where $(i_1, i_2, \dots i_n)$ is a permulation of numbers $\{1, 2, \dots n\}$.

Exchangeable random variables

The components of the random vector $\underline{X} = (X_1, \dots X_n)$ is called a exchangeable if for every permulation $(i_1, \dots, i_n)$ of $(1, 2, \dots n)$, we have

$$(X_1, X_2, X_3, \dots, X_n) \stackrel{d}{=} (X_{i_1}, X_{i_2}, \dots, X_{i_n})$$

Remark A finite collection of i.i.d. random variables is always exchangeable. But there are exchangeable random variables which are i.i.d.

Consider a random vector $(X, Y, Z) : (\Omega, \mathcal{F}, P) \to (\mathbb{R}^3, \mathcal{B}(\mathbb{R}^3))$ with probability density function

$$f(x, y, z) = \begin{cases} \frac{2}{3}(x + y + z), & (x, y, z) \in [0, 1]^3 \\ 0 & \text{otherwise.} \end{cases}$$

Exercise Show that (X, Y, Z) are not independent.

Moments and moment generating functions.

Let $(X_1, X_2, \dots X_n) : (\Omega, \mathcal{F}, \mathbb{P}) \longrightarrow (\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ be a ~~randoo~~ random vector ~~such that~~ and $g : \mathbb{R}^n \to \mathbb{R}$ such that $g^{-1}(B) \in \mathcal{B}(\mathbb{R}^n)$ for every $B \in \mathcal{B}(\mathbb{R})$.

Then we can compute

"Expected value of $\mathbb{E}(g(X_1, X_2, \dots X_n))$

if $\mathbb{E}(|g(X_1, \dots X_n)|) < \infty$"

As before, we shall focus on the random vectors which are of either continuous or discrete type.

Some of the interesting functions are as follows:

$$g(x_1, x_2, \dots, x_n) = \sum_{j=1}^{n} x_j$$

$$g(x_1, x_2, \dots x_n) = \prod_{j=1}^{n} x_j^{k_j} \quad \text{where } (k_1, k_2, \dots, k_n) \in \mathbb{N}^n.$$

$$g(x_1, x_2, \dots, x_n) = e^{\sum_{j=1}^{n} t_j x_j} \quad \text{where } (t_1 \dots t_n) \in \mathbb{R}^n.$$

For the ~~joit~~ joint moments, the following function turns out to be the most important.

$$g(x_1, x_2, \dots x_n) = \prod_{j=1}^{n} x_j^{k_j} \quad \text{where } (k_1, k_2, \dots k_n) \in \mathbb{N}^n.$$

Definition (Moment generating function of a random vector)

Let $(X_1, X_2, \dots X_n) : (\Omega, \mathcal{F}, \mathbb{P}) \longrightarrow (\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$

be a random vector. If

$$\mathbb{E}\left(e^{\sum_{j=1}^{n} t_j X_j}\right) < \infty \quad \text{for } |t_j| \le h_j \,;\, j = 1, 2 \ldots \text{ for some}$$

$$(h_1, h_2, \ldots, h_n) \in (0, \infty)^n,$$

then we write

$$M(t_1, t_2, \ldots t_n) = \mathbb{E}\left(e^{\sum_{j=1}^{n} t_j X_j}\right)$$

and call it the moment generating function (m.g.f.) of the random vector $(X_1, X_2, \ldots, X_n)$.

* We can define the characteristic function of a random vector in similar way. But we are not ~~goint~~ going to do that here.

We state some properties of moment generating functions without proof.

Theorem

The moment generating function of a random vector uniquely determines the joint distribution of the random vector and conversely, if moment generating function exists, it is unique.

As the joint distribution of a random vector leads to the marginal distributions of its components, it is natural that the marginal distributions ~~can~~ can be uniquely determined from the joint moment generating function.

Note that if

$$M_{B}(t_1, t_2 \ldots t_n) = E\left(\exp\left\{\sum_{i=1}^{n} t_i X_i\right\}\right).$$

then,

$$M(t_1, 0, 0, \ldots) = E\left(\exp\{t_1 X_1\}\right).$$

In general,

$$M(0, 0, 0 \ldots 0, t_j, 0, \ldots 0) = E\left(\exp\{t_j X_j\}\right)$$

turns out to be the moment generating function of the random variable X_j and hence, uniquely characterizes the marginal distribution of X_j.

<u>Theorem</u> (Joint moments and moment generating $f^{\underline{n}}$)

If $M(t_1 \ldots t_n)$ exists, then the moments of all orders of $(X_1 \ldots, X_n)$ exist and may be obtained from

$$\left.\frac{\partial^{k_1+k_2+\cdots+k_n}}{\partial t_1^{k_1}\, \partial t_2^{k_2} \cdots \partial t_n^{k_n}}\right|_{t_1=t_2=\cdots=t_n=0} = E\left(\prod_{i=1}^{n} X_i^{k_i}\right)$$

Theorem

Let $(X_1, X_2, \ldots, X_n)$ be a ~~collection of independent random variable~~ random vector (that is, defined on the same probability space). Then its components are independent if and only if

$$M(t_1, t_2, \ldots t_n) = \prod_{j=1}^{n} M(0, 0, \ldots t_j, 0, 0, \ldots 0).$$

whenever $M(t_1, t_2, \ldots t_n)$ exists.

Theorem

Let $(X_1, X_2, \ldots X_n)$ be a random vector with independent components and moment generating function of X_i is given by $M_i(s)$. for $i = 1, 2, \ldots n$. Then the moment generating function of $Y = \sum_{i=1}^{n} a_i X_i$ for real numbers $(a_1, a_2, \ldots a_n)$ is given by

$$M_Y(s) = \prod_{i=1}^{n} M_i(a_i s).$$

Moments and moment inequalities of random vectors

Recall that if it exists, the correlation function of X and Y is given by

~~Exercise~~ $\rho = \rho_{X,Y} = \dfrac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)\,\mathrm{Var}(Y)}}$.

where $(X,Y):(\Omega,\mathcal{F},\mathbb{P}) \longrightarrow (\mathbb{R}^2,\mathcal{B}(\mathbb{R}^2))$ be a random vector.

Remark. ρ is known as a measure of the linear dependence between the random variables X and Y.

Remark It follows easily that $|ab| \le 2^{-1}(a^2+b^2)$. and so we can say $\mathbb{E}(|XY|)$ exists if both $\mathbb{E}(X^2)$ and $\mathbb{E}(Y^2)$ exist.

Remark We say X and Y to be uncorrelated if

$$\rho_{X,Y} = 0.$$

Remark The correlation coefficient ρ between two random variables X and Y satisfies $|\rho| \le 1$. The equality $|\rho| = 1$ holds if and only if there exist constants a and b such that $\mathbb{P}(Y = aX+b) = 1$.

Exercise. Let $\mathbb{E}(X^2) < \infty$, $\mathbb{E}(Y^2) < \infty$ and let $U = aX+b$ and $V = cY+d$. Then

$$\rho_{X,Y} = \pm \rho_{U,V}$$

where (X,Y) is a random vector.

<u>Exercise</u> Let $(X_1, X_2, \ldots, X_n)$ be a random vector such that $\mathbb{E}|X_i| < \infty$ for $i = 1, 2, \ldots n$. Consider $(a_1, a_2, \ldots, a_n) \in \mathbb{R}^n$ ~~such that~~ and define

$$S = \sum_{i=1}^{n} a_i X_i.$$

Then $\mathbb{E}(S)$ exists and we have $\mathbb{E}(S) = \sum_{i=1}^{n} a_i \mathbb{E}(X_i)$.

<u>Exercise.</u>

Let $(X_1, X_2, \ldots, X_n)$ be a random vector with independent components. ~~Then~~ such that $\mathbb{E}(|X_i|) < \infty$. Then $\mathbb{E}\left(\prod_{i=1}^{n} X_i\right)$ exists and

$$\mathbb{E}\left(\prod_{i=1}^{n} X_i\right) = \prod_{i=1}^{n} \mathbb{E}(X_i).$$

<u>~~Exer~~ Theorem</u>

Let $(X_i : 1 \le i \le n)$ be a random vector composed of independent random variables. Consider $g_i : \mathbb{R} \to \mathbb{R}$ such that $g_i^{-1}(B) \in \mathcal{B}(\mathbb{R})$ for every $B \in \mathcal{B}(\mathbb{R})$, for $i = 1, 2, \ldots n$. ~~Then~~ such that $\mathbb{E}(|g_i(X_i)|) < \infty$ for $i = 1, 2, \ldots n$. Then we have

$$\mathbb{E}\left[\prod_{i=1}^{n} g_i(X_i)\right] = \prod_{i=1}^{n} \mathbb{E}(g_i(X_i)).$$

Exercise

Let $(X_1, X_2, \ldots X_n)$ be a random vector such that $\mathbb{E}(X_i^2) < \infty$ for $i = 1, 2, \ldots n$. Consider a deterministic n-tuple $(a_1, \ldots a_n)$ and define

$$S = \sum_{i=1}^{n} a_i X_i.$$

Then it follows that

$$\mathrm{Var}(S) = \sum_{i=1}^{n} a_i^2 \,\mathrm{Var}(X_i) + 2\sum_{i=1}^{n} \sum_{j=1, j>i}^{n} a_i a_j \,\mathrm{Cov}(X_i, X_j).$$

Hence it follows that

$$\mathrm{Var}(S) = \sum_{i=1}^{n} a_i^2 \,\mathrm{Var}(X_i)$$

if X_i's are independent random variables.

Exercise

Let $(X_1, X_2, \ldots, X_n)$ be a finite collection of independent random variables. Consider a deterministic n-tuple $(a_1, a_2, \ldots, a_n)$ such that

$$\sum_{i=1}^{n} a_i = 1.$$

Assume that $\mathbb{E}(X_i^2) < \infty$ for $1 \le i \le n$. and $\mathrm{Var}(X_i) = \sigma_i^2$ for $i = 1, 2 \ldots n$. Define

$$S_n = \sum_{i=1}^{n} a_i X_i \quad \text{and so} \quad \mathrm{Var}(S_n) = \sum_{i=1}^{n} a_i^2 \sigma_i^2.$$

Find the weights $(a_1, \ldots a_n)$ such that $\mathrm{Var}(S_n)$ is minimum.

An inequality

$$|a+b|^r \le c_r\left(|a|^r+|b|^r\right)$$

where
$$c_r = \begin{cases} 1 & \text{for } 0\le r\le 1 \\ 2^{r-1} & \text{for } r>1 \end{cases}$$

Note that it is enough to consider $0<a\le b$. Consider $x=a/b$. Then

$$\frac{(a+b)^r}{a^r+b^r} = \frac{(1+x)^r}{1+x^r}.$$

Define $f(x) = \dfrac{(1+x)^r}{1+x^r}$ for $x\in(0,1]$.

$$f'(x) = \frac{(1+x^r)\,r(1+x)^{r-1} - (1+x)^r\,r\,x^{r-1}}{(1+x^r)^2}.$$

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

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

$$= \frac{r(1+x)^{r-1}}{(1+x^r)^2}\left[1-x^{r-1}\right]. \quad \text{for } x\in(0,1].$$

It follows that $f'(x)>0$ if $r>1$ $=0$ if $r=1$ and <0 if $r<1$. Thus,

$$\max_{x\in[0,1]} f(x) = f(0) = 1 \quad \text{if } r\le 1.$$

$$\max_{0\le x\le 1} f(x) = f(1) = 2^{r-1} \quad \text{if } r>1.$$

and the desired inequality follows.

Theorem

Let $(X, Y) : (\Omega, \mathcal{F}, \mathbb{P}) \to (\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2))$ be a random vector. Let $r > 0$ be a fixed number such that $\mathbb{E}(|X|^r) < \infty$ and $\mathbb{E}(|Y|^r) < \infty$. Then $\mathbb{E}(|X+Y|^r) < \infty$ and furthermore,

$$\mathbb{E}(|X+Y|^r) \leq c_r \left(\mathbb{E}[|X|^r] + \mathbb{E}[|Y|^r] \right).$$

Holder's inequality.

Consider two positive real numbers p and q such that $1/p + 1/q = 1$. Then we have

$$|xy| \leq \frac{1}{p}|x|^p + \frac{1}{q}|y|^q.$$

The inequality follows from convexity of the function $x \mapsto -\log x$. It follows for any two positive real numbers λ_1, λ_2 such that $\lambda_1 + \lambda_2 = 1$, we have

$$-\log(\lambda_1 x + \lambda_2 y) \leq -\lambda_1 \log x - \lambda_2 \log y$$

that is, $\log(\lambda_1 x + \lambda_2 y) \geq \log x^{\lambda_1} y^{\lambda_2}$

that is, $\lambda_1 x + \lambda_2 y \geq x^{\lambda_1} y^{\lambda_2}$

for $x, y > 0$. Take $x = |x|^p$ and $y = |y|^q$ and $\lambda_1 = \frac{1}{p}$ and $\lambda_2 = \frac{1}{q}$ to get Holder's inequality.