MA 2103 Assignment 08

Abhay Kshirsagar 19MS172 ()

November 23, 2020

Question 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 LoLN, about what total score for the semester are we likely to see?

(b) Using the CLT calculate the approximate probability the student scores enough points for a guaranteed A+ score.

Solution :

Let, $\{X_n\}_n$ a sequence of random variables, expressing the marks the student might get in the nth part of the exam.
So,

$$X_n = \begin{cases} 5, & \text{if the student completes the part} \\ 0, & \text{if the student doesn't complete the part} \end{cases}$$

then, $\mu_X = E(X_n) = \frac{9}{10} \times 5 + \frac{1}{10} \times 0 = \frac{9}{2}$
Now the Variance is given as σ standard deviation

$$\sigma_X^2 = Var(X_n) = \frac{9}{10}(5 - \frac{9}{2})^2 + \frac{1}{10}(0 - \frac{9}{2})^2 = \frac{1}{10}(\frac{9}{4} + \frac{81}{4}) = \frac{9}{4} \tag{1}$$

$$\sigma_X = \frac{3}{2}$$

Let,

$$S = \sum_{n=1}^{100} X_n$$

Now

$$\mu_S = 100 \times \mu_X = 100 \times \frac{9}{2} = 450$$

Now,

$$\sigma_S = \sqrt{100} \times \sigma_X = 10 \times \frac{3}{2}$$

$$\sigma_S = 15$$

(a) :
Let,

$$\bar{X} = \frac{S}{100} = \frac{1}{100}\sum_{i=1}^{100} X_i$$

Now, as n = 100 is pretty large using LoLN, we find that the average marks the student might get in any part is

$$\mu_X = \frac{9}{2} = 4.5$$

As, $\lim_{i\to 0} P(|\bar{X}_n - \mu_X| < \epsilon) = 1$ for large n
So, the expected total marks is , $100 \times 4.5 = 450$

(b) :
Using CLT we want to find,

$$P(425 \leq S \leq 500) = P(\frac{425-450}{15} \leq \frac{S-\mu_S}{\sigma_S} \leq \frac{500-450}{15})$$
$$= P(-\frac{5}{3} \leq \frac{S-\mu_S}{\sigma_S} \leq \frac{10}{3})$$

as the cutoff for passing is,

$$\frac{85}{100} \times 500 = 450$$

if $S \geq 450$ the student-wise get A+, Let Z be the standardized variable

$$P(-\frac{5}{3} \leq \frac{S-\mu_S}{\sigma_S} \leq \frac{10}{3}) = 0.9518$$

the required probability is 0.9518

Question 2.

Following are the announced awards for a dice game:

(a) Roll and 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, $\{X_n\}_n$ a finite sequence of random variables, representing the prize one gets for the result of the casting die at a time
So,

$$X_n = \begin{cases} 0, & \text{Die shows a odd number} \\ 2, & \text{Die shows 2 or 4} \\ 26, & \text{Die shows a 6} \end{cases}$$

$$E_1 = \text{Casting an odd number}$$
$$E_2 = \text{Casting number 2 or 4}$$
$$E_3 = \text{Casting number 6}$$

Then,

$$P(E_1) = \frac{3}{6} = \frac{1}{2}$$
$$P(E_2) = \frac{2}{6} = \frac{1}{3}$$
$$P(E_3) = \frac{1}{6}$$

Now mean is

$$\mu_X = E(X_n) = \frac{1}{2} \times 0 + \frac{1}{3} \times 2 + \frac{1}{6} \times 26 = 5$$
$$\sigma_X^2 = Var(X_n) = \frac{1}{2}(0-5)^2 + \frac{1}{3}(2-5)^2 + \frac{1}{6}(26-5)^2 = \frac{1}{6}(75 + 18 + 441) = 89 \tag{2}$$
$$\sigma_X = 9.933$$

Let,

$$S_30 = \sum_{i=1}^{3} 0X_i$$

S_30 is the possible sum of the minimas.

$$\sigma_S = \sqrt{30} \times \sigma_X = Rs.51.66$$

$$\mu_S = 30 \times \mu_X = Rs.150$$

expected value of cummulative minimas $= \mu_S = Rs150$
Standard deviation of cummulative winnings $\sigma_S = Rs.51.66$
We want the probability of winning atleast 200 in total. The maximum money one can win is $30 \times 26 = Rs.780$
So, our required probability is

$$P(200 \leq S_30 \leq 780) = 0.1666$$

The required probability is 0.1666

Question 3.

Chebyshev inequality states that the probability of a deviation of a discrete random variable X with expected value μ, and variance Var(X) is given by the following, for any positive real number ϵ, $P(|X - \mu| \geq \epsilon) \leq Var(X)/\epsilon^2$
Show that the probability of a deviation from the mean of more than k standard deviations is $\leq 1/k^2$. This can be thought as a corollary of CLT.

Solution :
Now it is given that,

$$P(|X - \mu| \geq \epsilon) \leq Var(X)/\epsilon^2 = \frac{\sigma_X^2}{\epsilon^2}$$

if the deviation of X wrt μ is more than kVar(X), then

$$\epsilon = k \times Var(X) = k \cdot \sigma_X$$

So,

$$P(|X-\mu| \geq \frac{1}{k^2}) =\leq \frac{\sigma_X^2}{k^2\sigma_X^2} = \frac{1}{k^2}$$

So,

$$P(|X-\mu| \geq \frac{1}{k^2}) = \frac{1}{k^2}$$

Hence Proved

Question 4.

Let $\{X_i\}$ be a trials process with probability 0.3 for success and 0.7 for 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 $P(0.2 \leq A_{1000} \leq 0.4)$. Hint: Use Chebyshev inequality.

Solution :
Let p = Probability of success and q =Probability of failure

$$X_j = \begin{cases} 1, \text{if j th outcome is success} \\ 0, \text{if j th outcome is failure} \end{cases}$$

$$p = 0.3; q = 0.7$$

here,

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

Now,

$$\begin{aligned} \mu_X &= E(X_n) = p \times 1 + q \times 0 = 0.3 \\ \sigma_X^2 &= Var(X_n) = p \cdot (1-\mu_X)^2 = p(1-p)^2 + (1-p)p^2 = p(1-p) = p \cdot q \\ \sigma_X &= \sqrt{p(1-p)} = \sqrt{pq} = \frac{21}{10} \end{aligned} \tag{3}$$

So,

$$\mu_A = \mu_X = p$$

$$\sigma_A = \frac{\sigma_X}{\sqrt{n}} = \sqrt{\frac{pq}{n}}$$

Now,

$$\begin{aligned} P(0.2 \leq A_100 \leq 0.4) &= P(0.2 - 0.3 \leq A_100 - 0.3 \leq 0.4 - 0.3) \\ &= P(|A_100 - \mu_A| \leq 0.1) \\ &= 1 - P(|A_100 - \mu_A| > 0.1) \end{aligned}$$

We find that

$$P(0.2 \leq A_100 \leq 0.4) > 1 - \frac{Var(A_100)}{(0.1)^2} = 1 - \frac{21}{100 \times 100 \times 0.1^2}$$

$$P(0.2 \leq A_100 \leq 0.4) > 1 - 0.21 = 0.79$$

Also by calulating we find that

$$P(0.2 \leq A_100 \leq 0.4) = 0.971$$

For A_1000

$$\mu_A = 0.3$$

$$\sigma^2_{A_{1000}} = \frac{pq}{1000} = \frac{0.21}{1000}$$

So,

$$\sigma_{A_{1000}} = 0.0144$$

Now,

$$P(0.2 \leq A_{1000} \leq 0.4) = P(|A_{1000} - 0.3| \leq 0.1) > 1 - P(|A_{1000} - 0.3| \geq 0.1)$$

$$P(0.2 \leq A_{1000} \leq 0.4) = 1 - \frac{\sigma^2_{A_{1000}}}{0.1^2} = 1 - \frac{0.21}{100 \times 100 \times 0.1^2} = 1 - 0.021 = 0.979$$

SO,

$$0.979 < P(|A_{1000} - 0.3| \leq 0.1) < 1$$

Also by calculator we can see that approximately

$$P(0.2 \leq A_{1000} \leq 0.4) = 1$$

Question 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 :
Let the size of the sample is expressed by n. For the population distribution we find, $\bar{X_n}$ =average distribution over n samples (sample mean) ,and μ = population mean
Let, σ_{X_n} be the standard deviation of the sample we want,

$$|\bar{X_n} - \mu| \leq \frac{1}{4} \cdot \sigma_{X_n}$$

The mentioned probability is

$$P(|\bar{X_n} - \mu| < \frac{\sigma_{X_n}}{4}) = 0.95$$

$$P(\frac{|\bar{X_n} - \mu|}{\frac{\sigma_{X_n}}{\sqrt{n}}} < \frac{\sqrt{n}}{4}) = 0.95$$

Let,

$$Z = \frac{|\bar{X_n} - \mu|}{\frac{\sigma_{X_n}}{\sqrt{n}}}$$

WE know that,

$$P(|Z| < 1.96) = 0.95$$

So,

$$\frac{\sqrt{n}}{4} = 1.96$$

$$n = 1.96^2 \times 4^2 = 61.4656$$

$$n > 60$$

OR

$$n \geq 61$$

So the sample size should be greater than equal to 61 Also by chebyshev's inequality , we see ,

$$P(|\bar{X_n} - \mu| < \frac{\sigma_{X_n}}{4}) = 1 - P(|\bar{X_n} - \mu| \geq \frac{\sigma_{X_n}}{4})$$

$$\geq 1 - \frac{Var(\bar{X_n})}{\sigma^2_{X_n}} \cdot 16$$

$$= 1 - \frac{\sigma^2_{X_n}}{n\sigma^2_{X_n}} \cdot 16$$

$$= 1 - \frac{16}{n}$$

$$1 - \frac{16}{n} \leq 0.95$$

$$n \leq \frac{16}{0.05} = 320$$

So$61 \leq n \leq 320$ The size should satisfy this inequality

Question 6.

Two random samples of size 100 are drawn from two populations P1 and P2 and their means X1 and X2. If $\mu_1 = 10, \sigma_1 = 2, \mu_2 = 8, \sigma_2 = 1$ find:

(a) $E(X_1 - X_2)$;

(b) $(X_1 - X_2)$;

(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 :
Size of the samples , $n_1 = n_2 = 100$, $\mu_1 = 10, \sigma_1 = 2, \mu_2 = 8, \sigma_2 = 1$

$$X_1 = \frac{1}{100}\sum_{i=1}^{1} 00X_i^1$$

$$X_2 = \frac{1}{100}\sum_{i=1}^{1} 00X_i^2$$

$\{X_i^1\}$ and $\{X_i^2\}$ forms the sample spaces for p_1 and p_2 So,

$$E($$

$$X_1 - X_2)$$

$$= E(X_1) - E(X_2) = \frac{1}{100}\sum E(X_i^1) - \frac{1}{100}\sum E(X_i^2)$$

$$= \frac{1}{100}\sum \mu_1 - \frac{1}{100}\sum \mu_2$$

$$= \frac{100}{100}(\mu_1 - \mu_2)$$

$$= \mu_1 - \mu_2$$

$$= 10 - 8$$

$$E(X_1 - X_2) = 2$$

(b) :

$$\begin{aligned}\sigma^2(X_1 - X_2) &= Var(X_1 - X_2)\\ &= E(X_1 - X_2 - \bar{X_1} + \bar{X_2})^2\\ &= E((X_1 - \bar{X_1}) - (X_2 - \bar{X_2}))^2\\ &= E((X_1 - \bar{X_1})^2 + (X_2 - \bar{X_2})^2\\ &\quad -2(X_1 - \bar{X_1})(X_2 - \bar{X_2}))\end{aligned}$$

Assuming X_1 and X_2 are not correlated, we see that

$$\begin{aligned}E[(X_1 - \bar{X_1})(X_2 - \bar{X_2}))] &= E(X_1 - \bar{X_1})E(X_2 - \bar{X_2}))\\ &= (E(X_1) - E(\bar{X_1}))(E(X_2) - E(\bar{X_2})))\\ &= [\bar{X_1} - \bar{X_1}][\bar{X_2} - \bar{X_2}]\\ &= 0\end{aligned}$$

So,

$$\begin{aligned}\sigma^2(X_1 - X_2) &= E((X_1 - \bar{X_1})^2 + (X_2 - \bar{X_2})^2 - 0\\ &= Var(X_1) + Var(X_2)\\ &= \sigma^2_{X_1} + \sigma^2_{X_2}\\ \sigma(X_1 - X_2) &= \sqrt{\frac{\sigma_1^2}{100} + \frac{\sigma_2^2}{100}}\\ &= \frac{\sqrt{4+1}}{10} = \frac{\sqrt{5}}{10} = \frac{\sigma'}{10}\end{aligned}$$

(c) :

$$P(|X_1 - X_2| < 1.5) = P(-1.5 < X_1 - X_2 < 1.5) \approx 0.0127$$

The probability is 0.0127.

(d): We want

$$\begin{aligned}P(1.75 < |X_1 - X_2| < 2.5) &= P(|X_1 - X_2) < 2.5) - P(|X_1 - X_2) < 1.75)\\ P(1.75 < |X_1 - X_2| < 2.5) &= 0.9873 - 0.1318\\ P(1.75 < |X_1 - X_2| < 2.5) &= 0.8553\end{aligned}$$

The required probability is 0.8553