MA 2103 : Mathematical Methods II

Satvik Saha, 19MS154 November 21, 2020

Problem 1. Grades of a class are based entirely on midterm and final exams. In all, the exams have 100 separate parts, each worth 5 points. A student scoring at least 85%, is guaranteed an A+ score. Throughout this problem, consider a particular student who, based on performance in other courses and amount of effort put into the class, is estimated to have a 90% chance to complete any particular part of exam correctly. Parts not completed correctly receive zero credit.

(a) Assume that the scores on different parts are independent. Based on the Law of Large Numbers, about what total score for the semester are we likely to see?

(b) Using the Central Limit Theorem, calculate the approximate probability the student scores enough points for a guaranteed A+ score.

Solution.

(a) Let X be the random variable denoting the student's score in a particular part. There are only two possible values: 0 (incorrect) and 5 (correct). For the student in question, his expected score is

$$\mu = \mathrm{E}[X] = 0 \cdot P(X = 0) + 5 \cdot P(X = 5) = 4.5.$$

Since each part is independent of the other, the expected total score is simply $100 \cdot 4.5 = 450$. Now, the Law of Large numbers tells us that the actual score should approach the expected score, given enough trials. Thus, given 100 trials (parts), we should expect a score of 450. The required score for an A+ is $0.85 \cdot 500 = 425$. Thus, we expect the student to score an A+.

(b) We use the Central Limit theorem to conclude that the distribution of the sum of scores for the parts, i.e. the total score is the normal distribution $N(n\mu, \sqrt{n}\sigma) = N(100\mu, 10\sigma)$. We calculate the standard deviation for a single part as

$$\sigma^2 = \mathrm{Var}(X) = \mathrm{E}[X^2] - \mathrm{E}[X]^2 = 0 \cdot P(X^2 = 0) + 5^2 \cdot P(X^2 = 5^2) - 4.5^2 = 22.5 - 20.25 = 2.25.$$

Thus, $\sigma = 1.5$, so the total score has the distribution $N(450, 15)$. The probability that the student scores enough to guarantee an A+ can be computed numerically.

$$P\left(\sum X \geq 425\right) \approx 0.9522 \approx 95.22\%.$$

Problem 2. Following are the announced awards for a dice game.

(a) Roll an odd number : Rs. 0.

(b) Roll a 2 or a 4 : Rs. 2.

(c) Roll a 6 : Rs. 26.

If you play the dice game 30 times, what is the expected value and standard deviation of your cumulative winnings? What is the probability you win at least Rs. 200?

Solution. Let X be the random variable denoting the money won in a round. Assuming a fair dice, each number has a probability 1/6 to turn up. Only 2, 4 and 6 win anything, thus

$$\mu = \mathrm{E}[X] = \frac{1}{6}(2 + 2 + 26) = \frac{30}{6} = 5.$$

The standard deviation σ is calculated from

$$\sigma^2 = \mathrm{Var}(X) = \mathrm{E}[X^2] - \mathrm{E}[X]^2 = \frac{1}{6}(2^2 + 2^2 + 26^2) - 5^2 = 114 - 25 = 89.$$

Thus, $\sigma = 9.43$

The Central Limit theorem gives the distribution of the cumulative winnings over $n = 30$ rounds as $N(n\mu, \sqrt{n}\sigma) = N(150, 51.67)$. Thus, the expected winnings is Rs. 150, with a standard deviation of Rs. 51.67.

The probability of winning at least Rs. 200 is calculated numerically as

$$P\left(\sum X \geq 200\right) \approx 0.1666 \approx 16.66\%.$$

Note that this is approximately the probability that the score is above one standard deviation form the mean.

Problem 3. Chebyshev's inequality states that the probability of a deviation of a discrete random variable X with expected value μ and variance $\mathrm{Var}(X)$ is given by

$$P(|X - \mu| \geq \epsilon) \leq \frac{\mathrm{Var}(X)}{\epsilon^2},$$

where ϵ is any positive real number. Show that the probability of a deviation from the mean of more than k standard deviations is less than or equal to $1/k^2$.

Solution. Simply substitute $\epsilon = k\sigma$, and note that $\sigma^2 = \mathrm{Var}(X)$ to write

$$P(|X - \mu| \geq k\sigma) \leq \frac{\mathrm{Var}(X)}{k^2\sigma^2} = \frac{1}{k^2}.$$

Problem 4. Let $\{X_i\}$ be a trials process with probability 0.3 of success and 0.7 of failure. Let $X_j = 1$ if the j^{th} outcome is a success and 0 otherwise. Find $P(0.2 \leq A_{100} \leq 0.4)$ and $P(0.2 \leq A_{1000} \leq 0.4)$ using Chebyshev's inequality.

Solution. We calculate the mean μ and standard deviation of each trial X as follows.

$$\mu = \mathrm{E}[X] = 0 \cdot 0.7 + 1 \cdot 0.3 = 0.3.$$

$$\sigma^2 = \mathrm{E}[X^2] - \mathrm{E}[X]^2 = 0^2 \cdot 0.7 + 1^2 \cdot 0.3 - 0.3^2 = 0.21.$$

Thus, $\sigma = 0.458$. We denote the average of n trials as

$$A_n = \frac{1}{n}\sum_{i=1}^{n} X_i.$$

By linearity of expectation, we have $\mathrm{E}[A_n] = \mu$. The Central Limit theorem gives the variance of the mean as $\mathrm{Var}(A_n) = \sigma^2/n$. Thus, assuming a normal distribution in the limiting case, we numerically calculate the approximate probabilities,

$$P(0.2 \leq A_{100} \leq 0.4) \approx 0.971, \qquad P(0.2 \leq A_{1000} \leq 0.4) \approx 1.$$

Chebyshev's inequality gives bounds on these probabilities. Setting $\epsilon = 0.1$, we see that

$$P(|A_{100} - 0.3| \geq 0.1) \leq \frac{0.21}{0.1^2 \cdot 100} = 0.21, \qquad P(|A_{1000} - 0.3| \geq 0.1) \leq \frac{0.21}{0.1^2 \cdot 1000} = 0.021.$$

The complements give

$$P(|A_{100} - 0.3| \leq 0.1) \geq 0.79, \qquad P(|A_{1000} - 0.3| \leq 0.1) \geq 0.979.$$

Problem 5. A researcher wishes to estimate the mean of a population using a sample large enough that the probability will be 0.95 that the sample mean will not differ from the population mean by more than 25% of the standard deviation. How large a sample should he take?

Solution. Suppose the sample has size n. The mean $\overline{X}$ is normally distributed by the Central Limit theorem, centred at μ with standard deviation $\sigma' = \sigma/\sqrt{n}$. Thus, we want

$$P(|\overline{X} - \mu| \leq 0.25\sigma) = 0.95.$$

On the other hand, we know that 95% of a normal distribution is contained within 1.96 standard deviations of the mean. Thus, we set

$$0.25\sigma = 1.96\sigma' = \frac{1.96\sigma}{\sqrt{n}}, \qquad n = (4 \cdot 1.96)^2 \approx 61.5.$$

Thus, our researcher should choose a sample size greater than or equal to 62.

Problem 6. Two random samples of size 100 are drawn from two populations P_1 and P_2 and their means X_1 and X_2. If $\mu_1 = 10$, $\sigma_1 = 2$ and $\mu_2 = 8$, $\sigma_2 = 1$, find

(a) $\mathrm{E}[X_1 - X_2]$.

(b) $\sigma(X_1 - X_2)$.

(c) The probability that the difference between a given pair of sample means is less than 1.5.

(d) The probability that the difference between a given pair of sample means is greater than 1.75 but less than 2.5.

Solution. Note that the means must have a standard deviation of $\sigma/\sqrt{100} = \sigma/10$.

(a) From linearity of expectation,

$$\mathrm{E}[X_1 - X_2] = \mathrm{E}[X_1] - \mathrm{E}[X_2] = \mu_1 - \mu_2 = 2.$$

(b) We know that

$$\mathrm{Var}(X_1 - X_2) = \mathrm{Var}(X_1) + \mathrm{Var}(X_2) = \frac{2^2}{100} + \frac{1^2}{100} = \frac{5}{100}.$$

Thus, $\sigma(X_1 - X_2) = 0.224$.

(c) We have already parametrized the normal distribution of $X_1 - X_2$ as $N(2, 0.224)$. Thus, we want

$$P(X_1 - X_2 < 1.5) = 0.0127.$$

(d) We want

$$P(1.75 < X_1 - X_2 < 2.5) = 0.856.$$