MATH 7241 F20

Problem Set #2

Solutions

MATH 7241: Problem Set #2

Due date: Friday October 2

Reading: relevant background material for these problems can be found in the class notes, and in Ross (Chapters 2,3,5) and in Grinstead and Snell (Chapters 1,2 3, 6).

Exercise 1 A typing firm has three typists A,B and C. The number of errors per 100 pages made by typist A is a Poisson random variable with mean 2.6; the number of errors per 100 pages made by typist B is a Poisson random variable with mean 3; the number of errors per 100 pages made by typist C is a Poisson random variable with mean 3.4. A manuscript of 300 pages is sent to the firm. Let X denote the number of errors in the typed manuscript.
a) Assume that one typist is randomly selected to do all the work. Find the mean and variance of X.
b) Assume instead that the work is divided into three equal parts which are given to the three typists. Find the mean and variance of X in this case.

a) $\mathbb{E}[X] = \mathbb{E}[X|A]\frac{1}{3} + \mathbb{E}[X|B]\frac{1}{3} + \mathbb{E}[X|C]\frac{1}{3}$

$\mathbb{E}[X^2] = \mathbb{E}[X^2|A]\frac{1}{3} + \mathbb{E}[X^2|B]\frac{1}{3} + \mathbb{E}[X^2|C]\frac{1}{3}$

Given A: Poisson, mean $= 3(2.6) = 7.8 = \lambda_A$.

$\Rightarrow \mathbb{E}[X|A] = \lambda_A$

$VAR[X|A] = \lambda_A \Rightarrow \mathbb{E}[X^2|A] = VAR[X|A] + (\mathbb{E}[X|A])^2$

$= \lambda_A + \lambda_A^2$

Similarly for $\lambda_B = 9$, $\lambda_C = 10.2$

$\Rightarrow \mathbb{E}[X] = \frac{1}{3}(\lambda_A + \lambda_B + \lambda_C) = 9$

$\mathbb{E}[X^2] = \frac{1}{3}(\lambda_A + \lambda_B + \lambda_C + \lambda_A^2 + \lambda_B^2 + \lambda_C^2) = 90.96$

$$\Rightarrow VAR[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = 9.96$$

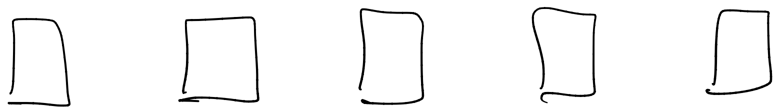
b) $X = X_A + X_B + X_C$

independent Poisson, rates $2.6, 3, 3.4$

$$\Rightarrow \mathbb{E}[X] = \mathbb{E}[X_A] + \mathbb{E}[X_B] + \mathbb{E}[X_C] = 9$$

$$VAR[X] = VAR[X_A] + VAR[X_B] + VAR[X_C] = 9$$

$\boxed{\textbf{Exercise 2}}$ In a variation of the classic Monty Hall game show, the host sets up five doors and hides prizes behind two of the doors. The contestant first guesses a door, and then the host opens one of the other four doors to show that it does not conceal a prize. The contestant is offered the opportunity to switch her guess to a different door. Should she switch or stay with her original choice? [Hint: see notes on the Monty Hall question, and try to imitate the solution provided there].



5 doors, 2 prizes

$R = \{\text{original guess was correct}\}$

$S = \{\text{correctly guesses after switching}\}$

$D = \{\text{correctly guesses without switching}\}$

$$\mathbb{P}(D) = \mathbb{P}(D \mid R)\,\mathbb{P}(R) + \mathbb{P}(D \mid R^c)\,\mathbb{P}(R^c)$$
$$= 1\left(\frac{2}{5}\right) + 0 = \frac{2}{5}$$

$$\mathbb{P}(S) = \mathbb{P}(S \mid R)\,\mathbb{P}(R) + \mathbb{P}(S \mid R^c)\,\mathbb{P}(R^c)$$
$$= \left(\frac{1}{3}\right)\left(\frac{2}{5}\right) + \left(\frac{2}{3}\right)\left(\frac{3}{5}\right)$$
$$= \frac{8}{15}$$

since $\mathbb{P}(S) > \mathbb{P}(D) \Rightarrow$ she should switch

$\boxed{\textbf{Exercise 3}}$ A fair die is rolled repeatedly. Let X_n be the result of the n^{th} roll. So X_n takes values $\{1, \ldots, 6\}$, each with probability 1/6, and the random variables $X_1, X_2, \ldots$ are all independent. Let

$$N = \min\{n : X_n = X_{n-1},\, n \geq 2\}$$

That is, N is the first roll where the result is equal to the previous roll. [e.g., if you roll the sequence $2, 3, 1, 4, 4, 6, \ldots$ then $N = 5$.] Find $E[N]$.
[Hint: argue that $\mathbb{E}[N|X_1 = a] = \mathbb{E}[N|X_1 = b]$ for all $a, b = 1, \ldots, 6$, and conclude that $\mathbb{E}[N] = \mathbb{E}[N|a]$ for any $a = 1, \ldots, 6$. Then condition on the outcomes of the first two rolls, and imitate the argument we used in class for the 'rat in a maze' problem to derive an equation for $\mathbb{E}[N]$.]

$\mathbb{E}[N \mid X_1 = a]$ is same for all $a = 1, \ldots, 6$

$$\Rightarrow \mathbb{E}[N] = \sum_{a=1}^{6} \mathbb{E}[N \mid X_1 = a]\, \mathbb{P}(X_1 = a) = \mathbb{E}[N \mid X_1 = 1]$$

Condition on X_2:

$$\mathbb{E}[N \mid X_1 = 1] = \sum_{b=1}^{6} \mathbb{E}[N \mid X_1 = 1, X_2 = b]\, \mathbb{P}(X_2 = b \mid X_1 = 1)$$

$$= \frac{1}{6} \mathbb{E}[N \mid X_1 = 1, X_2 = 1] + \frac{1}{6} \sum_{b=2}^{6} \mathbb{E}[N \mid X_1 = 1, X_2 = b]$$

$$= \frac{1}{6}(2) + \frac{1}{6} \sum_{b=2}^{6} \left(1 + \mathbb{E}[N \mid X_1 = b]\right)$$

$$= \frac{1}{3} + \frac{5}{6}\left(1 + \mathbb{E}[N \mid X_1 = 1]\right)$$

$$\Rightarrow \mathbb{E}[N] = \mathbb{E}[N \mid X_1 = 1] = 7$$

$\boxed{\textbf{Exercise 4}}$ Suppose that $\{X_i\}$ are IID uniform random variables on the interval $[-1, 1]$. Let Z be a standard normal random variable. Using the CLT, find the number a so that

$$\lim_{n\to\infty} P\Big(\sum_{i=1}^{n} X_i \geq \sqrt{n}\Big) = P(Z \geq a)$$

[Hint: you will need to find the mean and variance of X, which is uniform on $[-1, 1]$].

let $$Z_n = \frac{\sum_{i=1}^{n} X_i - n\mu}{\sqrt{n}\,\sigma}$$

CLT: $Z_n \to Z \sim N(0,1)$ as $n \to \infty$

$\mu = \mathbb{E}[X] = 0$ (clear by symmetry)

$\sigma^2 = \mathbb{E}[X^2] = \int_{-1}^{1} x^2 \cdot \frac{1}{2}\, dx = \frac{1}{3}$

$$\Rightarrow \mathbb{P}\Big(\sum_{i=1}^{n} X_i \geq \sqrt{n}\Big) = \mathbb{P}\Big(Z_n \geq \frac{\sqrt{n} - 0}{\sqrt{n}\ 1/\sqrt{3}}\Big)$$

$$= \mathbb{P}(Z_n \geq \sqrt{3})$$

$$\Rightarrow \lim_{n\to\infty} \mathbb{P}\Big(\sum_{i=1}^{n} X_i \geq \sqrt{n}\Big) = \mathbb{P}(Z \geq \sqrt{3})$$

$$\Rightarrow a = \sqrt{3}$$

$\boxed{\textbf{Exercise 5}}$ Randomly distribute r balls in n boxes so that the sample space consists of n^r equally likely elements. Let N_n be the number of empty boxes. Suppose that $r, n \to \infty$ in such a way that their ratio r/n converges to a constant value c. Show that $n^{-1}\,\mathbb{E}[N_n]$ converges as $n \to \infty$ and find the limiting value in terms of c.
[Hint: define $X_i = 1$ if the ith box is empty, and $X_i = 0$ if the ith box is not empty].

$$X_i = \begin{cases} 1 & \text{if } i^{\text{th}} \text{ box is empty} \\ 0 & \text{else} \end{cases}$$

$$\Rightarrow \quad N_n = \sum_{i=1}^{n} X_i$$

$$\Rightarrow \quad \mathbb{E}[N_n] = \sum_{i=1}^{n} \mathbb{E}[X_i] = n\,\mathbb{P}(X_i = 1)$$

$$= n\left(1 - \frac{1}{n}\right)^r$$

$$\Rightarrow \quad \frac{1}{n}\mathbb{E}[N_n] = \left(1 - \frac{1}{n}\right)^r = \left[\left(1 - \frac{1}{n}\right)^n\right]^{\frac{r}{n}}$$

<u>Fact:</u> $\left(1 + \frac{x}{n}\right)^n \to e^x$ as $n \to \infty$

also $\frac{r}{n} \to c$

$$\Rightarrow \quad \left[\left(1 - \frac{1}{n}\right)^n\right]^{\frac{r}{n}} \to \left(e^{-1}\right)^c = e^{-c}$$

$$\Rightarrow \quad \lim_{n \to \infty} \frac{1}{n}\mathbb{E}[N_n] = e^{-c}$$

Exercise 6 Suppose X is an exponential random variable. One of the following three formulas is correct:

$$\begin{aligned}
(a) \quad \mathbb{E}[X^2 \mid X > 1] &= \mathbb{E}[(X+1)^2] \\
(b) \quad \mathbb{E}[X^2 \mid X > 1] &= \mathbb{E}[X^2] + 1 \\
(c) \quad \mathbb{E}[X^2 \mid X > 1] &= (\mathbb{E}[X] + 1)^2
\end{aligned}$$

Without doing computations, use the memoryless property of the exponential distribution to explain which answer is correct.

Given $X > 1$, memoryless property says

$$X = 1 + X'$$

where $X' \sim$ exponential, same rate as X.

$$\Rightarrow \mathbb{E}[X^2 \mid X > 1] = \mathbb{E}[(1+X')^2 \mid X > 1]$$

$$= \mathbb{E}[(1+X')^2] \quad \text{as } X' \text{ indep. of } \{X > 1\}$$

$$= \mathbb{E}[(1+X)^2]$$