Stat 134 lec 16

Midterm review

Emma Shearin

5:02pm

It would be helpful if you would review calculations of expectation/variance for geometric-ish distributions (like Pitman 3.4.10). I'm having trouble with the complicated series manipulations required to obtain a closed-form answer on problems like this.

Reply (2 likes)

Justin Han

Yesterday

Joint distributions (i.e. probability of the sum of two RVs where RVs are certain kinds of probability distributions)

Reply (1 like)

Lindsey Chung

Yesterday

Hard problems that use inclusion/exclusion

Reply (1 like)

Tony Wang

6:03pm

Draw cards from a standard deck until three Aces have appeared. Let X = number of cards drawn.

Find: i. $P(X > x)$

ii. $P(X = x)$

iii. EX as a simple fraction

iv. $Var(X)$ using the method of indicators

Q6. From the conceptual review posted on the website.

Balaji Veeramani

12:02am

A population contains G good and B bad elements. Elements are drawn one by one at random without replacement. Suppose the first good element appears on draw number X. Find a simple formula for SD(X).

(From Pitman 3.6.9)

Also, a review of the matching problem would be good.

Reply

Emma 3.4.10

10. Let X be the number of Bernoulli (p) trials required to produce at least one success and at least one failure. Find:

$X=2$: pq or qp

$X=3$: p^2q or q^2p not pqp

a) the distribution of X;

$P(X=x) = q^{x-1}p + p^{x-1}q$ on $\{2,3,\dots\}$

b) $E(X)$;

$$E(X) = \sum_{x=2}^{\infty} xP(X=x) = \sum_{x=2}^{\infty} x\cdot(q^{x-1}p + p^{x-1}q)$$

$$= \sum_{x=2}^{\infty} x\cdot q^{x-1}p + \sum_{x=2}^{\infty} x\cdot p^{x-1}q$$

$$= \left(\sum_{x=1}^{\infty} x\cdot q^{x-1}p\right) - p + \left(\sum_{x=1}^{\infty} x\cdot p^{x-1}q\right) - q = \frac{1}{p} + \frac{1}{q} - 1$$

$Y \sim Geom(P)$

$Var(Y) = E(Y^2) - E(Y)^2$

$E(Y^2) = E(Y)^2 + Var(Y)$

$// \frac{1}{p^2} + \frac{1-p}{p^2} = E(Y)^2 + Var(Y)$

c) $Var(X)$.

$$E(X^2) = \sum_{x=2}^{\infty} x^2P(X=x) = \sum_{x=2}^{\infty} x^2\cdot(q^{x-1}p + p^{x-1}q)$$

$$= \sum_{x=2}^{\infty} x^2\cdot q^{x-1}p + \sum_{x=2}^{\infty} x^2\cdot p^{x-1}q$$

$$= \left(\sum_{x=1}^{\infty} x^2\cdot q^{x-1}p\right) - p + \left(\sum_{x=1}^{\infty} x^2\cdot p^{x-1}q\right) - q = \frac{2-p}{p^2} + \frac{2-q}{q^2} - 1$$

$$Var(X) = E(X^2) - [E(X)]^2 = \frac{2-p}{p^2} + \frac{2-q}{q^2} - 1 - \left(\frac{1}{p} + \frac{1}{q} - 1\right)^2$$

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

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

Justin: convolution formula:

p160.

16. **Discrete convolution formula.** Let X and Y be independent random variables with non-negative integer values. Show that:

a) $P(X+Y=n)=\sum_{k=0}^{n}P(X=k)P(Y=n-k)$.

$$P(X+Y=n)=P(X=0,Y=n)+P(X=1,Y=n-1)+\cdots+P(X=n,Y=0)$$
$$=\sum_{k=0}^{n}P(X=k,Y=n-k)=\sum_{k=0}^{n}P(X=k)P(Y=n-k)$$

b) Find the probability that the sum of numbers on four dice is 8, by taking X to be the sum on two of the dice, Y the sum on the other two.

t	2	3	4	5	6	7	8	9	10	11	12
$P(T=t)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

$\frac{1}{36}$	12											
$\frac{2}{36}$	11											
$\frac{3}{36}$	10											
$\frac{4}{36}$	9											
$\frac{5}{36}$	8											
$\frac{6}{36}$	7											
$\frac{5}{36}$	6											
$\frac{4}{36}$	5											
$\frac{3}{36}$	4											
$\frac{2}{36}$	3											
$\frac{1}{36}$	2											
		2	3	4	5	6	7	8	9	10	11	12
		$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

$$P(X+Y=8)=\sum_{k=0}^{8}P(X=k)P(Y=8-k)=\sum_{k=2}^{6}P(X=k)P(Y=8-k)$$
$$=\begin{cases}P(X=2,Y=6)+P(X=3,Y=5)+P(X=4,Y=4)\\ \qquad +P(X=5,Y=3)+P(X=6,Y=2)\end{cases}$$
$$=2[P(X=2,Y=6)+P(X=3,Y=5)]+P(X=4,Y=4)$$
$$=2\left[\frac{1}{36}\cdot\frac{5}{36}+\frac{2}{36}\cdot\frac{4}{36}\right]+\frac{3}{36}\cdot\frac{3}{36}=\frac{35}{36^2}$$

1 − (at least one person gets hat back)

Lindsey

1. 10 people throw their hats into a box and randomly redistribute the hats among themselves. Assume every permutation of the hats is equally likely. Let N be the number of people who get their own hats back. Find the following:

(a) $\mathbb{E}[N^2]$

Let N_i be the indicator for the event that the i-th person gets their own hat back.

$$\mathbb{E}[N^2] = \mathbb{E}[(\sum_{i=1}^{10} N_i)^2] = \sum_{i=1}^{10}\sum_{j=1}^{10} \mathbb{E}[N_i N_j] = 90(1/90) + 10(1/10) = 2$$

(b) $P(N = 8)$

If 8 people got their own hat back, that means only 2 people have hats that are not their own. There are $\binom{10}{2}$ ways to pick such a pair.

$$P(N = 8) = \frac{\binom{10}{2}}{10!}$$

(c) $P(N = 0)$

Let A_i represent the event that person i got their own hat back.

$$P(N = 0) = 1 - P(A_1 \cup A_2 \cup ... \cup A_{10})$$

By the Principle of Inclusion Exclusion:

$$P(A_1 \cup A_2 \cup ... \cup A_{10}) = \sum_{i=1}^{10} P(A_i) - \sum_{i<j} P(A_i \cap A_j) + \sum_{i<j<k} P(A_i \cap A_j \cap A_k) - ...$$

$$= 10\frac{1}{10} - \binom{10}{2}\frac{1}{10*9} + \binom{10}{3}\frac{1}{10*9*8} - ... = 1 - \frac{1}{2!} + \frac{1}{3!} - ... - \frac{1}{10!}$$

Thus

$$P(N = 0) = 1 - P(A_1 \cup A_2 \cup ... \cup A_{10}) = 1 - 1 + \frac{1}{2!} - \frac{1}{3!} + \frac{1}{4!} - \frac{1}{5!} + \frac{1}{6!} - \frac{1}{7!} + \frac{1}{8!} - \frac{1}{9!} + \frac{1}{10!}$$

Tony → A without replacement version of negative binomial

6. Draw cards from a standard deck until three Aces have appeared. Let X = number of cards drawn. Find:

i. $P(X > x)$

ii. $P(X = x)$

iii. $\mathbb{E}X$ as a simple fraction

iv. $\mathrm{Var}(X)$ using the method of indicators

(i) $P(X > x)$ = chance 3^{rd} ace not before x^{th} draw

$= P(2 \text{ aces in } x \text{ draws}) + P(1 \text{ ace in } x \text{ draws}) + P(0 \text{ ace in } x \text{ draws})$

$$= \frac{\binom{48}{x-2}\binom{4}{2}}{\binom{52}{x}} + \frac{\binom{48}{x-1}\binom{4}{1}}{\binom{52}{x}} + \frac{\binom{48}{x}}{\binom{52}{x}}$$

ii) $P(X=x) = P(X>x-1) - P(X>x)$

and use formula above.

(or) $P(X=x) = P\left(\begin{matrix}2 \text{ aces in} \\ x-1 \text{ draws}\end{matrix}, \begin{matrix}\text{ace in} \\ x^{th} \text{ draw}\end{matrix}\right)$

$= P\left(\begin{matrix}2 \text{ aces in} \\ x-1 \text{ draws}\end{matrix}\right) P\left(\begin{matrix}1 \text{ ace in} \\ x^{th} \text{ draw}\end{matrix} \middle| \begin{matrix}2 \text{ aces in} \\ x-1 \text{ draws}\end{matrix}\right)$

$$= \frac{\binom{48}{x-3}\binom{4}{2}}{\binom{52}{x-1}} \cdot \frac{2}{52-x+1}$$

(iii) label non aces $1, 2, 3, \ldots, 48$

$\_A_1\_A_2\_A_3\_A_4\_$

$p_1 = \frac{3}{5}$

$X = \underset{\text{3 aces.}}{3} + I_1 + \cdots + I_{48}$

where $I_1 = \begin{cases} 1 & \text{if } 1^{st} \text{ non ace before } 3^{rd} \text{ ace.} \\ 0 & \text{else.} \end{cases}$

$\Rightarrow E(X) = 3 + 48\left(\frac{3}{5}\right) = \boxed{28.8}$

(iv) $\text{Var}(X) = \text{Var}(I_1 + \cdots + I_{48})$

2^{nd} non ace can go in any of 4 out of 6 slots

$q_{12} = \frac{3}{5} \cdot \frac{4}{6}$

$E((I_1 + \cdots + I_{48})^2) = 48 E(I_1) + 48 \cdot 47\, E(I_{12})$

$I_{12} = \begin{cases} 1 & \text{if } 1^{st} \text{ and } 2^{nd} \text{ non ace before } 3^{rd} \text{ ace.} \\ 0 & \text{else} \end{cases}$

$\Rightarrow E((I_1 + \cdots + I_{48})^2) = 48 \cdot \frac{3}{5} + 48 \cdot 47\left(\frac{3}{5} \cdot \frac{4}{6}\right) = 931.2$

$\Rightarrow \text{Var}(X) = E((I_1 + \cdots + I_{48})^2) - [E(I_1 + \cdots + I_{48})]^2 = 931.2 - \left(48\left(\frac{3}{5}\right)\right)^2$

$= \boxed{101.76}$

Balaji 3.6.9

$-A_1 - {}^{n}A - A_2 - A_3 - A_4 -$

9. A population contains G good and B bad elements, $G+B=N$. Elements are drawn one by one at random without replacement. Suppose the first good element appears on draw number X. Find simple formulae, not involving any summation from 1 to N, for:

a) $E(X)$;

Let I_i indicate the i^{th} bad element is drawn before the 1^{st} success. The I_i's are identically distributed.

$$E(I_i) = \frac{1}{G+1}$$

$$E(I_iI_j) = \frac{1}{G+1}\cdot\frac{2}{G+2} = \frac{2}{(G+1)(G+2)}$$

Let $W_1 = \#$ of bad elements before the first good element.

$$W_1 = I_1 + \cdots + I_B$$

$$X = W_1 + 1$$

$$E(W_1) = B\cdot E(I_1) = \frac{B}{G+1}$$

$$E(X) = E(W_1) + 1 = \frac{B}{G+1} + \frac{G+1}{G+1} = \frac{N+1}{G+1}$$

---

b) $SD(X)$.

$$E(W_1^2) = \sum_i E(I_i^2) + \sum_{i\neq j} E(I_iI_j) = B\cdot E(I_1) + B(B-1)E(I_1I_2)$$

$$= \frac{B}{G+1} + B(B-1)\frac{2}{(G+1)(G+2)} = \frac{B}{G+1}\left(1+\frac{2B-2}{G+2}\right)$$

$$= \frac{B}{G+1}\left(\frac{G+2+2B-2}{G+2}\right) = \frac{B(G+2B)}{(G+1)(G+2)}$$

$$Var(X) = Var(W_1) = E(W_1^2) - [E(W_1)]^2 = \frac{B(G+2B)}{(G+1)(G+2)} - \left(\frac{B}{G+1}\right)^2$$

$$= \frac{B}{G+1}\left(\frac{G+2B}{G+2} - \frac{B}{G+1}\right) = \frac{B}{G+1}\left[\frac{(G^2+G+2GB+2B)-(GB+2B)}{(G+2)(G+1)}\right]$$

$$= \frac{B}{G+1}\left[\frac{G^2+G+GB}{(G+2)(G+1)}\right] = \frac{GB(N+1)}{(G+1)^2(G+2)}$$

$$SD(X) = \sqrt{\frac{GB(N+1)}{(G+1)^2(G+2)}}$$

Extra indicator problem

Suppose you have 100 balls of color red, blue, and green for a total of 300 balls. There are 100 bins and 3 balls are placed in each bin. Let X be number of bins that have three balls of the same color. Find a) E(X) and b) Var(X).

X = the number of bins that have 3 balls of the same color

$0, 1, \ldots, 33$

$X = I_1 + \cdots + I_{100}$

$I_1 = \begin{cases} 1 & \text{if } 1^{st} \text{ bin has 3 balls same color} \\ 0 & \text{else} \end{cases}$ ← p_1

3 choices of color

$p_1 = \dfrac{\binom{3}{1}\binom{100}{3}}{\binom{300}{3}}$

$\Rightarrow E(X) = 100p_1 = \boxed{\dfrac{100\binom{3}{1}\binom{100}{3}}{\binom{300}{3}}}$

$I_{12} = \begin{cases} 1 & \text{if } 1^{st} \text{ and } 2^{nd} \text{ bin has 3 balls same color} \\ 0 & \text{else} \end{cases}$ ← q_{12}

$q_{12} = \dfrac{\binom{3}{1}\binom{100}{3}}{\binom{300}{3}} \cdot \left[\dfrac{\binom{97}{3} + \binom{2}{1}\binom{100}{3}}{\binom{297}{3}}\right]$

$E(X^2) = 100p_1 + 100 \cdot 99\, q_{12}$

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

plug in values.