Stat 394 Probability I

Lecture 7

Richard Li

July 13, 2017

More about expectation and variance operation

$$E(aX+b) = aE(X) + b$$

$$Var(aX+b) = a^2 Var(X)$$

$$E(X_1 + X_2) = E(X_1) + E(X_2)$$

$$Var(X_1 + X_2) \neq Var(X_1) + Var(X_2)$$

unless . . . .

Expectation of sums

For random variables $X_1, X_2, ..., X_n$, $E(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} E(X_i)$

See pf in textbook.

Example.

$X \sim Bin(n, p) \iff X = X_1 + X_2 + \cdots + X_n$

$X_i \sim Ber(p)$

$E(X) = np$

$E(X_1 + \cdots + X_n)$

$= E(X_1) + E(X_2) + \cdots + E(X_n)$

$= p + p + \cdots + p = np$

Geometric distribution

Suppose independent trials, each of which results in a *success* with probability p and *failure* with probability $1 - p$, are performed until a success occurs. Let X represent the number of trials required, then X is said to be a **geometric random variable** with parameter p, i.e., $X \sim Geom(p)$, and

$$p(X = k) = (1 - p)^{k-1}p, k = 1, 2, ...$$

Expectation and variance

$$\underbrace{E(X) = 1/p}$$

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

$$\sum_{x=0}^{\infty} (1-p)^x = \frac{1}{p} \quad (*)$$

take derivative of $(*)$ w.r.t p.

Coupon collector problem

Suppose you are at the entrance of Odegaard library and ask the birthday of each person entering the library. How many people do you expect that you need to ask, in order to collect all $N = 365$ different days (ignoring Feb 29).

Negative binomial distribution

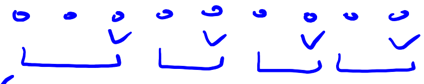
Suppose independent trials, each of which results in a *success* with probability p and *failure* with probability $1 - p$, are performed until r success occurs. Let X represent the number of trials required, then X is said to be a **negative binomial random variable** with parameter (r, p), i.e., $X \sim NB(r, p)$, and

$$p(X = k) = \binom{k-1}{r-1} p^r (1-p)^{k-r}, k = r, r+1, ...$$

Expectation and variance

$$E(X) = \frac{r}{p} \qquad Var(X) = \frac{r(1-p)}{p^2}$$

Not a proof.



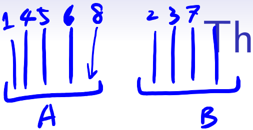
$$X = X_1 + X_2 + \cdots + X_r$$

X_i : # trials since last success before i-th success.

$$X_i \sim Geo(p) \Rightarrow E(X_i) = 1/p$$

$$E(X) = E(X_1 + \cdots + X_r) = E(X_1) + \cdots + E(X_r) = \frac{r}{p}$$

FYI $\left(Var(X) = \cdots = Var(X_1) + \cdots + Var(X_r) = \frac{r(1-p)}{p^2} \right)$



The Banach match problem

A mathematician carries two matchboxes, each originally containing n matches. Each time he needs a match, he is equally likely to take it from either box. What is the probability that, upon reaching for a box and finding it empty, there are exactly k matches still in the other box? Here, $0 \leq k \leq n$.

X : # of draws when you find out A is empty.

Y : # of matches in box B when you — —.

$\Rightarrow X \sim NB(n+1, 1/2)$

$$P(Y=k) = P(X = n+(n-k)+1) = \binom{2n-k+1-1}{n+1-1}\left(\frac{1}{2}\right)^{2n-k+1}$$

$$\Rightarrow 2P(Y=k) = \binom{2n-k}{n}\left(\frac{1}{2}\right)^{2n-k}$$

Hypergeometric distribution

Suppose a sample of size n is to be chosen randomly (without replacement) from an urn containing N balls, of which m are white and $N - m$ are black. Let X denote the number of white balls selected, then X is said to be a **hypergeometric random variable**,

$$p(X = k) = \frac{\binom{m}{k}\binom{N-m}{n-k}}{\binom{N}{n}}, k = 0, 1, ..., n$$

and we can show (see textbook) that

$$E(X) = \frac{mn}{N}, \quad Var(X) = \frac{mn}{N}(1 - \frac{m}{N})(1 - \frac{n-1}{N-1})$$