Lecture 3

1/22/15

Galton - Watson Process - index case $\rightarrow$ new cases

$\hookrightarrow$ next lecture or 2 lectures - but working towards

Immigration - Death Process - today

---

correlation:

$$\text{cov}(X, Y) = E[(X - EX)(Y - EY)]$$

$$= E[XY] + E[-Y\,EX] + [-X\,EY] + E[EX\,EY]$$

$$= \underbrace{E[XY] - (EX)(EY)}$$

if X, Y indep $\Rightarrow$ $\overset{\text{uncorrelated}}{\text{cor}(X,Y) = 0}$

★ uncorrelated variables may NOT! be independent!

Example of this:

x	y	P(X=x, Y=y)	xy
0	1	1/4	0
0	-1	1/4	0
-1	0	1/4	0
1	0	1/4	0

$E[XY] = 0 \qquad EX = 0 \qquad EY = 0$

$\text{cov}(x, y) = 0$ uncorrelated

x	P(x)
-1	1/4
0	1/2
1	1/4

y	P(y)
-1	1/4
0	1/2
1	1/4

IF indep

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

$P(0, 0) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$

<u>But</u> really $P(0,0) = 0$

<u>So not indep!</u>

$$r = \frac{\text{cor}(X,Y)}{\sqrt{\text{var}X \ \text{var}Y}} \qquad -1 < r < 1$$

Proof in notes

Binomial Process

Binomial Distribution

N indep Bernouilli trial — sum of successes

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

Different situation - Geometric Distribution

o o o o o when do I get the 1st success

1 2 3 4 5

1 - p

2 - $p(1-p)$

3 - $p(1-p)^2$

$\vdots$

i - $p(1-p)^{i-1}$

$$\boxed{P(Y=i) = (1-p)^{i-1} p}$$

* Parameterization varies in literature - BE AWARE!

Think of this as a waiting time (in discrete units)

Binomial counts in fixed window
- N indep Bernouilli trials

goes with

Geometric waiting time between successes
- Indep Bernouilli trials until a success

$$\sum_{i=1}^{\infty} p(1-p)^{i-1}$$ has to be 1 - probability of the sample space!

Next, we prove this —

Proof geometric distribution sums to 1

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

Change of variables:

i	$j = i-1$
1	0
∞	∞

$$p \sum_{j=0}^{\infty} (1-p)^{j}$$ what is this?

$$L = \sum_{j=0}^{\infty} a^j = a^0 + a^1 + a^2 + a^3 + \ldots$$

$$= 1 + a + a^2 + a^3 + \ldots$$

$$= 1 + a(1 + a + a^2 + \ldots)$$

$$= 1 + aL$$

$$\Rightarrow L = \frac{1}{1-a}$$

$$\sum_{j=0}^{\infty} a^j = \frac{1}{1-a}$$ when $0 \le a < 1$

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

Generating Fxn

$E[u^Y]$ where $0 \leq u \leq 1$

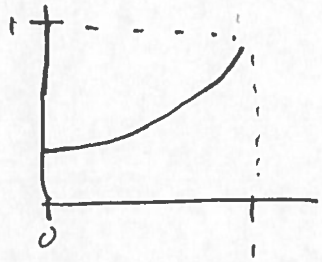
For a geometric Y

$$E[u^Y] = \sum_{y=1}^{\infty} u^y p(1-p)^{y-1} = pu \sum_{y=1}^{\infty} (u(1-p))^{y-1}$$

$$= pu \sum_{y=0}^{\infty} (u(1-p))^y = pu \frac{1}{1-u(1-p)}$$

$$= \frac{up}{1-u(1-p)}$$

probability generating function for the geometric or PGF



We will need this for the next lecture on Galton-Watson Process

$$\frac{d}{du} E[u^Y] = \sum_{i=0}^{\infty} P(Y=i) \, i u^{i-1}$$

$u \to 1$

$$\frac{d}{du} E[u^Y] \Big|_{u=1} = EY$$

The derivative of the PGF at $u=1$ is the expectation of the RV!

This is useful!

Mean of a geometric is $\frac{1}{p}$

recover w/ $p = 0.1$ each day, waiting time for 1 person on avg is $\frac{1}{0.1} = 10$ day

Property of memorylessness – unique to geometric!

Back to Binomial

$$\binom{N}{x} p^x (1-p)^{N-x}$$

$\lambda = Np$ freeze the mean! Take $N \to \infty$, $p \to 0$

$\Rightarrow$ Poisson

PGF of Binomial $(up + (1-p))^N$

$$= (up + q)^N$$

$$\lim_{\substack{N\to\infty \\ p\to 0 \\ Np=\lambda}} (up+q)^N = e^{\lambda(u-1)}$$

which is the PGF of a Poisson RV

$$= e^{-\lambda} e^{\lambda u} = -e^{-\lambda}\left[1 + \lambda u + \frac{1}{2!}(\lambda u)^2 + \frac{1}{3!}(\lambda u)^3 + \dots\right]$$

$$P(Y=i) = \frac{e^{-\lambda}\lambda^i}{i!} \quad \text{Poisson Distribution}$$

* Poisson is a limit of a binomial as mean stays the same but N of trials increases

Bernoulli Trials:

# of events in a fixed window : Binomial $\xrightarrow[Np=\lambda]{N\to\infty,\ p\to 0}$ Poisson

waiting times : Geometric $\longrightarrow$ Exponential

If we have a

PGF: $g(u)$ $\longrightarrow$

$g'(1) = EY$

$g''(1) = E[Y^2] - EY$

$\text{var}(Y) = g''(1) + g'(1) - [g'(1)]^2$

This can be easier than calculating from the definition

Immigration-Death Problem

$$X_{t+1} = Y_t + \text{Binom}(X_t, a) = Y_t + \underbrace{a \circ X_t}$$

# of cases at $t+1$

new cases of disease
Poisson

Old cases that survive

X_0

$X_1 = Y_0$

$X_2 = Y_1 + a \circ X_1$

Binomial given fixed #

Poisson

$\leadsto$ So the whole thing is a random # of random trials

~ sort of binomially thinning a Poisson

$\longrightarrow$ what is it? Use PGFs

$$E[u^{a \circ X_1}]$$

Y is Poisson

X is binomially thinned ~~an~~ Y $\quad a \circ Y$

$X = a \circ Y$

$$E[u^X] = E[\underbrace{E[u^X \mid Y]}]$$

$$E\left[[ua + (1-a)]^Y \right]$$

$$= \sum_{y=0}^{\infty} (ua + (1-a))^y P(Y=y)$$

$$= e^{\lambda(ua + (1-a) - 1)}$$

$$= e^{\lambda a (u-1)}$$

<u>Useful Fact:</u>

PGF of a Poisson

$$g(u) = \sum_{y=0}^{\infty} u^y P(Y=y)$$

$$= e^{\lambda(u-1)}$$

$a \circ Y$ is a Poisson w/ a smaller mean

mean λ goes to λa

This means the immigration-death process is always Poisson!

$$E[X_{t+1}] = \cancel{E[X} \; E[Y_t] + E[a \circ X_t]$$

$$= E[Y_t] + a E[X_t]$$

we can rewrite:

$$\mu_{t+1} = \theta + a \mu_t \quad \Rightarrow \quad \underset{\text{Equilibrium}}{\bar{\mu}} = \theta \frac{1}{1-a}$$

Prevalence is incidence times duration!