# Week 1

**Preliminaries:** moment generating function, joint distribution, conditional probability, random sum and basic concepts of stochastic processes.

## 1 Random variable (vector), moment generate function and joint distribution

Let $\Omega$ denote the set of all outcomes in an experiment.

- A random variable (r.v.) $X$ is defined to be a function $X$: $\Omega \to R$.
- We define

  $$F(t) = P(X \le t) = P\{\omega : X(\omega) \le t\}, \quad t \in R$$

  the distribution function of $X$.
- $X$ is called discrete if there is a finite or denumerable set of real values $x_1, x_2, ...$ such that

  $$p_i = P(X = x_i) > 0, \quad i = 1, 2, ... \quad \text{and} \quad \sum_{i=1}^{\infty} p_i = 1.$$

  Binomial variable $X \sim B(n, p)$ :

  Poisson variable $X \sim Poisson(\lambda)$:

- $X$ is called continuous if there is a non-negative function $p(s)$ such that
$$F(t) = \int_{-\infty}^{t} p(s)ds.$$
$p(s)$ is called density function of $X$.

Normal variable $X \sim N(0,1)$:

Exponential variables $X \sim Exp(\lambda)$:

- The m-th moment of a r.v. $X$:
$$EX^m = \sum_{i=1}^{\infty} x_i^m P(X = x_i) \quad \text{or} \quad EX^m = \int_{-\infty}^{\infty} x^m\, p(x)dx$$
Note that $var(X) = EX^2 - (EX)^2$.
$EX$ is called the expectation of $X$. $EX^m$ exists if $E|X|^m < \infty$.

If $X \sim Poisson(\lambda)$, then $EX = \lambda$ and $var(X) = \lambda$.

If $X \sim Exp(\lambda)$, then $EX = 1/\lambda$ and $var(X) = 1/\lambda^2$.

- Moment generating function (mgf) of $X$: $g(s) = Es^X, s \geq 0$.

  Another definition: $g(t) = Ee^{tX}, t \in R$.

  If $X$ only takes non-negative integers, $g(x)$ is also called the probability generating function (pgf) of $X$, i.e.,

$$g(s) = Es^X = \sum_{j=0}^{\infty} s^j \, p_j, \quad \text{where } p_j = P(X = j).$$

  If $X \sim Possion(\lambda)$, then $g(s) = Es^X = e^{\lambda(s-1)}$;

  if $X \sim Exp(\lambda)$, then $g(t) = Ee^{tX} = (1 - t/\lambda)^{-1}$ for $t < \lambda$.

- The joint distribution function of $X_1, X_2, ..., X_n$ [random vector $(X_1, ..., X_n)$] is defined as

$$F(x_1, x_2, ..., x_n) = P(X_1 \le x_1, X_2 \le x_2, ..., X_n \le x_n), \quad x_j \in R.$$

  - Joint distribution of discrete $(X, Y)$:

$$P(X = k, Y = j) \;=\; ....$$

  - Joint density of continuous $(X, Y)$: $p(s, t) \ge 0, s, t \in R$,

$$P(X \le x, Y \le y) \;=\; \int_{-\infty}^{x} \int_{-\infty}^{y} p(s, t) ds dt$$

- $X_1, X_2, ..., X_n$ are said to be independent if, for all $x_j \in R$,

$$F(x_1, x_2, ..., x_n) = P(X_1 \le x_1)\, P(X_2 \le x_2) \ldots P(X_n \le x_n).$$

  If $X_1, X_2, ..., X_n$ are independent, then, for $A_1, A_2, ..., A_n \subset R$,

$$\begin{aligned} & P(X_1 \in A_1, X_2 \in A_2, ..., X_n \in A_n) \\ = \;& P(X_1 \in A_1)\, P(X_2 \in A_2) \ldots P(X_n \in A_n). \end{aligned}$$

  Discrete random variable:

$$P(X = x_k, Y = y_j) \;=\; P(X = x_k)\, P(Y = y_j), \quad \text{all } k, j$$

  Continuous random variable:

$$p(s, t) = p_X(s)\, p_Y(t), \quad \text{all } s, t \in R.$$

- Three useful theorems

  **Th1 .1.** Suppose that $X_1, X_2, ..., X_n$ are independent with mgfs $g_1(s), ..., g_n(s)$, respectively. Then the mgf of $S_n = \sum_{k=1}^n X_k$ is given by

$$g(s) \;=\; Es^{S_n} = \Pi_{k=1}^n g_k(s).$$

  *Proof.*

  **Th1.2.** (adv) Let $X$ be a r.v. only takes integer values. Then the dist of $X$ is uniquely determined by its pgf $g(s) = Es^X, 0 \le s \le 1$.

  **Th1.3.** (adv) Let $g(t) = Ee^{tX}$. If $var(X) < \infty$, then

$$EX = g'(t)\big|_{t=0}, \quad var(X) = g''(t)\big|_{t=0} - \left[g'(t)\big|_{t=0}\right]^2.$$

# 2 Conditional probability and conditional expectation

For events $A$ and $B$, we define

$$P(A|B) = \frac{P(A \cap B)}{P(B)}, \quad \text{if } P(B) > 0.$$

If $X$ is a discrete RV, the **conditional distribution** of $X$ given the event $B$ is

$$P(X = x|B) \;=\; \frac{P(X = x, B)}{P(B)}, \quad x = x_1, x_2, ...$$

and the **conditional expectation** of $X$ given $B$ is

$$E[X|B] \;=\; \sum_{j=1}^{\infty} x_j P(X = x_j \mid B).$$

Note that $E[X|B] = E(XI_B)/P(B)$ if $P(B) > 0$.

- Suppose that $X$ and $Y$ are two discrete rvs, taking values $x = x_i, i = 1, 2, ...$ and $y = y_j, j = 1, 2, ...$, respectively. If $P(Y = y) \neq 0$, then

  $$P(X = x|Y = y) \;=\; \frac{P(X = x, Y = y)}{P(Y = y)}, \quad x = x_1, x_2, ...$$

  **The law of total probability**:

  $$P(X = x) \;=\; \sum_{j=1}^{\infty} P(X = x \mid Y = y_j) P(Y = y_j).$$

**Total expectation theorem**:

$$EX = \sum_{j=1}^{\infty} E(X \mid Y = y_j) P(Y = y_j).$$

*Proof.*

- Suppose that $X$ and $Y$ are two continuous rvs having a joint density $f(x,y)$. Let $f_X(y) = \int f(x,y)dy$ and $f_Y(y) = \int f(x,y)dx$ be marginal densities of $X$ and $Y$, respectively. The **conditional density** of $X$ given by $Y = y$ is

$$f_{X|Y}(x \mid y) \;=\; \frac{f(x,y)}{f_Y(y)}.$$

**Note that:** If $X$ and $Y$ are independent, then $f_{X|Y}(x \mid y) = f_X(x)$ or $f(x,y) = f_X(x)f_Y(y)$.

We have:

$$P(a \leq X \leq b \mid Y = y) \;=\; \int_a^b f_{X|Y}(x \mid y)dx,$$

$$E(X \mid Y = y) \;=\; \int x f_{X|Y}(x \mid y)dx,$$

$$\begin{aligned} EX &= \int\int x f(x,y)dxdy \\ &= \int E(X \mid Y = y) f_Y(y)dy \end{aligned}$$

- Conditional expectation as a Random Variable

  Conditional expectations such as $E(X|Y=2)$ or $E(X|Y=5)$ are numbers. So $\psi(y) := E(X \mid Y=y)$ is a function of $y$

  If we consider $\psi(Y) = E(X|Y)$, it is a random variable, which is usually called **the conditional expectation** of $X$ given $Y$.

  It has the following properties:

  – If $X$ and $Y$ are independent, then $E(X \mid Y) = EX$;

  – If $E|X| < \infty$, then $EX = E\big[E(X \mid Y)\big]$;

  – For any functional $g(X,Y)$ of $(X,Y)$, we have

  $$E\big[g(X,Y) \mid Y=y\big] = E\big[g(X,y) \mid Y=y\big].$$

  In particular, if $X$ and $Y$ are independent, then

  $$E\big[g(X,Y) \mid Y=y\big] = Eg(X,y).$$

  – If $Z = g(Y)$, then $E(XZ \mid Y) = Z\,E(X \mid Y)$.

- Example : Let $X$ and $Y$ be independent having the same distribution with $E|X| < \infty$. Let $Z = X+Y$ . Find $E(Z|X), E(X|Z), E(XZ|X), E(XZ|Z)$.

# 3 Random sum and basic concepts of random processes

Very often in applications, we need to consider random sum: $S_N = \sum_{k=1}^{N} X_k$, where

- $N$ is a random variable taking non-negative integer values with $P(N = j) = p_j, j = 0, 1, 2, ...$ ;
- $X_k, k \geq 1$, is a sequence of random variables.

**Examples:**

- Queueing: $N :=$ the number of customers arriving in a specific time period, $X_k :=$ the service time required by the ith customer. Then $S_N$ denotes the total service time required by customers arriving in that time period.
- Risk Theory: $N :=$ the number of claims arriving at an insurance company in a given week, $X_k :=$ the amount of the kth claim. Then $S_N$ denotes the total liability for that week.

We have the following result.

- **Wald's identity**: If $X_k$ are iid with $E|X_1| < \infty$ and $N$ is non-negative integer value rv such that $\{N = n\}$ is independent of $X_1, X_2, ...$, then $ES_N = ENEX_1$ provided that $EN < \infty$.

**Proof.** By using the total expectation theorem: $(S_0 := 0)$

$$\begin{aligned}
ES_N &= \sum_{j=0}^{\infty} E\big(S_N \mid N = j\big)\, P(N = j) \\
&= \sum_{j=0}^{\infty} E\Big(\sum_{k=1}^{j} X_k\Big)\, P(N = j) \quad \text{(due to the independence between } N \text{ and } X_k\text{)} \\
&= \sum_{j=0}^{\infty} (jEX_1)\, P(N = j) \\
&= EX_1 \sum_{j=0}^{\infty} jP(N = j) = EX_1 EN.
\end{aligned}$$

A **stochastic process** is a collection of random variables

$$(X_t, t \in T) = (X_t(\omega), t \in T, \omega \in \Omega),$$

defined on some probability space $(\Omega, \mathcal{F}, P)$, where $T$ is an index set such as $T = \{0, 1, 2, ...\}$ or $T = [0, \infty)$, etc. We denote **state space** by $S$, the set of all possible values of $X_t, t \in T$.

**Examples:**

- Let $X_k, k \geq 1$ be i.i.d r.vs only taking integer values. Let $S_0 = 0$ and $S_n = \sum_{k=1}^{n} X_k$. $\{S_n, n \geq 0\}$ is a random process with state space $S = \{... -2, -1, 0, 1, 2, ...\}$, This process is commonly called **Random walk**.

- Let $X_t$ denote the total number of some events (telephone calls, customer arrivals, etc) that occurs within the time interval $[0, t]$. $\{X_t, t \in [0, \infty)\}$ is a random process with state space $S = \{0, 1, 2, ...\}$. This process will become the well-known **Poisson process** under some additional conditions.

- Let $X_k$ denote the S&P 500 index in the k-th day in 1979. $\{X_k, k = 1, 2, ..., 365\}$ is a random process with state space $S = (0, \infty)$.

**Remarks:**

- A stochastic process is actually a function of two variables $t$ and $\omega$:
  (a) for a fixed time $t$, it is a r.v. $X_t = X_t(\omega), \omega \in \Omega$;
  (b) for a fixed $\omega$, it is a function of time $X_t = X_t(\omega), t \in T$, which is called a **realization**, a **trajectory**, or **a sample** path of the process $X$.

  It may be useful for the intuition to think of $t$ as "time" and each $\omega$ as an individual "particle", "stock price" or "experiment". Then the picture $X_t(\omega)$ would represent the position (or the result) at time $t$ of the particle (stock price, or experiment) $\omega$. The idea is that if we run an experiment and observe the random values of $X_t(\omega)$ as time evolves, we are in fact looking at a sample path $\{X_t(\omega) : t \geq 0\}$ for some fixed $\omega \in \Omega$. If we rerun the experiment, we will in general observe a different sample path.

- The **finite-dimensional distributions (fdds)** of the stochastic process $X_t, t \in T$, are the distribution of the finite-dimensional vectors

$$(X_{t_1}, ..., X_{t_n}), \qquad t_1, ..., t_n \in T,$$

  for all possible choices of times $t_1, ..., t_n \in T$ and every $n \geq 1$.

  The family of fdds determines many (but not all) important properties of a stochastic process, describing the relationships among the random variables, $X_t, t \in T$.

- The process $\{X_t, t \in T\}$ is **strictly stationary** if the fdds are invariant under shifts of the index $t$:

$$(X_{t_1}, ..., X_{t_n}) =_d (X_{t_1+h}, ..., X_{t_n+h}),$$

for all possible $t_1, t_2, ..., t_n \in T$, $n \geq 1$, and $h$ such that $t_1 + h, t_2{+}h, ..., t_n + h \in T$, where $=_d$ denotes the same in distribution.

**Examples:**

- The process $\{X_t, t \in T\}$ is **weakly stationary** if its mean and covariance functions are invariant under shifts of the index $t$:

$$EX_{t+h} = EX_t, \qquad Cov(X_{s+h}, X_{t+h}) = Cov(X_s, X_t)$$

  for all possible $s, t \in T$, and $h$ such that $s+h, t+h \in T$.

- The two modes of stationarity do not imply each other. Clearly, weak stationarity does not imply strictly stationarity. On the other hand, if $\{X_t, t \in T\}$ is strictly stationary, it may not have first and/or second moments, which is required for weak stationarity. Of course, if the second moments of $X_t$ exist, then strictly stationarity does imply weak stationarity.

  **Examples:**

- Many important processes deal with properties of increments. Let $\{X_t, t \in T\}$ be a stochastic process and $T \in R$ be an interval.
  - $X_t$ is said to have **stationary increments** if

    $$X_t - X_s =_d X_{t+h} - X_{s+h}, \quad \text{for all } t, s \text{ and } h \text{ with } t+h, s+h \in T.$$

  - $X_t$ is said to have **independent increments** if for for all $t_i \in T$ with $t_1 < t_2 < ... < t_n$ and $n \geq 1$

    $$X_{t_2} - X_{t_1}, X_{t_3} - X_{t_2}, \cdots, X_{t_n} - X_{t_{n-1}}$$

    are independent random variables.

  Independence and stationarity are clearly two distinct concepts. Processes with increments satisfying both requirements are an important class of processes in stochastic calculus.

  **Examples:**

A random process may be considered as a well-defined process once its state space $S$, index set $T$ and the joint distribution family are prescribed. Depending on the nature of the state space $S$ and the index set $T$, we may classify random processes into four classes:

**1.** Both $S$ and $T$ discrete: discrete valued random processes. The random walk and Markov chain are in this class.

**2.** $S$ discrete and $T$ continuous: discrete valued continuous parametric random processes. The Poisson process and point processes are in this class.

**3.** $S$ continuous and $T$ discrete: continuous random process with discrete parameter. Many financial random processes, such as the process in Ex 1.3, in this class.

**4.** Both $S$ and $T$ continuous: continuous random processes. Brownian motion, etc, are in this class.