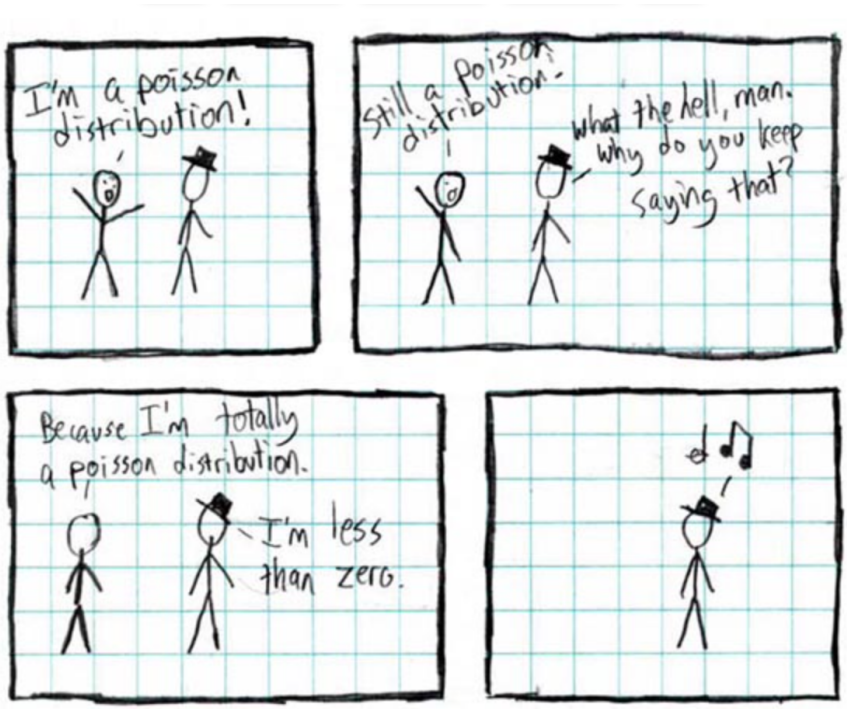
Stat 88: Probability & Mathematical Statistics in Data Science



https://xkcd.com/12/

Lecture 11: 2/12/2021

Waiting times, exponential approximations, the Poisson distribution

Sections 4.3, 4.4 (and finish 4.2)

Geometric distribution

$P(S) = p$, $P(F) = 1-p$

- Recall that T_1 has the **geometric distribution**, denoted $\boldsymbol{T_1 \sim Geom(p)}$ on {1, 2, 3,…}, when we have k-1 failures, and then first success is on k[th] trial.
- $f(k) = P(T_1 = k) = P(FFF \ldots FS) =$ $(1-p)^{k-1}\, p$.
  pmf
- $F(k) = P(T_1 \le k) = 1 - P(T_1 > k) =$ $1-(1-p)^k$
  c.d.f $\quad$ $(1-p)^k$ ←
  $P(T_1 > K)$: Right tailed probabilities

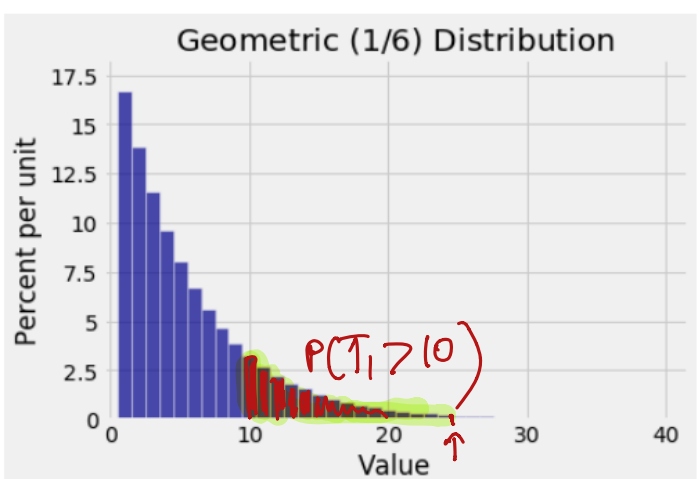
Roll a die until ***first*** ace (1 spot):

$P(\boxdot) = \frac{1}{6}$ $\quad P(\text{not } \boxdot) = \frac{5}{6}$

$P(T_1 = 1) = \frac{1}{6}$

$P(T_1 = 2) = \frac{5}{6} \cdot \frac{1}{6}$

$P(T_1 = 3) = \left(\frac{5}{6}\right)^2 \left(\frac{1}{6}\right) = \frac{5}{6}\left(\frac{5}{6} \cdot \frac{1}{6}\right)$ $\quad P(T_1 = 2)$



Waiting time until r^{th} success

- Say we roll a 8 sided die.

$P(\boxed{8}) = \frac{1}{8}$ $P(\text{not } \boxed{8}) = \frac{7}{8}$

- What is the chance that the first time we roll an eight is on the 11^{th} try?

$= P(FFFFFFFFFFS) = \left(\frac{7}{8}\right)^{10}\left(\frac{1}{8}\right)$

10 F

$T_4 \leftarrow$ Waiting time until the 4th S.

- What is the chance that it takes us 15 times until the 4^{th} time we roll eight? (That is, the waiting time until the 4^{th} time we roll an eight is 15)

$= P(\leftarrow 14 \text{ rolls} \rightarrow S) = \binom{14}{3}\left(\frac{1}{8}\right)^4\left(\frac{7}{8}\right)^{11}$ (15th)

3 S, 11 F; $\frac{1}{8}$

This prob can be computed using a bin. dsn & is $\binom{14}{3}\left(\frac{1}{8}\right)^3\left(\frac{7}{8}\right)^{11}\cdot\frac{1}{8}$

- What is the chance that we need **more** than 15 rolls to roll an eight 4 times?

$P(T_4 > K)$ = Prob that in 1st k rolls we have FEWER than 4 S

- Notice that the ***right-tail*** probability of T_4 is a left hand (cdf) of the Binomial distribution for (15,1/8), and where k=3.

want r^{th} success on k^{th} trial, $k \geq r$

- In general, $P(T_r = k) =$ $P(\underbrace{\text{— — .. - —}}_{k-1 \text{ trials have } r-1 \text{ S}} \cdot S) = \binom{k-1}{r-1}p^{r-1}(1-p)^{k-r}\cdot p$

(k-1) - (r-1) F = k - r

$= P(\text{at most 3 S in k rolls})$

F(3), F is cdf for Bin(k,p) dsn

- And $P(T_r > k) =$

$P(T_r = k) = \binom{k-1}{r-1}p^r(1-p)^{k-r}$

$P(T_r > k)$ = Prob that we have AT MOST $r-1$ successes in first k trials

$P(X \leq r-1)$, $X \sim Bin(k, p)$

$= F(r-1)$

$$= \sum_{i=0}^{r-1} \binom{k}{i} p^i (1-p)^{k-i} = P(X=0) + P(X=1) + \cdots + P(X=r-1)$$

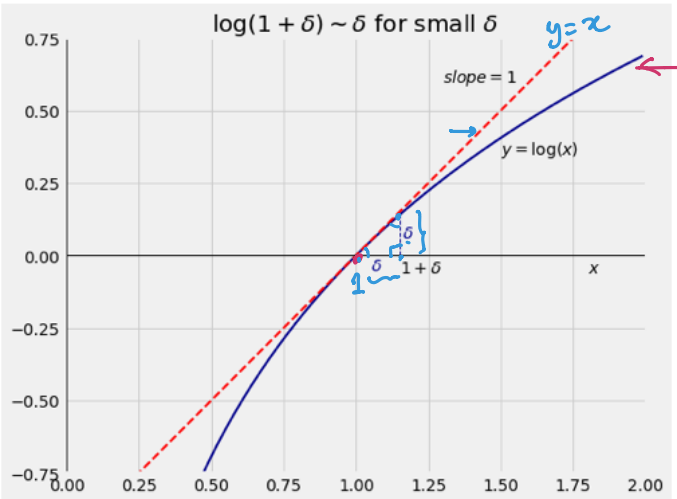
T_r is called a NEGATIVE BINOMIAL R.V.

Waiting time until r^{th} S.

The geom. dsn is a special case for $r=1$

4.3 Exponential Approximations

log ↔ $\log_e$ or ln.



← $y = \log(x)$ (ln x)

slope of tangent at $x=1$

$\frac{d(\log x)}{dx}\Big|_{x=1}$

$\frac{1}{x}\Big|_{x=1} = 1$

Very useful approximation: $\log(1+\delta) \approx \delta$, for δ close to 0

Taylor's theorem, let a be close to x

$f(x) = f'(a)\cdot(x-a) + \text{Rem} \to 0$, so $f(x) \approx f'(a)\cdot(x-a)$

$f(x) = \log(1+x)$, $a = 0$

$\log(1+x) \approx \log(1+0) + \frac{1}{1+0}\cdot(x-0) \approx x$

When x is SMALL

$\log(1-\delta) \approx -\delta$ when δ very small

How to use this approximation

$\log(1+x) \approx x$

$\log(1-x) \approx -x$

- Approximate the value of $x = \left(1 - \frac{3}{100}\right)^{100}$

$x = \left(1 - \frac{3}{100}\right)^{100} \rightarrow \log x = \log\left(\left(1 - \frac{3}{100}\right)^{100}\right)$

$= 100 \cdot \log\left(1 - \frac{3}{100}\right)$

$x \approx e^{-3}$

$\log x \approx 100 \cdot \left(-\frac{3}{100}\right) = -3$

- $x = \left(1 - \frac{2}{1000}\right)^{5000}$

$\log x = 5000 \cdot \log\left(1 - \frac{2}{1000}\right) \approx 5000\left(-\frac{2}{1000}\right) = -10$

$x \approx e^{-10}$

- $x = (1-p)^n$, for large n and small p

$(1-p)^n = P(\text{all F in } n \text{ trials})$

Suppose $n >> 0$ (n is very large)

$p << 1$ (p is very small)

$x = (1-p)^n$

$\log x = n \log(1-p) \approx -n \cdot p \Rightarrow x \approx e^{-np}$

$\log(1-p) \approx -p$

"This implies"

Example

- A book chapter n = 100,000 words and the chance that a word in the chapter has a typo (independently of all other words) is very small : p = 1/1,000,000 = 10^{-6}. Give an approximation of the chance the chapter ***doesn't*** have a typo. (Note: A typo is a *rare event*)

$n = 10^5 \quad p = 10^{-6}$

prob. of a word having a typo $= 10^{-6} = p$

$$P(\text{no typo}) = (1-p)^n \approx e^{-np} = e^{-\frac{1}{10}}$$

#of S, F

Counting Just that Counting is difficult when n is large & p is small

Bootstraps and probabilities

- Bootstrap sample: sample of size n drawn with replacement from original sample of n individuals

- Suppose one particular individual in the original sample is called ~~Ali~~ John. What is the probability that ~~Ali~~ John is chosen at least once in the bootstrap sample?

- Use the complement.

P(John chosen on any draw) $= \frac{1}{n} \approx \frac{1}{99}$

P(John is picked at least once) $= 1 - P(\text{John is not picked at all})$

$= 1 - \left(\frac{n-1}{n}\right)^n$ ← # of draws

$= 1 - e^{-1}$

P(not John)

$\left(\frac{n-1}{n}\right)^n = \left(1 - \frac{1}{n}\right)^n$

$\approx e^{-1}$

$x = \left(1 - \frac{1}{n}\right)^n$

$\log x = n \log\left(1 - \frac{1}{n}\right)$

$\approx n\left(-\frac{1}{n}\right) = -1$

The Poisson Distribution

rare ⇒ prob. is very small.

- Used to model rare events. X is the number of times a rare event occurs, X = 0, 1, 2, …

- We say that a random variable X has the Poisson distribution if

$$\boldsymbol{P(X = k) = e^{-\mu}\frac{\mu^k}{k!}}, \quad k = 0, 1, 2, 3..$$

- The parameter of the distribution is μ

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

$$P(X=k) = e^{-\mu}\frac{\mu^k}{k!}$$

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

$$P(X=1) = e^{-\mu}\frac{\mu}{1} = \mu e^{-\mu}$$

$$P(X=2) = e^{-\mu}\mu^2/2!$$

Relationship between Poisson and Binomial distributions

- ***The Law of Small Numbers***: when n is large and p is small, the binomial (n,p) distribution is ***well approximated*** by the Poisson(μ) distribution where $\mu=np$.

You can think of μ as a rate.