Ch03_Random Processes

Prof. Cheolsoo Park

Random Variable (RV)

❖ Concept and example of Random Variable

* Concept(Fig. 1)

- Random Variable X is a function to map each events or experiments results to numerical values on real number line
- ζ(zeta) : events or experiments results such as spots on a dice when throw it
- S : Sample space including all events or result of experiments

* Example(Fig. 2)

-> When throw a coin, event head can be mapped to real number 1 and event tail to real number 2

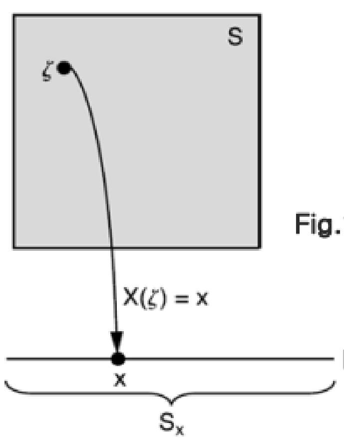


Fig.1 concept of RV

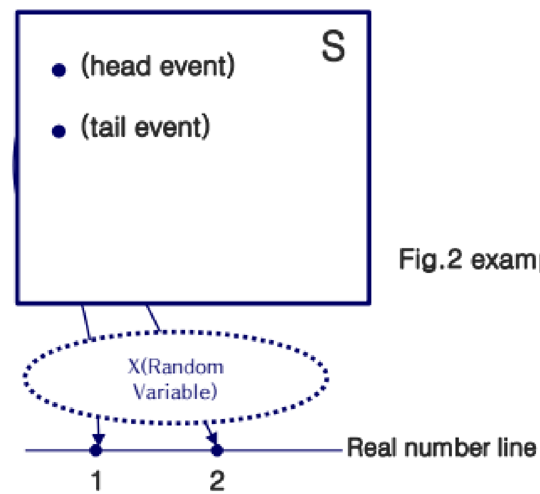


Fig.2 example of RV

Probability Distribution

❖ Probability distribution

-> probability distribution : distribution of probability values(P(X=x)) of each random variables which are mapped to real numbers

- Example 1) When we experiments throwing two coins, happening total head events can be mapped to number 0, 1 and 2 and each probabilities are $\frac{1}{4}$, $\frac{1}{2}$, $\frac{1}{4}$

x	0	1	2
$P(\mathbf{X}=x)$	1/4	1/2	1/4

- Example 2) When we experiments throwing two dices, the sum of each pips can be mapped 2,3,4, … 11, 12 and each probabilities are below

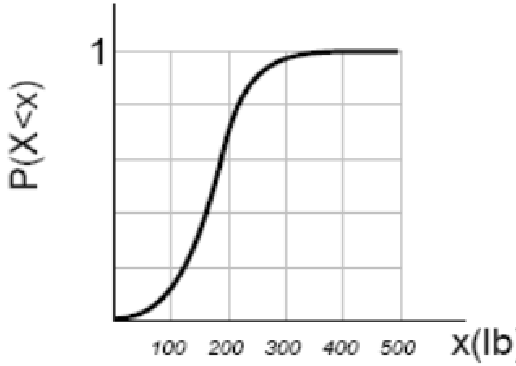
y	2	3	4	5	6	7	8	9	10	11	12
$P(\mathbf{Y}=y)$	1/36	2/36	3/36	4/36	5/36	6/36	5/36	4/36	3/36	2/36	1/36

Cumulative Distribution Function

❖ Cumulative Distribution Function(cdf)

Cumulative distribution function $F_X(x)$ is probability function which accumulates(integrals) X's probabilities from $-\infty$ to specific value x

$$F_X(x) \triangleq P[X \le x] \quad \text{for} \quad -\infty < x < +\infty$$



cdf for a person's weight

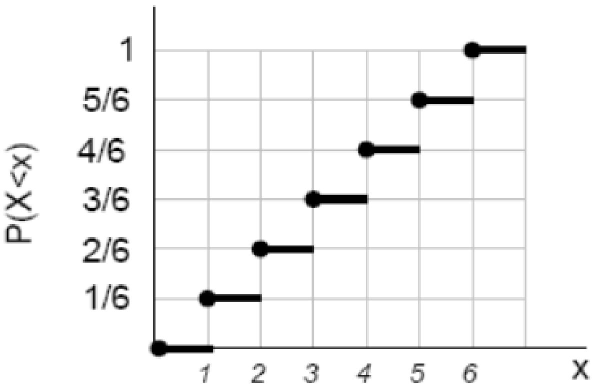
* Property of cdf

$$0 \le F_X(x) \le 1$$

$$\lim_{x\to\infty} F_X(x) = 1$$

$$\lim_{x\to-\infty} F_X(x) = 0$$

$$F_X(a) \le F_X(b) \quad \text{if } a \le b$$



cdf for rolling a dice

Probability Density Function

Probability Density Function, pdf $f_X(x)$ is differential value of continuous cdf $F_X(x)$

To discrete case, it is called **Probability Mass Function (pmf)**

$$f_X(x) = \frac{dF_X(x)}{dX}$$

<pdf>

$$f_X(x) = \frac{\Delta F_X(x)}{\Delta X}$$

<pmf>

* Property of pdf

$$f_X(x) > 0$$

$$P[a < x < b] = \int_a^b f_X(x)dx$$

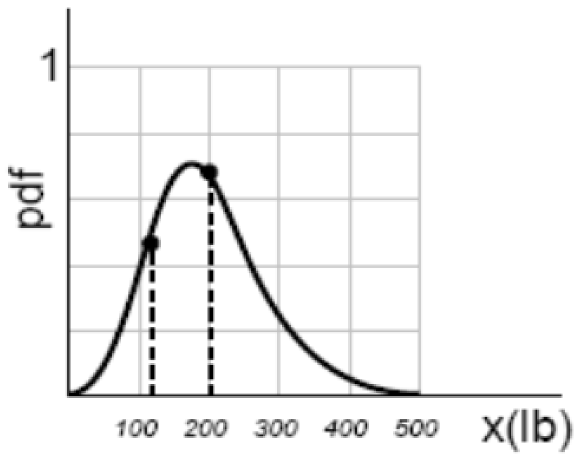
$$F_X(x) = \int_{-\infty}^{x} f_X(x)dx$$

$$1 = \int_{-\infty}^{+\infty} f_X(x)dx$$

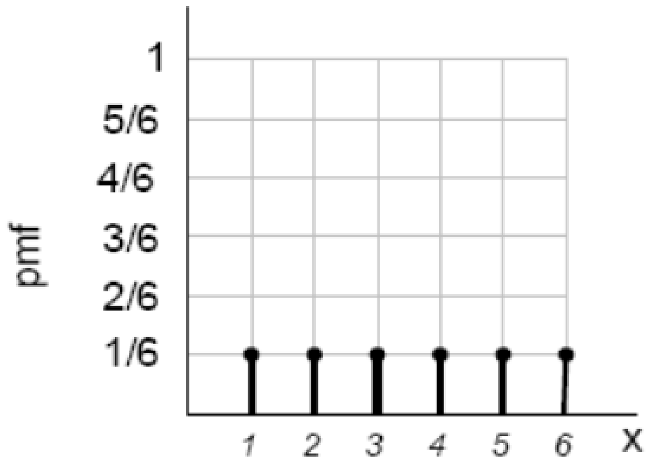
$$f_X(x \mid A) = \frac{d}{dx} F_X(x \mid A) \quad \text{where } F_X(x \mid A) = \frac{P[\{X < x\} \cap A]}{P[A]} \quad \text{if} \quad P[A] > 0$$

Probability Density Function

❖ To get probability from pdf and pmf



pdf for a person's weight



pmf for rolling a (fair) dice

To get probability, it should integral certain range

pmf itself represents probability

Average and Deviation of Random Variable

❖ Expectation : Average of random variable

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

uniform distribution

General data's average : $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ (where n is total number of data)

then, $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{1}{n}\sum_{all\ x} n_x \cdot x = \sum_{all\ x} x \cdot \frac{n_x}{n}$ (where n_x is numbers of data x, n_x/n is probability of data x)

So in continuous case, $E[X] = \int_{-\infty}^{\infty} \boldsymbol{x} \cdot \boldsymbol{f_X(x) dx}$

Average and Deviation of Random Variable

❖ Deviation process of induction is similar as expectation

$$s^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2 = \sum_{all\ x}(x - \bar{x})^2 \frac{n_x}{n}$$

where $\bar{X}$ is average of x and n_x/n is probability of x

In othere words, variation of random variable X can be represented as

$$\sigma^2 = \sum_{all\ x}(x - \mu)^2 p(x)$$

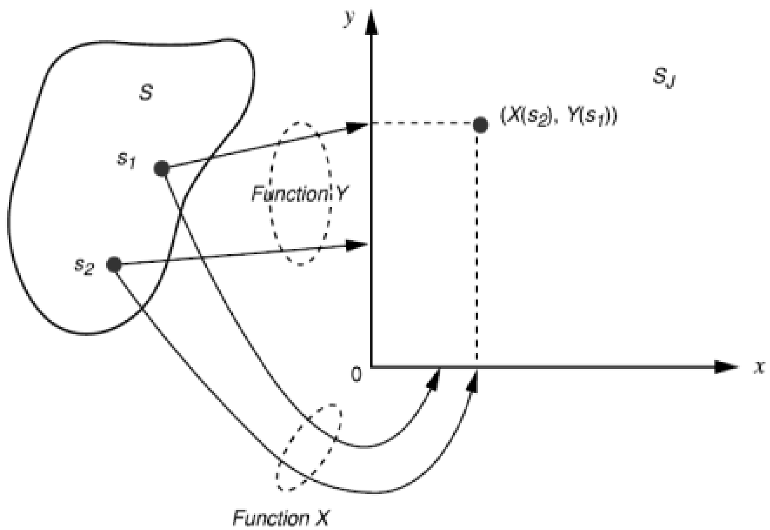
In continuous case,

$$\sigma^2 = \int_{-\infty}^{\infty}(x-\mu)^2\, p(x)\ (where\ \mu\ is\ expectation\ and\ p(x)\ is\ probablitiy\ of\ x)$$

Vector Random Variable

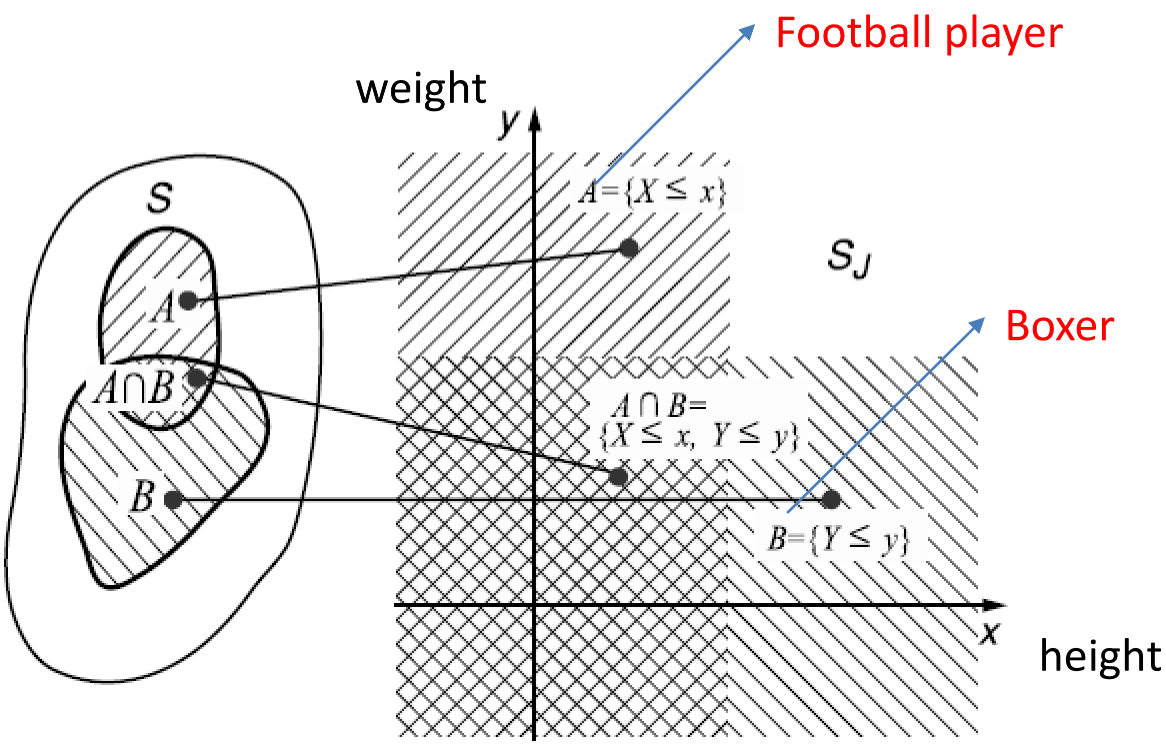
❖ Concept

- Vector random variable is to consider more than two random variable
- It can be defined as column vector
- In particular, the case of considering two random variable is called double random variable



mapping from sample space S to joint sample space S_J

Vector Random Variable



Marginal PDF

- Marginal Probability Density Function

$$f_{x_1}(x_1) = \int_{x_2=-\infty}^{x_2=+\infty} f_{x_1 x_2}(x_1 x_2) dx_2$$

Covariance Matrix

- Expectation of a Random Vector

  - $\mu = \begin{pmatrix} E(y_1) \\ E(y_2) \\ \vdots \\ E(y_n) \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_n \end{pmatrix}$
  - Average vector of Y, $E(y) = \mu$

- Random Matrix

  - $E(Y) = \begin{pmatrix} E(Y_{11}) & \cdots & E(Y_{1c}) \\ \vdots & \ddots & \vdots \\ E(Y_{r1}) & \cdots & E(Y_{rc}) \end{pmatrix} = \begin{pmatrix} \mu_{11} & \cdots & \mu_{1c} \\ \vdots & \ddots & \vdots \\ \mu_{r1} & \cdots & \mu_{rc} \end{pmatrix}$

Covariance Matrix

- Covariance Matrix

$$E\{(\boldsymbol{Y}-\boldsymbol{\mu})(\boldsymbol{Y}-\boldsymbol{\mu})'\} = \begin{pmatrix} E(Y_1-\mu_1)^2 & E(Y_1-\mu_1)(Y_2-\mu_2) & \cdots & E(Y_1-\mu_1)(Y_n-\mu_n) \\ E(Y_2-\mu_2)(Y_1-\mu_1) & E(Y_2-\mu_2)^2 & \cdots & E(Y_2-\mu_2)(Y_n-\mu_n) \\ \vdots & \vdots & \ddots & \vdots \\ E(Y_n-\mu_n)(Y_1-\mu_1) & E(Y_n-\mu_n)(Y_2-\mu_2) & \cdots & E(Y_n-\mu_n)^2 \end{pmatrix}$$

$$= \begin{pmatrix} \sigma_1^2 & \cdots & c_{1N} \\ \vdots & \vdots & \vdots \\ c_{1N} & \cdots & \sigma_N^2 \end{pmatrix}$$

- The matrix $\sum$ is called the **covariance matrix / variance-covariance matrix of Y**.

Correlation

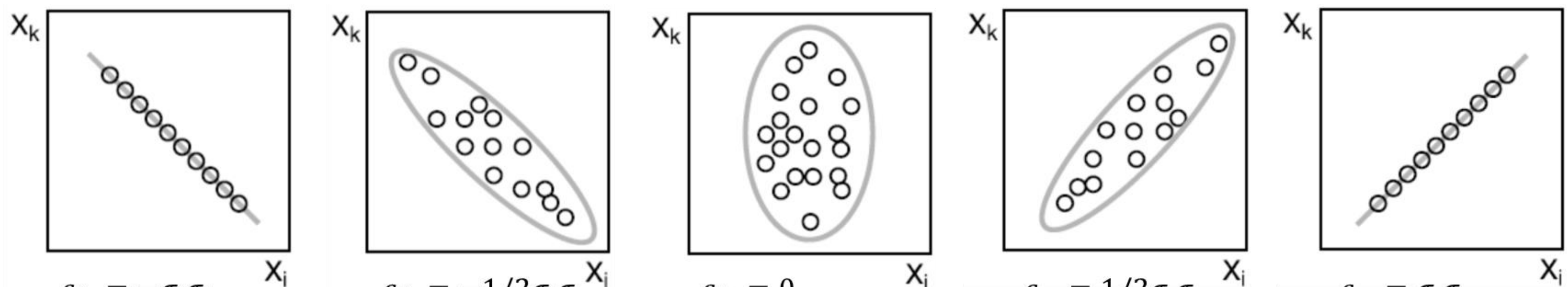
$$\sum = \Gamma R \Gamma = \begin{bmatrix} \sigma_1 & \cdots & 0 \\ 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_N \end{bmatrix} \cdot \begin{bmatrix} 1 & \cdots & \rho_{1N} \\ \rho_{12} & \cdots & \vdots \\ \vdots & \ddots & \vdots \\ \rho_{1N} & \cdots & 1 \end{bmatrix} \cdot \begin{bmatrix} \sigma_1 & \cdots & 0 \\ 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_N \end{bmatrix}$$

Covariance Matrix

- Covariance Matrix

- Characteristic of covariance

- $|c_{ik}| < \sigma_i \sigma_k$, $\sigma_i : x_i's\ standard\ deviation$
- $c_{ii} = \sigma_i^2 = VAR(x_i)$
- $c_{ik} = \rho_{ik}\sigma_i\sigma_k$

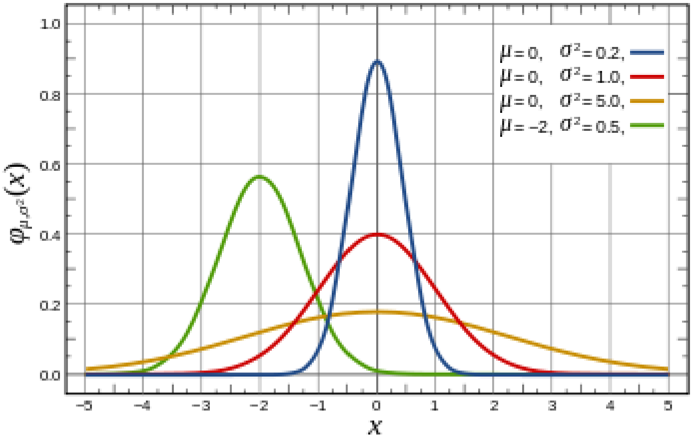


$c_{ik} = -\sigma_i\sigma_k$	$c_{ik} = -1/2\sigma_i\sigma_k$	$c_{ik} = 0$	$c_{ik} = 1/2\sigma_i\sigma_k$	$c_{ik} = \sigma_i\sigma_k$
$\rho_{ik} = -1$	$\rho_{ik} = -1/2$	$\rho_{ik} = 0$	$\rho_{ik} = 1/2$	$\rho_{ik} = 1$

Gaussian Probability Distribution

- Gaussian Probability Density Function
  - the most popular distribution to be applied in science and engineering
  - known as normal distribution
- Gaussian distribution for a single variable x,

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



Gaussian Probability Distribution

- Gaussian distribution for multi-dimensional vector **x**,

$$g_{(\mu,\Sigma)}(\mathbf{x}) = \frac{1}{\sqrt{2\pi}^{d}\sqrt{\det(\Sigma)}} e^{-\frac{1}{2}(\mathbf{x}-\mu)^{T}\Sigma^{-1}(\mathbf{x}-\mu)}$$

Gaussian Probability Distribution

- Central limit theorem
  - The sum of a set of random variables has a distribution increasing Gaussianity as the number of terms in the sum increases
  - In practice, central limit theorem says that

    the expectation of the mean of arbitrary selected samples has a normal distribution, no matter what the population is

$$\mu_N \approx Normal\left(\mu, \frac{\sigma}{\sqrt{N}}\right)$$

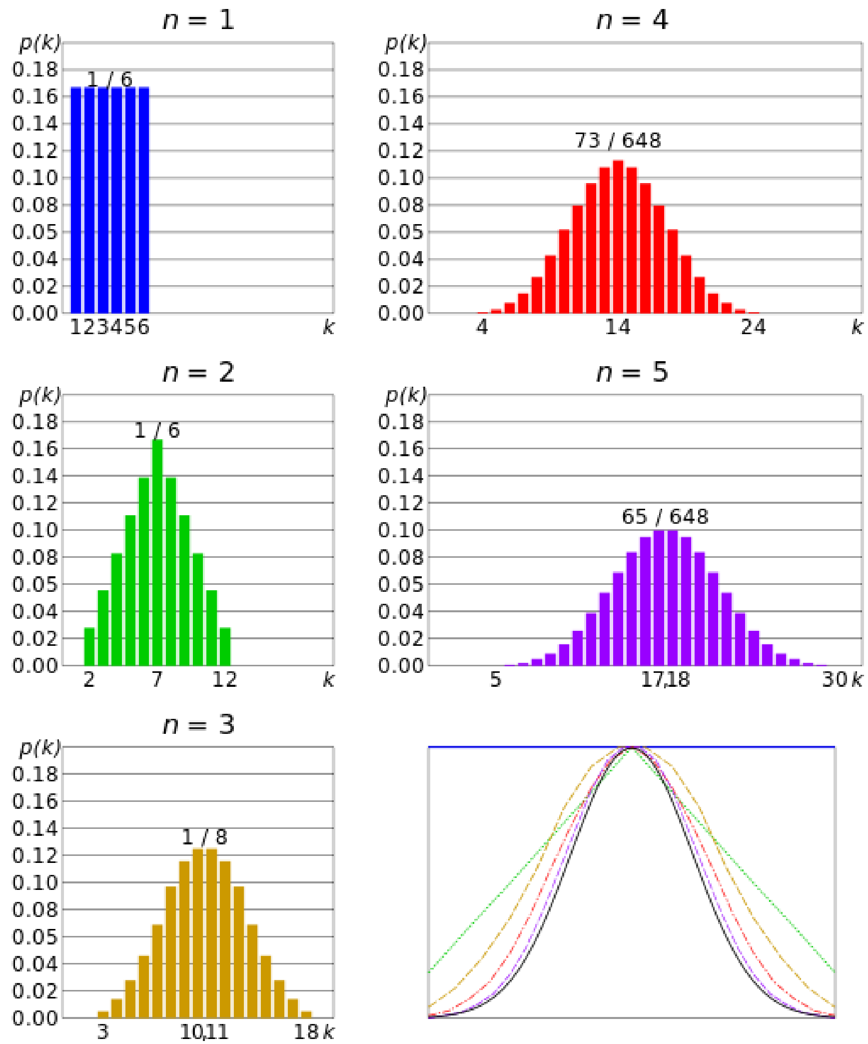
, where the μ_N is a mean of each sample, N is a size of sample and μ, σ are the properties of the population

➔ We can use the 'normal distribution' in many cases

https://www.youtube.com/watch?v=jvoxEYmQHNM

Gaussian Probability Distribution

- p(k) for the sum of n fair 6-sided dice to show their convergence to a normal distribution with increasing n, in accordance to the central limit theorem (wikipedia)
- k denotes the score obtained by n times throwing of the dice

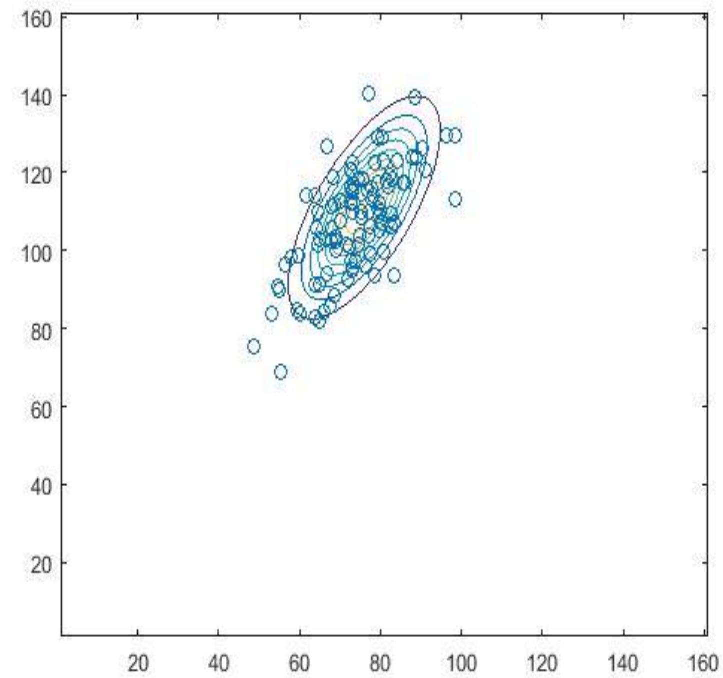


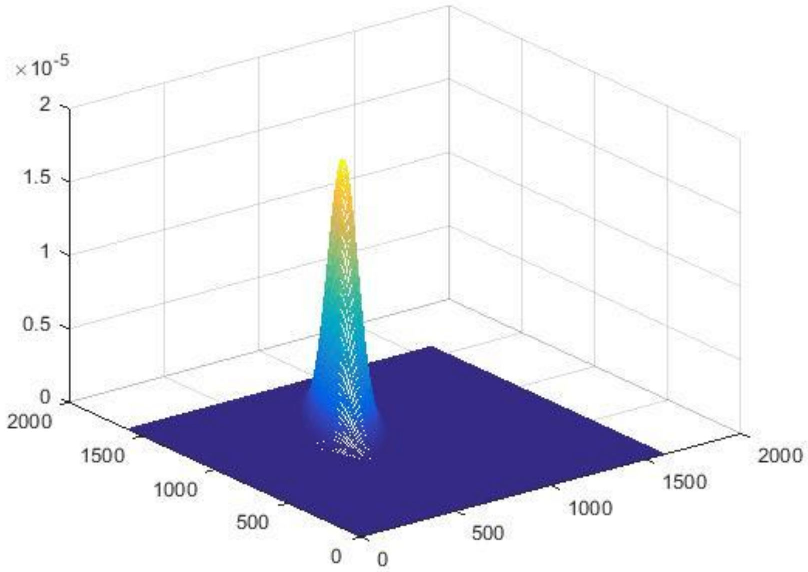
Examples of Bivariate Gaussian

mean = [750 1100],

Cov = [8000 8400;

8400 18500]



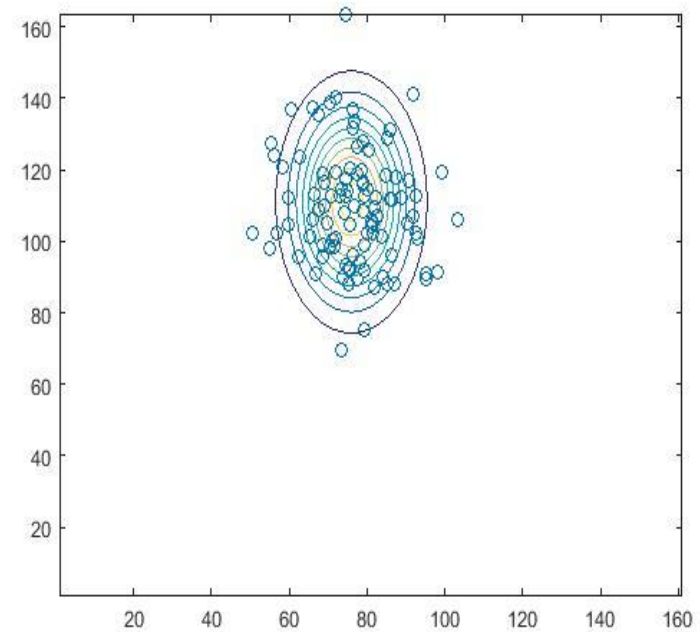


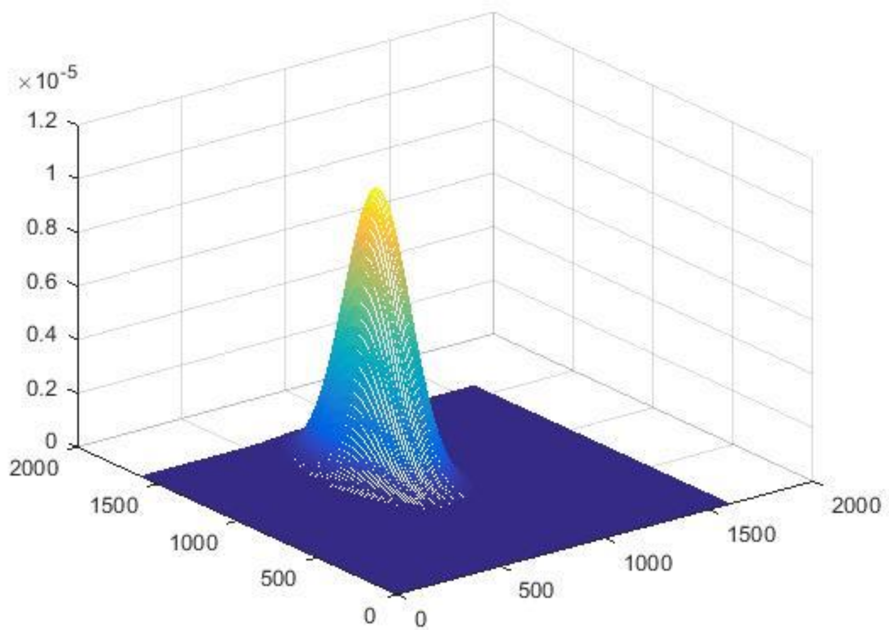
Examples of Bivariate Gaussian

mean = [750 1100],

Cov = [8000 0;

0 28500]





Examples of Bivariate Gaussian

mean = [750 1100],

Cov = [8000 0;

0 8000]