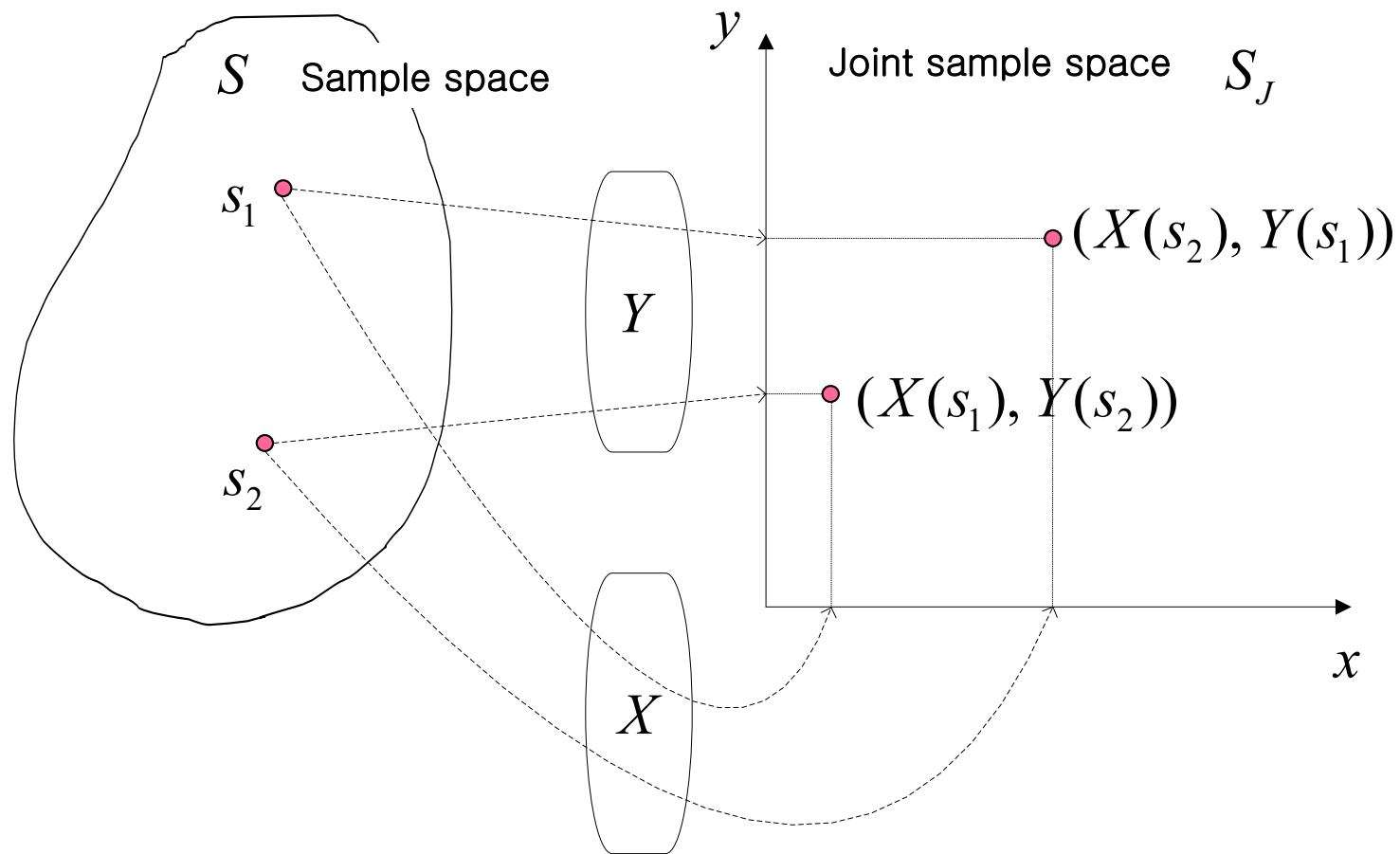


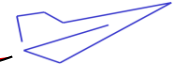
# Expectation and Estimation

# Multiple Random Variables

## □ 2차원 랜덤변수 사상



# Joint Cumulative Distribution Function



## □ Single Cumulative Distribution Function

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

$$F_Y(y) = P\{Y \leq y\}$$

## □ Joint Cumulative Distribution Function

$$F_{X,Y}(x, y) = P\{X \leq x \text{ and } Y \leq y\}$$

## □ Ex 4.2-1)

$$S_J = \{(1,1), (2,1), (3,3)\} \text{ and } P(1,1) = 0.2, P(2,1) = 0.3, P(3,3) = 0.5$$

$$\Rightarrow F_{X,Y}(0,1) = P\{X \leq 0, Y \leq 1\} = 0$$

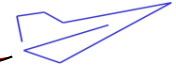
$$F_{X,Y}(1,1) = P\{X \leq 1, Y \leq 1\} = P(1,1) = 0.2$$

$$F_{X,Y}(2,1) = P\{X \leq 2, Y \leq 1\} = P(1,1) + P(2,1) = 0.5$$

$$F_{X,Y}(2,2) = P\{X \leq 2, Y \leq 2\} = P(1,1) + P(2,1) = 0.5$$

$$F_{X,Y}(3,3) = P\{X \leq 3, Y \leq 3\} = P(1,1) + P(2,1) + P(3,3) = 1.0$$

# Joint Cumulative Distribution Function



## □ Properties of Joint Cumulative Distribution Function

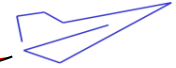
- $F_{X,Y}(-\infty, -\infty) = F_{X,Y}(-\infty, y) = F_{X,Y}(x, -\infty) = 0$
- $F_{X,Y}(\infty, \infty) = 1$
- $0 \leq F_{X,Y}(x, y) \leq 1$
- $F_{X,Y}(x, y)$  is a non-decreasing function.
- $F_{X,Y}(x_2, y_2) + F_{X,Y}(x_1, y_1) - F_{X,Y}(x_1, y_2) - F_{X,Y}(x_2, y_1)$   
 $= P\{x_1 < X \leq x_2, y_1 < Y \leq y_2\} \geq 0$
- $F_{X,Y}(x, \infty) = F_X(x), F_{X,Y}(\infty, y) = F_Y(y)$

Let  $A = \{X \leq x\}$ ,  $B = \{Y \leq y\}$ , then  $F_{X,Y}(x, y) = P\{X \leq x, Y \leq y\} = P[A \cap B]$ .

For  $y \rightarrow \infty$ ,  $B = \{Y \leq \infty\} = S \Rightarrow A \cap B = A$

Therefore,  $F_{X,Y}(x, \infty) = P[A \cap S] = P[A] = P\{X \leq x\} = F_X(x)$

# Joint Probability Density Function



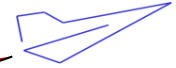
## □ Joint Probability Density Function

$$f_{X,Y}(x,y) = \frac{\partial^2 F_{X,Y}(x,y)}{\partial x \partial y}$$

## □ Properties of Joint pdf

- $f_{X,Y}(x,y) \geq 0$
- $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(x,y) dx dy = 1$
- $F_{X,Y}(x,y) = \int_{-\infty}^y \int_{-\infty}^x f_{X,Y}(\xi_1, \xi_2) d\xi_1 d\xi_2$
- $F_X(x) = \int_{-\infty}^x \int_{-\infty}^{\infty} f_{X,Y}(\xi_1, \xi_2) d\xi_2 d\xi_1, \quad F_Y(y) = \int_{-\infty}^y \int_{-\infty}^{\infty} f_{X,Y}(\xi_1, \xi_2) d\xi_1 d\xi_2$
- $P\{x_1 < X \leq x_2, y_1 < Y \leq y_2\} = \int_{y_1}^{y_2} \int_{x_1}^{x_2} f_{X,Y}(\xi_1, \xi_2) d\xi_1 d\xi_2$
- $f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y) dy, \quad f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y) dx \quad (*)$

# Joint Probability Density Function

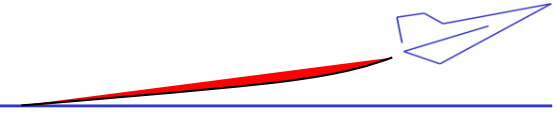


## □ Proof of (\*)

$$\begin{aligned} f_X(x) &= \frac{dF_X(x)}{dx} = \frac{d}{dx} \int_{-\infty}^x \underbrace{\int_{-\infty}^{\infty} f_{X,Y}(\xi_1, \xi_2) d\xi_2}_{\triangleq k(\xi_1)} d\xi_1 \\ &= \frac{d}{dx} [K(x) - K(-\infty)], \quad K = \int k(\xi_1) d\xi_1 \\ &= k(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, \xi_2) d\xi_2 = \int_{-\infty}^{\infty} f_{X,Y}(x, y) dy \end{aligned}$$

In similar way,  $f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x, y) dx$

# Conditional Joint Probability



## □ Conditional Probability of X on B

$$F_X(x | B) = P\{X \leq x | B\} = \frac{P\{X \leq x \cap B\}}{P[B]} \text{ for } P[B] \neq 0$$

$$\Rightarrow f_X(x | B) = \frac{dF_X(x | B)}{dx}$$

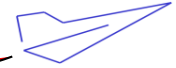
## □ Conditional Probability of X

Let  $B = \{y - \Delta y < Y \leq y + \Delta y\}$

then  $F_X(x | B) = F_X(x | y - \Delta y < Y \leq y + \Delta y)$

$$= \frac{P[X \leq x \cap \{y - \Delta y < Y \leq y + \Delta y\}]}{P\{y - \Delta y < Y \leq y + \Delta y\}}.$$

# Conditional Joint Probability



## □ Conditional Probability of X

Recall that  $P\{x_1 < X \leq x_2, y_1 < Y \leq y_2\} = \int_{y_1}^{y_2} \int_{x_1}^{x_2} f_{X,Y}(\xi_1, \xi_2) d\xi_1 d\xi_2$ .

$$F_X(x | B) = \frac{\int_{y-\Delta y}^{y+\Delta y} \int_{-\infty}^x f_{X,Y}(\xi_1, \xi_2) d\xi_1 d\xi_2}{\int_{y-\Delta y}^{y+\Delta y} f_Y(\xi) d\xi}$$
$$\approx \frac{\int_{-\infty}^x f_{X,Y}(\xi_1, y) d\xi_1 \cdot 2\Delta y}{f_Y(y) \cdot 2\Delta y}$$

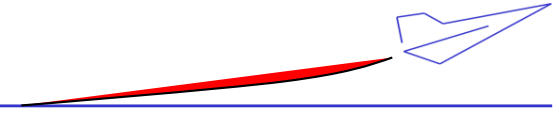
$$\text{If } \Delta y \rightarrow 0, F_X(x | B) \rightarrow F_X(x | Y = y) = \frac{\int_{-\infty}^x f_{X,Y}(\xi_1, y) d\xi_1}{f_Y(y)}.$$

$$\Rightarrow f_X(x | Y = y) = \frac{dF_X(x | Y = y)}{dx} = \frac{f_{X,Y}(x, y)}{f_Y(y)}$$

$$\text{Simply, } f_X(x | y) = f_X(x | Y = y) = \frac{f_{X,Y}(x, y)}{f_Y(y)}, f_Y(y | x) = \frac{f_{X,Y}(x, y)}{f_X(x)}$$



# Conditional Joint Probability



## □ Statistical Independence

It is said that  $A$  is statistically independent on  $B$ ,  
if  $P[A \cap B] = P[A]P[B]$ .

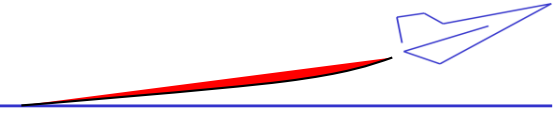
Let  $A = \{X \leq x\}$ ,  $B = \{Y \leq y\}$ , then  $P\{X \leq x, Y \leq y\} = P\{X \leq x\}P\{Y \leq y\}$ .  
 $\Rightarrow F_{X,Y}(x, y) = F_X(x)F_Y(y)$

$$\begin{aligned} F_X(x | Y \leq y) &= \frac{P\{X \leq x \cap Y \leq y\}}{P\{Y \leq y\}} = \frac{F_{X,Y}(x, y)}{F_Y(y)} = \frac{F_X(x)F_Y(y)}{F_Y(y)} \\ &= F_X(x) \end{aligned}$$

In similar way,  $F_Y(y | X \leq x) = F_Y(y)$

Also,  $f_X(x | Y \leq y) = f_X(x)$ ,  $f_Y(y | X \leq x) = f_Y(y)$

# Conditional Joint Probability



## □ Example

Let the joint density of two random variables  $X$  and  $Y$  be given by

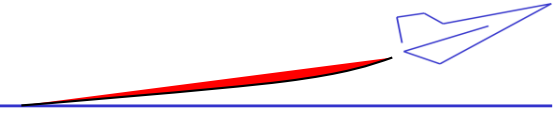
$$f_{X,Y}(x,y) = \begin{cases} \frac{1}{6}(x+4y), & 0 < x < 2, \ 0 < y < 1 \\ 0, & \text{otherwise} \end{cases}$$

1)  $f_{X,Y}(x,y) \geq 0$  and

$$\begin{aligned} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(x,y) dx dy &= \int_0^1 \int_0^2 \frac{1}{6}(x+4y) dx dy \\ &= \int_0^1 \left[ \frac{1}{6} \left( \frac{1}{2} x^2 + 4xy \right) \right]_{x=0}^{x=2} dy = \int_0^1 \frac{1}{6} (2 + 8y) dy \\ &= \left[ \frac{1}{6} (2y + 4y^2) \right]_{y=0}^{y=1} = 1 \end{aligned}$$

$\Rightarrow$  Probability Density function

# Conditional Joint Probability



## □ Example

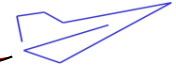
$$2) f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)dy = \int_0^1 \frac{1}{6}(x+4y)dy = \frac{1}{6}(xy + 2y^2) \Big|_0^1 = \frac{1}{6}(x+2)$$

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)dx = \int_0^2 \frac{1}{6}(x+4y)dx = \frac{1}{6}\left(\frac{1}{2}x^2 + 4xy\right) \Big|_0^2 = \frac{1}{6}(2+8y) = \frac{1}{3}(1+4y)$$

$$f_X(x|y) = f_X(x|Y=y) = \begin{cases} \frac{f_{X,Y}(x,y)}{f_Y(y)} = \frac{\frac{1}{6}(x+4y)}{\frac{1}{3}(1+4y)} = \frac{x+4y}{2(1+4y)}, & 0 < x < 2 \\ 0, & \text{otherwise} \end{cases}$$

$$\begin{aligned} 3) F_X(1|0.5) &= F_X(x|Y=0.5) = \frac{\int_{-\infty}^x f_{X,Y}(\xi_1, y)d\xi_1}{f_Y(y)} = \frac{\int_{-\infty}^1 f_{X,Y}(\xi_1, 0.5)d\xi_1}{f_Y(0.5)} \\ &= \frac{\int_0^1 \frac{1}{6}(x+2)dx}{1} = \frac{1}{6}\left(\frac{1}{2}x^2 + 2x\right) \Big|_0^1 = \frac{5}{12} \end{aligned}$$

# Expected Value



## □ Expected value or Mean of a discrete random variable X

$$m_X = E[X] = \sum_{x \in S_X} xp_X(x) = \sum_k x_k p_X(x_k)$$

- The expected value (or expectation) refers, intuitively, to the value of a random variable one would "expect" to find if one could repeat the random variable process an infinite number of times and take the average of the values obtained.

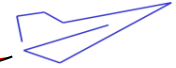
- The expected value is a weighted average of all possible values.

## □ Ex) 주사위를 1회 던졌을 때 나타나는 눈의 기대값

$$X = \{1, 2, 3, 4, 5, 6\}, p_X(x_i) = 1/6, i = 1, 2, \dots, 6$$

$$m_X = E[X] = \sum_k x_k p_X(x_k) = 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + \dots + 6 \cdot \frac{1}{6} = \frac{21}{6} = \frac{7}{2}$$

# Expected Value



## □ Expected value a random variable X (General meaning)

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

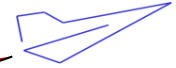
## □ Ex 3.1-2)

$$f_X(x) = \begin{cases} \frac{1}{b} e^{-(x-a)/b}, & x > a \\ 0, & x < a \end{cases}$$

$$E[X] = \frac{1}{b} \int_a^{\infty} x e^{-\frac{(x-a)}{b}} dx = \frac{1}{b} e^{\frac{a}{b}} \int_a^{\infty} x e^{-\frac{x}{b}} dx = \left( \frac{1}{b} e^{\frac{a}{b}} \right) \underbrace{\left[ e^{-\frac{x}{b}} \left( \frac{x}{-1/b} - \frac{1}{1/b^2} \right) \right]_a^{\infty}}_{\text{from C-46}}$$

$$= \left( \frac{1}{b} e^{\frac{a}{b}} \right) e^{-\frac{a}{b}} [ab + b^2] = a + b$$

# Expected Value



## □ Ex) Mean of normal pdf

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma_X^2}} e^{-\frac{(x-a_X)^2}{2\sigma_X^2}}$$

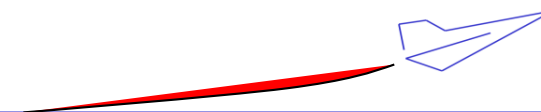
$$E[X] = \frac{1}{\sqrt{2\pi\sigma_X^2}} \int_{-\infty}^{\infty} x e^{-\frac{(x-a_X)^2}{2\sigma_X^2}} dx$$

$$= \frac{1}{\sqrt{2\pi\sigma_X^2}} \int_{-\infty}^{\infty} \left[ x e^{-\frac{(x-a_X)^2}{2\sigma_X^2}} - a_X e^{-\frac{(x-a_X)^2}{2\sigma_X^2}} + a_X e^{-\frac{(x-a_X)^2}{2\sigma_X^2}} \right] dx$$

$$= \frac{1}{\sqrt{2\pi\sigma_X^2}} \left[ \underbrace{\int_{-\infty}^{\infty} (x - a_X) e^{-\frac{(x-a_X)^2}{2\sigma_X^2}} dx}_{=0 \because \text{odd function}} + \int_{-\infty}^{\infty} a_X e^{-\frac{(x-a_X)^2}{2\sigma_X^2}} dx \right]$$

$$= a_X \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma_X^2}} e^{-\frac{(x-a_X)^2}{2\sigma_X^2}} dx = a_X \underbrace{\int_{-\infty}^{\infty} f_X(x) dx}_{=1} = a_X$$

# Moments



## □ Expected value functions of a random variable $X$

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

## □ Moments about the origin

Let  $g(X) = X^n$ ,  $n = 0, 1, 2, \dots$

then the  $n$ -th moment of  $X$  is given by  $m_n = E[X^n] = \int_{-\infty}^{\infty} x^n f_X(x) dx$

## □ Central Moments

Let  $g(X) = (X - \bar{X})^n$ ,  $n = 0, 1, 2, \dots$

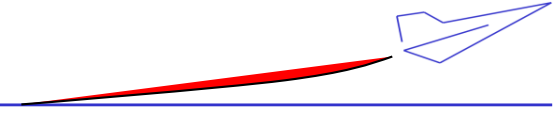
where  $\bar{X}$  is the mean of  $X$ , then the  $n$ -th central moment of  $X$

is given by  $\mu_n = E[(X - \bar{X})^n] = \int_{-\infty}^{\infty} (x - \bar{X})^n f_X(x) dx$ .

$$\text{cf) } \mu_0 = E[(X - \bar{X})^0] = \int_{-\infty}^{\infty} f_X(x) dx = 1$$

$$\mu_1 = E[(X - \bar{X})] = \int_{-\infty}^{\infty} (x - \bar{X}) f_X(x) dx = 0$$

# Moments

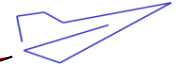


## □ Properties of Mean

- $E[g(X) + h(X)] = E[g(X)] + E[h(X)]$
- $E[ag(X)] = aE[g(X)]$
- $E[g(X) + c] = E[g(X)] + c$
- $E[c] = c$
- $E[\sum g_k(X)] = \sum E[g_k(X)]$



# Moments



□ **Ex)** 주사위를 1회 던졌을 때 나타나는 눈의 모멘트

$$X = \{1, 2, 3, 4, 5, 6\}, p_X(x_i) = 1/6, i = 1, 2, \dots, 6$$

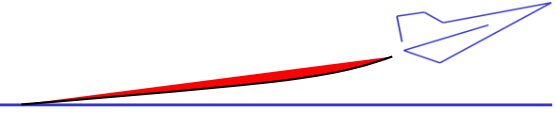
$$\text{Mean (the 1}^{st} \text{ moment): } \bar{X} = E[X] = \sum_k x_k p_X(x_k) = 1 \cdot \frac{1}{6} + \dots + 6 \cdot \frac{1}{6} = \frac{21}{6} = \frac{7}{2}$$

$$\text{The 2}^{nd} \text{ moment: } E[X^2] = \sum_k k^2 p_X(x_k) = 1^2 \cdot \frac{1}{6} + 2^2 \cdot \frac{1}{6} + \dots + 6^2 \cdot \frac{1}{6} = \frac{91}{6}$$

Variance(the 2<sup>nd</sup> central moment):

$$\begin{aligned} E[(X - \bar{X})^2] &= \sum_k \left(k - \frac{7}{2}\right)^2 p_X(x_k) \Leftarrow \text{Too tedious} \\ &= E[X^2 - 2X\bar{X} + \bar{X}^2] \\ &= E[X^2] - 2\bar{X} E[X] + \bar{X}^2 = E[X^2] - \bar{X}^2 \\ &= \frac{91}{6} - \left(\frac{7}{2}\right)^2 \end{aligned}$$

# Moments



## □ Ex) Gaussian Random Variable

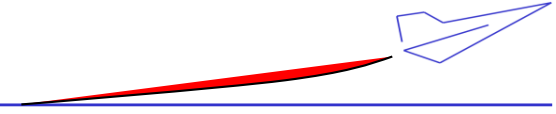
$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma_X^2}} e^{-\frac{(x-a_X)^2}{2\sigma_X^2}}$$

Mean (the 1<sup>st</sup> moment):  $\bar{X} = E[X] = a_X$

Variance (the 2<sup>nd</sup> central moment):  $E[(X - \bar{X})^2] = \sigma_X^2$

⇒ The pdf of the gaussian RV is represented by its mean and variance.

# Moments



## □ Joint moments about the origin

For two random variables  $X$  and  $Y$ ,

the joint moment is given by  $m_{nk} = E[X^n Y^k] = \int_{-\infty}^{\infty} x^n y^k f_{X,Y}(x, y) dx dy$ ,

where  $n + k$  is the order of the joint moment.

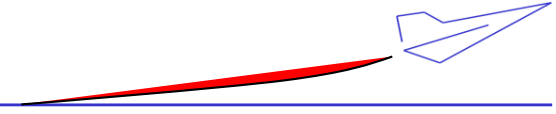
## □ Correlation

$$R_{XY} = m_{11} = E[XY] = \int_{-\infty}^{\infty} x y f_{X,Y}(x, y) dx dy$$

If  $R_{XY} = E[X]E[Y]$  is satisfied, we say that there is **no correlation** between  $X$  and  $Y$ . Or, it is said that  $X$  is statistically independent on  $Y$ .

If  $R_{XY} = 0$ , two random variables  $X$  and  $Y$  are **orthogonal** each other.

# Moments



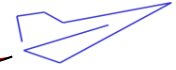
## □ Joint central moments

For two random variables  $X$  and  $Y$ ,

$$\mu_{nk} = E[(X - \bar{X})^n (Y - \bar{Y})^k] = \int_{-\infty}^{\infty} (x - \bar{X})^n (y - \bar{Y})^k f_{X,Y}(x, y) dx dy$$

is called the **joint central moment**.

# Moments



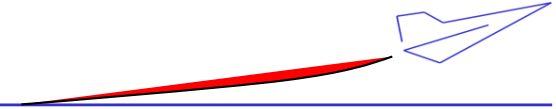
## □ Covariance

$$\begin{aligned} C_{XY} &= \mu_{11} = E[(X - \bar{X})(Y - \bar{Y})] = \int_{-\infty}^{\infty} (x - \bar{X})(y - \bar{Y}) f_{X,Y}(x, y) dx dy \\ &= \underbrace{\int_{-\infty}^{\infty} xy f_{X,Y}(x, y) dx dy}_{=E[XY]} - \underbrace{\bar{X} \int_{-\infty}^{\infty} y f_{X,Y}(x, y) dx dy}_{=E[\bar{X}Y] = \bar{X}E[Y] = \bar{X}\bar{Y}} - \underbrace{\bar{Y} \int_{-\infty}^{\infty} x f_{X,Y}(x, y) dx dy}_{=E[X\bar{Y}] = \bar{Y}E[X] = \bar{X}\bar{Y}} \\ &\quad + \underbrace{\bar{X}\bar{Y} \int_{-\infty}^{\infty} f_{X,Y}(x, y) dx dy}_{=1} = R_{XY} - \bar{X}\bar{Y} \end{aligned}$$

If  $C_{XY} = 0 (\Rightarrow R_{XY} = E[X]E[Y])$ , there is **no correlation or independent** each other between  $X$  and  $Y$ .

If  $X$  and  $Y$  are orthogonal ( $R_{XY} = 0$ ),  $C_{XY} = -E[X]E[Y]$ .

# Moments



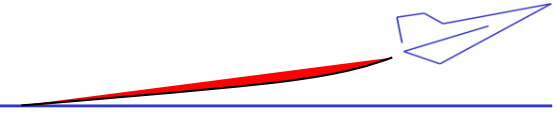
## □ Correlation coefficient

$$\rho = E \left[ \frac{(X - \bar{X})}{\sigma_X} \frac{(Y - \bar{Y})}{\sigma_Y} \right]$$

where  $\sigma_X^2 = E[(X - \bar{X})^2]$ ,  $\sigma_Y^2 = E[(Y - \bar{Y})^2]$ .

HW) Prove  $-1 \leq \rho \leq 1$ .

# Multivariate random variables



## □ Multivariate random variables (Random vectors)

- Vectors with random variables

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

- Mean

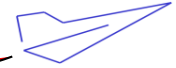
$$\mu = E[x] = \begin{bmatrix} E[x_1] \\ E[x_2] \\ \vdots \\ E[x_n] \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_n \end{bmatrix}$$

- Covariance

$$\Sigma = E[(x - \mu)(x - \mu)^T] =$$

$$\begin{bmatrix} E[(x_1 - \mu_1)(x_1 - \mu_1)] & E[(x_1 - \mu_1)(x_2 - \mu_2)] & \cdots & E[(x_1 - \mu_1)(x_n - \mu_n)] \\ E[(x_2 - \mu_2)(x_1 - \mu_1)] & E[(x_2 - \mu_2)(x_2 - \mu_2)] & \cdots & E[(x_2 - \mu_2)(x_n - \mu_n)] \\ \vdots & \vdots & \ddots & \vdots \\ E[(x_n - \mu_n)(x_1 - \mu_1)] & E[(x_n - \mu_n)(x_2 - \mu_2)] & \cdots & E[(x_n - \mu_n)(x_n - \mu_n)] \end{bmatrix} = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_2^2 & \cdots & \sigma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{n1} & \sigma_{n2} & \cdots & \sigma_n^2 \end{bmatrix}$$

# Multivariate random variables



## □ Mean and covariance under linear transformation

- Suppose the mean and the covariance of  $x$  is given by

Mean:  $E[x] = \bar{x}$

Covariance:  $E[(x - \bar{x})(x - \bar{x})^T] = \Sigma$

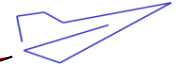
- A new variable via linear transformation  $y = Ax$ . Then,

Mean: 
$$\begin{aligned} E[y] &= E[Ax] \\ &= AE[x] \\ &= A\bar{x} \end{aligned}$$

Covariance: 
$$\begin{aligned} E[(y - \bar{y})(y - \bar{y})^T] &= E[(Ax - A\bar{x})(Ax - A\bar{x})^T] \\ &= E[A(x - \bar{x})(x - \bar{x})^T A^T] \\ &= AE[(x - \bar{x})(x - \bar{x})^T]A^T \\ &= A\Sigma A^T \end{aligned}$$



# Multivariate Gaussian Random Variables

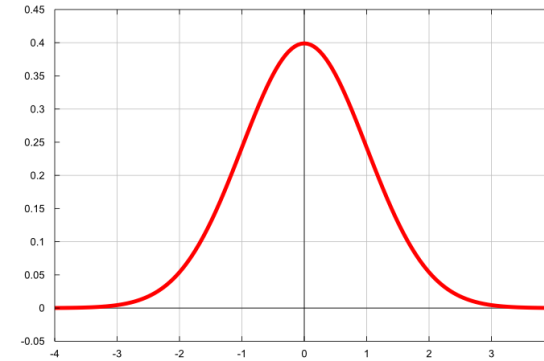


## □ Multivariate Gaussian Probability Density Function

- Univariate Gaussian pdf

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}}$$

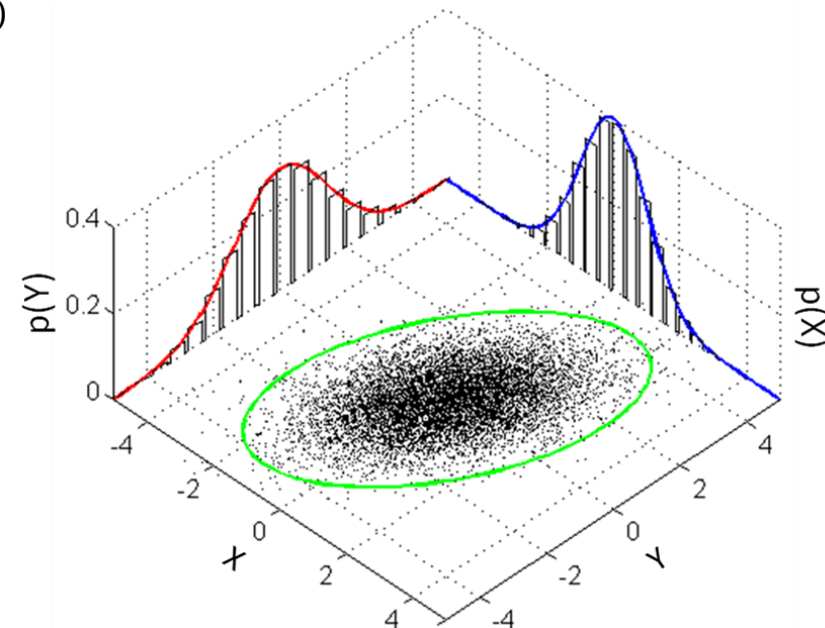
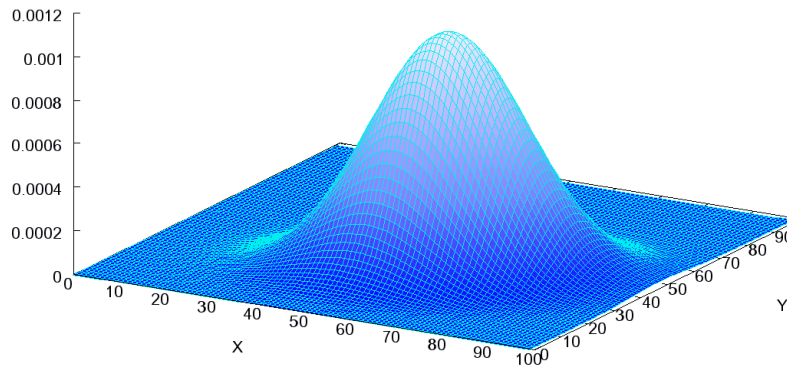
$$x \sim N(\mu, \sigma^2)$$



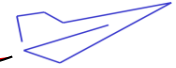
- Multivariate Gaussian pdf

$$f_X(x) = \frac{1}{\sqrt{(2\pi)^n |\Sigma|}} e^{-\frac{1}{2}(x-\mu)^T \Sigma^{-1} (x-\mu)}$$

$$x \sim N(\mu, \Sigma)$$

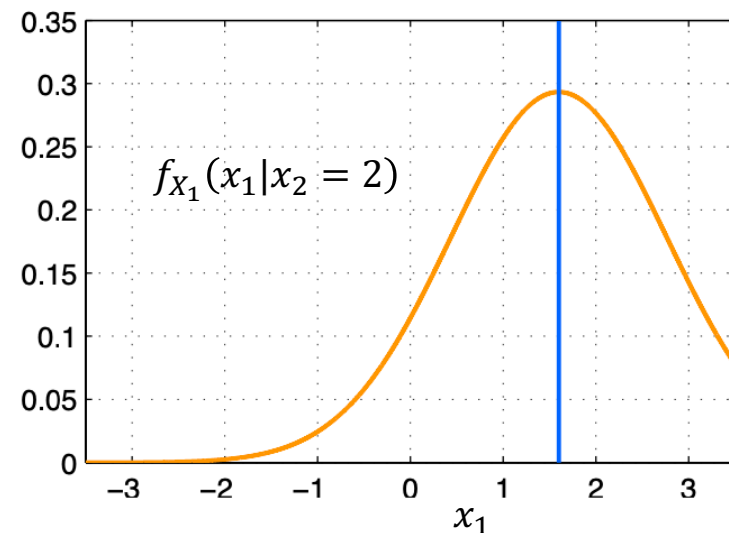
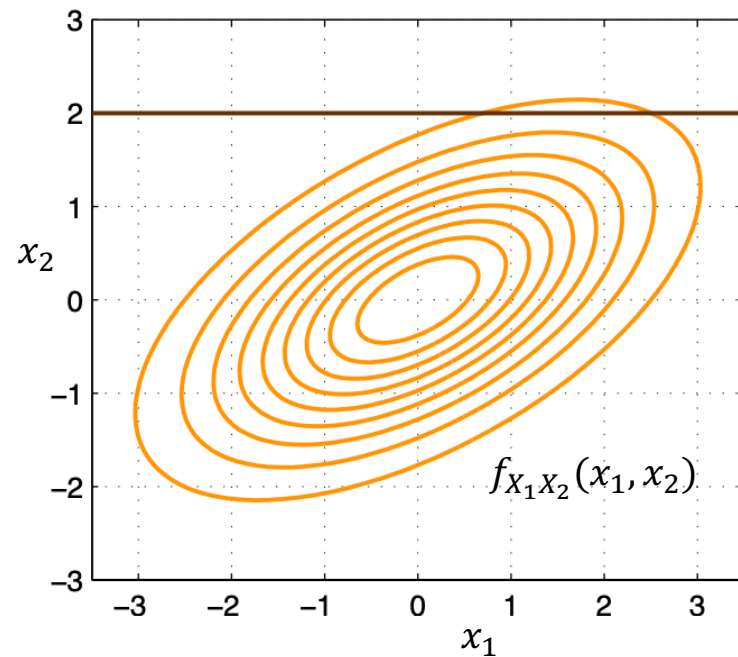


# Multivariate Gaussian Random Variables

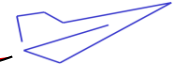


## □ Conditioning

- Given a joint pdf  $f_{X_1X_2}(x_1, x_2)$  on  $\mathbb{R}^2$
- Measure  $x_2$ ;  
we would like to find  
the conditional pdf of  $x_1$ ,
- For example,  
when we know  $x_2 = 2$ ,  
what is  $f_{X_1}(x_1|x_2 = 2)$ ?



# Multivariate Gaussian Random Variables



## □ Conditioning

- Suppose  $x \sim N(0, \Sigma)$  with

$$x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \quad \Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{12}^T & \Sigma_{22} \end{bmatrix}$$

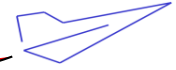
- Suppose we measure  $x_2 = y$ .
- Conditional pdf  $x_1$  given  $x_2 = y$ :

$$\begin{aligned} f_{X_1}(x_1 | x_2 = y) \\ = \frac{1}{\sqrt{(2\pi)^n |\Sigma_{11} - \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{12}^T|}} e^{-\frac{1}{2} (x - \Sigma_{12} \Sigma_{22}^{-1} y)^T (\Sigma_{11} - \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{12}^T)^{-1} (x - \Sigma_{12} \Sigma_{22}^{-1} y)} \end{aligned}$$

$$\Leftrightarrow f_{X_1}(x_1 | x_2 = y) \sim N(\Sigma_{12} \Sigma_{22}^{-1} y, \Sigma_{11} - \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{12}^T)$$

$$\Rightarrow E[x_1 | x_2 = y] = \Sigma_{12} \Sigma_{22}^{-1} y \quad \text{cov}[x_1 | x_2 = y] = \Sigma_{11} - \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{12}^T$$

# Multivariate Gaussian Random Variables



## □ Conditioning

- Example

$x \sim N(0, \Sigma)$  with

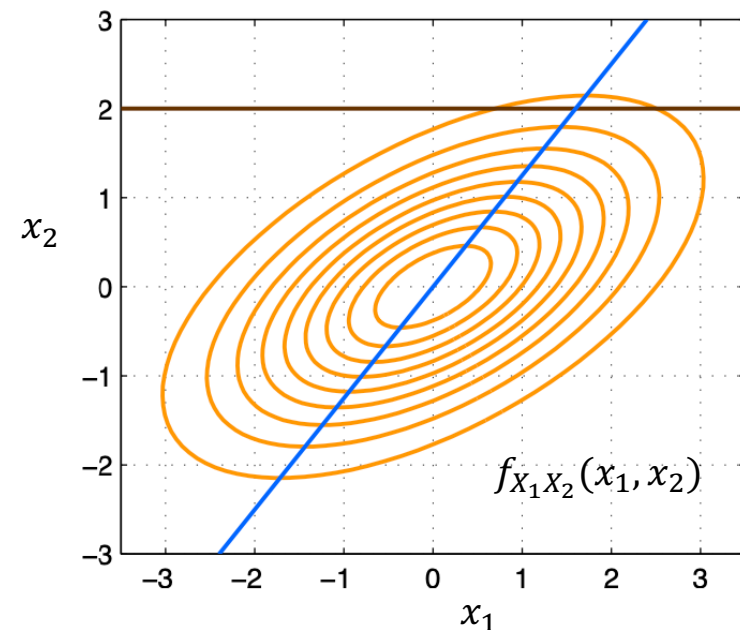
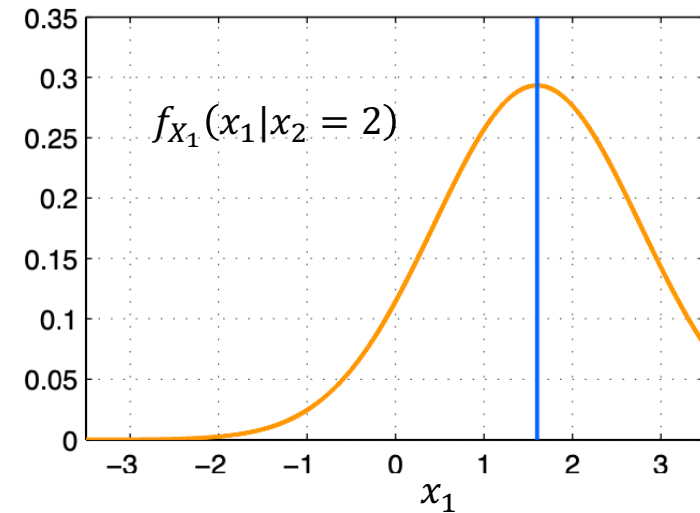
$$x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \quad \Sigma = \begin{bmatrix} 2 & 0.8 \\ 0.8 & 1 \end{bmatrix}$$

$$\Sigma_{12}\Sigma_{22}^{-1}y = 0.8y$$

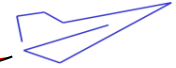
$$\Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{12}^T = 1.36$$

$$f_{X_1}(x_1|y) \sim N(0.8y, 1.36)$$

$$f_{X_1}(x_1|y = 2) \sim N(1.6, 1.36)$$



# Multivariate Gaussian Random Variables



## □ Estimation

- Suppose  $x \sim N(0, \Sigma_x)$ ,  $w \sim N(0, \Sigma_w)$  uncorrelated, and a measurement  $y = Ax + w$  is given:
- We want the best estimate on  $x$  given  $y$ , i.e.,  $E[x|y]$

- Let  $z = \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} I & 0 \\ A & I \end{bmatrix} \begin{bmatrix} x \\ w \end{bmatrix}$

- Then

$$\begin{aligned} \Sigma_z = \begin{bmatrix} x \\ y \end{bmatrix} &= \begin{bmatrix} I & 0 \\ A & I \end{bmatrix} \begin{bmatrix} \Sigma_x & 0 \\ 0 & \Sigma_w \end{bmatrix} \begin{bmatrix} I & A^T \\ 0 & I \end{bmatrix} \\ &= \begin{bmatrix} \Sigma_x & \Sigma_x A^T \\ A \Sigma_x & A \Sigma_x A^T + \Sigma_w \end{bmatrix} \end{aligned}$$

$$\Rightarrow E[x|y] = \Sigma_x A^T (A \Sigma_x A^T + \Sigma_w)^{-1} y$$

- If  $\Sigma_w = I$  and  $\Sigma_x \rightarrow \infty$ , this approaches to the least squares approximate solution

$$E[x|y] \rightarrow (A^T A)^{-1} A^T y$$