Question 1

Let $X \sim U(1,5)$ be a discrete uniform random variable, and define $Y = 3^X$.

1. What is the expected value of X?
2. What is the expected value of Y?
3. What is the probability mass function of Y?

Solution

1. The mean of the discrete uniform distribution $U(a,b)$ is $\frac{a+b}{2}$, therefore $E[X] = \frac{1+5}{2} = 3$.
2. The expected value of Y is $E[3^X] = \sum_{x=1}^{5} 3^x \cdot \frac{1}{5} = \frac{363}{5} = 72.6$.
3. The probability mass function of Y is $P(Y = y) = \frac{1}{5}$ for $y \in \{3, 9, 27, 81, 243\}$ and $P(Y = y) = 0$ otherwise.

Question 2

Consider a business that conducts a certain number of transactions per weekday. The number of transactions per weekday is a Poisson random variable, with a mean of 2.

1. What is the probability that on a given day there will be no transactions?
2. What is the probability that at least four transactions will be performed on a given day?

Solution

1. Denote by X the number of transactions per weekday. Since the mean of a $Pois(\lambda)$ random variable is λ, we have that $X \sim Pois(2)$. Using the PMF of the Poisson distribution, it follows that $\Pr[X = 0] = \frac{2^0 e^{-2}}{0!} = e^{-2} \approx 0.135$. Therefore, there is approximately a 13.5% chance that there will be no transactions on a given day.
2. We can use the PMF of the Poisson distribution to sum the probabilities of having up to three transactions, and take the complement: $\Pr[X \geq 4] = 1 - \Pr[X < 4] = 1 - \sum_{k=0}^{3} \frac{2^k e^{-2}}{k!} \approx 0.143$. Therefore, there is approximately a 14.3% chance that at least four transactions will be performed on a given day.

Question 3

Give a simple example of a random variable X and function g to show that in general, $E[g(X)] \neq g(E[X])$.

Solution

Let X be a discrete random variable with $P(X = 1) = P(X = 2) = \frac{1}{2}$. Let $g(X) = X^2$. Then $E[g(X)] = E[X^2] = \frac{1}{2} \cdot 1^2 + \frac{1}{2} \cdot 2^2 = \frac{5}{2}$, while $g(E[X]) = g(1.5) = 1.5^2 = \frac{9}{4}$.

Question 4

In a multiple-choice exam, there are 10 questions. Each has four possible answers (with only one correct). Alice didn't prepare for the exam, so she guessed all her answers. Let X denote the number of her correct answers.

1. What is the distribution of X? Write its PMF.
2. To pass the test, a student should get 55%. What is the probability that Alice passed the test?

Solution

1. The distribution of X is binomial with $n = 10$ and $p = \frac{1}{4}$. The PMF of X is given by $P(X = k) = \binom{10}{k}\left(\frac{1}{4}\right)^k\left(\frac{3}{4}\right)^{10-k}$ for $k = 0, 1, \ldots, 10$.
2. The probability that Alice passed the test is given by $P(X \geq 6) = 1 - P(X \leq 5) = 1 - \sum_{k=0}^{5}\binom{10}{k}\left(\frac{1}{4}\right)^k\left(\frac{3}{4}\right)^{10-k} \approx 0.0197$. In other words, there is approximately a 1.97% chance that Alice passed the test.

Question 5

You are advising an institution planning to offer a \$10,000 loan. There are two applicants for the loan: a low-risk applicant and a high-risk applicant. The probability of a high-risk applicant defaulting on the loan is 0.1, while the probability of a low-risk applicant defaulting is 0.025. However, the company can offer a high-risk applicant an interest rate of 15%, while it can only offer a low-risk applicant an interest rate of 5%.

1. For each applicant, calculate the expectation and variance of the revenue. Is it necessarily better to choose the option with the higher expected revenue?
2. Please refer to this YouTube video to learn about the geometric distribution. You can also refer to the lecture notes and Wikipedia.

   Following the success of the first loan, the institution started offering further loans, one after the other. Each loan is offered to a low-risk applicant with probability 0.8 and to a high-risk applicant with probability 0.2, independently of the previous loans.

   Denote by D the number of loans until the first default. For example, $D = 10$ means nine loans did not default and the tenth one did. Calculate the expectation and variance of D. What is the probability that $D \geq 4$?

Solution

1. For a low-risk applicant, the institution loses \$10,000 with probability 0.025 and gains \$500 with probability 0.975. Therefore, the expectation of the revenue is given by $0.975 \cdot 500 + 0.025 \cdot (-10000)$, which equals \$237.5. The variance is given by $0.975 \cdot (500 - 237.5)^2 - 0.025 \cdot (-10000 - 237.5)^2 = 2687343.75$, and so we have a standard deviation of approximately \$1639.

   For a high-risk applicant, the institution loses \$10,000 with probability 0.1 and gains \$1500 with probability 0.9. Therefore, the expectation of the revenue is given by $0.9 \cdot 1500 + 0.1 \cdot (-10000)$, which equals \$350. The variance is given by $0.9 \cdot (1500 - 350)^2 - 0.1 \cdot (-10000 - 350)^2 = 11902500$, and so we have a standard deviation of \$3450.

   In this case the applicant with higher expected revenue also has higher risk of default, and so it is not necessarily better (nor necessarily worse) to choose the option with the higher expected revenue.
2. We are given that $D \sim \text{Geo}(p)$, where p is the overall probability of default in a given loan. To compute p, we can use the law of total probability: $p = 0.8 \cdot 0.025 + 0.2 \cdot 0.1 = 0.04$. Therefore, $D \sim \text{Geo}(0.04)$. It follows that $E[D] = \frac{1}{0.04} = 25$ and $\text{Var}(D) = \frac{1-0.04}{0.04^2} = 600$. Finally, $\Pr[D \geq 4] = (1-0.04)^3 = 0.884736$, and so the probability of at least three no-default loans until the first default is approximately 88.5%.