Tutorial 2

AID-521: Mathematics for Data Science

Abhishek Samantray, Vimal Kumar

August 10, 2021

Qn. 1.

Suppose that a random system of police patrol is devised so that a patrol officer may visit a given beat location $Y = 0, 1, 2, 3, ...$ times per half-hour period, with each location being visited an average of once per time period. Assume that Y possesses, approximately, a Poisson probability distribution.

(a) Calculate the probability that the patrol officer will miss a given location during a half-hour period.

(b) What is the probability that it will be visited once? Twice?

(c) What is the probability that it will be visited at least once?

Ans/Sol.

Qn. 2.

Arrivals of customers at a checkout counter follow a Poisson distribution. It is known that, during a given 30-minute period, one customer arrived at the counter.

(a) Find the probability that the customer arrived during the last 5 minutes of the 30-minute period.

Ans/Sol.

Qn. 3.

A local supermarket has three checkout counters. Two customers arrive at the counters at different times when the counters are serving no other customers. Each customer chooses a counter at random, independently of the other. Let Y_1 denote the number of customers who choose counter 1 and Y_2, the number who select counter 2.

(a) Find the joint probability function of Y1 and Y2.

(b) Find the values of joint cdf $F(y_1, y_2)$ at $(-1, 2), (1.5, 2)$, and $(5, 7)$.

Ans/Sol.

Qn. 4.

Gasoline is to be stocked in a bulk tank once at the beginning of each week and then sold to individual customers. Let Y_1 denote the proportion of the capacity of the bulk tank that is available after the tank is stocked at the beginning of the week. Because of the limited supplies, Y_1 varies from week to week. Let Y_2 denote the proportion of the capacity of the bulk tank that is sold during the week. Because Y_1 and Y_2 are both proportions, both variables take on values between 0 and 1. Further, the amount sold, y_2, cannot exceed the amount available, y_1. Suppose that the joint density function for Y_1 and Y_2 is given by

$$f(y_1, y_2) = \begin{cases} 3y_1, & 0 \le y_2 \le y_1 \le 1, \\ 0, & \text{elsewhere.} \end{cases}$$

(a) Find the probability that less than one-half of the tank will be stocked and more than one-quarter of the tank will be sold.

(b) Find the covariance between the amount in stock Y_1 and amount of sales Y_2.

Ans/Sol.

Qn. 5.

From a group of three Republicans, two Democrats, and one independent, a committee of two people is to be randomly selected. Let Y_1 denote the number of Republicans and Y_2 denote the number of Democrats on the committee.

(a) Find the joint probability function of Y_1 and Y_2.

(b) Find the marginal probability function of Y_1.

(c) Find the conditional distribution of Y_1 given that $Y_2 = 1$. That is, given that one of the two people on the committee is a Democrat, find the conditional distribution for the number of Republicans selected for the committee.

(d) Is the number of Republicans in the sample independent of the number of Democrats? (Is Y_1 independent of Y_2?)

Ans/Sol.

Qn. 6.

A soft-drink machine has a random amount Y_2 in supply at the beginning of a given day and dispenses a random amount Y_1 during the day (with measurements in gallons). It is not resupplied during the day, and hence $Y_1 \le Y_2$. It has been observed that Y_1 and Y_2 have a joint density given by

$$f(y_1, y_2) = \begin{cases} 1/2, & 0 \le y_1 \le y_2 \le 2, \\ 0, & \text{elsewhere.} \end{cases}$$

That is, the points (y_1, y_2) are uniformly distributed over the triangle with the given boundaries.

(a) Find the conditional density of Y_1 given $Y_2 = y_2$.

(b) Evaluate the probability that less than 1/2 gallon will be sold, given that the machine contains 1.5 gallons at the start of the day.

(c) Find the conditional expectation of the amount of sales, Y_1, given that $Y_2 = 1.5$.

Ans/Sol.

Qn. 7.

$$f(y_1, y_2) = \begin{cases} 6y_1y_2^2, & 0 \le y_1 \le 1,\ 0 \le y_2 \le 1, \\ 0, & \text{elsewhere.} \end{cases}$$

$$g(y_1, y_2) = \begin{cases} 2, & 0 \le y_2 \le y_1 \le 1, \\ 0, & \text{elsewhere.} \end{cases}$$

(a) Show that Y_1 and Y_2 are independent if their joint distribution is given by $f(y_1, y_2)$.

(b) Show that Y_1 and Y_2 are dependent if their joint distribution is given by $g(y_1, y_2)$.

Ans/Sol.

Qn. 8.

A process for producing an industrial chemical yields a product containing two types of impurities. For a specified sample from this process, let Y_1 denote the proportion of impurities in the sample and let Y_2 denote the proportion of type I impurities among all impurities found. Suppose that the joint distribution of Y_1 and Y_2 can be modeled by the following probability density function:

$$f(y_1, y_2) = \begin{cases} 2(1-y_1), & 0 \le y_1 \le 1,\ 0 \le y_2 \le 1, \\ 0, & \text{elsewhere.} \end{cases}$$

(a) Find the expected value of the proportion of type I impurities in the sample.

Ans/Sol.

Qn. 9.

Let Y_1 and Y_2 denote random variables. Then prove that

$$E[Y_1] = E[\ E[\ Y1|Y2\ ]\ ],$$

where on the right-hand side the inside expectation is with respect to the conditional distribution of Y_1 given Y_2 and the outside expectation is with respect to the distribution of Y_2.

Ans/Sol.

Qn. 10.

A quality control plan for an assembly line involves sampling $n = 10$ finished items per day and counting Y, the number of defectives. If p denotes the probability of observing a defective, then Y has a binomial distribution, assuming that a large number of items are produced by the line. But p varies from day to day and is assumed to have a uniform distribution on the interval from 0 to 1/4.

(a) Find the expected value of Y .

Ans/Sol.

Qn. 11.

Let $Y_1, ..., Y_n$ be independent random variables with $E[Y_i] = \mu$ and $Var(Y_i) = \sigma^2$. (These variables may denote the outcomes of n independent trials of an experiment.) Define

$$\bar{Y} = \frac{1}{n} \sum_{i=1}^{n} Y_i$$

(a) Show that $E[\bar{Y}] = \mu$.

(b) Show that $Var(\bar{Y}) = \sigma^2/n$.

Qn. 12.

The number of defectives Y in a sample of $n = 10$ items selected from a manufacturing process follows a binomial probability distribution. An estimator of the fraction defective in the lot is the random variable $\hat{p} = Y/n$.

(a) Find the expected value of $\hat{p}$.

(b) Find the variance of $\hat{p}$.

Ans/Sol.

1) $P(X=x) = \frac{e^{-\lambda} \lambda^{x}}{x!}$; $x = 0, 1, 2 \dots$

2) Uniform prob. for each 5 min block.

3) (a)

		Y_2	
Y_1	0	1	2
0		..	
1		..	
2		..	

→ equally likely

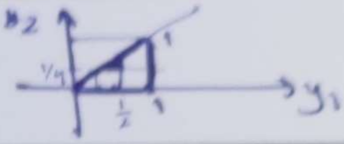
$S = \{(1,1), (1,2), (1,3), \dots, (3,3)\}$
↳ equally likely ... 1/9

$Y_1 \rightarrow 2, 1, 1, 1, 0, 0, 1, 0, 0$

$Y_2 \rightarrow 0, 1, 0, 1, 2, 1, 0, 1, 0$

Y_1 \ Y_2	0	1	2
0	1/9		1/9
1		2/9	0
2	1/9	0	0

(b) $(0,0) + (0,1) + (0,2)$
$+ (1,0) + (1,1) + (1,2)$



4) (a) $P(0.25 \le Y_1 \le 0.5, \ Y_2 > 0.25)$

$$= \int_0^{0.5}\left(\int_{0.25}^{y_1} 3y_1 \, dy_2\right) dy_1$$

$$= \int_{0.25}^{0.5} \int_{0.25}^{y_1} 3y_1 \, dy_2 \, dy_1$$

$$= \int_{0.25}^{0.5} 3y_1 \left(\int_{0.25}^{y_1} dy_2\right) dy_1$$

$$= \int_{1/4}^{1/2} 3y_1 \left(y_1 - \frac{1}{4}\right) dy_1 = \dots$$

$$= 5/128$$

(b) $E(Y_1 Y_2) = \int_0^1 \int_0^{y_1} y_1 y_2 \, 3y_1 \, dy_2 \, dy_1$

$= \dots = 3/10$

$E[Y_1] = 3/4 \quad E[Y_2] = 3/8$

$Cov(Y_1, Y_2) = \frac{3}{10} - \frac{3}{4} \cdot \frac{3}{8} = 0.02$

$\binom{4}{2} = \frac{4 \times 3 \times \cancel{2!}}{2 \times \cancel{2!}}$ $\quad \frac{6!}{2!\,4!} = \frac{6 \times 5}{2} = 15$ $\quad \frac{1!}{0!\,(1-0)!} = 1$

5) 3R, 2D, 1i | 6 $\gg$ 2

(a)

		Y_1		
		0	1	2
Y_2	0	●	3/15	3/15
	1	2/15	4/15	●
	2	1/15	●	●

Combinatorics: so, no need for enlisting outcomes.

$$P(Y_1 = 1, Y_2 = 1) = \frac{\binom{3}{1}\binom{2}{1}}{\binom{6}{2}}$$ ← sample space size

(d) $p(y_1, y_2) = p(y_1) \cdot p(y_2) \quad \forall y_1, y_2$

$p(0,0) = 0$

$p_1(0) = \frac{\binom{3}{2}}{15} = \frac{3}{15}, \quad p_2(0) = 6/15$

Get a feeling!!

6) <a> use = Y_1 $\leq$ tot = Y_2

$f(y_1 | Y_2 = y_2) = \cancel{\int_{y_2} f(y_1, y_2)}$

$= \dfrac{f(y_1, y_2)}{\int_{-\infty}^{\infty} f(y_1, y_2)\, dy_1} = \dfrac{1/2}{\begin{cases} 1/2\, y_2 ; & 0 \leq y_2 \leq 2 \\ 0 ; & \text{else} \end{cases}} = 1/y_2$

$[0 < y_2 \leq 2]$

<b> $P(Y_1 \leq 1/2 | Y_2 = 1.5)$

$= \int_0^{1/2} f(y_1 | y_2 = 1.5)\, dy_1$

$= \int_0^{1/2} \left(\frac{1}{y_2} \middle| y_2 = 1.5\right) dy_1 = \int_0^{1/2} \frac{1}{1.5}\, dy_1$

$= \frac{1}{2} \cdot \frac{1}{1.5} = \frac{1}{3}$

<c> $E[Y_1 | y_2 = 1.5]$

$= \int_0^{1.5 = y_2} y_1 f(y_1 | y_2 = 1.5)\, dy_1$

$= \int_0^{y_2} \frac{y_1}{y_2}\, dy_1 = \frac{1}{y_2} * \frac{y_2^2}{2} = \frac{y_2}{2}$

$= 0.75$

7) <a> $f_{Y_1}(y_1) = \int_0^1 6 y_1 y_2^2 \, dy_2$ $\quad 0 < y_1 \le 1$

$= 2y_1$ $\quad$ else 0

$f_{Y_2}(y_2) = \int_0^1 6 y_1 y_2^2 \, dy_1$

$= 3y_2^2$

<b> $f_{Y_1}(y_1)$ $\quad 0 \le y_1 \le 1$

$= \begin{cases} \int_0^{y_1} 2 \, dy_2 = 2y_1 \\ 0 \quad \text{else} \end{cases}$

$f_{Y_2}(y_2) = \begin{cases} \int_{y_2}^1 2 \, dy_1 = 2(1-y_2) \\ 0 \end{cases}$

$\therefore f(y_1, y_2) \ne f_{Y_1}(\cdot) f_{Y_2}(\cdot)$

8) Prop^n of Type I impurities in the entire sample $= Y_1 \cdot Y_2$

$\therefore E[Y_1 Y_2] = \int_0^1 \int_0^1 y_1 y_2 \, 2(1-y_1) \, dy_2 \, dy_1$

$= 1/6.$

11) Do yourself ... just apply definit^n.

9) Thm ✓ [SOL]

10) (*) $E[Y_1] = \underset{(Y_2)}{E}\left[\underset{(Y_1)}{E[Y_1 | Y_2]}\right]$

$E[Y]$, given p, $= np$
(binomial dist^n)

$E[Y] = E\left[E[Y|p]\right] = E(np)$

$= n \cdot E(p) = n \cdot \left(\frac{0 + 1/4}{2}\right)$

↳ uniform dist^n

$= n/8$

$= \frac{10}{8}$ (as $n = 10$).

$= 1.25$

↳ Long Run expected # defect / day

12) $E(\hat{p}) = E\left(\frac{Y}{n}\right)$

$= \frac{1}{n} E(Y)$ $\{$ $\because$ n is constant

$= \frac{1}{n} \times np$ $\{$ $\because$ Exp. value of binomial $= np$.

$= p$

$Var(\hat{p}) = Var\left(\frac{Y}{n}\right)$

$= \left(\frac{1}{n}\right)^2 Var(Y)$ $\{$ $\because$ n is constant (use property of variance)

$= \frac{1}{n^2} \cdot np(1-p)$

$= \frac{p(1-p)}{n}$.