Stat 134 lec 15

Warmup: 10:00–10:10

Lucy and two friends each have a p-coin and toss it independently at the same time.

a) What is the probability it takes Lucy more than n tosses to get heads?

X = # tosses until Lucy gets heads

$X \sim Geom(p)$ on $1, 2, 3, \dots$

$$P(X > n) = q^n p + q^{n+1} p + \cdots = q^n p \left[1 + q + q^2 + \cdots\right] = \frac{q^n p}{1-q} = \boxed{q^n}$$

b) What is the probability that the first person to get a head has to toss more than n times.

$$P(\min(X_1, X_2, X_3) > n)$$

$$= P(X_1 > n, X_2 > n, X_3 > n)$$

$$= \underbrace{P(X_1 > n)}_{= q^n}{}^3$$ by independence and $X_1, X_2, X_3 \sim Geom(p)$,

$$= \boxed{q^{3n}} = (q^3)^n$$

Fact $X \sim Geom(p)$ iff $P(X > n) = q^n$

It follows that

$$\min(X_1, X_2, X_3) \sim Geom(1 - q^3)$$

Announcement: midterm review materials are on website.
Will start review Friday.

last time

Sec 3.4 Geometric distribution (Geom(p))

ex Coupon Collector's problem

You have a collection of boxes each containing a coupon. There are n different coupons. Each box is equally likely to contain any coupon independent of the other boxes.

X = # boxes needed to get all n different coupons.

ex $n = 3$ $X = X_1 + X_2 + X_3$

$$\underbrace{c}_{X_1}\underbrace{c\,c\,c}_{X_2}\underbrace{c\,c\,c\,c}_{X_3}$$

a) What is the distribution of X_1, X_2, X_3? Are they independent? – yes

$X_1 \sim \text{Geom}\left(\frac{3}{3}\right)$, $X_2 \sim \text{Geom}\left(\frac{2}{3}\right)$, $X_3 \sim \text{Geom}\left(\frac{1}{3}\right)$

b) What is $E(X) = \frac{1}{\frac{3}{3}} + \frac{1}{\frac{2}{3}} + \frac{1}{\frac{1}{3}} = 3\left(1 + \frac{1}{2} + \frac{1}{3}\right)$

$\text{Var}(X_1) = \frac{q}{p^2}$

c) What is $\text{Var}(X)$? $\frac{\frac{0}{3}}{\left(\frac{3}{3}\right)^2} + \frac{\frac{1}{3}}{\left(\frac{2}{3}\right)^2} + \frac{\frac{2}{3}}{\left(\frac{1}{3}\right)^2} = 3\left(0 + \frac{1}{2^2} + \frac{2}{1^2}\right)$

soln for n coupons:

X_1 = # boxes to 1^{st} coupon $\sim \text{Geom}\left(\frac{n}{n}\right)$

$X_1 + X_2$ = # boxes to 2^{nd} coupon so $X_2 \sim \text{Geom}\left(\frac{n-1}{n}\right)$

$\vdots$

$X_1 + \dots + X_n$ = # boxes to n^{th} coupon so $X_n \sim \text{Geom}\left(\frac{1}{n}\right)$

$X = X_1 + \dots + X_n$ sum of indep Geom with diff. p.

$$E(X) = \underset{\frac{n}{n}}{E(X_1)} + \underset{\frac{n}{n-1}}{E(X_2)} + \underset{\frac{n}{n-2}}{E(X_3)} + \dots + \underset{\frac{n}{1}}{E(X_n)}$$

$$\boxed{E(X) = n\left(1 + \frac{1}{2} + \frac{1}{3} + \dots + \frac{1}{n}\right)}$$

$$\boxed{Var(X) = n\left(\frac{0}{n^2} + \frac{1}{(n-1)^2} + \dots \frac{n-1}{1^2}\right)}$$

Today

1. Finish Sec 3.4 Minimum of independent geometrics
2. Sec 3.5 Poisson distribution
3. Poisson random scatter (PRS) AKA Poisson Process

(1) Sec 3.4 Minimum of independent geometrics

Adam, Beth and John independently flip a p_1, p_2, p_3 coin respectively.
let X = # trials until Adam, Beth or John get a heads.

ex

ex	A	TTT	$X_1 \sim Geom(p_1)$
	B	TTT	$X_2 \sim Geom(p_2)$
	J	$\underbrace{TTH}_{X=3}$	$X_3 \sim Geom(p_3)$

a) what is probability Adam, Beth or John get a head?

$$P = \text{Prob}\,[A \text{ or } B \text{ or } J \text{ get heads}]$$
$$= 1 - \text{Prob}\,(A, B, J \text{ don't get heads})$$
$$= 1 - q_1 q_2 q_3$$

b) what distribution is X?

$$X = \min(X_1, X_2, X_3) \sim Geom(1 - q_1 q_2 q_3)$$

② sec 3.5 Poisson distribution (Pois(μ))

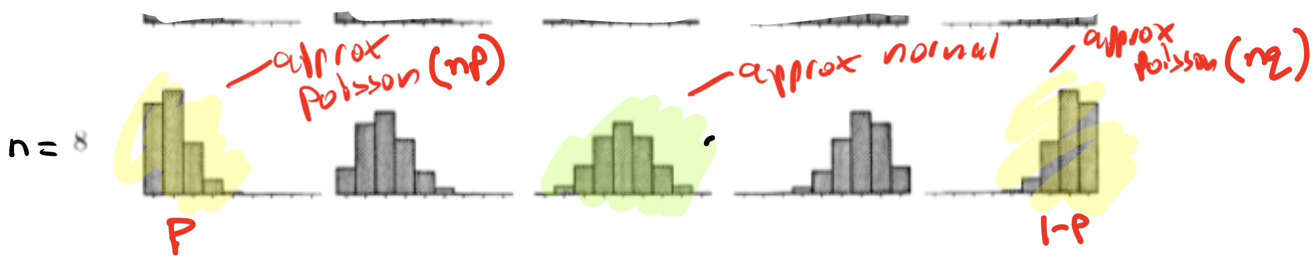
$X \sim \text{Pois}(\mu)$

$$P(X=k) = \frac{e^{-\mu}\mu^k}{k!} \quad k=0,1,2,\dots$$

Intuitively, we know $E(X)=\mu$ and $\text{Var}(X)=\mu$

since,

$$\text{Bin}(n,p) \to \text{Pois}(\mu) \quad \text{when} \quad \begin{array}{l} n\to\infty \\ p\to 0 \\ np\to\mu. \end{array}$$



Also we expect $npq \to \mu q \approx \mu$ so $\text{var}(X)$ should be μ. See appendix for a proof.

ex Let $X \sim \text{Pois}(\mu)$

Find $E(\underbrace{X(X+1)}_{=X^2+X}) = E(X^2) + E(X) = \mu + \mu^2 + \mu = \boxed{2\mu + \mu^2}$

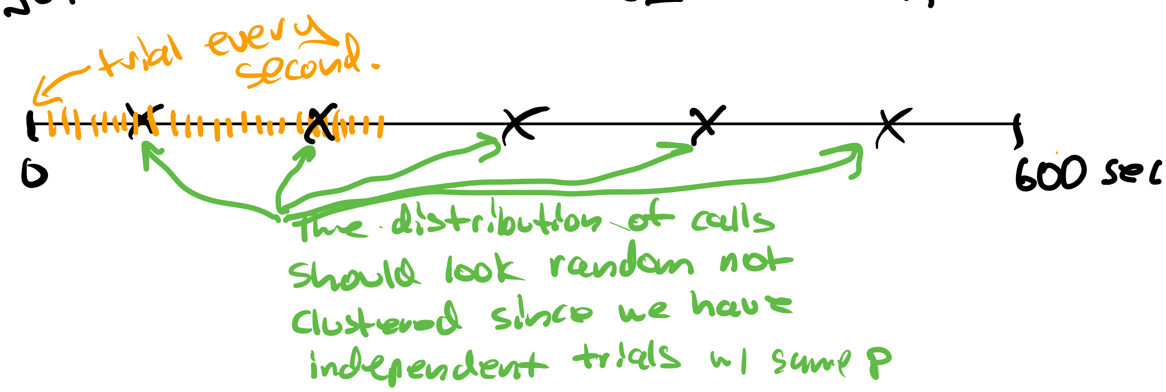
$E(X^2) = \text{Var}(X) + (E(X))^2$

③ Poisson Random Scatter (PRS)

A random scatter of points in a time line is an example of a Poisson random scatter.

ex X = number of calls coming into a hotel reservation center in 600 seconds

Choose an interval of time so no time interval gets more than one call (ex seconds).



PRS assumptions

1) No time interval gets more than one call

2) Have n iid Bernoulli p trials with $\mu = np$ large n, small p.

(i.e all calls are independent of each other with the same probability) let $X = \#$ calls in $\underbrace{t \text{ seconds}}_{\text{time of } n \text{ trials}}$

Then $X \sim \text{Poi}(\mu) \leftarrow$ limit of $\text{Bin}(n,p)$ as $n \to \infty$, $p \to 0$, $np \to \mu$.

Say on average there are $\mu = 5$ calls in 600 seconds

Let $\lambda > 0$ be the rate (or intensity) of calls per second

ex $\lambda = 5/600$ calls/sec in above example.

Since λ is the same every time interval (PRS assumptions) $\mu = \lambda t$.

λ has units calls/sec so $\mu = \lambda t$ has units calls in t sec

ex $\mu = \lambda t = \frac{5}{600} \cdot 600 = 5$ calls in 600 sec.

Stat 134

1. Which of the following can be modeled as a Poisson Random Scatter with intensity $\lambda > 0$?

   a ~~The number of blueberries in a 3 cubic inch blueberry muffin~~

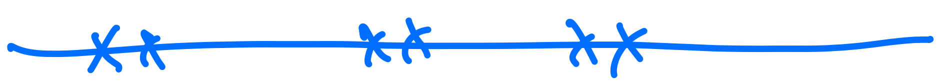
   different λ at 3pm and 3am.

   b The number of patients entering a doctor's office in a 24 hour period.

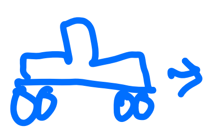
   c The number of times a day a person feels hungry

   d The number of air pulses counted every second from cars driving over an empty rubber hose lying across a highway between noon and 1pm.

   e more than one of the above



air pulses



come in pairs from front and back tires. this isn't random scatter.

Appendix

Let $X \sim \text{Pois}(\mu)$

Then $E(X) = \mu$ and

$Var(X) = \mu$

Pf/

Recall $e^{\mu} = 1 + \mu + \frac{\mu^2}{2!} + \cdots = \sum_{k=0}^{\infty} \frac{\mu^k}{k!}$ Taylor series.

$$E(X) = \sum_{k=0}^{\infty} k \cdot P(X=k) = \sum_{k=0}^{\infty} k e^{-\mu} \frac{\mu^k}{k!}$$

$$= \sum_{k=1}^{\infty} \cancel{k} e^{-\mu} \frac{\mu^{k-1} \mu}{(k-1)! \cancel{k}}$$

(note $0 \cdot e^{-\mu} \frac{\mu^0}{0!} = 0$)

$$= \mu e^{-\mu} \sum_{k=1}^{\infty} \frac{\mu^{k-1}}{(k-1)!}$$

$$= \mu e^{-\mu} \underbrace{\left(1 + \mu + \frac{\mu^2}{2!} + \cdots\right)}_{e^{\mu}} = \boxed{\mu}$$

Next we show $Var(X) = \mu$:

$$Var(X) = E(X^2) - E(X)^2$$

$$= E(X^2) - E(X) + E(X) - E(X)^2$$

$$= E(X(X-1)) + E(X) - E(X)^2$$

$$E(X(X-1)) = \sum_{k=0}^{\infty} \cancel{k}(\cancel{k-1})\underbrace{P(X=k)}_{= \frac{e^{-\mu}\mu^{k}}{\cancel{k}(\cancel{k-1})(k-2)!}}$$

$$= e^{-\mu} \sum_{k=2}^{\infty} \frac{\mu^{k}}{(k-2)!}$$

$$= e^{-\mu}\mu^{2} \overbrace{\sum_{k=2}^{\infty} \frac{\mu^{k-2}}{(k-2)!}}^{= e^{\mu}} = \mu^{2}$$

$$\Rightarrow Var(X) = \mu^{2} + \mu - \mu^{2} = \boxed{\mu}$$

$\square$