Information and Communication Theory (UEC-310)

Tutorial-2

Dr. Amit Mishra

Question-1

I. Suppose someone gives you a coin and claims that this coin is biased; that it lands on heads only 48% of the time. You decide to test the coin for yourself. If you want to be 95% confident that this coin is indeed biased, how many times must you flip the coin?

Solution-1

Law of Large Numbers

The **law of large numbers** has a very central role in probability and statistics. It states that if you repeat an experiment independently a large number of times and average the result, what you obtain should be close to the expected value. There are two main versions of the law of large numbers. They are called the **weak** and **strong** laws of the large numbers.

let us define the *sample mean*.

Definition 7.1. For i.i.d. random variables $X_1, X_2, \ldots, X_n$, the **sample mean**, denoted by $\overline{X}$, is defined as

$$\overline{X} = \frac{X_1 + X_2 + \ldots + X_n}{n}.$$

Another common notation for the sample mean is M_n. If the X_i's have CDF $F_X(x)$, we might show the sample mean by $M_n(X)$ to indicate the distribution of the X_i's.

Note that since the X_i's are random variables, the sample mean, $\overline{X} = M_n(X)$, is also a random variable. In particular, we have

$$\begin{aligned} E[\overline{X}] &= \frac{EX_1 + EX_2 + \ldots + EX_n}{n} && \text{(by linearity of expectation)} \\ &= \frac{nEX}{n} && (\text{since } EX_i = EX) \\ &= EX. \end{aligned}$$

Also, the variance of $\overline{X}$ is given by

$$\begin{aligned} \mathrm{Var}(\overline{X}) &= \frac{\mathrm{Var}(X_1 + X_2 + \ldots + X_n)}{n^2} && (\text{since } \mathrm{Var}(aX) = a^2\mathrm{Var}(X)) \\ &= \frac{\mathrm{Var}(X_1) + \mathrm{Var}(X_2) + \ldots + \mathrm{Var}(X_n)}{n^2} && (\text{since the } X_i\text{'s are independent}) \\ &= \frac{n\mathrm{Var}(X)}{n^2} && (\text{since } \mathrm{Var}(X_i) = \mathrm{Var}(X)) \\ &= \frac{\mathrm{Var}(X)}{n}. \end{aligned}$$

Now let us state and prove the **weak law of large numbers (WLLN).**

The weak law of large numbers (WLLN)

Let X_1, X_2 , ... , X_n be i.i.d. random variables with a finite expected value $EX_i = \mu < \infty$. Then, for any $\epsilon > 0$,

$$\lim_{n\to\infty} P(|\overline{X} - \mu| \geq \epsilon) = 0.$$

Proof

The proof of the weak law of large number is easier if we assume $\mathrm{Var}(X) = \sigma^2$ is finite. In this case we can use Chebyshev's inequality to write

$$\begin{aligned} P(|\overline{X} - \mu| \geq \epsilon) &\leq \frac{\mathrm{Var}(\overline{X})}{\epsilon^2} \\ &= \frac{\mathrm{Var}(X)}{n\epsilon^2}, \end{aligned}$$

which goes to zero as $n \to \infty$.

Continued…

Using WLLN :

Let X be the random variable such that $X = 1$ if the coin lands on heads and $X = 0$ for tails.

Thus $\mu = 0.48 = p$ and $\sigma^2 = p(1-p) = 0.48 \times 0.52 = 0.2496$.

To test the coin we flip it n times and allow for a 2% error of precision, i.e. $\epsilon = 0.02$.

This means we are testing the probability of the coin landing on heads being between $(0.46, 0.50)$.

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

$$P\left(\left\{\left|\frac{k}{n} - p\right| > \varepsilon\right\}\right) < \frac{pq}{n\varepsilon^2}.$$

$$\boxed{P(|\overline{X} - \mu| \geq \epsilon) \leq \frac{pq}{n\,\varepsilon^2}.}$$

Continued…

By the Law of Large Numbers, we want n such that

$$P[|\bar{X} - 0.48| > 0.02] \leq \frac{0.2496}{n(0.02)^2}$$

So for a 95% confidence interval we need

$$\frac{0.2496}{n(0.02)^2} = 0.05$$

Thus n should be

$$0.2496 \times 2500 \times 20 = 12,480.$$

That is quite a lot of coin flips!

Question-1

II. A survey of **1500** people is conducted to determine whether they prefer Pepsi or Coke. The results show that **27%** of people prefer Coke while the remaining **73%** favor Pepsi. Estimate the Margin of error in the poll with a confidence of **90%**.

Solution-1-II

Let

$$X_n = \begin{cases} 1 & \text{if } n^{th} \text{ person is in favour of Coke} \\ 0 & \text{otherwise} \end{cases}$$

for $n = 1, 2, ..., 1500$. Then $X_1, X_2, ...X_{1500}$ are i.i.d. Bernoulli random variables and $\hat{p} = \frac{S_{1500}}{1500} = 0.27$.

So we let X_n be the Bernoulli random variable as described above with

$$P(X_1 = 1) = 0.27.$$

We know that:

$$\mu = p \text{ and } \sigma^2 = p(1-p)$$

Then $\mu = E(X_1) = 0.27$ and $\sigma^2 = 0.27 \times 0.73 = 0.1971.$

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

Then by the LLN with $n = 1500$ we have

$$P\left(\left|\frac{X_1 + ... + X_{1500}}{1500} - 0.27\right| \geq \epsilon\right) \leq \frac{0.1971}{1500\epsilon^2}, \ \epsilon > 0.$$

So if we set $\frac{1}{10} = \frac{0.1971}{1500\epsilon^2}$, we get

$$\epsilon = \sqrt{\frac{0.1971 \times 10}{1500}} = 0.036$$

Thus we have that the margin of error is less than 4% with 90% confidence.

The Law of Large Numbers is heavily used in fields of finance and insurance to assess risks as well as predict economic behaviour.

Question-1

III. Two new companies are being assessed. Company **A** has a total market value of **$60** Million and company **B** has a total market value of **$4** Million. Which company is more likely to increase its total market value by **50%** within the next few years?

Solution-1-III

Intuition based solution:

In order for Company A to increase its total market value by 50% it requires an increase of \$30 million whereas company B only requires a \$2 million increase.

The Law of Large Numbers implies that it is much more likely for company B to expand by 50% than company A.

This makes sense because if company A was just as likely to expand by 50% as company B, company A could quickly have a higher market value than the entire economy.

Similarly, insurance providers use the LLN in order to determine effective prices for their clients.

By analyzing and recording the number accidents among men aged 23-25, they can ascertain with a high degree of accuracy, the probability of X amount of males aged 23 that will be the cause of an accident in any given year.

This allows them to set an effective price for clients that fall into the range.

We may also consider an example of how to apply the LLN to gambling situations, where probabilities are extremely important

Question-1

IV. A 1-dollar bet on craps has an expected winning of **E(X) = -0.141**. What does the Law of Large Numbers say about your winnings if you make a large numbers of **1-dollar** bets at the craps table?
Does the Law of Large Numbers guarantee your loses will be negligible?
Does the Law of Large Numbers guarantee you will lose if **n** is large?

Craps: A gambling game played with two dice, chiefly in North America.
A throw of 7 or 11 is a winning throw, 2, 3, or 12 is a losing throw.

Solution-1-IV

By the LLN your losses will not be small, they will average to 0.141 per game which would imply that your losses will, on average, become large for a large number of games.

The Law of Large Numbers also does not guarantee you will lose for a large n. It says that for a large number n, it is very unlikely that you will lose. However for a fixed n finite, no assumption can be made whether you will win or lose.

Question-2

- The Random variable X has PDF

$$f(x) = \begin{cases} \frac{1}{4}x\,; & 1 \le x \le 3 \\ 0; \, otherwise \end{cases}$$

Find the Cumulative Distribution Function F(x).

- A random variable X has PDF given by

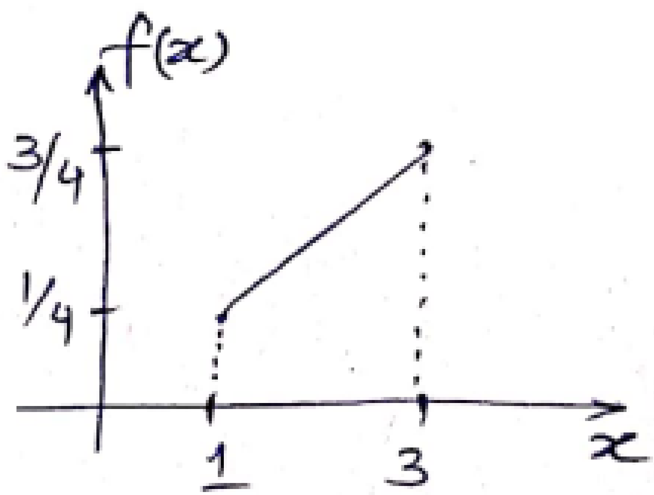
$$f(x) = \begin{cases} \frac{1}{3}; & 0 \le x < 1 \\ \frac{2}{7}x^2; & 1 \le x \le 2 \\ 0; & Otherwise \end{cases}$$

Find the cumulative distribution function F(x).

Solution-2-I

The random variable (rv) X has Pdf

$$\text{Pdf} = f(x) = \begin{cases} \frac{1}{4}x & ; 1 \le x \le 3 \\ 0 & ; \text{otherwise} \end{cases}$$



CDF =

$$f_X(x) = F(x) = \int_1^x f(x)\,dx$$

$$= \int_1^x \frac{1}{4}x\,dx = \frac{1}{4}\left[\frac{x^2}{2}\right]_1^x$$

$$= \frac{1}{4}\left[\frac{x^2}{2} - \frac{1}{2}\right] = \frac{x^2-1}{8}$$

Apply limits

at $x = 1$

$$F(x) = \frac{x^2-1}{8} = 0$$

at $x = 3$

$$f(x) = \frac{x^2-1}{8} = \frac{8}{8} = 1$$

at $1 < x < 3$

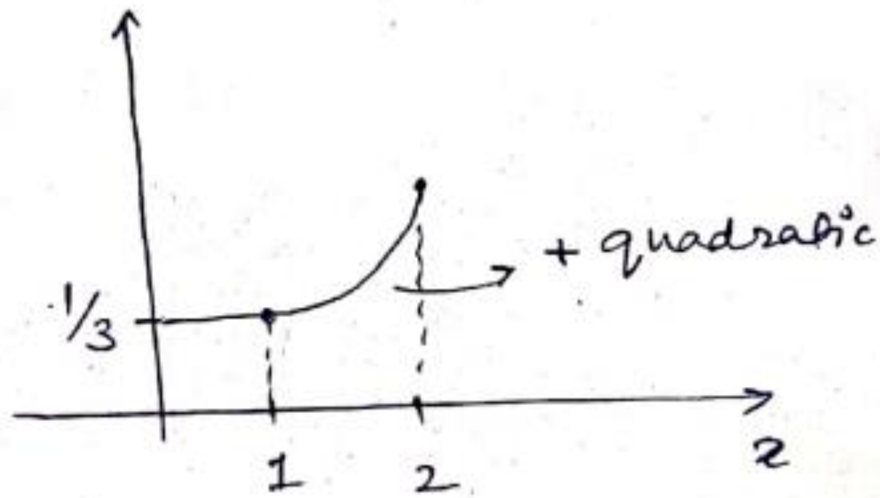
$$f(x) = \frac{x^2-1}{8}$$

$$f(x) = \begin{cases} 0 \; ; & x \leq 1 \\ \dfrac{x^2-1}{8} \; ; & 1 < x < 3 \\ 1 \; ; & x \geq 3 \end{cases}$$

Solution-2-II

A random variable X has the pdf as -

$$f(x) = \begin{cases} \frac{1}{3} ; & 0 \le x < 1 \\ \frac{2}{7}x^2 ; & 1 \le x \le 2 \\ 0 ; & \text{otherwise} \end{cases}$$



CDF

$$f_X(x) = f(x)$$

calculate each part separately -

For X < 0 ;   F(x)=0

For 0 ≤ X ≤ 1 ;   F(x)=?

PCP

$$= \int_0^x \frac{1}{3}\,dx = \left[\frac{x}{3}\right]_0^x = \frac{x}{3}$$

For $1 \le X \le 2$; $F(x)$=?

1nd In cumulative distribution function (CDF) area under 0 to 1 will also be consider - Therefore -

$$= \int_1^x \frac{2}{7} x^2\, dx + \frac{1}{3} = \frac{2}{7}\left[\frac{x^3}{3}\right]_1^x + \frac{1}{3}$$

$$= \frac{2(x^3-1)}{21} + \frac{1}{3} = \frac{2x^3+5}{21}$$

For $X = 2$; $F(x)$=?

$$= \frac{2x^2+5}{21} = \frac{2\times 2^3+5}{21} = \frac{21}{21} = 1$$

$$f(x) = \begin{cases} 0 & x < 0 \\ \frac{x}{3} & 0 \leq x < 1 \\ \frac{2x^2+5}{21} & 1 \leq x \leq 2 \\ 1 & x = 2 \end{cases}$$

Question-3

Consider an experiment with only two possible outcomes '**success**' or '**failure**'. Let ***p*** be the probability of success, ***n*** as the total number ***f*** trials and ***Xn*** be the number of successes.

Prove that for any

$$\in > 0, P\left(\left|\frac{X_n}{n} - p\right| < \in\right) \rightarrow\rightarrow\rightarrow 1 \quad as\, n \rightarrow \infty$$

Solution-3

Define:

- sample space Ω to consist of all possible infinite binary sequences of coin tosses;
- event H_1 - head on **first** toss;
- event E - first head on even numbered toss.

We want $P(E)$: using the Theorem of Total Probability, and the partition of Ω given by $\{H_1, H_1'\}$

$$P(E) = P(E|H_1)P(H_1) + P(E|H_1')P(H_1').$$

Now clearly, $P(E|H_1) = 0$ (given H_1, that a head appears on the first toss, E cannot occur) and also $P(E|H_1')$ can be seen to be given by

$$P(E|H_1') = P(E') = 1 - P(E),$$

(given that a head does **not** appear on the first toss, the required conditional probability is merely the probability that the sequence concludes after a further **odd** number of tosses, that is, the probability of E'). Hence $P(E)$ satisfies

$$P(E) = 0 \times p + (1 - P(E)) \times (1 - p) = (1 - p)(1 - P(E)),$$

so that

$$P(E) = \frac{1 - p}{2 - p}.$$

Alternatively, consider the partition of E into $E_1, E_2, \ldots$ where E_k is the event that the first head occurs on the $2k$th toss. Then $E = \bigcup_{k=1}^{\infty} E_k$, and

$$P(E) = P\left(\bigcup_{k=1}^{\infty} E_k\right) = \sum_{k=1}^{\infty} P(E_k).$$

Now $P(E_k) = (1-p)^{2k-1}p$ (that is, $2k-1$ tails, then a head), so

$$
\begin{aligned}
P(E) &= \sum_{k=1}^{\infty} (1-p)^{2k-1} p \\
&= \frac{p}{1-p} \sum_{k=1}^{\infty} (1-p)^{2k} = \frac{p}{1-p} \frac{(1-p)^2}{1-(1-p)^2} \\
&= \frac{1-p}{2-p}.
\end{aligned}
$$

Thanks !