TA Session 1

16/08/2019

Agenda :

- Review of probability : Sample space, event space, probability measure,
  Inequalities : Markov, Chebyshev, Chernoff bound
  Limit theorems: law of large numbers,
  Central limit theorem
- Example simulations

Sample Space, Event space and probability measure

- Let E denote an experiment. Let Ω be the collection of all possible outcomes of this experiment.
- The set Ω is known in the probability literature as the Sample space

Examples :

① Coin toss : $\Omega = \{H, T\}$

Die throw : $\Omega = \{1, 2, \ldots, 6\}$

Coin toss until the first head is observed : $\Omega = \{H, TH, TTH, \cdots\}$

Def^n (event): Any subset of Ω for which we are interested in assigning probability is known as an event.

Def^n (event space): The collection of all events is called an event space. We shall denote an event space by $\mathcal{F}$. Some of the important properties that an event space has to satisfy are :

- $\Omega \in \mathcal{F}$
- $A \in \mathcal{F} \Rightarrow A^c \in \mathcal{F}$
- $A_1, A_2, \cdots \in \mathcal{F} \Rightarrow \bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$

<u>Examples of event space</u> :

1. $\mathcal{F} = \{\phi, \Omega\}$, $\mathcal{F} = 2^{\Omega}$
2. If $\Omega = \{1, 2, \ldots, 6\}$, $\mathcal{F} = \{\phi, \Omega, \{1\}, \{2,3,4,5,6\}\}$

<u>Remark</u> : An event space is also referred to as a <u>σ-algebra</u>. The pair $(\Omega, \mathcal{F})$ is referred to as a <u>measurable space</u>.

<u>Def^n (probability measure)</u> :

Let $(\Omega, \mathcal{F})$ be a measurable space. A function $P : \mathcal{F} \to [0,1]$ is known as a probability measure if it satisfies the following conditions :

1. $P(\Omega) = 1$
2. If $A_1, A_2, \cdots$ is a collection of pairwise disjoint sets, then

$$P\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n).$$

<u>Example</u> : ① $\Omega = \{1, 2, \ldots, 6\}$

$\mathcal{F} = 2^{\Omega}$

$P(\{i\}) = \frac{1}{6} \quad \forall \ i = 1, \ldots, 6.$

② $\Omega = \{1, \dots, p\}$, where p is prime.

$\mathcal{F} = 2^{\Omega}$

$$P(A) = \frac{|A|}{p}, \quad A \subseteq \Omega \qquad (|A| = \#\text{ elements in } A)$$

Check that P is a valid probability measure.

Random Variables and Expectations

Let $(\Omega, \mathcal{F})$ be a measurable space. A function $X : \Omega \to \mathbb{R}$ is known as a random variable if :

for every $x \in \mathbb{R}$,

$$\{\omega \in \Omega : X(\omega) \leq x\} \in \mathcal{F}$$

$$\begin{tikzcd}
(\Omega, \mathcal{F}) \arrow[rr, "X"] & & \mathbb{R} \arrow[d] \\
\{\omega \in \Omega : X(\omega) \leq x\} \arrow[u] & & x \arrow[ll]
\end{tikzcd}$$

Talk about how $P(X \leq x)$ makes sense for each x. Introduce CDF here.

Example : ① $\Omega = \{H, T\}$

$\mathcal{F} = 2^{\Omega}$

$$X(\omega) = \begin{cases} 1, & \omega = H \\ 0, & \omega = T \end{cases}$$

Then, $$\{\omega \in \Omega : X(\omega) \leq x\} = \begin{cases} \phi, & x < 0 \\ \{T\}, & 0 \leq x < 1 \\ \Omega, & x \geq 1. \end{cases}$$

Thus, X is a random variable.

② $\Omega = \{1, 2, 3, 4\}$

$\mathcal{F} = \{\phi, \Omega, \{1\}, \{2,3\}, \{2,3,4\}, \{1,4\}, \{4\}, \{1,2,3\}\}$

$$X(\omega) = \begin{cases} -\pi, & \omega = 2,3 \\ \pi, & \omega = 1 \\ 0, & \omega = 4. \end{cases}$$

Check that the above is a random variable.

Suppose now that

$$X(\omega) = \begin{cases} 1, & \omega = 1,3,4 \\ -\sqrt{2}, & \omega = 2. \end{cases}$$

Then, $\{\omega \in \Omega : X(\omega) \leq -\sqrt{2}\} = \{2\} \notin \mathcal{F}$. Hence this function is not a random variable.

Remark : From the above examples, it is clear that random variables are very closely associated with the underlying event space $\mathcal{F}$ wrt which they are defined. For this matter, a random variable X defined wrt $(\Omega, \mathcal{F})$ is also referred to as an $\mathcal{F}$-measurable function.

Def^n (discrete random variable)

A random variable X is said to be a discrete random variable if it takes at most countably many values. If X is a discrete random variable, we can define a function

$$p_X : \mathbb{R} \to [0,1]$$

as

$$\begin{aligned} p_X(x) &= P(\{\omega \in \Omega : X(\omega) = x\}) \\ &= P(X = x), \quad x \in \mathbb{R}. \end{aligned}$$

p_X is known as the probability mass function (or pmf) of X.

Clearly, it follows that for a discrete rv X taking values $x_1, x_2, x_3, \dots$, we have

$$F(x) = \sum_{n : x_n \leq x} p_X(x_n)$$

$$= \sum_{n : x_n \leq x} P(X = x_n).$$

<u>Def^n (Continuous random variable)</u>

A random variable X is said to be a Continuous random variable if its CDF F_X can be expressed as

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

for some integrable function f_X.

If X is a continuous rv, f_X is called its probability density function (pdf).

<u>Expectations</u>

<u>Def^n (expectation of a discrete rv)</u>

Let $(\Omega, \mathcal{F}, P)$ be given. Let X be a discrete random variable taking values $x_1, x_2, x_3, \dots$. Then, its <u>expectation</u> is denoted by $E[X]$ and is defined as

$$E[X] = \sum_{n=1}^{\infty} x_n\, p_X(x_n)$$

$$= \sum_{n=1}^{\infty} x_n\, P(X = x_n),$$

provided the infinite sum above is not of the form $\infty - \infty$.

Def^n (expectation of a continuous random variable)

Let $(\Omega, \mathcal{F}, P)$ be given. Let X be a continuous random variable with pdf f_X. Then, its expectation is denoted by $E[X]$ and is defined as

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

provided the integral on the RHS is not of the form $\infty - \infty$.

Remark: (1) The value of $E[X]$ can be $-\infty$, $+\infty$ or any finite real value.

(2) In the case when we arrive at $\infty - \infty$ for either the infinite sum in the case of discrete rvs or the integral in the case of continuous rvs, we say that expectation is not defined.

More generally, given a continuous function $g : \mathbb{R} \to \mathbb{R}$, we define

$$E[g(X)] = \sum_{n=1}^{\infty} g(x_n)\, p_X(x_n) \qquad \text{(discrete)}$$

or

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

provided the RHSs make sense (i.e., are not of the form $\infty - \infty$).

Inequalities

(1) Markov's inequality

For any non-negative random variable X,

$$P(X > t) \leq \frac{E[X]}{t}, \qquad t > 0.$$

② <u>Chebyshev's inequality</u>:

Let X be a random variable with mean X. Define variance of X as the quantity

$$Var(X) := E\left[(X - E[X])^2\right].$$

Then, Chebyshev's inequality states that

$$P\left(|X - E[X]| > t \right) \leq \frac{Var(X)}{t^2} \quad \text{for any } t > 0.$$

③ <u>Chernoff bound</u>

Let X be a random variable defined on $(\Omega, \mathcal{F}, P)$. The moment generating function of X, denoted by M_X, is a function

$$M_X : \mathbb{R} \to [0, \infty]$$

defined as

$$M_X(t) = E\left[e^{tX}\right], \quad t \in \mathbb{R}.$$

The Chernoff bound states that

$$P(X > a) \leq \frac{E\left[e^{tX}\right]}{e^{ta}} \quad \forall \ t > 0, \ a \in \mathbb{R}.$$

<u>Example</u>:

Let $X \sim N(0,1)$. Here, $E[X] = 0$ and $Var(X) = 1$.

Let us bound $P(|X| > t)$ using the above inequalities.

(1) Markov inequality:

$$P(|X|>t) \leq \frac{E[|X|]}{t}, \quad t>0.$$

Now, $$E[|X|] = \int_{-\infty}^{\infty} |x| \cdot \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \, dx$$

$$= \frac{2}{\sqrt{2\pi}} \int_{0}^{\infty} x \, e^{-x^2/2} \, dx$$

$$= \sqrt{\frac{2}{\pi}}.$$

Thus,

$$P(|X|>t) \leq \frac{1}{t}\sqrt{\frac{2}{\pi}}, \quad t>0 \longrightarrow (*)$$

Chebyshev's inequality:

$$P(|X|>t) \leq \frac{E[X^2]}{t^2}$$

$$= \frac{Var(X)}{t^2} \quad (\because E[X]=0)$$

$$= \frac{1}{t^2}, \quad t>0 \longrightarrow (**)$$

Chernoff bound

$$m_X(s) = E[e^{sX}] = \int_{-\infty}^{\infty} e^{sx} \cdot \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \, dx$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{(x-s)^2}{2} + s^2/2} \, dx$$

$$= e^{s^2/2} \cdot \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-(x-s)^2/2} \, dx$$

$$= e^{s^2/2}.$$

Then,

$$P(|X|>t) = P(X>t) + P(X<-t)$$
$$= 2P(X>t) \quad (\because \text{ Gaussian dist}^n \text{ is symmetric})$$
$$\leq 2 \frac{E[e^{sX}]}{e^{st}} \quad \forall\ s>0,\ t>0$$
$$= 2\exp\left(-st + s^2/2\right) \quad \forall\ s>0,\ t>0$$

$\Rightarrow$ We can minimize the RHS over $s>0$ to get the tightest bound

Doing so, we get

$$P(|X|>t) \leq 2e^{-t^2/2}, \quad t>0.$$

Limit Theorems

① Weak law of large numbers

Suppose $X_1, \ldots, X_n$ are iid rvs with mean $E[X]$ and variance $Var(X)$. Then,

$$Var\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{Var(X)}{n} \longrightarrow 0 \text{ as } n\to\infty$$

Thus, intuitively, $\frac{1}{n}\sum_{i=1}^{n} X_i \to$ Constant rv as $n\to\infty$. This is the formal statement of the weak law of large numbers (WLLN)

<u>Thm (WLLN)</u>: Suppose $X_1, X_2, \ldots$ are iid rvs with $E[X]<\infty$ and $Var(X)<\infty$. Then,

$$\frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow[n\to\infty]{P} E[X],$$

i.e., $\forall\ t > 0$,

$$P\left(\left| \frac{1}{n} \sum_{i=1}^{n} X_i - E[X] \right| > t \right) \xrightarrow[n\to\infty]{} 0 .$$

<u>Thm (CLT)</u>: Suppose $X_1, X_2, \ldots$ are iid rvs with $E[X] < \infty$ and $\mathrm{var}(X) < \infty$. Then, denoting by S_n the sum $S_n = \sum_{i=1}^{n} X_i$, CLT says that

$$\frac{S_n - ES_n}{\sqrt{\mathrm{var}(S_n)}} \xrightarrow[d]{n\to\infty} N(0,1),$$

i.e., $$P\left(\frac{S_n - ES_n}{\sqrt{\mathrm{var}(S_n)}} \leq x \right) \xrightarrow[n\to\infty]{} \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-t^2/2}\, dt \qquad \forall\ x \in \mathbb{R}.$$