Topic 6: Continuous Random Variables

Lecture Outline

- Continuous Random Variables
  - Probability Density Functions (PDF)
  - Typical Continuous RV: *Uniform*, *Exponential, Gaussian*
- Expectation and Variance of Continuous RVs
  - Expected Value Rule
- Cumulative Distribution Functions (CDF)
- Normal (*Gaussian*) Random Variable (常態分佈、高斯分佈)

Reading : Textbook 3.1- 3.3

Continuous Random Variable

A random variable X defined on an experiment in such a way that X can take on a continuum of values is a *continuous random variable*, e.g.,

- the voltage across a heated resistor
- the unknown phase of a random electromagnetic wave
- the noise level in the recorded signal of an ECG plot

Definition:

A random variable X is called continuous if its probability law can be described in terms of a nonnegative function f_X, called the ***probability density function*** of X, or PDF (pdf) for short, which satisfies

$$P(X \in B) = \int_B \underbrace{f_X(x)}_{\text{PDF}} dx$$

for every subset B of the real line.

Therefore, the PDF completely describes the behavior of a continuous RV.

Analogy/Difference between PMF and PDF

Probability Mass Function (PMF)	Probability Density Function (PDF)
• PMF is used to describe discrete random variables • For a discrete random variable X and a set B with finite number of real elements, the probability that $X \in B$ is $$P(X \in B) = \sum_{x \in B} p_X(x) = \sum_{x \in B} P(X = x)$$ • PMF $p_X(x)$ itself represents the probability of the event $X=x$	• PDF is used to describe continuous random variables • For a continuous random variable X every subset B of the real line, the probability that $X \in B$ is $$P(X \in B) = \int_B f_X(x)dx$$ • PDF is *NOT probability* of any event.

Takeaway points:

1. The summation in discrete cases becomes integral in the continuous cases
2. The measure $f_X(x)dx$ bears similar interpretations to $P(X=x)$, *i.e.* they're probabilities

Why? How to Interpret PDF from PMF?

we explain ***why using integral*** with the help of ***calculus*** to evaluate probabilities involved with a continuous RV.

For a discrete random variable X, we know

$$P(a \leq X \leq b) \;=\; \sum_{a \leq x \leq b} P(X = x)$$

But we cannot do the same for continuous *X, because P(X=x) is **zero** for any continuous X and any possible value x.*

However, the idea can still be extended from regarding X as discrete and pushing to the limit as a discrete X ***approaches to*** a continuous random variable. We can let X take uncountably infinite number of values in $[a,b]$, and evaluate the probability $P(a \leq X \leq b)$ using the concept of Riemann sum.

Detail is provided in the supplementary material.

Probability Density Function

By definition, $P(X \in B) = \int_B f_X(x)dx$ **(計算以連續RV所描述事件發生機率時最常使用做法)**

- The probability of the event $a \le X \le b$ is

$$P(a \le X \le b) = \int_a^b f_X(x)dx$$

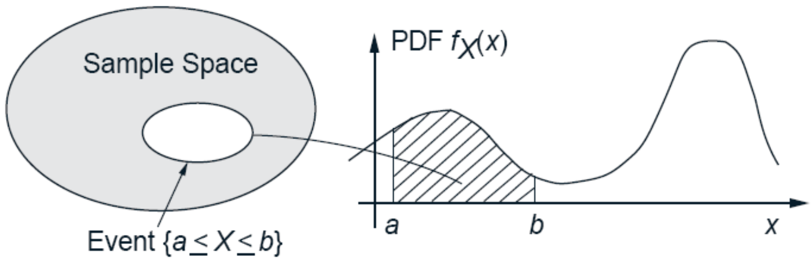
即 PDF 圖形下在 a 和 b 之間的面積

  - So, the probability of X=c for any constant c is $P(X = c) = \int_c^c f_X(x)dx = 0$

但這不表示 X=c 不會發生，只是它發生的機率極低，數學上以0表示

- PDF $f_X(x)$ must be nonnegative since for any subset *B*

$$P(X \in B) = \int_B f_X(x)dx \ge 0$$



Probability Density Function

- PDF $f_X(x)$ must satisfy normalization equation, *i.e.*

$$P(-\infty < X < \infty) = \int_{-\infty}^{\infty} f_X(x)dx = 1$$

- Unlike PMF of a discrete RV, PDF is ***not*** the probability of any event. Hence, in particular, it needs ***not*** be less than 1 for all x

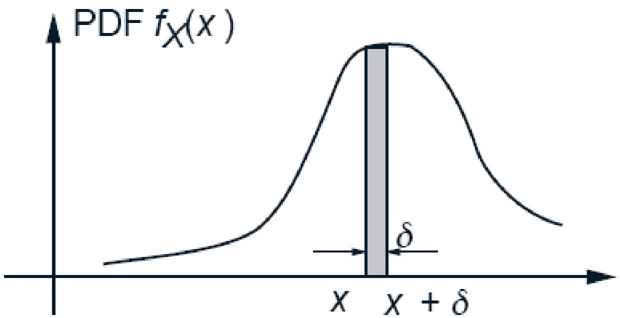
Example:

The function $f_X(x)=2$ for $0 \leq x \leq 1/2$ is a legitimate PDF whose value is larger than 1.

Interpretation of Probability Density Function

For an interval $[x, x + \delta]$ with very small length δ, we have

$$P(x < X < x + \delta) = \int_x^{x+\delta} f_X(x)dx \approx f_X(x) \cdot \delta$$



PDF的**物理意義**:

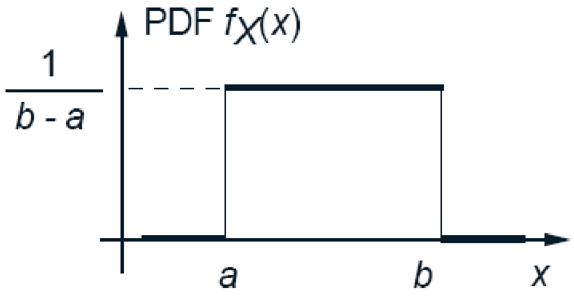
- We can view $f_X(x)$ as the "probability per unit length" near x
- PDF是密度的概念。觀念上，$f_X(3)$ 並非是 X=3的機率值，而是約略正比於X落在3附近的機率大小。事實上，X=3的機率為 0
- What does $f_X(x_1) > f_X(x_2)$ mean?

Example - Uniform Random Variable

Continuous Uniform PDF:

Let X be a uniformly distributed continuous random variable within [a,b], its PDF is given by

$$f_X(x) = \begin{cases} c & \text{if } a \le x \le b \ , \\ 0 & \text{otherwise.} \end{cases}$$



Example - Exponential Random Variable

Exponential PDF:

（定義）A random variable is said to be exponential if its PDF takes the form

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & \text{for } x \geq 0, \\ 0 & \text{otherwise.} \end{cases}$$

where λ is a positive parameter that characterizes the PDF

（應用）Exponential random variable is often used to model the *amount of time* we need to wait (waiting time) until a certain rare event occurs

Examples:

- The amount of time we wait in the station until the next bus arrives
- The amount of time it takes until the next earthquake strikes

(Why?) The rationale of why exponential random variable is related to waiting time is explained in the supplementary material.

Expectation

- The expected value or mean of a continuous random variable X is defined by

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

- **Expected Value Rule**

  In complete analogy to the discrete case, the mean of $Y = g(X)$ for any real-valued function $g(\cdot)$ is given by

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

  You can find a proof in the textbook (page 185).

Variance

- The variance of a (continuous) random variable X is

$$\begin{aligned}\mathrm{var}(X) &\triangleq E[(X - E[X])^2] \\ &= E[X^2] - (E[X])^2\end{aligned}$$

- If $Y=aX+b$, where a and b are given scalars, then we have

$$\begin{aligned}E[Y] &= aE[X] + b \\ \mathrm{var}(Y) &= a^2\mathrm{var}(X)\end{aligned}$$

上述結果和 discrete random variables 的情況一樣!

Examples

- Mean and variance of a random variable X uniformly distributed in $[a,b]$

$$E[X] = \frac{a+b}{2}$$
$$\text{var}(X) = \frac{(b-a)^2}{12}$$

- Mean and variance of an exponential random variable X with parameter λ

$$E[X] = \frac{1}{\lambda}$$
$$\text{var}(X) = \frac{1}{\lambda^2}$$

Example

Example 3.5, page 148:

The time until a small meteorite first lands anywhere in the Sahara desert is modeled as an exponential random variable with a mean of 10 days. The time is currently midnight. What is the probability that a meteorite first lands some time between 6am and 6pm of the first day?

Cumulative Distribution Functions (CDF)

- For discrete RV's, we use PMF's; for continuous RV's, we use PDF's, to complete characterize RVs

- Some RVs are mixed, having both discrete and continuous components, e.g., *packets waiting time in a router* (applications in networking)
  - If the router buffer is empty, the incoming packet does not have to wait (waiting time=0, discrete, with nonzero probability)
  - A packet may wait for a random amount of time (continuous) until it can be processed

  See Problem 9 in the textbook for another example!

- The cumulative distribution function or CDF $F_X(x)$ of a random variable can characterize all these cases — discrete, continuous, and mixed.

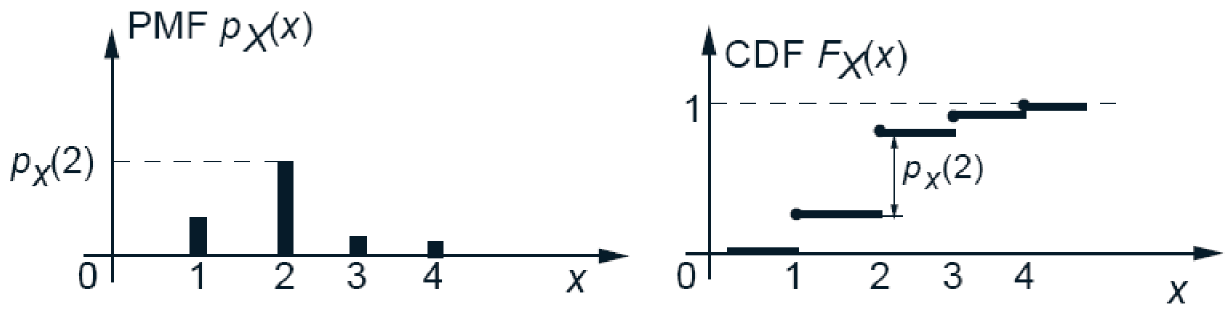
  The CDF of a random variable X is defined by

$$F_X(x) \triangleq P(X \leq x)$$

Obtain CDF from PMF and PDF

- For discrete random variables, the CDF is related to PMF in terms of

$$F_X(x) \triangleq P(X \leq x) = \sum_{k \leq x} p_X(k)$$



- For continuous random variables, the CDF is related to PMF in terms of

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

Obtain PDF and PMF from CDF

- Suppose that the RV X is continuous and is described by a pdf $f_X(x)$. Then,

$$
\begin{aligned}
P(a < X \leq b) &= F_X(b) - F_X(a) \\
&= \int_b^a f_X(x)dx
\end{aligned}
$$

From the fundamental theorem of calculus (微積分基本定理), we have, for continuous X,

$$
\boxed{f_X(x) = \frac{dF_X(x)}{dx}}
$$

這個結果經常被使用。當我們想找尋一個隨機變數 X 的 PDF 時，典型作法即是先找其 CDF，再微分之!

- If X is discrete and takes integer values,

$$
\begin{aligned}
p_X(k) &= P(X \leq k) - P(X \leq k-1) \\
&= F_X(k) - F_X(k-1)
\end{aligned}
$$

Properties of CDF (1)

- CDF is the probability of a certain event. Hence, CDF is nonnegative and less than 1. Also we have

$$F_X(\infty) = 1$$

- The probability that X falls within an interval can be computed from CDF

$$\begin{aligned} P(a < X \leq b) &= P(X \leq b) - P(X \leq a) \\ &= F_X(b) - F_X(a) \end{aligned}$$

  - The previous property proves that $F_X(x)$ is a non-decreasing function of its argument, *i.e.*, either it grows or it stays the same
  - Since $P(X=c)=0$ for any constant c, we actually have

$$P(a \leq X \leq b) = P(a < X \leq b) = P(a \leq X < b) = P(a < X < b)$$

Properties of CDF (2)

- **Finding PDF from CDF**

  As just mentioned, it is typical to find the PDF of a continuous RV by first finding the CDF and then differentiating. That is

$$\begin{aligned} f_X(x) &= \frac{dF_X(x)}{dx} \\ &= \frac{d}{dx} P(X \leq x) \end{aligned}$$

  Example:

  Suppose you are allowed to observe a continuous RV independently n times $X_1, \ldots, X_n$, and define

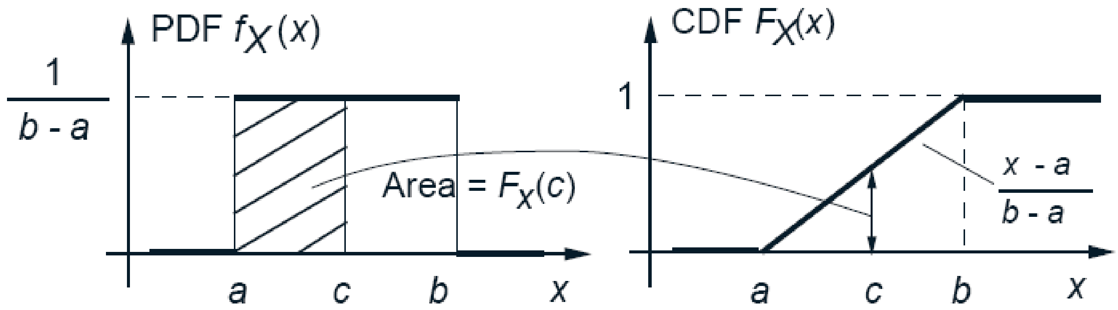
$$X = \max\{X_1, X_2, \cdots, X_n\}$$

  Assuming the X_i are mutually independent, what is the PDF of X?

Example - CDF of Continuous Uniform

Let X be a uniformly distributed continuous random variable within $[a,b]$, its CDF is given by

$$F_X(x) = \begin{cases} 0 & \text{if } x \le a \ , \\ \frac{x-a}{b-a} & \text{if } a \le x \le b, \\ 1 & \text{if } x \ge b. \end{cases}$$

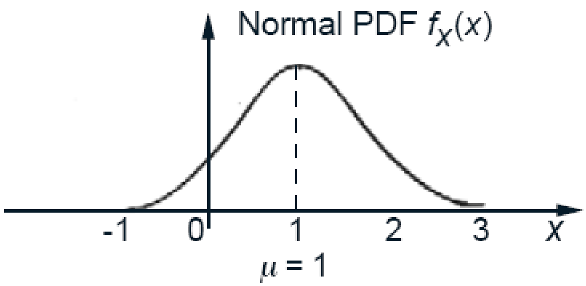


Normal (Gaussian) Random Variable 常態、高斯分布

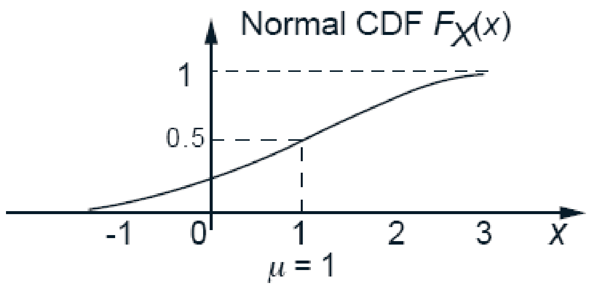
Normal PDF:

A continuous random variable X is said to be normal (or Gaussian) with parameter μ and σ^2 , often denoted by $X \sim N(\mu, \sigma^2)$, if it has a PDF of the form

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



bell-shaped



Importance of Gaussian Random Variable

- The normal random variable plays an important role in a broad range of probabilistic models

  The main reason is

  it models well the *additive effect* of many independent factors, in a variety of engineering, physical, and statistical contexts

  This is a result from the central limit theorem (中央極限定理)

- ***Central Limit Theorem***

  Mathematically, the key fact is that the sum of a large number of independent and identically distributed (i.i.d.) random variables $X_1 \ldots X_n$ with **ANY** probability distributions has an approximately **Gaussian CDF**, regardless of the actual distributions of the random variables. We will talk more about this in **chapter 5**.

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} X_i \xrightarrow{n \to \infty} \text{Gaussian}$$

Properties of Gaussian

- We see the Gaussian PDF is a bell-shaped curve that is symmetric about μ

- The mean and variance of a Gaussian random variable with parameter μ and σ^2 are

$$E[X] = \mu$$
$$\mathrm{Var}(X) = \sigma^2$$

- A Gaussian random variable Y with zero mean and unit variance is said to be standard Gaussian

- **(定理: 高斯的線性轉換仍為高斯)**
  Let X be a Gaussian random variable with mean μ and variance σ^2, and if a, b are scalars, then the random variable

$$Y=aX+b$$

  is also Gaussian with mean $a\mu+b$ and variance $a^2\sigma^2$.

  You need to find the PDF of Y to prove this statement.

Standard Normal

- A normal random variable Y with zero mean and unit variance is said to be a standard normal. The CDF $\Phi(y)$ of a standard normal is

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

  No closed-form expression exists for $\Phi(y)$.

  - The value of $\Phi(y)$ for various y can be looked up from the normal table
  - $\Phi(-y)=1-\Phi(y)$

- We can standardize a normal random variable $X \sim N(\mu, \sigma^2)$ by

$$Y \triangleq \frac{X - \mu}{\sigma}$$

Example – A Simple Communication System

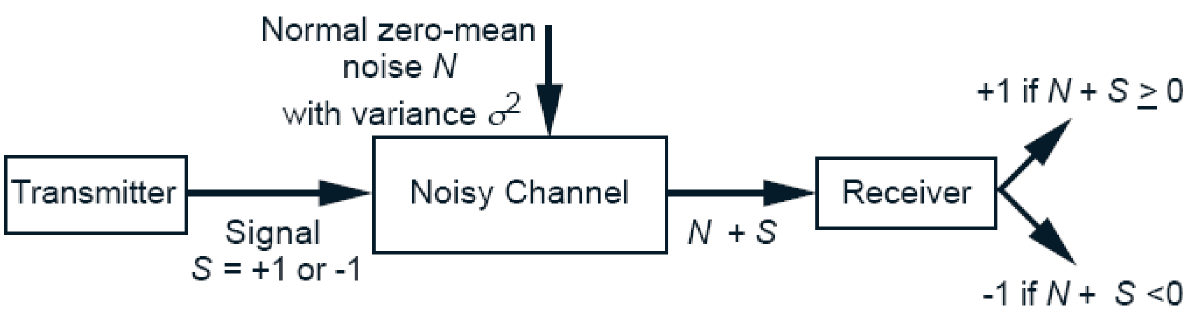
A binary message is transmitted as a signal S that is either −1 or +1. The communication channel corrupts the transmission with additive Gaussian noise N having mean μ = 0 and variance σ^2. The model for the received signal R is

$$R = S + N$$

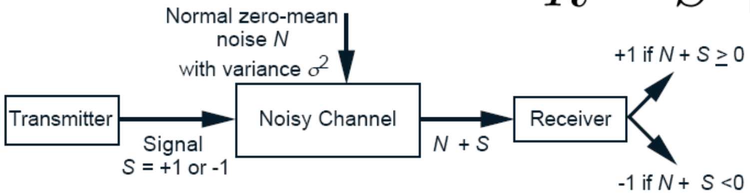
(Receiver 訊號偵測機制) Suppose that the receiver concludes that the signal −1 (or +1) was transmitted if the value of R is < 0 (or ≥ 0), respectively.

Question:

What is the probability that the receiver makes an error decision?



$$R = S + N$$



What is the probability that the receiver makes an error decision?