Markov and Chebyshev, Convergence in Probability, Weak Law of Large Numbers

Q1. You place an electrode on your head and you measure the voltage.

Will you get the exact same number any time you measure?

Can we think of the Voltage measured as a Random Variable?

How can we figure out (estimate) the expected value of the RV Voltage?

A1.

No.

Yes.

Average many measurements -> this works because of the Week Law of Large Numbers.

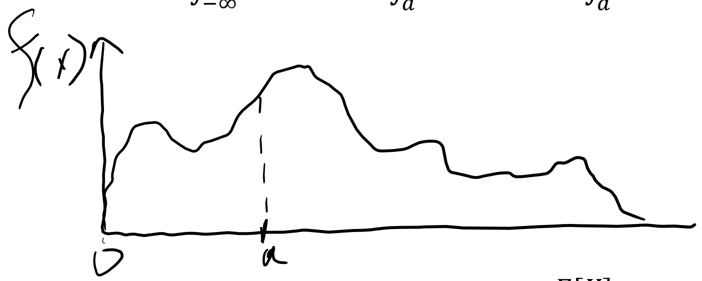
Markov Inequality

Q2. Let's say I tell you that the expected value of the Voltage RV is 10μV (small, i.e. close to zero) and that due to the way you've placed the electrode you will only read positive values. If I ask you whether you will ever measure a value of 10mV, what can you tell me? How about 1V?

A2.

E[X] is small and X>=0

$$E[X] = \int_{-\infty}^{\infty} xf(x)dx \geq \int_{a}^{\infty} xf(x)dx \geq \int_{a}^{\infty} af(x)dx = a\int_{a}^{\infty} f(x)dx = aP(X \geq a)$$



$$E[X] \geq \ aP(X \geq a) => \ P(X \geq a) \leq \frac{E[X]}{a}$$

OR

Y = 0 if X < a

Y = a if X >= a

$$E[Y] \leq E[X] \quad => aP(X \geq a) \leq E[X]$$

Q3. Can we use the Markov inequality to <u>estimate</u> the probability that a RV with an exponential distribution with λ=1 will be bigger or equal than a number a?

A3.

$$P(x \geq a) = 1/a$$

but we know that $P(x \geq a) = e^{-a}$

So the M.I. gives an upper bound (not a very good one). It is though a first rough estimation.

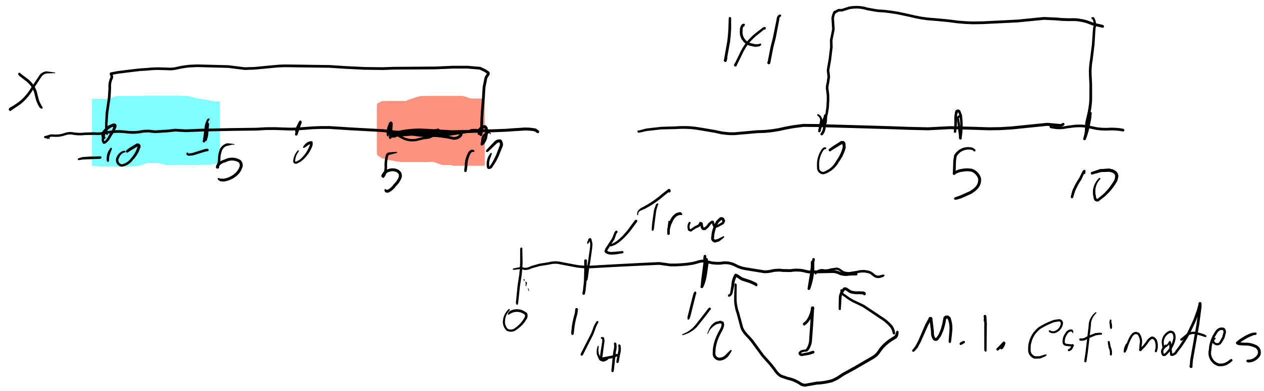
Q4. We have a uniform RV between -10 and 10. What is the true P(X>=5) for this distribution? What is the M.I. for a=5?

A4.

P(X>=5) = 5/20 = ¼

$$P(|X| \geq 5) = \frac{E[|X|]}{5} = \frac{5}{5} = 1$$

But $P(|X| \geq 5) = 2 * P(X \geq 5)$



Talk about the probability of the square of a variable being larger than the square of a number $(P(X \geq c) = P(X^2 \geq c^2))$

Chebyshev Inequality

Q5. Let's think about the variance of a probability distribution. For small variance what are the chances of a number being far away from the mean? Write the probability of the event that a value of the RV is away from its mean by a number c. Then show that:

$$P(|X - \mu| \geq c) \leq \frac{\sigma^2}{c^2}$$

A5.

Markov

$$P(|X - \mu| \geq c) = P((X - \mu)^2 \geq c^2) \leq \frac{E[(X - \mu)^2]}{c^2} = \frac{\sigma^2}{c^2}$$

Q6. Calculate the C.I. for c = k σ

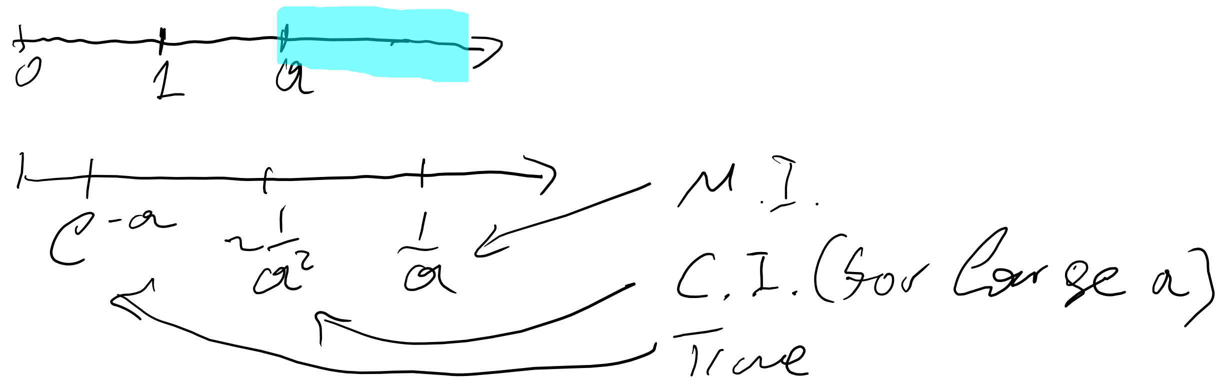
A6.

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

Q7. Calculate the C.I. for the exponential distribution with λ=1 for a > 1.

A7.

$$P(X \geq a) = P\big((X-1) \geq a-1\big) \leq P(|X-1| \geq a-1) \leq \frac{1}{(a-1)^2}$$



So, for large a the C.I. is a stronger bound. Is this true for smaller values of a (how about a=1.2 in the above example)?

Weak Law of Large Numbers

Back to **Q1.** Why does averaging many samples of the same RV result in an estimate of the RV's expectation value with an ever increasing accuracy?

A1 extended.

Assume you have measured the voltage from the electrode n times. Each measurement can be considered a RV. They all have the same mean and the same Variance (why?) and let's assume they are independent (is that true in this case?)

Define the RVs as $X_1, X_2, X_3, \ldots, X_n$

(You can also think of each RV as $X_i = \mu + W_i \;\; with\ E[W_i] = 0\ and\ indipendent\ W_i$)

Define Sample mean $M_n = \frac{X_1 + X_2 + X_3 + \ldots + X_n}{n}$ and call $\mu = E[X_i]$

What is M_n (another RV)? What does $E[M_n]$ mean (the theoretical expected value of the sample mean if we were to do multiple experiments).

I would like to show that as n becomes really big, M_n becomes closer and closer to μ!

Q8. Show that $E[M_n] = \mu$

A8.

$$E[M_n] = \frac{E[X_1 + X_2 + X_3 + \ldots + X_n]}{n} = \frac{n\mu}{n} = \mu$$

Q9. Calculate the $Var(M_n)$

A9.

$$Var(M_n) = \frac{Var(X_1 + X_2 + X_3 + \ldots + X_n)}{n^2} = \frac{n\,\sigma^2}{n^2} = \frac{\sigma^2}{n}$$

Q10. Use C.I. to calculate a bound of the probability that the M_n is not μ by ε. What happens if n goes to infinity?

A10.

$$P(|M_n - \mu| \geq \epsilon) \leq \frac{Var(M_n)}{\epsilon^2} = \frac{\sigma^2}{n\epsilon^2} \xrightarrow{for\, n \to \infty} 0$$

(Maybe talk here about the frequentist view of probability defining event A, p =P(A) and Xi indicator of A so the Mn becomes the empirical frequency of event A)

Accuracy of estimations using the C.I.

Q11. You use your voltage measuring technology to build a brain computer interface. But you realise that your algorithm works only for some people. Before you market the product you need to know the percentage of people (p) the BCI will work on. You try your tech on n people to find out. Can you say with absolute confidence what that p is? Also what do we need to do to keep the iid assumption alive here while we are trying this on some people?

A11.

No. One gets unlucky (or lucky some times) since the number of people you can sample is limited.

Do not the same person more than once.

Q12. Can you guarantee a very small error in the p (say 1%).

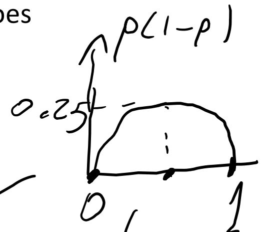
A12.

No. See A11.

Q13. Given an n (let's say 5000) what is probability that the error of p will be 5%?

A13.

Xi = 0 if the algo doesn't work on the ith person and Xi = 1 if it does



$$M_n = \frac{X_1 + X_2 + X_3 + \ldots + X_n}{n}$$

We want $|M_n - p| < 0.05$

Xi is Bernoulli so the variance is p(1-p)

C.I. : $P(|M_n - p| \geq 0.05) \leq \frac{\sigma^2}{n0.05^2} = \frac{p(1-p)}{5*10^3*0.05^2} \leq 0.02$

ask them to do this in Python

Q14. How many people should you try the algo on to have a probability of being wrong by 5% equal to 1%?

A14.

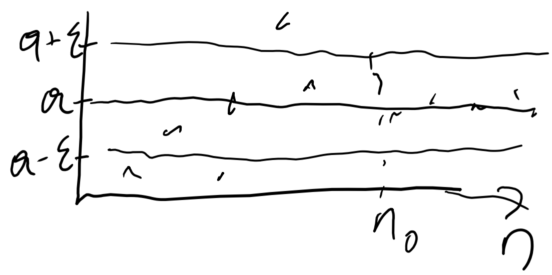
$$\frac{0.25}{n * 5 * 10^{-4}} \leq 0.01 => n \geq 50000$$

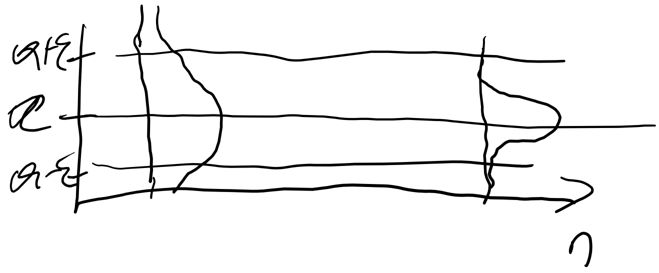
Convergence in probability

A sequence Y_n converges in probability to a number a if:

For any $\varepsilon > 0$, $\lim_{n \to \infty} P(|Y_n - a| \geq \varepsilon) = 0$

So the WLLN says that $M_n \xrightarrow[n \to \infty]{} \mu \ in\ probability$

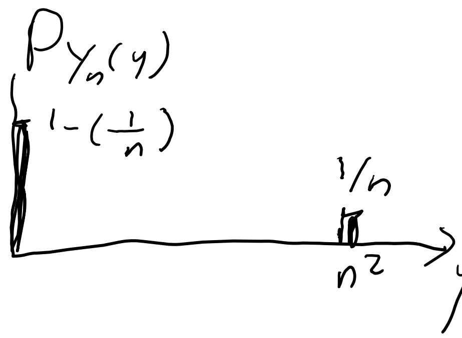




Here the limit is not a number but a distribution with an ever decreasing variance!

Q15. Find an example where a sequence Yn converges in probability to a number but the expectation value of any Yn (E[Yn]) converges to a different number

A15.



Yn = 0 with p = 1 – (1/n) and n^2 with p = 1/n

$P(|Y_n - 0| \geq \epsilon) = \frac{1}{n} \ \to \ 0 \ in\ prob.$

But $E[Y_n] = \ n^2 \frac{1}{n} = n \xrightarrow[n \to \infty]{} \infty$