Cramer's thm

Thm Let

- $X_1, \dots, X_n$ iid
- $\mathbb{E}[X_1] = \mu$ exists

Then

$$\boxed{\begin{aligned} &\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}(S_n \geq an) = -I(a), \quad a \geq \mu \\ &\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}(S_n \leq an) = -I(a), \quad a \leq \mu \end{aligned}}$$

where

$$I(a) = \sup_t \{ at - \log(\underbrace{\varphi(t)}_{\mathbb{E}[e^{tX_1}]}) \}$$

Reduced proof Under the assumption

$$\varphi(t) = \mathbb{E}[e^{tX_1}] < \infty$$

for all $t \in \mathbb{R}$.

Legendre transform

set

$$L(t) = \log \varphi(t)$$

Then

Legendre transform

$$I(x) = \sup \{ xt - L(t) ; t \in \mathbb{R} \}$$

Thus I convex.

Convexity of L

we have

$$\varphi(\theta t_1 + (1-\theta) t_2) = E[e^{\theta t_1 X_1} e^{(1-\theta) t_2 X_2}]$$

$$\leq E[e^{t_1 X_1}]^{\theta} E[e^{t_2 X_1}]^{1-\theta}$$

$$= \varphi(t_1)^{\theta} \varphi(t_2)^{1-\theta}$$

Thus

$$L(\theta t_1 + (1-\theta) t_2)$$

$$\leq \theta L(t_1) + (1-\theta) L(t_2)$$

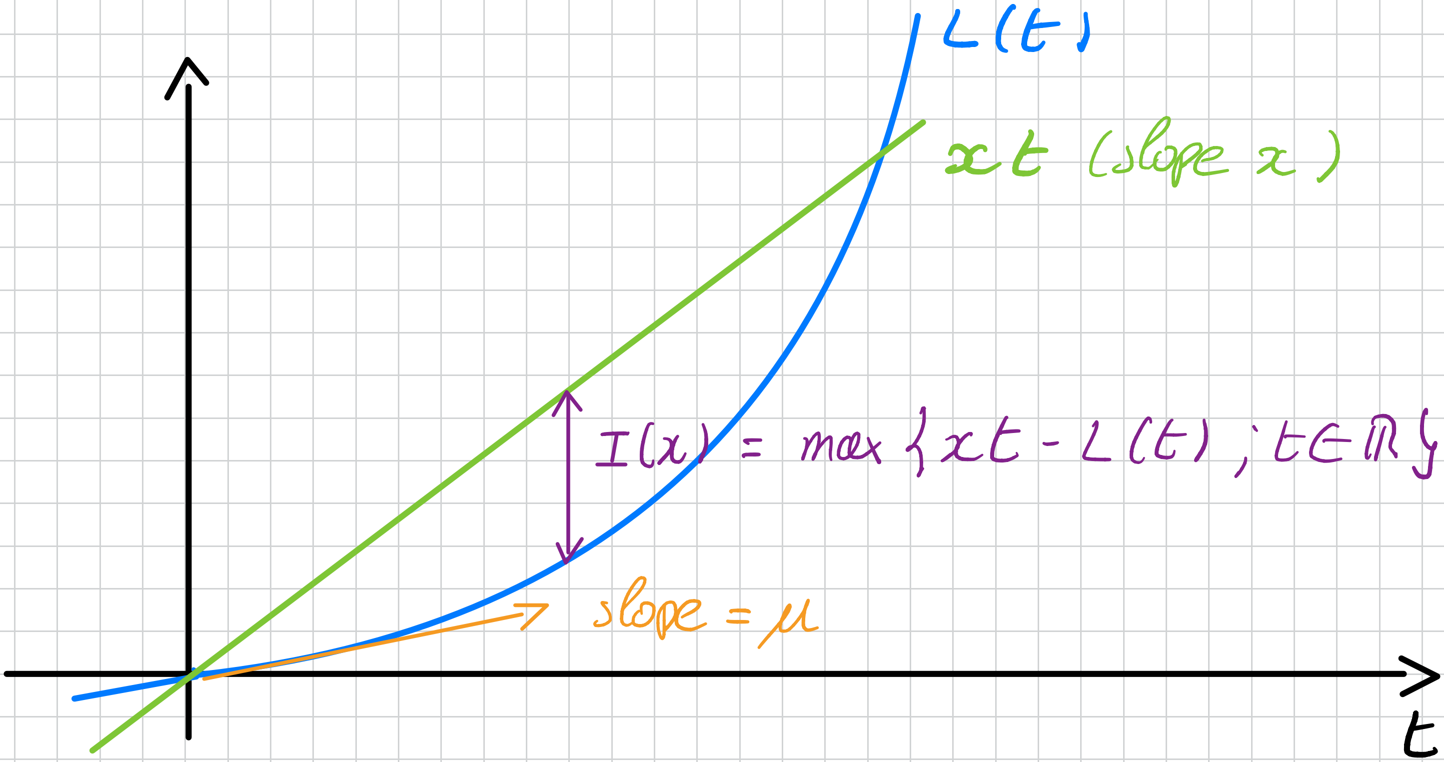
$$\Rightarrow \boxed{L \text{ convex}}$$

Function L and moments

We have

$$\boxed{L(0) = 0, \quad L'(0) = \mu, \quad L''(0) = \mathrm{Var}(X_1)}$$

Construction of I



We have

- $x > \mu \Rightarrow$ sup attained at $t \geq 0$

  $I(x) = \sup\{xt - L(t);\ t \geq 0\}$

- $x < \mu \Rightarrow$ sup attained at $t \leq 0$

  $I(x) = \sup\{xt - L(t);\ t \leq 0\}$

Upper bound for LDP   Write

$$\mathbb{P}(S_n \geq na) = \mathbb{P}(e^{tS_n} \geq e^{tna})$$

$$\leq \frac{\mathbb{E}[e^{tS_n}]}{e^{nta}}$$

Here $t \geq 0$, which means this is for $a > \mu$

$$= \frac{(\varphi(t))^n}{e^{nta}}$$

$$= \exp(n \ln(\varphi(t)) - nta)$$

$$= \exp(-n(at - L(t)))$$

Hence

$$\frac{1}{n} \ln \mathbb{P}(S_n \geq na)$$

$$\leq -(at - L(t)) \quad \forall t \geq 0$$

$$\Rightarrow \frac{1}{n} \ln \mathbb{P}(S_n \geq a) \leq - \sup_{t \geq 0}\{at - L(t)\}$$

for $a > \mu$

$$\Rightarrow \frac{1}{n} \ln \mathbb{P}(S_n \geq a) \leq -I(a)$$

1- Lower bound for Cramer

Lower bound idea

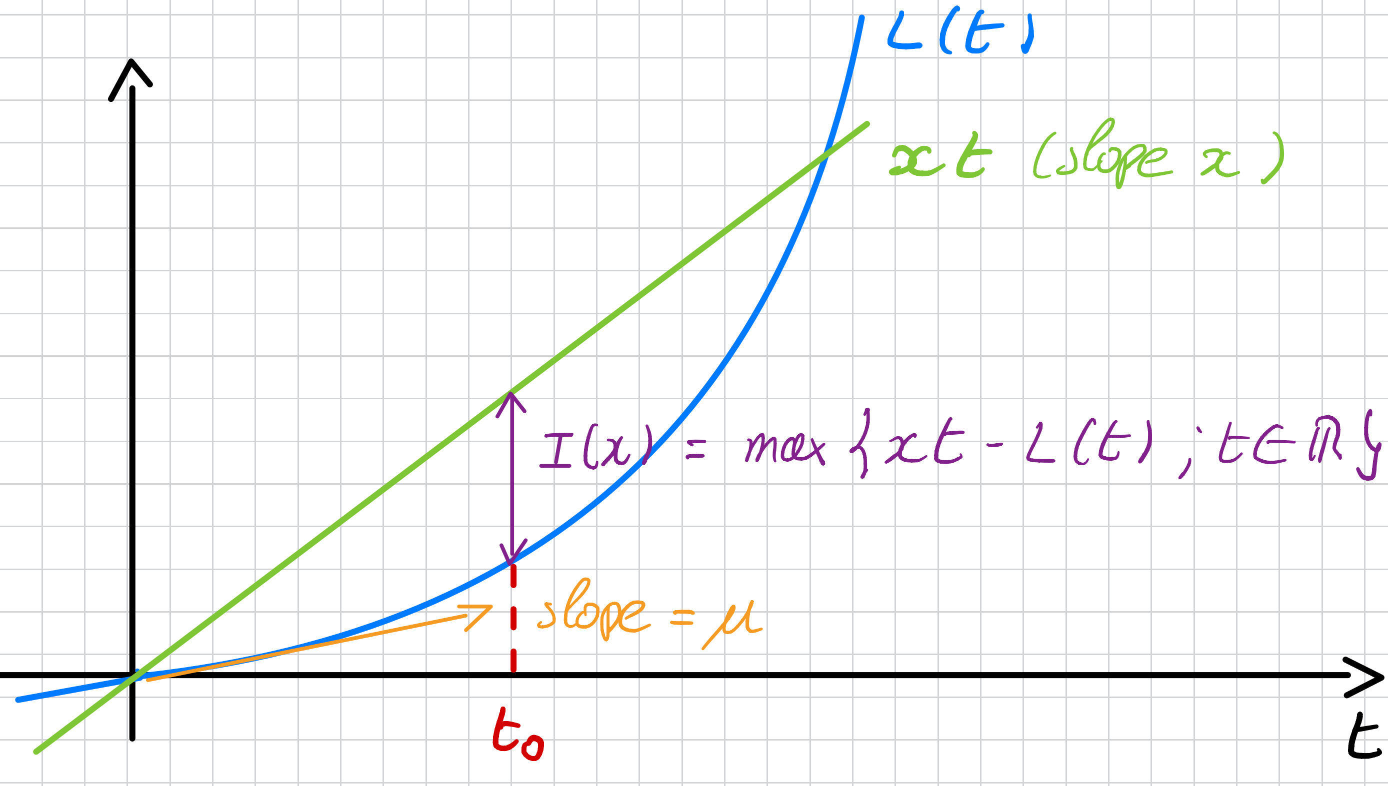
(i) Usually for a r.v X,

$a \mapsto \mathbb{P}(|X - a| > \varepsilon)$ max when $a = \mu$

(ii) we use I to transform $\mathcal{L}(x)$ into a distribution with mean a $\longrightarrow$ $\boxed{\text{rare event becomes typical}}$

(iii) The cost to change the measure will be quantified by $e^{-nI(a)}$

Lower bound setting



Hyp: $I(a) = a t_0 - L(t_0)$

Derivative of L

If

$$I(a) = \sup\{\overbrace{at - L(t)}^{K(t)} ; t \in \mathbb{R}\}$$

$$= a t_0 - L(t_0)$$

then $K'(t_0) = 0$

$\Rightarrow$ $L'(t_0) = a$

New distribution

Define $\mathbb{Q} \ll \mathbb{P}$ with

$$\boxed{\frac{d\mathbb{Q}}{d\mathbb{P}}(x) = \frac{e^{t_0 x}}{\varphi(t_0)}}$$

Mean under Q we have

$$\mathbb{E}_Q[X]$$

$$= \mathbb{E}_P\left[X \frac{dQ}{dP}(x)\right]$$

$$= \frac{1}{\varphi(t_0)} \quad \mathbb{E}_P\left[X e^{t_0 X}\right]$$

$$= \frac{\varphi'(t_0)}{\varphi(t_0)}$$

$$= d(\ln(\varphi(t_0))) = L'(t_0) = a$$

$$\Rightarrow \quad \boxed{\mathbb{E}_Q[X] = a}$$

$\longrightarrow$

Variance under $\mathbb{Q}$ we have

$$\mathbb{E}_{\mathbb{Q}}[X^2] = \frac{\mathbb{E}_{\mathbb{P}}[X^2 e^{t_0 X}]}{\varphi(t_0)}$$

$$= \frac{\varphi''(t_0)}{\varphi(t_0)}$$

Thus

$$\mathrm{Var}_{\mathbb{Q}}(X) = \mathbb{E}_{\mathbb{Q}}[X^2] - (\mathbb{E}_{\mathbb{Q}}[X])^2$$

$$= \frac{\varphi''(t_0)}{\varphi(t_0)} - \frac{(\varphi'(t_0))^2}{(\varphi(t_0))^2}$$

$$= \frac{1}{(\varphi(t_0))^2}\left\{\varphi''(t_0)\,\varphi(t_0) - (\varphi'(t_0))^2\right\}$$

$$= \left(\frac{\varphi'}{\varphi}\right)'(t_0) = L''(t_0)$$

Hence

$$\boxed{\mathrm{Var}_{\mathbb{Q}}(X) = L''(t_0) \equiv \hat{\sigma}^2 > 0}$$

L convex

Notation for multidimensional dist we set

(i) $\mathbb{P}_n = \mathcal{L}(X_1, \dots, X_n) = \mathbb{P}^{\otimes n}$,

so that

$$\mathbb{P}(S_n \geq na) = \mathbb{P}_n(S_n \geq an)$$

(ii) $\mathbb{Q}_n \equiv$ probability with

$$\frac{d\mathbb{Q}_n}{d\mathbb{P}_n} = \frac{e^{t_0(x_1+\dots+x_n)}}{(\varphi(t_0))^n} = \prod_{i=1}^{n} \frac{e^{t_0 x_i}}{\varphi(t_0)}$$

Thus under $\mathbb{Q}_n$ we have

$\{X_i ; 1 \leq i \leq n\}$ i.i.d with $X_1 \sim \mathbb{Q}$

Hyp we treat the case

$$\boxed{a > \mu \quad (\Rightarrow \quad t_0 \geq 0)}$$

computation - heuristics write

$$\mathbb{P}\left(\frac{S_n}{n} > a\right)$$

$$= \mathbb{E}_Q\left\{(\varphi(t_0))^n e^{-t_0 S_n} \mathbb{1}\left(\frac{S_n}{n} \geq a\right)\right\}$$

$$= (\varphi(t_0))^n \mathbb{E}_Q\left[e^{-n t_0 S_n/n} \mathbb{1}(S_n/n \geq a)\right]$$

justify!

$$\simeq \varphi(t_0)^n e^{-n a t_0} Q\left(\frac{S_n}{n} \geq a\right)$$

of order $c, \approx 1$

$$\simeq \exp(-n(a t_0 - L(t_0)))$$

$$\simeq \exp(-n I(a))$$

computation for lower bound we have

$$\mathbb{P}(S_n \geq an)$$
$$= \mathbb{E}_{P_n}[\ \mathbb{1}(an \leq S_n)\ ]$$
$$= (\varphi(t_0))^n\ \mathbb{E}_{Q_n}[\ e^{-t_0 S_n}\ \mathbb{1}(an \leq S_n \leq an + \sqrt{n})]$$
$$\geq (\varphi(t_0))^n \exp(-t_0(an + \sqrt{n}))$$
$$\times\ Q_n(an \leq S_n \leq an + \sqrt{n})$$

In addition

$$Q_n(an \leq S_n \leq an + \sqrt{n})$$
$$= Q_n\left(0 \leq \frac{S_n - an}{\hat{\sigma}\sqrt{n}} \leq \frac{1}{\hat{\sigma}}\right)$$
$$\xrightarrow[\text{CLT}]{n \to \infty} \mathbb{P}\left(0 \leq Z \leq \frac{1}{\hat{\sigma}}\right) \quad Z \sim \mathcal{N}(0,1)$$

Hence

n large enough

$$Q_n(an \leq S_n \leq an + \sqrt{n}) \geq c_1 > 0$$

$$\Rightarrow \liminf \frac{1}{n} \ln(\mathbb{P}(S_n \geq an))$$
$$\geq \varphi(t_0) - a t_0 = -I(a)$$

2 - condition on t_0

<u>Hyp</u> we have assumed

$$I(a) = a t_0 - L(t_0)$$

Is this realistic ?

<u>Main case</u> Assume $\boxed{a > \mu}$ and

$$\boxed{\mathbb{P}(X_1 < a) > 0 \quad , \quad \mathbb{P}(X_1 > a) > 0}$$

Then

$$at - L(t) = -\ln\left(\mathbb{E}\left[e^{t(X_1 - a)}\right]\right)$$

$$= -\ln \mathbb{E}\Big\{ \underbrace{e^{t(X_1-a)} \mathbb{1}_{(X_1 \le a)}}_{\substack{\downarrow\, t\to+\infty \\ \mathbb{1}_{(X_1 = a)} \text{ [dominated cvgce]}}} + \underbrace{e^{t(X_1-a)} \mathbb{1}_{(X_1 > a)}}_{\substack{\downarrow\, t\to\infty \\ +\infty \text{ [monotone cvgce]}}} \Big\}$$

Hence $\lim\limits_{t\to\infty} at - L(t) = -\infty$

Similarly $\lim\limits_{t\to-\infty} at - L(t) = -\infty$

Also L concave

$\Rightarrow$ $\boxed{\text{Maximum achieved for some } t_0 \in \mathbb{R}}$

Degenerate case 1 Assume $a > \mu$ and

$$\boxed{\mathbb{P}(X_1 \geq a) = 0}$$

Then

$$I(a) = \sup\{ at - \overbrace{\ln(\mathbb{E}[e^{tX_1}])}^{< at \text{ for } t \geq 0} ; t \geq 0 \}$$

$$= \infty$$

Hence upper bound becomes

$$\frac{1}{n} \ln \underbrace{\mathbb{P}(S_n/n \geq a)}_{\text{true, since this is } 0} \leq -\infty$$

Degenerate case 2 Assume $a > \mu$ and

$\mathbb{P}(X_1 > a) = 0$, $\mathbb{P}(X_1 = a) > 0$

Then

(i) we have

$$\frac{1}{n} \ln \mathbb{P}(S_n \geq an)$$
$$= \frac{1}{n} \ln \left(\mathbb{P}(X_1 = a)\right)^n$$
$$= \ln\left(\mathbb{P}(X_1 = a)\right) \qquad (1)$$

(ii) One computes

$$I(a) = \sup\{at - L(t);\ t \geq 0\}$$
$$= \sup\{-\ln E[e^{t(X_1 - a)}];\ t \geq 0\}$$
$$= -\inf\{\ln E[e^{t\overbrace{(X_1-a)}^{\leq 0}}];\ t \geq 0\}$$
$$= -\lim_{t\to\infty}\{\ln E[e^{t(X_1-a)}];\ t \geq 0\}$$
$$= -\ln\left(\mathbb{P}(X_1 = a)\right) \qquad (2)$$

Thus gathering (1) and (2) we get

$$\boxed{\lim_n \frac{1}{n} \ln \mathbb{P}(S_n \geq an) = -I(a)}$$

3- Example of rate function

Exponential Consider $X_1 \sim \mathcal{E}(\lambda)$. Then

$$\varphi(t) = \int_0^\infty e^{tx} \lambda e^{-\lambda x} \, dx$$

$$= \lambda \int_0^\infty e^{-(\lambda - t)x} \, dx$$

$$\boxed{\varphi(t) = \begin{cases} \dfrac{\lambda}{\lambda - t} & \text{if } t < \lambda \\ \infty & \text{if } t \geq \lambda \end{cases}}$$

Derivative of φ we have

$$\varphi'(t) = \frac{\lambda}{(\lambda - t)^2} \qquad t < \lambda$$

Computing $I(a)$

$$K(t) = at - \ln\left(\frac{\lambda}{\lambda - t}\right)$$

$$K'(t) = a - \frac{\varphi'(t)}{\varphi(t)}$$

$$= a - \frac{\lambda}{(\lambda-t)^2} \frac{(\lambda-t)}{\lambda} = a - \frac{1}{\lambda - t}$$

Thus

$$K'(t) = 0 \iff \frac{1}{\lambda - t} = a$$

$$\iff \lambda - t = \frac{1}{a} \iff t = \lambda - \frac{1}{a} \equiv t_0$$

and

$$I(a) = K(t_0)$$

$$= a\left(\lambda - \frac{1}{a}\right) - \ln\left(\frac{\lambda}{\lambda - \left(\lambda - \frac{1}{a}\right)}\right)$$

$$\boxed{I(a) = \lambda a - 1 - \ln(\lambda a)}$$