Oct 4th

## Geometric R.V.

An experiment consists of repeating trials <u>until</u> the <u>first</u> success.

- success prob $p$
- failure " $1-p$

$X$ = # of trials to first success

pmf: $P(X=x) = (1-p)^{x-1} p$

made 3 free throws
first 2 → fail
3rd → goes in

ex Products produced by a machine have a 3% defective rate. What is the prob that the 1st defective occurs in the 5th item inspected?

defective → "success"

5th item → first 4 are not defective

$$P(X=5) = (1-.03)^4 (.03) = (.97)^4(.03)$$
$$= .0265 \quad 2.7\%$$

# $E[X]$ & $Var(X)$?

Recall (Calc 2) Geometric Series

$$f(x) = \sum_{k=0}^{\infty} x^k = \frac{1}{1-x} \quad \text{if} \quad |x|<1$$

$\underbrace{\frac{1}{1-x}}_{(1-x)^{-1}}$

$= 1 + x + x^2 + x^3 + \dots$

$$f'(x) = \sum_{k=1}^{\infty} k x^{k-1} = \frac{1}{(1-x)^2}$$

$$f''(x) = \sum_{k=2}^{\infty} (k-1) k x^{k-2} = \frac{2}{(1-x)^3}$$

rewrite

shift index

$$\sum_{k=2}^{\infty} (k-1) k \, x^{k-2} = \sum_{k=1}^{\infty} (k-1+1)(k+1) \, x^{k-2+1}$$

$$= \sum_{k=1}^{\infty} \underbrace{k(k+1)}_{k^2+k} x^{k-1}$$

$$= \sum_{k=1}^{\infty} k^2 x^{k-1} + \sum_{k=1}^{\infty} k x^{k-1}$$

$$\frac{2}{(1-x)^3} = \underbrace{\sum_{k=1}^{\infty} k^2 x^{k-1}}_{\text{solve for this}} + \frac{1}{(1-x)^2}$$

$$\Rightarrow \sum_{K=1}^{\infty} K^2 x^{K-1} = \frac{2}{(1-x)^3} - \frac{1}{(1-x)^2} = \frac{2-(1-x)}{(1-x)^3}$$

$$= \frac{1+x}{(1-x)^3}$$

Back to $E[x]$:

$$E[x] = 1 \cdot \underbrace{P(X=1)}_{p} + 2\underbrace{P(X=2)}_{(1-p)p} + 3\underbrace{P(X=3)}_{(1-p)^2 p} + \dots \quad + (1-p)^3 p$$

(1) $E[x] = p + 2p(1-p) + 3p(1-p)^2 + 4p(1-p)^3 + \dots$

multiply by $(1-p)$

(2) $(1-p)E[x] = p(1-p) + 2p(1-p)^2 + 3p(1-p)^3 + 4p(1-p)^4 + \dots$

Subtract (2) from (1)

$$E[x] - (1-p)E[x] = p - p(1-p) + 2p(1-p) - 2p(1-p)^2 + 3p(1-p)^2 - 3p(1-p)^3 + \dots$$

$$\cancel{E[x]} - \cancel{E[x]} + pE[x] = p + p(1-p) + p(1-p)^2 + p(1-p)^3 + \dots$$

divide by $p$ $\quad E[x] = \underset{\downarrow\,(1-p)^0}{1} + (1-p)^1 + (1-p)^2 + (1-p)^3 + \dots$

$$= \sum_{K=0}^{\infty} (1-p)^K$$ Geometric Series!

$(1-p) < 1$

$$= \frac{1}{1-(1-p)} = \frac{1}{p}$$ E[X] Geo R.V.

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

$\frac{1}{p}$

$$E[X] = p + 2p(1-p) + 3p(1-p)^2 + 4p(1-p)^3 + \dots$$

$$E[X^2] = 1^2 p + 2^2 p(1-p) + 3^2 p(1-p)^2 + 4^2 p(1-p)^3 + \dots$$

$$= \sum_{K=1}^{\infty} K^2 p(1-p)^{K-1} = p \sum_{K=1}^{\infty} K^2 (1-p)^{K-1}$$ above

constant

$$= p \cdot \frac{(1+(1-p))}{(1-(1-p))^3}$$

$$= p \cdot \frac{(2-p)}{p^3} = \frac{2-p}{p^2}$$

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

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

# Poisson Discrete R.V.

↳ an approximation to Binomial R.V. in certain conditions (large n, small prob p)

↳ the average # of successes ($\lambda$) that occurs in a specific region / interval

lambda

constant

↳ # of ppl entering building

↳ # of cars on highways

↳ # calls recieved

pmf $P(X=k) = \dfrac{e^{-\lambda}\lambda^{k}}{k!}$ $\lambda = \text{avg}$

$\lambda = 5$

ex Suppose on average there are 5 car crashes per month. What is the prob that there will be at most 1 car crash in a certain month.

at most 1: $P(X=0) + P(X=1) = \dfrac{e^{-5}\,5^{0}}{\underbrace{0!}_{1}} + \dfrac{e^{-5}\,5^{1}}{1!}$ ($5^0 = 1$)

no crash OR 1 crash

$= e^{-5} + 5e^{-5}$

$= 6e^{-5} = \dfrac{6}{e^{5}}$

ex The number of emails you get are an average of 0.2 emails per minute.

a) What is the prob that you do not get an email in the length of 5 minutes?

$\hookrightarrow$ scale $\lambda$

$\lambda = (0.2)5 = 1$

$$P(X=0) = \frac{e^{-1} 1^0}{0!} = \frac{e^{-1} \cdot 1}{1} = \frac{1}{e} \approx .36$$

b) What is the prob that you get more than 3 emails in 10 minutes?

$\hookrightarrow \lambda = (0.2)10 = 2$

$$\begin{aligned}
P(X>3) &= P(X=4) + P(X=5) + P(X=6) + \ldots \\
&= 1 - P(X \leq 3) \\
&= 1 - \left(P(X=0) + P(X=1) + P(X=2) + P(X=3)\right) \\
&= 1 - \frac{e^{-2} 2^0}{0!} - \frac{e^{-2} 2^1}{1!} - \frac{e^{-2} 2^2}{2!} - \frac{e^{-2} 2^3}{3!} \\
&= 1 - e^{-2}\left(1 + 2 + 2 + \frac{8}{6}\right) \\
&= 1 - \frac{19}{3e^2} \approx .1429
\end{aligned}$$

$E[x]$ of Poisson $\longrightarrow$ $\lambda$ is the average

$$E[x] = \sum_{x=1}^{\infty} x\,P(X=x) = \sum_{x=1}^{\infty} x\,\frac{e^{-\lambda}\lambda^{x}}{x!}$$

$x! \rightarrow x\cdot(x-1)!$

first term is zero

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

shift index by 1

$$= \sum_{x=0}^{\infty} e^{-\lambda}\frac{\lambda^{x+1}}{x!}$$

$$= \sum_{x=0}^{\infty} \frac{e^{-\lambda}\lambda^{x}\lambda^{1}}{x!}$$

constant w.r.t. x

Calc 2 Taylor Series

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

$x=\lambda$
$k=x$

$$= e^{-\lambda}\lambda \underbrace{\sum_{x=0}^{\infty} \frac{\lambda^{x}}{x!}}$$

$$= e^{-\lambda}\lambda\, e^{\lambda}$$

$$= \lambda e^{-\lambda+\lambda}$$

$$= \lambda e^{0} = \lambda \cdot 1 = \lambda \checkmark$$

$\hookrightarrow E[x]$ Poisson

Variance $(X) = \lambda$ also (will skip derivation)

Relate Poisson to Binomial:

ex $X$ = # of cars that pass in an hour

We want count # cars for the next hour.

$\lambda = n \cdot p$ $\quad$ $n$ = # of trials

$\hookrightarrow p = \frac{\lambda}{n}$ $\quad$ $p$ = prob success in each trial

(Binomial)

Binomial R.V. $\quad$ interval is 60 min $= n$

$$P(X=K) = \binom{n}{K} p^K (1-p)^{n-K}$$

$$= \binom{60}{K} \left(\frac{\lambda}{60}\right)^K \left(1-\frac{\lambda}{60}\right)^{60-K}$$

To make approximation better, use smaller intervals min $\to$ seconds

$$P(X=K) = \binom{3600}{K}\left(\frac{\lambda}{3600}\right)^K \left(1-\frac{\lambda}{3600}\right)^{3600-K}$$

make interval smaller:

$$= \lim_{n\to\infty} \binom{n}{K}\left(\frac{\lambda}{n}\right)^K\left(1-\frac{\lambda}{n}\right)^{n-K}$$

$\frac{10!}{7!} = 10\cdot 9\cdot 8$

$$= \lim_{n\to\infty} \frac{n!}{(n-K)!\,K!}\,\frac{\lambda^K}{n^K}\left(1-\frac{\lambda}{n}\right)^n\left(1-\frac{\lambda}{n}\right)^{-K}$$

$\frac{10!}{(10-3)!}$ $\hookrightarrow K$

$$= \lim_{n\to\infty} \frac{n(n-1)\cdot\ldots\cdot(n-K+1)}{K!}\,\frac{\lambda^K}{n^K}\left(1-\frac{\lambda}{n}\right)^n\left(1-\frac{\lambda}{n}\right)^{-K}$$

$$= \lim_{n\to\infty} \frac{\overbrace{n(n-1)\cdots(n-k+1)}^{k \text{ \# of terms}}}{n^k} \frac{\lambda^k}{k!}\left(1-\frac{\lambda}{n}\right)^n\left(1-\frac{\lambda}{n}\right)^{-k}$$

$$= \lim_{n\to\infty}\left[\frac{n^k + \cdots}{n^k} \frac{\lambda^k}{k!}\left(1-\frac{\lambda}{n}\right)^n\left(1-\frac{\lambda}{n}\right)^{-k}\right]$$

$n^k \leftarrow$ highest power

A B C D

$\lim_{n\to\infty} A = 1$ same power top & bottom

$$\lim_{n\to\infty} B = \lim_{n\to\infty} \frac{\lambda^k}{k!} = \frac{\lambda^k}{k!}$$

$$\lim_{n\to\infty} C = \lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^n = e^{-\lambda}$$ L'Hopital's Rule

$$\lim_{n\to\infty} D = \lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^{-k} = (1-0)^{-k} = 1^{-k} = 1$$

$$= 1 \cdot \frac{\lambda^k}{k!} e^{-\lambda} \cdot 1 = \frac{e^{-\lambda}\lambda^k}{k!}$$ $\boxed{\text{Poisson R.V.}}$

$$\frac{e^{-\lambda}\lambda^x}{x!}$$

ex A certain disease occurs in 1.2% of the population. If 100 people are selected randomly, compute the prob that no people have the disease.

Binomial R.V.

success → disease

n trials → independent
(ppl)

$n = 100 \quad p = 0.012$

$$P(X=0) = \binom{n}{k} p^k (1-p)^{n-k}$$

$$= \binom{100}{0} (0.012)^0 (1-.012)^{100}$$

$$= .299$$

almost the same

Poisson (n "large", p "small") $\lambda = np$

$= (100)(.012)$

$= 1.2$

$$P(X=0) = \frac{e^{-\lambda} \lambda^0}{0!} = \frac{e^{-1.2} \lambda^0}{0!} = e^{-1.2}$$

$= .301194$

Poisson can be used to approx Binomial (w/ large n, small prob p)

— end discrete R.V.

end Ch 4

Start Ch 5 next