CS2 107 2-27-24

- midterm 2: Thur Feb. 29
- skip 4.4

## 4.5 sum of a random # of r.v.s

Let $X_1, X_2, X_3, \ldots$ be an infinite seq. of independent, identically distributed r.v.s abbreviated: (iid)

This means they all have the same PDF, (or PMF) which we write

$$f_X(x) \quad \text{or} \quad P_X(x)$$

Let $N$ be a r.v. with values in positive integers $\{1, 2, 3, \dots\}$.

Assume: $\{N, X_1, X_2, \dots\}$ is independent. Now define

$$Y = X_1 + X_2 + \dots + X_N$$

we denote by $E[X]$ and $Var(X)$ the common mean & variance of all $X_i$.

Goal: find $E[Y]$ and $Var(Y)$.

<u>Note</u>:

$$E[Y|N=n]$$
$$= E[X_1+X_2+\cdots+X_N \mid N=n]$$
$$= E[X_1+X_2+\cdots+X_n \mid N=n]$$
$$= E[X_1+X_2+\cdots+X_n] \quad \{\text{by indep.}$$
$$= n\,E[X]$$

Replace $n$ by $N$ to get

$$E[Y|N] = N \cdot E[X]$$

By law of iterated expectation

$$E[Y] = E\big[E[Y|N]\big]$$

$$= E[N \cdot E[X]]$$

$$= E[N] \cdot E[X] \leftarrow$$

Also

$$\mathrm{Var}(Y \mid N=n)$$

$$= \mathrm{Var}(X_1 + X_2 + \cdots + X_N \mid N=n)$$

$$= \mathrm{Var}(X_1 + \cdots + X_n \mid N=n)$$

$$= \mathrm{Var}(X_1 + \cdots + X_n) \quad \text{by indep.}$$

$$= \mathrm{Var}(X_1) + \cdots + \mathrm{Var}(X_n) \quad \text{by indep.}$$

$$= n\,\mathrm{Var}(X).$$

$$\therefore \quad \mathrm{Var}(Y \mid N) = N \cdot \mathrm{Var}(X)$$

By the law of total variance

$$\mathrm{Var}(Y) = E[\mathrm{Var}(Y|N)] + \mathrm{Var}(E[Y|N])$$

$$= E[N\,\mathrm{Var}(X)] + \mathrm{Var}(N\,E[X])$$

$$= E[N]\cdot \mathrm{Var}(X) + E[X]^2 \cdot \mathrm{Var}(N)$$

<u>EX.</u> Suppose have 2 coins

coin 1: $P(\text{head}) = p$

coin 2: $P(\text{head}) = q$

Flip coin 1 until $1^{st}$ head. let $N$ be # flips. note $N$ is geometric with param. $p$.

Flip coin 2 $N$ times and count # heads. let $Y$ be # heads in $2^{nd}$ sequence.

observe

$$Y = X_1 + X_2 + \cdots + X_N$$

where each $X_i$ is Bernoulli with parameter $q$. also the $X_i$ are indep. of each other and of $N$.

let $E[X]$, $Var(X)$ denote their common mean & var.

we have

$$E[N] = \frac{1}{p}, \quad Var(N) = \frac{1-p}{p^2}$$

and

$$E[X] = q, \quad Var(X) = q(1-q).$$

Thus

$$E[Y] = E[N] \cdot E[X] = \frac{q}{p}$$

and

$$\mathrm{Var}(Y) = E[N]\,\mathrm{Var}(X) + E[X]^2\,\mathrm{Var}(N)$$

$$= \frac{q(1-q)}{p} + \frac{q^2(1-p)}{p^2}$$

# 6.1 The Bernoulli Process

**Defn**

The Bernoulli Process is a sequence

$$X_1, X_2, \ldots\ldots$$

of indep. Bernoulli r.v.s with

$P(X_t = 1) = p$ (heads, success, arrival)

$P(X_t = 0) = 1 - p$ (tails, failure, non-arrival) we can think of this as

- one experiment performed $\infty$-many times: $\Omega = \{0, 1\}$

or

- a single experiment with $\Omega = \{\text{infinite bit strings}\}$

Have 2 Basic questions:

① given # of trials $n$, how many successes?

answer: Binomial $(n, p)$

(2) how many trials until 1st success ?

answer : Geometric ( $p$ ) .

## Memorylessness

Let $T$ be time # trials to 1st success. Then $T$ is Geometric($p$), i.e.

$$P_T(t) = (1-p)^{t-1} p$$

$$E[T] = \frac{1}{p} \quad , \quad Var(T) = \frac{1-p}{p^2}$$

we observe

$$X_1, X_2, \dots, X_n$$

are all failures. Then the event $\{T > n\}$ has occurred.

What is distribution of $T-n$, the time left until 1st head.

$$P(T-n=t \mid T>n) = \frac{P(T-n=t,\ T>n)}{P(T>n)}$$

$$= \frac{P(T=n+t,\ T>n)}{P(T>n)}$$

$$= \frac{P(T = n+t)}{P(T > n)} \quad \begin{cases} \text{if } A \subseteq B \\ \text{then } A \cap B = A \end{cases}$$

$$= \frac{(1-p)^{n+t-1} \cdot p}{(1-p)^{n}}$$

$$= \frac{\cancel{(1-p)^{n}} \cdot (1-p)^{t-1} \cdot p}{\cancel{(1-p)^{n}}}$$

$$= (1-p)^{t-1} \cdot p$$

$$= P(T = t) = p_T(t)$$

as expected, independence implies memorylessness.

# Inter-arrival times

let

$T_1$ = time to $1^{st}$ arrival

$T_2$ = from $1^{st}$ arrival to $2^{nd}$

$\vdots$

$T_K$ = " $(K-1)^{th}$ " " $K^{th}$

Then

$$Y_K = T_1 + T_2 + \cdots + T_K$$

$= $ time to $K^{th}$ arrival