MTH202: Solutions (Selected exercises)

April 21, 2019

1. Let X be a random variable such that $E[X] = 0$ and $P(-3 < X < 2) = 1/2$. Find a lower bound for $Var(X)$.

   Solution: By Chebyshev's inequality: $P(|X - E[X]| \geq 2) \leq \frac{Var(X)}{4}$. So,

$$\begin{aligned} Var(X) &\geq 4P(|X - E[X]| \geq 2) \\ &= 4P(|X| \geq 2) \;\; \text{since} E[X] = 0 \\ &= 4[1 - P(-2 < X < 2)] \\ &\geq 4[1 - P(-3 < X < 2)] \;\; \text{since} \{-3 < X > 2\} \subset \{-2 < X < 2\} \\ &= 4[1 - 1/2] \\ &= 2 \end{aligned}$$

   Thus, $Var(X) \geq 2$.

---

2. The mean or expected score of a test is 75, with a variance of 25. What is the probability that the test score is between 50 and 100?

   Solution: Let X denote the test score. We are given that $E[X] = 75, Var(X) = 25$. Then,

$$\begin{aligned} P(50 < X < 100) &= P(-25 < X - 57 < 25) \\ &= 1 - P(|X - 75| \geq 25) \\ &\geq 1 - \frac{Var(X)}{25^2} \;\; \text{by Chebyshev's inequality} \\ &= 1 - \frac{25}{25^2} \\ &= \frac{24}{25} \end{aligned}$$

---

3. A fair coin is tossed repeatedly. Suppose that HEADS appears for the first time after X tosses and TAILS appears first time after Y tosses. Find the joint probability mass function of X and Y. Compute the corresponding marginals.

Solution: Note that:

$$P_{XY}(1, n) = P(X = 1, Y = n) = 1/2^n$$

$$P_{XY}(n, 1)) = P(X = n, Y = 1) = 1/2^n$$

The marginals are given by:

$$P_X(n) = 1/2^n \quad \text{and} \quad P_Y(n) = 1/2^n$$

---

4. You have invited 64 guests to a party. You need to make sandwiches for the guests. You believe that a guest might need $0, 1$ or 2 sandwiches with probabilities $1/4, 1/2$ and $1/4$ respectively. You assume that the number of sandwiches each guest needs is independent from other guests. How many sandwiches should you make so that you are 95 percent sure that there is no shortage?
(**Hint:** Let Y be the total number of sandwiches needed. You need to determine y (or a lower bound on y) such that $P(Y \leq y) \geq 0.95$. For a good approximation of y, use CLT. Assume $\phi^{-1}(0.95) \approx 1.65$.)

Solution: Let X_i denote the number of sandwiches i^{th} guest requires. Then, the p.m.f. of X_i is given by:

$$X_i = \begin{cases} 0 & \text{with probability } 1/4 \\ 1 & \text{with probability } 1/2 \\ 2 & \text{with probability } 1/4 \end{cases}$$

Define $Y = \sum_{i=1}^{64} X_i$. Then, Y is the total number of sandwiches needed. We want to find y such that $P(Y \leq y) = 0.95$. Since X_i's are i.i.d., by central limit theorem, we have that $P\left(\frac{\frac{\sum_i^n X_i}{n} - E[X_1]}{Var(X_1)} \leq y\right) \to P(Z \leq y)$ as $n \to \infty$ (where $Z \sim \mathcal{N}(0, 1)$.

Assuming that $n = 64$ is large enough, we can approximate the distribution $\frac{\frac{\sum_i^n X_i}{n} - E[X_1]}{Var(X_1)} = \frac{Y - E[Y]}{\sqrt{Var(Y)}}$ by a standard normal distribution. We first compute $E[Y]$ and $Var(Y)$.

$$E[Y] = \sum_{i=1}^{6} 4E[X_i] = \sum_{i=1}^{6} 4(0 \times 1/4 + 1 \times 1/2 + 2 \times 1/4) = 64$$

Since X_i's are independent.

$$Var(Y) = \sum_{i=1}^{6} 4Var(X_i) = \sum_{i=1}^{6} 41/2 = 32$$

Thus,

$$\begin{aligned} P(Y \leq y) &= P\left(\frac{Y - E[Y]}{\sqrt{Var(Y)}} \leq y\right) \\ &= P\left(\frac{Y - 64}{\sqrt{32}} \leq \frac{y - 64}{4\sqrt{2}} \leq y\right) \\ &= \phi\left(\frac{y - 64}{4\sqrt{2}}\right) \end{aligned}$$

Now, recall that we want to compute y such tat $P(Y \leq y) = 0.95$. So,

$$\begin{aligned} \phi\left(\frac{y - 64}{4\sqrt{2}}\right) &= 0.95 \\ \text{Or, } \frac{y - 64}{4\sqrt{2}} &= \phi^{-1}(0.95) \\ \text{i.e., } \frac{y - 64}{4\sqrt{2}} &\approx 1.65 \end{aligned}$$

This implies $y \approx 73.3$. This implies 74 sandwiches is a good estimate.

---

5. Let $X_1, X_2, \ldots, X_n$ be i.i.d. $Exp(1)$ random variables. Let

$$Y = \frac{X_1 + X_2 + \ldots + X_n}{n}$$

How large n should be such that $P(0.9 \leq Y \leq 1.1) \geq 0.9$?

Solution: Same as above. In this case, you are given $Y = \sum_{i=1}^{n} X_i$ such that $E[Y] = Var(Y) = n$ (Check this). You need to find n such that

$P\left(0.9 \leq \frac{Y}{n} \leq 1.1\right)$. Again by CLT,

$$\begin{aligned} P\left(0.9 \leq \frac{Y}{n} \leq 1.1\right) &= P\left(0.9n \leq Y \leq 1.1n\right) \\ &= P\left(\frac{0.9n-n}{\sqrt{n}} \leq \frac{Y-n}{\sqrt{n}} \leq \frac{1.1n-n}{\sqrt{n}}\right) \\ &= P\left(-0.1\sqrt{n} \leq \frac{Y-n}{\sqrt{n}} \leq 0.1\sqrt{n}\right) \\ &\approx \phi(0.1\sqrt{n}) - \phi(-0.1\sqrt{n}) \\ &= 2\phi(0.1\sqrt{n}) - 1 \end{aligned}$$

Thus, to find n such that

$$P\left(0.9 \leq \frac{Y}{n} \leq 1.1\right) \geq 0.9$$

we need to solve

$$2\phi(0.1\sqrt{n}) - 1 \geq 0.9$$

This gives, $n \geq 272.25$ (using $\phi^{-1}(0.95) \approx 1.65$).

---

6. Let X_i be i.i.d. $Unif(0,1)$. We define the sample mean as

$$M_n = \frac{X_1 + X_2 + \ldots + X_n}{n}$$

Then:

- Find $E[M_n]$ and $Var(M_n)$ as a function of n.
- Using Chebyshev's inequality, find an upper bound on $P\left(|M_n - 1/2| \geq 1/100\right)$.

Solution: $E[M_n] = \frac{1}{n}\sum_{i=1}^{n} E[X_i] = 1/2$ and,

$$Var(M_n) = \frac{1}{n^2}\sum_{i=1}^{n} = \frac{1}{12n}$$

Chebyshev's inequality implies:

$$P\left(\left|M_n - 1/2\right| \geq \frac{1}{100}\right) \leq \frac{10^4}{12n}$$

---

7. Let X and Y be independent random variables such that $Var(X) = 2$ and $Y \sim Poi(1)$. Given $E[X^2Y + Y^2X] = 1$, compute $E[X]$.

   Solution: Recall that if X, Y are independent, $E[h(X)g(Y)] = E[h(X)]E[g(Y)]$ for any real-valued functions h, g.

   Let $\mu = E[X]$, we get:

$$\begin{aligned} E[X^2Y + Y^2X] &= E[X^2]E[Y] + E[Y^2]E[X] \\ &= (Var(X) + (E[X])^2)E[Y] + (Var(Y) + (E[Y])^2)E[X] \\ &= (2 + \mu^2) \times 1 + (1 + 1^2) \times \mu \\ &= \mu^2 + 2\mu + 2 \end{aligned}$$

   Solve $\mu^2 + 2\mu + 2 = 1$ to get $\mu = -1$.