STA258H5
University of Toronto Mississauga

Al Nosedal and Omid Jazi

Winter 2023

LAW OF LARGE NUMBERS

Convergence in Probability

infinite seq of RV's

The sequence of random variables $\overbrace{X_1,\ X_2,\ X_3, ..., X_n, ...}^{\text{infinite seq of RV's}}$ is said to **converge in probability** to the constant c, if for every $\epsilon > 0$,

$$\lim_{n\to\infty} P\left(|X_n - c| \leq \epsilon\right) = 1$$

OR

$$\lim_{n\to\infty} P\left(|X_n - c| > \epsilon\right) = 0$$

equivalent

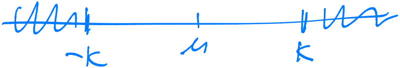
Notation: $X_n \xrightarrow{P} c$

seq

const

Chebyshev's Inequality

Special case of Markov's Ineq.



If X is a random variable with finite mean μ and variance σ^2, then for any value $k > 0$,

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

using complements

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

Another common way to express Markov

STA 256

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

The Weak Law of Large Numbers (WLLN)

$\bar{X} \xrightarrow{p} \mu$

Let $X_1, \; X_2, \; \ldots$ be a sequence of independent and identically distributed random variables, each having finite mean $E(X_i) = \mu$ and $Var(X_i) = \sigma^2$. Then, for any $\epsilon > 0$,

$$P\left[\left|\frac{X_1 + X_2 + \cdots + X_n}{n} - \mu\right| \geq \epsilon\right] \to 0 \;\text{ as }\; n \to \infty$$

We write
$\bar{X}_n \xrightarrow{\mathcal{P}} \mu$.

as $n \uparrow$ $\bar{X}$ gets closer to μ

proof:

[want to show

$$\lim_{n\to\infty} P(|\bar{X}_n - \mu| > \varepsilon) = 0 \qquad \forall \varepsilon > 0$$

$\bar{X}_n$ ← sample mean with size n]

we have $X_1, X_2, \ldots$ iid with $E(X_i) = \mu$, $Var(X_i) = \sigma^2$

Let $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$

By CLT, we know

$$\bar{X}_n \sim N\left(\mu_{\bar{x}} = \mu, \ \sigma_{\bar{x}}^2 = \frac{\sigma^2}{n}\right)$$

Chebyshev's states

For RV with mean μ, and variance σ^2

$$P(|X - \mu| > k) \leq \frac{\sigma^2}{k^2} \qquad k > 0$$

Applying Chebyshev for $\bar{x}_n$, for $k = \varepsilon > 0$

$$P(|\bar{X}_n - \mu| > \varepsilon) \leq \frac{\sigma^2/n}{\varepsilon^2}$$

$$P(|\bar{X}_n - \mu| > \varepsilon) \leq \frac{\sigma^2}{n\varepsilon^2}$$

$$\lim_{n\to\infty} P(|\bar{X}_n - \mu| > \varepsilon) \le \boxed{\lim_{n\to\infty} \frac{\sigma^2}{n\varepsilon^2}} \longrightarrow 0$$

$$\underbrace{\lim_{n\to\infty} P(|\bar{X}_n - \mu| > \varepsilon)}_{\text{probabilities must be non-neg}} \le 0$$

$$\lim_{n\to\infty} P(|\bar{X}_n - \mu| > \varepsilon) = 0$$

By def^{n} of convergence in prob

$$\bar{X}_n \xrightarrow{p} \mu$$

Proof

As
$E\left[\frac{X_1+X_2+\cdots+X_n}{n}\right]=\mu$ and $Var\left[\frac{X_1+X_2+\cdots+X_n}{n}\right]=\frac{\sigma^2}{n}$
it follows from Chebyshev's inequality that

$$P\left[\left|\frac{X_1+X_2+\cdots+X_n}{n}-\mu\right|\geq\epsilon\right]\leq\frac{\sigma^2}{n\epsilon^2}$$

and the result is proved.

Example

$$P(X=x) = \frac{e^{-\lambda}\lambda^x}{x!} \quad x=0,1,2,\ldots$$

Let X_i for $i = 1, 2, 3, \ldots$ be independent Poisson random variables, with rate parameter $\lambda = 3$. Prove $\bar{X}_n \xrightarrow{\mathcal{P}} 3$.

properties (Poisson)

$E(X) = Var(X) = \lambda$

In example $\lambda = 3$, $E(X) = 3$

Solution

Note that $E\left[\frac{X_1+X_2+\cdots+X_n}{n}\right] = 3$ and $Var\left[\frac{X_1+X_2+\cdots+X_n}{n}\right] = \frac{3}{n}$.
It follows from Chebyshev's inequality that for any $\epsilon > 0$,

$$P\left[\left|\overbrace{\frac{X_1+X_2+\cdots+X_n}{n}}^{\bar{X}_n} \overbrace{- 3}^{-\mu}\right| \geq \epsilon\right] \leq \frac{3}{n\epsilon^2}.$$

Therefore,

$$P\left[\left|\frac{X_1+X_2+\cdots+X_n}{n} - 3\right| \geq \epsilon\right] \to 0 \quad \text{as} \quad n \to \infty$$

and $\bar{X}_n \xrightarrow{\mathcal{P}} 3$.

Simulation

```
n=10;

trial=seq(1,n,by=1);

sample=rbinom(n,1,1/2);

plot(trial,cumsum(sample)/trial,type="l",ylim=c(0,1),
col="blue");

points(trial,cumsum(sample)/trial,col="red");

abline(h=0.5,lty=2, col="black");
```

1, 2, ..., 10

n = 10

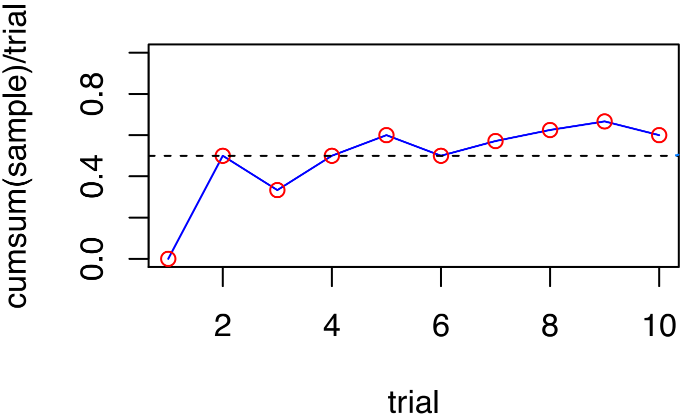
p = 1/2

$\underbrace{0, 1, 0, 1, 1, \ldots, ?}_{10 \text{ random}}$

10 samples

1 trial each

Simulation

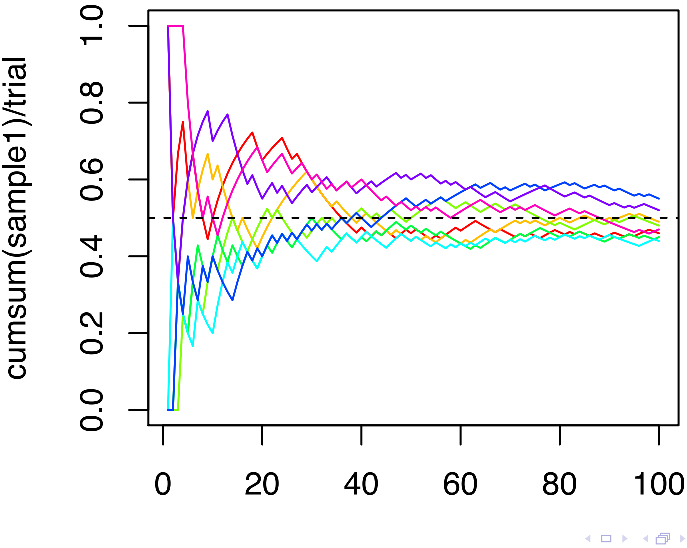


Simulation

```
n=100;

trial=seq(1,100,by=1);

sample1=rbinom(n,1,1/2);
sample2=rbinom(n,1,1/2);
sample3=rbinom(n,1,1/2);
sample4=rbinom(n,1,1/2);
sample5=rbinom(n,1,1/2);
sample6=rbinom(n,1,1/2);
sample7=rbinom(n,1,1/2);
sample8=rbinom(n,1,1/2);

colors=rainbow(8);
```

Simulation

```
plot(trial,cumsum(sample1)/trial,type="l",
col=colors[1],ylim=c(0,1));
lines(trial,cumsum(sample2)/trial,col=colors[2]);
lines(trial,cumsum(sample3)/trial,col=colors[3]);
lines(trial,cumsum(sample4)/trial,col=colors[4]);
lines(trial,cumsum(sample5)/trial,col=colors[5]);
lines(trial,cumsum(sample6)/trial,col=colors[6]);
lines(trial,cumsum(sample7)/trial,col=colors[7]);
lines(trial,cumsum(sample8)/trial,col=colors[8]);
abline(h=0.5,lty=2, col="black");
```


cumsum(sample1)/trial
0.0
0.2
0.4
0.6
0.8
1.0
0
20
40
60
80
100

Empirical Probability

The Law of Large Numbers gives us Empirical Probabilities for events.
Example: Toss a coin. Define the random variable:

$$X = \begin{cases} 1 & \text{heads up} \\ 0 & \text{tails up} \end{cases}$$

Think about sampling n of observations of X, and finding the sample mean $\bar{X}_n$. Let $n \to \infty$, and $\bar{X}_n \xrightarrow{\mathcal{P}} P(\text{heads up})$.