1 If $X \sim B(n_1, p)$, $Y \sim B(n_2, p)$, and X and Y are independent each other, what is the distribution of $X + Y$? Answer the question using the moment generating function of the Binomial distribution.

Solution. Use the fact that the moment generating function of binomial distribution $B(n, p)$ is

$$(1 - p + pe^t)^n$$

Since X and Y are independent each other, the moment generating function of $X + Y$ is

$$\begin{aligned} M_{X+Y}(t) &= M_X(t) \times M_Y(t) \\ &= (1 - p + pe^t)^{n_1} \times (1 - p + pe^t)^{n_2} \\ &= (1 - p + pe^t)^{n_1 + n_2} \end{aligned}$$

which means that

$$X + Y \sim B(n_1 + n_2, p)$$

2

20 points

(a) The number of people entering the intensive care unit at a hospital on any single day possesses a Poisson distribution with a mean equal to five persons per day. Let X be the number of people entering the intensive care unit on a particular day. Given that $X > 1$, what is the probability of $X > 2$?

(b) If events are occurring in time according to a Poisson distribution with mean λ, then the interarrival times between events have an exponential distributions with mean $1/\lambda$. If calls come into a police emergency center at the rate of 10 per hour, what is the probability that more than 15 minutes will elapse between the next two calls?

Solution. (a) Since $X \sim Pois(5)$,

$$\begin{aligned} P(X > 2 \mid X > 1) &= \frac{P(X > 2, X > 1)}{P(X > 1)} \\ &= \frac{P(X > 2)}{P(X > 1)} \\ &= \frac{1 - \frac{5^0 e^{-5}}{0!} - \frac{5^1 e^{-5}}{1!} - \frac{5^2 e^{-5}}{2!}}{1 - \frac{5^0 e^{-5}}{0!} - \frac{5^1 e^{-5}}{1!}} \\ &= 0.9122 \quad \textbf{(+10 points)} \end{aligned}$$

(b) Let X be the interarrival time between next two calls. Then since the events are occurring with Poisson distribution with mean 10, X has an exponential distribution with mean $1/10$ which means that $X \sim Exp(10)$. Therefore, the desired probability is

$$P(X > 1/4) = e^{-10/4} = 0.0821 \quad \textbf{(+10 points)}$$

- In 2(a), one who gets the equation $P(X > 2)/P(X > 1)$ can get **(+3 points)**.
- In 2(b), one who finds the distribution of the interarrival time can get **(+5 points)**.

3
20 points

According to Zimmels, the sizes of particles used in sedimentation experiments often have a uniform distribution. In sedimentation involving mixtures of particles of various sizes, the larger particles hinder the movements of the smaller ones. Thus, it is important to study both the mean and the variance of particle sizes. Suppose that spherical particles have diameters that are uniformly distributed between .01 and .05 centimeters. Find the mean and variance of the volumes of these particles.

Solution. Let S and V be the diameter and volume of a spherical particle respectively. Then $S \sim U(0.01, 0.05)$ and

$$V = \frac{4}{3}\pi\left(\frac{S}{2}\right)^3 = \frac{1}{6}\pi S^3$$

Then the mean of the volumes of the particles is

$$\mathbb{E}(V) = \mathbb{E}\left(\frac{\pi}{6}S^3\right) = \frac{\pi}{6}\int_{0.01}^{0.05} s^3 \frac{1}{0.05-0.01} ds = 6.5 \cdot 10^{-6}\pi \quad \textbf{(+10 points)}$$

Also, since

$$\mathbb{E}(V^2) = \mathbb{E}\left(\frac{\pi^2}{36}S^6\right) = \frac{\pi^2}{36}\int_{0.01}^{0.05} s^6 \frac{1}{0.05-0.01} ds = 7.750 \cdot 10^{-11}\pi^2 \quad \textbf{(+5 points)}$$

, the variance of the volumes of the particles is

$$\mathbb{V}(V) = \mathbb{E}(V^2) - \mathbb{E}(V)^2 = 3.525 \cdot 10^{-11}\pi^2 \quad \textbf{(+5 points)}$$

- One who calculates the correct values of the mean and variance through other ways also can get full points.
- One who writes the correct integral equations of the mean and variance with wrong final values can get **(+5 points)** each.

4
20 points

Scores on a certain nationwide college entrance examination follow a normal distribution with a mean of 500 and a standard deviation of 100.

(a) If a school only admits students who score over 680, what proportion of the student pool would be eligible for admission?

(b) What limit would you set that makes 50% of the students eligible?

(c) What should be the limit if only the top 15% are to be eligible?

Solution. (a) Let X be an entrance examination score of the student and $Z = \frac{X-500}{100}$. Then $X \sim N(500, 100^2)$ and $Z \sim N(0,1)$ (standard normal distribution). Then the proportion of the students who score over 680 is

$$P(X > 680) = P\left(Z > \frac{680-500}{100}\right) = P(Z > 1.8) = 1 - 0.9641 = 0.0359 \quad \textbf{(+6 points)}$$

(b) Let the limit which makes 50% of the students eligible is A. Then A should satisfy

$$P(X > A) = P\left(Z > \frac{A-500}{100}\right) = 0.5$$

which implies

$$\frac{A-500}{100} = 0$$

Therefore, $A = 500$ **(+7 points)**

(c) Let the limit which makes 15% of the students eligible is B. Then B should satisfy

$$P(X > B) = P\left(Z > \frac{B-500}{100}\right) = 0.15$$

and

$$P\left(Z < \frac{B-500}{100}\right) = 0.85$$

which implies

$$\frac{B-500}{100} = 1.04$$

Therefore, $B = 604$ **(+7 points)**.

- One who changes the probability into the probability related to Z (standard normal distribution) can get **(+3 points)** each.

5 (20 points) The daily amount of money spent on maintenance and repairs by a company is observed, over a long period of time, to be normal distributed with mean \$200 and standard deviation \$100. Assume that each day spending is independent and \$210 is budgeted for each business day (from Monday to Friday)

(a) What is the probability that the actual cost on a randomly selected day will exceed the budgeted amount?

(b) Find the probability that the mean daily cost (of five business days) of a randomly selected week exceeds the budgeted daily amount.

(c) Assume that there are four weeks in October. What is the probability that the mean daily cost exceeds the budgeted daily amount for at least one week?

(d) Assume that this budget plan will be sustained next year (52 weeks). What is the probability that the mean daily cost will exceed the budgeted daily amount in more than 25 weeks during the next year?

Solution. In this problem, the standard normal distribution will be denoted as Z.

(a) Let X be the actual cost on a day. Since $X \sim N(200, 100^2)$,

$$P(X > 210) = P\left(Z > \frac{210-200}{100}\right) = P(Z > 0.1) = 0.4602 \quad \textbf{(+5 points)}$$

(b) Let $\overline{X}$ be the mean daily cost of five days. Then since $\overline{X} \sim N(200, 100^2/5)$,

$$P(\overline{X} > 210) = P\left(Z > \frac{210-200}{100/\sqrt{5}}\right) = P\left(Z > \frac{\sqrt{5}}{10}\right) = 0.4129 \quad \textbf{(+5 points)}$$

(c) Let Y be the number of weeks in October whose mean daily cost exceeds \$210. Then by the answer of (b), Y has a binomial distribution $B(4, 0.4129)$. Therefore, the probability that the mean daily cost exceeds the budget for at least one week is

$$P(Y \geq 1) = 1 - (1 - 0.4129)^4 = 0.8812 \quad \textbf{(+5 points)}$$

(d) Let Y be the number of weeks in next week whose mean daily cost exceeds \$210. Then by the answer of (b), Y has a binomial distribution $B(52, 0.4129)$ which can be approximated as $N(21.47, 3.55^2)$. Therefore, the probability that the mean daily cost exceeds the budget for more thatn 25 weeks is

$$P(Y > 25) = P(Y \geq 25.5) = P(Z \geq 1.14) = 0.1271 \quad \textbf{(+5 points)}$$

- One who gets correct z values in 5(a),(b) can get **(+3 points)** each.
- In 5(d), one who correctly approximates the distribution into the normal distribution can get **(+3 points)**.

6
20 points

Suppose that (X_1, X_2, X_3, X_4) is a random sample from a normal distribution with mean μ and variance σ^2.

(a) Show that

$$U = \frac{(X_1 - X_2)^2}{2\sigma^2}$$

follows a χ^2 distribution with 1 df.

(b) Show that

$$V = \frac{(X_3 - X_4)}{\sqrt{(X_1 - X_2)^2}}$$

follows a t distribution with 1 df.

(c) Show that

$$W = \frac{(X_3 - X_4)^2}{(X_1 - X_2)^2}$$

follows a F distribution with $(1, 1)$ df.

Solution. (a) Since $X_1 - X_2 = N(0, 2\sigma^2)$,

$$\frac{X_1 - X_2}{\sqrt{2}\sigma} = N(0, 1)$$

Therefore,

$$U = \frac{(X_1 - X_2)^2}{2\sigma^2} = \left(\frac{X_1 - X_2}{\sqrt{2}\sigma}\right)^2 = \chi_1^2 \quad \textbf{(+6 points)}$$

(b) By (a),

$$\frac{(X_1 - X_2)^2}{2\sigma^2} = \chi_1^2 \text{ and } \frac{X_3 - X_4}{\sqrt{2}\sigma} = N(0, 1)$$

Therefore,

$$V = \frac{(X_3 - X_4)}{\sqrt{(X_1 - X_2)^2}} = \frac{\frac{(X_3 - X_4)}{\sqrt{2}\sigma}}{\sqrt{\frac{(X_1 - X_2)^2}{2\sigma^2}/1}} = t_1 \quad \textbf{(+7 points)}$$

(c) By (a),

$$\frac{(X_1 - X_2)^2}{2\sigma^2} = \chi_1^2 \text{ and } \frac{(X_3 - X_4)^2}{2\sigma^2} = \chi_1^2$$

Therefore,

$$W = \frac{(X_3 - X_4)^2}{(X_1 - X_2)^2} = \frac{\frac{(X_3 - X_4)^2}{2\sigma^2}/1}{\frac{(X_1 - X_2)^2}{2\sigma^2}/1} = F_{1,1} \quad \textbf{(+7 points)}$$

7 (5.12) The number of times that an individual contracts a cold in a given year is a Poisson random variable with parameter $\lambda = 3$. Suppose a new wonder drug (based on large quantities of vitamin C) has just been marketed that reduces the Poisson parameter to $\lambda = 2$ for 75 percent of the population. For the other 25 percent of the population, the drug has no appreciable effect on colds. If an individual tries the drug for a year and has 0 colds in that time, how likely is it that the drug is beneficial for him or her?

Solution. Let A be the event that the drug is beneficial for a individual and X be the number of times that a individual contracts a cold. Then,

$$\begin{aligned}
P(A \mid X = 0) &= \frac{P(A, X = 0)}{P(X = 0)} \\
&= \frac{P(A, X = 0)}{P(A, X = 0) + P(A^C, X = 0)} \\
&= \frac{P(X = 0 \mid A)P(A)}{P(X = 0 \mid A)P(A) + P(X = 0 \mid A^C)P(A^C)} \\
&= \frac{P(X = 0 \mid \lambda = 2)P(A)}{P(X = 0 \mid \lambda = 2)P(A) + P(X = 0 \mid \lambda = 3)P(A^C)} \\
&= \frac{e^{-2} \cdot 0.75}{e^{-2} \cdot 0.75 + e^{-3} \cdot 0.25} = 0.8908
\end{aligned}$$

7 (5.18) A contractor purchases a shipment of 100 transistors. It is his policy to test 10 of these transistors and to keep the shipment only if at least 9 of the 10 are in working condition. If the shipment contains 20 defective transistors, what is the probability it will be kept?

Solution. The desired probability is

$$
\begin{aligned}
&\frac{\text{the number of cases of at least 9 of the 10 are in working condition}}{\text{the number of cases of test 10 of 100 transistors}}\\
=&\frac{\text{9 of 10 are in working condition} + \text{10 of 10 are in working condition}}{\text{the number of cases of test 10 of 100 transistors}}\\
=&\frac{\binom{80}{10}+\binom{80}{9}\binom{20}{1}}{\binom{100}{10}} = 0.3630
\end{aligned}
$$

7 (5.26) The weekly demand for a product approximately has a normal distribution with mean 1000 and standard deviation 200. The current on hand inventory is 2200 and no deliveries will be occurring in the next two weeks. Assuming that the demands in different weeks are independent,

(a) what is the probability that the demand in each of the next 2 weeks is less than 1100?

(b) what is the probability that the total of the demands in the next 2 weeks exceeds 2200?

Solution. (a) Let W_1 and W_2 be the demand in the first week and the second week respectively. Then since

$$\begin{aligned} P(W_1 < 1100) &= P\left(\frac{W_1 - 1000}{200} < \frac{1100 - 1000}{200}\right) \\ &= P(Z < 0.5) = 0.6915 \end{aligned}$$

and $P(W_2 < 1100) = 0.6915$ (similarly),

$$P(W_1 < 1100, W_2 < 1100) = P(W_1 < 1100)P(W_2 < 1100) = 0.9615^2 = 0.4782$$

(b) Since W_1 and W_2 are independent and

$$E[W_1 + W_2] = 2000, \ Var[W_1 + W_2] = 200^2 + 200^2 = 80000$$

, $W_1 + W_2$ is a normal distribution $N(2000, 80000)$. Therefore,

$$\begin{aligned} P(W_1 + W_2 > 2200) &= P\left(\frac{W_1 + W_2 - 2000}{\sqrt{80000}} > \frac{2200 - 2000}{\sqrt{8000}}\right) \\ &= P(Z > 2/\sqrt{8}) = 0.2399 \end{aligned}$$

7 (5.41) Earthquakes occur in a given region in accordance with a Poisson process with rate 5 per year.

(a) What is the probability there will be at least two earthquakes in the first half of 2015?

(b) Assuming that the event in part (a) occurs, what is the probability that there will be no earthquakes during the first 9 months of 2016?

(c) Assuming that the event in part (a) occurs, what is the probability that there will be at least four earthquakes over the first 9 months of the year 2015?

Solution. (a) Let A be the number of earthquakes in the first half of 2015. Then A follows the Poisson distribution with $\lambda = 5/2 = 2.5$. Therefore, the probability is

$$P(A \geq 2) = 1 - P(A = 0) - P(A = 1) = 1 - e^{-2.5} - 2.5e^{-2.5} = 0.7127$$

(b) Let B be the number of earthquakes in the first 9 months of 2016. Then B follows the Poisson distribution with $\lambda = 15/4 = 3.75$. Then since A and B are independent,

$$P(B = 0 \mid A \geq 2) = P(B = 0) = e^{-3.75} = 0.7127$$

(c) Let C be the number of earthquakes in the first 9 months of 2015. Then C follows the Poisson distribution with $\lambda = 15/4 = 3.75$ and $C - A$ follow the Poisson distribution with $\lambda = 5/4 = 1.25$. Therefore, since A and $C - A$ are independent,

$$\begin{aligned}
&P(C \geq 4 \mid A \geq 2) \\
=&P\left(\frac{C \geq 4, A \geq 2}{A \geq 2}\right) \\
=&\frac{P(A = 2, C - A \geq 2) + P(A = 3, C - A \geq 1) + P(A \geq 4)}{P(A \geq 2)} \\
=&\frac{e^{-2.5}2.5^2/2 \cdot (1 - e^{-1.25} - 1.25e^{-1.25}) + e^{-2.5}2.5^3/6 \cdot (1 - e^{-1.25}) + 1 - e^{-2.5}(1 + 2.5 + 2.5^2/2 + 2.5^3/6)}{1 - e^{-2.5} - 2.5e^{-2.5}} \\
=&0.6821
\end{aligned}$$

8 (6.15) A club basketball team will play a 60-game season. Thirty-two of these games are against class A teams and 28 are against class B teams. The outcomes of all the games are independent. The team will win each game against a class A opponent with probability .5, and it will win each game against a class B opponent with probability .7. Let X denote its total number of victories in the season.

(a) Is X a binomial random variable?

(b) Let X_A and X_B denote, respectively, the number of victories against class A and class B teams. What are the distributions of X_A and X_B?

(c) What is the relationship between X_A, X_B, and X?

(d) Approximate the probability that the team wins 40 or more games.

Solution. (a) Since X can be increased by 1 with probability either 0.5 or 0.7, X is not a binomial random variable.

(b) X_A and X_B are both binomial random variable such that

$$X_A \sim B(32, 0.5),\ X_B \sim B(28, 0.7)$$

(c) $X = X_A + X_B$

(d) X_A and X_B can be approximated to the normal variables $N(16, 8)$ and $N(19.6, 5.88)$ respectively. Then since X_A and X_B are independent and

$$E[X_A + X_B] = 16 + 19.6 = 35.6,\ Var[X_A + X_B] = 8 + 5.88 = 13.88$$

, $X_A + X_B$ can be approximated to the normal variable $N(35.6, 13.88)$. Therefore,

$$\begin{aligned} P(X > 39.5) &= P\left(\frac{X - 35.6}{\sqrt{13.88}} > \frac{39.5 - 35.6}{\sqrt{13.88}}\right) \\ &= P(Z > 3.9/\sqrt{13.88}) = 0.148 \end{aligned}$$