Lecture 21 - Multivariate Random Variables

If $X_1, \dots, X_n$ are r.v.s. then

$$\underset{\sim}{X} = \begin{bmatrix} X_1 \\ \vdots \\ X_n \end{bmatrix} = (X_1, \dots, X_n)'$$

is called a <u>multivariate random variable</u>

or a <u>random vector</u>

<u>Defn:</u> PMFs/PDFs

If X_i's are <u>discrete</u>, then the <u>joint pmf</u>

$$f(\underset{\sim}{x}) = f(x_1, x_2, \dots, x_n)$$

$$= \mathbb{P}(X_1 = x_1, X_2 = x_2, X_3 = x_3, \dots, X_n = x_n)$$

So $f : \mathbb{R}^n \to \mathbb{R}$ $\qquad \underset{\sim}{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$

<u>Cts case:</u>

If the X_i are continuous then the

<u>joint pdf</u> is defined as

$$f : \mathbb{R}^n \longrightarrow \mathbb{R}$$

So that $A \subset \mathbb{R}^n$, then

$$\mathbb{P}(\underset{\sim}{X} \in A) = \mathbb{P}((X_1, X_2, \ldots, X_n) \in A)$$

$$= \int_A f(x_1, x_2, x_3, \ldots, x_n) \, d\underset{\sim}{x}$$

$$= \int\int_A \cdots \int f(x_1, x_2, \ldots, x_n) \, dx_1 \, dx_2 \cdots dx_n$$

Expectation

If $g : \mathbb{R}^n \longrightarrow \mathbb{R}$ then

$$\mathbb{E}\left[g(\underset{\sim}{X})\right] = \begin{cases} \sum_{x_1} \sum_{x_2} \cdots \sum_{x_n} g(x_1, \ldots, x_n) f(x_1, \ldots, x_n) & \text{(discrete)} \\ \int\int \cdots \int g(x_1, \ldots, x_n) f(x_1, \ldots, x_n) \, dx_1 \cdots dx_n & \text{(cts)} \end{cases}$$

$\underset{\sim}{X} = (X_1, \ldots, X_n)'$

Theorem: Marginal Distribution

Theorem: Marginal Distribution

The marginal pmf/pdf of X_i is

$$f_{X_i}(x_i) = \begin{cases} \underbrace{\sum_{x_1}\sum_{x_2}\cdots\sum_{x_{i-1}}\sum_{x_{i+1}}\cdots\sum_{x_n}}_{\text{sum over all other}} f(x_1, x_2, \ldots, x_n) \\ \int \cdots\cdots \int f(x_1, \ldots x_n)\, dx_1 \cdots dx_{i-1}\, dx_{i+1} \cdots dx_n \end{cases}$$

Conditional Distributions

Two seq. of indices

$\underbrace{i_1, \ldots, i_m}$ and $\underbrace{j_1, \ldots, j_k}$

the conditional pmf/pdf of $X_{i_1}, \ldots, X_{i_m}$ given $X_{j_1}, \ldots, X_{j_k}$ is

$$\boxed{f(x_{i_1}, \ldots, x_{i_m} \mid x_{j_1}, \ldots, x_{j_k}) = \frac{f(x_{i_1}, \ldots, x_{i_m}, x_{j_1}, \ldots, x_{j_k})}{f(x_{j_1}, \ldots, x_{j_k})}}$$

Ex. $X_1, \dots, X_4$ with a joint pdf of

$$\boxed{f(x_1, x_2, x_3, x_4) = \frac{3}{4}(x_1^2 + x_2^2 + x_3^2 + x_4^2) \quad \text{for } 0 < x_i < 1}$$

(a) $P(X_1 < \tfrac{1}{2}, X_2 < \tfrac{3}{4}, X_4 > \tfrac{1}{2})$

$$= \int_0^1 \int_{1/2}^1 \int_0^{3/4} \int_0^{1/2} \frac{3}{4}(x_1^2 + x_2^2 + x_3^2 + x_4^2)\, dx_1\, dx_2\, dx_4\, dx_3$$

$$= \cdots \text{Calc III} = 3/256$$

(b) Joint dist of X_1 and X_2?

$$f(x_1, x_2) = \int_0^1 \int_0^1 \underbrace{f(x_1, x_2, x_3, x_4)}_{\frac{3}{4}(x_1^2 + x_2^2 + x_3^2 + x_4^2)}\, dx_3\, dx_4$$

$$= \cdots \text{Calc III} = \boxed{\frac{1}{2} + \frac{3}{4}(x_1^2 + x_2^2)}$$

(c) $E[X_1 X_2] = \int_0^1\int_0^1 x_1 x_2 \left(\frac{1}{2}+\frac{3}{4}(x_1^2+x_2^2)\right) dx_1 dx_2$

$= \cdots$ Calc III $= 5/16$

(d) Conditional Dist. of X_3 and X_4 given X_1 and X_2

$$f(x_3, x_4 | x_1, x_2) = \frac{f(x_1, x_2, x_3, x_4)}{f(x_1, x_2)}$$

$$\boxed{= \frac{\frac{3}{4}(x_1^2+x_2^2+x_3^2+x_4^2)}{\frac{1}{2}+\frac{3}{4}(x_1^2+x_2^2)}}$$

Mutual Independence

We say $X_1, X_2, X_3, \ldots, X_n$ are mutually independent if for any sets $A_1, \ldots, A_n \subset \mathbb{R}$

$$P(X_1 \in A_1, X_2 \in A_2, X_3 \in A_3, \ldots, X_n \in A_n)$$

$$= P(X_1 \in A_1) P(X_2 \in A_2) \cdots P(X_n \in A_n).$$

Theorem: $\{X_i\}_{i=1}^n$ are independent iff

Theorem: $\{X_i\}_{i=1}^{n}$ are independent iff

(a) $f(x_1, x_2, \ldots, x_n) = f_{X_1}(x_1) f_{X_2}(x_2) \cdots f_{X_n}(x_n)$

(b) $F(x_1, \ldots, x_n) = F_{X_1}(x_1) F_{X_2}(x_2) \cdots F_{X_n}(x_n)$

Theorem: If $\{X_i\}$ are independent and $g_i : \mathbb{R} \to \mathbb{R}$

then
$$E[g_1(X_1) g_2(X_2) \cdots g_n(X_n)]$$
$$= E[g_1(X_1)] E[g_2(X_n)] \cdots E[g_n(X_n)].$$

Corollary: $\{g_i(X_i)\}$ are independent

Corollary: MGF of Sum of Independent

If $\{X_i\}$ are independent and $Z = \sum_{i=1}^{n} X_i$ then
$$\boxed{M_Z(t) = \prod_{i=1}^{n} M_{X_i}(t)}$$

Pf. $M_Z(t) = E[e^{tZ}] = E[e^{t \sum_{i=1}^{n} X_i}]$
$$= E\left[\prod_{i=1}^{n} e^{tX_i}\right] \qquad g_i(X_i) = e^{tX_i}$$

$$= \prod_{i=1}^{n} \underbrace{E[e^{tX_i}]}_{M_{X_i}(t)}$$

Follow on theorem:

If a_i and b_i are constants then

$$Z = \sum_{i=1}^{n} (a_i + b_i X_i)$$

then

$$\boxed{M_Z(t) = e^{t \sum_{i=1}^{n} b_i} \prod_{i=1}^{n} M_{X_i}(a_i t)}$$

Ex, $X_i \sim N(\mu_i, \sigma_i^2)$ and the X_i are independent then

$$Y = \sum_{i=1}^{n} (a_i + b_i X_i) \sim N\left(\sum_{i=1}^{n} (b_i \mu_i + a_i), \sum_{i=1}^{n} b_i^2 \sigma_i^2\right)$$

How to show: Use prev. theorem.

Multivariate Transformation

Consider

$g : \mathbb{R}^n \to \mathbb{R}^n$ and $\underset{\sim}{X} = (X_1, \ldots, X_n)$

is a random vector.

Define
$$\boxed{U = g(\underset{\sim}{X})}$$

i.e. $U_i = g_i(X_1, \dots, X_n)$

$g_i : \mathbb{R}^n \to \mathbb{R}$ is the i^{th} component fn of g

Theorem: In the continuous case, if g is invertible and differentiable then

$$f_U(u) = f_{\underset{\sim}{X}}(g^{-1}(u)) \left| \det\left(\frac{\partial g^{-1}}{\partial u}\right) \right|$$

$$\frac{\partial g^{-1}}{\partial u} = \begin{bmatrix} \frac{\partial g_1^{-1}}{\partial u_1} & \frac{\partial g_1^{-1}}{\partial u_2} & \cdots & \frac{\partial g_1^{-1}}{\partial u_n} \\ \frac{\partial g_2^{-1}}{\partial u_1} & & \cdots & \\ & & & \frac{\partial g_n^{-1}}{\partial u_n} \end{bmatrix}$$

Jacobian matrix $(n \times n)$

Ex. Let $X_i \overset{iid}{\sim} \text{Exp}(\lambda)$

iid = independent identically distributed

Let $U = g(\underset{\sim}{X})$

$U = (u_1, \dots, u_n)$

Specifically,

$$\begin{cases} u_1 = x_1 & \Leftrightarrow x_1 = u_1 = g_1^{-1}(u) \\ u_2 = x_1 + x_2 & \Leftrightarrow x_2 = u_2 - u_1 = g_2^{-1}(u) \\ u_3 = x_1 + x_2 + x_3 & \Leftrightarrow x_3 = u_3 - u_2 = g_3^{-1}(u) \\ \vdots & \vdots \\ u_n = x_1 + \dots + x_n & \Leftrightarrow x_n = u_n - u_{n-1} = g_n^{-1}(u) \end{cases}$$

inverse fn

$J = \frac{\partial g^{-1}}{\partial u}$ is an $n \times n$ matrix

$$\boxed{J_{ij} = \frac{\partial g_i^{-1}}{\partial u_j}}$$

$$J = \begin{bmatrix} 1 & 0 & & & & \\ -1 & 1 & & & 0 & \\ 0 & -1 & 1 & & & \\ \vdots & \vdots & -1 & \ddots & & \\ 0 & 0 & 0 & \cdots & -1 & 1 \end{bmatrix}$$

$|\det J| = 1$ b/c J triangler mtx
det is prod. of diagonal.

So $\boxed{f_U(u) = f_{\underset{\sim}{X}}(g^{-1}(u))}$

pdf of

recall: $X_i \overset{iid}{\sim} Exp(\lambda)$

pdf of $Exp(\lambda)$

$f(x) = \lambda e^{-\lambda x}$

$$f_{\underset{\sim}{X}}(\underset{\sim}{x}) = \prod_{i=1}^{n} f_{X_i}(x_i) = \prod_{i=1}^{n} f(x_i)$$

b/c independence

$$= \prod_{i=1}^{n} \lambda e^{-\lambda x_i} = \lambda^n \prod_{i=1}^{n} e^{-\lambda x_i}$$

$$f_U(u) = f_{\underset{\sim}{X}}(u_1, u_2 - u_1, u_3 - u_2, u_4 - u_3, \dots, u_n - u_{n-1})$$

$$= \lambda^n e^{-\lambda u_1} e^{-\lambda(u_2 - u_1)} e^{-\lambda(u_3 - u_2)} \cdots e^{-\lambda(u_n - u_{n-1})}$$

$$= \lambda^n e^{-\lambda u_n}$$

expand exponents:

$$-\cancel{\lambda u_1} - \cancel{\lambda u_2} + \cancel{\lambda u_1} - \cancel{\lambda u_3} + \cancel{\lambda u_2} + \dots + \dots - \lambda u_n$$

$$\boxed{f_U(u) = \lambda^n e^{-\lambda u_n} \quad \text{for } 0 < u_1 < u_2 < u_3 < \dots < u_n}$$

recall $u_1 = X_1 < X_1 + X_2 = u_2$

so $u_1 < u_2$

Means / Variance in multivar. case

Univariate: $E[X] \in \mathbb{R}$

$$Var(X) = \mathbb{E}[(X - E[X])^2] \in \mathbb{R}$$

Multivariate: $\underset{\sim}{X} = (X_1, \ldots, X_n)'$

$$\mu = \mathbb{E}[\underset{\sim}{X}] = \begin{bmatrix} \mathbb{E}[X_1] \\ E[X_2] \\ \vdots \\ E(X_n] \end{bmatrix} \in \mathbb{R}^n$$

Expected Value

Covariance Matrix

$$Cov(\underset{\sim}{X}) = \Sigma \in \mathbb{R}^{n \times n}$$

where $\Sigma_{ij} = Cov(X_i, X_j)$

notice $\Sigma_{ii} = Cov(X_i, X_i) = Var(X_i)$

$$\Sigma = \begin{bmatrix} Var(X_1) & Cov(X_1, X_2) & \\ Cov(X_2, X_1) & Var(X_2) & \\ & & \ddots \end{bmatrix}$$

symmetric

Aside: $\Sigma = \mathbb{E}[(\underset{\sim}{X} - E[\underset{\sim}{X}])(\underset{\sim}{X} - \mathbb{E}[\underset{\sim}{X}])']$

direct generalization

$n \times 1$ $\quad$ $1 \times n$

direct generalization of variance

$$\text{(similar } Var(X) = E[(X - E(X))(X - E(X))]$$

($n \times 1$, $1 \times n$)

Theorem:

If $a \in \mathbb{R}^m$ and $B \in \mathbb{R}^{m \times n}$ and $\underset{\sim}{X}$ is a n-comp. r. vector then

(a) $E[\underbrace{a + B\underset{\sim}{X}}_{\text{Vector in } \mathbb{R}^m}] = a + B\,E[\underset{\sim}{X}]$ (linearity of E)

(b) $Cov(\underbrace{a + B\underset{\sim}{X}}_{\mathbb{R}^m}) = \underbrace{B\,\overbrace{Cov(\underset{\sim}{X})}^{n \times n}\,B'}_{m \times m}$

(analogy univariate case:

$E[a + bX] = a + b\,E[X]$

and $Var(a + bX) = b^2\,Var(X)$)

Multivariate Normal

$\underset{\sim}{X} = (X_1, \ldots, X_n) \sim N(\mu, \Sigma)$

$\mu \in \mathbb{R}^n$ and $\Sigma \in \mathbb{R}^{n \times n}$

the pdf

$\underset{\sim}{X} = (x_1, .., x_n) \in \mathbb{R}$

$$f(\underset{\sim}{X}) = (2\pi)^{-n/2} \det(\Sigma)^{-1/2} \exp\Big(\underbrace{-\tfrac{1}{2}(\underset{\sim}{X}-\mu)'\Sigma^{-1}(\underset{\sim}{X}-\mu)}_{\text{a number}}\Big)$$

Univariate:

$$f(x) = (2\pi)^{-1/2}(\sigma^2)^{-1/2} \exp\Big(-\tfrac{1}{2}(x-\mu)(\sigma^2)^{-1}(x-\mu)\Big)$$

Special Case:

$\mu = 0 \in \mathbb{R}^n$ and $\Sigma = I_n$

then called standard MVN

Theorem: If $\underset{\sim}{X} \sim N(\mu, \Sigma)$ and $a \in \mathbb{R}^m$, $B \in \mathbb{R}^{n \times m}$

if $\underset{\sim}{Y} = a + B\underset{\sim}{X}$ a m-comp. r. vector

then

$$Y \sim N(a + B\mu,\ B\Sigma B') \quad (*)$$

Follow-up: $\underset{\sim}{X}$ is MVN $\Leftrightarrow$ every linear comb of its components is normal

Useful Corollaries:

Marginals of MVN are normal

Caution: If marginally X_i are normal no guarentee, that jointly they are normal.