Talk about concentration bound

L Entropy (Jianing talked about this)

Bernoulli measure.

L Herbst argument

— Alexander's bound on nonrandom fluctuations. (brilliant)

Concentration about the mean

Let's recall the Chernoff trick for indep sums.

Let $S_n = X_1 + \cdots + X_n$ $\{X_i\}$ are iid

Center: replace X_i by $X_i - \mathbb{E}[X_i]$

$$\frac{S_n}{n} \xrightarrow{LLN} 0 \text{ a.s. by assuming } E[X_1] = 0$$

Assume $M(\theta) = E[e^{\theta X}] < \infty$ for $\theta \in (-\delta, \delta)$ (Moment generating functions)

$$X \sim Exp(\lambda) : M(\theta) = \int_0^\infty e^{\theta x} \lambda e^{-\lambda x} dx = \frac{\lambda}{\lambda - \theta} \quad (\theta < \lambda)$$

Markov.

$$P(X > s) \leq \frac{1}{s} E[X]$$

Then $\mathbb{P}(S_n > nt) \overset{\text{arbitrary +ve}}{=} \mathbb{P}(e^{\theta S_n} > e^{\theta n t})$

$$\leq e^{-\theta n t} (M(\theta))^n \quad \text{Markov inequality}$$

$$\mathbb{E}[e^{\theta S_n}] = \mathbb{E}[e^{\theta(X_1 + X_2 \cdots X_n)}] = \prod_{i=1}^n E[e^{\theta X_1}]$$

Optimize over θ to get

$$\mathbb{P}(S_n > t) \leq e^{-n \underbrace{\sup_\theta (\theta t - \log M(\theta))}} = e^{-I(t)}$$

$$e^{-n(\theta t - \log M(\theta))}$$

$$\boxed{I(t) = \sup_\theta (\theta t - \log M(\theta)) \quad \text{Legendre transform of } \log M(\theta)}$$

convex fn.

Ex: prove this.

If $X_i \sim N(0,1)$

$$M(\theta) = e^{\theta^2/2} \qquad I(t) = \sup_\theta (\theta t - \theta^2/2) \qquad t = \frac{\theta}{2}$$

$$\log M(\theta) = \frac{\theta^2}{2} \qquad = \frac{t^2}{2}$$

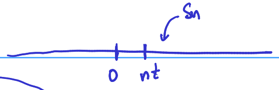
$$\Rightarrow \mathbb{P}\left(\frac{S_n}{n} > t\right) \leq e^{-n \textcircled{t^2/2}} \rightarrow \text{Gaussian concentration} \quad -(\#2b)$$

Ex: If X is Gaussian, show

$$M(\theta) = \mathbb{E}[e^{\theta X}] = e^{\theta^2/2}$$

(Complete the square in the exponent)

Called GAUSSIAN concentration.



How about other distributions? Suppose X_i has mean 0 and variance 1. Then

$$\log M(\theta) = \log \mathbb{E}[e^{\theta X}] \approx \log \mathbb{E}\left[1 + \cancel{\theta X}^{\,0} + \frac{\theta^2 X^2}{2}\right] \approx \log\left(1 + \frac{\theta^2}{2}\right)$$

$$\approx \frac{\theta^2}{2} + O(\theta^4) \quad \text{for small enough } \theta.$$

$$\log M(\theta) \leq C\theta^2 \quad -(\#2)$$

$$1 + \theta E[\cancel{X}^{\,0}] + \frac{\theta^2}{2} E[X^2]$$

$$\log(1+x) \approx x - \frac{x^2}{2} \cdots$$

Thus infact $I(t) \approx \frac{t^2}{2}$ for small t.

So one can expect the Gaussian concentration to hold in a small interval $t \leq C$

Ex: Prove (#2) under reasonable assumptions on the distribution.

Ex: What's $I(t)$ for an exponential?

$$I(t) = \sup_{\theta<\lambda} \left(\theta t - \log \frac{\lambda}{\lambda-\theta}\right)$$

(#2b) $P(S_n > nt) \leq e^{-nt^2/2}$ (Gaussian tail)

$P(S_n > nt) \leq e^{-nt}$ (Exponential tail)

$$\to P(S_n > \sqrt{n}\, t) \leq e^{-t^2/2}$$

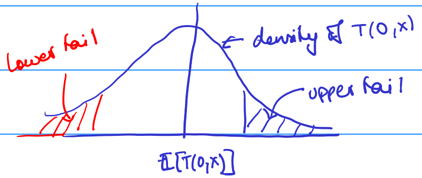
For the passage time, we establish

Theorem Let $d \geq 2$, $F(0) < p_c$ (no percolation) and $M(\alpha) < \infty$ → $\mathbb{E}\left[e^{\alpha \tau_{e(1)}}\right]$ (exponential moments) for some $\alpha > 0$. Then $\exists\ C_1\ C_2$ st

$$P\left(T(0,x) - \mathbb{E}\, T(0,x) \geq t\sqrt{|x|_1} \right) \leq e^{-C_1 t^2} \quad \forall x$$

Gaussian concentration for the passage time.

$$\text{for } t \in \left(0, C_2\sqrt{|x|_1}\right)$$

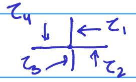


Tells you scale of fluctuation

$\chi \leq \frac{1}{2}$

$$|T - \mathbb{E}T(0,x)| \approx |x|^{1/3}$$

Called an UPPER TAIL estimate.



If $f(0) < p_c$ and $E\left(\min\{\tau_1, \dots \tau_d\}\right)^2 < \infty$ then

$$P\left(T(0,x) - \mathbb{E}\, T(0,x) \leq -t\sqrt{|x|_1} \right) \leq e^{-C_1 t^2} \quad t \geq 0$$

LOWER TAIL estimate.

Pf: Let $\log M(\theta) = \log \mathbb{E}\, e^{\theta(T - \mathbb{E}T)}$ $\quad T = T(0,x)$

$$\mathbb{P}\left(T - \mathbb{E}T \geq t\sqrt{|x|_1} \right) = \mathbb{P}\left(\underset{X}{\boxed{e^{\lambda(T-\mathbb{E}T)}}} \geq \underset{a}{e^{t\lambda\sqrt{|x|_1}}} \right) \quad \lambda > 0$$

← exponential is a monotone function

$$\underset{\substack{\uparrow \\ \text{Markov}}}{\leq} e^{-t\lambda\sqrt{|x|_1}}\, e^{\log M(\lambda)} \quad — ①$$

$$P(X > a) \leq \frac{1}{a} E[X]$$

Conc. of Meas.

Milman, Dvoretsky, Gromov.

Talagrand (1995 paper)
Isoperimetric inequalities ... in product spaces Pub. IHES.

(1994) Ann of Prob.
· Talagrand L1-L2 bound

POLL: Do you know Markov's ineq. YES or NO.

$$\boxed{\log M(\theta) \le C\theta^2}$$

$$\le e^{-\sup_\lambda \left(\lambda\sqrt{|x|_1} - C\lambda^2|x|_1\right)}$$

$\hookrightarrow \lambda = \frac{1}{2\sqrt{|x|_1}C}$

$$\mathbb{P}(\underset{\uparrow}{X} > a) = \mathbb{E}\left[1_{\{X>a\}}\right] \le \mathbb{E}\left[\frac{X}{a}\underbrace{1_{\{X>a\}}}\right] \le \mathbb{E}\left[\frac{X}{a}\right]$$

X is nonnegative

If $\log M(\theta) \le C|x|_1\theta^2 \longrightarrow \boxed{\#1}$

$$\le e^{-t\theta\sqrt{|x|_1} + C|x|_1\theta^2}$$

and simply optimize like we did for the LLN and get a Gaussian bound.

$$\left(\theta = \frac{t}{2\sqrt{|x|_1}C}\right) \to = e^{-\frac{t^2}{2C}}$$

$$\mathbb{P}\left(T - \mathbb{E}T \ge t\sqrt{|x|_1}\right) \le e^{-\frac{t^2}{2c}}$$ (Gaussian concentration)

$\hookrightarrow$ Non random fluctuations in First passage percolation.

To show this bound, we use the famous Herbst argument.

Consider

$$\frac{d}{d\lambda}\left(\frac{\log M(\lambda)}{\lambda}\right) = \frac{1\,M'(\lambda)}{\lambda M(\lambda)} - \frac{\log M(\lambda)}{\lambda^2}$$

$$= \frac{\lambda M'(\lambda) - M(\lambda)\log M(\lambda)}{\lambda^2 M(\lambda)} \qquad \lambda M'(\lambda) = \lambda\mathbb{E}\left[T e^{\lambda\hat{T}}\right]$$

$$= \frac{\mathbb{E}\left[\lambda\hat{T}e^{\lambda\hat{T}}\right] - \mathbb{E}\left[e^{\lambda\hat{T}}\right]\log\mathbb{E}\left[e^{\lambda\hat{T}}\right]}{\lambda^2 M(\lambda)} \overset{?}{\le} \boxed{C|x|_1}$$

$$M(\lambda) = \mathbb{E}\left[e^{\lambda(T-\mathbb{E}T)}\right. \qquad M'(\lambda) = \mathbb{E}\left[(T-\mathbb{E}T)e^{\lambda(T-\mathbb{E}T)}\right]$$

$$\hat{T} = T - \mathbb{E}T$$

$$\int \frac{d}{d\lambda}\left(\frac{\log M(\lambda)}{\lambda}\right) \le \int C|x|_1 \, d\lambda$$

$$\Rightarrow \frac{\log M(\lambda)}{\lambda} \le C|x|_1\lambda$$

Then $\frac{\log M(\lambda)}{\lambda} \le C|x|_1\lambda \Rightarrow \boxed{\log M(\lambda) \le C\lambda^2|x|_1}$

which is what we want.

Let us look at

$$\boxed{\mathbb{E}\left[\lambda\hat{T}e^{\lambda\hat{T}}\right] - \mathbb{E}\left[e^{\lambda\hat{T}}\right]\log\mathbb{E}\left[e^{\lambda\hat{T}}\right]} \le C|x|_1\lambda^2 M(\lambda)$$

$$\mathrm{Ent}(X) = \mathbb{E}\left[X\log X\right] - \mathbb{E}[X]\log\mathbb{E}[X]$$

On $\{1,\dots,n\}$ $\mu = \{p_1,\dots p_n\}$ $\sum p_i = 1$

$$\mathrm{Ent}(\mu) = -\left(\sum p_i\log p_i - \sum p_i \overset{0}{\log}\sum p_i\right)$$

and set $X = e^{\lambda T}$

Then this quantity is something called the ENTROPY.

$$\mathrm{Ent}(X) = \mathbb{E}[X \log X] - \mathbb{E}[X] \log \mathbb{E}[X] \qquad X \geq 0$$

This definition is the same as the usual entropy of a measure.

$\mathbb{E}[e^{\lambda \hat{T}}]$

We need to prove that

$$\boxed{\mathrm{Ent}(e^{\lambda \hat{T}}) \leq C|x|_1 \lambda^2 M(\lambda)} \longrightarrow \#1a$$

POLL: which of the following is true

A	B	C
$\mathrm{Ent}(X) \geq 0$	$\mathrm{Ent}(X) \leq 0$	Can be positive or -ve depending on X

$\mathrm{Ent}(X) \geq 0$

$\varphi(t) = t \log t$ and apply Jensen.

$\varphi'(t) = \log t + 1$

$\varphi''(t) = \frac{1}{t} \qquad t \geq 0 \Rightarrow \varphi'' \geq 0 \Rightarrow \varphi$ is convex

$$\mathbb{E}[\varphi(X)] = \mathbb{E}[X \log X] \geq \varphi(\mathbb{E}[X])$$
$$= \mathbb{E}[X] \log \mathbb{E}X.$$

$$\Rightarrow \mathrm{Ent}(e^{\lambda T}) \leq \bar{2}\, \mathbb{E}[e^{\lambda T}] C_2 \lambda^2 |x|_1$$

The entropy features in a so-called Log Sobolev inequality (discovered by Gross). It, like the Poincare inequality has a clean form for the Gaussian (and other similar continuous distribution)

Poincare : $\mathrm{Var}(f) \leq C \|Df\|_2^2$ ← Jiaming proved this

LSI : $\mathrm{Ent}(f^2) \leq C \|Df\|_2^2$ ← using the "heat equation"

$n \to \infty$ limit of these

Analytic inequalities called Sobolev inequalities

POLL: Have you heard of these?

YES or NO

PDEs or functional analysis you will learn about them on manifolds or $\mathbb{R}^n$

On $\mathbb{R}^n$ $\qquad \|f\|^{\cdots} \leq \|\nabla f\|_{,} \|f\|$

↳ Lebesgue meas.

α, β, C

(Guionnet - Zegarlinsky)

Proving the LSI is not hard. We will do this in some simple case if there is interest.

We have done this. To prove it in the case of Bernoulli measure $\Omega = \{0,1\}^n$ $\mathbb{P}^{\otimes n}$

$X \sim N(0,1)$ and

$\mathrm{Ent}(f^2) \leq C \, \| \partial f \|_2^2$

$\mathbb{E}[f] = \int f(x) \frac{e^{-\frac{x^2}{2}} dx}{\sqrt{2\pi}}$

$(X_1, \ldots, X_n)$, want to find a version of log sobolev for n iid copies,

The entropy has a nice property called "tensorization".

let $(w_{e(1)}, w_{e(2)}, \ldots)$ be an enumeration of the edge weights. on the lattice Given some function

$f(w_{e(1)}, w_{e(2)}, \ldots)$ $f: \mathbb{R}^\infty \to \mathbb{R}$ (like the passage time)

define the conditional entropy

$$\mathrm{Ent}_i(f) = \int f \log f \, dF_i - \int f dF_i \log \int f dF_i$$

= a fn of $w_{e(j)}$ $j \neq i$

$\int f(w_{e(1)}, w_{e(2)} \cdots, w_{e(i)}, \cdots) \log f \; dF_i(w(e_i))$

conditional entropy

Its still random

Then

$$\mathrm{Ent}(f) \leq \sum_{i=1}^{\infty} \mathbb{E} \, \mathrm{Ent}_i(f) = \sum_{i=1}^{\infty} \int \mathrm{Ent}_i(f) \prod_{j \neq i} dF_j$$

Tensorisation

integrate over all other variables that I had not integrated over.

Young's Inequality ⇒ Variational description of Entropy

$$\mathrm{Ent}(f) = \sup_g \left\{ \int f g \, dF : \int e^g dF \leq 1 \right\}$$

(Integration over a single variable)

$uv \leq \frac{u^p}{p} + \frac{v^q}{q}$

POLL: Have you seen Hölder proved this way?

YES OR NO

$\|x\| = \sup \{ \langle x, y \rangle : \|y\| \leq 1 \}$

The usual L^p decoupling of a product $\boxed{uv = \frac{u^p}{p} + \frac{v^q}{q}}$

Can use this to prove Hölder's inequality "log version or exponential decoupling version"

$uv \le \frac{u^p}{p} + \frac{v^q}{q}$

Prob(Young):

$$\boxed{\text{let } u > 0 \text{ then } uv \le u\log u - u + e^v}$$

Assume $\int f = 1$ wlog

$\Rightarrow \mathrm{Ent}(f) = \int f \log f - \int f \log \overset{=0}{\int f} = \int f \log f$

$fg \le f\log f - f + e^g$

$$\boxed{\mathrm{Ent}(f) = \sup\left\{ \int fg\, dF : \int e^g dF \le 1 \right\}}$$

Then $\int fg \le \int f \log f - \overset{=1}{\int f} + \overset{\le 1}{\int e^g}$, $f > 0$

$\le \int f \log f = \mathrm{Ent}(f)$ $\longrightarrow$ RHS $\le \mathrm{Ent}(f)$

So choose $g = \log f$ and this completes the proof $\longrightarrow$ $\int e^g = \int e^{\log f} = \int f = 1$

Proof of tensorization:

$\mathrm{Ent}(f) \le \sum_{i=1}^{\infty} \mathbb{E}\, \mathrm{Ent}_i(f) = \sum_{i=1}^{\infty} \int \mathrm{Ent}_i(f) \prod_{j \ne i} dF_j$

let $g_i(w_i, w_{i+1}, \dots w_n)$

$:= \log \frac{\mathbb{E}[e^g \mid w_i \dots w_n]}{\mathbb{E}[e^g \mid w_{i+1} \dots w_n]}$

↓ keeping fixed

$= \log \frac{\int e^g dF_1 \dots dF_{i-1}}{\int e^g dF_1 \dots dF_i}$

I will prove this for $f : \mathbb{R}^n \to \mathbb{R}$ and I'll let $n \to \infty$ and use some convergence theorem.

$g : \mathbb{R}^n \to \mathbb{R}$

(Do you recognize this?) $\longrightarrow$ logarithmic version of a Martingale difference

$\mathbb{E}[h \mid w_i \dots w_n] - \mathbb{E}[h \mid w_{i+1}, \dots, w_n]$

Then since $\sum_{i=1}^{n} g_i = \log e^g - \log \mathbb{E}[e^g] = \log \mathbb{E}[e^g \mid w_1 \dots w_n] - \log \mathbb{E}[e^g]$

$g = \sum_{i=1}^{n} g_i + \log \mathbb{E}[e^g]$ $\longrightarrow$ will give conditional entropy

$$\text{Ent}(f) = \sup_g \left\{ \int fg \;\middle|\; \text{st} \int e^g \le 1 \right\}$$

Using $\mathbb{E}[e^g] \le 1$ we get $\boxed{g \le \sum_{i=1}^n g_i}$

Thus returning to

$$\text{Ent}(f) = \sup_g \{ \mathbb{E}[fg] : \mathbb{E}[e^g] \le 1 \}$$

we have

$$\mathbb{E}[fg] \le \sum_{i=1}^n \mathbb{E}[f g_i]$$

$f \geq 0$

$$= \sum_{i=1}^n \int \int f g_i \, dF_i \prod_{j \ne i} dF_j$$

$\int_{\text{over } j \ne i} \int_{\text{over } i}$

$$\text{Ent}_i(f) = \sup_{g_i} \left\{ \int f g_i \;\middle|\; \text{st} \int e^{g_i} \le 1 \right\}$$

$$\le \sum_{i=1}^n \int \text{Ent}_i(f) \prod_{j \ne i} dF_j$$

$$\sup \; \mathbb{E}[fg] \le \sum_{i=1}^n \int \text{Ent}_i(f) \prod_{j \ne i} dF_j$$

Can I get equality here?

Do you "lose" a lot when applying tensorization.

Question: Is this inequality sharp for some f?

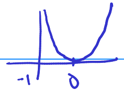
Pf of Young's inequality:

Follows from $\log(1+x) \le x \qquad -1 < x < \infty$

$\varphi(x) = x - \log(1+x)$, $\varphi(0) = 0$ $\quad \varphi'(x) = 1 - \frac{1}{1+x}$

$\varphi''(x) = \frac{1}{(1+x)^2} \geq 0 \Rightarrow \varphi$ is convex

$\varphi'(x) = 0 \quad 1+x = 1 \Rightarrow x = 0$ so its minimum value is at 0



$\boxed{\text{let } u > 0 \text{ then } uv \le u\log u - u + e^v}$

$$\text{So} \quad uv = u \log \frac{e^v}{u} + u\log u \le u\left(\frac{e^v}{u} - 1\right) + u\log u$$

$uv - u\log u + u\log u$

$$= u\log u - u + e^v$$

It's worth noting that the proof of the original Young's inequality is quite similar. So is the application to Hölder's, in fact!

Boucheron (Generalized LSI)

Lemma (Boucheron-Lugosi-Massart) Let $q(x) = x(e^x - 1)$

If X is an rv and X' indep copy then

$$\text{Ent}\, e^{tX} \leq \mathbb{E}\left[e^{tX} q(\lambda(X'-X)_+)\right]$$

X' is an independent copy of X

We will Apply this to Ent_i.

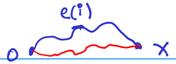
We will not prove this inequality, but it very similar in spirit to the Poincare - Efron Stein connection: the derivative in Poincare is replaced by a "difference between independent copies" in Efron - Stein.

Boucheron et al

$x(e^x - 1) \approx x^2$ $|T_i' - T|^2$ Efron-Stein

$$\mathbb{E}\,\text{Ent}_i(e^{\lambda T}) \leq \mathbb{E}\left[e^{\lambda T} q(\lambda(T_i' - T)_{\boxed{+}})\right]$$

$T_i' = T$ with just the i^{th} weight $\omega(i)$ is replaced by $\omega'(i)$



As before we only know $(T_i - T)_+ > 0$ only when $e(i)$ is in $\underline{GEO}(0,x)$ (in the original weights)

$q(\lambda \tau_e')$ does not depend on τ_e'

$$\text{Ent}\, e^{\lambda T} \leq \sum_{i=1}^{\infty} \mathbb{E}\left[e^{\lambda T} q(\lambda(T_i' - T)_+) \mathbb{1}_{\{e(i) \in \underline{GEO}(0,x)\}}\right]$$

Used discrete log Sobolev.

Poincare

$$\text{Var}(f) \leq C \|\nabla f\|_2^2$$

uses a derivative and f need not be diff. or the underlying distribution is discrete.

Efron-Stein

$$\text{Var}(f) \leq \frac{1}{2} \sum_{i=1}^{n} \mathbb{E}\left[(f - f_i)^2\right]$$

$$f(X_1 \ldots X_n), \quad f_i = f(X_1, \ldots, \hat{X}_i, \ldots X_n)$$

i^{th} variable was replaced by an independ copy. ⇒ derivative was replaced by a difference.

$$\text{Ent}(f^2) \leq C \|\nabla f\|_2^2$$

Log-Sobolev for general distributions that doesn't need a derivative.

Plan:

1) Concentration inequality
2) Alexander bound (brilliant)
3) Tuesday

POLL

Semigroup/generator/log Sobolev for Bernoulli

A

$1/3$ fluctuations exponent $\frac{1/2}{\text{log correction}}$

B

Jamik's theorem.

$$\text{Ent}\, e^{\lambda T} \leq \sum \mathbb{E}\, \text{Ent}_i\, e^{\lambda T}$$

(tensorization)

$q(x) = x(e^{x}-1)$

Notice 1) $q(x)$ is increasing

2) $T_i' - T \leq \tau'_{e(i)} - \tau_{e(i)} \leq \tau'_{e(i)}$ when $e(i) \in \underline{GEO}(0,x)$

↳ nonnegative

$$\Rightarrow q((T_i'-T)_+) \leq q(\tau'_{e(1)})$$

$$\leq \sum_{i=1}^{\infty} \mathbb{E}[\underbrace{q(\lambda\tau_e')}_{}] \, \underbrace{\mathbb{E}\left[e^{\lambda T} \mathbb{1}_{\{e(i)\in \underline{GEO}(0,x)\}}\right]}_{}$$

same for all τ_e' (iid)

move the sum in

Split into product using independences.

and sum over indicators to count $GEO(0,X)$

$$= \underbrace{\mathbb{E}[q(\lambda\tau_e')] \, \mathbb{E}[e^{\lambda T} \, |\underline{GEO}(0,x)|]}_{}$$

How to control $\mathbb{E}[XY]$?

POLL

A Hölder

B Cauchy Schwartz

C Other → Exponential version of Young's inequality

$$\mathrm{Ent}(X) = \sup\{XY \mid \mathbb{E}[e^{Y}] \leq 1\}$$

Of course we can just use $\mathbb{E}[XY] \leq \mathrm{Ent}(X)$ if $\mathbb{E}[e^{Y}] \leq 1$ and $X>0$ – (#2)

| we may not have $\mathbb{E}[e^{Y}] \leq 1$ (normalize Y)

So for any $Y+C$ satisfies $\mathbb{E}[e^{Y}]e^{C} = 1$

$$\Rightarrow C = -\log \mathbb{E}[e^{Y}]$$

and so $\mathbb{E}[X(Y+C)] \leq \mathrm{Ent}(X)$ (because $\mathbb{E}[e^{Y+C}] = 1$)

$$\Rightarrow \mathbb{E}[XY] \leq \mathrm{Ent}(X) + \mathbb{E}[X] \log \mathbb{E}[e^{Y}] \longrightarrow \mathbb{E}[XC]$$

↳ for ANY Y

$$\int ab \le \int \frac{a^p}{p} + \frac{b^p}{q}$$

Thus

$$\mathbb{E}[\underbrace{e^{\lambda T}}_{x}\underbrace{|GEO(0,x)|}_{Y}] = \mathbb{E}\left[ae^{\lambda T}\frac{|GEO(0,x)|}{a}\right]$$

its useful to have this constant

why are we introducing this constant?

$$\le a\,\mathrm{Ent}(e^{\lambda T}) + a\mathbb{E}e^{\lambda T}\log\mathbb{E}\frac{|Geo(0,x)|}{a} \quad \text{—(\#3)}$$

$$\mathrm{Ent}(ae^{\lambda T}) + \mathbb{E}[ae^{\lambda T}]\log\mathbb{E}\left[\frac{|GEO(0,x)|}{a}\right]$$

decoupled from two!

Is it an inequality or equality

$$\mathrm{Ent}(aY) = \mathbb{E}[aY\log(aY)] - \mathbb{E}(aY)\log\mathbb{E}(aY)$$

$$a\left[\mathbb{E}[\cancel{Y}(\log a + \log Y)] - \mathbb{E}[\cancel{Y}](\log a + \log\mathbb{E}Y)\right]$$

$$= a\,\mathrm{Ent}(Y)$$

POLL: Does $\mathrm{Ent}(aY) = a\,\mathrm{Ent}(Y)$?

[✓ YES] NO

$$\underline{\mathrm{Ent}(e^{\lambda T})} \le \mathbb{E}[q(\lambda\tau_e)]\;\mathbb{E}[e^{\lambda T}|GEO(0,x)|]$$

$$\mathbb{E}[q(\lambda\tau_e)]\left(a\,\underline{\mathrm{Ent}(e^{\lambda T})} + a\mathbb{E}e^{\lambda T}\log\mathbb{E}e^{\frac{|GEO(0,x)|}{a}}\right)$$

Use (#3) and obtain

$$\mathrm{Ent}(e^{\lambda T})(1 - a\mathbb{E}[q(\lambda\tau_e)])$$

$$\le a\mathbb{E}[q(\lambda\tau_e))\log\mathbb{E}e^{\frac{|Geo(0,x)|}{a}}\mathbb{E}[e^{\lambda T}] \quad \text{(\#3a)}$$

So as long as $a\,\mathbb{E}[\underbrace{q(\lambda\tau_e)}]<1$ then we can do this. $\approx \lambda^2\tau_e^2$

assuming $a\,\mathbb{E}[q(\lambda\tau_e)]$ → (#3) call it L

Which gives

$$\mathrm{Ent}(e^{\lambda T}) \le \frac{L}{1-L}\log\mathbb{E}e^{\frac{|GEO(0,x)|}{a}}\mathbb{E}[e^{\lambda T}]$$

"τ_e cannot be too big"

Lemma Assuming $\mathbb{E}[e^{\alpha\tau_e}]<\infty$ $\forall\alpha>0$ $\exists\, a, c_1>0$

$$\log\mathbb{E}e^{\frac{|Geo(0,x)|}{a}} \le c_1|x|_1$$

to make this exp moment converge

If you assume $\tau_e \in \{m,n\}$

$$|GEO(0,x)| \le \frac{|x|n}{m}$$

τ_e is "eventually bounded"

$X \sim \text{Exp}(1)$

$$\mathbb{E}[e^{\lambda X}] = \int_0^\infty e^{\lambda t} e^{-t} dt$$

$$= \int_0^\infty e^{-(1-\lambda)t} dt$$

$\lambda < 1$

So this determines a. ← fixes our parameter a

Now we are at

$$\text{End}\, e^{\lambda T} \le \frac{L}{\underbrace{1 - L}_{\le 2}} \mathbb{E}[e^{\lambda T}]\, C_1 |x|_1$$

$$L = \mathbb{E}[q(\lambda\tau_e')] \approx \lambda^2 \mathbb{E}[\tau_e'^2] << 1$$

We will choose λ so small st $L = a\,\mathbb{E}[q(\lambda\tau_e)] < \frac{1}{2}$

Since $\mathbb{E}\left[\frac{q(\lambda\tau_e)}{\lambda^2}\right] \longrightarrow \mathbb{E}[\tau_e^2]$ → Exercise using dominated convergence.

So

$$\text{End}(e^{\lambda T}) \le \bar{2}\, \mathbb{E}[e^{\lambda T}]\, C_2 \lambda^2 |x|_1$$

This is what we wanted to prove in (#1a)