Random Variables

A **random variable** X on a probability space (Ω, p)

if a function $X : \Omega \to \mathbb{R}$.

$\{X = a\}$ is the event $\{\omega \in \Omega \mid X(\omega) = a\}$

The **expectation** of a discrete r.v. X is defined as

$$E(X) = \sum_{a \in \underbrace{A}_{\substack{\text{set of values} \\ X \text{ can take}}}} a \cdot Pr(X = a)$$

(range of X)

Example:

$n = 2\ell$ packets sent over Internet.

Consider 3 models for packet loss

each packet lost independently with prob p $\qquad X_1$: # packets lost

all n packets on same path & with probability p some link fails $\Rightarrow$ all packets lost. O.w. all received $\qquad X_2$: # packets lost

divide into 2 equal size groups

each lost independently with prob p. $\qquad X_3$: # packets lost

$$\Pr(X_1 = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

$$\Pr(X_3 = k) = \begin{cases} (1-p)^2 & k = 0 \\ 2p(1-p) & k = \ell = \frac{n}{2} \\ p^2 & k = n \end{cases}$$

$$\Pr(X_2 = k) = \begin{cases} 1-p & k = 0 \\ 0 & 0 < k < n \\ p & k = n \end{cases}$$

$E(X_2) = 0 \cdot (1-p) + n \cdot p \;= np$

$E(X_3) = 0 \cdot (1-p)^2 + \frac{n}{2} \cdot 2p(1-p) + n\,p^2 \;= np$

$E(X_1) = \sum_{k=0}^{n} k \binom{n}{k} p^k (1-p)^{n-k}$

Random variable X_1 very important type of r.v.

$\boxed{\text{Binomial r.v. with params } n, p}$

defined by 2 parameters n, p

X_1 : # of Heads in n indep coin tosses each with probability p of coming up H.

Linearity of expectation

Thm: $\forall$ 2 random vars X & Y on same probability space,

$$E(X+Y) = E(X) + E(Y)$$

Proof

Observe: $E(X) = \sum_{a \in \underbrace{A}_{\text{set of values r.v. takes}}} a \cdot Pr(X=a)$

$$= \sum_{a \in A} a \cdot \sum_{w \mid X(w)=a} Pr(w)$$

$w \in \Omega$ s.t. $X(w)=a$

$$= \sum_{w \in \Omega} X(w) Pr(w)$$

$$E(X+Y) = \sum_{w \in \Omega} [X(w) + Y(w)] Pr(w) = \sum_{w \in \Omega} X(w) Pr(w) + \sum_{w \in \Omega} Y(w) Pr(w)$$

$$= E(X) + E(Y)$$

and by induction on n

Corollary $X_1, X_2, .., X_n$ r.v.'s on same prob space

then $E(X_1+X_2+..+X_n) = E(X_1)+E(X_2)+\ldots+E(X_n)$

☆ ☆ Super important!

Also $\boxed{E(cX) = c\,E(X) \quad \text{for } c \text{ a constant}}$ check this!

Not true for other operations

$E(XY) \neq E(X)E(Y)$ in general

$E\left(\frac{1}{X}\right) \neq \frac{1}{E(X)}$ in general

etc.

Importance of linearity of expectation:

say you want to compute $E(X)$ for some r.v. X

break it up into simpler r.v.'s

$X = X_1 + X_2 + .. + X_n$

compute each $E(X_i)$ and then sum.

$E(X)$ $\quad X \sim Bin(n,p)$ $\quad$ binomial params n & p

X: # H's in n indep coin tosses, prob p of heads

$X = X_1 + X_2 + .. + X_n$

where $X_i = \begin{cases} 1 & \text{if } i^{th} \text{ toss H's} \\ 0 & \text{if } i^{th} \text{ toss T's} \end{cases}$

X_i is called an $\boxed{\text{indicator r.v.}}$ or $\boxed{\text{Bernoulli r.v.}}$

indicates success/failure

H/T

etc

$E(X_i) = 0 \cdot (1-p) + 1 \cdot p = p$

$$\boxed{E(X) = \sum_{i=1}^{n} E(X_i) = np}$$

n homeworks returned to n students shuffled randomly

X: # students that receive own homework

$$E(X) = \sum_{k=0}^{n} k \underbrace{\Pr(X=k)}_{\text{complicated!}}$$

use linearity of expectation

$$X_i = \begin{cases} 1 & \text{if } i^{th} \text{ student receives their HW back} \\ 0 & \text{o.w.} \end{cases}$$

$$X = \sum_{i=1}^{n} X_i \quad \underset{\text{Lin of Exp}}{\Longrightarrow} \quad E(X) = \sum_{i=1}^{n} E(X_i)$$

$$E(X_i) = 0 \cdot \left(1 - \tfrac{1}{n}\right) + 1 \cdot \tfrac{1}{n} = \tfrac{1}{n} \quad ; \Rightarrow E(X) = 1$$

A deep idea: What might be surely possible

- n prisoners at a jail
- of $\binom{n}{2}$ possible pairs, k are risky

- cafeteria can hold all n prisoners, but to reduce chance of fight, schedule 2 lunches

  A: 12-1 B: 1-2

Theorem For any set R of risky pairs, it is possible to assign the prisoners to lunch slots so that at least $\frac{1}{2}$ the risky pairs are broken up.

Prove this using "the probabilistic method"

- assign each prison at random → A or B
- argue that with positive probability, at least $\frac{1}{2}$ risky pairs broken up.

⇒ $\exists$ way to achieve this!

$\Omega = \{A, B\}^n$ (for each prisoner, which lunch)

p: uniform dist'n

X: # risky pairs broken up

Claim: $E(X) = \frac{|R|}{2}$ ⇒ $\Pr(\text{at least } \frac{|R|}{2} \text{ pairs broken}) > 0$

⇒ $\exists$ way

$$X = \sum_{(i,j) \in R} X_{ij} \qquad X_{ij} = \begin{cases} 1 & i \,\&\, j \to \text{diff lunch slots} \\ 0 & \text{otherwise} \end{cases}$$

$$E(X_{ij}) = \frac{1}{2}$$

$$E(X) = \sum_{(i,j) \in R} E(X_{ij}) = \frac{|R|}{2} \Rightarrow \exists\, \omega \text{ s.t. } X(\omega) \geq \frac{|R|}{2}$$