Quick Review

$X \sim$ Bernoulli(p) if

$$p_X(x) = \begin{cases} p & \text{if } x = 1 \\ 1-p & \text{if } x = 0 \end{cases}$$

- Can be thought of as a biased coin toss
- Indicators

$$\mathbb{E}[X] = p \quad , \quad Var(X) = p(1-p)$$

$X \sim$ Binomial(n,p) if for $k \in \{0, \dots, n\}$,

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

- Can be thought of as a sum of n independent Bernoulli(p) RVs

$$\mathbb{E}[X] = np, \quad Var(X) = np(1-p)$$

$X \sim$ Geometric(p) if for any positive integer k,

$$p_X(k) = (1-p)^{k-1} p$$

- Distribution of the # of coin tosses until the first head
- # of independent trials until a success

- "Memoryless" – behavior conditioned on $X > k$ is identical:

$$\mathbb{P}[X = s] = \mathbb{P}[X = s+k \mid X > k].$$

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

$X \sim$ Poisson (λ) if for all nonnegative integers k,

$$p_X(k) = e^{-\lambda} \frac{\lambda^k}{k!}$$

- Surprisingly, turns out to be a good model for reality
- # of packets per hour to a router, e.g.
- Distribution is determined by the avg.

$$\mathbb{E}[X] = \lambda, \quad Var(X) = \lambda$$

Derivations:

Bernoulli(p):

$$\mathbb{E}[X] = \mathbb{P}[X=1] \cdot 1 + \mathbb{P}[X=0] \cdot 0$$

$$= p$$

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

$$= \mathbb{P}[X^2 = 1] \cdot 1 + \mathbb{P}[X^2 = 0] \cdot 0 - p^2$$

same as $\mathbb{P}[X=1]$ ↗ same as $\mathbb{P}[X=0]$ ⤴

$$= p - p^2 = p(1-p).$$

Binomial (n,p):

Recall that if $X \sim \text{Binomial}(n,p)$, then we can write $X = Y_1 + \cdots + Y_n$, where the Y_i are i.i.d. Bernoulli(p).

$$\mathbb{E}[X] = \mathbb{E}[Y_1 + Y_2 + \cdots + Y_n]$$

$$= \sum_{i=1}^{n} \mathbb{E}[Y_i] \quad \leftarrow \text{linearity}$$

$$= \sum_{i=1}^{n} p = np$$

Bernoulli expectation

$$\text{Var}(X) = \text{Var}(Y_1 + Y_2 + \cdots + Y_n)$$

$$= \sum_{i=1}^{n} \text{Var}(Y_i) \quad \leftarrow \text{variance is linear over independent RVs}$$

$$= \sum_{i=1}^{n} p(1-p)$$

Bernoulli variance

$$= np(1-p)$$

Disclaimer: the next two are fairly involved, and involve some calculus...

Geometric (p):

$$\mathbb{E}[X] = \sum_{i=1}^{\infty} \mathbb{P}[X \geq i] \quad \leftarrow \text{Tail sum formula}$$

(only holds for RVs with range $\mathbb{N}$)

$\mathbb{P}[X \geq i]$ is the probability that we fail in the first $i-1$ steps, so

$$\mathbb{P}[X \geq i] = (1-p)^{i-1},$$

hence

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

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

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

$$\mathbb{E}[X^2] = \sum_{i=1}^{\infty} i^2 \cdot \mathbb{P}[X = i]$$

$$= p \sum_{i=1}^{\infty} i^2 (1-p)^{i-1}.$$

To compute this sum, it helps to consider the power series

$$f(x) = \frac{1}{1-x} = \sum_{i=1}^{\infty} x^{i-1}$$

defined when $|x| < 1$. Observe that

$$\frac{d}{dx}\left[x f(x)\right] = \frac{d}{dx}\left[\sum_{i=1}^{\infty} x^i\right]$$

$$= \sum_{i=1}^{\infty} i x^{i-1}$$

(we can interchange summation, differentiation bc $xf(x)$ is "well-behaved" - see math 104 for this rigorized)

One can calculate the LHS as

$$\frac{d}{dx}\left[x \cdot \frac{1}{1-x}\right] = \frac{1}{(1-x)^2},$$

so for $|x| < 1$,

$$\frac{1}{(1-x)^2} = \sum_{i=1}^{\infty} i x^{i-1}.$$

If we do this again, we have

$$\frac{d}{dx}\left[\frac{x}{(1-x)^2}\right] = \sum_{i=1}^{\infty} i^2 x^{i-1},$$

so, evaluating the LHS again, we see that

$$\frac{1+x}{(1-x)^3} = \sum_{i=1}^{\infty} i^2 x^{i-1}.$$

With this formula, plugging in $x = (1-p) < 1$, we see that

$$\mathbb{E}[X^2] = p \sum_{i=1}^{\infty} i^2 (1-p)^{i-1} = \frac{2-p}{p^2},$$

so

$$\mathrm{Var}(X) = \frac{2-p}{p^2} - \frac{1}{p^2} = \frac{1-p}{p^2}$$

Poisson (λ):

$$\mathbb{E}[X] = \sum_{i=0}^{\infty} i \cdot \mathbb{P}[X=i] = \sum_{i=0}^{\infty} i \cdot e^{-\lambda} \cdot \frac{\lambda^i}{i!}$$

$$= \sum_{i=1}^{\infty} e^{-\lambda} \cdot \frac{\lambda^i}{(i-1)!}$$

Note that we're indexing at 1 now

$$= \lambda e^{-\lambda} \sum_{i=1}^{\infty} \frac{\lambda^{i-1}}{(i-1)!}$$

$$= \lambda e^{-\lambda} (e^{\lambda})$$

Taylor series for e^x

$$= \lambda.$$

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

$$\mathbb{E}[X^2] = \sum_{i=0}^{\infty} i^2 \cdot \mathbb{P}[X=i] = \sum_{i=0}^{\infty} i^2 \cdot e^{-\lambda} \cdot \frac{\lambda^i}{i!}$$

Again note the index change

$$= \sum_{i=1}^{\infty} i \cdot e^{-\lambda} \cdot \frac{\lambda^i}{(i-1)!}$$

$$\sum_{i=1}^{\infty} i \cdot e^{-\lambda} \cdot \frac{\lambda^i}{(i-1)!}$$

split up the i into $1 + (i-1)$

$$= \sum_{i=1}^{\infty} e^{-\lambda} \cdot \frac{\lambda^i}{(i-1)!} + \sum_{i=1}^{\infty} (i-1) e^{-\lambda} \cdot \frac{\lambda^i}{(i-1)!}$$

$$= \sum_{i=1}^{\infty} e^{-\lambda} \cdot \frac{\lambda^i}{(i-1)} + \sum_{i=2}^{\infty} e^{-\lambda} \cdot \frac{\lambda^i}{(i-2)!}$$

$$= \lambda e^{-\lambda} \sum_{i=0}^{\infty} \frac{\lambda^i}{i!} + \lambda^2 e^{-\lambda} \sum_{i=0}^{\infty} \frac{\lambda^i}{i!} \leftarrow \text{Taylor series for } e^{\lambda}$$

$$= \lambda (e^{-\lambda})(e^{\lambda}) + \lambda^2 e^{-\lambda}(e^{\lambda})$$

$$= \lambda + \lambda^2$$

Hence,

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

$$= \lambda$$

CS 70 Discrete Mathematics and Probability Theory
Summer 2020 Course Notes **DIS 6A**

1 Warm-up

For each of the following parts, you may leave your answer as an expression.

(a) You throw darts at a board until you hit the center area. Assume that the throws are i.i.d. and the probability of hitting the center area is $p = 0.17$. What is the probability that you hit the center on your eighth throw?

(b) Let $X \sim \text{Geometric}(0.2)$. Calculate the expectation and variance of X.

(c) Suppose the accidents occurring weekly on a particular stretch of a highway is Poisson distributed with average number of accidents equal to 3 cars per week. Calculate the probability that there is at least one accident this week.

(d) Consider an experiment that consists of counting the number of α particles given off in a one-second interval by one gram of radioactive material. If we know from past experience that, on average, 3.2 such α-particles are given off per second, what is a good approximation to the probability that no more that 2 α-particles will appear in a second?

2 Coupon Collector Variance

It's that time of the year again - Safeway is offering its Monopoly Card promotion. Each time you visit Safeway, you are given one of n different Monopoly Cards with equal probability. You need to collect them all to redeem the grand prize.

Let X be the number of visits you have to make before you can redeem the grand prize. Show that $\mathrm{Var}(X) = n^2 \left(\sum_{i=1}^{n} i^{-2}\right) - \mathbb{E}(X)$. *[Hint: Try to express the number of visits as a sum of geometric random variables as with the coupon collector's problem. Are the variables independent?]*

3 Boutique Store

Consider a boutique store in a busy shopping mall. Every hour, a large number of people visit the mall, and each independently enters the boutique store with some small probability. The store owner decides to model X, the number of customers that enter her store during a particular hour, as a Poisson random variable with mean λ.

Suppose that whenever a customer enters the boutique store, they leave the shop without buying anything with probability p. Assume that customers act independently, i.e. you can assume that they each flip a biased coin to decide whether to buy anything at all. Let us denote the number of customers that buy something as Y and the number of them that do not buy anything as Z (so $X = Y + Z$).

(a) What is the probability that $Y = k$ for a given k? How about $\mathbb{P}[Z = k]$? *Hint*: You can use the identity

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}.$$

(b) State the name and parameters of the distribution of Y and Z.

(c) Prove that Y and Z are independent. In particular, prove that for every pair of values y, z, we have $\mathbb{P}[Y = y, Z = z] = \mathbb{P}[Y = y]\mathbb{P}[Z = z]$.