Maxwell Hanks – n12405710

MXB101 Problem Solving Task 4 (10%)

Question 1a.

Find $\mathbb{E}(N)$, the expected number of people an infectious person interacts with before infecting someone with the disease.

Solution

Given $N|P = p \sim \text{Geo}(p)$, where $N = 1,2,...$ with PMF:

$$\Pr(N = n|P = p) = (1-p)^{n-1}p$$

Using the total law of expectation:

$$\mathbb{E}(N) = \mathbb{E}\big(\mathbb{E}(N|P)\big)$$

Given $N|P = p \sim \text{Geo}(p)$, therefore:

$$\mathbb{E}(N|P = p) = \frac{1}{p}$$

Hence,

$$\mathbb{E}(N) = \mathbb{E}\left(\frac{1}{P}\right)$$

Given $f(p) = \frac{3}{2}p(2-p)$ where $0 < p < 1$:

$$\begin{aligned}
\mathbb{E}\left(\frac{1}{p}\right) &= \int_0^1 \frac{1}{p} f(p)\mathrm{d}p \\
&= \int_0^1 \frac{1}{p}\left(\frac{3}{2}p(2-p)\right)\mathrm{d}p \\
&= \frac{3}{2}\int_0^1 (2-p)\mathrm{d}p \\
&= \frac{3}{2}\left[2p - \frac{1}{2}p^2\right]_0^1 \\
&= \frac{3}{2}\left[2(1) - \frac{1}{2}(1)^2\right] \\
&= \frac{9}{4}
\end{aligned}$$

Hence $\mathbb{E}(N) = \frac{9}{4}$.

Question 1b.

What is the expected number of individuals infected from 100 interactions with the infectious person?

Solution

Given $\mathbb{E}(P) = \frac{5}{8}$ and there are 100 interactions, the expected number of infected individuals Q is given by:

$$\begin{aligned}
Q &= N \cdot \mathbb{E}(P) \\
&= 100 \cdot \frac{5}{8} \\
&= 62.5
\end{aligned}$$

Hence there are 62.5 expected infected individuals.

Question 2.

Consider now that the duration of the infection (i.e., the length of time that a person has the disease), T, is dependent upon the infectivity such that

$$f(t|p) = \frac{2-pt}{2(2-p)}, \qquad 0 < p < 1, \qquad 0 < t < 2$$

Question 2a.

Find $f(t,p)$.

Solution

From our initial context, we have:

$$f(p) = \frac{3}{2}p(2-p), \qquad 0 < p < 1$$

and also

$$f(t|p) = \frac{2-pt}{2(2-p)}, \qquad 0 < t < 2$$

Hence, the joint density function is:

$$\begin{aligned} f(t,p) &= f(t|p)f(p) \\ &= \left(\frac{2-pt}{2(2-p)}\right) \times \frac{3}{2}p(2-p) \\ &= \frac{2-pt}{2} \times \frac{3}{2}p \\ \therefore\ f(t,p) &= \frac{3}{4}p(2-pt) \end{aligned}$$

Where $0 < p < 1,\ 0 < t < 2$.

Question 2b.

Show that $\text{Cov}(T,P) = -\frac{1}{48}$.

Solution:

To find *covariance* we show that:

$$\text{Cov}(T,P) = \mathbb{E}(TP) - \mathbb{E}(T)\mathbb{E}(P)$$

Given $\mathbb{E}(P) = \frac{5}{8}$, $\mathbb{E}(TP)$ and $\mathbb{E}(T)$ can be determined. Let us define the bounds given by $0 < p < 1,\ 0 < t < 2$ on some region $\mathbb{R}^2$.

$$\begin{aligned} \mathbb{E}(T) &= \iint_{\mathbb{R}^2} tf(t,p)\mathrm{d}p\mathrm{d}t \\ &= \frac{3}{4}\int_0^2\int_0^1 tp(2-pt)\mathrm{d}p\mathrm{d}t \\ &= \frac{3}{4}\int_0^2 t\left[p^2 - \frac{1}{3}tp^3\right]_{p=0}^{p=1}\mathrm{d}t \\ &= \frac{3}{4}\int_0^2 t\left(1-\frac{1}{3}t\right)\mathrm{d}t \\ &= \frac{3}{4}\left[\frac{1}{2}t^2 - \frac{1}{6}t^3\right]_{t=0}^{t=2} \\ \therefore\ \mathbb{E}(T) &= \frac{5}{6} \end{aligned}$$

To calculate $\mathbb{E}(TP)$:

$$
\begin{aligned}
\mathbb{E}(TP) &= \iint_{\mathbb{R}^2} tpf(t,p)\mathrm{d}p\mathrm{d}t \\
&= \frac{3}{4}\int_0^2 \int_0^1 tp^2(2-pt)\mathrm{d}p\mathrm{d}t \\
&= \frac{3}{4}\int_0^2 t\left[\frac{2}{3}p^3 - \frac{1}{4}tp^4\right]_{p=0}^{p=1}\mathrm{d}t \\
&= \frac{3}{4}\int_0^2 \frac{2}{3}t - \frac{1}{4}t^2\mathrm{d}t \\
&= \frac{3}{4}\left[\frac{1}{3}t^2 - \frac{1}{12}t^3\right]_{t=0}^{t=2} \\
\therefore \mathbb{E}(TP) &= \frac{1}{2}
\end{aligned}
$$

Substituting values:

$$
\begin{aligned}
&\mathrm{Cov}(T,P) = \mathbb{E}(TP) - \mathbb{E}(T)\mathbb{E}(P) \\
&= \frac{1}{2} - \frac{5}{8}\left(\frac{5}{6}\right) \\
&\therefore \mathrm{Cov}(T,P) = -\frac{1}{48}
\end{aligned}
$$

Hence, as given previously, $\mathrm{Cov}(T,P) = -\frac{1}{48}$.

Question 2c.

Are infections that last longer more infectious on average?

Solution

Since $\mathrm{Cov}(T,P) = -\frac{1}{48}$ which is negative covariance, this indicates a negative association between duration T and infectivity P, hence, infections that last longer tend to on average have *lower* infectivity.

Question 3a.

Write the transition probability matrix for the Markov chain.

Solution

Since there are two possible states a person can be in – infected, and not infected, let us denote a 2x2 transition probability matrix:

$$\mathbf{P} = \begin{array}{cc} & X_{t+1} \\ X_t & \begin{bmatrix} p_{n,n} & p_{y,n} \\ p_{y,n} & p_{y,y} \end{bmatrix} \end{array}$$

Where n indicates non-infectious, and y indicates infectious. The probability of being infected one year and infected the next, hence $p_{y,y} = 0.4$. The probability of being not infected one year and also not infected the next, $p_{n,n} = k$. Given the property of a Markov chain where $\sum_j p_{i,j} = 1$, $p_{y,n} = 1 - p_{y,y} = 0.6$, and $p_{n,y} = 1 - p_{n,n} = 1 - k$. Hence:

$$\mathbf{P} = \begin{array}{cc} & X_{t+1} \\ X_t & \begin{bmatrix} k & 1-k \\ 0.6 & 0.4 \end{bmatrix} \end{array}$$

Question 3b.

While the researchers could not measure k directly, they were able to determine that, in the long run, 25% of individuals are infected in any given year. By considering the steady-state of the Markov chain, determine k.

Solution

Let $\tilde{\pi}^T$ be the stationary distribution denoted $\tilde{\pi}^T = \langle \pi_0, \pi_1 \rangle$ such that $\tilde{\pi}^T = \tilde{\pi}^T \mathbf{P}$, and $\sum_i \pi_i = \pi_0 + \pi_1 = 1$. Given that $\pi_1 = 0.25$, hence, $\pi_0 = 0.75$, from the first component of $\tilde{\pi}^T = \tilde{\pi}^T \mathbf{P}$:

$$\begin{aligned} \pi_0 &= \pi_0 k + \pi_1 \cdot 0.6 \\ 0.75 &= 0.75k + 0.25(0.6) \\ \therefore k &= 0.8 \end{aligned}$$

So in the long run, the probability of staying uninfected given being infectious the previous year is $k = 0.8$.