MTH 102: Probability and Statistics

Quiz 2

07/06/2022

Sanjit K. Kaul

No books, notes, or devices are allowed. Just a pen and eraser. Any exchange of information related to the quiz with a human or machine will be deemed as cheating. Institute rules will apply. Explain your answers. Show your steps. Approximate calculations are fine as long as the approximations are reasonable. You have 60 minutes.

Question 1. ~~20 marks~~ There are a total of l lectures that are to be attended by a class of k students. The instructor is working on an attendance strategy. He decides that, given the length of a lecture, he will only roll call exactly $j < k$ students. Every lecture the instructor will choose j students randomly to call from the list of k students. Answer the following questions. Show all your steps.

(a) (5 marks) Assume that a certain student attends any lecture, independently of the other, with probability q. What is the probability p that the student's name is called during a lecture?

(b) (5 marks) Let X_i be the random variable that governs whether the student's name is called during lecture i, where $i \in \{1, 2, \ldots, l\}$. $X_i = 1$ in case the student's name is called during lecture i. Else $X_i = 0$. Let $M = \sum_{i=1}^{l} X_i$ be the random number of lectures during which the student's name is called. Derive the PMF of X_i and that of M.

(c) (5 marks) Let Y_i be the random variable that governs whether the student attends lecture i. $Y_i = 1$ in case the student attends lecture i. Else $Y_i = 0$. Let $N = \sum_{i=1}^{l} Y_i$. Derive the distribution of the random variable N.

(d) (10 marks) Consider the random variable $Z = \sum_{i=1}^{l} X_i Y_i$. Derive its distribution.

(e) (20 marks) Derive the conditional distribution of the random variable N, given the event $Z = z$.

(f) (10 marks) Use the above derived conditional distribution to derive the corresponding conditional expectation.

Next we will derive the conditional expectation $E[N|Z = z]$ of N, given the event $Z = z$, in an alternate manner. To do so, answer the following questions.

(aa) (5 marks) Consider the claim $E[N|Z = z] \geq z$. Is the claim correct? Explain your answer.

(bb) (10 marks) The z lectures are those that were attended by the student and during which the student's name was also called. Consider the other $l - z$ lectures, given that $Z = z$. For any of the other $l - z$ lectures, derive the probability $P[Y_i = 1|Z_i = 0]$.

(cc) (20 marks) Use the expressions derived in the above two parts to derive $E[N|Z = z]$. [Hint: What random variable models each of the $l - z$ lectures, given $Z = z$?]

(dd) (10 marks) Derive $E[N]$ starting with your expression for $E[N|Z = z]$.

(a) P[Student's name is called during a lecture]

$$= \frac{{}^{k-1}C_{j-1}}{{}^{k}C_{j}}$$

5

$$= \frac{j}{k}.$$

[j students can be chosen from k in a total of ${}^{k}C_{j}$ ways.
Given a student is amongst the selected j, the no. of ways is ${}^{k-1}C_{j-1}$]

(b) X_i is a Bernoulli RV with parameter p.

We have l RV(s), $X_1, X_2, \dots, X_l$, each of which is Bern(p).

2

$$P[X_i = x] = \begin{cases} p & x=1 \\ (1-p) & x=0 \end{cases}$$

Given $M = \sum_{i=1}^{l} X_i$.

$S_M = \{0, 1, 2, \ldots, m\}$

$P[M=0] = P[$Student's name isn't called in all of the l lectures$]$

$= (1-p)^l$.

$P[M=1] = P[$Student's name is called in exactly 1 out of l lectures$]$

$= {}^{l}C_1 (1-p)^{l-1} p$.

$\vdots$

3

We have a binomial RV.

$M \sim \text{Binomial}(l, p)$

$$P[M=m] = \begin{cases} {}^{l}C_m (1-p)^{l-m} p^m, & m = 0, \ldots, l \\ 0 & \text{, otherwise.} \end{cases}$$

(c)

Y_i, $i = 1, 2, \ldots, l$ are Bern(q) 2

N is Binomial(l, q). 3

Steps as in part (b)

(d) $$Z = \sum_{i=1}^{l} X_i Y_i$$

$X_i Y_i$ is either 1 or 0.

$$P[X_i Y_i = 1] = P[X_i = 1, Y_i = 1]$$
$$= P[X_i = 1] P[Y_i = 1] = pq.$$

3

$\therefore$ $X_i Y_i$, $i = 1, 2, \ldots, l$ are Bern(pq)

Z is a sum of the Bern RV(s).

Z is Binomial(l, pq)

One could derive the PMF as in part (b).

For convenience, define $Z_i = X_i Y_i$.

$$P[Z=0] = P[\text{All } Z_i \text{ take a value } 0]$$

$$= \left(P[Z_1=0]\right)\left(P[Z_2=0]\right) \cdots\cdots \left(P[Z_\ell=0]\right)$$

$$= (1-pq)^\ell$$

7

$\vdots$

$$P[Z=z] = \begin{cases} {}^{\ell}C_z \, (1-pq)^{\ell-z} \, (pq)^z, & z=0,1,\ldots,\ell, \\ 0, & \text{otherwise.} \end{cases}$$

(e) We want to calculate

$$P[N=n \mid Z=z] \text{ for all } n \in R.$$

$$= \begin{cases} \dfrac{P[N=n, Z=z]}{P[Z=z]} & n \geq z \\ 0 & \text{otherwise} \end{cases}$$

5

$P[N=n, Z=z] = P[$Student attends n lectures and the attendance is $z]$

$= P[$Out of l lectures, student attends z during which attendance is recorded and $(n-z)$ during which attendance is not recorded$]$

10 ?

$$= {}^{l}C_{z} \; {}^{l-z}C_{n-z} \; (pq)^{z} \, (q(1-p))^{n-z} \, (1-q)^{l-n}$$

$$= \frac{\lfloor l}{\lfloor l-z \, \lfloor z} \; \frac{\lfloor l-z}{\lfloor n-z \, \lfloor l-n} \; (pq)^{z} (1-q)^{l-n} (q(1-p))^{n-z}$$

$$= \frac{l!}{z!\,(l-n)!\,(n-z)!}\,(pq)^{z}\,(1-q)^{l-n}\,\big(q(1-p)\big)^{n-z},$$

$$n \geq z.$$

$$P[Z = z] = \frac{l!}{z!\,(l-z)!}\,(1-pq)^{l-z}\,(pq)^{z}$$

$\therefore$

$$P[N=n \mid Z=z] = \begin{cases} \dfrac{(l-z)!}{(l-n)!\,(n-z)!}\,\dfrac{(1-q)^{l-n}\,\big(q(1-p)\big)^{n-z}}{(1-pq)^{l-z}}, & n \geq z \\[2ex] 0, & \text{otherwise.} \end{cases}$$

5

(f)

$$E[N \mid Z = z] =$$

$$\sum_{n=z}^{\ell} n \frac{\lfloor \ell - z}{\lfloor \ell - n \lfloor n - z} \frac{(1-q)^{\ell-n} (q(1-p))^{n-z}}{(1-pq)^{\ell-z}}.$$

8

+ 2 [If there is an attempt to simplify] (aa) Consider:

$$E[N \mid Z = z]$$

It is given that the recorded attendance is z. Therefore the number of lectures attended $N \geq z$.

5

$$\therefore E[N \mid Z = z] \geq z.$$

(bb) $$P[Y_i = 1 \mid Z_i = 0]$$

$$= \frac{P[Y_i = 1, Z_i = 0]}{P[Z_i = 0]}$$

3

$$= \frac{P[Y_i = 1, Z_i = 0]}{P[Y_i = 1, Z_i = 0] + P[Y_i = 0, Z_i = 0]}$$

7

$$= \frac{q(1-p)}{q(1-p) + (1-q)} = \frac{q(1-p)}{1-pq} \quad //$$

Note that we could calculate $1 - P[Z_i = 1]$.

(cc)

$E[N \mid Z = z] = z +$ Expected value of the sum of $\ell - z$ RV(s), where each is Bern$(P[Y_i = 1 \mid Z_i = 0])$

5

$= z +$ Sum of the expected value . . .

$= z + (\ell - z) P[Y_i = 1 \mid Z_i = 0]$

$= z + (\ell - z) \dfrac{q(1-p)}{1-pq}.$

15

(dd) $E[N] = \sum_{z=0}^{\ell} E[N \mid Z = z] P[Z = z]$ 2

$= \sum_{z=0}^{\ell} \left[z \left(1 - \dfrac{q(1-p)}{1-pq}\right) + \ell \dfrac{q(1-p)}{1-pq} \right] P[Z = z]$

$$= \sum_{z=0}^{\ell} z\, P[Z=z]\left(1 - \frac{q(1-p)}{1-pq}\right)$$

$$+ \frac{\ell\, q(1-p)}{1-pq}$$

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

$$= \ell q\left[\frac{p - pq + 1 - p}{1-pq}\right] = \ell q.$$

8