Experimental Statistics for Engineers I

Kamala Dadashova

Problem 1:
The joint probability distribution of the number X of cars and the number Y of buses per signal cycle at a proposed left-turn lane is displayed in the accompanying joint probability table.

	Y=0	Y=1	Y=2	Total
X=0	0.025	0.015	0.01	0.05
X=1	0.050	0.03	0.02	0.1
X=2	0.125	0.075	0.05	0.25
X=3	0.150	0.09	0.06	0.3
X=4	0.100	0.06	0.04	0.2
X=5	0.050	0.03	0.02	0.1
Total	0.5	0.3	0.2	1

Table 1: Joint probability table

(a) What is the probability that there is exactly one car and exactly one bus during a cycle?

$$p(1,1) = P(X = 1 \; and \; Y = 1) = 0.030$$

(b) What is the probability that there is at most one car and at most one bus during a cycle?

$$P(X \leq 1 \; and \; Y \leq 1) = p(0,0) + p(0,1) + p(1,0) + p(1,1) = 0.025 + 0.015 + 0.050 + 0.030 =$$

0.12 (c) What is the probability that there is exactly one car during a cycle? Exactly one bus?

$$P(X = 1) = p(1,0) + p(1,1) + p(1,2) = 0.050 + 0.030 + 0.020 = 0.1$$

$$P(Y = 1) = p(0,1) + p(1,1) + p(2,1) + p(3,1) + p(4,1) + p(5,1) =$$

$$0.015 + 0.030 + 0.075 + 0.090 + 0.060 + 0.030 = 0.3$$

(d) Suppose the left-turn lane is to have a capacity of five cars, and that one bus is equivalent to three cars. What is the probability of an overflow during a cycle?

$$P(X + 3Y > 5) = 1 - P(X + 3Y \leq 5) =$$

$$1-(p(0,0)+p(1,0)+p(2,0)+p(3,0)+p(4,0)+p(5,0)+p(0,1)+p(1,1)+p(2,1)) = 1-0.62 = 0.38$$

(e) Are X and Y independent rv's? Explain.

Marginal probabilities for X:

$$p_X(0) = 0.05,\ p_X(1) = 0.1,\ p_X(2) = 0.25,\ p_X(3) = 0.3,\ p_X(4) = 0.2,\ p_X(5) = 0.1$$

Marginal probabilities for Y:

$$p_Y(0) = 0.5,\ ,p_Y(1) = 0.3,\ p_Y(2) = 0.2,$$

$$0.025 = p(0,0) = p_X(0)p_Y(0) = 0.050.5 = 0.025$$
$$0.050 = p(1,0) = p_X(1)p_Y(0) = 0.10.5 = 0.05$$
$$0.125 = p(2,0) = p_X(2)p_Y(0) = 0.250.5 = 0.125$$
$$0.15 = p(3,0) = p_X(3)p_Y(0) = 0.30.5 = 0.15$$
$$0.1 = p(4,0) = p_X(4)p_Y(0) = 0.20.5 = 0.1$$
$$0.05 = p(5,0) = p_X(5)p_Y(0) = 0.10.5 = 0.05$$

$$0.015 = p(0,1) = p_X(0)p_Y(1) = 0.050.3 = 0.015$$
$$0.03 = p(1,1) = p_X(1)p_Y(1) = 0.10.3 = 0.03$$
$$0.075 = p(2,1) = p_X(2)p_Y(1) = 0.250.3 = 0.075$$
$$0.09 = p(3,1) = p_X(3)p_Y(1) = 0.30.3 = 0.09$$
$$0.06 = p(4,1) = p_X(4)p_Y(1) = 0.20.3 = 0.06$$
$$0.03 = p(5,1) = p_X(5)p_Y(1) = 0.10.3 = 0.03$$

$$0.01 = p(0,2) = p_X(0)p_Y(2) = 0.050.2 = 0.01$$
$$0.02 = p(1,2) = p_X(1)p_Y(2) = 0.10.2 = 0.02$$
$$0.05 = p(2,2) = p_X(2)p_Y(2) = 0.250.2 = 0.05$$
$$0.06 = p(3,2) = p_X(3)p_Y(2) = 0.30.2 = 0.06$$
$$0.04 = p(4,2) = p_X(4)p_Y(2) = 0.20.2 = 0.04$$
$$0.02 = p(5,2) = p_X(5)p_Y(2) = 0.10.2 = 0.02$$

So, X and Y are independent.

Problem 2:
You have two lightbulbs for a particular lamp. Let X= the lifetime of the first bulb and Y= the lifetime of the second bulb (both in 1000s of hours). Suppose that X and Y are independent and that each has an exponential distribution with parameter $\lambda = 1$.

(a) What is the joint pdf of X and Y?

Since the events are independent, using pdf of exponential distribution, we get

$$f(x,y) = f_X(x)f_Y(y) = \begin{cases} e^{-x-y} & x \geq 0, y \geq 0 \\ 0 & otherwise \end{cases}$$

(b) What is the probability that each bulb lasts at most 1000 hours (i.e., $X \leq 1 and Y \leq 1$)?

Using independence of events and cdf for exponential distribution,

$$P(X \leq 1 \text{ and } Y \leq 1) = P(X \leq 1)P(Y \leq 1) = (1 - e^{-1})(1 - e^{-1}) \approx 0.4$$

(c) What is the probability that the total lifetime of the two bulbs is at most 2? [Hint: Draw a picture of the region A=(x, y): $x \geq 0$, $y \geq 0$, $x + y \leq 2$ before integrating.]

$$P(X + Y \leq 2) = \int_0^2 \int_0^{2-x} e^{-x-y} \, dy \, dx = \int_0^2 e^{-x} - e^{-2} \, dx = -e^{-2} - 2e^{-2} + 1 = 0.594$$

(d) What is the probability that the total lifetime is between 1 and 2?

$$P(1 \leq X + Y \leq 2) = P(X + Y \leq 2) - P(X + Y \leq 1) = 0.594 - 0.264 = 0.33$$

where

$$P(X + Y \leq 1) = \int_0^1 \int_0^{1-x} e^{-x-y} \, dy \, dx = \int_0^1 e^{-x} - e^{-1} \, dx = -e^{-1} - e^{-1} + 1 = 0.264$$

Problem 3:
Consider a small ferry that can accommodate cars and buses. The toll for cars is \$3, and the toll for buses is \$10. Let X and Y denote the number of cars and buses, respectively, carried on a single trip. Suppose the joint distribution of X and Y is as given in the table of Exercise 7. Compute the expected revenue from a single trip.
Let $R(X,Y) = 3X + 10Y$ be a revenue function. Using Table 1, we get the following,

$$E[h(X,Y)] = \sum_{X=0}^{5} \sum_{Y=0}^{2} (3X + 10Y)p(x,y) = 15.4$$

Problem 4: Let X be the number of packages being mailed by a randomly selected customer at a certain shipping facility. Suppose the distribution of X is as follows:

x	1	2	3	4
p(x)	0.4	0.3	0.2	0.1

Table 2: Distribution of X

(a) Consider a random sample of size n=2 (two customers), and let $\bar{X}$ be the sample mean number of packages shipped. Obtain the probability distribution of $\bar{X}$.

x_1	x_2	$p(x_1, x_2)$	$\bar{x}$	R(range)
1	1	0.16	1	0
1	2	0.12	1.5	1
1	3	0.08	2	2
1	4	0.04	2.5	3
2	1	0.12	1.5	1
2	2	0.09	2	0
2	3	0.06	2.5	1
2	4	0.03	3	2
3	1	0.08	2	2
3	2	0.06	2.5	1
3	3	0.04	3	0
3	4	0.02	3.5	1
4	1	0.04	2.5	3
4	2	0.03	3	2
4	3	0.02	3.5	1
4	4	0.01	4	0

Table 3

In table 3, we provide all 16 cases along with probabilities, $\bar{x}$, and range values. The probability distribution of $\bar{X}$ is given in Table 4.

$\bar{x}$	1	1.5	2	2.5	3	3.5	4
p($\bar{x}$)	0.16	0.24	0.25	0.2	0.1	0.04	0.01

Table 4: Probability distribution of $\bar{X}$

(b) Refer to part (a) and calculate $P(\bar{X} \leq 2.5)$.

$$P(\bar{X} \leq 2.5) = 0.16 + 0.24 + 0.25 + 0.2 = 0.85$$

(c) Again consider a random sample of size n=2, but now focus on the statistic R= the sample range (difference between the largest and smallest values in the sample). Obtain the distribution of R. [Hint: Calculate the value of R for each outcome and use the probabilities from part (a).]

Using Table 3, we construct Table 5 for R

r	0	1	2	3
p(r)	0.3	0.4	0.22	0.08

Table 5: Probability distribution of sample range

(d) If a random sample of size n=4 is selected, what is $P(\bar{X} \leq 1.5)$? [Hint: You should not have to list all possible outcomes, only those for which $\bar{x} \leq 1.5$.]

$$P(\bar{X} \leq 1.5) = P(1,1,1,1) + P(2,1,1,1) + P(1,2,1,1) + P(1,1,2,1) + P(1,1,1,2)$$
$$+P(1,1,3,1) + P(3,1,1,1) + P(1,3,1,1) + P(1,1,1,3)+$$
$$+P(1,1,2,2) + P(1,2,1,2) + P(2,2,1,1) + P(2,1,2,1) + P(1,2,2,1) + P(2,1,1,2) =$$
$$0.4^4 + 4 \cdot 0.3 \cdot 0.4^3 + 4 \cdot 0.2 \cdot 0.4^3 + 6 \cdot 0.4^2 \cdot 0.3^2 = 0.24$$

Problem 5:
Let $X_1, X_2,$ *and* X_3 represent the times necessary to perform three successive repair tasks at a certain service facility. Suppose they are independent, normal rv's with expected values μ_1, μ_2, μ_3 and variances σ_1^2, σ_2^2, and σ_3^2, respectively.

(a) If $\mu_1 = \mu_2 = \mu_3 = 60$, and $\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = 15$, calculate $P(T_0 \leq 200)$ and $P(150 \leq T_0 \leq 200)$.

$$E(X_1 + X_2 + X_3) = 3 \cdot 60 = 180$$
$$V(X_1 + X_2 + X_3) = 3 \cdot 15 = 45$$
$$SD(X_1 + X_2 + X_3) = \sqrt{45}$$

$$P(X_1 + X_2 + X_3 \leq 200) = pnorm(200, 180, sqrt(45)) = 0.9986$$

$$P(150 \leq X_1+X_2+X_3 \leq 200) = pnorm(200, 180, sqrt(45)) - pnorm(150, 180, sqrt(45)) \approx 0.9986$$

(b) Using μ_i and σ_i given in part(a), calculate both $P(\bar{X} \geq 55)$ and $P(58 \leq \bar{X} \leq 62)$

$$\mu_{\bar{X}} = \mu = 60 \, \sigma_{\bar{X}} = \frac{\sigma_X}{\sqrt{n}} = \frac{\sqrt{15}}{\sqrt{3}}$$

$$P(\bar{X} \geq 55) = 1 - pnorm(55, 60, sqrt(15/3)) = 0.9873$$
$$P(58 \leq \bar{X} \leq 62) = pnorm(62, 60, sqrt(15)/sqrt(3)) - pnorm(58, 60, sqrt(15)/sqrt(3)) = 0.6289$$

(c) Using μ_i and σ_i given in part(a), calculate and interpret $P(-10 \leq X_1 - 0.5X_2 - 0.5X_3 \leq 5)$

$$E(X_1 - 0.5X_2 - 0.5X_3) = \mu_1 - 0.5\mu_2 - 0.5\mu_3 = 0$$
$$V(X_1 - 0.5X_2 - 0.5X_3) = \sigma_1^2 + 0.25\sigma_2^2 + 0.25\sigma_3^2 = 1.5 * 15 = 22.5$$
$$SD(X_1 - 0.5X_2 - 0.5X_3) = \sqrt{22.5}$$

$$P(-10 \le X_1 - 0.5X_2 - 0.5X_3 \le 5) = pnorm(5, 0, sqrt(22.5)) - pnorm(-10, 0, sqrt(22.5)) \approx 0.8366$$

Here, we have three random variables, namely, X_1, X_2, X_3. The probability we found above represent the linear combination of these random variables. We know that expectation value of random variables forming linear combination is the sum/difference of those random variables expectation value,i.e $E(a_1X_1 + a_2X_2 + a_3X_3) = a_1E(X_1) + a_2E(X_2) + a_3E(X_3)$. Also, $Var(a_1X_1 + a_2X_2 + a_3X_3) = a_1^2Var(X_1) + a_2^2Var(X_2) + a_3^2Var(X_3)$. In order to find, probability of linear combination $X_1 - 0.5X_2 - 0.5X_3$, we firstly find expectation value and standard deviation of this linear combination. Finally, we find $P(-10 \le X_1 - 0.5X_2 - 0.5X_3 \le 5)$ using R.

(d) If $\mu_1 = 40,\ \mu_2 = 50, \mu_3 = 60$, and $\sigma_1^2 = 10, \sigma_2^2 = 12, \sigma_3^2 = 14$, calculate $P(X_1 + X_2 + X_3 \le 160)$ and $P(X_1 + X_2 \ge 2X_3)$.

$$E(X_1 + X_2 + X_3) = 150$$
$$V(X_1 + X_2 + X_3) = 36$$
$$SD(X_1 + X_2 + X_3) = 6$$

$$P(X_1 + X_2 + X_3 \le 160) = pnorm(160, 150, 6) = 0.9522$$

Then,

$$P(X_1 + X_2 \ge 2X_3) = P(X_1 + X_2 - 2X_3 \ge 0)$$

$$E(X_1 + X_2 - 2X_3) = 40 + 50 - 120 = -30$$
$$V(X_1 + X_2 - 2X_3) = \sigma_1^2 + \sigma_2^2 + 4\sigma_3^2 = 10 + 12 + 4 \cdot 14 = 78$$
$$SD(X_1 + X_2 + X_3) = \sqrt{78}$$

$$P(X_1 + X_2 - 2X_3 \ge 0) = pnorm(0, -30, sqrt(78), lower.tail = F) = 0.0003.$$