Chapter 4. Multivariate Distributions

- ♣ Joint p.m.f. (p.d.f.)
- ♣ Independent Random Variables
- ♣ Covariance and Correlation Coefficient
- ♣ Expectation and Covariance Matrix
- ♣ Multivariate (Normal) Distributions
- ♣ Matlab Codes for Multivariate (Normal) Distributions
- ♣ Some Practical Examples

□ *The Joint Probability Mass Functions and p.d.f.*

- Let X and Y be two discrete random variables and let R be the corresponding space of X and Y. The joint p.m.f. of $X = x$ and $Y = y$, denoted by $f(x, y) = P(X = x, Y = y)$, has the following properties:

  (a) $0 \leq f(x, y) \leq 1$ for $(x, y) \in R$.

  (b) $\sum_{(x,y)\in R} f(x, y) = 1$,

  (c) $P(A) = \sum_{(x,y)\in A} f(x, y)$, where $A \subset R$.

- The marginal p.m.f. of X is defined as $f_X(x) = \sum_y f(x, y)$, for each $x \in R_x$.

- The marginal p.m.f. of Y is defined as $f_Y(y) = \sum_x f(x, y)$, for each $y \in R_y$.

- The random variables X and Y are independent iff (if and only if) $f(x, y) \equiv f_X(x)f_Y(y)$ for $x \in R_x$, $y \in R_y$.

Example 1. $f(x, y) = (x + y)/21$, $x = 1, 2, 3$; $y = 1, 2$, then X and Y are not independent.

Example 2. $f(x, y) = (xy^2)/30$, $x = 1, 2, 3$; $y = 1, 2$, then X and Y are independent.

The Joint Probability Density Functions

- Let X and Y be two continuous random variables and let R be the corresponding space of X and Y. The joint p.d.f. of $X = x$ and $Y = y$, denoted by $f(x, y) = P(X = x, Y = y)$, has the following properties:

  (a) $f(x, y) \geq 0$ for $-\infty < x, y < \infty$.

  (b) $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) dx dy = 1$.

  (c) $P(A) = \int \int_A f(x, y)$, where $A \subset R$.

- The marginal p.d.f. of X is defined as $f_X(x) = \int_{-\infty}^{\infty} f(x, y) dy$, for $x \in R_x$.

- The marginal p.d.f. of Y is defined as $f_Y(y) = \int_{-\infty}^{\infty} f(x, y) dx$, for $y \in R_y$.

- The random variables X and Y are independent iff (if and only if) $f(x, y) \equiv f_X(x) f_Y(y)$ for $x \in R_x$, $y \in R_y$.

Example 3. Let X and Y have the joint p.d.f.

$$f(x, y) = \frac{3}{2} x^2 (1 - |y|), \quad -1 < x < 1. \quad -1 < y < 1.$$

Let $A = \{(x, y) | 0 < x < 1,\ 0 < y < x\}$. Then

$$\begin{aligned} P(A) &= \int_0^1 \int_0^x \tfrac{3}{2} x^2 (1 - y) dy dx &&= \int_0^1 \tfrac{3}{2} x^2 \left[y - \tfrac{y^2}{2} \right]_0^x dx \\ &= \int_0^1 \tfrac{3}{2} \left[x^3 - \tfrac{x^4}{2} \right] dx &&= \tfrac{3}{2} \left[\tfrac{x^4}{4} - \tfrac{x^5}{10} \right]_0^1 = \tfrac{9}{40} \end{aligned}$$

Example 4. Let X and Y have the joint p.d.f.

$$f(x, y) = 2, \quad 0 \leq x \leq y \leq 1.$$

Thus $R = \{(x, y) | 0 \leq x \leq y \leq 1\}$. Let $A = \{(x, y) | 0 \leq x \leq \frac{1}{2},\ 0 \leq y \leq \frac{1}{2}\}$. Then

$$\begin{aligned} P(A) &= P\left(0 \leq X \leq \tfrac{1}{2},\ 0 \leq Y \leq \tfrac{1}{2}\right) &&= P\left(0 \leq X \leq Y,\ 0 \leq Y \leq \tfrac{1}{2}\right) \\ &= \int_0^{1/2} \int_0^y 2 dx dy &&= \tfrac{1}{4} \end{aligned}$$

Furthermore,

$$f_X(x) = \int_x^1 2 dy = 2(1 - x), \quad 0 \leq x \leq 1 \quad \textit{and} \quad f_Y(y) = \int_0^y 2 dx = 2y, \quad 0 \leq y \leq 1.$$

Independent Random Variables

The random variables X and Y are independent iff their joint probability function is the product of their marginal distribution functions, that is,

$$f(x,y) = f_X(x)f_Y(y), \quad \forall \;\; x, y$$

More generally, the random variables $X_1, X_2, \cdots, X_n$ are mutually independent iff their joint probability function is the product of their marginal probability (density) functions, i.e.,

$$f(x_1, x_2, \cdots, x_n) = f_{X_1}(x_1)f_{X_2}(x_2)\cdots f_{X_n}(x_n), \quad \forall \;\; x_1, x_2, \cdots, x_n$$

(1) Let X_1 and X_2 be independent Poisson random variables with respective means $\lambda_1 = 2$ and $\lambda_2 = 3$. Then

(a) $P(X_1 = 3, X_2 = 5) = P(X_1 = 3)P(X_2 = 5) = \frac{e^{-2}2^3}{3!} \times \frac{e^{-3}3^5}{5!}$.

(b) $P(X_1 + X_2 = 1) = P(X_1 = 1)P(X_2 = 0) + P(X_1 = 0)P(X_2 = 1) = \frac{e^{-2}2^1}{1!} \times \frac{e^{-3}3^0}{0!} + \frac{e^{-2}2^0}{0!} \times \frac{e^{-3}3^1}{1!}$.

(2) Let $X_1 \sim b(3, 0.8)$ and $X_2 \sim b(5, 0.7)$ be independent binomial random variables. Then

(a) $P(X_1 = 2, X_2 = 4) = P(X_1 = 2)P(X_2 = 4) = \binom{3}{2}(0.8)^2(1-0.8)^{3-2} \times \binom{5}{4}(0.7)^4(1-0.7)^{5-4}$

(b) $P(X_1 + X_2 = 7) = P(X_1 = 2)P(X_2 = 5) + P(X_1 = 3)P(X_2 = 4) = \binom{3}{2}(0.8)^2(1-0.8)^{3-2} \times \binom{5}{5}(0.7)^5(1-0.7)^{5-5} + \binom{3}{3}(0.8)^3(1-0.8)^{3-3} \times \binom{5}{4}(0.7)^4(1-0.7)^{5-4}$

(3) Let X_1 and X_2 be two independent randome variables having the same exponential distribution with p.d.f. $f(x) = 2e^{-2x}, \;\; 0 < x < \infty$. Then

(a) $E[X_1] = E[X_2] = 0.5$ and $E[(X_1 - 0.5)^2] = E[(X_2 - 0.5)^2] = 0.25$.

(b) $P(0.5 < X_1 < 1.0, \;\; 0.7 < X_2 < 1.2) = \left(\int_{0.5}^{1.0} 2e^{-2x}dx\right) \times \left(\int_{0.7}^{1.2} 2e^{-2x}dx\right)$

(c) $E[X_1(X_2 - 0.5)^2] = E[X_1]E[(X_2 - 0.5)^2] = 0.5 \times 0.25 = 0.125$.

Covariance and Correlation Coefficient

For artibrary random variables X and Y, and constants a and b, we have

$$E[aX + bY] = aE[X] + bE[Y]$$

Proof: We'll show for the continuous case, the discrete case can be similarly proved.

$$\begin{aligned}
E[aX + bY] &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(ax + by)f(x, y)dxdy \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}axf(x, y)dxdy + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}byf(x, y)dxdy \\
&= \int_{-\infty}^{\infty}ax\left[\int_{-\infty}^{\infty}f(x, y)dy\right]dx + \int_{-\infty}^{\infty}by\left[\int_{-\infty}^{\infty}f(x, y)dx\right]dy \\
&= a\int_{-\infty}^{\infty}xf_X(x)dx + b\int_{-\infty}^{\infty}yf_Y(y)dy \\
&= aE[X] + bE[Y]
\end{aligned}$$

Similarly,

$$E\left(\sum_{i=1}^{n}a_iX_i\right) = \sum_{i=1}^{n}a_iE(X_i)$$

Furthermore,

$$E[XY] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}xyf(x, y)dxdy$$

[*Example*] Let $f(x, y) = \frac{1}{3}(x + y), \quad 0 < x < 1,\ 0 < y < 2$, and $f(x, y) = 0$ elsewhere.

$$E[XY] = \int_0^1\int_0^2 xyf(x, y)dydx = \int_0^1\int_0^2 xy\frac{1}{3}(x + y)dydx = \frac{2}{3}$$

- Let X and Y be independent random variables, then

$$E(XY) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}xyf_X(x)f_Y(y)dxdy = \left[\int_{-\infty}^{\infty}xf_X(x)dx\right]\cdot\left[\int_{-\infty}^{\infty}yf_Y(y)dy\right] = E(X)\cdot E(Y)$$

- The covariance between r.v.'s X and Y is defined as

$$Cov(X, Y) = E[(X-\mu_X)(Y-\mu_Y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x-\mu_X)(y-\mu_Y)f(x, y)dydx = E(XY)-\mu_X\mu_Y$$

- If X and Y are independent r.v.s, then $Cov(X, Y) = 0$.
- The correlation coefficient is defined by $\rho(X, Y) = \frac{Cov(X,Y)}{\sigma_X\sigma_Y}$

Expectation and Covariance Matrix

Let $X_1, X_2, \ldots, X_n$ be random variables such that the expectation, variance, and covariance are defined as follows.

$$\mu_j = E(X_j), \quad \sigma_j^2 = Var(X_j) = E[(X_j - \mu_j)^2]$$

$$Cov(X_i, X_j) = E[(X_i - \mu_i)(X_j - \mu_j)] = \rho_{ij}\sigma_i\sigma_j$$

Suppose that $\mathbf{X} = [X_1, X_2, \ldots, X_n]^t$ is a random vector, then the expected mean vector and covariance matrix of $\mathbf{X}$ is defined as

$$\begin{aligned} E(\mathbf{X}) &= [\mu_1, \mu_2, \ldots, \mu_n]^t = \mu \\ Cov(\mathbf{X}) &= E[(\mathbf{X} - \mu)(\mathbf{X} - \mu)^t] \\ &= [E((X_i - \mu_i)(X_j - \mu_j))] \end{aligned}$$

Theorem 1: Let $X_1, X_2, \ldots, X_n$ be n independent r.v.'s with respective means $\{\mu_i\}$ and variances $\{\sigma_i^2\}$, then $Y = \sum_{i=1}^n a_i X_i$ has mean $\mu_Y = \sum_{i=1}^n a_i \mu_i$ and variance $\sigma_Y^2 = \sum_{i=1}^n a_i^2 \sigma_i^2$, respectively.

Theorem 2: Let $X_1, X_2, \ldots, X_n$ be n independent r.v.'s with respective moment-generating functions $\{M_i(t)\}$, $1 \leq i \leq n$, then the moment-generating function of $Y = \sum_{i=1}^n a_i X_i$ is $M_Y(t) = \prod_{i=1}^n M_i(a_i t)$.

Multivariate (Normal) Distributions

♢ (Gaussian) Normal Distribution: $X \sim N(u, \sigma^2)$

$$f_X(x) = f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp^{-(x-u)^2/2\sigma^2} \quad for - \infty < x < \infty$$

$$mean \;\; and \;\; variance: \;\; E(X) = u, \;\; Var(X) = \sigma^2$$

♢ (Gaussian) Normal Distribution: $X \sim N(\mathbf{u}, C)$

$$f_X(\mathbf{x}) = f(\mathbf{x}) = \frac{1}{(2\pi)^{d/2}[det(C)]^{1/2}} e^{-(\mathbf{x}-\mathbf{u})^t C^{-1}(\mathbf{x}-\mathbf{u})/2} \quad for \;\; \mathbf{x} \in R^d$$

$$mean \; vector \;\; and \;\; covariance \; matrix: \;\; E(X) = \mathbf{u}, \;\; Cov(X) = C$$

♢ Simulate $\mathbf{X} \sim N(\mathbf{u},\ C)$

(1) $C = LL^t$, where L is lower-Δ.

(2) Generate $\mathbf{y} \sim N(\mathbf{0}, I)$.

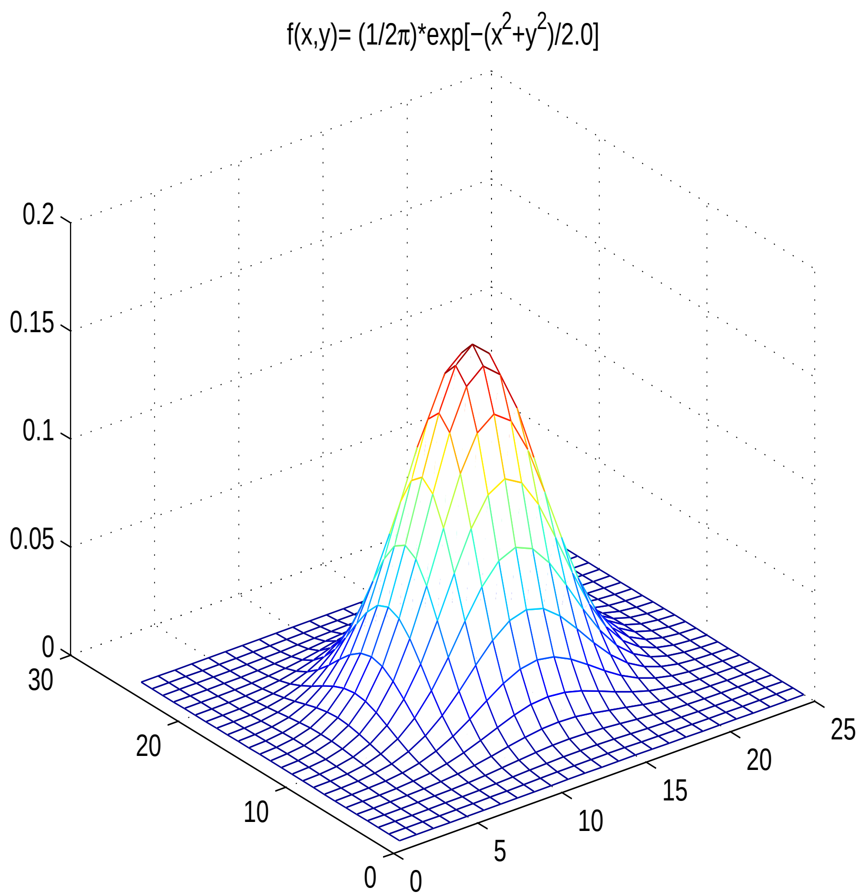
(3) $\mathbf{x} = \mathbf{u} + L * \mathbf{y}$

(4) Repeat Steps $(\mathbf{2}) and (\mathbf{3})$ M times.

```
%  Simulate N([1 3]', [4,2; 2,5])
%
n=30;
X1=random('normal',0,1,n,1);
X2=random('normal',0,1,n,1);
Y=[ones(n,1), 3*ones(n,1)]+[X1,X2]*[2 1; 0, 2];
Yhat=mean(Y)  % estimated mean vector
Chat=cov(Y)   % estimated covariance matrix
% Z=[X1, X2];
```

Plot a 2D standard Gaussian Distribution

```
x=-3.6:0.3:3.6;
y=x';
X=ones(length(y),1)*x;
Y=y*ones(1,length(x));
Z=exp(-(X.^2+Y.^2)/2+eps)/(2*pi);
mesh(Z);
title('f(x,y)= (1/2\pi)*exp[-(x^2+y^2)/2.0]')
```



Some Practical Examples

(1) Let X_1, X_2, and X_3 be independent r.v.s from a geometric distribution with p.d.f.

$$f(x) = \left(\frac{3}{4}\right)\left(\frac{1}{4}\right)^{x-1}, \quad x = 1, 2, \cdots$$

Then

(a)

$$\begin{aligned} P(X_1 = 1, X_2 = 3, X_3 = 1) &= P(X_1 = 1)P(X_2 = 3)P(X_3 = 1) &= f(1)f(3)f(1) \\ &= \left(\tfrac{3}{4}\right)^3 \left(\tfrac{1}{4}\right)^2 &= \tfrac{27}{1024} \end{aligned}$$

(b)

$$\begin{aligned} P(X_1 + X_2 + X_3 = 5) &= 3P(X_1 = 3, X_2 = 1, X_3 = 1) + 3P(X_1 = 2, X_2 = 2, X_3 = 1) \\ &= \tfrac{81}{512} \end{aligned}$$

(c) Let $Y = max\{X_1, X_2, X_3\}$, then

$$\begin{aligned} P(Y \le 2) &= P(X_1 \le 2)P(X_2 \le 2)P(X_3 \le 2) \\ &= \left(\tfrac{3}{4} + \tfrac{3}{4} \cdot \tfrac{1}{4}\right)^3 \\ &= \left(\tfrac{15}{16}\right)^3 \end{aligned}$$

(2) Let the random variables X and Y have the joint density function

$$\begin{aligned} f(x, y) &= xe^{-xy-x}, \quad x > 0,\ y > 0 \\ f(x, y) &= 0 \ \ elsewhere \end{aligned}$$

Then

(a) $f_X(x) = \int_0^\infty xe^{-xy-x}dy = e^{-x},\ x > 0;\quad \mu_X = 1,\ \sigma_X^2 = 1.$

(b) $f_Y(y) = \frac{1}{(1+y)^2},\ y > 0;\ \mu_Y = \lim_{y\to\infty}[\ln(1 + y) - 1]$ does not exist.

(c) X and Y are *not independent* since $f(x, y) \neq f_X(x)f_Y(y)$.

(d)

$$
\begin{aligned}
P(X+Y\le 1) &= \int_0^1 \left(\int_0^{1-x} xe^{-xy-x}dy\right)dx \\
&= \int_0^1 (e^{-x}-e^{-2x+x^2})dx \\
&= \int_0^1 e^{-x}dx - e^{-1}\times[\int_0^1 e^{1-2x+x^2}dx] \\
&= 1-e^{-1}-e^{-1}\times(\int_0^1 e^{t^2}dt)
\end{aligned}
$$

(3) Let (X,Y) be uniformly distributed over the *unit circle* $\{(x,y):(x^2+y^2)\le 1\}$. Its joint p.d.f is given by

$$
\begin{aligned}
f(x,y) &= \tfrac{1}{\pi},\ \ x^2+y^2\le 1 \\
f(x,y) &= 0\ \ elsewhere
\end{aligned}
$$

(a) $P(X^2+Y^2\le \frac{1}{4})=\frac{\pi}{4}\cdot\frac{1}{\pi}$.

(b) $\{(x,y):\ (x^2+y^2)\le 1,\ x>y\}$ is a semicircle, so $P(X>Y)=\frac{1}{2}$.

(c) $P(X=Y)=0$.

(d) $\{(x,y):\ (x^2+y^2)\le 1,\ x<2y\}$ is a semicircle, so $P(Y<2X)=\frac{1}{2}$.

(e) Let $R=X^2+Y^2$, then $F_R(r)=P(R\le r)=r$ if $r<1$, and $F_R(r)=1$ if $r\ge 1$.

(f) Compute $f_X(x)$ and $f_Y(y)$ and show that $Cov(X,Y)=0$ but X and Y *are not independent*.

Stochastic Process

Definition: A Bernoulli trials process is a sequence of independent and identically distributed (*iid*) Bernoulli r.v.'s $X_1, X_2, \cdots, X_n$. It is the mathematical model of n repetitions of an experiment under identical conditions, with each experiment producing only two outcomes called *success/failure*, *head/tail*, etc. Two examples are described below.

(i) Quality control: As items come off a production line, they are inspected for defects. When the *ith* item inspected is defective, we record $X_i = 1$ and write down $X_i = 0$ otherwise.

(ii) Clinical trials: Patients with a disease are given a drug. If the *ith* patient recovers, we set $X_i = 1$ and set $X_i = 0$ otherwise. are mutually independent.

A Bernoulli trials process is a sequence of independent and identically distributed (*iid*) random variables $X_1, X_2, \cdots, X_n$, where each X_i takes on only one of two values, 0 or 1. The number $p = P(X_i = 1)$ is called the probability of *success*, and the number $q = 1 - p = P(X_i = 0)$ is called the probability of *failure*. The sum $T = \sum_{i=1}^{n} X_i$ is called the number of successes in n Bernoulli trials, where $T \sim b(n, p)$ has a *binomial distribution*.

Definition: $\{X(t),\ t \geq 0\}$ is a Poisson process with intensity $\lambda > 0$ if

(i) For $s \geq 0$ and $t > 0$, the random variable $X(s + t) - X(s)$ has the Poisson distribution with parameter λt, i.e.,

$$P[X(t + s) - X(s) = k] = \frac{e^{-\lambda t}(\lambda t)^k}{k!}, \quad k = 0, 1, 2, \cdots$$

and

(ii) For any time points $0 = t_0 < t_1 < \cdots < t_n$, the random variables

$$X(t_1) - X(t_0),\ X(t_2) - X(t_1),\ \cdots,\ X(t_n) - X(t_{n-1})$$

are mutually independent.

The Poisson process is an example of a *stochastic process*, a collection of random variables indexed by the time parameter t.