Lecture

Last time

- random variable: $X : \Omega \to \mathbb{R}$
  - ↳ associated to r.v. X is its distribution
  - → discrete r.v. : distribution <=> $\mathbb{P}_x$ (proof)
- expectation ($\mathbb{E}[X] = \sum_x x \mathbb{P}_x(x)$)
  - → linear operator on r.v.'s
  - ↳ ie, if X, Y r.v's, then $\mathbb{E}[X+Y] = \mathbb{E}[X] + \mathbb{E}[Y]$

    (don't need to compute the PMF of $(X+Y)$ → just need marginal PMFs of X & Y).

Example (continuing from last lec)

$\Omega = \{0,1\}^n$

$P(\omega) = p^{\#\{i : \omega_i = 1\}} (1-p)^{\#\{i : \omega_i = 0\}}$

$X(\omega) = \#\{i : \omega_i = 1\} =$ "# heads" , $X \sim \text{Binomial}(n, p)$

↳ can write X as

$$X = \sum_{i=1}^{n} x_i \qquad \text{s.t.} \quad x_i = \begin{cases} 1 & \text{if flip} = H \\ 0 & \text{o/w} \end{cases}$$

$$\mathbb{E}[X] = \mathbb{E}\left[\sum_{i=1}^{n} x_i\right] = \sum_{i=1}^{n} \mathbb{E}[x_i] = \sum_{i=1}^{n} (p-1)^0 (p)^1 = \sum_{i=1}^{n} p = np$$

General strategy of this example is important

- we introduce "indicator" rv's & express r.v. of interest in terms of indicators, then use properties of expectation
- in general, for $A \in \mathcal{F}$, the indicator is

$$\mathbb{1}_A(\omega) = \begin{cases} 0 & \text{if } \omega \notin A \\ 1 & \text{if } \omega \in A \end{cases}$$

$\mathbb{1}_A \sim \text{Bern}(P(A))$

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

Ex: n people put hats in a bucket & each draw 1 hat (permutation of hats)

X = # people who get their own hat back

↳ w/o replacement

$\mathbb{E}[X] = ?$

$$X = \sum_{i=1}^{n} \mathbb{1}_{A_i} \qquad A_i = \{\text{person } i \text{ gets own hat back}\}$$

these events are NOT independent (bc we're drawing w/o replacement)

$$\mathbb{E}[X] = \sum_{i=1}^{n} \mathbb{E}[\mathbb{1}_{A_i}] = \sum_{i=1}^{n} \underbrace{P(A_i)}_{\frac{1}{n}} = n\left(\frac{1}{n}\right) = 1$$

↳ underlying probability space → all possible permutations

Ex: Coupon collector problem

- Suppose boxes of cereal contain 1 of N distinct coupons.

Q/ How many boxes do I need to purchase before collecting all N coupons?

A/ Lets use LoE!

→ Suppose sequence of 4 coupons is

A A B A B B C C A B C C C A D → $x = 14$ coupons

$x_1 = 1$ $x_2 = 2$ $x_3 = 3$ $x_4 = 7$

$x_i = \{$# boxes to buy to get i^{th} coupon after getting $(i-1)^{th}$ coupon$\}$

$$x_i \sim \text{Geom}\left(\frac{N-(i-1)}{N}\right)$$

if you're on the last coupon, $\frac{1}{N}$ prob. of new coupon. If you're on the first, prob of new is 1.

$$\mathbb{E}[X] = \sum_{i=1}^{n} \mathbb{E}[x_i]$$
$$= \sum_{i=1}^{n} \frac{N}{N-(i-1)}$$
$$= N \underbrace{\sum_{k=1}^{N} \frac{1}{k}}_{\text{harmonic}}$$
$$\approx N\log(N)$$

Another Way of Computing Expectation: Tail Sum Formula

- Tail sum formula: for non-negative integer-valued r.v. X:

$$\mathbb{E}[X] = \sum_{k=1}^{\infty} P\{x = k]$$

Proof: write $P\{x \geq k\} = \sum_{x \geq k} P\{X = x\}$

$\rightsquigarrow$ exchange sums over k & x.

$X_1 \sim \text{geom}(p)$, $X_2 \sim \text{geom}(q)$ } indep.

★ recompute by looking at M as a geom. r.v.

$$\underline{M} = \min\{X_1, X_2\} \longrightarrow \text{geom}\left(\frac{1}{}\right.$$

$$P\{M \geq k\} = P\{X_1 \geq k \cap X_2 \geq k\} = P\{X_1 \geq k\} P\{X_2 \geq k\}$$

$$\mathbb{E}[M] = \sum_{k \geq 1} P\{X_1 \geq k\} P\{X_2 \geq k\} = \frac{1}{1-(1-p)(1-q)}$$

Variance

- $\text{Var}(X)$: quantitative notion of X around its $\mathbb{E}[X]$

$$\text{Var}(X) = \mathbb{E}[\underbrace{(X - \mathbb{E}[X])^2}_{g(x)}] = \mathbb{E}[X^2] - \mathbb{E}[X]^2$$

↳ Variance is nothing but the expectation of another r.v. $(g(x))$

$$\mathbb{E}[(X - \mathbb{E}[X])^2] = \mathbb{E}[X^2 - 2X\mathbb{E}[X] + \mathbb{E}[X]^2)$$
$$= \mathbb{E}[X^2] - 2\mathbb{E}[X]\mathbb{E}[X] + \mathbb{E}[X]^2$$
$$= \mathbb{E}[X^2] - 2\mathbb{E}[X]^2 + \mathbb{E}[X]^2$$
$$= \mathbb{E}[X^2] - \mathbb{E}[X]^2$$

- Note: there are other ways to measure variability:
  - p^{th} moment: $\mathbb{E}[|x|^p]$ → 2^{nd} moment very common
  - entropy: $H(x) := \sum_x P_x(x) \log \frac{1}{P_x(x)} = \mathbb{E}[\log(\frac{1}{P_x(x)})]$

- Unlike $\mathbb{E}[\cdot]$, $\text{Var}(\cdot)$ is not linear in its variables:

$\Rightarrow$ $\text{Var}(X+Y) \neq \text{Var}(X) + \text{Var}(Y)$

$$\text{Var}(X+Y) = \mathbb{E}[X + Y - \mathbb{E}[X] - \mathbb{E}[Y])^2$$

$$= Var(X) + Var(Y) + 2\underbrace{\mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])]}_{Cov(X,Y)}$$

↳ positive means X & Y are more correlated with each other (ie, they more in the same direction)

↳ when they're uncorrelated, $Cov(X,Y) = 0$

- If X, Y uncorrelated (ie, $Cov(X,Y) = 0$) then $Var(X+Y) = Var(X) + Var(Y)$
- independence ⇒ uncorrelated ($Cov(X,Y) = 0$) but not vice versa
- def: correlation coefficient:

$$\rho(X,Y) = \frac{Cov(X,Y)}{\sqrt{Var(X)\,Var(Y)}}$$

This normalization allows us to see the relationship btwn X & Y.

↳ this coeff. is always btwn -1 & +1

Recall Cauchy-Schwartz

$$\frac{x^T y}{\|x\|\,\|y\|} \in [-1, 1]$$

→ C-S can be formulate for rv's

Cauchy-Schwartz:

$$|\mathbb{E}[XY]| \le (\mathbb{E}X^2)^{\frac{1}{2}} (\mathbb{E}Y^2)^{\frac{1}{2}}$$

$$\mathbb{E}[XY] = \sum_{x,y} \underbrace{P_{xy}(x,y)^{\frac{1}{2}} x \; y P_{xy}(x,y)^{\frac{1}{2}}}_{\text{inner product}}$$

$$\le \left(\sum_{x,y} x^2 P_{xy}(x,y)\right)^{\frac{1}{2}} \left(\sum_{x,y} y^2 P_{xy}(x,y)\right)^{\frac{1}{2}}$$

What's this summation over?

Example: Computing Variance of Binomial

$X \sim Bin(n, p)$

$Var(X) = ?$

$X = \sum_{i=1}^{n} X_i \qquad X_i \sim_{iid} Bern(p)$

By independence, we have this equality

$$Var(X) = \sum_{i=1}^{n} Var(X_i) = n Var(X_i) = np(1-p)$$

fish → good way to measure arrivals

Poisson Random Variables

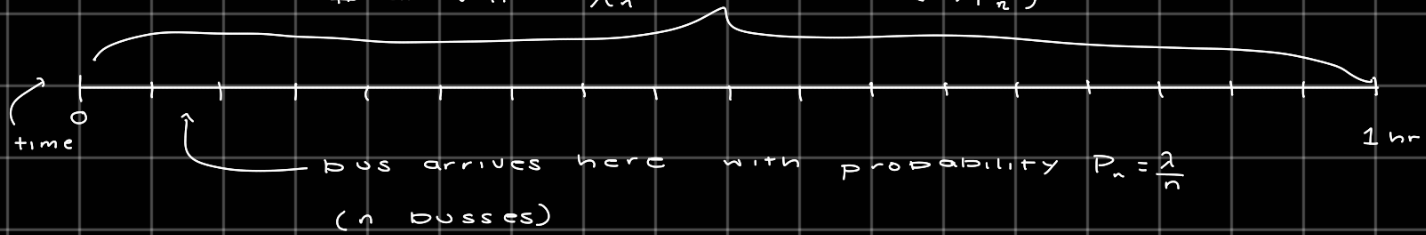
$X \sim Poisson(\lambda)$

λ: rate

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

Q/ Where does Poisson come from?

arrivals = $X_n \sim \text{Binomial}(n, p_n)$



↳ What's probability that k buses arrive?

$$P(X_n = k) = \binom{n}{k}\left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k} \quad \leftarrow \text{binomial PMF}$$

$$= \underbrace{n(n-1)\cdots(n-k+1)}_{\text{as } n \to \infty,\ \text{this} = n^k} \left(\frac{1}{\underbrace{n-\lambda}_{\text{this} = n^k}}\right)^k \frac{\lambda^k}{k!} \left(1 - \frac{\lambda}{n}\right)^n$$