RV Key Concepts:

- a RV is a real-valued func of the outcome of an exp
- a func of a RV is another RV
- we can associate w/ a RV parameters/attributes like means & variances
- a RV can be conditioned on an event or another RV
- notion of independence of a RV from another RV or event

Ex: Chess match

Anand plays Kasparov: the 1st to win a game wins the match; match is drawn if 10 consecutive draws

$P(\text{A wins a game}) = 0.3$

$P(\text{K wins a game}) = 0.4$

$P(\text{draw}) = 0.3$

1/ What's PMF of the duration of the match L?

a) $L=10$ iff 1st 9 games drawn

$\Rightarrow P_L(10) = (0.3)^9$

b) $L=k<10$ if 1st $k-1$ are drawn & kth has a result

$\Rightarrow P_L(k) = (0.3)^{k-1} \cdot 0.7$ for $k=1,\dots,9$

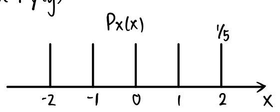
2/ $P(\text{A wins match}) = \sum_{k=0}^{9} (0.3)^k (0.3)$

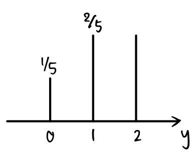
Functions of a RV

Let $P_X(x) = \frac{1}{5}$; $x = -2,-1,0,1,2$

Let $Y = |X|$.

Find $P_Y(y)$





$\boxed{\text{If } Y=g(X),\ P_Y(y) = \sum P_X(x)\ \{x : g(x)=y\}}$

Some popular discrete RVs:

① Bernoulli - $B(p)$ or $Bern(p)$

$X = \begin{cases} 0 & \text{if tails} \\ 1 & \text{if heads} \end{cases}$

(coin toss)

② Binomial - $Binom(n,p)$

$P_X(k) = \binom{n}{2} p^k (1-p)^{n-k}$, $k=0,1,2,\dots$

(# of successes, i.e. heads, in n indep coin tosses)

Note: $\sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} = 1$ (Binomial Formula)

③ Geometric - $Geom(p)$

$P_X(k) = (1-p)^{k-1} \cdot p$, $k=1,2,\dots$

(# of tosses until 1st success)

④ Poisson - $Pois(\lambda)$

$P_X(k) = \frac{e^{-\lambda}\lambda^k}{k!}$, $k=0,1,2,\dots$

limiting distr of the $Bin(n,p)$ as $n\uparrow$, $p\downarrow$, and $np=\lambda$

⑤ Uniform - Unif$\{1,2,\dots,n\}$

$P_X(k) = \frac{1}{n}$

equally likely to take on any value

Expectation

- "Summary" of your distr
- center of distr

Def: $\boxed{\mathbb{E}[X] = \sum_{w\in\Omega} X(w)P(w) \qquad \mathbb{E}[X] = \sum_{x} x \cdot P(X=x)}$

Remark: $\sum_{x\in X} |x| P_X(x) < \infty$ to be well-defined

! **Theorem (LOE):** For any 2 RVs X, Y defined on the same prob space,

(i) $\mathbb{E}[X+Y] = \mathbb{E}[X] + \mathbb{E}[Y]$

(ii) $\mathbb{E}[cX] = c\,\mathbb{E}[X]$

Proof: Follows from linearity of sums

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

$= \sum_{w} X(w)P(w) + \sum_{w} Y(w)P(w)$

$= \mathbb{E}[X] + \mathbb{E}[Y]$

Def: Indicator RV

Let $A \subseteq \Omega$ be an event. We define the indicator of A, $\mathbb{1}_A$ or $\mathbb{1}\{A\}$

$$\mathbb{1}_A(w) = \begin{cases} 0; & w \notin A \\ 1; & w \in A \end{cases}$$

Observe that $\mathbb{1}_A$ follows Bern(p) where $p = P(A)$.

* An important consequence of indicator RVs is that $X^2 = X$, i.e. $X^k \ \forall k \geq 1$ has the same distr as X.

$\boxed{\mathbb{E}[\mathbb{1}_A] = P(A)}$

Ex: Seats on a plane (Hat matching)

Full flight, n passengers w/ assigned seats but passengers sit randomly

Q/ $\mathbb{E}$[# of passengers who sit in assigned seats]

Let X = # of passengers who sit in assigned seats

$$X_i = \begin{cases} 1 & \text{if passenger } i \text{ sits in their seat} \\ 0 & \text{else} \end{cases}$$

$$X = \sum_{i=1}^{n} X_i = X_1 + \dots + X_n$$

$\mathbb{E}[X] = \mathbb{E}[X_1] + \dots + \mathbb{E}[X_n]$ Note: NO independence, but linearity of exp still holds

$= n\mathbb{E}[X_1]$

$= nP(X_1 = 1)$

$= n\left(\frac{1}{n}\right)$

$= 1$

$$\mathbb{E}[f(X)] = \sum_{w\in\Omega} f(X(w))\cdot P(w)$$

$$= \sum_{x} f(x)\, P(X=x)$$

extended:

$$\mathbb{E}[f(X_1,\dots,X_n)] = \sum_{x_1,\dots,x_n} f(x_1,\dots,x_n)\, P_{X_1,\dots,X_n}(x_1,\dots,x_n)$$

Theorem: Expectation of indep RVs

Def: X, Y are indep RVs iff $P(X=x)P(Y=y) = P(X=x, Y=y)$

$$\mathbb{E}[XY] = \mathbb{E}[X]\,\mathbb{E}[Y]$$

Proof: $\mathbb{E}[XY] = \sum_{x,y} xy\, P_{XY}(x,y)$

$$= \sum_{x} x\cdot P_X(x)\cdot \sum_{y} y\cdot P_Y(y)$$

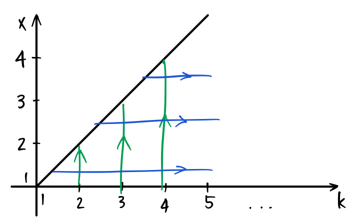
Tail Sum Formula

Theorem: Let X be a RV that takes values only in $\mathbb{N}$, then, $\mathbb{E}[X] = \sum_{x=1}^{\infty} P(X\geq x)$

Proof: $\mathbb{E}[X] = \sum_{k=1}^{\infty} k P(X=k)$

$$= \sum_{k=1}^{\infty} \underbrace{\sum_{x=1}^{k} 1}_{x\leq k}\, P(X=k)$$ ① count vertically

$$= \sum_{x=1}^{\infty} \underbrace{\sum_{k=x}^{\infty} P(X=k)}_{P(X\geq x)}$$ ② count horizontally



Revisit Popular Distributions

① Bern(p):

$$\boxed{\mathbb{E}[X] = p}$$

② Bin(n,p):

$$\mathbb{E}[X] = \sum_{x=0}^{n} \binom{n}{x} p^x (1-p)^{n-k}$$

$X = X_1 + \dots + X_n$ where $X_i = 0$ or 1, $X_i \sim \text{Bern}(p)$

$$\boxed{\mathbb{E}[X] = \sum \mathbb{E}[X_i] = np}$$

③ Geom(p):

Note: $\sum_{x=1}^{\infty} (1-p)^{x-1}\cdot p = 1$

$P(X=x) = (1-p)^{x-1}\cdot p$

$P(X>x) = (1-p)^{x}$ fail x times

↳ CDF (cumulative distribution function): $P_X(x) = P(X\leq x) = 1 - P(X>x)$

$$\boxed{\mathbb{E}[X] = \frac{1}{p}}$$

① Use TSF: $\mathbb{E}[X] = \sum_{x=1}^{\infty} P(X \geq x)$ } change bounds

$= \sum_{x=0}^{\infty} P(X > x)$

$= \sum_{x=0}^{\infty} (1-p)^x$

$= \frac{1}{p}$

Another way is to use memoryless property:

Memoryless Property:

The Geom distr satisfies $\boxed{P(X > s+t \mid X > s) = P(X > t)}$

Proof: $P(X > s+t \mid X > s) = \frac{P(X > s+t, X > s)}{P(X > s)}$

$= \frac{P(X > s+t)}{P(X > s)}$

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

$= (1-p)^t$

$= P(X > t)$

Moral: Get no breaks for failing the first s times.

② Use memorylessness:

$$\mathbb{E}[X] = \underbrace{\mathbb{E}[X \mid X=1]}_{1} \cdot \underbrace{P(X=1)}_{p} + \underbrace{\mathbb{E}[X \mid X>1]}_{1+\mathbb{E}[X]} \cdot \underbrace{P(X>1)}_{1-p}$$

renewal / memorylessness

already failed once

$\mathbb{E}[X] = p + (1-p)(1 + \mathbb{E}[X])$

$\Rightarrow \mathbb{E}[X] = \frac{1}{p}$

③ Brute force:

$\mathbb{E}[X] = \sum_{x=1}^{\infty} x(1-p)^{x-1} p = p \sum_{x=0}^{\infty} (x+1)(1-p)^x$

Recall: $\sum_{k=0}^{\infty} x^k = \frac{1}{1-x}$

$\Rightarrow \sum_{k=0}^{\infty} x^{k+1} = \frac{x}{1-x}$

differentiate wrt x: $\sum_{k=0}^{\infty} (k+1) x^k = \frac{1}{(1-x)^2}$

Set $1-x = p \Rightarrow \mathbb{E}[X] = p \frac{1}{(1-(1-p))^2} = p \cdot \frac{1}{p^2} = \frac{1}{p}$

Ex: Let $X \sim Geom(p)$, $Y \sim Geom(q)$ be indep RVs $\Rightarrow \min(X,Y) \sim Geom(p+q-pq)$

If min of X, Y > z, then both X & Y > z

$P(\min\{X,Y\} > z) = P(X > z, Y > z)$

$= P(X > z)\, P(Y > z)$

$= (1-p)^z (1-q)^z$

$= (1-p-q+pq)^z$

$= (1-(p+q-pq))^z$

↳ tail prob of Geom(p+q-pq)

Ex: Coupon Collector Problem

There are n coupons you want to collect. Each time you buy an item from the store, you collect a random coupon (each is equally likely to appear).

Q/ What is the expected num of items you must buy before you collect every coupon?

Let Ti be the num of items it requires to collect the ith new coupon.

Starting from when you have seen $i-1$ distinct coupons, T_i represents num of items you must buy before you see a new coupon.

$\rightarrow$ Total time to collect all coupons $= T = \sum_{i=1}^{n} T_i$

$\mathbb{E}[T_i]$ = exp time to get ith distinct coupon

Prob you have not seen a coupon on your next buy is $\frac{n-i+1}{n}$.

$T_i \sim \text{Geom}(p)$, where $p = \frac{n-i+1}{n}$

By linearity,

$$\mathbb{E}[T] = \sum_{i=1}^{n} \mathbb{E}[T_i] = \sum_{i=1}^{\infty} \frac{n}{n-i+1}$$
$$= n \sum_{i=1}^{\infty} \frac{1}{i}$$
$$= nH_n \quad \text{where } H_n = \text{nth harmonic sum} \approx \ln n$$

$$\boxed{\mathbb{E}[T] \approx n \ln n}$$

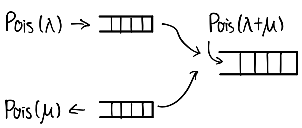
④ $\text{Pois}(\lambda)$:

$$\mathbb{E}[X] = \sum_{x=0}^{\infty} x \frac{e^{-\lambda}\lambda^x}{x!} = \sum_{x=1}^{\infty} \frac{xe^{-\lambda}\lambda^x}{x!}$$
$$= e^{-\lambda}\lambda \underbrace{\sum_{x=1}^{\infty} \frac{\lambda^{x-1}}{(x-1)!}}_{e^{\lambda}\text{ (Taylor)}}$$

$$\boxed{\mathbb{E}[X] = \lambda}$$

Sums of Poisson RVs

Theorem: Let $X \sim \text{Pois}(\lambda)$, $Y \sim \text{Pois}(\mu)$ be indep RVs, then $X+Y \sim \text{Pois}(\lambda+\mu)$



Poisson merging

Theorem (Poisson Splitting): If $X \sim \text{Pois}(\lambda)$ & conditioned on X, $Y \sim \text{Bin}(X, p)$
$\hookrightarrow Y \sim \text{Pois}(\lambda p)$ unconditioned