Monday, September 5, 2022 3:51 PM

CSCE 750
9/15/22

Geometric & Binomial distributions

A Bernoulli trial is a 2-outcome random experiment ("success" & "failure")

Let p = prob of success
Let $q := 1-p$ = prob of failure

Binomial distribution
Fixed number n of Bernoulli trials, each with success prob p, indep
Given k where $0 \le k \le n$
want Pr[exactly k successes]

$$\text{Prob}[=k \text{ successes}] = \binom{n}{k} p^k q^{n-k}$$

$\binom{n}{k} = {}_nC_k$ = "n choose k"
(a binomial coefficient)

$\binom{n}{k}$ = # of ways of choosing k items from n items (without replacement & without regard to the order)
= # of k-element subsets of an n-element set.

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} \quad (n! = 1 \cdot 2 \cdots n)$$

Pascal's Triangle

$\binom{n}{k}$

n \ k	0	1	2	3	4	...
0	1					
1	1	1				
2	1	2	1			
3	1	3	3	1		
4	1	4	6	4	1	

$$\Pr[k \text{ successes}] = b(n,p;k) := \binom{n}{k} p^k q^{n-k}$$

If X is binom dist
$X \sim b(n,p)$
$\Pr[X=k] = b(n,p;k)$

$$E(X) = \sum_{k=0}^{n} k \Pr[X=k] = \sum_{k=0}^{n} k \binom{n}{k} p^k q^{n-k} = pn$$

Binomial theorem $\forall x,y \in \mathbb{R}$, $\forall n \ge 0$,

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}$$

[$0! = 1$ by convention]

Geometric distribution

Sequence of Bernoulli trials (indep, all with success prob p), sequence ends with the first success.

The geometric distribution is the prob distribution on the length of this sequence. For $k = 1, 2, \ldots$

Pr[1st success on the k'th trial]
$= q^{k-1} p$ ←

[Some folks take the geo dist instead to be the # of failures before the first success.]

Let X be geo. distributed with success prob p.

$$E(X) = \sum_{k=1}^{\infty} k \cdot \Pr[X=k] = \sum_{k=1}^{\infty} k q^{k-1} p = \frac{p}{q} \sum_{k=1}^{\infty} k q^k$$

$$= \frac{p}{q}\left(\frac{q}{(1-q)^2}\right) = \frac{p}{p^2} = \boxed{\frac{1}{p}}$$

$[1-q=p]$

Suppose you have a biased coin
Pr[Heads] = h
Pr[Tails] = $1-h = t$
$h \ne \frac{1}{2}$.

Want to simulate a fair coin flip.

Sequence of Bernoulli trials
Each B-trial does this:

Flip coin twice, getting
HH, HT, TH, TT
with probs h^2, ht, th, t^2 respectively ($ht = th$)

"Success" means HT or TH
"Failure" " HH or TT

Success prob $= \boxed{2ht} =: p$
failure prob $= 1 - 2ht = h^2 + t^2 =: q$

Expected number of trials is $\frac{1}{p} = \frac{1}{2ht} = \frac{1}{2h(1-h)}$