LARGE DEVIATION THEORY

(THE THEORY OF VERY UNLIKELY EVENTS)

LET $X_1, X_2, \ldots$ INDEPENDENT AND IDENTICALLY DISTRIBUTED (I.I.D.) BERNOULLI RANDOM VARIABLES (R.V.'S):

$\boxed{\mathbb{P}(X_i = 1) = p}$, $\boxed{\mathbb{P}(X_i = 0) = 1-p}$

X_i IS AN INDICATOR R.V., $\boxed{X_i \sim \mathrm{BER}(p)}$

$\boxed{S_n = X_1 + \ldots + X_n}$ S_n HAS BINOMIAL DISTRIBUTION

$\boxed{S_n \sim \mathrm{BIN}(n, p)}$ $\mathbb{P}(S_n = r) = \binom{n}{r} \cdot p^r \cdot (1-p)^{n-r}$

WEAK LAW OF LARGE NUMBERS:

$\boxed{\frac{S_n}{n} \xrightarrow{\mathbb{P}} p}$ $\frac{S_n}{n}$ CONVERGES IN PROBABILITY TO p

$\forall \varepsilon > 0 : \lim_{n \to \infty} \mathbb{P}\left(\left| \frac{S_n}{n} - p \right| \geq \varepsilon \right) = 0$

QUESTION: IF $\boxed{x > p}$, HOW SMALL IS THE PROBABILITY $\mathbb{P}\left(\frac{S_n}{n} \geq x\right)$ IF $n \gg 1$?

DEF: IF $a_1, a_2, \ldots \in \mathbb{R}_+$, $b_1, b_2, \ldots \in \mathbb{R}_+$

WE SAY $\boxed{a_n \approx b_n}$ IF $\boxed{\lim_{n \to \infty} \frac{1}{n} \ln\left(\frac{a_n}{b_n}\right) = 0}$

IN WORDS: THESE SEQUENCES GROW/DECAY AT THE SAME EXPONENTIAL RATE.

THM: $\boxed{p \leq x \leq 1} \Rightarrow \boxed{\mathbb{P}\left(\frac{S_n}{n} \geq x\right) \approx e^{-n \cdot I(x)}}$

$\boxed{0 \leq x \leq p} \Rightarrow \boxed{\mathbb{P}\left(\frac{S_n}{n} \leq x\right) \approx e^{-n \cdot I(x)}}$

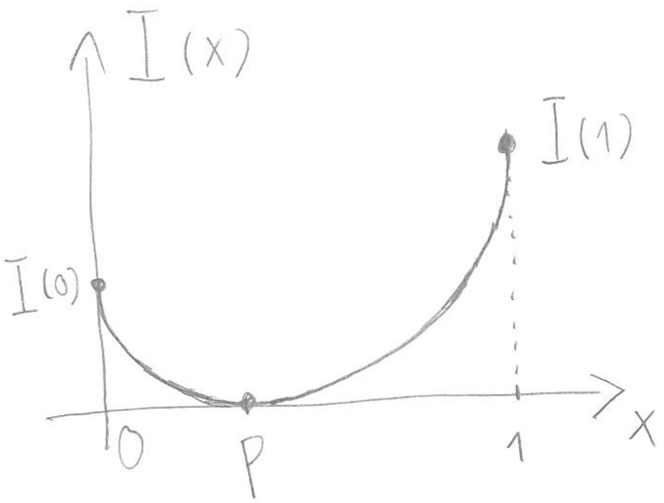
WHERE $\boxed{I(x) = x \cdot \ln\left(\frac{x}{p}\right) + (1-x) \cdot \ln\left(\frac{1-x}{1-p}\right)}$

IN INFORMATION THEORY: $I(x)$ IS KNOWN AS THE KULLBACK-LEIBLER DIVERGENCE FROM BER(p) TO BER(x).

ALSO KNOWN AS: RELATIVE ENTROPY.

WE CALL $x \mapsto I(x)$ THE LARGE DEVIATION RATE FUNCTION OF THE BER(p) DISTRIBUTION.

BEFORE PROOF, SOME COMMENTS: $\boxed{0 \le x \le 1}$



$$I'(x) = \ln\left(\frac{x}{p}\right) - \ln\left(\frac{1-x}{1-p}\right)$$

$$I''(x) = \frac{1}{x} + \frac{1}{1-x}$$

$\boxed{I(p) = 0}$ $\boxed{I'(p) = 0}$ $\boxed{I''(x) > 0}$

$\boxed{I(1) = \ln\left(\frac{1}{p}\right)}$

MAKES SENSE:

$$\lim_{n\to\infty} \mathbb{P}\left(\frac{S_n}{n} \ge p\right) = \frac{1}{2}$$ BY CENTRAL LIMIT THM

THUS $\lim_{n\to\infty} \frac{1}{n} \ln\left(\mathbb{P}\left(\frac{S_n}{n} \ge p\right)\right) = 0$, THUS $\boxed{I(p) = 0}$

$\mathbb{P}\left(\frac{S_n}{n} \ge 1\right) = \mathbb{P}(S_n = n) = p^n$, THUS

$\lim_{n\to\infty} \frac{1}{n} \ln\left(\mathbb{P}\left(\frac{S_n}{n} \ge 1\right)\right) = \ln(p)$, THUS

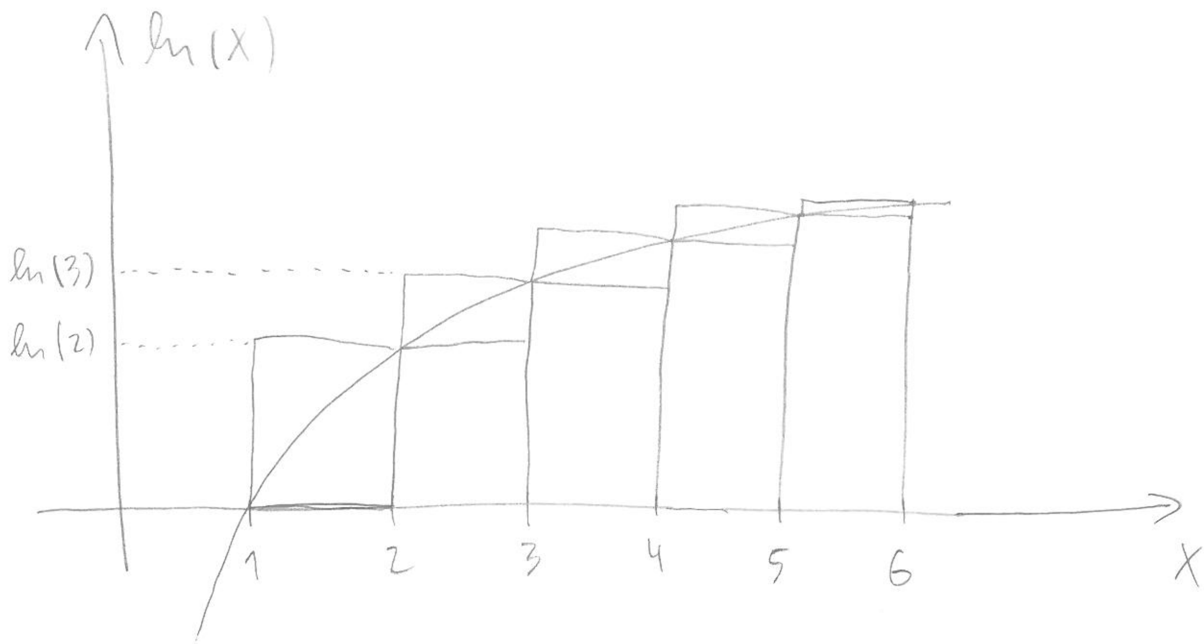
$\mathbb{P}\left(\frac{S_n}{n} \ge 1\right) \approx e^{-n \cdot I(1)}$, WHERE $\boxed{I(1) = \ln\left(\frac{1}{p}\right)}$

LEMMA: $\boxed{n! \approx n^n \cdot e^{-n}}$

(THIS IS A VERY CRUDE VERSION OF STIRLING'S FORMULA)

PROOF OF LEMMA:

$$\ln(n!) = \ln(1) + \ln(2) + \ldots + \ln(n)$$



THUS $\underbrace{\int_1^n \ln(x)\,dx}_{a_n = \boxed{n \cdot \ln(n) - (n-1)}} \le \ln(n!) \le \underbrace{\int_1^{n+1} \ln(x)\,dx}_{b_n = \boxed{(n+1)\cdot \ln(n+1) - n}}$

THUS $n^n \cdot e^{-(n-1)} = e^{a_n} \le n! \le e^{b_n} = (n+1)^{n+1} \cdot e^{-n}$

WE WILL SHOW THAT $\boxed{e^{a_n} \overset{A}{\approx} e^{b_n} \overset{B}{\approx} n^n \cdot e^{-n}}$

WHICH WILL IMPLY THE CLAIM OF LEMMA.

$$\lim_{n\to\infty} \frac{1}{n} \ln\left(\frac{e^{b_n}}{e^{a_n}}\right) = \lim_{n\to\infty} \frac{1}{n}\left(\int_n^{n+1} \ln(x)\,dx\right) = 0$$

$$\lim_{n\to\infty} \frac{1}{n} \ln\left(\frac{e^{a_n}}{n^n \cdot e^{-n}}\right) = \lim_{n\to\infty} \frac{1}{n} = 0$$

PROOF OF THM STATED ON PAGE 2:

FIRST WE WILL SHOW THAT

$$\boxed{\mathbb{P}\left(S_n = \lceil n \cdot x \rceil\right) \approx e^{-n \cdot I(x)}} \qquad 0 \leq x \leq 1$$

$$\mathbb{P}\left(S_n = \lceil n \cdot x \rceil\right) = \frac{n!}{\lceil n \cdot x \rceil! \left(n - \lceil n \cdot x \rceil\right)!} \cdot p^{\lceil n \cdot x \rceil} \cdot (1-p)^{n - \lceil n \cdot x \rceil} \approx$$

$$\approx \frac{n^n \cdot e^{-n}}{(n \cdot x)^{n \cdot x} \cdot e^{-n \cdot x} \cdot (n \cdot (1-x))^{n \cdot (1-x)} \cdot e^{-n \cdot (1-x)}} \cdot p^{n \cdot x} \cdot (1-p)^{n \cdot (1-x)}$$

$$= \frac{n^n}{n^{n \cdot x} \cdot x^{n \cdot x} \cdot n^{n \cdot (1-x)} \cdot (1-x)^{n \cdot (1-x)}} \cdot p^{n \cdot x} \cdot (1-p)^{n \cdot (1-x)}$$

$$= \frac{p^{n \cdot x} \cdot (1-p)^{n \cdot (1-x)}}{x^{n \cdot x} \cdot (1-x)^{n \cdot (1-x)}} = \left(\left(\frac{p}{x}\right)^{x} \cdot \left(\frac{1-p}{1-x}\right)^{(1-x)}\right)^{n}$$

$$= \exp\left(n \cdot \left(x \cdot \ln\left(\frac{p}{x}\right) + (1-x) \cdot \ln\left(\frac{1-p}{1-x}\right)\right)\right)$$

$$= e^{-n \cdot I(x)} \quad \checkmark$$

IN ORDER TO COMPLETE THE PROOF OF THM FROM PAGE 2, IT REMAINS TO SHOW THAT...

$$\boxed{\mathbb{P}\left(S_n = \lceil n\cdot x\rceil\right) \approx \mathbb{P}\left(S_n \geq \lceil n\cdot x\rceil\right)}$$ FOR $p \leq x \leq 1$

PROOF OF THIS : LET $P_r = \mathbb{P}(S_n = r)$

WE WILL SHOW THAT

$$\boxed{P_{\lceil n\cdot x\rceil} \geq P_r} \quad (\star)$$ FOR ANY $\lceil n\cdot x\rceil \leq r \leq n$

THEN IT FOLLOWS THAT

$$\underbrace{\mathbb{P}\left(S_n = \lceil n\cdot x\rceil\right)}_{P_{\lceil n\cdot x\rceil}} \leq \underbrace{\mathbb{P}\left(S_n \geq \lceil n\cdot x\rceil\right)}_{= \sum_{r=\lceil n\cdot x\rceil}^{n} P_r} \leq \underbrace{n\cdot\mathbb{P}\left(S_n = \lceil n\cdot x\rceil\right)}_{n\cdot P_{\lceil n\cdot x\rceil}}$$

AND WE HAVE $\boxed{P_{\lceil n\cdot x\rceil} \approx n\cdot P_{\lceil n\cdot x\rceil}}$, since

$$\lim_{n\to\infty} \frac{1}{n} \ln\left(\frac{n\cdot P_{\lceil n\cdot x\rceil}}{P_{\lceil n\cdot x\rceil}}\right) = \lim_{n\to\infty} \frac{1}{n} \ln(n) = 0 \quad \checkmark$$

IT REMAINS TO SHOW $(\star)$: IT IS ENOUGH TO SHOW THAT $\lceil n\cdot x\rceil \leq r \Rightarrow \boxed{\frac{P_{r+1}}{P_r} \leq 1}$

INDEED:

$$\frac{P_{r+1}}{P_r} = \frac{n-r}{r+1}\cdot\frac{p}{1-p} = \underbrace{\left(\frac{1-\frac{r}{n}}{1-p}\right)}_{\leq 1}\cdot\underbrace{\left(\frac{p}{\frac{r}{n}+\frac{1}{n}}\right)}_{\leq 1} \quad \checkmark$$

SINCE $\boxed{p \leq x \leq \frac{r}{n}}$

THE PROOF OF $\boxed{0 \leq x \leq p} \Rightarrow \boxed{\mathbb{P}\left(\frac{S_n}{n} \leq x\right) \approx e^{-n \cdot I(x)}}$

IS SIMILAR AND WE OMIT IT.

NOW $X_1, X_2, \ldots$ ARE GENERAL I.I.D. R.V.'S

IF $m = \mathbb{E}(X_i)$ THEN $\forall x > m$ WE HAVE

$\lim_{n \to \infty} \mathbb{P}\left(\frac{S_n}{n} \geq x\right) = 0$ BY WEAK LAW OF LARGE NUMBERS.

BUT HOW FAST DOES IT CONVERGE TO ZERO?

<u>DEF</u>: $\lambda \in \mathbb{R}$: $\boxed{Z(\lambda) := \mathbb{E}\left(e^{\lambda \cdot X_i}\right)}$

THE MOMENT GENERATING FUNCTION OF X_i

LET $\boxed{\hat{I}(\lambda) := \ln(Z(\lambda))}$

THE LOGARITHMIC MOMENT GEN. FUNCTION OF X_i

NOTE: $\mathbb{E}\left(e^{\lambda S_n}\right) = \mathbb{E}\left(e^{\lambda \cdot X_1} \cdot e^{\lambda \cdot X_2} \cdot \ldots \cdot e^{\lambda \cdot X_n}\right) =$

$= \mathbb{E}\left(e^{\lambda X_1}\right) \cdot \mathbb{E}\left(e^{\lambda \cdot X_2}\right) \cdot \ldots \cdot \mathbb{E}\left(e^{\lambda \cdot X_n}\right) = (Z(\lambda))^n$

THE EXPECTED VALUE OF THE PRODUCT OF INDEPENDENT R.V.'S IS EQUAL TO THE PRODUCT OF THEIR EXPECTATIONS.