Carleton University
Department of Systems and
Computer Engineering

Stochastic Processes
94.553 Fall 2000
A. H. Banihashemi

Chapter 3 Summary
(Random Variables)

1 The Notion of a Random Variable

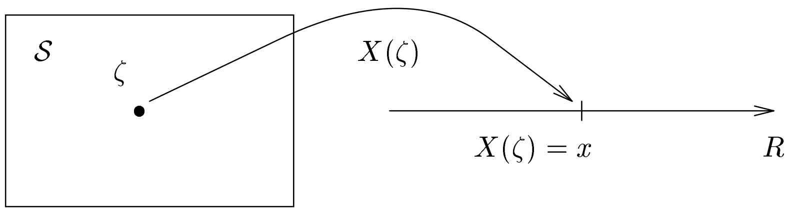


Figure 1. X maps outcomes in a sample space to numbers on the real line.

A random variable is a *function* that assigns a real number $X(\zeta)$ to each outcome ζ in the sample space of a random experiment. The range of X is denoted by S_X.

2 The Cumulative Distribution Function (cdf)

$$\boxed{F_X(x) \triangleq P[X \le x]}$$

Properties of the cdf:

1. $0 \le F_X(x) \le 1$;
2. $\lim_{x\to\infty} F_X(x) = 1$;
3. $\lim_{x\to-\infty} F_X(x) = 0$;
4. $F_X(x)$ is a nondecreasing function of x;
5. $F_X(x)$ is continuous from the right;
6. $P[a < X \le b] = F_X(b) - F_X(a)$;
7. $P[X = b] = F_X(b) - F_X(b^-) =$ height of jump discontinuity at b;
8. $P[X > x] = 1 - F_X(x)$.

Three types of random variables:

discrete: These types of random variables have a cdf that is a stair-case function, $F_X(x) = \sum_{x_k \in \mathcal{S}_\mathcal{X}} p_X(x_k)u(x - x_k)$, where $P_X(x_k) \triangleq P[X = x_k]$ is the *probability mass function* (pmf) of X and

$$u(x) = \begin{cases} 1 & x \geq 0 \\ 0 & x < 0 \end{cases}$$

is the unit step function.

continuous: $F_X(x)$ is continuous, and smooth enough so that for some function $f_X(t)$,

$$F_X(x) = \int_{-\infty}^{x} f_X(t)dt.$$

For continuous random variables, $P[X = x] = 0$ for all x, since there are no jump discontinuities in the cdf.

mixed-type: These random variables are a mixture of discrete and continuous random variables, and have a cdf of the form

$$F_X(x) = pF_D(x) + (1 - p)F_C(x)$$

where $0 < p < 1$, $F_D(x)$ is the cdf of a discrete random variable, and $F_C(x)$ is the cdf of a continuous random variable.

3 The Probability Density Function (pdf)

$$\boxed{f_X(x) \triangleq \frac{d}{dx}F_X(x)}$$

Properties of the pdf:

1. $P[x < X \leq x + \Delta x] \simeq f_X(x)\Delta x$;
2. $f_X(x) \geq 0$;
3. $P[a < X \leq b] = \int_a^b f_X(x)dx = F_X(b) - F_X(a)$;
4. $F_X(x) = \int_{-\infty}^{x} f_X(t)dt$;
5. $\int_{-\infty}^{\infty} f_X(x)dx = 1$;
6. for discrete random variables, we can define

$$f_X(x) = \sum_{x_k \in \mathcal{S}_\mathcal{X}} P_X(x_k)\delta(x - x_k) \ ,$$

where $\delta(t)$ is the delta function defined by $u(x) = \int_{-\infty}^{x} \delta(t)dt$.

Conditional cdf's and pdf's:

If A is an event with $P[A] > 0$,

$$F_X(x|A) \triangleq P[X \leq x|A] = \frac{P[\{X \leq x\} \cap A]}{P[A]}$$

and

$$f_X(x|A) \triangleq \frac{d}{dx} F_X(x|A).$$

Conditional cdf's and pdf's have all the properties of ordinary cdf's and pdf's.

4 Some Important Random Variables

A. Discrete Random Variables

1. Bernoulli

$$\mathcal{S}_X = \{0, 1\}$$
$$p_X(0) = 1 - p; p_X(1) = p$$
$$E[X] = p; \quad VAR[X] = p(1-p); \quad \Phi_X(\omega) = (1 - p + pe^{j\omega})$$

2. Binomial

$$\mathcal{S}_X = \{0, 1, 2, \ldots, n\}$$
$$p_X(k) = \binom{n}{k} p^k (1-p)^{n-k}, k = 0, 1, \ldots, n$$
$$E[X] = np; \quad VAR[X] = np(1-p); \quad \Phi_X(\omega) = (1 - p + pe^{j\omega})^n$$

3. Geometric

   (a)

$$\mathcal{S}_X = \{1, 2, 3, \ldots\}$$
$$p_X(k) = p(1-p)^{k-1}, k = 1, 2, 3, \ldots$$
$$E[X] = \frac{1}{p}; \quad VAR[X] = \frac{1-p}{p^2}; \quad \Phi_X(\omega) = \frac{pe^{j\omega}}{1 - (1-p)e^{j\omega}}$$

   (b)

$$\mathcal{S}_X = \{0, 1, 2, 3, \ldots\}$$
$$p_X(k) = p(1-p)^k, k = 0, 1, 2, 3, \ldots$$
$$E[X] = \frac{1-p}{p}; \quad VAR[X] = \frac{1-p}{p^2}; \quad \Phi_X(\omega) = \frac{p}{1 - (1-p)e^{j\omega}}$$

   This is the only discrete random variable with the *memoryless property*:

$$P[X \geq k + j | X > j] = P[X \geq k]; \quad \forall j, k.$$

4. Poisson

$$\mathcal{S}_X = \{0, 1, 2, 3, \ldots\}$$
$$p_X(k) = \frac{\alpha^k e^{-\alpha}}{k!}, k = 0, 1, 2, \ldots \text{ and } \alpha > 0$$
$$E[X] = \alpha; \quad VAR[X] = \alpha; \quad \Phi_X(\omega) = exp(\alpha(e^{j\omega} - 1))$$

B. Continuous Random Variables

1. Uniform

$$\mathcal{S}_X = [a, b]$$

$$f_X(x) = \begin{cases} \frac{1}{b-a} & a \le x \le b; \\ 0 & \text{otherwise.} \end{cases}$$

$$E[X] = \frac{a+b}{2}; \quad VAR[X] = \frac{(b-a)^2}{12}; \quad \Phi_X(\omega) = \frac{e^{j\omega b} - e^{j\omega a}}{j\omega(b-a)}$$

2. Exponential

$$\mathcal{S}_X = [0, \infty)$$

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & x \ge 0; \\ 0 & \text{otherwise.} \end{cases} \quad \text{where } \lambda > 0$$

$$E[X] = \frac{1}{\lambda}; \quad VAR[X] = \frac{1}{\lambda^2}; \quad \Phi_X(\omega) = \frac{\lambda}{\lambda - j\omega}$$

This is the only continuous random variable with the memoryless property:

$$P[X > t + h | X > t] = P[X > h]; \quad t, h \ge 0.$$

3. Gaussian (or Normal)

$$\mathcal{S}_X = (-\infty, \infty)$$

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

$$E[X] = m; \quad VAR[X] = \sigma^2; \quad \Phi_X(\omega) = e^{jm\omega - \sigma^2\omega^2/2}$$

The cdf of a Gaussian can be expressed in terms of the tabulated function $\Phi(z)$:

$$F_X(x) = \Phi(\frac{x-m}{\sigma}) \text{ , where } \Phi(z) \triangleq \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt$$

(Do not confuse $\Phi(z)$ with $\Phi_X(\omega)$; they are two different functions.) For electrical engineering applications, we are often interested in the complement, $Q(z)$, of $\Phi(z)$, defined as

$$Q(z) \triangleq 1 - \Phi(z) = \int_{z}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt.$$

5 Functions of a Random Variable

If X is a random variable and $g(x)$ is a real function, $Y = g(X)$ is a random variable.

Let $g(x)$ be a differentiable real function, and take $Y = g(X)$.

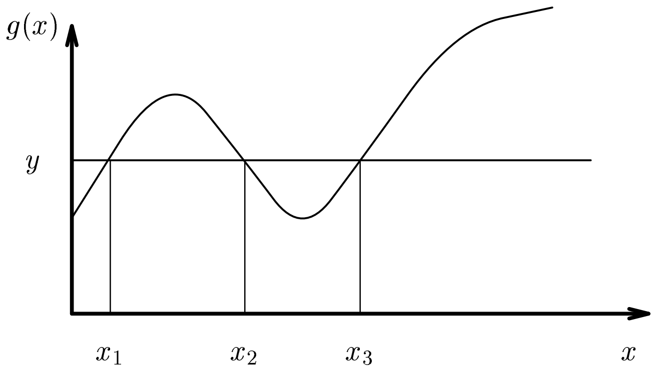


Figure 2. Example of a function.

Fix y as in Fig. 2. If the equation $g(x) = y$ has n solutions, $x_1, x_2, \ldots, x_n$, then

$$f_Y(y) = \sum_{k=1}^{n} \left[\frac{f_X(x)}{|g'(x)|} \right]_{x=x_k}$$

where $g'(x) \triangleq \frac{d}{dx} g(x)$.

6 Expected Value

The *expected value* or *mean* of a random variable X is defined as:

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

For a discrete random variable, this reduces to $\bar{x} = \sum_{x \in \mathcal{S}_X} x \, p_X(x)$.
For a function of a random variable,

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

which yields $E[g(X)] = \sum_{x \in \mathcal{S}_X} g(x) p_X(x)$, for a discrete X.
Notes:

1. $E[c] = c$ for any constant c;
2. $E[c \, g(X)] = c \, E[g(X)]$ for any constant c;
3. $E[\sum_{k=1}^{n} g_k(X)] = \sum_{k=1}^{n} E[g_k(X)]$.

Moments:
$m_n[X] \triangleq E[X^n]$ is the nth *moment* of X. We have $m_0[X] = 1$; $m_1[X] = E[X]$; $m_2[X] = E[X^2] =$ mean squared value.

$\mu_n[X] \triangleq E[(X-\bar{x})^n]$ is the nth *central moment* of X. We have $\mu_0[X]=1$; $\mu_1[X]=0$;

$$\mu_2[X] \triangleq VAR[X] = E[(X-\bar{x})^2] = \overline{x^2} - (\bar{x})^2.$$

$VAR[X]$ is the *variance* of X; $STD[X] = \sigma_X \triangleq \sqrt{VAR[X]}$ is the *standard deviation* of X.

Notes:

For any constant c:

1. $VAR[c] = 0$;
2. $VAR[X+c] = VAR[X]$;
3. $VAR[cX] = c^2 VAR[X]$.

7 The Markov and Chebyshev Inequalities

Let X be any non-negative random variable. For any $a>0$, the *Markov inequality* states:

$$P[X \geq a] \leq \frac{\bar{x}}{a}$$

We can use the Markov inequality to prove the *Chebyshev inequality*, which states that, for *any* random variable X,

$$P[\,|X-\bar{x}| \geq b\,] \leq \frac{\sigma_X^2}{b^2}.$$

8 Transform Methods

The *characteristic function* of a random variable X is

$$\begin{aligned}\Phi_X(\omega) &\triangleq E[e^{j\omega X}] \\ &= \int_{-\infty}^{\infty} f_X(x) e^{j\omega x}\,dx\ ,\end{aligned}$$

where $j=\sqrt{-1}$. The characteristic function of X and the pdf of X constitute a Fourier transform pair. The pdf can be recovered as

$$f_X(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \Phi_X(\omega) e^{-j\omega x}\,d\omega.$$

For an integer-valued discrete random variable X,

$$\Phi_X(\omega) = \sum_{k=-\infty}^{\infty} p_X(k) e^{j\omega k}$$

is a periodic function of ω with period 2π. The pmf can be recovered as

$$p_X(k) = \frac{1}{2\pi}\int_0^{2\pi} \Phi_X(\omega) e^{-j\omega k}\,d\omega, \quad k = 0, \pm 1, \pm 2, \ldots.$$

Moment Theorem:

$$E[X^n] = \frac{1}{j^n}\left[\frac{d^n}{d\omega^n}\Phi_X(\omega)\right]_{\omega=0}$$

9 Computer Generation of Random Variates

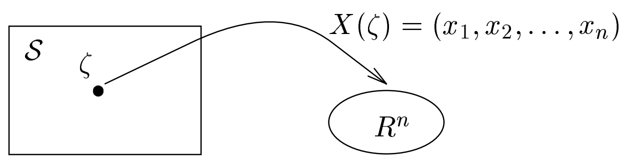
Many computer languages have a function (`rand` or `rnd` or something else) that returns a pseudo-random floating point number, uniformly distributed in $[0,1]$.

To generate a random variable with cdf $F(x)$ by the "transformation method," let U be uniformly distributed in $[0,1]$, and let $Z = F^{-1}(U)$. Then Z has the desired distribution. This method is effective when F^{-1} can be easily computed.

Chapter 4 Summary

(Multiple Random Variables)

10 Vector Random Variables



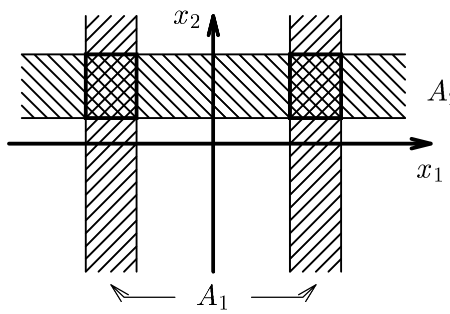
X maps outcomes in a sample space to vectors in R^n.

A vector random variable is a *function* that assigns a vector $X(\zeta) = (x_1, x_2, \ldots, x_n)$ to each outcome ζ in the sample space of a random experiment.

Product-Form Events: An event $A \subset R^n$ has product-form if

$$A = A_1(X_1) \cap A_2(X_2) \cap \cdots \cap A_n(X_n)$$

where $A_i(X_i)$ is a one-dimensional event involving the random variable X_i only. Product form events are "rectangular," involving the intersection of "strips:"



$A_1 \cap A_2$ is a product-form event.

Independence: the random variables $X_1, X_2, \ldots, X_n$ are independent if and only if, for *every* product form event $A_1(X_1) \cap \cdots \cap A_n(X_n)$, we have

$$P[A_1(X_1) \cap \cdots \cap A_n(X_n)] = P[A_1(X_1)] \times \cdots \times P[A_n(X_n)].$$

11 Pairs of Random Variables, (X, Y)

A. Discrete Random Variables

$$\left.\begin{array}{r}X \in \mathcal{S}_X = \{x_1, x_2, \ldots, x_k, \ldots\} \\ Y \in \mathcal{S}_Y = \{y_1, y_2, \ldots, y_j, \ldots\}\end{array}\right\} \text{finite or countably infinite}$$

The joint probability mass function (pmf) of X and Y is

$$p_{X,Y}(x_k, y_j) \triangleq P[X = x_k, Y = y_j].$$

Note:

$$\sum_{x_k \in \mathcal{S}_X} \sum_{y_j \in \mathcal{S}_Y} p_{X,Y}(x_k, y_j) = 1.$$

For any event A,

$$P[A] = \sum_{(x_k, y_j) \in A} p_{X,Y}(x_k, y_j).$$

Marginal pmf's:

$$\begin{aligned} p_X(x_k) \triangleq P[X = x_k] &= \sum_{y_j \in \mathcal{S}_Y} p_{X,Y}(x_k, y_j) \\ p_Y(y_j) \triangleq P[Y = y_j] &= \sum_{x_k \in \mathcal{S}_X} p_{X,Y}(x_k, y_j), \end{aligned}$$

i.e., sum over the undesired component.

The Joint Cumulative Distribution Function (cdf):

$$\boxed{F_{X,Y}(x, y) \triangleq P[X \leq x, Y \leq y]}$$

Properties of the joint cdf:

1. $0 \leq F_{X,Y}(x, y) \leq 1$;
2.

$$\begin{aligned} \lim_{x \to \infty} \lim_{y \to \infty} F_{X,Y}(x, y) &= F_{X,Y}(\infty, \infty) = 1 \\ \lim_{y \to \infty} F_{X,Y}(x, y) &= F_{X,Y}(x, \infty) = F_X(x) \\ \lim_{x \to \infty} F_{X,Y}(x, y) &= F_{X,Y}(\infty, y) = F_Y(y) \end{aligned}$$

3.

$$\begin{aligned}\lim_{x\to-\infty} F_{X,Y}(x,y) &= 0\\ \lim_{y\to-\infty} F_{X,Y}(x,y) &= 0\end{aligned}$$

4. $F_{X,Y}(x,y)$ is nondecreasing in the "northeast direction," i.e., if $x_1 \le x_2$ and $y_1 \le y_2$ then $F_{X,Y}(x_1,y_1) \le F_{X,Y}(x_2,y_2)$.

5. $F_{X,Y}(x,y)$ is continuous from the "north" and from the "east," i.e.,

$$\begin{aligned}\lim_{x\to a^+} F_{X,Y}(x,y) &= F_{X,Y}(a,y)\\ \lim_{y\to b^+} F_{X,Y}(x,y) &= F_{X,Y}(x,b)\end{aligned}$$

6. $P[x_1 < X \le x_2, y_1 < Y \le y_2] = F_{X,Y}(x_2,y_2) - F_{X,Y}(x_2,y_1) - F_{X,Y}(x_1,y_2) + F_{X,Y}(x_1,y_1)$.

B. Jointly Continuous Random Variables

The joint pdf of X and Y, if it exists, satisfies

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

Properties of the joint pdf:

1. $f_{X,Y}(x,y) \ge 0$;
2. $P[(X,Y) \in \mathbf{A}] = \int\int_{\mathbf{A}} f_{X,Y}(x,y)\,dx\,dy$; in particular,

$$\begin{aligned}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dy\,dx &= 1\\ \int_{u=-\infty}^{x}\int_{v=-\infty}^{y} f_{X,Y}(u,v)\,dv\,du &= F_{X,Y}(x,y)\\ \int_{x=a}^{b}\int_{y=c}^{d} f_{X,Y}(x,y)\,dy\,dx &= P[a < X \le b, c < Y \le d]\end{aligned}$$

3. $P[x < X \le x+dx, y < Y \le y+dy] \simeq f_{X,Y}(x,y)dxdy$;
4. $\left.\begin{array}{l} f_X(x) = \int_{y=-\infty}^{\infty} f_{X,Y}(x,y)dy \\ f_Y(y) = \int_{x=-\infty}^{\infty} f_{X,Y}(x,y)dx \end{array}\right\}$ marginal pdf's.

C. Joint Random Variables that Differ in Type

If the types of the two random variables are different, that is, e.g., X is discrete and Y is continuous, it is usually easier to work with the joint cdf, $F_{X,Y}(x,y)$, or events such as $\{X = k, Y \le y\}$.

12 Independence of Two Random Variables

X and Y are independent if and only if:

Discrete	**Either**	**Continuous**
$p_{X,Y}(x_k,y_j) = p_X(x_k)p_Y(y_j)$ for all x_k, y_j	$F_{X,Y}(x,y) = F_X(x)F_Y(y)$ for all x, y	$f_{X,Y}(x,y) = f_X(x)f_Y(y)$ for all x, y

If X and Y are independent, $g(X)$ and $h(Y)$ are independent.

13 Conditional Probability and Conditional Expectation

Discrete	**Continuous**
$p_Y(y_j\|x_k) = \frac{p_{X,Y}(x_k,y_j)}{p_X(x_k)} = F_Y(y_j\|x_k) - F_Y(y_j^-\|x_k)$ $p_X(x_k\|y_j) = \frac{p_{X,Y}(x_k,y_j)}{p_Y(y_j)} = F_X(x_k\|y_j) - F_X(x_k^-\|y_j)$	$f_Y(y\|x) = \frac{f_{X,Y}(x,y)}{f_X(x)} = \frac{d}{dy}F_Y(y\|x)$ $f_X(x\|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)} = \frac{d}{dx}F_X(x\|y)$
If X and Y are independent, $p_Y(y_j\|x_k) = p_Y(y_j)$ and $p_X(x_k\|y_j) = p_X(x_k)$.	If X and Y are independent, $f_Y(y\|x) = f_Y(y)$ and $f_X(x\|y) = f_X(x)$.
$P[Y \in A\|x_k] = \sum_{y_j \in A} p_Y(y_j\|x_k)$ $P[X \in B\|y_j] = \sum_{x_k \in B} p_X(x_k\|y_j)$	$P[Y \in A\|x] = \int_{y \in A} f_Y(y\|x)\, dy$ $P[X \in B\|y] = \int_{x \in B} f_Y(x\|y)\, dx$
$P[Y \in A] = \sum_{x_k} P[Y \in A\|x_k] p_X(x_k)$ $P[X \in B] = \sum_{y_j} P[X \in B\|y_j] p_Y(y_j)$	$P[Y \in A] = \int_{-\infty}^{\infty} P[Y \in A\|x] f_X(x)\, dx$ $P[X \in B] = \int_{-\infty}^{\infty} P[X \in B\|y] f_Y(y)\, dy$
$E[h(Y)\|x_k] = \sum_{y_j} h(y_j) p_Y(y_j\|x_k)$	$E[h(Y)\|x] = \int_{-\infty}^{\infty} h(y) f_Y(y\|x) dy$
$E[h(Y)] = E[E[h(Y)\|X]]$ $= \sum_{x_k} E[h(Y)\|x_k] p_X(x_k)$	$E[h(Y)] = E[E[h(Y)\|X]]$ $= \int_{-\infty}^{\infty} E[h(Y)\|x] f_X(x) dx$

14 Multiple Random Variables

	Discrete	**Either**	**Continuous**
Description:	Joint pmf $p_{X_1,\ldots,X_n}(x_1,\ldots,x_n) = p_{\mathbf{X}}(\mathbf{x})$	Joint cdf $F_{X_1,\ldots,X_n}(x_1,\ldots,x_n) = F_{\mathbf{X}}(\mathbf{x})$	Joint pdf $f_{X_1,\ldots,X_n}(x_1,\ldots,x_n) = f_{\mathbf{X}}(\mathbf{x})$
Properties:	Total sum = 1	$F_{X_1,\ldots,X_n}(\infty,\ldots,\infty) = 1$	Total integral = 1
$P[A]$:	Sum over A		Integrate over A
Marginals:	Sum over undesired component(s)	Set undesired component(s) to ∞	Integrate over undesired component(s)
Conditionals:	$p_{X_n}(x_n\|x_1,\ldots,x_{n-1}) = \frac{p_{X_1,\ldots,X_n}(x_1,\ldots,x_n)}{p_{X_1,\ldots,X_{n-1}}(x_1,\ldots,x_{n-1})}$	$F_{X_n}(x_n\|x_1,\ldots,x_{n-1}) = \int_{-\infty}^{x_n} f_{X_n}(u\|x_1,\ldots,x_{n-1}) du$	$f_{X_n}(x_n\|x_1,\ldots,x_{n-1}) = \frac{f_{X_1,\ldots,X_n}(x_1,\ldots,x_n)}{f_{X_1,\ldots,X_{n-1}}(x_1,\ldots,x_{n-1})}$
Independence:	$X_1,\ldots,X_n$ are independent if and only if $p_{X_1,\ldots,X_n}(x_1,\ldots,x_n) = p_{X_1}(x_1) \times \cdots \times p_{X_n}(x_n)$	$X_1,\ldots,X_n$ are independent if and only if $F_{X_1,\ldots,X_n}(x_1,\ldots,x_n) = F_{X_1}(x_1) \times \cdots \times F_{X_n}(x_n)$	$X_1,\ldots,X_n$ are independent if and only if $f_{X_1,\ldots,X_n}(x_1,\ldots,x_n) = f_{X_1}(x_1) \times \cdots \times f_{X_n}(x_n)$

15 Functions of Several Random Variables

A. One Function of Several Random Variables

Let $Y = g(X_1,\ldots,X_n)$. Then

$$\begin{aligned}
F_Y(y) &= P[Y \le y] \\
&= P[g(X_1,\ldots,X_n) \le y] \\
&= \underbrace{\int \cdots \int}_{\substack{\text{all } (x_1,\ldots,x_n) \text{ such that} \\ g(x_1,\ldots,x_n) \le y}} f_{X_1,\ldots,X_n}(x_1,\ldots,x_n)\, dx_1 \cdots dx_n \\
&= \int_{\mathbf{x}:g(\mathbf{x}) \le y} f_{\mathbf{X}}(\mathbf{x}) d\mathbf{x}\,.
\end{aligned}$$

Also, $f_Y(y) = \frac{dF_Y(y)}{dy}$.
For example, if $Y = X_1 + X_2$, then

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X_1,X_2}(x, y-x)\,dx.$$

If X_1 and X_2 are independent,

$$f_Y(y) = \underbrace{\int_{-\infty}^{\infty} f_{X_1}(x) f_{X_2}(y-x)\,dx}_{\text{convolution!}}\,.$$

As another method for finding the pdf of a function of several random variables, one can use the conditional pdf. For example, let $Z = g(X, Y)$, then

$$f_Z(z) = \int_{-\infty}^{\infty} f_Z(z|y) f_Y(y) dy = \int_{-\infty}^{\infty} f_Z(z|x) f_X(x) dx\ .$$

B. Transformations of Random Vectors

Let $\mathbf{X} = (X_1, \ldots, X_n)$ and $\mathbf{Z} = (Z_1, \ldots, Z_n) = (g_1(\mathbf{X}), \ldots, g_n(\mathbf{X})) = \mathbf{g}(\mathbf{X})$. Then

$$F_{\mathbf{Z}}(\mathbf{z}) = \int_{\mathbf{x}: g_k(\mathbf{x}) \le z_k, \forall k} f_{\mathbf{X}}(\mathbf{x}) d\mathbf{x}\ .$$

Pdf of Linear Transformations

Let $\mathbf{Z} = A\mathbf{X}$. Then $f_{\mathbf{Z}}(\mathbf{z}) = \frac{f_{\mathbf{X}}(A^{-1}\mathbf{z})}{|\det(A)|}$.

Pdf of General Transformations

Let the set of equations $\mathbf{z} = \mathbf{g}(\mathbf{x})$ have a unique solution given by $\mathbf{x} = (x_1, \ldots, x_n) = (h_1(\mathbf{z}), \ldots, h_n(\mathbf{z})) = \mathbf{h}(\mathbf{z})$. Then

$$f_{\mathbf{Z}}(\mathbf{z}) = \frac{f_{\mathbf{X}}(\mathbf{h}(\mathbf{z}))}{|J(\mathbf{x})|}\ , \tag{1}$$

or equivalently

$$f_{\mathbf{Z}}(\mathbf{z}) = f_{\mathbf{X}}(\mathbf{h}(\mathbf{z})) |J(\mathbf{z})|\ ,$$

where

$$J(\mathbf{x}) = \det \begin{pmatrix} \frac{\partial g_1}{\partial x_1} & \cdots & \frac{\partial g_1}{\partial x_n} \\ \cdot & \cdots & \cdot \\ \frac{\partial g_n}{\partial x_1} & \cdots & \frac{\partial g_n}{\partial x_n} \end{pmatrix} \quad \text{and} \quad J(\mathbf{z}) = \det \begin{pmatrix} \frac{\partial h_1}{\partial z_1} & \cdots & \frac{\partial h_1}{\partial z_n} \\ \cdot & \cdots & \cdot \\ \frac{\partial h_n}{\partial z_1} & \cdots & \frac{\partial h_n}{\partial z_n} \end{pmatrix}$$

are the *Jacobians* of the transformation and its inverse, respectively.

If the equation $\mathbf{z} = \mathbf{g}(\mathbf{x})$ has more than one solution, the pdf is equal to the sum of terms of the form (1), with each solution providing one such term.

16 Expected Value of Functions of Random Variables

$$Y = g(X_1, \ldots, X_n)$$

Discrete: $E[Y] = \sum_{x_1} \cdots \sum_{x_n} g(x_1, \ldots, x_n) p_{X_1,\ldots,X_n}(x_1, \ldots, x_n)$
Continuous: $E[Y] = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} g(x_1, \ldots, x_n) f_{X_1,\ldots,X_n}(x_1, \ldots, x_n) d\mathbf{x}$

- $E[\sum_{i=1}^{n} X_i] = \sum_{i=1}^{n} E[X_i]$ (independence not required).
- If $X_1, \ldots, X_n$ are independent **and** $g(X_1, \ldots, X_n)$ is separable, i.e.,

$$g(X_1, \ldots, X_n) = g_1(X_1) g_2(X_2) \cdots g_n(X_n),$$

  then

$$E[g(X_1, \ldots, X_n)] = E[g_1(X_1)] E[g_2(X_2)] \cdots E[g_n(X_n)].$$

Moments, Correlation, Covariance:

- The j,kth joint moment of X and Y is

$$m_{j,k}[X,Y] = E[X^j Y^k].$$

- The j,kth joint central moment of X and Y is

$$\mu_{j,k}[X,Y] = E\left[(X - E[X])^j (Y - E[Y])^k\right].$$

- $m_{1,1}[X,Y] = E[XY]$ is the *correlation* of X and Y.
- If $E[XY] = E[X]E[Y]$, then X and Y are said to be *uncorrelated.*
- If $E[XY] = 0$, then X and Y are said to be *orthogonal.*
- $\mu_{1,1}[X,Y] = E\left[(X - E[X])(Y - E[Y])\right] \triangleq COV[X,Y] = \sigma_{XY}$ is the *covariance* of X and Y.
- $COV[X,Y] = E[XY] - E[X]E[Y]$ (=0 if X and Y are uncorrelated.)
- The *correlation coefficient* $\rho_{X,Y}$ is

$$\rho_{X,Y} \triangleq \frac{COV[X,Y]}{\sigma_X \sigma_Y} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}.$$

  It can be shown that $-1 \leq \rho_{X,Y} \leq 1$, and the bounds are achieved if X and Y are related linearly.

Facts:

- If X and Y are independent, then X and Y are uncorrelated.
- If X and Y are uncorrelated, then X and Y may or may not be independent. However if X and Y are jointly Gaussian, and X and Y are uncorrelated, then X and Y are independent.

17 Jointly Gaussian Random Variables

A. Two Random Variables

Let X and Y be random variables with

$$\begin{array}{ll} E[X] = m_X; & VAR[X] = \sigma_X^2; \\ E[Y] = m_Y; & VAR[Y] = \sigma_Y^2; \end{array} \quad \frac{COV[X,Y]}{\sigma_X \sigma_Y} = \rho.$$

Then X and Y are said to be *jointly Gaussian* if and only if

$$f_{X,Y}(x,y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} \times \exp\left\{ \frac{-1}{2(1-\rho^2)} \left[\frac{(x-m_X)^2}{\sigma_X^2} - \frac{2\rho(x-m_X)(y-m_Y)}{\sigma_X\sigma_Y} + \frac{(y-m_Y)^2}{\sigma_Y^2} \right] \right\}.$$

Properties:

- X and Y are Gaussian random variables with means m_X and m_Y, and variances σ_X^2 and σ_Y^2, respectively. (Marginal distributions are Gaussian.)
- If $\rho = 0$, X and Y are independent.
- The conditional pdf's $f_X(x|y)$ and $f_Y(y|x)$ are also Gaussian, e.g., $f_X(x|y)$ is Gaussian with mean $m_X + \rho(\sigma_X/\sigma_Y)(y - m_Y)$ and variance $\sigma_X^2(1-\rho^2)$.

B. n Random Variables

The random variables $X_1, \ldots, X_n$ are called jointly Gaussian if their joint pdf is given by

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{\exp\{-\frac{1}{2}(\mathbf{x}-\mathbf{m})^T K^{-1}(\mathbf{x}-\mathbf{m})\}}{\sqrt{(2\pi)^n \det(K)}},$$

where $\mathbf{x}$ and $\mathbf{m} = (E[X_1], \ldots, E[X_n])^T$ are column vectors, and K is the *covariance matrix* defined by

$$K = \begin{pmatrix} \sigma_{X_1}^2 & \sigma_{X_1X_2} & \cdots & \sigma_{X_1X_n} \\ \sigma_{X_2X_1} & \sigma_{X_2}^2 & \cdots & \sigma_{X_2X_n} \\ \cdot & \cdot & \cdots & \cdot \\ \sigma_{X_nX_1} & \cdots & & \sigma_{X_n}^2 \end{pmatrix}.$$

Properties:

- The pdf is completely specified by the individual means and variances and pairwise covariances.
- The marginal distributions are all (jointly) Guassian.
- The linear transformation of a set of jointly Guassian random variables results in another set of jointly Guassian random variables. In particular, the linear combination of a set of jointly Gaussian random variables is Gaussian.
- All the conditional distributions are also (jointly) Guassian.