the Capital Asset Pricing Model (CAPM) in which the distribution of the error component is well-known to be flat-tailed. Their simulation shows that the Bayesian estimator is robust and superior to the least squares estimator when the CAPM is contaminated by large normal and/or non-normal disturbances, especially by Cauchy disturbances. In their empirical study, we find that the robust Bayesian estimate is uniformly more efficient than the least squares estimate in terms of the relative efficiency of one-step ahead forecast mean square error, especially for small samples.

In addition, readers may refer in Matsumura, Tsui and Wong (1990) use a multinomial
distribution model within the dollar-unit sampling framework, with a Dirichlet prior distribution to develop the extended model and a different Dirichlet prior to generate upper and lower bounds and two-sided confidence intervals for situations in which both understatement and overstatement errors are possible.

6.3 Portfolio estimation

Another area in decision sciences that the approaches discussed in our paper can be used is to estimate portfolio return that Markowitz (1952) introduces the theory in which investors select portfolios to maximize profit subject to achieving a specified level of calculated risk or, equivalently, minimize variance subject to obtaining a predetermined level of expected gain.

Bai, Liu, and Wong (2009a) prove that the estimates proposed by Markowitz (1952) is seriously depart from its theoretic optimal return and they call this phenomenon ”over-prediction.” To circumvent this over-prediction problem, they introduce the bootstrap-corrected estimates for the optimal return and its asset allocation, and prove that the estimates can correct the over-prediction and reduce the error drastically. They also prove that the estimates are proportionally consistent with their counterpart parameters. Leung, Ng, and Wong (2012) extend the theory by developing a new estimator for the optimal portfolio return based on an unbiased estimator of the inverse of the covariance matrix and its related terms, and derive explicit formulae for the estimator of the optimal portfolio return. Li, Bai, McAleer, and Wong (2016) further improve the estimation by using the spectral distribution of the sample covariance.


6.4 Stochastic Dominance estimation

Another important area in decision sciences that the theory discussed in our paper can be used is to get the estimation of stochastic dominance (SD) for different types of investors. Readers may refer to Wong and Li (1999), Li and Wong (1999), Wong (2007), Sriboonchitta, Wong, Dhompongs, and Nguyen (2009), Levy (2015), Chan, Clark, and Wong (2016), Guo
and Wong (2016) for the SD theory for risk averters and risk seekers; refer to Levy and Levy (2002, 2004) and Wong and Chan (2008) for the prospect SD (PSD) and Markowitz SD (MSD) to link to investors with the corresponding S-shaped and reverse S-shaped utility functions; and refer to Leshno and Levy (2002), Guo, Zhu, Wong, and Zhu (2013), Guo, Post, Wong, and Zhu (2014), and Guo, Wong, Zhu (2016) for the theory of almost SD. For example, Bai, Li, McAleer, and Wong (2015) extend the SD test statistics developed by Davidson and Duclos (2000) to get SD tests for risk averters and risk seekers, Bai, Li, Liu, and Wong (2011) develop the SD test statistics MSD and PSD, and Ng, Wong, and Xiao (2017) develop the SD test by using quantile regressions. In addition, Lean, Wong, Zhang (2008) have conducted simulation and show that SD tests introduced by Davidson and Duclos (2000) has better size and power performances than two alternative tests. The approaches discussed in our paper is useful to their SD test statistics.

The approaches discussed in our paper is useful to the SD theory because there are several SD tests that can be used the approaches discussed in our paper to use moment generating function, expectation and variance of different distributions. What’s more, SD itself is to compare the distributions of different aspect. Thus, the approaches discussed in our paper can directly use to the SD theory.

The SD theory can be used in many areas, including indifference curves (Wong, 2006, 2007; Ma and Wong, 2010; Broll, Egozcue, Wong, and Zitikis, 2010), two-moment decision model (Broll, Guo, Welzel, and Wong, 2015; Guo, Wagener, and Wong, 2018), moment rule (Chan, Chow, Guo, and Wong, 2018), economic growth (Chow, Vieito, and Wong, 2018), diversification (Egozcue and Wong, 2010; Egozcue, Fuentes García, Wong, and Zitikis, 2011; Lozza, Wong, Fabozzi, and Egozcue, 2018). It can also be applied to many different assets, including stock (Fong, Lean, and Wong, 2008), fund (Gasbarro, Wong, and Zumwalt, 2007, 2012; Wong, Phoon, Lean, 2008), futures (Lean, McAleer, Wong, 2010; Lean, Phoon, Wong, 2012; Qiao, Clark, Wong, 2012; Qiao, Wong, Fung, 2013; Lean, McAleer, Wong, 2015; Clark, Qiao, Wong, 2016), Warrant (Chan, de Peretti, Qiao, Wong, 2012; Wong, Lean, McAleer, Tsai, 2018), Option (Abid, Mroua, and 2009), wine (Bouri, Gupta, Wong, and Zhu, 2018), warrants (Chan, de Peretti, Qiao, and Wong, 2012), gold (Hoang, Wong, and Zhu, 2015, 2018; Hoang, Zhu, El Khamlichi, and Wong, 2019), property market (Qiao, Wong, 2015; Tsang, Wong, Horowitz, 2016).

In addition, it can also be used to test for anomaly and market efficient (Lean, Smyth, Wong, 2007; Qiao, Qiao, Wong, 2010), examine different trading strategies (Fong, Wong,
and Lean, 2005; Wong, Thompson, Wei, Chow, 2006), banking performance (Broll, Wong, and Wu, 2011), study the effects of financial crisis (Vieito, Wong, Zhu, 2015; Zhu, Bai, Vieito, Wong, 2019), and international trade (Broll, Wahl, and Wong, 2006). In addition, it can also be used to measure income inequality (Valenzuela, Wong, and Zhu, 2019). All of these applications are related to decision science.

6.5 Risk Measure Estimation

Risk measure estimation is another important area that the approaches discussed in our paper can be used. We include mean-variance rule as one of the risk measures, especially because the approaches discussed in our paper include estimating mean and variance. Readers may refer to Markowitz (1952) and Wong (2007) for the MV rule for risk averters and risk seekers, respectively, refer to Leung and Wong (2008), Wong, Wright, Yam, and Yung (2012), and the references therein for the Sharpe ratio, refer to Ma and Wong (2010) and the references therein for VaR and conditional-VaR (CVaR), refer to Guo, Jiang, and Wong (2017), Guo, Chan, Wong, and Zhu (2018), and the references therein for the Omega ratio, refer to Niu, Wong, and Xu (2017) and the references therein for the n-order Kappa ratio, refer to Guo, Niu, and Wong (2019) and the references therein for the Farinelli and Tibiletti ratio, and refer to Niu, Guo, McAleer, and Wong (2018), Lu, Yang, Wong (2018), Lu, Hoang, and Wong (2019).

Furthermore, the economic performance measure of risk and the economic index of riskiness, refer to Bai, Wang, Wong (2011), Bai, Hui, Wong, Zitikis (2012) for the mean-variance ratio test, refer to Tang, Sriboonchitta, Ramos, Wong (2014), Ly, Pho, Ly, Wong (2019a,b) for Copulas. The approaches discussed in our paper is useful to the theory of risk measure estimation because most, if not all, of the risk measure estimation will use distribution function, moment generating function, mean, and variance. There are other risk measures, for example, Guo, Li, McAleer, Wong, (2018), etc. In addition, there are many applications for the risk measures in decision sciences, see, for example, our discussion in Sections 6.3 and 6.4 for the applications.

6.6 Behavioral Models

The approaches discussed in our paper can be used in many behavioral models. We first review the utility functions that are the basics of the behavioral models. Utility starts


The approaches discussed in our paper can be used in many behavioral models because after one develops any behavioral model, one may then develop the corresponding econometric models so that the behavioral models can be estimated. For example, Fabozzi, Fung, Lam, and Wong (2013) extend the models developed by Lam, Liu, and Wong (2010, 2012), Guo, McAlleer, Wong, and Zhu (2017) and others by developing 3 tests to test for the magnitude effect of short-term underreaction and long-term overreaction that can use the approaches discussed in our paper to get optimization solutions. On the other hand, Wong, Chow, Hon, and Woo (2018) conduct a questionnaire survey to examine whether the theory developed by Lam, Liu, and Wong (2008, 2010), and Guo, McAlleer, Wong, and Zhu (2017) and others that can use the approaches discussed in our paper to get the moment generating function, expectation and variance of different distributions for the behavior models estimators.

There are many other behavior models also. For example, Egozcue and Wong (2010a) and Egozcue, Fuentes García, Wong, and Zitikis (2012a) develop an analytical theory to explain the behavior of investors with extended value functions in segregating or integrating multiple outcomes when evaluating mental accounting. Guo, Wong, Xu, and Zhu (2015), Egozcue, Guo, and Wong (2015), and Guo, and Wong (2019) develop models to investigate
regret-averse firms’ production and hedging behaviors while Guo, Egozcue, and Wong (2019) develop several properties of using disappointment aversion to model production decision.

6.7 Economic and Financial Indicators

Most of economic and financial indicators could be related to decision sciences and can use the approaches discussed in our paper to get the moment generating function, expectation and variance of different distributions for the economic and financial indicators. There are many economic and financial indicators that can use the approaches discussed in our paper to get the moment generating function, expectation and variance of different distributions for the economic and financial indicators. We only discuss those related to our work.

We have developed some financial indicators and have applied some economic indicators to study some important economic issues that could be related to decision sciences and can use the approaches discussed in our paper to get the moment generating function, expectation and variance of different distributions for the economic and financial indicators. For example, Wong, Chew, and Sikorski (2001) develop a new financial indicator to test the performance of stock market forecasts by using the E/P ratios and bond yields. They also develop two test statistics to utilize the indicator and illustrate the tests in several stock markets. Exploring the characteristics associated with the formation of bubbles that occurred in the Hong Kong stock market in 1997 and 2007 and the 2000 dot-com bubble of Nasdaq, McAleer, Suen, and Wong (2016) establish trading rules that not only produce returns significantly greater than buy-and-hold strategies, but also produce greater wealth compared with TA strategies without trading rules.

In addition, Chong, Cao, and Wong (2017) develop a new market sentiment index for the Hong Kong stock market by using the turnover ratio, short-selling volume, money flow, HIBOR, and returns of the U.S. and Japanese markets, the Shanghai and Shenzhen Composite indices. Thereafter, they incorporate the threshold regression model with the sentiment index as a threshold variable to capture the state of the Hong Kong stock market. Sethi, Wong, and Acharya (2018) examine the sectoral impact of disinflationary monetary policy by calculating the sacrifice ratios for several OECD and non-OECD countries. Sacrifice ratios calculated through the episode method reveal that disinflationary monetary policy has a differential impact across three sectors in both OECD and non-OECD countries.
6.8 Cointegration and Causality

Most of the cointegration and causality estimation and testing statistics could be related to decision sciences and can use the approaches discussed in our paper to get the moment generating function, expectation and variance of different distributions. There are many cointegration and causality estimation and testing statistics that can use the approaches discussed in our paper to get the moment generating function, expectation and variance of different distributions. We only discuss those related to our work.

Tiku and Wong (1998) develop a unit root test to take care of data follow an AR(1) model. Penm, Terrell, Wong (2003) present simulations and an application that demonstrates the usefulness of the zero-non-zero patterned vector error-correction models (VECMs). Lam, Wong, and Wong (2006) develop some properties on the autocorrelation of the k-period returns for the general mean reversion (GMR) process in which the stationary component is not restricted to the AR(1) process but takes the form of a general ARMA process. Bai, Wong, and Zhang (2010) develop a nonlinear causality test in multivariate settings. Bai, Li, Wong, and Zhang (2011) first discuss linear causality tests in multivariate settings and thereafter develop a nonlinear causality test in multivariate settings.

Bai, Hui, Jiang, Lv, Wong, Zheng (2018) revisit the issue by estimating the probabilities and reestablish the CLT of the new test statistic. Hui, Wong, Bai, and Zhu (2017) propose a quick and efficient method to examine whether a time series possesses any nonlinear feature by testing a kind of dependence remained in the residuals after fitting the dependent variable with a linear model. All the above models are be related to decision sciences and can use the approaches discussed in our paper to get the moment generating function, expectation and variance of different distributions.

Most of statistical and econometric models could be related to decision sciences and can use the approaches discussed in our paper to get the moment generating function, expectation and variance of different distributions. There are many statistical and econometric models that can use the approaches discussed in our paper to get the moment generating function, expectation and variance of different distributions. We only discuss those related to our work.

We have been developing or applying some other statistical and econometric models that can use the approaches discussed in our paper to get the moment generating function, expectation and variance of different distributions. We state a few here. First, Wong and Miller (1990) develop a theory and methodology for repeated time series (RTS) measurements on autoregressive integrated moving average-noise (ARIMAN) process. Second, Bian, McAleer, and Wong (2011) develop a new test, the trinomial test, for pairwise ordinal data samples to improve the power of the sign test by modifying its treatment of zero differences between observations, thereby increasing the use of sample information. The models in the above papers can use the approaches discussed in our paper to get the moment generating function, expectation and variance of different distributions.

Raza, Sharif, Wong, and Karim (2016) have used maximal overlap discrete wavelet transform (MODWT), wavelet covariance, wavelet correlation, continuous wavelet power spectrum, wavelet coherence spectrum and wavelet-based Granger causality analysis to investigate the empirical influence of tourism development (TD) on environmental degradation in a high-tourist-arrival economy (i.e. United States), using the wavelet transform framework. Xu, Wong, Chen, and Huang (2017) analyze the relationship among stock networks by focusing on the statistically reliable connectivity between financial time series, which accurately reflects the underlying pure stock structure.

Furthermore, readers may refer in Tsendsuren, Li, Peng, and Wong (2018) examine the relationships among three health status indicators (self-perceived health status, objective health status, and future health risk) and life insurance holdings in 16 European countries. Mou, Wong, and McAleer (2018) analyze core enterprise credit risks in supply chain finance by means of a 'fuzzy analytical hierarchy process' to construct a supply chain financial credit risk evaluation system, making quantitative measurements and evaluation of core enterprise credit risk.
In addition, Pham, Wong, Moslehpour, and Musyoki (2018) suggest an outsourcing hierarchy model based on the concept of the analytic hierarchy process with four levels of the most concerned attributes: competitiveness, human resources, business environment, and government policies and compare between the analytic hierarchy process (AHP) and Fuzzy AHP show some significant differences but lead to similar conclusions. They provide decision makers an outsourcing hierarchy model based on the AHP and Fuzzy AHP approach with the most concerned factors.

We note that it is not only statistical and econometric models related to decision sciences that can use the approaches discussed in our paper to get the moment generating function, expectation and variance of different distributions. There are many other models, for example, probability and mathematical models that can use the approaches discussed in our paper to get the moment generating function, expectation and variance of different distributions.

Over here, we give a few examples. Egozcue, Fuentes García, and Wong (2009) derive some covariance inequalities for monotonic and non-monotonic functions. Egozcue, Fuentes García, Wong, and Zitikis (2010) sharpen the upper bound of a Grüss-type covariance inequality by incorporating a notion of quadrant dependence between random variables and also utilizing the idea of constraining the means of the random variables. Egozcue, Fuentes García, Wong, and Zitikis (2011a) show that Grüss-type probabilistic inequalities for covariances can be considerably sharpened when the underlying random variables are quadrant dependent in expectation (QDE).

Moreover, Egozcue and Wong (2010a) extend prospect theory, mental accounting, and the hedonic editing model by developing an analytical theory to explain the behavior of investors with extended value functions in segregating or integrating multiple outcomes when evaluating mental accounting. Egozcue, Fuentes García, Wong, and Zitikis (2012a) develop decision rules for multiple products, which generally call ‘exposure units’ to naturally cover manifold scenarios spanning well beyond ‘products’. All the above models could use the approaches discussed in our paper to get the moment generating function, expectation and variance of different distributions.

Last, we note that there are many other areas in decision sciences that can use the approaches discussed in our paper to get the moment generating function, expectation and variance of different distributions, in this paper we also review some as discussed in the above. For more applications in decision sciences that can use the approaches discussed
in our paper to get the moment generating function, expectation and variance of different distributions, readers may refer to Chang, McAleer, and Wong (2015, 2016, 2018, 2018a, 2018b, 2018c) and Pho, Tran, Ho, and Wong (2019) for more information.

7 Conclusion

In this paper, we present the theory of some important distribution functions and their moment generating functions. We also introduce two approaches to derive the expectations and variances for all the distribution functions being studied in our paper. The first approach is to use the first and second derivatives of the moment generating function to calculate the expectation and variance of the corresponding distribution while the second approach is to use direct calculation. We discuss the advantages and disadvantages of each approach in our paper.

In addition, we display the diagrams of the probability mass function, probability density function, and cumulative distribution function for each distribution function being investigated in this paper. For each distribution, we show how to construct the corresponding regression models. We also discuss the difficulty when the outcome of the variables have much more zeros than expected and how to overcome the difficulty. In addition, we review the applications of the theory discussed and developed in this paper to decision sciences.

Moreover, we have checked many books and papers. So far, we cannot find any book or paper present the detail of the theory discussed in our paper. Thus, we strongly believe though some or even all theories developed in our paper are well-known, our paper is the first paper discussing the details of the theory for some important distribution functions with applications, and thus, our paper could still have important some contributions to the literature.
References


Cain, M.: The moment-generating function of the minimum of bivariate normal random


Chan, R.H., Chow, S.C., Guo, X., Wong, W.K.: Central Moments, Stochastic Dominance, Moment Rule, and Diversification with Application, the International Conference on Scientific Computing, in honor of Professor Raymond Chan’s 60th birthday, to be held on December 5-8, 2018 in the Chinese University of Hong Kong (2018).


Chin, H. C., Quddus, M. A.: Applying the random effect negative binomial model to examine traffic accident occurrence at signalized intersections. Accident Analysis & Prevention, 35(2), 253-259 (2003).


Egozcue M., Wong W.K.: Gains from Diversification on Convex Combinations: A Majoriza-


Wong, W.K., Thompson H.E., Wei S., Chow Y.F.: Do Winners perform better than Losers? A Stochastic Dominance Approach, Advances in Quantitative *Analysis of Finance and


