Random Variables & distributions

(Lecture-6-contd)

We are interested not in the experiment, but in the consequence of the outcome.

e.g. loss/gain more important that the outcome that gives them.

Example:- Fair coin tossed twice

$$\Omega = \{HH, HT, TH, TT\}$$

$$X : \Omega \to \mathbb{R}$$

$$X(\omega) = \text{\# of heads}$$

$$X(HH) = 2, \quad X(HT) = X(TH) = 1, \quad X(TT) = 0$$

Gambler wagers 1 unit on the result. Fortune doubles if head & is ruined on a fail.

$$W(HH) = 4, \quad W(HT) = W(TH) = W(TT) = 0$$

After an outcome, X takes some value.

This value is more likely to be in certain subsets of $\mathbb{R}$ than others, depending on $(\Omega, \mathcal{F}, P)$ & X.

For coin-tossing example,

"$f(x) =$ Probability that $X = x$"

can be used to describe the distribution of likelihoods of possible values of X.

In general, this is not enough.

(e.g. real intervals. $P(X = x) = 0$)

So, use distribution function

$F : \mathbb{R} \to [0,1]$ defined as

$$F(x) = P(X \leq x).$$

A more general definition of random variable (r.v.)

Def: A r.v. $X : \Omega \to \mathbb{R}$ with the property that $\{\omega \in \Omega \mid X(\omega) \leq x\} \in \mathcal{F} \ \forall x \in \mathbb{R}$.

We call "X is $\mathcal{F}$-measurable".

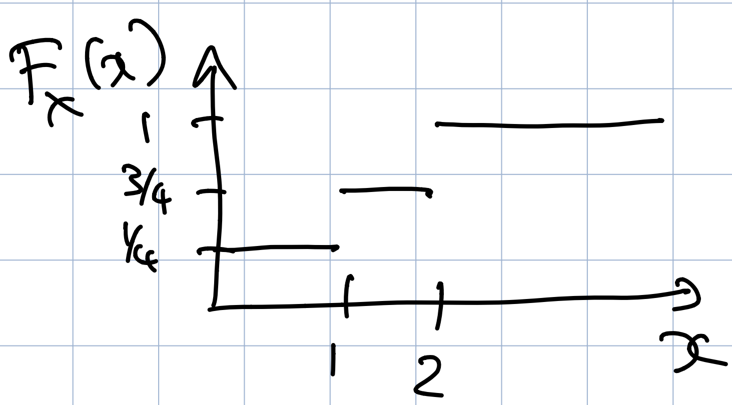
Def: Distribution of a r.v. X is

$F : \mathbb{R} \to [0,1]$ given by

$$F(x) = P(X \leq x)$$

Note: $\{X \leq x\} = \{\omega \in \Omega \mid X(\omega) \leq x)$

Example:



$$F_X(x) = \begin{cases} 0 & x < 0 \\ 1/4 & 0 \leq x < 1 \\ 3/4 & 1 \leq x < 2 \\ 1 & x \geq 2 \end{cases}$$

Properties of distributions:

A distribution F satisfies

(a) $\lim_{x \to -\infty} F(x) = 0$, $\lim_{x \to \infty} F(x) = 1$

(b) If $x < y$, then $F(x) \leq F(y)$

(c) F is right-continuous, i.e.,

$$F(x+h) \to F(x) \text{ as } h \downarrow 0$$

Proof:

(a) $B_n = \{ \omega \in \Omega \mid X(\omega) \leq -n \}$

$$= \{ X \leq -n \}$$

$$B_1 \supseteq B_2 \supseteq B_3 \supseteq - - -$$

$$B = \bigcap_i B_i = \phi$$

$P(B) = \lim_{n \to \infty} P(B_n) = 0$ | $P(B_n) = F(-n)$

So, $\lim_{x \to -\infty} F(x) = 0$

$A_n = \{X \leq n\}$

$A_1 \subseteq A_2 \subseteq A_3 \subseteq \cdots$

$A = \bigcup_i A_i = \Omega$

$P(A) = \lim_{n \to \infty} P(A_n) = 1 \quad \Big| \; P(A_n) = F(n)$

So, $\lim_{x \to \infty} F(x) = 1.$

(b) $A = \{X \leq x\}, \; B = \{X \leq y\}$

$A \subseteq B$

$\Rightarrow F(x) = P(A) \leq P(B) = F(y)$

Lecture-7

15/10/19

(c) $\lim_{n\to\infty} F\left(x+\frac{1}{n}\right)$

$$= \lim_{n\to\infty} P\left(X \leq x+\frac{1}{n}\right)$$

$$= P\left(\bigcap_{n=1}^{\infty} \left\{X \leq x+\frac{1}{n}\right\}\right)$$

$$= P(X \leq x)$$

$$= F(x) \quad \blacksquare$$

Note: ① CDF (Cumulative distribution function ≡ distribution) is not necessarily continuous.

② Any function satisfying properties (a)–(c) is the CDF of some r.v. X.

Examples:

① Constant r.v.

$$X(\omega) = c \quad \forall \omega \in \Omega$$

$$F(x) = \begin{cases} 0 & \text{if } x < c \\ 1 & \text{if } x \geq c \end{cases}$$

② Bernoulli r.v. (coin with bias p)

$X : \Omega \to \mathbb{R}$ s.t. $X(H) = 1$, $X(T) = 0$

$$F(x) = \begin{cases} 0 & x < 0 \\ 1-p & 0 \leq x < 1 \\ 1 & x \geq 1 \end{cases}$$

③ Indicator r.v.

A be any event

$$I_A(\omega) = \begin{cases} 1 & \text{if } \omega \in A \\ 0 & \text{else} \end{cases}$$

I_A = Bernoulli r.v. taking values 1 and 0 with probabilities $P(A)$ & $P(A^c)$, respectively.

In general, $\{B_i, i \in I\}$ disjoint set of events & $A \subseteq \bigcup_{i \in I} B_i$

$$I_A = \sum_{i \in I} I_{A \cap B_i}$$

Lemma (do your own proof):

a) $P(X>x) = 1 - F(x)$

b) $P(x < X \leq y) = F(y) - F(x)$

c) $P(X = x)$

$$= F(x) - \lim_{y \uparrow x} F(y)$$

Example:

① Let X be a r.v. with distribution

$$F(x) = \begin{cases} 0 & x < -1 \\ 1-p & -1 \leq x < 0 \\ 1-p+\frac{1}{2}xp & 0 \leq x \leq 2 \\ 1 & x > 2 \end{cases}$$

Find

(a) $P(X=-1) = 1-p$

(b) $P(X=0) = (1-p)-(1-p) = 0$

(c) $P(X \geq 1) = 1 - P(X<1)$

$$= 1-\left(1-p+\frac{p}{2}\right) = \frac{p}{2}$$

② Suppose X has distribution F.

What is the distribution of

$$Y = aX+b, \quad a,b \in \mathbb{R}?$$

Soln Case $a>0$

$$P(Y \leq y) = P(aX+b \leq y)$$

$$= P\left(X \leq \frac{y-b}{a}\right) = F\left(\frac{y-b}{a}\right)$$

Case $a < 0$

$$P(Y \leq y) = P(aX + b \leq y)$$

$$= P\left(X \geq \frac{y-b}{a}\right)$$

$$= 1 - P\left(X < \frac{y-b}{a}\right)$$

$$= 1 - \lim_{x \uparrow \frac{y-b}{a}} F(x)$$

Case $a = 0$

$$P(Y \leq y) = \begin{cases} 0 & \text{if } y < b \\ 1 & \text{if } y \geq b \end{cases}$$

A few subtleties:

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

F is a distribution (Verify!)

F has no jumps $\Rightarrow$ F is the CDF of a r.v. that is not discrete

---

Conceptual difficulties in uncountable spaces.

$\Omega = [0,1]$ uncountable space.

Draw a number at random in $[0,1]$

No interval is preferrable

$$P((a,b)) = b-a$$

Ω is uncountable, so consider basic subsets & assign probabilities.

For complex subsets, use the axioms ($\cap^n$, $\cup^n$, complement) & apply the rules of probability.

Example: $P((a,b)) = b-a$

$$P([a,b])$$

$$= P\left(\bigcup_{k=1}^{\infty}\left(a-\frac{1}{k}, b-\frac{1}{k}\right)\right)$$

$$= \lim \left(b-a-\frac{2}{k}\right)$$

$k \to \infty$

$$= (b-a)$$

$$P(\{x\}) = 0$$

But, if $P((a,b)) = \sqrt{b-a}$

$$P((0,1)) = 1$$

$$(0,1) = (0,0.5) \cup (0.5,1) \cup \{0.5\}$$

$$P((0,0.5)) + P((0.5,1)) + P(\{0.5\})$$

$$= \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}} + 0 = \sqrt{2} > 1$$

"Cannot assign arbitrary probabilities to basic subsets".

Alternative: Use distribution.

Example:- Pick any distribution F.

Assign $P((a,b]) = F(b) - F(a)$, $a < b$

Lecture-8 Law of averages

Repeated experimentation

N trials, A occurs $N(A)$ times

$$\frac{N(A)}{N} \longrightarrow P(A) \text{ as } N \to \infty$$

Is the theory built so far consistent with the above requirement?

Suppose $A_1, A_2, \dots$ sequence of independent events with $P(A_i) = p$, $\forall i$, $0 < p < 1$.

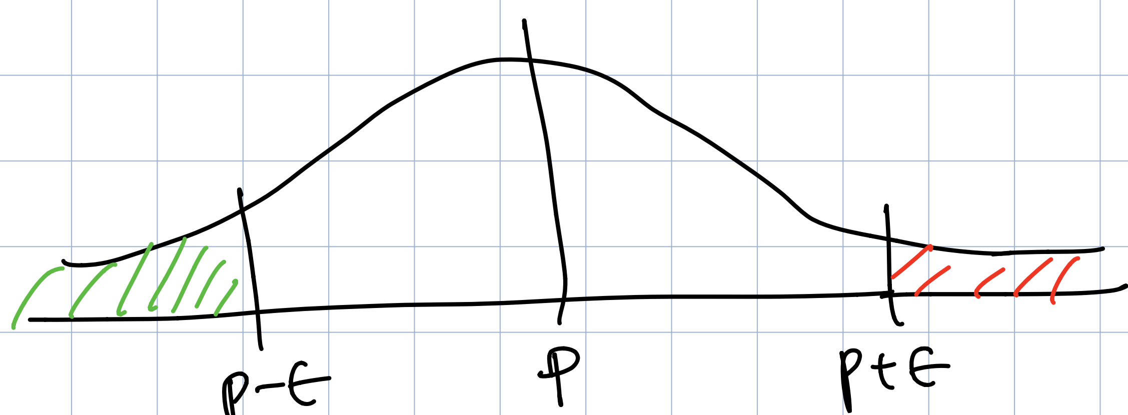
$$S_n = \sum_{i=1}^{n} I_{A_i}$$

Theorem (Bernoulli, 1692)

$\frac{S_n}{n} \to p$ as $n \to \infty$ in the following sense:

$$P\left(p - \epsilon \le \frac{S_n}{n} \le p + \epsilon \right) \to 1 \text{ as}$$

$n \to \infty$, $\forall \epsilon > 0$



Proof: Toss a coin repeatedly, heads occurs with probability p.

S_n has the same distribution as the nuber of heads H_n which occur during first n trials.

$$P(S_n = k) = P(H_n = k) \quad \forall k$$

$$P\left(\frac{S_n}{n} > p+\epsilon\right)$$

$$\leq \sum_{k \geq n(p+\epsilon)} P(H_n = k)$$

$$P(H_n = k) = \binom{n}{k} p^k q^{n-k} \quad (q = 1-p)$$

$$0 \leq k \leq n$$

$$P\left(\frac{S_n}{n} > p+\epsilon\right) \leq \sum_{k=m}^{n} \binom{n}{k} p^k q^{n-k},$$

$$m = \lceil n(p+\epsilon) \rceil$$

$\lambda > 0$, $\quad e^{\lambda k} \geq e^{\lambda n(p+\epsilon)}$, if $k \geq m$

So,

$$P\left(\frac{S_n}{n} > p+\epsilon\right) \leq \sum_{k=m}^{n} e^{\lambda(k-n(p+\epsilon))} \binom{n}{k} p^k q^{n-k}$$

$$= e^{-\lambda n\epsilon} \sum_{k=m}^{n} \binom{n}{k} \left(pe^{\lambda q}\right)^k \left(qe^{-\lambda p}\right)^{n-k}$$

$$\boxed{\begin{aligned} &\lambda q k - \lambda p(n-k) \\ &= \lambda k - \lambda p n \end{aligned}}$$

Binomial theorem

$$\leq e^{-\lambda n\epsilon} \left(pe^{\lambda q} + qe^{-\lambda p}\right)^n$$

Use $\boxed{e^x \leq x + e^{x^2}}$ to get

$$P\left(\frac{S_n}{n} > p+\epsilon\right) \le e^{-\lambda n\epsilon}\left(pe^{\lambda^2 q^2} + qe^{\lambda^2 p^2}\right)^n$$

$$\le e^{\lambda^2 n - \lambda n \epsilon}$$

Optimizing over λ, we obtain $\lambda^* = \frac{\epsilon}{2}$

and

$$P\left(\frac{S_n}{n} > p+\epsilon\right) \le e^{-n\epsilon^2/4}$$

So, $P\left(\frac{S_n}{n} > p+\epsilon\right) \to 0$ as $n \to \infty$.

Similarly,

$$P\left(\frac{S_n}{n} < p-\epsilon\right) \le e^{-n\epsilon^2/4}$$

& hence $\to 0$ as $n\to\infty$

$\blacksquare$

$$P\left(p-\epsilon \le \frac{S_n}{n} \le p+\epsilon \,,\ \forall\, n \ge m\right)$$

$$\to 1 \quad \text{as} \quad m \to \infty$$

Lecture-9

Law of averages (contd)

Proof of $\quad e^x \le x + e^{x^2}$

$$f(x) = x + e^{x^2} - e^x$$

$$f'(x) = 1 + 2x e^{x^2} - e^x$$

$$f''(x) = 2e^{x^2} + 4x^2 e^{x^2} - e^x$$

$$= 2(1+2x^2)e^{x^2} - e^x$$

$$\ge e^x\left(2e^{x^2-x} - 1\right)$$

$$\ge e^x\left(2e^{-1/4} - 1\right)$$

$$> 0$$

since $x^2 - x \ge -\frac{1}{4}$

$f'' > 0 \implies f$ is convex

$f'(0) = 0 \implies f$ is minimized at 0

$\implies f(x) \geq f(0)$

$\implies x + e^{x^2} \geq e^x$

$\blacksquare$

Earlier theorem says, for a given n, $\frac{S_n}{n}$ lies between $p-\epsilon$ & $p+\epsilon$ with high probability.

$$A_n = \left\{ p - \epsilon \leq \frac{S_n}{n} \leq p + \epsilon \right\}$$

$$P\left(\bigcap_{n=m}^{\infty} A_n \right)$$

Consider $\bigcup_{n=m}^{\infty} A_n^c$

$$P\left(\bigcup_{n=m}^{\infty} A_n^c\right) \leq \sum_{n=m}^{\infty} P(A_n^c)$$

$$\leq \sum_{n=m}^{\infty} 2e^{-n\epsilon^2/4}$$

$$\xrightarrow[\text{as } m\to\infty]{} 0$$

$$P\left(p-\epsilon \leq \frac{S_n}{n} \leq p+\epsilon, \ \forall n \geq m\right) \to 1$$

as $m\to\infty$

Types of random variables:

(1) A r.v. X is discrete if it takes values in some countable subset $\{x_1, x_2, \ldots\}$ of $\mathbb{R}$.

Discrete r.v. X has probability mass function $f : \mathbb{R} \to [0,1]$ given by $f(x) = P(X = x)$

<u>Note</u>: Distribution of a discrete r.v. has jump discontinuities at $x_1, x_2, \ldots$ & is a constant in between.

x_1 x_2 ---- x_n

(2) A r.v. X is continuous if its distribution function can be expressed as

$$F(x)=\int_{-\infty}^{x} f(u)\,du, \quad \forall x\in\mathbb{R}, \quad (*)$$

for some integrable function

$f:\mathbb{R}\to[0,\infty)$.

f is the probability density function of X

H.W. (1) Check $(*)$ is a valid distribution

(2) $P(X=x)=0$ for a continuous r.v. X.

Random vectors:

For a r.v. X, we had $F_X(x) = P(X \le x)$

For a pair (X, Y) of r.v.s,

one defines their joint distribution as

$$F_{X,Y}(x, y) = P(X \le x, Y \le y)$$

In general, the joint distribution of a random vector $(X_1, \dots, X_n)$

is $P(X_1 \le x_1, \dots\dots, X_n \le x_n)$

Lemma: The joint distribution $F_{X,Y}$ of (X, Y) satisfies

① $\lim\limits_{x,y \to -\infty} F_{X,Y}(x, y) = 0$

$$\lim_{x,y\to\infty} F_{X,Y}(x,y) = 1$$

(2) If $(x_1, y_1) \le (x_2, y_2)$,

then $F_{X,Y}(x_1,y_1) \le F_{X,Y}(x_2,y_2)$

(3) $F_{X,Y}$ is continuous from the above

$$F_{X,Y}(x+u,\ y+v) \to F_{X,Y}(x,y)$$

as $u, v \downarrow 0$

It may be seen without great difficulty that

$$\lim_{y\to\infty} F_{X,Y}(x,y) = F_X(x)\left(= P(X\le x)\right)$$

and

$$\lim_{x \to \infty} F_{X,Y}(x,y) = F_Y(y) \; (= P(Y \le y))$$

Note:

Given the joint distribution $F_{X,Y}$, it is possible to find the marginals F_X, F_Y.

The converse is not true in general.

Example

Flip a fair coin twice & record the outcomes

$$(X_D, Y_D)$$

Flip the coin once & record the same outcome twice

$$(X_L, Y_L)$$

X_D, Y_D, X_L, Y_L have the same distribution

$$P(X_D = Y_D = \text{heads}) = \frac{1}{4}$$

$$P(X_L = Y_L = \text{heads}) = \frac{1}{2}$$