SECTION 5 - EXPECTATION AND OTHER DISTRIBUTION PARAMETERS

Expected value of a random variable:

A random variable is a numerical outcome from an experiment or from a random procedure. If it is possible to repeat the experiment many times, the numerical outcomes will fluctuate from one experiment to the next because of the variability inherent in the behavior of a random variable.

Although successive numerical outcomes of the random variable will fluctuate, as more and more random outcomes are observed, the numerical average of those outcomes will tend to stabilize.

For instance, if we repeatedly toss a fair die, each successive outcome will be an integer from 1 to 6, but as the number of successive tosses n gets large, if we calculate the average outcome of the n tosses, it will tend toward a constant limit. This is the average value, or **expected value**, or the **mean** of the random variable.

For a random variable X, the expected value (also called the **expectation**) is denoted $\boldsymbol{E[X]}$, or $\boldsymbol{\mu_X}$ or $\boldsymbol{\mu}$. The mean is interpreted as the "average" of the random outcomes.

Mean of a Discrete Random Variable

For a discrete random variable, the expected value of X is $\boldsymbol{\sum x \times p(x) = x_1 \times p(x_1) + x_2 \times p(x_2) + \cdots}$, where the sum is taken over all points x at which X has non-zero probability. For instance, if X is the result of one toss of a fair die, then $E[X] = 1 \times \frac{1}{6} + 2 \times \frac{1}{6} + \cdots + 6 \times \frac{1}{6} = \frac{7}{2}$. The meaning of this value is that if the die is tossed many times, then the "long-run" average of the numbers turning up is $\frac{7}{2}$.

We can see this from another point of view. Suppose that we toss the die n times. Then, "on average", we expect that there will be $\frac{n}{6}$ tosses that are 1, and the same for 2, 3, 4, 5, and 6. Therefore, the average outcome would be the total of all the tosses, divided by n, which is

$$\frac{\frac{n}{6}\times 1+\frac{n}{6}\times 2+\frac{n}{6}\times 3+\frac{n}{6}\times 4+\frac{n}{6}\times 5+\frac{n}{6}\times 6}{n} = 1 \times \frac{1}{6} + 2 \times \frac{1}{6} + \cdots + 6 \times \frac{1}{6} = \frac{7}{2},$$

which is the same as the formal definition of the mean.

Note that the mean of a random variable X is not necessarily one of the possible outcomes for X ($\frac{7}{2}$ is not a possible outcome when tossing a die).

Mean of a Continuous Random Variable

For a continuous random variable, the expected value is $\int_{-\infty}^{\infty} x \times f(x)\,dx$.

Although this integral is written with lower limit $-\infty$ and upper limit ∞, the interval of integration is the interval of non-zero-density for X. For instance, if $f(x) = \begin{cases} 2x \text{ for } 0<x<1 \\ 0, \text{ elsewhere} \end{cases}$,

then $E[X] = \int_0^1 x \times 2x\,dx = \int_0^1 2x^2\,dx = \frac{2}{3}$. Note that even though this random variable is defined on the interval $(0,1)$, the mean is not the midpoint of that interval. This could have been anticipated since the density is higher for x values near 1 than it is for x values near 0. The distribution is weighted more heavily toward 1 than 0. This can also be seen, for instance, by noting that $P[0 < X < \frac{1}{2}] = \frac{1}{4}$, and $P[\frac{1}{2} < X < 1] = \frac{3}{4}$.

The expected value is the "average" over the range of values that X can be, in the sense of the average being the "weighted center" (not necessarily the "geographic center") of the distribution. In the case of the die toss example above, all outcomes were equally likely, so it is not surprising that the mean was in the middle of the possible outcomes. In the continuous example in the previous paragraph, the mean of $\frac{2}{3}$ reflected the higher density for the values close to 1.

Expectation of $h(x)$: If h is a function, then $E[h(X)]$ is equal to $\sum_x h(x) \times p(x)$

if X is a discrete random variable, and it is equal to $\int_{-\infty}^{\infty} h(x) \times f(x)\,dx$ if X is a continuous random variable. Suppose that $h(x) = \sqrt{x}$. Then for the die toss example,

$E[h(X)] = E\left[\sqrt{X}\right] = \sqrt{1} \times \frac{1}{6} + \sqrt{2} \times \frac{1}{6} + \cdots + \sqrt{6} \times \frac{1}{6} = 1.805$.

For the continuous random variable with density function $f(x) = \begin{cases} 2x \text{ for } 0<x<1 \\ 0, \text{ elsewhere} \end{cases}$,

we have $E[h(X)] = E\left[\sqrt{X}\right] = \int_0^1 \sqrt{x} \times 2x\,dx = \frac{4}{5}$.

Example 5-1:

Let X equal the number of tosses of a fair die until the first "1" appears.
Find $E[X]$.

Solution:

X is a discrete random variable that can take on an integer value ≥ 1. The probability that the first 1 appears on the x-th toss is $p(x) = (\frac{5}{6})^{x-1} \times \frac{1}{6}$ for $x \geq 1$

($x-1$ tosses that are not 1 followed by a 1). This is the probability function of X. Then

$E[X] = \sum_{k=1}^{\infty} k \times f(k) = \sum_{k=1}^{\infty} k \times (\frac{5}{6})^{k-1} \times \frac{1}{6} = \frac{1}{6} \times [1 + 2 \times \frac{5}{6} + 3 \times (\frac{5}{6})^2 + \cdots]$.

We use the general increasing geometric series relation $1 + 2r + 3r^2 + \cdots = \frac{1}{(1-r)^2}$,

so that $E[X] = \frac{1}{6} \times \frac{1}{(1-\frac{5}{6})^2} = 6$. □

Example 5-1 used an identity involving an infinite increasing geometric series. It is worthwhile knowing this identity. There are a number of ways to derive the infinite increasing geometric series formula used in Example 5-1. For instance, from the equation $1 + r + r^2 + r^3 + \cdots = \frac{1}{1-r}$, if we differentiate both sides of the equation, we get $1 + 2r + 3r^2 + \cdots = \frac{1}{(1-r)^2}$.

Example 5-2:

A fair die is tossed until the first 1 appears. Let x equal the number of tosses required, $x = 1, 2, 3, \ldots$ You are to receive $(0.5)^x$ dollars if the first 1 appears on the x-th toss. What is the expected amount that you will receive?

Solution:

This is the same distribution as in Example 5-1 above, with the probability that the first 1 appears on the x-th toss being $(\frac{5}{6})^{x-1} \times \frac{1}{6}$ for $x \geq 1$ ($x-1$ tosses that are not 1, followed by a 1), and the amount received in that case is $h(x) = (0.5)^x$. Then, the expected amount received is

$E[h(X)] = E[(0.5)^X] = \sum_{k=1}^{\infty} (.5)^k (\frac{5}{6})^{k-1} (\frac{1}{6}) = (\frac{1}{12})[1 + (\frac{5}{12}) + (\frac{5}{12})^2 + \cdots] = \frac{1}{7}$. □

Moments of a random variable: If $n \geq 1$ is an integer, the ***n*-th moment of *X*** is $\boldsymbol{E[X^n]}$. If the mean of X is μ, then the ***n*-th central moment of *X* (about the mean μ)** is $E[(X-\mu)^n]$.

Example 5-3:

You are given that $\theta > 0$ is a constant, and the density function of X is $f(x) = \theta e^{-x\theta}$, for $x > 0$, and 0 elsewhere. Find the n-th moment of X, where n is a non-negative integer (assuming that $\theta > 0$).

Solution:

The n-th moment of X is $E[X^n] = \int_0^\infty x^n \cdot \theta e^{-x\theta} dx$. Applying integration by parts, this can be written as

$$\int_0^\infty x^n \, d(-e^{-x\theta}) = -x^n e^{-x\theta}\Big|_{x=0}^{x=\infty} - \int_0^\infty -nx^{n-1} e^{-x\theta} \, dx = \int_0^\infty n x^{n-1} e^{-x\theta} dx$$

Repeatedly applying integration by parts results in $E[X^n] = \frac{n!}{\theta^n}$.

An alternative to integration by parts is the method mentioned on page 23.

It is worthwhile noting the general form of the integral that appears in this example; if $k \geq 0$ is an integer and $a > 0$, then by repeated applications of integration by parts, we have $\int_0^\infty t^k \times e^{-at} \, dt = \frac{k!}{a^{k+1}}$.

In this example $\int_0^\infty x^n \theta e^{-x\theta} dx = \theta \int_0^\infty x^n e^{-x\theta} dx = \theta \times \frac{n!}{\theta^{n+1}} = \frac{n!}{\theta^n}$. □

Symmetric Distribution:

If X is a continuous random variable with pdf $f(x)$, and if c is a point for which $f(c+t) = f(c-t)$ for all $t > 0$, then X is said to have a symmetric distribution about the point $x = c$. For such a distribution, the mean will be the point of symmetry, $E[X] = c$. This will be shown in more detail later in the notes, and we will review a couple of specific symmetric distributions.

Variance of X:

Let us go back to the die toss example again. We saw that the mean of X, the outcome of the fair die toss, was $\frac{7}{2}$. Each time the die is tossed, the actual outcome is from 1 to 6, so there will be some "deviation" or distance in the outcome from the mean of $\frac{7}{2}$. Sometimes the deviation will be $\frac{1}{2}$ (for tosses that are 3 or 4), sometimes the deviation will be $\frac{3}{2}$ and sometimes it will be $\frac{5}{2}$.

Now suppose that we consider a modification to this die toss example in which we use a modified die which has three sides each with 1 on them, and 3 sides each with 6 on them. The random variable Y that represents the outcome of the modified die has the following probability distribution: $P[Y=1]=\frac{1}{2},\ P[Y=6]=\frac{1}{2}$. The mean of Y is $\frac{7}{2}$, but each time the modified die is tossed, the deviation of the outcome from the mean is always $\frac{5}{2}$. We see that both the original die and the modified die have the same mean, but the modified die tends to have outcomes that have larger deviation from the mean.

In probability theory, there is a generally accepted way of measuring the deviation from the mean that occurs in a random variable. This is called the **variance** of the random variable. The variance of X is denoted $\boldsymbol{Var[X]}$, $\boldsymbol{V[X]}$, $\boldsymbol{\sigma_X^2}$ or $\boldsymbol{\sigma^2}$. The variance is defined as follows:

$\boldsymbol{Var[X]=E[(X-\mu_X)^2]}$ (the variance is the **2nd central moment of X about its mean**).

It is possible to show that $\boldsymbol{Var[X]=E[X^2]-(E[X])^2=E[X^2]-\mu_X^2}$.

This is usually the most efficient or practical way to calculate variance.

The variance is a measure of the "dispersion" of X about the mean. A large variance indicates significant levels of probability or density for points far from $E[X]$. The variance is always ≥ 0. The variance of X is equal to 0 only if X has a discrete distribution with a single point and probability 1 at that point (which means it is not random at all).

For the original standard die toss example, we have

$$E[X^2]=1^2\times\tfrac{1}{6}+2^2\times\tfrac{1}{6}+\cdots+6^2\times\tfrac{1}{6}=\tfrac{91}{6},\text{ and}$$

$$Var[X]=E[X^2]-(E[X])^2=\tfrac{91}{6}-(\tfrac{7}{2})^2=\tfrac{35}{12}.$$

For the continuous random variable with density function $f(x)=\begin{cases}2x \text{ for } 0<x<1\\ 0, \text{elsewhere}\end{cases}$, we have $E[X^2]=\int_0^1 x^2\times 2x\,dx=\frac{1}{2}$, and $Var[X]=E[X^2]-(E[X])^2=\frac{1}{2}-(\frac{2}{3})^2=\frac{1}{18}$.

Standard deviation of X:

The standard deviation of the random variable X is the square root of the variance, and is denoted $\sigma_X=\sqrt{Var[X]}$. The **coefficient of variation** of X is $\frac{\sigma_X}{\mu_X}$.

Example 5-4:

A continuous random variable X has density function $f_X(x) = \begin{cases} 1-|x| \text{ if } |x|<1 \\ 0, \text{ elsewhere} \end{cases}$.

The continuous random variable W has density function $f_W(w) = \begin{cases} 0.5-0.25|w| \text{ if } |w|<2 \\ 0, \text{ elsewhere} \end{cases}$.

Find the mean and variance of X and W.

Solution:

The density of X is symmetric about 0 (since $|-x| = |x|$, it follows that $f_X(x) = f_X(-x)$), so that $E[X] = 0$. This can be verified directly:

$E[X] = \int_{-1}^{1} x(1-|x|)\,dx = \int_{-1}^{0} x(1+x)\,dx + \int_{0}^{1} x(1-x)\,dx = -\frac{1}{6} + \frac{1}{6} = 0$. Then,

$$Var[X] = E[X^2] - (E[X])^2 = E[X^2] = \int_{-1}^{1} x^2(1-|x|)\,dx = \int_{-1}^{0} x^2(1+x)\,dx + \int_{0}^{1} x^2(1-x)\,dx = \frac{1}{6}.$$

The pdf of W is also symmetric about 0 for the same reason, and has mean $E[W] = 0$, or

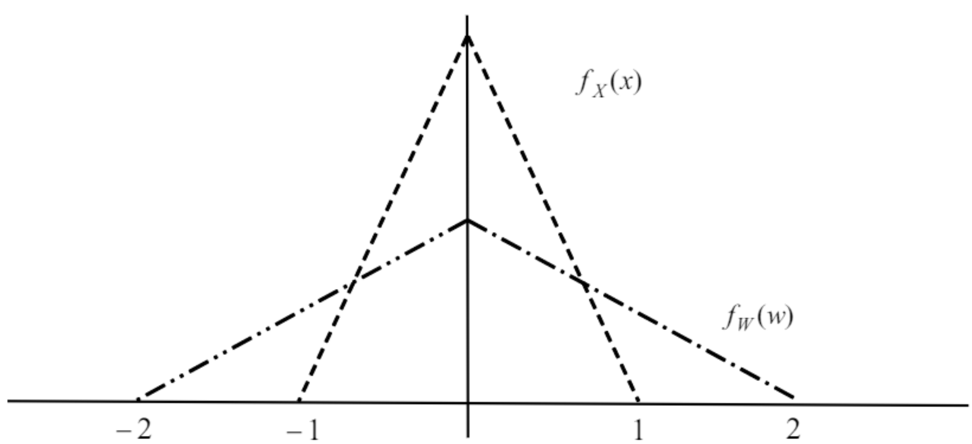
$$E[W] = \int_{-2}^{2} x(0.5-0.25|x|)\,dx = \int_{-2}^{0} x(0.5+0.25x)\,dx + \int_{0}^{2} x(0.5-0.25x)\,dx = -\frac{1}{3} + \frac{1}{3} = 0.$$

Then

$$\begin{aligned} Var[W] &= E[W^2] - (E[W])^2 = E[W^2] = \int_{-2}^{2} x^2(0.5-0.25|x|)\,dx \\ &= \int_{-2}^{0} x^2(0.5+0.25x)\,dx + \int_{0}^{2} x^2(0.5-0.25x)\,dx = \frac{1}{3} + \frac{1}{3} = \frac{2}{3}. \end{aligned}$$

The graphs of the pdfs of X and W are in the diagram on the following page. We see that the pdf of W is more widely dispersed about its mean than the pdf of X is, and so we would anticipate a larger variance for W by comparing the graphs. Comparison of pdf's to determine relative size of variance might not always be as straightforward as it is in this example.

The graphs of the pdfs of X and W are in the following diagram:



□

Moment generating function of random variable X: The moment generating function of X (mgf) is denoted $M_X(t)$, $m_X(t)$, $M(t)$ or $m(t)$, and it is defined to be $\boldsymbol{M_X(t) = E[e^{tX}]}$.

Probability generating function of a random variable X: The probability generating function of X (pgf) is denoted $P_X(t)$, and is defined to be $\boldsymbol{P_X(t) = E[t^X]}$. Probability generating functions usually arise in the context of a discrete random variable.

If X is discrete then $M_X(t) = \Sigma e^{tx} p(x)$ and $P_X(t) = \Sigma t^x p(x)$

If X is continuous then $M_X(t) = \int_{-\infty}^{\infty} e^{tx} f(x)\,dx$ and $P_X(t) = \int_{-\infty}^{\infty} t^x f(x)\,dx$.

Note that $P_X(t) = M_X(ln\, t)$ (for $t > 0$) and $M_X(u) = P_X(e^u)$.

Some important properties of moment generating functions and probability generating functions are:

(i) It is always true that $M_X(0) = 1$ and $P_X(1) = 1$. For instance, in the continuous case,
$M_X(0) = \int_{-\infty}^{\infty} e^{0 \cdot x} \times f(x)\,dx = \int_{-\infty}^{\infty} f(x)\,dx = 1$ and $P_X(1) = \int_{-\infty}^{\infty} 1^x \times f(x)\,dx = 1$.

(ii) The moments of X can be found from the successive derivatives of $M_X(t)$.
$\boldsymbol{M_X'(0) = E[X]}$, $\boldsymbol{M_X''(0) = E[X^2]}$, $M_X^{(n)}(0) = E[X^n]$, $\frac{d^2}{dt^2} ln[M_X(t)]\Big|_{t=0} = Var[X]$
For instance, in the continuous case,
$M_X'(t) = \frac{d}{dt} M_X(t) = \frac{d}{dt} \int_{-\infty}^{\infty} e^{tx} f(x)\,dx = \int_{-\infty}^{\infty} \frac{d}{dt} e^{tx} f(x)\,dx = \int_{-\infty}^{\infty} x\, e^{tx} f(x)\,dx$
so that
$M_X'(0) = \int_{-\infty}^{\infty} x\, e^{0 \times x} f(x)\,dx = \int_{-\infty}^{\infty} x\, f(x)\,dx = E[X]$
$M_X''(t) = \frac{d^2}{dt^2} M_X(t) = \frac{d^2}{dt^2} \int_{-\infty}^{\infty} e^{tx} f(x)\,dx = \int_{-\infty}^{\infty} \frac{d^2}{dt^2} e^{tx} f(x)\,dx = \int_{-\infty}^{\infty} x^2\, e^{tx} f(x)\,dx$
so that
$M_X''(0) = \int_{-\infty}^{\infty} x^2\, e^{0 \times x} f(x)\,dx = \int_{-\infty}^{\infty} x^2\, f(x)\,dx = E[X^2]$
Similar derivations apply for higher moments and also in the discrete case.
(Notice that we have implicitly assumed that we can take the derivative "inside" the integral $\frac{d}{dt} \int_{-\infty}^{\infty} = \int_{-\infty}^{\infty} \frac{d}{dt}$; this requires some conditions on the function $f(x)$ which are generally satisfied by the standard functions we will encounter.).

(iii) If X has a discrete non-negative integer distribution with $p_k = P[X = k]$, then the probability generating function is $P_X(t) = p_0 + p_1 \times t + p_2 \times t^2 + p_3 \times t^3 + \cdots = \sum_{k=0}^{\infty} p_k \times t^k$.
The probabilities of individual integer outcomes can be found as follows.
$P_X(0) = p_0$. $P_X'(t) = p_1 + 2p_2 \cdot t + 3p_3 t^2 + \cdots$, so that $P_X'(0) = p_1$
$P_X''(t) = 2! \times p_2 + 3 \times 2 \times p_3 t + \cdots$, so that $P_X''(0) = 2p_2$ and $p_2 = \frac{1}{2!} \times P_X''(0)$
Continuing in this way, we get $p_k = \frac{1}{k!} \times P_X^{(k)}(0)$ (k-th derivative evaluated at $t = 0$).

Note that for a continuous random variable X, $P_X'(t) = \int_{-\infty}^{\infty} x \times t^{x-1} \times f(x)\,dx$, so that $P_X'(1) = \int_{-\infty}^{\infty} x \times f(x)\,dx = E[X]$. This is also valid for a discrete random variable.
In a similar way it can be shown that $P_X''(1) = E[X^2 - X] = E[X^2] - E[X]$.

(iv) The moment generating function of X might not exist for all real numbers, but usually exists on some interval of real numbers.

Example 5-5:
The pdf of X is $f(x) = 5e^{-5x}$ for $x > 0$. Find the moment generating function of X and use it to find the first and second moments of X, and the variance of X.
Solution:
$M_X(t) = \int_0^{\infty} e^{tx} \times 5e^{-5x}\,dx = 5\int_0^{\infty} e^{-(5-t)x}\,dx = \frac{5}{5-t}$
(we have used the integration rule $\int_0^{\infty} e^{-at}\,dt = \frac{1}{a}$ if $a > 0$).
This integration is valid if $5 - t > 0$, or equivalently, if $t < 5$.
Then $\frac{d}{dt} M_X(t)\Big|_{t=0} = M_X'(0) = \frac{5}{(5-0)^2} = \frac{1}{5} = E[X]$
and $\frac{d^2}{dt^2} M_X(t)\Big|_{t=0} = M_X''(0) = \frac{2\times 5}{(5-0)^3} = \frac{2}{25} = E[X^2]$.
The variance of X is $Var[X] = E[X^2] - (E[X])^2 = \frac{1}{25}$. □

Example 5-6:
Find the moment generating function and probability generating function for each of these two random variables.
(i) X = outcome of a die toss, $p(x) = P[X = x] = \frac{1}{6}$ for $x = 1, 2, 3, 4, 5, 6$.
(ii) X is a continuous random variable with density function $f(x) = \begin{cases} 2x \text{ for } 0<x<1 \\ 0, \text{ elsewhere} \end{cases}$.
Solution:
(i) $M_X(t) = E[e^{tX}] = \sum_{x=1}^{6} e^{tx} \times p(x)$
$= e^t \times \frac{1}{6} + e^{2t} \times \frac{1}{6} + e^{3t} \times \frac{1}{6} + e^{4t} \times \frac{1}{6} + e^{5t} \times \frac{1}{6} + e^{6t} \times \frac{1}{6} = \frac{1}{6}e^t \times (\frac{e^{6t}-1}{e^t-1})$
Note that $M_X'(t) = e^t \times \frac{1}{6} + 2e^{2t} \times \frac{1}{6} + 3e^{3t} \times \frac{1}{6} + 4e^{4t} \times \frac{1}{6} + 5e^{5t} \times \frac{1}{6} + 6e^{6t} \times \frac{1}{6}$ and
$\frac{d}{dt} M_X(t)\Big|_{t=0} = M_X'(0) = 1 \times \frac{1}{6} + 2 \times \frac{1}{6} + 3 \times \frac{1}{6} + 4 \times \frac{1}{6} + 5 \times \frac{1}{6} + 6 \times \frac{1}{6} = \frac{7}{2} = E[X]$. Also
$\frac{d^2}{dt^2} M_X(t)\Big|_{t=0} = M_X''(0) = 1^2 \times \frac{1}{6} + 2^2 \times \frac{1}{6} + 3^2 \times \frac{1}{6} + 4^2 \times \frac{1}{6} + 5^2 \times \frac{1}{6} + 6^2 \times \frac{1}{6} = \frac{91}{6} = E[X^2]$
The variance of X is $E[X^2] - (E[X])^2 = \frac{91}{6} - (\frac{7}{2})^2 = \frac{35}{12}$.
$P_X(t) = E[t^X] = [t^1 + t^2 + t^3 + t^4 + t^5 + t^6] \times \frac{1}{6}$.
Note that $P_X(0) = 0 = p_0$, $P_X'(t) = [1 + 2t + 3t^2 + 4t^3 + 5t^4 + 6t^5] \times \frac{1}{6}$
so that $P_X'(0) = \frac{1}{6} = p_1$, and $P_X''(t) = [2 + 6t + 12t^2 + 20t^3 + 30t^4] \times \frac{1}{6}$
so that $P_X''(0) = \frac{2}{6} = 2! \times p_2$ and $p_2 = \frac{1}{6}$.

Note also that $P'_X(1) = [1+2+3+4+5+6] \times \frac{1}{6} = \frac{7}{2} = E[X]$, and
$P''_X(1) = [2+6+12+20+30] \times \frac{1}{6} = \frac{35}{3} = E[X^2] - E[X] = E[X^2] - \frac{7}{2}$, so that $E[X^2] = \frac{91}{6}$
and $Var[X] = \frac{91}{6} - (\frac{7}{2})^2 = \frac{35}{12}$.

(ii) $M_X(t) = \int_0^1 e^{tx} \times f(x)\,dx = \int_0^1 e^{tx} \times 2x\,dx = 2(\frac{xe^{tx}}{t} - \frac{e^{tx}}{t^2})\Big|_{x=0}^{x=1} = 2(\frac{e^t}{t} - \frac{e^t-1}{t^2}) = M_X(t)$
Then $M'_X(t) = 2(\frac{te^t-e^t}{t^2} - \frac{t^2e^t-2te^t+2t}{t^4}) = 2(\frac{t^2e^t-2te^t+2e^t-2}{t^3})$.
Note that $M'_X(0)$ is found as a limit, $\lim_{t\to 0} 2(\frac{t^2e^t-2te^t+2e^t-2}{t^3}) = \frac{2}{3} = E[X]$ (by l'Hospital's rule; note that in some references spell "l'Hospital" without an "s" as "l'Hôpital").

The antiderivative of $e^{tx} \times 2x$ was found by integration by parts. A useful point to note is the general antiderivative $\boldsymbol{\int xe^{cx}\,dx = \frac{xe^{cx}}{c}} - \boldsymbol{\frac{e^{cx}}{c^2}}$, if c is a constant. This antiderivative has come up from time to time on previous exams. It is much more straightforward to find $E[X]$ and $E[X^2]$ directly as $\int_0^1 x \times f(x)\,dx = \int_0^1 x \times 2x\,dx = \int_0^1 2x^2\,dx = \frac{2}{3}$ and $\int_0^1 x^2 \times f(x)\,dx$
for the random variable in part (ii).

$P_X(t) = \int_0^1 t^x \times f(x)\,dx = \int_0^1 t^x \times 2x\,dx$. We can do an integration by parts to find the integral, or we can use the relationship $P_X(t) = M_X(ln\,t) = 2(\frac{(ln\,t)^2 \times t - 2(ln\,t) \times \cdot t + 2t - 2}{(ln\,t)^3})$. The Probability generating function is less convenient to work with than the moment generating function in this case.

□

Example 5-7:

The moment generating function of X is given as $\frac{\alpha}{\alpha-t}$ for $t < \alpha$, where $\alpha > 0$.
Find $Var[X]$.

Solution:

$Var[X] = E[X^2] - (E[X])^2$ and $E[X] = M'_X(0) = \frac{\alpha}{(\alpha-t)^2}\Big|_{t=0} = \frac{1}{\alpha}$
and $E[X^2] = M''_X(0) = \frac{2\alpha}{(\alpha-t)^3}\Big|_{t=0} = \frac{2}{\alpha^2} \rightarrow Var[X] = \frac{2}{\alpha^2} - (\frac{1}{\alpha})^2 = \frac{1}{\alpha^2}$.
Alternatively, $ln\,M_X(t) = ln(\frac{\alpha}{\alpha-t}) = ln\,\alpha - ln(\alpha - t) \rightarrow \frac{d}{dt}\,ln[M_X(t)] = \frac{1}{\alpha-t}$
and $\frac{d^2}{dt^2}\,ln[M_X(t)] = \frac{1}{(\alpha-t)^2}$ so that $Var[X] = \frac{d^2}{dt^2}\,ln[M_X(t)]\Big|_{t=0} = \frac{1}{\alpha^2}$.
This is like Example 5-5 above, with 5 replaced by α. □

Example 5-8:

The probability generating function of a non-negative integer-valued random variable X is given as

$$P_X(t) = [\tfrac{t}{3} + \tfrac{2}{3}]^3$$

Find the probability function, the mean and the variance of X.

Solution:

$P_X(0) = p_0 = (\frac{2}{3})^3 = \frac{8}{27}$, $P_X'(t) = 3 \times [\frac{t}{3} + \frac{2}{3}]^2 \times \frac{1}{3} = [\frac{t}{3} + \frac{2}{3}]^2 \to P_X'(0) = p_1 = (\frac{2}{3})^2 = \frac{4}{9}$

$\frac{d}{dt^2} P_X(t) = 2 \cdot [\frac{t}{3} + \frac{2}{3}] \times \frac{1}{3} = \frac{2t}{9} + \frac{4}{9} = 2p_2 \to p_2 = \frac{2}{9}$

$\frac{d}{dt^3} P_X(t) = \frac{2}{9} = 6p_3 \to p_3 = \frac{1}{27}$

and $\frac{d}{dt^k} P_X(t) = 0$ for $k = 4, 5, \ldots$ so that $p_k = 0$ for $k = 4, 5, \ldots$

$E[X] = \frac{8}{27} \times 0 + \frac{4}{9} \times 1 + \frac{2}{9} \times 2 + \frac{1}{27} \times 3 = 1$, or $E[X] = P_X'(1) = [\frac{1}{3} + \frac{2}{3}]^2 = 1$

$E[X^2] = \frac{8}{27} \times 0^2 + \frac{4}{9} \times 1^2 + \frac{2}{9} \times 2^2 + \frac{1}{27} \times 3^2 = \frac{5}{3}$ and $Var[X] = \frac{5}{3} - 1^2 = \frac{2}{3}$, or

$E[X^2] - E[X] = P_X''(1) = \frac{2}{9} + \frac{4}{9} = \frac{2}{3} \to E[X^2] = \frac{2}{3} + E[X] = \frac{5}{3}$ as before. $\square$

Percentiles of a distribution:

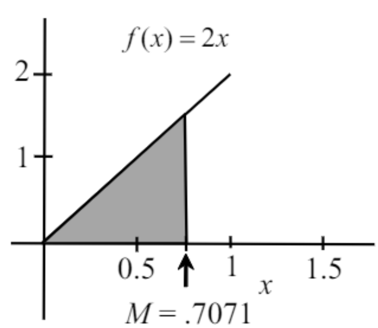
If $0 < p < 1$, then the $100p$-th percentile of the distribution of X is the number c_p which satisfies both of the following inequalities: $P[X \leq c_p] \geq p$ and $P[X \geq c_p] \geq 1 - p$

For a continuous random variable, it is sufficient to find the c_p for which $P[X \leq c_p] = p$. If $p = .5$, **the 50-th percentile of a distribution is referred to as the median of the distribution**; it is the point M for which $P[X \leq M] = 0.5$. The median M is the 50% probability point, half of the distribution probability is to the left of M and half is to the right. If X has a symmetric distribution about the point $x = c$, then the mean and the median of X will be equal to c.

For the continuous random variable with density function $f(x) = \begin{cases} 2x \text{ for } 0<x<1 \\ 0, \text{ elsewhere} \end{cases}$,

the median is M, where $\int_0^M 2x\,dx = M^2 = 0.5$, so that $M = \sqrt{0.5} = 0.7071$.

This is illustrated in the graph below. The shaded area below has probability 0.5, and is to the left of $M = 0.7071$, the median.



For trial use only.
Content may not
represent the most
current version

Example 5-9:

The continuous random variable X has pdf $f(x) = \frac{1}{2} \times e^{-|x|}$ for $-\infty < x < \infty$. Find the 87.5-th percentile of the distribution.

Solution:

The 87.5-th percentile is the number b for which

$0.875 = P[X \leq b] = \int_{-\infty}^{b} f(x)\,dx = \int_{-\infty}^{b} \frac{1}{2} \times e^{-|x|}\,dx$.

Note that this distribution is symmetric about 0, since $f(-x) = f(x)$, so the mean and median are both 0. Thus, $b > 0$, and so

$\int_{-\infty}^{b} \frac{1}{2} \times e^{-|x|}\,dx = \int_{-\infty}^{0} \frac{1}{2} \times e^{-|x|}\,dx + \int_{0}^{b} \frac{1}{2} \times e^{-|x|}\,dx = 0.5 + \int_{0}^{b} \frac{1}{2} \times e^{-x}\,dx$

$= 0.5 + \frac{1}{2}(1 - e^{-b}) = 0.875 \Rightarrow b = -ln(0.25) = ln\,4$. □

The mode of a distribution:

The mode is any point m at which the probability or density function $f(x)$ is maximized. The mode of the distribution in the graph at the bottom of the previous page is 1, since the maximum value of $f(x)$ occurs at $x = 1$. The mode of the distribution in Example 5-9 occurs at $x = 0$.

The skewness of a distribution:

If the mean of random variable X is μ and the variance is σ^2 then the skewness is defined to be $E[(X-\mu)^3]/\sigma^3$. If skewness is positive, the distribution is said to be skewed to the right, and if skewness is negative it is skewed to the left.

Some results and formulas relating to distribution moments:

(i) The mean of a random variable X might not exist, it might be $+\infty$ or $-\infty$, and the variance of X might be $+\infty$. For example, the continuous random variable X with pdf $f(x) = \begin{cases} \frac{1}{x^2} \text{ for } x \geq 1 \\ 0, \text{ otherwise} \end{cases}$ has expected value $\int_1^{\infty} x \times \frac{1}{x^2}\,dx = +\infty$.

(ii) For any constants a_1, a_2 and b and functions h_1 and h_2,
$\boldsymbol{E[a_1 h_1(X) + a_2 h_2(X) + b] = a_1 E[h_1(X)] + a_2 E[h_2(X)] + b}$.
As a special case, $\boldsymbol{E[aX + b] = aE[X] + b}$.

(iii) If X is a random variable defined on the interval $[a, \infty)$ ($f(x) = 0$ for $x < a$), then $E[X] = a + \int_a^{\infty} [1 - F(x)]\,dx$, and if X is defined on the interval $[a, b]$, where $b < \infty$, then $E[X] = a + \int_a^{b} [1 - F(x)]\,dx$. This relationship is valid for any random variable, discrete, continuous or with a mixed distribution. As a special, case, if X is a non-negative random variable (defined on $[0, \infty)$ or $(0, \infty)$) then $E[X] = \int_0^{\infty} [1 - F(x)]\,dx$.

(iv) Jensen's inequality: If h is a function and X is a random variable such that $\frac{d^2}{dx^2}\, h(x) = h''(x) \geq 0$ at all points x with non-zero density or probability for X, then $E[h(X)] \geq h(E[X])$, and if $h'' > 0$ then $E[h(X)] > h(E[X])$. The inequality reverses if $h'' \leq 0$. For example, if $h(x) = x^2$, then $h''(x) = 2 \geq 0$ for any x, so that $E[X^2] \geq (E[X])^2$ (this is also true since $Var[X] = E[X^2] - (E[X])^2 \geq 0$ for any random variable X). As another example, if X is a positive random variable (i.e., X has non-zero density or probability only for $x \geq 0$), and $h(x) = \sqrt{x}$, then $h''(x) = \frac{-1}{4x^{3/2}} < 0$ for $x > 0$, and it follows from Jensen's inequality that $E\left[\sqrt{X}\right] < \sqrt{E[X]}$.

(v) If a and b are constants, then $Var[aX+b] = a^2 Var[X]$.

(vi) Chebyshev's inequality: If X is a random variable with mean μ_X and standard deviation σ_X, then for any real number $r > 0$, $P[|X - \mu_X| > r\sigma_X] \leq \frac{1}{r^2}$.

(vii) Suppose that for the random variable X, the moment generating function $M_X(t)$ exists in an interval containing the point $t = 0$. Then $\boldsymbol{\frac{d^n}{dt^n}\, M_X(t)\Big|_{t=0} = M_X^{(n)}(0) = E[X^n]}$, the n-th moment of X, and $\frac{d}{dt}\, ln[M_X(t)]\Big|_{t=0} = \frac{M_X'(0)}{M_X(0)} = E[X]$, and $\frac{d^2}{dt^2}\, ln[M_X(t)]\Big|_{t=0} = Var[X]$.

The Taylor series expansion of $M_X(t)$ expanded about the point $t = 0$ is

$$M_X(t) = \sum_{k=0}^{\infty} \frac{t^k}{k!}\, E[X^k] = 1 + t \times E[X] + \frac{t^2}{2} \times E[X^2] + \frac{t^3}{6} \times E[X^3] + \cdots$$

Therefore, if we are given a moment generating function and we are able to formulate the Taylor series expansion about the point $t = 0$, we can identify the successive moments of X.

If X has a discrete distribution with probability space $\{x_1, x_2, x_3, ...\}$ and probability function $P(X = x_k) = p_k$, then the moment generating function is $M_X(t) = e^{tx_1} \times p_1 + e^{tx_2} \times p_2 + e^{tx_3} \times p_3 + \cdots$.

Conversely, if we are given a moment generating function in this form (a sum of exponential factors), then we can identify the points of probability and their probabilities. This is illustrated in Example 15-18 below.

If X_1 and X_2 are random variables, and $M_{X_1}(t) = M_{X_2}(t)$ for all values of t in an interval containing $t = 0$, then X_1 and X_2 have identical probability distributions. The same is true for the probability generating function.

(viii) The median (50-th percentile) and other percentiles of a distribution are not always unique. For example, if X is the discrete random variable with probability function $f(x) = .25$ for $x = 1, 2, 3, 4$, then the median of X would be any point from 2 to 3, but the usual convention is to set the median to be the midpoint between the two "middle" values of X, $M = 2.5$.

(ix) The distribution of the random variable X is said to be **symmetric about the point c** if $f(c+t)=f(c-t)$ for any value of t. It follows that the expected value of X and the median of X is c. Also, for a symmetric distribution, any odd-order central moments about the mean are 0, this means that $E[(X-\mu)^k]=0$ if k is an odd integer ≥ 1.

To see that the mean is c

$E[X]=\int_{-\infty}^{\infty} x\times f(x)\,dx=\int_{-\infty}^{c} x\times f(x)\,dx\,+\,\int_{c}^{\infty} x\times f(x)\,dx.$

If we apply the change of variable $t=x-c$ to the each integral on the right, the first becomes $\int_{-\infty}^{0}(c+t)\times f(c+t)\,dt\,=c\int_{-\infty}^{0}f(c+t)\,dt+\int_{-\infty}^{0}t\times f(c+t)\,dt$, and the second becomes $\int_{0}^{\infty}(c+t)\times f(c+t)\,dt\,=c\int_{0}^{\infty}f(c+t)\,dt+\int_{0}^{\infty}t\times f(c+t)\,dt$.

Then, $\int_{-\infty}^{0}f(c+t)\,dt+\int_{0}^{\infty}f(c+t)\,dt=\int_{-\infty}^{\infty}f(c+t)\,dt=1$ (this can be seen if we change the variable to $u=c+t$, and the integral becomes $\int_{-\infty}^{\infty}f(u)\,du=1$ since f is a pdf). Also, $\int_{-\infty}^{0}t\times f(c+t)\,dt\,+\int_{0}^{\infty}t\times f(c+t)\,dt=0$ (this can be seen if we change the variable in the first integral to $u=\,-t$, so the first integral becomes $-\int_{0}^{\infty}u\times f(c+u)\,du$, which is the negative of the second integral.)

(x) If $E[X]=\mu$, $Var[X]=\sigma^2$ and $Z=\frac{X-\mu}{\sigma}$, then $E[Z]=0$ and $Var[Z]=1$.

(xi) **A "mixture" of distributions:** Given any finite collection of random variables, $X_1,\,X_2,\ldots,\,X_k$ with density or probability functions, say $f_1(x),\,f_2(x),\,\ldots f_k(x)$, where k is a non-negative integer, and given a set of "weights", $\alpha_1,\,\alpha_2,\,\ldots,\,\alpha_k$, where $0\leq\alpha_i\leq 1$ for each i and $\sum_{i=1}^{k}\alpha_i=1$, it is possible to construct a new density function:

$f(x)=\alpha_1 f_1(x)+\alpha_2 f_2(x)+\cdots+\alpha_k f_k(x)$, which is a "weighted average" of the original density functions. It then follows that the resulting distribution X, whose density/probability function is f, has moments and moment generating function which are weighted averages of the original distribution moments and moment generating functions:

$E[X^n]=\alpha_1 E[X_1^n]+\alpha_2 E[X_2^n]+\cdots+\alpha_k E[X_k^n]$ and

$M_X(t)=\alpha_1 M_{X_1}(t)+\alpha_2 M_{X_2}(t)+\cdots+\alpha_k M_{X_k}(t)$.

A mixed distribution was defined in Section 4. The definition given here also applies in that case. This Mixtures of distributions will be considered again later in the study guide.

Example 5-10:

The skewness of a random variable is defined to be $\frac{E[(X-\mu)^3]}{\sigma^3}$, where $\mu=E[X]$ and $\sigma^2=Var[X]$. Find the skewness of the random variable X with pdf

$f(x)=\begin{cases}2x \text{ for } 0<x<1\\ 0\text{, elsewhere}\end{cases}$.

Solution:

$\mu = E[X] = \int_0^1 x \times (2x)\,dx = \frac{2}{3}$, $E[X^2] = \int_0^1 x^2 \times (2x)\,dx = \frac{1}{2}$.

$\sigma^2 = Var[X] = E[X^2] - (E[X])^2 = \frac{1}{2} - (\frac{2}{3})^2 = \frac{1}{18}$.

$E[(X-\mu)^3] = E[(X-\frac{2}{3})^3] = \int_0^1 (x-\frac{2}{3})^3 \times 2x\,dx$

$= \int_0^1 (x^3 - 2x^2 + \frac{4}{3}x - \frac{8}{27}) \times 2x\,dx = -\frac{1}{135}$.

Then $\frac{E[(X-\mu)^3]}{\sigma^3} = \frac{-1/135}{(1/18)^{3/2}} = -0.566$.

Since the skewness is negative, this random variable is said to be skewed to the left (of its mean). □

Example 5-11:

In Neverland there is a presidential election every year. The president must put his investments into a blind trust that earns compound interest at rate 10% (compounded annually). Neverland has no term limits for elected officials. The current president was just elected and has a blind trust worth $1,000,000 right now. The current president is very popular and he assesses his chance of being re-elected to be .75 each year from now on. The next election is one year from now, and elections will continue every year. Assuming the re-election probability stays the same year after year, find the expected value of the blind trust when the current president first loses an election in the future.

Solution: Let X denote the number of years until the current president is not re-elected.

The distribution of X is

X :	1	2	3	$\dots$	n	$\dots$
$p(x)$:	0.25	0.75×0.25	$(0.75)^2 \times 0.25$	$\dots$	$(0.75)^{n-1} \times 0.25$	$\dots$
Blind Trust	$1.1M$	$(1.1)^2 M$	$(1.1)^3 M$		$(1.1)^n M$	

The expected value of the blind trust at the time the current president first loses an election is

$$E[(1.1)^X] = (1.1) \times 0.25 + (1.1)^2 \times 0.75 \times 0.25 + (1.1)^3 \times (0.75)^2 \times 0.25 + \cdots$$

$$= (1.1) \times 0.25 \times [1 + (1.1) \times 0.75 + ((1.1)(.75))^2 + \cdots] = (1.1)(0.25)[\tfrac{1}{1-(1.1)(0.75)}] = 1.57$$

(million). Note that there is implicit assumption that there is no upper limit on how long the current president can survive. □

Example 5-12:

Smith finds the carnival game "over-under-seven" irresistible. The game involves the random toss of two fair dice. If a player bets 1 on "over" and the total on the dice is over 7, then the player wins 1 (otherwise he loses the 1 he bet). If he bets 1 on "under" and the total on the dice is under 7 then he wins 1 (otherwise he loses). If he bets 1 on "seven" and the total on the dice is 7 then he wins 4 (otherwise he loses 1). Smith devises the following strategy. His first bet is 1 on "under." If he wins, he walks away with his net gain of 1. If he loses, he doubles his bet to 2 on "under". If he wins, he walks away with his net gain of 1. If he loses, he doubles his bet to 4 and bets on "under", etc. Smith walks away as soon as he wins. Find his expected gain, and the number of bets it will take to win, on average.

Solution:

Smith's net gain is 1 as soon as he first wins. Therefore, whenever he wins, his net gain is 1, so his expected gain is 1. The probability of winning when betting "under" is $\frac{15}{36} = \frac{5}{12}$, since there are 15 winning dice combinations out of a total of $6 \times 6 = 36$ dice combinations.

1-1 , 1-2 , 1-3 , 1-4 , 1-5 , 2-1 , 2-2 , 2-3 , 2-4 , 3-1 , 3-2 , 3-3 , 4-1 , 4-2 , 5-1

are winners for "under". Let X be the bet number of his first win. The probability function for X is

X:	1	2	3	...	n
$p(x)$	$\frac{5}{12}$	$(\frac{7}{12})(\frac{5}{12})$	$(\frac{7}{12})^2(\frac{5}{12})$	...	$(\frac{7}{12})^{n-1}(\frac{5}{12})$

Then $E[X] = \sum_{n=1}^{\infty} n \times (\frac{7}{12})^{n-1}(\frac{5}{12}) = (\frac{5}{12}) \times [1 + 2(\frac{7}{12}) + 3(\frac{7}{12})^2 + \cdots]$

$= (\frac{5}{12}) \times [\frac{1}{(1-\frac{7}{12})^2}] = \frac{12}{5} = 2.4.$

We have used the identity $1 + 2a + 3a^2 + \cdots = \frac{1}{(1-a)^2}$ for $|a| < 1$. □

Example 5-13:

The pdf of X_1 is $f_1(x) = e^{-x}$ for $x > 0$ and the pdf for X_2 is
$f_2(x) = 2e^{-2x}$ for $x > 0$. We define a new random variable X with pdf
$f(x) = 0.5e^{-x} + e^{-2x}$ for $x > 0$. Find the mean and variance of X.

Solution:

$f(x) = (0.5)(e^{-x}) + (0.5)(2e^{-2x})$, which shows that X is a mixture of X_1 and X_2 with mixing weights .5 for X_1 and .5 for X_2. The first and second moments of X_1 and X_2 are

$E[X_1] = \int_0^\infty x \times e^{-x}\,dx = -xe^{-x} - e^{-x}\Big|_{x=0}^{x=\infty} = 1$

$E[X_1^2] = \int_0^\infty x^2 \times e^{-x}\,dx = -x^2e^{-x} - 2xe^{-x} - 2e^{-x}\Big|_{x=0}^{x=\infty} = 2$

$E[X_2] = \int_0^\infty x \times 2e^{-2x}\,dx = -xe^{-2x} - \frac{1}{2}e^{-x}\Big|_{x=0}^{x=\infty} = \frac{1}{2}$

$E[X_2^2] = \int_0^\infty x^2 \times 2e^{-2x}\,dx = -x^2e^{-2x} - xe^{-2x} - \frac{1}{2}e^{-2x}\Big|_{x=0}^{x=\infty} = \frac{1}{2}$

Then $E[X] = \int_0^\infty x \times f(x)\,dx = \int_0^\infty x \times [0.5 \times f_1(x) + 0.5 \times f_2(x)]\,dx$

$= 0.5 \times E[X_1] + 0.5 \times E[X_2] = 0.5 \times 1 + 0.5 \times \frac{1}{2} = 0.75$

and similarly, $E[X^2] = 0.5 \times E[X_1^2] + 0.5 \times E[X_2^2] = 0.5 \times 2 + 0.5 \times \frac{1}{2} = 1.25$.

Then, $Var[X] = E[X^2] - (E[X])^2 = 1.25 - (0.75)^2 = 0.6875.$

Note that $\boldsymbol{Var[X]}$ is <u>not</u> $\boldsymbol{0.5 \times Var[X_1] + 0.5 \times Var[X_2]}$. □

Example 5-14:

The pdf of X is $f_X(x) = \begin{cases} 2x \text{ for } 0<x<1 \\ 0, \text{ otherwise} \end{cases}$.

Find the mean and cdf of the conditional distribution of X given that $X \leq \frac{1}{2}$.

Solution:

If A is the event that $X \leq \frac{1}{2}$, then $P(A) = P[X \leq \frac{1}{2}] = \frac{1}{4}$,

and $f_{X|A}(x|X \leq \frac{1}{2}) = \frac{2x}{1/4} = 8x$, for $0 < x \leq \frac{1}{2}$, and $f_{X|A}(x|X \leq \frac{1}{2}) = 0$, otherwise.

$E[X|X \leq \frac{1}{2}] = \int_0^{1/2} x \times f_{X|A}(x|X \leq \frac{1}{2})\,dx = \int_0^{1/2} x \times 8x\,dx = \frac{1}{3}$.

The cdf of the conditional distribution is

$F_{X|A}(t) = P[X \leq t|X \leq \frac{1}{2}] = \int_0^t f_{X|A}(x)\,dx = \int_0^t 8x\,dx = 4t^2$ for $0 \leq t \leq \frac{1}{2}$

and $F_{X|A}(t) = 1$ for $t > \frac{1}{2}$. Note also that we can find the cdf of the conditional distribution from

$F_{X|A}(t) = P[X \leq t|X \leq \frac{1}{2}] = \frac{P[X \leq t \cap X \leq \frac{1}{2}]}{P[X \leq \frac{1}{2}]} = \frac{t^2}{1/4}$ for $t \leq \frac{1}{2}$. $\square$

Example 5-15:

The continuous random variable X has pdf $f(x) = 1$ for $0 < x < 1$.

X_1, X_2 and X_3 are independent random variables, all with the same distribution as X.

$Y = max\{X_1, X_2, X_3\}$ and $Z = min\{X_1, X_2, X_3\}$. Find $E[Y - Z]$.

Solution:

When dealing with the maximum or minimum of a collection of random variables, it is usually most efficient to work with the cdf F or the survival function $S = 1 - F$.

The cdf of Y is $F_Y(y) = P(Y \leq y) = P(max\{X_1, X_2, X_3\} \leq y)$.

In order for the inequality $max\{X_1, X_2, X_3\} \leq y$ to be true, it must be true that each of $X_1, X_2,$ and X_3 are $\leq y$.

Therefore, $P(max\{X_1, X_2, X_3\} \leq y) = P(X_1 \leq y \cap X_2 \leq y \cap X_3 \leq y)$.

Since X_1, X_2 and X_3 are independent random variables, this is equal to

$P(X_1 \leq y) \times P(X_2 \leq y) \times P(X_3 \leq y) = [F_X(y)]^3$.

From the definition of the distribution of X, we have $F_X(x) = P(X \leq x) = \int_0^x 1\,dt = x$.

It follows that $F_Y(y) = y^3$. Since X is defined on the interval $0 < x < 1$, the same is true for Y. We can find $E[Y]$ by first finding $f_Y(y) = F'_Y(y) = 3y^2$, and then

$E[Y] = \int_0^1 y \times f_Y(y)\,dy = \int_0^1 y \times 3y^2\,dy = \int_0^1 3y^3\,dy = \frac{3}{4}$.

Since $Y \geq 0$ (Y is a non-negative random variable), we can also formulate the mean of Y as

$E[Y] = \int_0^\infty [1 - F_Y(y)]\,dy$. We note that $F_Y(y) = y^3$ for $y \geq 1$, and therefore,

$E[Y] = \int_0^1 (1 - y^3)\,dy = \frac{3}{4}$.

The survival function of Z is $S_Z(z) = P(Z > z)] = P(min\{X_1, X_2, X_3\} > z)$.

In order for the inequality $min\{X_1, X_2, X_3\} > z$ to be true, it must be true that each of $X_1, X_2,$ and X_3 are $> z$.

Therefore, $P(min\{X_1, X_2, X_3\} > z) = P(X_1 > z \cap X_2 > z \cap X_3 > z)$.

Since X_1, X_2 and X_3 are independent random variables, $S_Z(z)$ is equal to

$S_Z(z) = P(X_1 > z) \times P(X_2 > z) \times P(X_3 > z) = [S_X(z)]^3 = [1 - F_X(z)]^3 = [1 - z]^3$.

Then $f_Z(z) = F'_Z(z) = -S'_Z(z) = 3(1-z)^2$ (using the chain rule), and Z is defined on the interval $0 < z < 1$.

Then, $E[Z] = \int_0^1 z \times f_Z(z)\,dz = \int_0^1 z \times 3(1-z)^2\,dz = \frac{1}{4}$.
Finally, $E[Y-Z] = E[Y] - E[Z] = \frac{3}{4} - \frac{1}{4} = \frac{1}{2}$.
Alternatively, since $F_Z(z) = 1$ for $z \geq 1$, we have
$E[Z] = \int_0^\infty [1 - F_Z(z)]\,dz = \int_0^1 [1 - F_Z(z)]\,dx = \int_0^1 S_Z(z)\,dz = \int_0^1 (1-z)^3\,dz = \frac{1}{4}$. □

Example 5-16:

Smith and Jones both love to gamble. Smith finds a casino that offers a game of chance in which a fair coin is tossed until the first head appears. If the first head appears on toss number X $(X = 1, 2, 3, ...)$, the game pays out \$3000$(1 - \frac{1}{2^X})$. The casino charges \$2100 to play the game. Jones realizes that, on average, the first head will occur on the 2nd toss, so he estimates that the on average, the payout on a game will be $3000(1 - \frac{1}{2^2}) = 2250$, and thinks this is a good game to play. Smith, who is a deeper thinker than Jones, uses Jensen's inequality to get an idea of what the expected payout will be. Smith then calculates the exact expected payout. Describe Smith's conclusions:

Solution:

The amount paid out in the game is $h(x) = 3000(1 - \frac{1}{2^x})$.

Smith notices that $h'(x) = 3000 \times \frac{1}{2^x} \times \ln 2$ and $h''(x) = -3000 \times \frac{1}{2^x} \times (\ln 2)^2 < 0$.

It follows from Jensen's inequality that $E[h(X)] < h(E[X])$.

$E[h(X)] = E[3000(1 - \frac{1}{2^X})]$ is the expected payout, and

$h(E[X]) = 3000(1 - \frac{1}{2^{E[X]}}) = 3000(1 - \frac{1}{2^2}) = 2250$

(we can find $E[X]$ in a way similar to the method applied in Example 5-1 above, where we found the expected toss number of the first "1" when a fair die is repeatedly tossed; for the toss of the fair coin, the expected toss number of the first head is $E[X] = 2$).

Jensen's inequality shows that the expected payout will be less than 2250.

The exact expected payout is $E[3000\,(1 - \frac{1}{2^X})] = 3000(1 - E[\frac{1}{2^X}])$, so we find

$$\begin{aligned} E[\tfrac{1}{2^X}] &= \tfrac{1}{2} \times P(X=1) + \tfrac{1}{2^2} \times P(X=2) + \tfrac{1}{2^3} \times P(X=3) + \cdots \\ &= \tfrac{1}{2} \times \tfrac{1}{2} + \tfrac{1}{2^2} \times \tfrac{1}{2^2} + \tfrac{1}{2^3} \times \tfrac{1}{2^3} + \cdots = \tfrac{1}{4} + \tfrac{1}{4^2} + \tfrac{1}{4^3} + \cdots = \tfrac{1}{3}. \end{aligned}$$

Then $E[3000\,(1 - \frac{1}{2^X})] = 3000(1 - \frac{1}{3}) = 2000$ is the expected payout. □

Example 5-17:

For each of the following random variables, use Chebyshev's inequality to get an upper bound on the probabilities
$P\Big[\,|X-E[X]| > \sqrt{Var[X]}\,\Big]$ and
$P\Big[\,|X-E[X]| > 2\sqrt{Var[X]}\,\Big]$, and also calculate the exact probabilities.

(a) X is the result of a fair die toss, 1,2,3,4,5,6, each with probability $\frac{1}{6}$.

(b) X is continuous with pdf $f_X(x)=1$ for $0<x<1$.

(c) X is continuous with pdf $f_X(x)=e^{-x}$ for $0<x$.

Solution:

According to Chebyshev's inequality, $P\Big[\,|X-E[X]| > r\sqrt{Var[X]}\,\Big] \le \frac{1}{r^2}$.

Therefore, with $r=1$, using Chebyshev's inequality, we have
$P\Big[\,|X-E[X]| > \sqrt{Var[X]}\,\Big] \le 1$ (since any probability must be ≤ 1, this will always be true).
With $r=2$, using Chebyshev's inequality, we have $P\Big[\,|X-E[X]| > 2\sqrt{Var[X]}\,\Big] \le \frac{1}{4}$.

(a) $E[X]=\frac{7}{2}$, $Var[X]=E[X^2]-(E[X])^2=\frac{35}{12}$.

Exact probabilities are

$$P\Big[\,|X-E[X]| > \sqrt{Var[X]}\,\Big] = P\Big[\,|X-\tfrac{7}{2}| > \sqrt{\tfrac{35}{12}}\,\Big] = 1-P\Big[\,|X-\tfrac{7}{2}| \le \,\Big]$$
$$=1-P\Big[-\sqrt{\tfrac{35}{12}} < X-\tfrac{7}{2} < \sqrt{\tfrac{35}{12}}\Big] = 1-P[1.79<X<5.2]$$
$$=1-P[X=2,3,4,5]=\tfrac{2}{6}\le 1, \text{ and}$$

$$P\Big[\,|X-E[X]| > 2\sqrt{Var[X]}\,\Big] = P\Big[\,|X-\tfrac{7}{2}| > 2\sqrt{\tfrac{35}{12}}\,\Big] = 1-P\Big[\,|X-\tfrac{7}{2}| \le 2\sqrt{\tfrac{35}{12}}\,\Big]$$
$$=1-P\Big[-2\sqrt{\tfrac{35}{12}} < X-\tfrac{7}{2} < 2\sqrt{\tfrac{35}{12}}\Big] = 1-P[0.08<X<6.9]$$
$$=1-P[X=1,2,3,4,5,6]=0\le\tfrac{1}{4}.$$

(b) $E[X]=\int_0^1 x\,dx=\frac{1}{2}$ and $E[X^2]=\int_0^1 x^2\,dx=\frac{1}{3}$, so $Var[X]=\frac{1}{3}-(\frac{1}{2})^2=\frac{1}{12}$.

Exact probabilities are

$$P\Big[\,|X-E[X]| > \sqrt{Var[X]}\,\Big] = P\Big[\,|X-\tfrac{1}{2}| > \sqrt{\tfrac{1}{12}}\,\Big] = 1-P\Big[|X-\tfrac{1}{2}| \le \sqrt{\tfrac{1}{12}}\,\Big]$$
$$=1-P\Big[-\sqrt{\tfrac{1}{12}} < X-\tfrac{1}{2} < \sqrt{\tfrac{1}{12}}\Big] = 1-P\Big[\tfrac{1}{2}-\sqrt{\tfrac{1}{12}} < X < \tfrac{1}{2}+\sqrt{\tfrac{1}{12}}\Big]$$
$$=1-2\sqrt{\tfrac{1}{12}}=.423\le 1$$

and

$$P\Big[\,|X-E[X]| > 2\sqrt{Var[X]}\,\Big] = P\Big[\,|X-\tfrac{1}{2}| > 2\sqrt{\tfrac{1}{12}}\,\Big] = 1-P\Big[\,|X-\tfrac{1}{2}| \le 2\sqrt{\tfrac{1}{12}}\,\Big]$$
$$=1-P\Big[-2\sqrt{\tfrac{1}{12}} < X-\tfrac{1}{2} < 2\sqrt{\tfrac{1}{12}}\Big] = 1-P\Big[\tfrac{1}{2}-2\sqrt{\tfrac{1}{12}} < X < \tfrac{1}{2}+2\sqrt{\tfrac{1}{12}}\Big]$$
$$=1-P[-0.077<X<1.077]=0\le\tfrac{1}{4}.$$

(c) $E[X] = \int_0^\infty x \cdot e^{-x}\,dx = 1$ and $E[X^2] = \int_0^\infty x^2 \cdot e^{-x}\,dx = 2$ (we use the general rule $\int_0^\infty x^k \cdot e^{-cx}\,dx = \frac{k!}{c^{k+1}}$ if k is an integer and $c > 0$).

Then $Var[X] = E[X^2] - (E[X])^2 = 1$.

Exact probabilities are

$P\Big[\,|X - E[X]| > \sqrt{Var[X]}\,\Big] = P[\,|X-1| > 1\,] = 1 - P[\,|X-1| \le 1\,]$

$= 1 - P[-1 < X - 1 < 1] = 1 - P[0 < X < 2] = 1 - \int_0^2 e^{-x}\,dx$

$= 1 - (1 - e^{-2}) = e^{-2} = 0.135 \le 1$

and

$P\Big[\,|X - E[X]| > 2\sqrt{Var[X]}\,\Big] = P[\,|X-1| > 2\,] = 1 - P[\,|X-1| \le 2\,]$

$= 1 - P[-2 < X-1 < 2] = 1 - P[-1 < X < 3] = 1 - P[0 < X < 3]$

$= 1 - \int_0^3 e^{-x}\,dx = 1 - (1 - e^{-3}) = e^{-3} = .050 \le \frac{1}{4}$. □

Example 5-18:

Each of the following is a moment generating function for a discrete non-negative integer-valued random variable. For each random variable, find the mean, variance and the probability $P(X \le 2)$.

(a) $\frac{1}{6}(1 + 2e^t + 3e^{2t})$

(b) $e^{2e^t - 2}$

Solution:

(a) $M'(t) = \frac{1}{6}(2e^t + 6e^{2t}) \to M'(0) = \frac{4}{3} = E[X]$

$M''(t) = \frac{1}{6}(2e^t + 12e^{2t}) \to M''(0) = \frac{7}{3} = E[X^2] \to Var[X] = \frac{7}{3} - (\frac{4}{3})^2 = \frac{5}{9}$.

$M(t) = p_0 + p_1 e^t + p_2 e^{2t} + p_3 e^{3t} + \cdots = \frac{1}{6}(1 + 2e^t + 3e^{2t}) = \frac{1}{6} + \frac{1}{3}e^t + \frac{1}{2}e^{2t}$

$\to p_0 = \frac{1}{6}$, $p_1 = \frac{1}{3}$, $p_2 = \frac{1}{2}$, and $P(X \le 2) = 1$.

(b) $ln[M(t)] = 2e^t - 2 \to \frac{d}{dt}\,ln[M(t)] = 2e^t \to \frac{d}{dt}\,ln[M(t)]\Big|_{t=0} = 2 = E[X]$.

$\frac{d^2}{dt^2}\,ln[M(t)] = 2e^t \to \frac{d}{dt^2}\,ln[M(t)]\Big|_{t=0} = 2 = Var[X]$

$M(t) = e^{-2}e^{2e^t}$. We use the Taylor expansion $e^y = 1 + y + \frac{y^2}{2!} + \frac{y^3}{3!} + \cdots$, so that

$e^{2e^t} = 1 + 2e^t + \frac{(2e^t)^2}{2!} + \frac{(2e^t)^3}{3!} + \cdots$, and then

$M(t) = e^{-2} \times [1 + 2e^t + \frac{(2e^t)^2}{2!} + \frac{(2e^t)^3}{3!} + \cdots] = e^{-2} + 2e^{-2}e^t + 2e^{-2}e^{2t} + \frac{4}{3}e^{-2}e^{3t} + \cdots$.

Therefore, $p_0 = e^{-2}$, $p_1 = 2e^{-2}$, $p_2 = 2e^{-2}$, $p_3 = \frac{4}{3}e^{-2}$, ...

and $P(X \le 2) = e^{-2} + 2e^{-2} + 2e^{-2} = 5e^{-2}$. □

Example 5-19:

The value of a piece of factory equipment after three years of use us $100(0.5)^X$ where X is a random variable having moment generating function $M_X(t) = \frac{1}{1-2t}$ for $t < \frac{1}{2}$.

Calculate the expected value of this piece of equipment after three years of use.

Solution:

We denote by V the value of the piece of equipment. We wish to find

$E[V] = E[100(0.5)^X] = 100 \times E[(0.5)^X]$, which can be written as $100 \times E[e^{X \times ln\,.5}]$.

But the MGF of X is $M_X(t) = E[e^{tX}]$, so that

$E[e^{X \times ln\,0.5}] = M_X(ln\,0.5) = \frac{1}{1-2 \times ln\,0.5} = 0.419$. Then, $E[V] = 41.9$. □