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

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

$\lim_{n \to \infty} \mathbb{P}\left(\frac{S_n}{n} \ge x\right) = 0$ 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.

NOW FOR ANY $\lambda \geq 0$: MARKOV'S INEQUALITY

$$\mathbb{P}\left(\frac{S_m}{m} \geq x\right) = \mathbb{P}\left(e^{\lambda \cdot S_m} \geq e^{\lambda \cdot m \cdot x}\right) \leq$$

$$\leq \frac{\mathbb{E}\left(e^{\lambda \cdot S_m}\right)}{e^{\lambda \cdot m \cdot x}} = \frac{\left(Z(\lambda)\right)^m}{e^{\lambda \cdot m \cdot x}} = e^{m \cdot \left(\hat{I}(\lambda) - \lambda \cdot x\right)}$$

NOW CHOOSE λ THAT GIVES THE BEST BOUND!

IF $\boxed{X_i \sim \mathrm{BER}(p)}$ THEN $\boxed{Z(\lambda) = p \cdot e^{\lambda \cdot 1} + (1-p) \cdot e^{\lambda \cdot 0}}$

THUS $\boxed{\hat{I}(\lambda) = \ln\left((1-p) + p \cdot e^{\lambda}\right)}$

GIVEN $x \geq \mathbb{E}(X_i) = p$, WHAT IS $\min\limits_{\lambda \geq 0} \left(\overbrace{\hat{I}(\lambda) - \lambda \cdot x}^{\psi(\lambda)}\right)$?

$$\psi'(\lambda) = \hat{I}'(\lambda) - x = \frac{p \cdot e^{\lambda}}{(1-p) + p \cdot e^{\lambda}} - x = 0$$ WANT

$\psi'(\lambda^*) = 0$ IF $\boxed{\lambda^* := \ln\left(\frac{x}{1-x} \cdot \frac{1-p}{p}\right)}$, PLUGGING THIS BACK INTO ψ:

$$\psi(\lambda^*) = \ln\left((1-p) + p \cdot e^{\lambda^*}\right) - \lambda^* \cdot x =$$

SEE PAGE 2

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

THUS $\boxed{\mathbb{P}\left(\frac{S_m}{m} \geq x\right) \leq e^{-m \cdot I(x)}}$ THIS IS GOOD NEWS!

OUR NEW METHOD GAVE ...

OUR NEW METHOD GAVE THE OPTIMAL UPPER BOUND ON $\mathbb{P}\left(\frac{S_n}{n} \geq x\right)$ IN THE CASE OF $\boxed{X_i \sim BER(p)}$, SO MAYBE IT IS ALSO OPTIMAL IN THE CASE OF GENERAL I.I.D. $X_1, X_2, \dots$

BACK TO THE GENERAL CASE:
WE HAVE SHOWN THAT IF $S_n = X_1 + \dots + X_n$ AND $x > \mathbb{E}(X_i)$ THEN $\mathbb{P}\left(\frac{S_n}{n} \geq x\right) \leq e^{-n \cdot I(x)}$

WHERE $\boxed{I(x) := \max_{\lambda \geq 0} \left(\lambda x - \hat{I}(\lambda)\right)}$

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

$Z(\lambda) = \mathbb{E}(e^{\lambda X_i})$

<u>DEF</u>: $f: \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$, LET US DEFINE

$$\boxed{\hat{f}(\lambda) = \sup_{x \in \mathbb{R}} \{\lambda \cdot x - f(x)\}}$$

FOR ANY $\lambda \in \mathbb{R}$

(NOTE: $\hat{f}: \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$)

$\hat{f}$ IS THE <u>LEGENDRE TRANSFORM</u> OF f

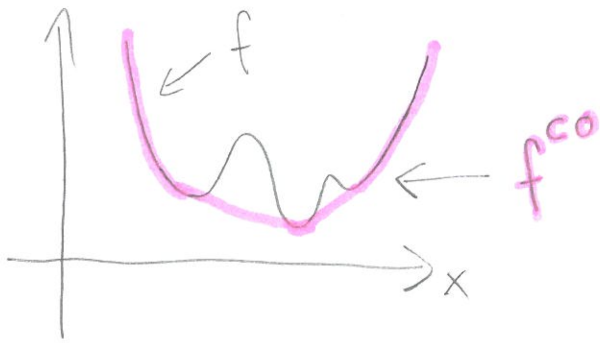
WE WILL NOW LEARN ABOUT THE PROPERTIES OF $\boxed{f \mapsto \hat{f}}$

CLAIM: $\hat{f}$ IS A **CONVEX** FUNCTION

PROOF: FOR ANY FIXED $x \in \mathbb{R}$ THE FUNCTION $\boxed{\lambda \longmapsto \lambda \cdot x - f(x)}$ IS AN AFFINE LINEAR FUNCTION, THUS $\hat{f}$ IS THE SUPREMUM OF — " — 'S, HENCE IT IS CONVEX. ✓

DEF: THE **LOWER CONVEX ENVELOPE** f^{co} OF A FUNCTION f IS THE SUPREMUM OF OF THOSE AFFINE LINEAR FUNCTIONS THAT LIE BELOW f:

$$f^{co}(x) = \sup\{g(x) : g(y) = a \cdot y + b,\ \forall y\ g(y) \le f(y)\}$$



FACT: IF f IS CONVEX AND CONTINUOUS, THEN $\boxed{f \equiv f^{co}}$

CLAIM: IF $\boxed{f \le g}$ THEN $\boxed{\hat{g} \le \hat{f}}$

PROOF: $\forall x\ \forall \lambda$: $\lambda \cdot x - f(x) \geq \lambda \cdot x - g(x)$, THUS

$$\hat{f}(\lambda) = \sup_x \{\lambda x - f(x)\} \geq \sup_x \{\lambda x - g(x)\} = \hat{g}(\lambda)$$

CLAIM: $\boxed{\hat{\hat{f}} \leq f^{co}}$ THE LEGENDRE TRANSFORM OF THE LEGENDRE TRANSFORM OF f

LIES BELOW THE LOWER CONVEX ENVELOPE OF f.

PROOF: ENOUGH TO SHOW THAT $\boxed{\hat{\hat{f}} \leq f}$ ✪,

SINCE $\hat{\hat{f}}$ IS THE SUPREMUM OF SOME AFFINE LINEAR FUNTIONS THAT LIE BELOW f BY ✪, BUT f^{co} IS THE SUP. OF ALL AFF. LIN. FUNCTIONS THAT LIE BELOW f, SO $\hat{\hat{f}} \leq f^{co}$.

NOW WE SHOW ✪:

$\forall \lambda \; \hat{f}(\lambda) = \sup\limits_{x} \{\lambda x - f(x)\}$, SO

$\forall \lambda \, \forall x \quad \hat{f}(\lambda) \geq \lambda x - f(x)$, SO

$\forall \lambda \, \forall x \quad f(x) \geq \lambda x - \hat{f}(\lambda)$, SO

$\forall x \quad f(x) \geq \sup\limits_{\lambda} \{\lambda x - \hat{f}(\lambda)\} = \hat{\hat{f}}(x)$

CLAIM: ACTUALLY WE HAVE $\boxed{\hat{\hat{f}} = f^{co}}$

(THUS, IN PARTICULAR, IF f IS CONVEX AND CONTINUOUS, THEN $\boxed{\hat{\hat{f}} = f}$)

PROOF: FIRST NOTE THAT IF $\boxed{g(x) = ax + b}$

THEN $\boxed{\hat{\hat{g}} = g}$. INDEED:

$$\hat{g}(\lambda) = \sup_x \{\lambda x - ax - b\} = \begin{cases} +\infty & \text{IF } \lambda \neq a \\ -b & \text{IF } \lambda = a \end{cases}$$

$$\hat{\hat{g}}(x) = \sup_\lambda \{\lambda x - \hat{g}(\lambda)\} = ax - \hat{g}(a) = ax + b = g(x)$$

NOW IN THE CASE OF GENERAL f, WE ONLY NEED TO SHOW $\boxed{\hat{\hat{f}} \geq f^{co}}$. ENOUGH TO SHOW THAT IF $g(x) = ax + b$ AND $g \leq f$ THEN $g \leq \hat{\hat{f}}$. AND INDEED:

$$\boxed{g \leq f} \Longrightarrow \boxed{\hat{g} \geq \hat{f}} \Longrightarrow \boxed{\hat{\hat{g}} \leq \hat{\hat{f}}}$$

BUT $\hat{\hat{g}} = g$ ✓

CLAIM: IF f'' IS POSITIVE AND CONTINUOUS, THEN $\boxed{\hat{f}' \overset{B}{=} (f')^{-1}}$ AND

$$\boxed{\hat{f}(x) \overset{A}{=} x \cdot (f')^{-1}(x) - f\left((f')^{-1}(x)\right)}$$

PROOF: $\hat{f}(x) = \sup_{\lambda} \{ \lambda x - f(\lambda) \}$

GIVEN x, LET $\psi(\lambda) = \lambda x - f(\lambda)$

THEN ψ IS STRICTLY CONCAVE, SO TO MAXIMIZE IT, WE NEED TO FIND THE UNIQUE λ^* FOR WHICH $\psi'(\lambda^*) = 0$.

$\psi'(\lambda) = x - f'(\lambda)$, THUS WANT: $f'(\lambda^*) = x$, THUS $\boxed{\lambda^* = (f')^{-1}(x)}$ C AND THUS

$$\hat{f}(x) = \psi(\lambda^*) = \lambda^* \cdot x - f(\lambda^*) = x \cdot (f')^{-1}(x) - f\left((f')^{-1}(x)\right)$$

NOW $\frac{d}{dx}(f')^{-1}(x) \underset{D}{=} \frac{1}{f''(\lambda^*)}$, SO

$$\hat{f}'(x) = \lambda^* + x \cdot \frac{1}{f''(\lambda^*)} - \underbrace{f'(\lambda^*)}_{x} \cdot \frac{1}{f''(\lambda^*)} = \lambda^* = (f')^{-1}(x)$$

NOW LET US CHECK FOR FUN THAT $\boxed{\hat{\hat{f}} = f}$

$$\hat{\hat{f}}(\lambda) = \lambda \cdot (\hat{f}')^{-1}(\lambda) - \hat{f}\left((\hat{f}')^{-1}(\lambda)\right) = \lambda \cdot f'(\lambda) - \hat{f}(f'(\lambda)) =$$

$$\lambda \cdot f'(\lambda) - \Big(f'(\lambda) \cdot \underbrace{(f')^{-1}(f'(\lambda))}_{\lambda} - \underbrace{f\left((f')^{-1}(f'(\lambda))\right)}_{f(\lambda)} \Big) = f(\lambda)$$