Lec 10/17

Monday, October 17, 2016 8:02 AM

$X = X_1 + \dots + X_n$

where $X_i \overset{ind}{\sim} Bern(\theta)$

$\Rightarrow X \sim Bin(n, \theta)$

$$p(x) = \begin{cases} \binom{n}{x} \theta^x (1-\theta)^{n-x} & \text{if } x = 0, .., n \\ 0 & o.w. \end{cases}$$

$E(X) = n\theta \qquad Var(X) = n\theta(1-\theta)$

5.5 Geometric & negative Binomial dists.

Consider a scenario where you conduct a sequence of bernoulli trials until you see 1st success.
$\overset{\wedge}{ind}$

Eg: shoot free throws until first make

$$p(x) = \begin{cases} \theta(1-\theta)^{x-1} & \text{if } x = 1, 2, \dots \\ 0 & o.w. \end{cases}$$

$X \sim Geom(\theta)$

$$\underbrace{\theta}_{1 \text{ success}} \; \underbrace{(1-\theta)^{(x-1)}}_{x-1 \text{ failures}}$$

$$M_X(t) = E(e^{tX}) = \sum_{x=1}^{\infty} e^{tx}(1-\theta)^{x-1}\theta = \theta e^t \sum_{x=1}^{\infty} (e^t(1-\theta))^{x-1}$$

$$= \theta e^t \sum_{x=1}^{\infty} (e^t(1-\theta))^{x} = \theta e^t \left(\frac{1}{1-(e^t(1-\theta))}\right) \quad \text{if} \quad e^t(1-\theta) < 1$$

$$e^t < \frac{1}{1-\theta}$$

$$t < -\log(1-\theta)$$

$$= \frac{\theta e^t}{1-(1-\theta)e^t}$$

$\psi(t) = \log(M_X(t))$ (cumulative generating function)

$\psi'(0) = \mu \qquad \psi''(0) = \sigma^2$

$$E(X) = \frac{1}{\theta}$$

$$Var(X) = \frac{1-\theta}{\theta^2}$$

Negative Binomial

X = # trials required to get k^{th} success:

$$P(x) = \begin{cases} \binom{x-1}{k-1}\theta^k(1-\theta)^{x-k} & \text{for } x = k, k+1, \dots \\ 0 & \text{o.w.} \end{cases}$$

$X \sim NegBin(k, \theta)$

Note: Geometric is negative binomial where $k=1$.

$$(e^t)^X = (e^t)^{x-k+k} = (e^t)^{x-k} e^{tk}$$

$$M_X(t) = E(e^{tX}) = \sum^{\infty} e^{tx}\binom{x-1}{k-1}\theta^k(1-\theta)^{x-k}$$

$$M_X(t) = E(e^{tX}) = \sum_{x=k}^{\infty} e^{tx} \binom{x-1}{k-1} \theta^k (1-\theta)^{x-k}$$

$$\text{ie. } = (e^t)^{x-k} e^{tk}$$

$$= \sum_{x=k}^{\infty} \binom{x-1}{k-1} (\theta e^t)^k \left((1-\theta)e^t\right)^{x-k}$$

$$= \frac{(\theta e^t)^k}{(1-(1-\theta)e^t)^k} \sum_{x=k}^{\infty} \binom{x-1}{k-1} \underbrace{[1-(1-\theta)e^t]}_{\theta'}^k \underbrace{[(1-\theta)e^t]}_{1-\theta'}^{x-k} \longrightarrow \text{this is sum of pmf values, so it is 1.}$$

$$= \left[\frac{\theta e^t}{1-(1-\theta)e^t}\right]^k$$

θ' must be in $[0,1]$

$$0 \le (1-\theta)e^t \le 1$$

$$0 \le e^t \le \frac{1}{1-\theta}$$

$$t < -\log(1-\theta)$$

So $E(X) = \dfrac{k}{\theta}$

$Var(X) = \dfrac{k(1-\theta)}{\theta}$

Ex: Suppose 6% of the U.S. pop has A^- blood type. Assume independence from donor to donor.

a) probability the first A^- donor is the 12th

b) probability the second A^- donor is the 20th

a) $X =$ # donors until 1st A^- donor.

$$P(X=12) = 0.06 \cdot 0.94^{11} = 0.0303\ldots$$

b) Let $Y =$ # donors until 2nd A^- donor

$$Y \sim \text{NegBin}(2, 0.06)$$

$$P(Y=20) = \binom{19}{1}(0.06)^2(0.94)^{18}$$

$$= 0.0225$$