Moments:

Def: The *rth* moment about the origin of a random variable *X*, denoted by μ'_r, is the expected value of X^r.

$$\mu'_r = E(X^r) = \sum_x x^r f(x) \qquad \text{for } r = 0, 1, 2, \ldots$$

when *X* is discrete, and when *X* is continuous

$$\mu'_r = E(X^r) = \int_{-\infty}^{\infty} X^r f(x)dx$$

When $r = 1$, $\mu'_1 = E(X)$ is called the mean of the distribution of *X*, and it is denoted by μ.

Def: The *rth* moment about the mean of a random variable *X*, denoted by μ_r, is the expected value of $(X - \mu)^r$.

$$\mu_r = E\left[(X-\mu)^r\right] = \sum_x (x-\mu)^r . f_X(x) \qquad \text{for } r = 0, 1, 2, \ldots$$

When *X* is discrete, and

$$\mu_r = E\left[(X-\mu)^r\right] = \int_{-\infty}^{\infty} (x-\mu)^r f_X(x)dx$$

when *X* is continuous.

Note that: $\mu_0 = 1$ and $\mu_1 = 0$.

$\mu_2 = E[(X - \mu)^2]$ is called the variance of the distribution of *X*, and it is denoted by σ^2, $Var(X)$ or $V(X)$. The positive square root of the variance $+\sqrt{Var(X)}$ is called the standard deviation, and it is denoted by σ.

Theorem:

$\sigma^2 = E(X^2) - \mu^2$

Proof:

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

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

$= E(X^2) - 2\mu^2 + \mu^2 = E(X^2) - \mu^2.$

It also follows immediately that

$E(X^2) = \sigma^2 + \mu^2$

Theorem:

If X has the variance σ^2, then: $Var(aX + b) = a^2\sigma^2$

Proof:

$Var(aX + b) = E[(aX + b) - \mu_{aX+b}]^2 = E[(aX + b) - E(aX+b)]^2$

$= E[aX + b - a\,E(X) - b]^2 = E[aX - aE(X)]^2 = a^2E[X - E(X)]^2$

$= a^2E[X - \mu]^2 = a^2\sigma^2.$

Example:

The daily demand for a certain tool has the following probability distribution, where X is the daily demand.

x	0	1	2
$f(x)$	0.1	0.5	0.4

1. Find $E(X)$.

$$E(X) = \sum_{x=0}^{2} xf(x) = 0(0.1) + 1(0.5) + 2(0.4) = 1.3$$

That is, the tool is used an average of 1.3 times per day.

2. Find $var(X)$.

$$E(X^2) = \sum_{x \in S} x^2 f(x) = 0(0.1) + 1(0.5) + 4(0.4) = 2.1$$

$var(X) = E(X^2) - [E(X)]^2 = 2.1 - 1.69 = 0.41$

3. If the daily cost of using the tool is $10X$. Find the average cost of using the tool and the variance of the costs.

$E(10X) = 10\,E(X) = 10(1.3) = 13$ *the average day cost* *var (10X) = 100 var(X)= 100 (0.41) = 41*

Example:

The distribution of the amount of gravel (in tons) sold by a particular construction company in a given week has the following *pdf*:

$$f(x)=\begin{cases}\frac{3}{2}\left(1-x^2\right) & 0\le x\le 1\\ 0 & elsewhere\end{cases}$$

Find: *1. F(X)* *2. E(X)* *3. $E(X^2)$* *4. var(X)*

1. $$F(X)=\int_0^x \frac{3}{2}\left(1-t^2\right)dt=\frac{3}{2}\left[t-\frac{t^3}{3}\right]_0^x=\frac{3}{2}\left[x-\frac{x^3}{3}\right]$$

2. $$\mu=E(X)=\int_0^1 x\frac{3}{2}\left(1-x^2\right)dx=\frac{3}{2}\int_0^1\left(x-x^3\right)dx=\frac{3}{2}\left[\frac{x^2}{2}-\frac{x^4}{4}\right]_0^1$$

$$=\frac{3}{2}\left[\frac{1}{2}-\frac{1}{4}\right]=\frac{3}{8}$$

3. $$E(X^2)=\int_0^1 x^2\frac{3}{2}\left(1-x^2\right)dx=\frac{3}{2}\int_0^1\left(x^2-x4\right)dx=\frac{3}{2}\left[\frac{x^3}{3}-\frac{x^5}{5}\right]_0^1$$

$$=\frac{3}{2}\left[\frac{1}{3}-\frac{1}{5}\right]=\frac{1}{5}$$

4. *var(X) = $E(X^2)$ – $[E(X)]^2$*

$$=\frac{1}{5}-\frac{9}{64}=\frac{64-45}{320}=\frac{19}{320}=0.059$$

Median

The median of a random variable *X* is defined as a number *m* such that $P(X\le m)\ge ½$ and $P(X\ge m)\ge ½$

So the median of *X* is any number that has half the mass of *X* to its right and the other half to its left.

If *X* is a continuous random variable, then the median of *X* satisfies

$$\int_{-\infty}^{med(X)} f_X(x)\,dx = \frac{1}{2} = \int_{med(X)}^{\infty} f_X(x)dx$$

Example:

Suppose that *X* has the following discrete distribution

$$f_X(x) = \begin{cases} 0.1 & x = 1 \\ 0.2 & x = 2 \\ 0.3 & x = 3 \\ 0.4 & x = 4 \end{cases}$$

Then,

P(X ≤ 3) =0.6 which is ≥ ½

and *P(X ≥ 3) = 1- P (X≤ 2)=1-0.3= 0.7 which is also ≥ ½*

Therefore, the median = 3

Example:

Suppose that *X* has the following pdf:

$$f_X(x) = \begin{cases} 3x^2 & 0 < x < 1 \\ 0 & elsewhere \end{cases}$$

Then

$$\int_0^m 3x^2\,dx = \int_m^1 3x^2 dx = \frac{1}{2}$$

Hence, $m = (1/2)^{1/3}$

Chebyshev's Theorem:

Chebysheve's inequality is helpful for understanding the significance of the standard deviation.

Theorem

If the random variable X has a finite mean μ and finite variance σ^2, then for every $k \geq 1$,

$$P(|X-\mu| \geq k\sigma) \leq \frac{1}{k^2}$$

Proof:

$$\sigma^2 = E[(X-\mu)^2] = \int_{-\infty}^{\infty} (X-\mu)^2 f_X(x)dx$$

Dividing the integral into three parts;

$$\sigma^2 = \int_{-\infty}^{\mu-k\sigma} (X-\mu)^2 f_X(x)dx + \int_{\mu-k\sigma}^{\mu+k\sigma} (X-\mu)^2 f_X(x)dx + \int_{\mu+k\sigma}^{\infty} (X-\mu)^2 f_X(x)dx$$

Since the integrand $(X-\mu)^2 f_X(x)$ is nonnegative. We can form the inequality

$$\sigma^2 \geq \int_{-\infty}^{\mu-k\sigma} (X-\mu)^2 f_X(x)dx\sigma^2 + \int_{\mu+k\sigma}^{\infty} (X-\mu)^2 f_X(x)dx$$

by deleting the second integral.

Now since $(X-\mu)^2 \geq k^2\sigma^2$ for $x \leq \mu - k\sigma$ or $x \geq \mu + k\sigma$, it follows that

$$\sigma^2 \geq \int_{-\infty}^{\mu-k\sigma} k^2\sigma^2 f_X(x)dx + \int_{\mu+k\sigma}^{\infty} k^2\sigma^2 f_X(x)dx$$

Hence,

$$\frac{1}{k^2} \geq \int_{-\infty}^{\mu-k\sigma} f_X(x)dx + \int_{\mu+k\sigma}^{\infty} f_X(x)dx$$

That is:

$$P(|X-\mu| \geq k\sigma) \leq \frac{1}{k^2}$$

It follows that:

$P(|X-\mu|<k\sigma) \geq 1-\frac{1}{k^2}$

This is equivalent to

$P(\mu-k\sigma \leq X \leq \mu+k\sigma) \geq 1-\frac{1}{k^2}$

For instance, let $k=2$, then the interval μ-2σ to $\mu+2\sigma$ must contain at least $1-\frac{1}{k^2}=1-\frac{1}{4}=0.75$ of the probability mass, and the probability is at least $1-\frac{1}{3^2}=\frac{8}{9}$ that a random variable X will take on a value within three standard deviations of the mean.

Chebyshev's inequality is used in various ways, as will be shown in the following examples:

Example:

If it is known that X has a mean 25 and a variance $\sigma^2=16$, then

1) $P(17<X<33)=P(|X-25|<8)=P(|X-\mu|<k\sigma)$

$k\sigma=8 \quad \therefore\ k=\frac{8}{\sigma}=\frac{8}{\sqrt{16}}=2$

$\therefore P(|X-\mu|<2\sigma) \geq 1-\frac{1}{2^2}$

Or

$P(17-25<X-\mu<33-25)=P(-8<X-\mu<8)=P(|X-\mu|<8) \geq 0.75$

2) $P(|X-25| \geq 12)=P(|X-\mu| \geq 3\sigma) \leq \frac{1}{3^2}=\frac{1}{9}$

Example:

The daily production of electric motors at a certain factory averaged 120 with standard deviation of 10.

a. What fraction of days will have a production level between 100 and 140?

b. Find the shortest interval certain to contain at least 90% of the daily production levels.

Solution

a) $\mu-k\sigma=120-10k=100 \ \rightarrow\ 10k=20 \quad \therefore k=2$

$\mu+k\sigma = 120+10k = 140 \rightarrow 10k = 20 \quad \therefore k = 2$

$$P(|X-\mu| < k\sigma) \geq 1-\frac{1}{k^2} = 1-\frac{1}{4} = 0.75$$

$\therefore$ at least *75*% of the days will have a production level in this interval.

b) To find k set $1-\frac{1}{k^2} = 0.90 \Rightarrow \frac{1}{k^2} = 0.10 \Rightarrow k^2 = 10 \therefore k = \sqrt{10} = 3.16$

The interval $\mu-3.16\sigma$ to $\mu+3.16\sigma$ is:

$120-3.16(10)$ to $120+3.16(10)$

88.4 to 151.6

will contain at least 90% of the daily production.

Example:

The daily cost for use of a certain tool had a mean \$13 and a variance of \$41. How often will this cost exceed \$30?

$$P(|X-\mu| \geq k\sigma) \leq \frac{1}{k^2}$$

$$k = \frac{30-\mu}{\sigma} = \frac{30-13}{\sqrt{41}} = 2.66$$

$\mu-2.66\sigma$ to $\mu+2.66\sigma$

$13-2.66\sqrt{41}$ to $13+2.66\sqrt{41}$

-4 to 30

must contain at least $1-\frac{1}{k^2} = 1-\frac{1}{(2.66)^2} = 1-0.14 = 0.84$

that is:

$$P(|X-\mu| < 2.66\sigma) \geq 1-\frac{1}{k^2} = 0.86$$

or

$$P(|X-\mu| \geq 2.66\sigma) \leq \frac{1}{k^2} = 0.14$$

Example:
The weekly amount Y spent for chemicals have a mean of \$445 and a variance of \$236. Within what interval would these weekly costs for chemicals be expected to lie at least 75% of the time?

The shortest interval containing at least 75% of the probability mass of Y gives $1-\frac{1}{k^2}=0.75$

$\frac{1}{k^2}=\frac{1}{4} \Rightarrow k^2=4 \therefore k=2$

Thus,

μ-2σ to $\mu+2\sigma$ will contain at least 75% of the probability

$445-2\sqrt{236}$ to $445+2\sqrt{236}$

445- 30.72 to 445+30.72

414.28 to 475.72

H.W.:

1. Find the mean and variance for the following discrete distribution:

$$f(x)=\frac{3!}{x!(3-x)!}(\frac{1}{4})^x(\frac{3}{4})^{3-x} \qquad x=0,1,2,3$$

2. Given $E(X+4)=10$, and $E[(X+4)^2]=116$, determine μ , σ^2 and $Var(X+4)$

3. If the random variable X has mean μ and standard deviation σ, let the random variable $Z=\frac{X-\mu}{\sigma}$. Find $E(Z)$ and $Var(Z)$.

4. Find the mean and variance for the random variable X whose *pdf* is:

$$f(x)=\begin{cases} 4x(1-x^2) & 0\le x\le 1 \\ 0 & otherwise \end{cases}$$

5. Let the random variable X has the following *pdf*:

$$f(x) = \begin{cases} x-8 & 8 \le x \le 9 \\ 10-x & 9 < x \le 10 \end{cases}$$

a. Find the mean and variance of *X*.
b. Let *Y = X/15 + 0.35,* find the mean and variance of *Y.*

6. Let X be a random variable with pdf

$$f_X(x) = \begin{cases} \frac{1}{8} & x = 0 \\ \frac{5}{8} & x = 1 \\ \frac{2}{8} & x = 2 \\ 0 & o.w \end{cases}$$

Find the median, $E(X)$, $E(X^2)$, $E(3X^2+5/X+1)$, $Var(X)$, and $Var(3X)$

7. The amount of time it takes a person to be served at a given gas station is a random variable with pdf

$$f_X(x) = \begin{cases} \frac{1}{4} e^{-\frac{x}{4}} & x > 0 \\ 0 & elsewhere \end{cases}$$

Find the median, the mean and the variance of this random variable.

Moment-Generating Functions:

<u>*Def*</u>. The moment-generating function of a random variable *X*, where it exists, is given by

$$M_X(t) = E(e^{tX}) = \sum_x e^{tx} . f(x)$$

when *X* is discrete, and

$$M_X(t) = E(e^{tX}) = \int_{-\infty}^{\infty} e^{tx} . f(x) dx$$

when *X* is continuous.

Where $-h < t < h$ for a positive number *h*, and we are usually interested in values of *t* in the neighborhood of *0*.

Let as substitute for e^{tx} its Maclurin's series expansion, namely,

$$e^{tx} = 1 + tx + \frac{t^2 x^2}{2!} + \frac{t^3 x^3}{3!} + \dots + \frac{t^r x^r}{r!} + \dots$$

For the discrete case, we thus get

$$M_X(t) = E\left(e^{tX}\right) = \sum_x e^{tx}.f(x) = \sum_x \left[1 + tx + \frac{t^2 x^2}{2!} + \dots + \frac{t^r x^r}{r!} + \dots\right] f(x)$$

$$= \sum_x f(x) + t.\sum_x xf(x) + \frac{t^2}{2!}\sum_x x^2 f(x) + \dots + \frac{t^r}{r!}\sum_x x^r f(x) + \dots$$

$$= 1 + tE(X) + \frac{t^2}{2!}E(X^2) + \frac{t^3}{3!}E(X^3) + \dots$$

$$= 1 + \mu t + \mu_2' \, .\frac{t^2}{2!} + \dots + \mu_r' . \frac{t^r}{r!} + \dots$$

It can be seen that in the Maclurin's series of the moment-generating function of *X*, the coefficient of $\frac{t^r}{r!}$ is μ_r', the *rth* moment about the origin. In the continuous case the argument is the same.

This gives a method of finding the moments.

Theorem:

If the mgf of the random variable *X* exist, then

$$\frac{d\mu_X(t)}{dt^r} \Big|_{t=0} = \mu_r{}' = E(X^r)$$

Thus

$$\mu_X'(t) = \sum x e^{tx} f(x)$$

$$\mu_X''(t) = \sum x^2 e^{tx} f(x)$$

and for each positive integer *r*,

$$\mu_X^{(r)}(t) = \sum x^r e^{tx} f(x)$$

Setting $t = 0$, we see that

$$\mu'_X(0) = \sum xf(x) = E(X) = \mu$$

$$\mu''_X(0) = \sum x^2 f(x) = E(X^2)$$

and, in general,

$$\mu^{(r)}{}_X(0) = \sum x^r f(x) = E(X^r)$$

In particular, if the moment-generating function exist

$\mu = M'(0)$ and $\sigma^2 = M''(0) - [M'(0)]^2$.

Example 1:

Find the moment generating function of the random variable whose *pdf* is given by:

$$f(x) = \begin{cases} e^{-x} & for \quad x > 0 \\ 0 & elsewhere \end{cases}$$

$$M_x(t) = E(e^{tx}) = \int_{-\infty}^{\infty} e^{tx} e^{-x} \quad dx$$

$$= \int_0^{\infty} e^{-x(1-t)}\, dx \quad = \frac{-1}{1-t} \int_0^{\infty} -(1-t)e^{-x(1-t)}\, dx$$

$$= -\frac{1}{1-t}[e^{-x(1-t)}]_0^{\infty} = -\frac{1}{1-t}[0-1] = \frac{1}{1-t} \qquad for \quad t < 1$$

Example 2:

Given that X has the probability distribution:

$$f(x) = \frac{C_x^3}{8} \qquad for \quad x = 0,1,2,3$$

Find the moment generating function and use it to determine μ_1' and μ_2'.

$$M_X(t) = E(e^{tx}) = \frac{1}{8}\sum_{x=0}^{3} e^{tx} C_x^3$$

$$= \frac{1}{8}(1 + 3e^t + 3e^{2t} + e^{3t}) = \frac{1}{8}(1 + e^t)^3$$

Then by theorem

$$\mu_1' = M'_X(0) = \frac{3}{8}(1+e^t)^2 e^t \Big|_{t=0} = \frac{3}{2}$$

$$\mu_2' = M''_X(0) = \frac{3}{4}(1+e^t)e^{2t} + \frac{3}{8}(1+e^t)^2 e^t \Big|_{t=0} = 3$$

Properties of mgfs:

Theorem

If $Y = aX + b$ then, $M_Y(t) = e^{bt} M_X(at)$

Proof

$$M_Y(t) = E(e^{tY}) = E\left(e^{t(aX+b)}\right) = E(e^{atX} e^{bt})$$

$$= e^{bt} E(e^{atX}) = e^{bt} M_X(at)$$

Note:

The r^{th} moment about the mean $E[(X - \mu)^r]$ can be computed using this property. Since

$M_{X-\mu}(t) = e^{-\mu t} M_X(t)$

$E\,[(X-\mu)^r] = \frac{d^r}{dt^r}\left[e^{-\mu t} M_X(t)\right]\big|_{t=0}$

Theorem: Uniqueness

Let X and Y be two random variables with respective *CDF's* $F_1(x)$ and $F_2(x)$, and *mgf's* $M_1(t)$ and $M_2(t)$, then $F_1(x) = F_2(x)$ for all x if and only if $M_1(t) = M_2(t)$ for all t in some open interval containing $0, -h < t < h$ for some $h > 0$.

H.W.

Let X be a random variable with pdf given by

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & for \quad x > 0 \\ 0 & elsewhere \end{cases}$$

Find $E(X)$ and $Var(X)$ using *mgf*.

<u>Characteristic Function</u>:

<u>Def</u>. The characteristic function of a random variable X, is defined as

$$\phi_X(t) = E\left(e^{itX}\right) \qquad -\infty < t < \infty$$

where $i = \sqrt{-1}$ is the pure imaginary quantity.

That is:

$$\phi_X(t) = E\left(e^{itX}\right) = \sum_{x \in S} e^{itx} . f_X(x) \qquad \text{when } X \text{ is discrete, and}$$

$$\phi_X(t) = E\left(e^{itX}\right) = \int_{-\infty}^{\infty} e^{itx} . f_X(x) dx \qquad \text{when } X \text{ is continuous.}$$

This expectation exists for every distribution. To see why it exists for all real t, let us define

$$e^{it} = \sum_{n=o}^{\infty} \frac{(it)^n}{n!} = 1 + it - \frac{t^2}{2} - \frac{it^3}{3!} + \frac{t^4}{4!} + \frac{it^5}{5!} - \dots$$

$$= \left(1 - \frac{t^2}{2!} + \frac{t^4}{4!} - \ \dots\right) + i\left(t - \frac{t^3}{3!} + \frac{t^5}{5!} - \ \dots\right)$$

$e^{it} = \cos t + i \sin t$ (1)

Using the fact that $\cos(-t) = \cos(t)$ *and* $\sin(-t) = -\sin t$, we see that

$e^{-it} = \cos t - i \sin t$ (2)

Solving (1) and (2) for $\cos t$ and $\sin t$,we get

$\cos t = \frac{e^{it}+e^{-it}}{2}$ and $\sin t = \frac{e^{it}-e^{-it}}{2i}$

It also follows from (1) that

$$\left|e^{it}\right| = (cos^2 t + sin^2 t)^{1/2} = 1$$

We note in the continuous case that

$$|\emptyset_X(t)| = \left|\int_{-\infty}^{\infty} e^{itx} f_X(x)\, dx\right| \le \int_{-\infty}^{\infty} \left|e^{itx} f_X(x)\, dx\right|$$

Or equivalently,

$$|\emptyset_X(t)| = \left|E\ e^{itx}\right| \le E\left|e^{itx}\right| = E\ 1 = 1$$

Note that $\phi_X(0) = E\ e^0 = 1$

Remark:

$\phi_X(t) = E\ e^{itx} = M_X(it)$ and $M_X(t) = Ee^{tx} = Ee^{\frac{itx}{i}} = \emptyset_X\left(\frac{t}{i}\right)$

Lemma:

$$\emptyset_X^{(r)}(t) = \frac{d^{(r)}}{dt^r} \emptyset_X(t)|_{t=0} = i^r E X^r$$

Proof:

$$\phi_X(t) = E\left(e^{itx}\right) = \sum_{n=0}^{\infty} \frac{(itx)^n}{n!} = E\left(1 + itx + \frac{(itx)^2}{2!} + \frac{(itx)^3}{3!} + \dots\right)$$

$$= 1 + itE(X) + \frac{i^2t^2}{2!}E(X^2) + \frac{i^3t^3}{3!}E(X^3) + \ldots$$

Now

$$\emptyset'_X(t) = \frac{d}{dt}\emptyset_X(t) = i\,E(X) + \frac{2ti^2}{2!}E(X^2) + \frac{3t^2i^3}{3!}\,E(X^3) + \ \ldots$$

$$\emptyset''_X(t) = \frac{d^2}{dt^2}\emptyset_X(t) = i^2\,E(X^2) + \frac{6t \quad i^3}{3!}\,E(X^3) + \ \ldots$$

.

Hence,

$$\emptyset'_X(t) = \frac{d}{dt}\emptyset_X(t)|_{t=0} = iE(X)$$

$$\emptyset''_X(t) = \frac{d^2}{dt^2}\emptyset_X(t)|_{t=0} = i^2E(X^2)$$

.

$$\emptyset^{(r)}_X(t) = \frac{d^{(r)}}{dt^r}\emptyset_X(t)|_{t=0} = i^rE(X^r)$$

Theorem

If $Y = aX + b$ then, $\emptyset_Y(t) = e^{itb}\emptyset_X(at)$

Proof

$$\emptyset_Y(t) = E\left(e^{itY}\right) = E\left(e^{it(aX+b)}\right) = E\left(e^{itaX}e^{itb}\right)$$

$$= e^{itb}E\left(e^{itaX}\right) = e^{itb}\emptyset_X(at)$$

Def:

The characteristic function about the mean is:

$$\emptyset_{X-\mu}(t) = e^{-it\mu}\emptyset_X(t)$$

Uniqueness Theorem:

Let X and Y be two random variables with respective *pdf's* $f_X(x)$ and $f_Y(y)$. Suppose that $\phi_X(t) = \phi_Y(t)$, then for all $t \in \boldsymbol{R}$, $f_X(x) = f_Y(y)$whenever $x = y$.

Inversion formula

One of the most useful properties of $\phi_X(t)$ is that it can be used to calculate the *pdf* of the random variable X. Specifically , we have the following inversion formula

$$f_X(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-ixt} \emptyset_X(t) dt$$

Expectation of a Sum of Random variables:

Theorem:

If X and Y are random variables, then

$$E(X+Y) = E(X) + E(Y)$$

Proof:

$$E(X+Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x+y) f(x,y) dxdy$$

$$= \int_{-\infty}^{\infty}\int_{=\infty}^{\infty} xf(x,y)dxdy \quad + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} yf(x,y)dxdy$$

$$= \int_{-\infty}^{\infty} x(\int_{-\infty}^{\infty} f(x,y)dy)dx + \int_{-\infty}^{\infty} y(\int_{=\infty}^{\infty} f(x,y)dx)dy$$

$$= \int_{-\infty}^{\infty} xf(x)dx \quad + \int_{-\infty}^{\infty} yf(y)dy \qquad = E(X) + E(Y)$$

A similar result can be shown in the discrete case.

In general, for any n,

$$E(X_1 + X_2 + ... + X_n) = E(X_1) + E(X_2) + ... + E(X_n)$$

Theorem*:*

If X and Y are two independent random variables, then

E(XY) = E(X)E(Y)

Proof:

$$E(XY) = \sum_x \sum_y xyf(x,y) = \sum_x \sum_y xyf(x)\, f(y)$$

$$= \sum_x xf(x) \quad \sum_y yf(y)$$

$$= E(X)E(Y)$$

Def: The conditional expectation of Y given $X=x$ is given by

$E[Y|x] = \int_{-\infty}^{\infty} y f_2(y|x)dy$ for continuous random variables

$E[Y|x] = \sum_y y f_2(y|x)$ for discrete random variables

Theorem:

Let X and Y be random variables such that Y has finite mean. Then

$E[E(Y|X)] = E(Y)$

Proof:

$$E[E(Y|X)] = \int_{-\infty}^{\infty} E[Y|x] f_1(x)dx = \iint_{-\infty\,-\infty}^{\infty\,\infty} y g_2(y|x)\, f_1(x)dydx$$

Since $g_2(y|x) = f(x,y)/f_1(x)$, it follows that

$$E[E(Y|X)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} y\, f(x,y)dydx = \int_{-\infty}^{\infty} y \int_{-\infty}^{\infty} f(x,y)dx\, dy$$

$$= \int_{-\infty}^{\infty} y\, g_2(y)dy = E(Y)$$

Covariance and Correlation Coefficient*:*

Def: If X and Y are two random variables with respective means $\bar{x}$and$\bar{y}$, then the covariance between X and Y is

$$\text{cov}(X,Y) = \sigma_{X,Y} = E(X-\overline{X})(Y-\overline{Y})$$

Note: If X and Y are independent random variables, then

$$\text{cov}(X,Y) = E(X-\overline{X})E(Y-\overline{Y}) = 0$$

Since

$$E(X-\overline{X}) = E(X) - E(\overline{X}) = \overline{X} - \overline{X} = 0$$

A useful expression for *cov(X,Y)* can be obtained by expanding the right side of the definition. These yields:

$$\begin{aligned}\operatorname{cov}(X,Y) &= E(X-\bar{X})(Y-\bar{Y}) = E(X-\mu_x)(Y-\mu_y)\\ &= E(XY-\mu_x Y-\mu_y X+\mu_x\mu_y)\\ &= E(XY)-\mu_x E(Y)-\mu_y E(X)+\mu_x\mu_y\\ &= E(XY)-2\mu_x\ \mu_y+\mu_x\mu_y = E(XY)-E(X)E(Y)\end{aligned}$$

From its definition, we see that:

1.cov(X, Y) = cov(Y, X)

2.cov(X, X) = var(X)

3. cov(aX, bY)=ab cov(X,Y)

4. var(aX, bY)= $a^2 var(X)+b^2 var(Y)+2ab\ cov(X,Y)$

5. var(aX, bY)= $a^2 var(X)+b^2 var(Y)$, if X, Y are indep. then

Def : The ratio

$$\rho_{xy} = \frac{\operatorname{cov}(X,Y)}{\sqrt{\operatorname{var}(X)\operatorname{var}(Y)}} = \frac{E(X-\bar{X})(Y-\bar{Y})}{\sqrt{E(X-\bar{X})^2}\ \sqrt{E(Y-\bar{Y})^2}}$$

is defined as the correlation coefficient between *X* and *Y*.

Note: If X and Y are indep. random variables, then $\rho_{X,Y} = 0$

Problem: Show that: $-1 \le \rho_{X,Y} \le 1$

Using Schwartz inequality: $[E(X,Y)]^2 \le E(X^2)E(Y^2)$

$$[E(X-\mu_X)(Y-\mu_Y)]^2 \le E(X-\mu_X)^2 E(X-\mu_Y)^2$$

$$[cov(x,y)]^2 \le \sigma_X^2\sigma_Y^2$$

$$\frac{[cov(x,y)]^2}{\sigma_X^2\sigma_Y^2} \le 1$$

$$\left[\frac{cov(x,y)}{\sigma_X^2\sigma_Y^2}\right]^2 \le 1$$

$$(\rho_{X,Y})^2 \le 1 \rightarrow -1 \le \rho_{X,Y} \le 1$$

Def: The *ij*-th moment about the origin of the random variables *X*and *Y*, denoted by

$m'_{ij} = E\left(X^i Y^j\right)$ $\quad i = 0,1, \ldots, j = 0,1, \ldots$

$m'_{11} = E\,(XY), m'_{10} = E\,(X), m'_{01} = E\,(Y)$

$m'_{20} = E\,(X^2)$

$m'_{02} = E\,(Y^2)$

Def: The *ij*-th moment about the mean of the random variables *X* and *Y*, denoted by

$m_{ij} = E\left[(X-\mu_X)^i (Y-\mu_Y)^j\right]$

Note that:

$m_{11} = cov(X,Y), m_{10} = E\,(X-\mu_X) = 0, m_{01} = E\,(Y-\mu_Y) = 0$

$m_{20} = E\,(X-\mu_X)^2 = var(X)$

$m_{02} = E\,(Y-\mu_Y)^2 = var(Y)$

H.W.

Let *(X,Y)* be a continuous random vector with pdf:

$$f_{X,Y}(x,y) = \begin{cases} e^{-x-y} & x > 0, y > 0 \\ 0 & o.w \end{cases}$$

1. Are *X* and *Y* indep.? 2. Find $\rho_{X,Y}$.

Discrete Distributions

Discrete Uniform Distribution

If the random variable *X* is equally likely to be each of the *n* integers *1, 2, …, n*. Then the pdf of *X* is:

$$f_X(x) = \begin{cases} \frac{1}{n} & x = 1,2,..,n \\ 0 & elsewhere \end{cases}$$

Example:
Throwing a die once yields the following pdf:

$$f_X(x) = \begin{cases} \frac{1}{6} & x = 1,2,...,6 \\ 0 & elsewhere \end{cases}$$

Then *X* has a discrete uniform distribution.

Example
One ball is selected at random from 5 balls numbered from 1 to 5. Find the pdf of the number of the selected ball.
Let *X* be the random variable of the selected ball, then

$$f_X(x) = \begin{cases} \frac{1}{5} & x = 1,2,3,4,5 \\ 0 & elsewhere \end{cases}$$

Lemma:
If *X* is has discrete uniform distribution, then

$$E(X) = \frac{n+1}{2}$$

and

$$Var(X) = \frac{n^2-1}{12}$$

Proof:

$$E(X) = \sum_{x\in S} x\, f_X(x) = \sum_{i=1}^{n} x\frac{1}{n} = \frac{1}{n}\sum_{i=1}^{n} x$$

$$= \frac{1}{n}\left(\frac{n(n+1)}{2}\right) = \frac{n+1}{2}$$

$$E(X^2) = \sum_{x\in S} x^2 f_X(x) = \sum_{i=1}^{n} x^2 \frac{1}{n} = \frac{1}{n}\sum_{i=1}^{n} x^2$$

$$= \frac{1}{n}\left[\frac{n(n+1)(2n+1)}{6}\right] = \frac{(n+1)(2n+1)}{6}$$

$$Var(X) = \frac{(n+1)(2n+1)}{6} - \frac{(n+1)^2}{4}$$

$$= \frac{n+1}{2}\left[\frac{4n-2-3n-3}{6}\right] = \frac{n+1}{2}\left[\frac{n-1}{6}\right] = \frac{n^2-1}{12}$$

Theorem

If $X_1, X_2, \ldots, X_n$ are independent random variables with *mgf's* $M_{x_i}(t)$, then the *mgf* of $Y = \sum_{i=1}^{n} x_i$ is

$M_Y(t) = M_{x_1}(t).M_{x_2}(t) \ldots M_{x_n}(t)$

Proof: H.W.

Example:

Expectation of a Sum of Random variables:

Theorm:

If X and Y are random variables, then

$E(X+Y) = E(X) + E(Y)$

Proof:

$$E(X+Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x+y)f(x,y)dxdy$$

$$= \int_{-\infty}^{\infty}\int_{=\infty}^{\infty} xf(x,y)dxdy \quad + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} yf(x,y)dxdy$$

$$= \int_{-\infty}^{\infty} x(\int_{-\infty}^{\infty} f(x,y)dy)dx + \int_{-\infty}^{\infty} y(\int_{=\infty}^{\infty} f(x,y)dx)dy$$

$$= \int_{-\infty}^{\infty} xf(x)dx \quad + \int_{-\infty}^{\infty} yf(y)dy \qquad = E(X)+E(Y)$$

A similar result can be shown in the discrete case.

In general, for any n,

$$E(X_1 + X_2 + ... + X_n) = E(X_1) + E(X_2) + ... + E(X_n)$$

Expectation of a Product of Independent Random Variables:

Theorem:

If X and Y are two independent random variables, then
$E_{XY}(XY) = E(X)E(Y)$
Proof:

$$E(XY) = \sum_x \sum_y xyf(x,y) = \sum_x \sum_y xyf(x)f(y)$$

$$= \sum_x xf(x) \quad \sum_y yf(y)$$

$$= E(X)E(Y)$$

Covariance and Correlation Coefficient:

Def: If X and Y are two random variables with respective means $\overline{X}$ and $\overline{Y}$, then the covariance between X and Y is

$$\mathbf{cov}(X,Y) = E(X-\overline{X})(Y-\overline{Y})$$

Note: If X and Y are independent random variables, then

$$\operatorname{cov}(X,Y) = E(X-\overline{X})E(Y-\overline{Y}) = 0$$

Since

$$E(X-\overline{X}) = E(X) - E(\overline{X}) = \overline{X} - \overline{X} = 0$$

A useful expression for $cov(X,Y)$ can be obtained by expanding the right side of the definition. This yields:

$$\operatorname{cov}(X,Y) = E(X-\overline{X})(Y-\overline{Y}) = E(X-\mu_x)(Y-\mu_y)$$

$$= E(XY - \mu_x Y - \mu_y X + \mu_x \mu_y)$$

$$= E(XY) - \mu_x E(Y) - \mu_y E(X) + \mu_x \mu_y$$

$$= E(XY) - 2\mu_x \mu_y + \mu_x \mu_y = E(XY) - E(X)E(Y)$$

From its definition, we see that:
$cov(X, Y) = cov(Y, X)$
and
$cov(X, X) = var(X)$

Def : The ratio

$$\rho_{xy} = \frac{\operatorname{cov}(X,Y)}{\sqrt{\operatorname{var}(X)\operatorname{var}(Y)}} = \frac{E(X-\overline{X})(Y-\overline{Y})}{\sqrt{E(X-\overline{X})^2}\ \sqrt{E(Y-\overline{Y})^2}}$$

is defined as the correlation coefficient between *X* and *Y*.