Seoul National University
Dept. of Civil and Environmental Engineering

Instructor: Junho Song
junhosong@snu.ac.kr

457.212 Statistics for Civil & Environmental Engineers
In-Class Material: Class 14
Multiple Random Variables – Part II (A&T: 3.3, Supp #2)

1. **Covariance** and **Correlation Coefficient** (Recall sample COV and sample corr. coeff.)

   (a) Covariance

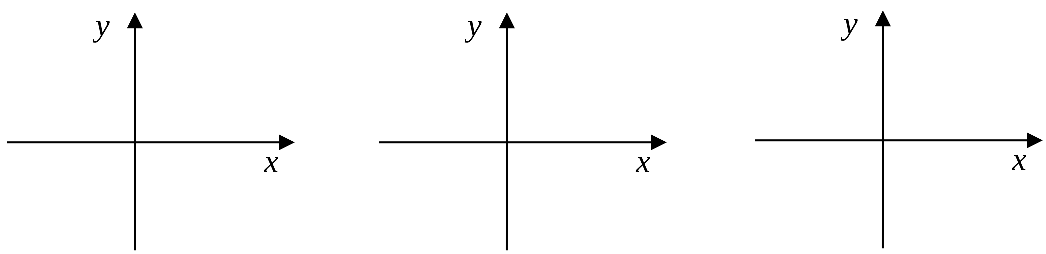
   $$\begin{aligned}\text{Cov}[X,Y] &= \text{E}[(X-\mu_X)(Y-\mu_Y)] \\ &= \text{E}[XY]-\mu_X\mu_Y\end{aligned}$$

   "mean of product minus product of means"

   (b) Covariance of discrete r.v.'s: $\text{Cov}[X,Y]=$

   (c) Covariance of continuous r.v.'s: $\text{Cov}[X,Y]=$

   (d) Sign of covariance:



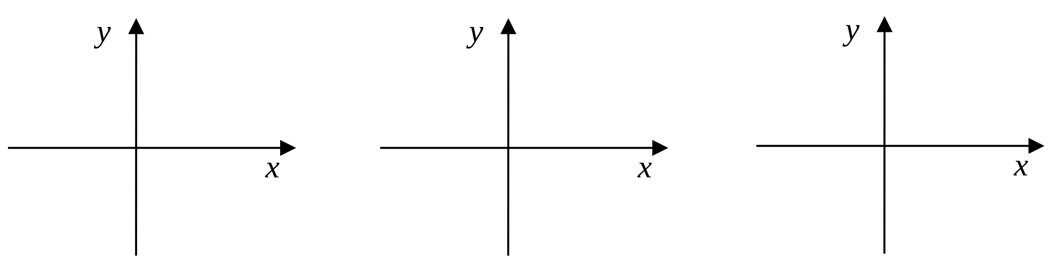
[ ]　　　　[ ]　　　　[ ]

   (e) Correlation coefficient:

   $\rho_{XY} = \overline{\qquad\qquad\qquad}$

   (f) Unlike covariance, correlation coefficient is bounded, i.e. $\quad \le \rho_{XY} \le$

   (g) Correlation coefficient can be used as a relative measure of the strength of statistical (　　　　)



[ ]　　　　[ ]　　　　[ ]

Instructor: Junho Song
junhosong@snu.ac.kr

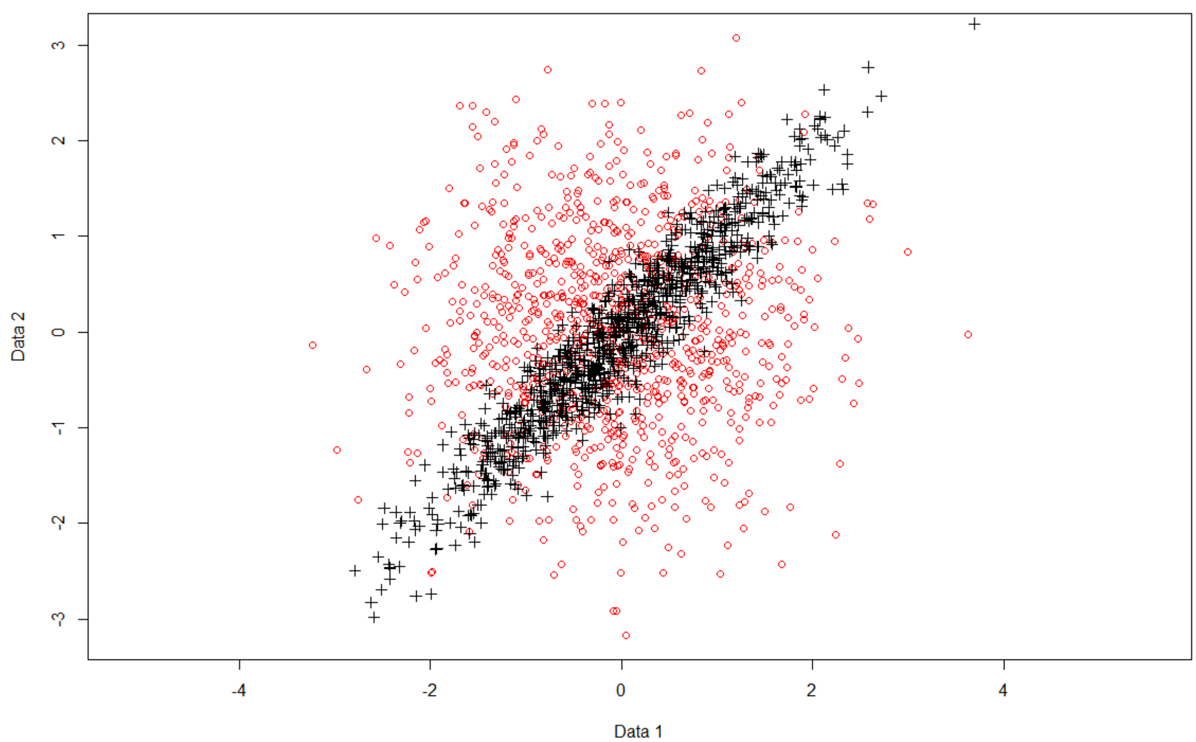
2. **Example** of joint PDF ~ two normal random variables (bi-variate joint PDF)

→ See Supplement # 2 on the course website

Example 1:

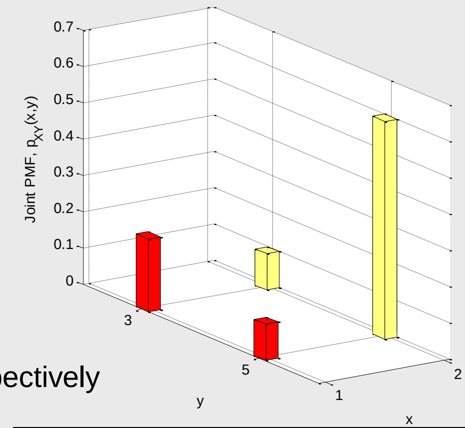
Create 1,000 pairs of bivariate standard normal random variables with correlation coefficient 0 and 0.95 respectively. Compare their scatter plots and sample correlation coefficients.

```
install.packages('mvtnorm')
library(mvtnorm)

# Generate 1000 pairs from bi-variate normal distribution (standard
   normals with zero correlation)
sample1 = rmvnorm(1000,c(0,0),matrix(c(1,0,0,1),2,2))

plot(sample1,type="p",col="red",asp=1,xlab='Data 1',ylab='Data 2')
cor(sample1) # sample correlation coefficient

# Generate 1000 pairs from binormal distribution (standard normals with
   high correlation)
sample2 = rmvnorm(1000,c(0,0),matrix(c(1,0.95,0.95,1),2,2)) # corr coeff =
   0.95

lines(sample2,type="p",col="black",pch=3) # 'lines' command overlaps a
   given plot
cor(sample2) # sample correlation coefficient
```



Instructor: Junho Song
junhosong@snu.ac.kr

Example 2: Consider two discrete random variables X and Y. Their joint PMF is given as



y \ x	1	2
3	0.2	0.1
5	0.1	0.6

(a) Determine the marginal PMF's of X and Y, respectively

x	$p_X(x)$

y	$p_Y(y)$

(b) Conditional probability mass function of X given $Y = 3$?

x	$p_{X\|Y}(x\|3)$

(c) Find the means of X and Y, respectively.

(d) Find the standard deviations of X and Y, respectively

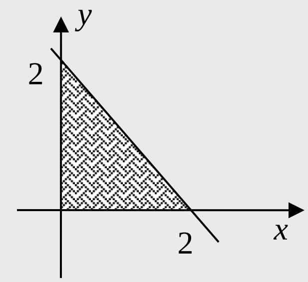
(e) Find the covariance between X and Y.

(f) Find the correlation coefficient between X and Y.

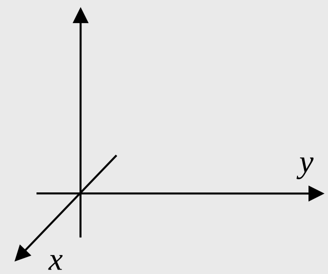
Seoul National University
Dept. of Civil and Environmental Engineering

Instructor: Junho Song
junhosong@snu.ac.kr

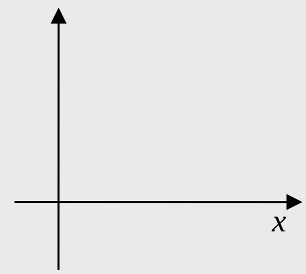
Example 3: Consider two random variables X and Y within the domain shown in the figure. It is assumed that the probability of each point in the domain is the same.



(a) Determine the joint PDF and plot.



(b) Determine the marginal PDFs of X and Y, respectively.



(c) Find the means of X and Y, respectively.

(d) Find the standard deviations of X and Y, respectively.

(e) Find the covariance of X and Y.

(f) Find the correlation coefficient between X and Y. Does the sign make sense to you? Why (not)?