Chapter 3

General Random Variables

Peng-Hua Wang

Graduate Institute of Communication Engineering

National Taipei University
Chapter Contents

3.1 Continuous Random Variables and PDFs
3.2 Cumulative Distribution Functions
3.3 Normal Random Variables
3.4 Joint PDFs of Multiple Random Variables
3.5 Conditioning
3.6 The Continuous Bayes’ Rule
3.7 Summary and Discussion
3.1 Continuous Random Variables and PDFs
Let $X$ be the arrival time of a bus. $X$ is a continuous random variable taking value between $p$ and $q$.

The sample space $\Omega = [p, q]$. For any point $c \in \Omega$, $P(X = a) = 0$.

We can find $P(x \leq a)$, $P(a < X \leq b)$, or $P(a < X \leq a + \delta)$. In general, we can find $P(X \in B)$ for $B \subset \Omega$.

Let $F_X(x) = P(X \leq x)$. We know that

$$F_X(-\infty) = P(X \leq -\infty) = 0, \quad F_X(\infty) = P(X \leq \infty) = 1$$
Let

\[ f_X(x) \triangleq \lim_{\delta \to 0} \frac{F_X(x + \delta) - F_X(x)}{\delta} = \frac{dF_X(x)}{dx} \]

We have

\[ P(X \leq x) = F_X(x) = \int_{-\infty}^{x} f_X(t)dt \]

Therefore, \( P(X \in B) \) can be evaluated in terms of integral of \( f_X(x) \). For example,

\[ P(a < X \leq b) = \int_{a}^{b} f_X(x)dx \]
**PDF**

- $f_X(x)$ is called the probability density function (PDF) of continuous random variable $X$.
- $F_X(x) = P(X \leq x)$ is called the cumulative distribution function (CDF).
  - $f_X(x) \geq 0$
  - $\int_{-\infty}^{\infty} f_X(x)dx = P(-\infty < X \leq \infty) = 1$
  - $P(x < X < x + \delta) \approx f_X(x) \cdot \delta$ if $\delta$ is small.
Example 3.1. Uniform distribution.

\[ f_X(x) = \begin{cases} 
  c, & a \leq x \leq b \\
  0, & \text{otherwise}.
\end{cases} \]

Find \( c \).
Example 3.3. Uniform distribution.

\[ f_X(x) = \frac{c}{\sqrt{x}}, \quad 0 < x \leq 1 \]

Find \( c \).
- A PDF can take arbitrarily large values.
Expectation

- The mean or expectation of a continuous random variable $X$ is defined by $E[X] = \int_{-\infty}^{\infty} xf_X(x)dx$.
- The $k$th moment is $E[X^k] = \int_{-\infty}^{\infty} x^k f_X(x)dx$.
- The variance of $X$ is
  \[
  \text{Var}(X) = E[(X - E[X])^2] = E[X^2] - (E[X])^2
  \]
- The mean of new RV $Y = g(X)$ is
  \[
  E[g(X)] = \int_{-\infty}^{\infty} g(x) f_X(x)dx.
  \]
Expectation

The expectation is well-defined if

\[ E[|X|] = \int_{-\infty}^{\infty} |x| f_X(x) \, dx < \infty. \]

A not-well-defined random variable: Cauchy RV. Its PDF is

\[ f_X(x) = \frac{c}{1 + x^2}, \quad -\infty < x < \infty. \]

(Please find \( c \)). \( E[X] \) is not well-defined.
Example 3.4. Uniformly RV

\[ f_X(x) = \frac{1}{b-a}, \quad a \leq x \leq b. \]

Find \( E[X] \) and \( \text{Var}(X) \).
Exponential RV

\[ f_X(x) = ke^{-\lambda x}, \quad 0 \leq x < \infty. \]

Find \( k \), \( E[X] \) and \( \text{Var}(X) \).
3.2 Cumulative Distribution Functions
Cumulative Distribution Functions

- The cumulative distribution function (CDF) of a rv $X$, denoted by $F_X(x)$, is defined by

$$F_X(x) \triangleq P(X \leq x) = \begin{cases} \sum_{k \leq x} p_X(k), & \text{if } X \text{ is discrete}, \\ \int_{-\infty}^{x} f_X(x) \, dx, & \text{if } X \text{ is continuous.} \end{cases}$$

- CDF is exactly probability. PDF is NOT probability.

- Note the “$\leq$” in the definition.

- “Any random variable associated with a given probability model has a CDF, regardless of whether it is discrete or continuous.”
Fig 3.6

CDFs of some discrete random variables
CDFs of some continuous random variables

Fig 3.7
Properties of a CDF

- **Definition:** $F_X(x) \triangleq P(X \leq x)$
- **Monotonically nondecreasing:** If $x \leq y$, then $F_X(x) \leq F_X(y)$.
- $F_X(-\infty) = 0, F_X(+\infty) = 1$
- If $X$ is discrete, $F_X(x)$ is piecewise constant. If $X$ is continuous, $F_X(x)$ is continuous.
- If $X$ is discrete,
  \[ p_X(k) = F_X(k) - F_X(k - 1) \]
- If $X$ is continuous,
  \[ f_X(x) = \frac{d}{dx} F_X(x) \]
Example 3.6

- Let $X_1$, $X_2$ and $X_3$ be 3 independent discrete random variables with identical PMFs.

  \[ X = \max \{X_1, X_2, X_3\} \]

  Find PMF of $X$.

- Let $X_1$, $X_2$ and $X_3$ be 3 independent continuous random variables with identical PDFs.

  \[ X = \max \{X_1, X_2, X_3\} \]

  Find PDF of $X$.  

3.3 Normal Random Variables
Definition

A continuous random variable $X$ is said to be normal or Gaussian if it has a PDF of the form

$$f_X(x) = ce^{-(x-\mu)^2/(2\sigma^2)}, -\infty < x < \infty.$$ 

- $c = \frac{1}{\sqrt{2\pi\sigma}}$
- $E[X] = \mu$
- $\text{Var}(X) = \sigma^2$
A standard normal rv $Y$ is the normal rv with $\mu = 0$ and $\sigma = 1$. Its CDF is denoted by $\Phi(y)$

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

$\Phi(-y) = 1 - \Phi(y)$ because the PDF of standard normal is even.

For normal rv $X$ with mean $\mu$ and variance $\sigma^2$, we know that $X = \sigma Y + \mu$. Thus,

$$
P(X \leq x) = P(\sigma Y + \mu \leq x) = P \left( Y \leq \frac{x - \mu}{\sigma} \right) = \Phi \left( \frac{x - \mu}{\sigma} \right)
$$
Example 3.7

The annual snowfall at a particular geographic location is modeled as a normal random variable with a mean of \( \mu = 60 \) inches and a standard deviation of \( \sigma = 20 \). What is the probability that this year’s snowfall will be at least 80 inches?

\[
Y = 20X + 60
\]

\[
P(Y \geq 80) = ?
\]
Example 3.8

A binary message is transmitted as a signal $s$, which is either $-l$ or $+1$. The communication channel corrupts the transmission with additive normal noise $N$ with mean $\mu = 0$ and variance $\sigma^2$. The receiver concludes that the signal $-1$ (or $+1$) was transmitted if the value received is $< 0$ (or $\geq 0$, respectively). What is the probability of error?
3.4 Joint PDFs Of Multiple Random Variables
Definitions

- $X$ and $Y$ are two continuous random variables. Their joint CDF is

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

- Joint PDF is

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

1. \( P((X,Y) \in B) = \iint_{(x,y) \in B} f_{X,Y}(x,y) \, dx \, dy \)

2. \( P(a < X \leq b, c < Y \leq d) = \int_x^b \int_y^d f_{X,Y}(x,y) \, dy \, dx \)

3. \( F_{X,Y}(x,y) = \int_{s=-\infty}^x \int_{t=-\infty}^y f_{X,Y}(s,t) \, dt \, ds \)

4. \( \int_{x=-\infty}^\infty \int_{y=-\infty}^\infty f_{X,Y}(x,y) \, dy \, dx = 1 \)

5. \( P(x < X \leq x + \delta, y < Y \leq y + \text{delta}) \approx f_{X,Y}(x,y) \delta^2 \)
Marginal PDF

\[ F_X(x) = P(X \leq x) = P(X \leq x, -\infty < Y < \infty) \]
\[ = \int_{s=-\infty}^{x} \int_{y=-\infty}^{\infty} f_{X,Y}(s,y) \, dy \, ds \]
\[ \Rightarrow f_X(x) = \frac{dF_X(x)}{dx} = \int_{y=-\infty}^{\infty} f_{X,Y}(x,y) \, dy \]

Similarly,

\[ f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y) \, dx \]
Example 3.9 Jointly Uniform PDF

\[ f_{X,Y}(x,y) = c, \]

- If \( a < x < b \) and \( c < y < d \), find \( c \).
- If \( |x| + |y| \leq r \), find \( c \).
- If \( \sqrt{x^2 + y^2} \leq r \), find \( c \).
Expectation

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

- \( E[aX + bY + c] = aE[X] + bE[Y] + c \)
- You can easily extend the results of two random variables to more joint random variables.
3.5 Conditioning
Conditioning an RV on an Event

- Conditioning CDF and PDF

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

\[ f_{X|A}(x) = \frac{d}{dx} F_{X|A}(x) \]

- Special case: \( A = \{X \in B\} \)

\[ F_{X|X \in B}(x) = P(X \leq x | X \in B) = \frac{P(\{X \leq x\} \cap \{X \in B\})}{P(X \in B)} \]

\[ = \frac{\int_{t \leq x, t \in B} f_X(t)dt}{P(X \in B)} \]

\[ f_{X|X \in B}(x) = \frac{d}{dx} F_{X|X \in B}(x) = \frac{f_X(x)}{P(X \in B)}, \quad x \in B. \]
Example 3.13. The Exponential RV

The time $T$ until a new light bulb burns out is an exponential rv with parameter $\lambda$. You turn the light on, leaves the room, and when you returns, $t$ time units later, finds that the light bulb is still on, which corresponds to the event $A = \{ T > t \}$. Let $X$ be the additional time until the light bulb burns out. What is the conditional CDF of $X$, given the event $A$?

**Hint.** $P(X > x|A) = P(T > T + x|T > t) = ?$
Example 3.14.

The metro train arrives at the station near your home every quarter hour starting at 6:00 a.m. You walk into the station every morning between 7:10 and 7:30 a.m., and your arrival time is a uniform random variable over this interval. What is the PDF of the time you have to wait for the first train to arrive?

Hint. Let $X$ be the time of your arrival. $X$ is a uniform random variable over the interval from 7:10 to 7:30. Let $Y$ be the waiting time. Let $A$ and $B$ be the events

$$A = \{ 7 : 10 \leq X \leq 7 : 15 \} = \{ \text{you board the 7:15 train} \},$$
$$B = \{ 7 : 15 < X \leq 7 : 30 \} = \{ \text{you board the 7:30 train} \}.$$

$$f_Y(y) = P(A)f_{Y|A}(y) + P(B)f_{Y|B}(y)$$
Let $X$ and $Y$ be continuous random variables with joint PDF $f_{X,Y}(x,y)$. The conditional PDF of $X$ given that $Y = y$, is defined by

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)}$$
Example 3.15.

\[ f_{X,Y}(x,y) = c, \quad x^2 + y^2 \leq r^2 \]

Find \( c, f_Y(y) \) and \( f_{X|Y}(x|y) \).
Conditional Expectations

Definition.

\[ E[X|A] = \int_{-\infty}^{\infty} xf_{X|A}(x) \, dx \]

\[ E[X|Y = y] = \int_{-\infty}^{\infty} xf_{X|Y}(x|y) \, dx \]

Total expectation. Let \( A_1, A_2, \ldots, A_n \) form a partition of the sample space.

\[ E[X] = \sum_{i} P(A_i) E[X|A_i] \]

\[ E[X] = \int_{-\infty}^{\infty} f_Y(y) E[X|Y = y] \, dy = E_Y[ E[X|Y = y] ] \]
Independence

- Two continuous random variables $X$ and $Y$ are independent if
  \[ f_{X,Y}(x,y) = f_X(x)f_Y(y) \]

- Three continuous random variables $X$, $Y$ and $Z$ are independent if
  \[ f_{X,Y,Z}(x,y,z) = f_X(x)f_Y(y)f_Z(z) \]

- If $X$ and $Y$ are independent, we have
  \[ E[XY] = E[X]E[Y] \]