Problem

There is a population of N people $\sum_{i=0}^{N} p_i = 1$

$$Pr(i \text{ good guys in pop}) = p_i \qquad i = 0, \ldots, N$$

Take a sample of n people. Each subset equally likely.

$$Pr(j \text{ good guys in pop} \mid k \text{ good guys in sample}) ?$$

Solution:

$\forall\, 0 \le i \le N$ define

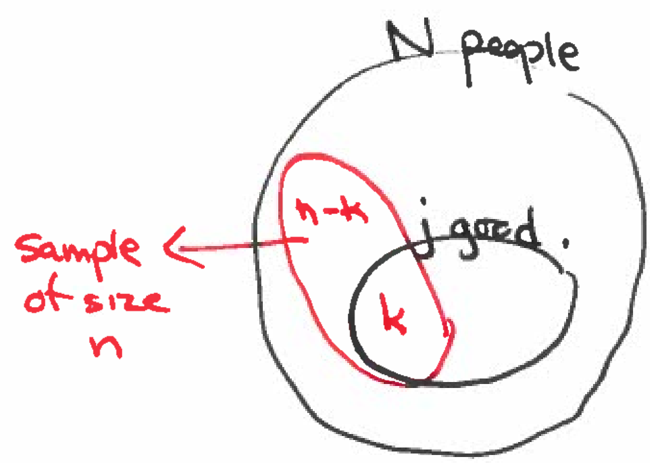
G_i : event that there i good guys in pop

S_i : event " " i " " in sample.

$$Pr(G_j \mid S_k) = \frac{Pr(G_j \cap S_k)}{Pr(S_k)} = \frac{Pr(S_k \mid G_j) Pr(G_j)}{Pr(S_k)}$$

Numerator: $Pr(G_j) = p_j$

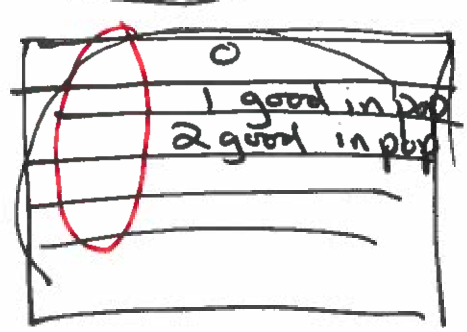
$$Pr(S_k \mid G_j) = \frac{\binom{j}{k}\binom{N-j}{n-k}}{\binom{N}{n}}$$



Denominator:

$$Pr(S_k) = \sum_{i=0}^{N} Pr(S_k \cap G_i)$$

$$= \sum_{i=0}^{N} Pr(S_k \mid G_i) Pr(G_i)$$



Random Variables

Oct 17, 2016

(S, p) prob space

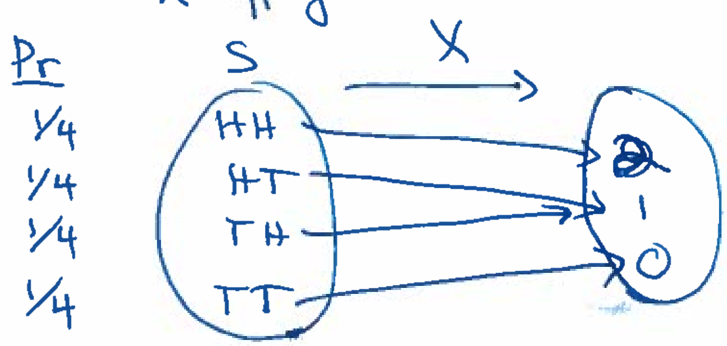
A <u>random variable</u> on (S, p) is a function

$$X : S \longrightarrow \mathbb{R}$$

Event $\{X = a\} = \{\omega \mid X(\omega) = a\}$

<u>Ex</u> 2 coin tosses equally likely outcomes

X : # of heads



$Pr(X=2) = Pr(HH) = \frac{1}{4}$

$Pr(X=1) = Pr(HT, TH) = \frac{1}{2}$

<u>Expectation of a r.v. X</u>, denoted $E(X)$

$$E(X) = \sum_{k \in Range(X)} k \, Pr(X=k) = \sum_{\omega \in S} X(\omega) Pr(\omega)$$

For example above

$$E(X) = 0 \cdot Pr(X=0) + 1 \cdot Pr(X=1) + 2 \cdot Pr(X=2)$$
$$= 0 \cdot \frac{1}{4} + 1 \cdot \frac{1}{2} + 2 \cdot \frac{1}{4} = 1$$

$\equiv$

$$E(X) = X(HH)Pr(HH) + X(HT)Pr(HT) + X(TH)Pr(TH) + X(TT)Pr(T$$
$$= 2 \cdot \frac{1}{4} + 1 \cdot \frac{1}{4} + 1 \cdot \frac{1}{4} + 0 \cdot \frac{1}{4}$$

Example:

homeworks of 3 people returned according to random permutation

Pr	S	$X(\omega)$
1/6	1 2 3	3
1/6	1 3 2	1
1/6	2 1 3	1
1/6	2 3 1	0
1/6	3 1 2	0
1/6	3 2 1	1

X # people who get their own homework back

$$Pr(X=3) = \frac{1}{6}$$

$$Pr(X=1) = \frac{3}{6} = \frac{1}{2}$$

$$Pr(X=0) = \frac{2}{6} = \frac{1}{3}$$

Example

$n = 2\ell$ packets sent over Internet

3 models for packet loss

Model I:

each packet lost independently with prob p

X_1 # of packets lost

X_1 takes values in $\{0, 1, 2, \dots, n\}$

$$\Pr(X_1 = k) = \Pr(\text{outcome where } k \text{ packets lost})$$

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

L	L	N	N	L	L
1					n

Model II: all lost with prob p

none lost with prob $1-p$.

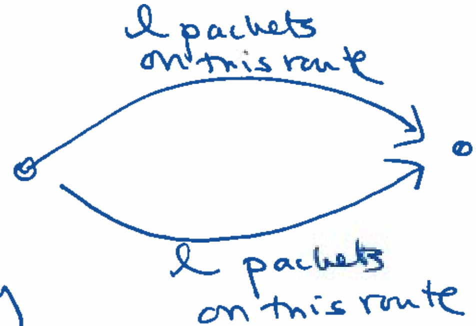
X_2 # of packets lost

X_2 takes values in $\{0, n\}$

$$\Pr(X_2 = 0) = 1 - p$$

$$\Pr(X_2 = n) = p$$

Model III :



each route "breaks" with prob p independently

X_3 # of packets lost

X_3 takes values in $\{0, \ell, n\}$

$$\Pr(X_3 = 0) = (1-p)^2$$

$$\Pr(X_3 = \ell) = 2p(1-p)$$

$$\Pr(X_3 = n) = p^2$$