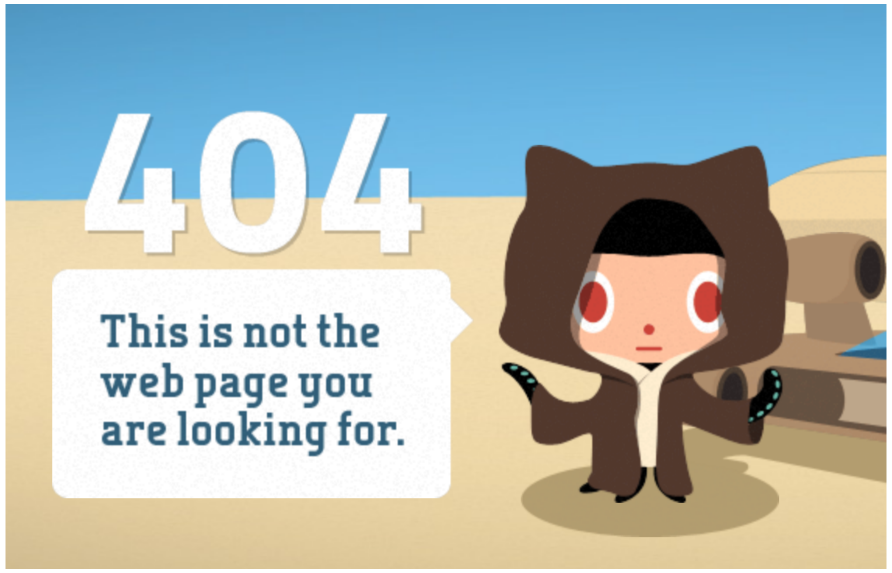
DCP 1206: Probability

Lecture 19 — Bivariate Normal Random Variables and Multivariate Random Variables

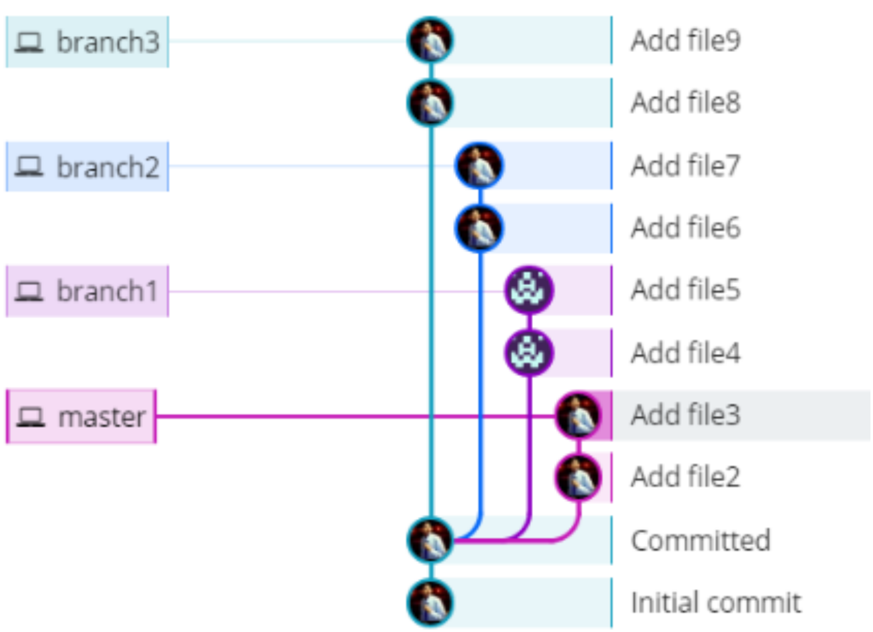
Ping-Chun Hsieh

November 22, 2019

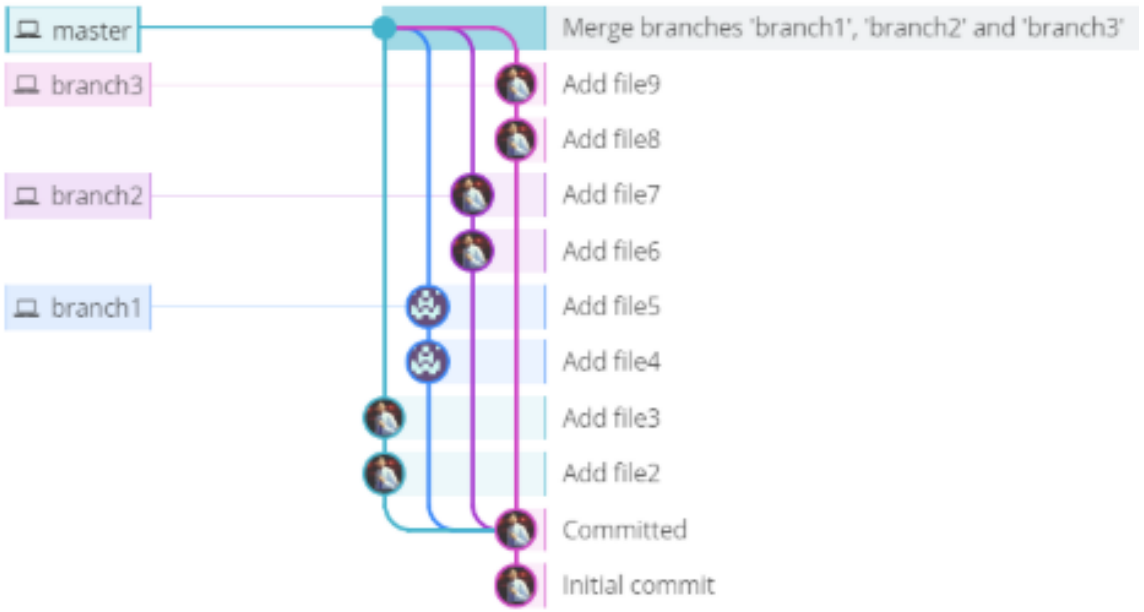
GitHub's Mascot: The Octocat



Octocat = Octopus-Cat



How to merge 3 branches simultaneously?



"Octopus-merge"

e.g.: git merge b1 b2 b3 b4

http://www.freblogg.com/2016/12/git-octopus-merge.html

This Lecture

1. Bivariate Normal Random Variables

2. Multivariate Random Variables

- Reading material: Chapter 10.5 and 9.1

1. Bivariate Normal Random Variables

Let’s do a quick review of “bivariate normal”!

Review: Bivariate Normal R.V.s (Formally)

- **Bivariate Normal**: X_1 and X_2 are said to be bivariate normal random variables if the joint PDF of X_1, X_2 is

$$f_{X_1X_2}(x_1, x_2) = \frac{1}{2\pi\sqrt{|\det(\Sigma)|}} \exp\left[-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\right]$$

where

$$\Sigma = \begin{bmatrix} \text{Cov}(X_1, X_1) & \text{Cov}(X_1, X_2) \\ \text{Cov}(X_2, X_1) & \text{Cov}(X_2, X_2) \end{bmatrix} = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}$$

$$x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \mu = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}$$

- Notation for bivariate normal: $\begin{bmatrix} X_1 \\ X_2 \end{bmatrix} \sim \mathcal{N}(\mu, \Sigma)$

Bivariate Normal R.V.s: Alternative Expression

- **Joint PDF of Bivariate Normal:**

$$f_{X_1X_2}(x_1, x_2) = \frac{1}{2\pi\sqrt{|\det(\Sigma)|}} \exp\left[-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu) \right]$$

$$\Sigma = \begin{bmatrix} \text{Cov}(X_1, X_1) & \text{Cov}(X_1, X_2) \\ \text{Cov}(X_2, X_1) & \text{Cov}(X_2, X_2) \end{bmatrix} = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}$$

$$x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \mu = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}$$

- **Alternative expression**:

$$f_{X_1X_2}(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left[-\frac{\left(\frac{(x_1-\mu_1)^2}{\sigma_1^2} - 2\rho\frac{(x_1-\mu_1)(x_2-\mu_2)}{\sigma_1\sigma_2} + \frac{(x_2-\mu_2)^2}{\sigma_2^2} \right)}{2(1-\rho^2)} \right]$$

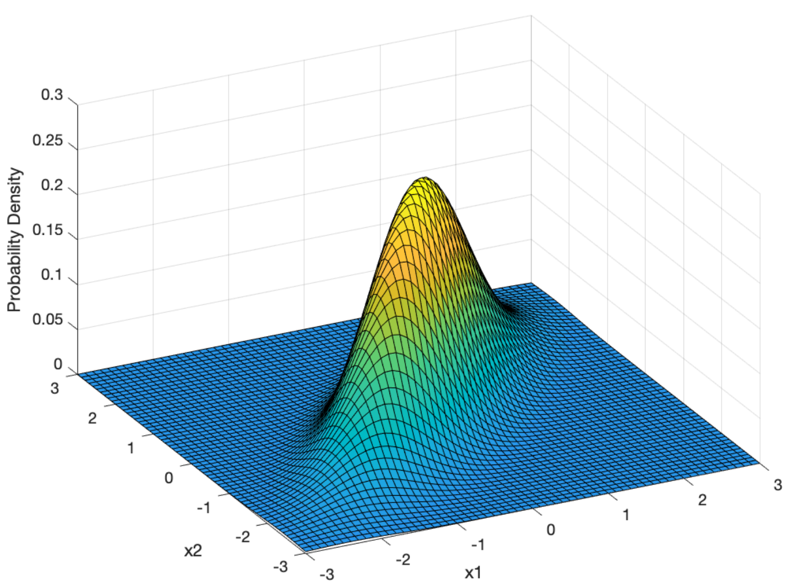
Plotting the Joint PDF Bivariate Normal

- **Joint PDF of Bivariate Normal**:

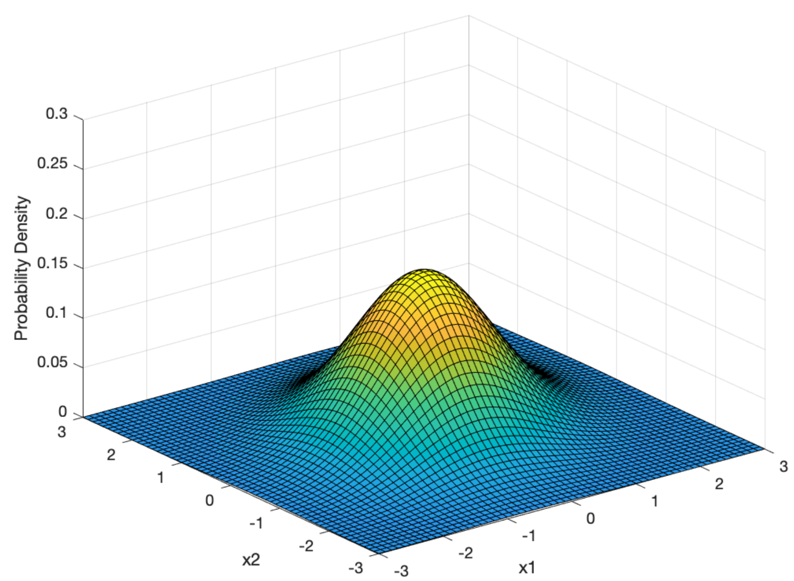
$$f_{X_1X_2}(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left[-\frac{\left(\frac{(x_1-\mu_1)^2}{\sigma_1^2} - 2\rho\frac{(x_1-\mu_1)(x_2-\mu_2)}{\sigma_1\sigma_2} + \frac{(x_2-\mu_2)^2}{\sigma_2^2} \right)}{2(1-\rho^2)} \right]$$

- **Example:** $\sigma_1 = \sigma_2 = 1, \ \mu_1 = \mu_2 = 0$

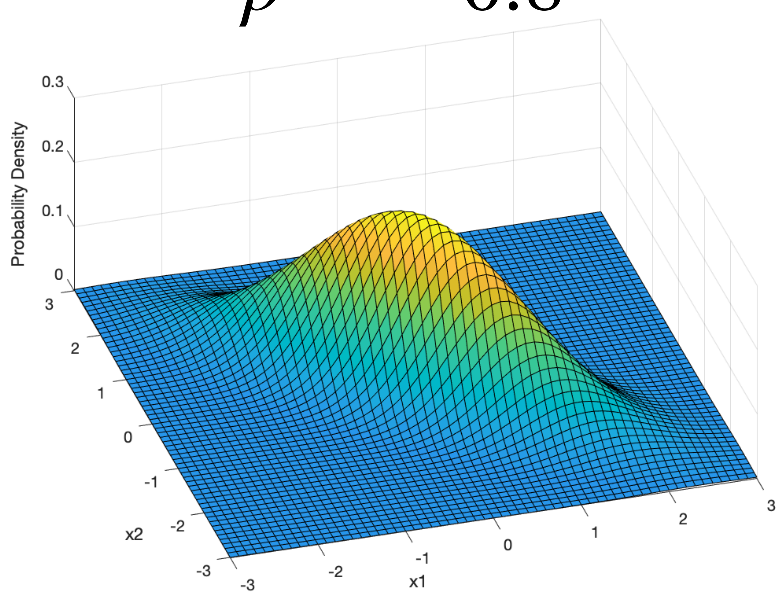
$\rho = 0.8$



$\rho = 0.2$



$\rho = -0.8$



X_1, X_2 Normal $\nRightarrow X_1, X_2$ Bivariate Normal

- Example: Let Y and Z be two independent standard normal r.v.s
  - $X_1 = |Y| \cdot \text{sign}(Z)$
  - $X_2 = Y$

- Question:
  - Are X_1 and X_2 normal?
  - Are X_1 and X_2 bivariate normal?

Now, let's study important properties of bivariate normal!

Properties of Bivariate Normal R.V.

- **If X_1, X_2 are bivariate normal, then we have**:

1. Marginal: $X_1 \sim \mathcal{N}(\mu_1, \sigma_1^2)$ and $X_2 \sim \mathcal{N}(\mu_2, \sigma_2^2)$
2. Conditional: $X_2 \,|\, X_1 = x_1 \sim \mathcal{N}\left(\mu_2 + \frac{\rho\sigma_2(x_1 - \mu_1)}{\sigma_1}, (1-\rho^2)\sigma_2^2\right)$
3. Correlation coefficient: $\rho(X_1, X_2) = \rho$
4. If X_1, X_2 are uncorrelated ($\rho = 0$), then X_1, X_2 are independent

1. Marginal: $X_1 \sim \mathcal{N}(\mu_1, \sigma_1^2)$ and $X_2 \sim \mathcal{N}(\mu_2, \sigma_2^2)$

$$f_{X_1X_2}(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left[-\frac{\left(\frac{(x_1-\mu_1)^2}{\sigma_1^2} - 2\rho\frac{(x_1-\mu_1)(x_2-\mu_2)}{\sigma_1\sigma_2} + \frac{(x_2-\mu_2)^2}{\sigma_2^2}\right)}{2(1-\rho^2)}\right]$$

$$\frac{\left(\frac{(x_1-\mu_1)^2}{\sigma_1^2} - 2\rho\frac{(x_1-\mu_1)(x_2-\mu_2)}{\sigma_1\sigma_2} + \frac{(x_2-\mu_2)^2}{\sigma_2^2}\right)}{2(1-\rho^2)} = \frac{(x_1-\mu_1)^2}{2\sigma_1^2} + \frac{1}{2\sigma_2^2}\left(\frac{(x_2-\mu_2)-\rho(x_1-\mu_1)}{\sqrt{1-\rho^2}}\right)^2$$

$$f_{X_1}(x_1) = \int_{-\infty}^{+\infty} f_{X_1X_2}(x_1, x_2)\, dx_2$$

$$= \frac{1}{\sqrt{2\pi}\,\sigma_1} \exp\left(-\frac{(x_1-\mu_1)^2}{2\sigma_1^2}\right) \cdot \underbrace{\int_{-\infty}^{+\infty} \left(\frac{1}{\sqrt{2\pi}\,\sigma_2 \cdot \sqrt{1-\rho^2}} \exp\left(-\frac{\left(\frac{(x_2-\mu_2)-\rho(x_1-\mu_1)}{\sqrt{1-\rho^2}}\right)^2}{2\sigma_2^2}\right)\right) dx_2}_{=1}$$

t

$$= \frac{1}{\sqrt{2\pi}\,\sigma_1} \exp\left(-\frac{(x_1-\mu_1)^2}{2\sigma_1^2}\right)$$

1. Marginal: $X_1 \sim \mathcal{N}(\mu_1, \sigma_1^2)$ and $X_2 \sim \mathcal{N}(\mu_2, \sigma_2^2)$

$$f_{X_1X_2}(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left[-\frac{\left(\frac{(x_1-\mu_1)^2}{\sigma_1^2} - 2\rho\frac{(x_1-\mu_1)(x_2-\mu_2)}{\sigma_1\sigma_2} + \frac{(x_2-\mu_2)^2}{\sigma_2^2}\right)}{2(1-\rho^2)} \right]$$

$$\frac{\left(\frac{(x_1-\mu_1)^2}{\sigma_1^2} - 2\rho\frac{(x_1-\mu_1)(x_2-\mu_2)}{\sigma_1\sigma_2} + \frac{(x_2-\mu_2)^2}{\sigma_2^2}\right)}{2(1-\rho^2)} = \frac{(x_2-\mu_2)^2}{2\sigma_2^2} + \frac{1}{2\sigma_1^2}\left(\frac{(x_1-\mu_1)-\rho(x_2-\mu_2)}{\sqrt{1-\rho^2}}\right)^2$$

$f_{X_2}(x_2) =$

(a) Home

Same as the last page

BJ4

2. Conditional: $X_2 \mid X_1 = x_1 \sim \mathcal{N}\left(\mu_2 + \frac{\rho\sigma_2(x_1 - \mu_1)}{\sigma_1}, (1-\rho^2)\sigma_2^2\right)$

$(1-\rho^2+\rho^2)\frac{(x_1-\mu_1)^2}{\sigma_1^2}$

$$f_{X_1X_2}(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left[-\frac{\left(\frac{(x_1-\mu_1)^2}{\sigma_1^2} - 2\rho\frac{(x_1-\mu_1)(x_2-\mu_2)}{\sigma_1\sigma_2} + \frac{(x_2-\mu_2)^2}{\sigma_2^2}\right)}{2(1-\rho^2)}\right]$$

$$f_{X_1}(x_1) = \frac{1}{\sqrt{2\pi}\sigma_1}\exp\left[-\frac{(x_1-\mu_1)^2}{2\sigma_1^2}\right]$$

$$f_{X_2|X_1}(x_2 \mid x_1) = \frac{\left(f_{X_1X_2}(x_1,x_2)\right)}{\left(f_{X_1}(x_1)\right)} \begin{array}{l} \leftarrow \text{joint} \\ \leftarrow \text{margrinal} \end{array}$$

$$= \frac{1}{\sqrt{2\pi}\cdot\sigma_2\cdot\sqrt{1-\rho^2}} \cdot \exp\left(-\frac{\frac{\rho^2(x_1-\mu_1)^2}{\sigma_1^2} - 2\rho\frac{(x_1-\mu_1)(x_2-\mu_2)}{\sigma_1\sigma_2} + \frac{(x_2-\mu_2)^2}{\sigma_2^2}}{2(1-\rho^2)}\right)$$

$$\left(\frac{\rho(x_1-\mu_1)}{\sigma_1} - \frac{x_2-\mu_2}{\sigma_2}\right)^2 = \frac{\rho^2(x_1-\mu_1)^2}{\sigma_1^2} - 2\rho\frac{(x_1-\mu_1)(x_2-\mu_2)}{\sigma_1\sigma_2} + \frac{(x_2-\mu_2)^2}{\sigma_2^2}$$

$$\sim \mathcal{N}\left(\mu_2 + \frac{\rho\sigma_2(x_1-\mu_1)}{\sigma_1}, (1-\rho^2)\sigma_2^2\right)$$

3. Correlation Coefficient: $\rho(X_1, X_2) = \rho$

$= Cov(X_1, X_2)/\sigma_1\sigma_2$

$\mathcal{N}\left(\mu_2 + \frac{\rho\sigma_2(x_1-\mu_1)}{\sigma_1}, (1-\rho^2)\sigma_2^2\right)$

$$f_{X_1X_2}(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left[-\frac{\left(\frac{(x_1-\mu_1)^2}{\sigma_1^2} - 2\rho\frac{(x_1-\mu_1)(x_2-\mu_2)}{\sigma_1\sigma_2} + \frac{(x_2-\mu_2)^2}{\sigma_2^2}\right)}{2(1-\rho^2)}\right]$$

$$\text{Cov}(X_1, X_2) = E[(X_1-\mu_1)(X_2-\mu_2)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \underbrace{(x_1-\mu_1)(x_2-\mu_2)}_{\text{function value}} \underbrace{f_{X_1X_2}(x_1, x_2)}_{\text{joint PDF}} dx_1 dx_2$$

Hint: $f_{X_2|X_1} = \frac{f_{X_1X_2}}{f_{X_1}} \Rightarrow \boxed{f_{X_1X_2} = f_{X_2|X_1} \cdot f_{X_1}}$

$$Cov(X_1, X_2) = \int\int (x_1-\mu_1)(x_2-\mu_2) \cdot f_{X_2|X_1} \cdot f_{X_1} \cdot dx_1 dx_2$$

$$= \int \{(x_1-\mu_1) \cdot f_{X_1}\} \left(\int ((x_2-\mu_2) \cdot f_{X_2|X_1}) dx_2\right) dx_1$$

$\int ((x_2-\mu_2) \cdot f_{X_2|X_1}) dx_2 = \frac{\rho\sigma_2(x_1-\mu_1)}{\sigma_1}$

$$= \frac{\rho\sigma_2}{\sigma_1} \underbrace{\int (x_1-\mu_1)^2 \cdot f_{X_1} \cdot dx_1}_{=\sigma_1^2} = \rho\sigma_1\sigma_2$$

3. Correlation Coefficient: $\rho(X_1, X_2) = \rho$ (Cont.)

(left blank intentionally for the proof)

4. Uncorrelated ($\rho = 0$) Implies Independence

$$f_{X_1X_2}(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left[-\frac{\left(\frac{(x_1-\mu_1)^2}{\sigma_1^2} - 2\rho\frac{(x_1-\mu_1)(x_2-\mu_2)}{\sigma_1\sigma_2} + \frac{(x_2-\mu_2)^2}{\sigma_2^2}\right)}{2(1-\rho^2)} \right]$$

- **If $\rho = 0$:**

$$f_{X_1X_2}(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2} \exp\left[-\frac{\left(\frac{(x_1-\mu_1)^2}{\sigma_1^2} + \left(\frac{(x_2-\mu_2)^2}{\sigma_2^2}\right)\right)}{2} \right]$$

$$= f_{X_1}(x_1) \cdot f_{X_2}(x_2).$$

Extension: Multivariate Normal R.V.

- **Multivariate Normal Random Variables:** $X_1, \cdots, X_n$ are said to be multivariate normal random variables if the joint PDF of $X_1, \cdots, X_n$ is

$$f_{X_1X_2\cdots X_n}(x_1, x_2, \cdots, x_n) = \underbrace{\frac{1}{2\pi\sqrt{|\det(\Sigma)|}} \exp\left[-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\right]}$$

where

$$\Sigma = \begin{bmatrix} \text{Cov}(X_1, X_1)\cdots & \text{Cov}(X_1, X_n) \\ \text{Cov}(X_2, X_1)\cdots & \text{Cov}(X_2, X_n) \\ \cdots & \\ \text{Cov}(X_n, X_1)\cdots & \text{Cov}(X_n, X_n) \end{bmatrix} \quad n$$

n

(i,j)-th $\Rightarrow \text{Cov}(X_i, X_j)$

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{bmatrix}, \mu = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \cdots \\ \mu_n \end{bmatrix}$$

<u>There is still one remaining question:</u>

Is it possible to construct

"<u>bivariate normal</u>" from "<u>normal</u>"?

Construction of Bivariate Normal R.V.

- Idea: Let Z, W be 2 independent standard normal r.v.s and define

$$X_1 = \sigma_1 Z + \mu_1$$

$$X_2 = \sigma_2\left(\rho Z + \sqrt{1-\rho^2}W\right) + \mu_2$$

- Result: X_1, X_2 are bivariate normal with joint PDF

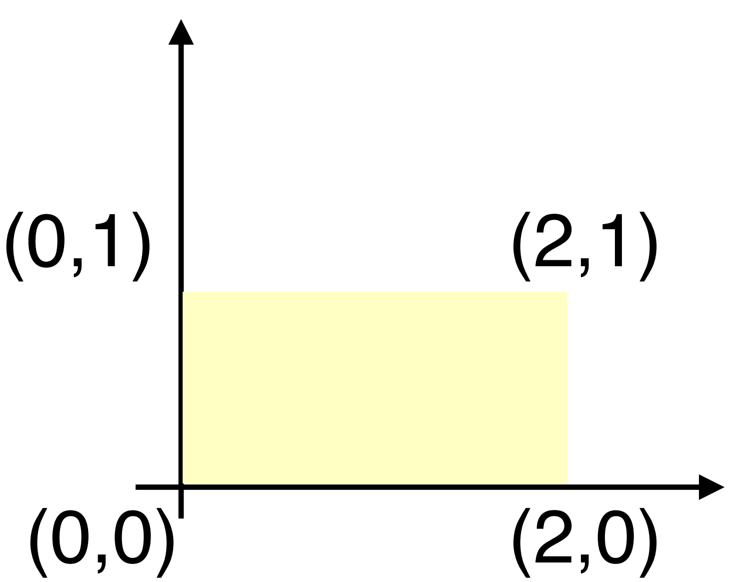
$$f_{X_1X_2}(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left[-\frac{\left(\frac{(x_1-\mu_1)^2}{\sigma_1^2} - 2\rho\frac{(x_1-\mu_1)(x_2-\mu_2)}{\sigma_1\sigma_2} + \frac{(x_2-\mu_2)^2}{\sigma_2^2}\right)}{2(1-\rho^2)}\right]$$

Linear Transformation of 2 Random Variables

- **Theorem**: Let U_1, U_2, V_1, V_2 be random variables that satisfy $V_1 = aU_1 + bU_2$ and $V_2 = cU_1 + dU_2$. Define the matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. Then, we have

$$f_{V_1V_2}(v_1, v_2) = \frac{1}{|\det(A)|} f_{U_1U_2}(A^{-1}[v_1, v_2]^T)$$

- **Intuition**:



$$A = \begin{bmatrix} 1 & -1 \\ 1 & 2 \end{bmatrix} \longrightarrow$$

For more details, please check: **https://www.stat.berkeley.edu/~aditya/resources/AllLectures2018Fall201A.pdf**

Joint PDF of X_1 and X_2

- For simplicity, assume $\mu_1 = \mu_2 = 0$ (can be handled via translation)

$$X_1 = \sigma_1 Z$$
$$X_2 = \sigma_2\left(\rho Z + \sqrt{1-\rho^2}W\right)$$

$$f_{X_1X_2}(x_1, x_2) = \frac{1}{|\det(A)|} f_{ZW}(A^{-1}[x_1, x_2]^T)$$

Quick Summary: Bivariate Normal

joint PDF (matrix form)

Marginal PDF

Conditional PDF

X_1, X_2 bivariate normal

uncorrelated vs independence?

Construction from 2 independent
standard normal r.v.s

2. Multivariate Random Variables

From Bivariate To Multivariate

- Key Idea: Bivariate definitions and properties can be directly extended to the "multivariate" cases

- For example:

  1. Joint CDF / PMF / PDF
  2. Expected value
  3. Marginal CDF / PMF / PDF
  4. Independence

Joint CDF

Joint CDF of 2 Random Variables: Let X and Y be two random variables defined on the same sample space Ω. The joint CDF $F_{XY}(t, u)$ is defined as

$$F_{XY}(t, u) = P(X \leq t, Y \leq u), \;\; \forall t, u \in \mathbb{R}$$

Joint CDF of n Random Variables: Let $X_1, \cdots, X_n$ be random variables defined on the same sample space Ω. The joint CDF $F(x_1, x_2 \cdots, x_n)$ is defined as

$$F(x_1, x_2 \cdots, x_n) = P(X_1 \leq x_1, X_2 \leq x_2, \cdots, X_n \leq x_n), \;\; \forall x_i \in \mathbb{R}$$

Joint PDF

Joint PDF of 2 Random Variables: Let X and Y be two continuous random variables. Then, $f_{XY}(x, y)$ is the joint PDF of X and Y if for every subset B of $\mathbb{R}^2$, we have

$$P((X, Y) \in B) = \iint_B f_{XY}(x, y)dxdy$$

Joint PDF of n Random Variables: Let $X_1, \cdots, X_n$ be n continuous random variables. Then, $f(x_1, x_2, \cdots, x_n)$ is the joint PDF of $X_1, \cdots, X_n$ if for every subset B of $\mathbb{R}^n$, we have

$$P((X_1, X_2, \cdots, X_n) \in B) = \int \cdots \int_B f(x_1, x_2, \cdots, x_n)dx_1 dx_2 \cdots dx_n$$

Expected Value

Expected Value of a Function of 2 Continuous RVs:
Let X, Y be 2 continuous random variables with joint PDF $f_{XY}(x, y)$. Let $g(\cdot, \cdot)$ be a function from $\mathbb{R}^2 \to \mathbb{R}$
The expected value of $g(X, Y)$ is

function value

joint PDF

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

Expected Value of a Function of n Continuous RVs:
Let $X_1, \cdots, X_n$ be n continuous random variables with joint PDF $f(x_1, x_2, \cdots, x_n)$. Let g be a function from $\mathbb{R}^n \to \mathbb{R}$. The expected value of $g(X_1, X_2, \cdots, X_n)$ is

$$E[g(X_1, X_2, \cdots, X_n)] = \int \cdots \int g(x_1, \cdots, x_n) \cdot f(x_1 \cdots x_n) dx_1 \cdots dx_n$$

Independence Property: Joint CDF

Independence $\equiv$ joint CDF is the product of the marginal CDFs:

Two random variables X, Y are **independent** if and only if

$$\underbrace{F_{XY}(t, u) = F_X(t) \cdot F_Y(u)}$$

Independence of n Random Variables $\equiv$ joint CDF is the product of the marginal CDFs:

Random variables $X_1, X_2, \cdots, X_n$ are **independent** if and only if

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

Independence Property: Joint PDF

Joint PDF is the product of the marginal PDFs under independence: Two continuous random variables X, Y are **independent** if and only if the joint PDF satisfies that

$$f_{XY}(t, u) = f_X(t) \cdot f_Y(u)$$

Joint PDF is the product of the marginal PDFs of n random variables under independence: n continuous random variables $X_1, X_2, \cdots, X_n$ are **independent** if and only if the joint PDF satisfies that

$$f_{X_1X_2\cdots X_n}(x_1, x_2, \cdots, x_n) = f_{X_1}(x_1) \cdots f_{X_n}(x_n)$$

Independence Property: Expected Value

- **Expected value under independence**: Suppose X, Y are independent random variables. Then, we have

$$\underbrace{E[g(X) \cdot h(Y)] = E[g(X)] \cdot E[h(Y)]}$$

- **Expected value of n random variables under independence**: Suppose $X_1, X_2, \cdots, X_n$ are independent random variables. Then, we have

$$E[g_1(X_1)\cdots g_n(X_n)] = E[g_1(X_1)] \cdots E[g_n(X_n)]$$

Next Lecture

1. Sum of n independent random variables

2. Moment generating function (MGF)

1-Minute Summary

1. Bivariate Normal Random Variables

- Joint PDF
- X_1, X_2 normal $\not\Rightarrow X_1, X_2$ bivariate normal
- 4 properties: marginal PDF / conditional PDF / ρ / uncorrelated
- Construction from 2 independent standard normal r.v.s

2. Multivariate Random Variables

- Joint CDF / PDF / PMF of n random variables
- Expected value regarding n random variables
- Independence properties of n random variables