Lecture 5

Revision of large-deviations.

$$\boxed{P\left(\frac{M_n}{n^\alpha} = m\right) \asymp e^{-n^\beta \phi(m)}}$$

Important: (1) for cases where rare events are important.

(2) In physics, $\phi(m)$ is a natural extension of the idea of ~~Bs~~ Landau free energy out side equilibrium.

(3) Many stat mech symmetry relations, fluctuation theorems, are stated in terms of large-deviations.

Ref: Review article of Hugo Touchett in Phys. Reports.

Extreme value statistics :

Ref.: Sanjib Sabhapandit
Arxiv. 1907.00944.
Extremes & Records

①

Another important example of limit distributions.

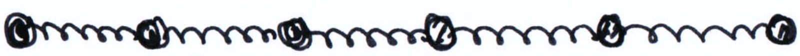
Q. Let $\{x_1, \cdots, x_N\}$ be random variables. Then what is the distribution of $x_{max} = \max\{x_1, \cdots, x_N\}$? (OR x_{min})

Examples: x_i's could be temperature at a place, magnitude of earthquakes, water level in a river, stock prices, eigenvalues of a random matrix, etc.

These are important in the context where such extreme values have drastic consequences.

Ref: Fisher & Tippett (1928).

Ex. 1.

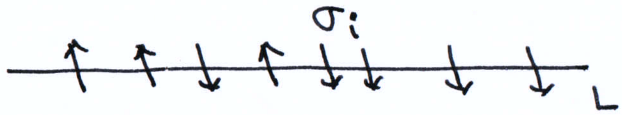


strength of each link is Random.

Then, what is the mimum force required to break the chain.

Ex 2. Random energy model. [Derrida, 1981]

very important for physics of glasses, replica symmetry.



There are $N = 2^L$ spin configs C.

Each config is assigned an energy E at random.

Then, what is the free energy

$$F(\beta) = -\frac{1}{\beta} \log \sum_{i=1}^{N} e^{-\beta E_i}$$

rugged energy landscape.

① at $\beta \to 0$ (high temp)

$$\sum_{i=1}^{N} e^{-\beta E_i} \simeq \sum_{i=1}^{N} (1 - \beta E_i) \longrightarrow N(1 - \beta \langle E \rangle)$$

$$\Rightarrow F(\beta) \simeq -\frac{\ln N}{\beta} + \langle E \rangle = -\frac{L \ln 2}{\beta} + \langle E \rangle$$

disorder not important.

(2) for $\beta \to \infty$ (low temp)

$$\sum_{i=1}^{N} e^{-\beta E_i} \simeq e^{-\beta \min\{E_i\}}$$

$\Rightarrow$ free energy $\boxed{F(\beta) = E_{min}}$ Ground state.

General question: Similar to central limit theorem, is there an asymptotic distribution for Prob (x_{max}) ?
(or x_{min})

Remark: unlike $\sum_i x_i$ in CLT, x_{max} is a highly non-linear function of $\{x_1 \cdots x_N\}$.

First, simple examples

for iid $\{x_1 \cdots\cdots x_N\}$ drawn from $p(x)$.

$$\text{Prob}\{\max\{x_1, \cdots, x_N\} < x) = \text{Prob}\left(x_1 < x \ \& \ x_2 < x \cdots\cdots \& \ x_N < x\right)$$

$$= \text{Prob}(x_1 < x)\,\text{Prob}(x_2 < x) \cdots \text{Prob}(x_N < x)$$

$$\Rightarrow \boxed{F^{(N)}(x) = \left(\mathcal{F}(x)\right)^N}$$

$$\mathcal{F}(x) = \int^{x} dy\, p(y) \equiv \text{cumulative Probability.}$$

$$\Rightarrow \text{Prob}(x_{max} = x) := P^{(N)}_{max}(x) = \partial_x F^{(N)}(x)$$

$$\Rightarrow \boxed{P^{(N)}_{max}(x) = N \cdot \left(\mathcal{F}(x)\right)' \left(\mathcal{F}(x)\right)^{N-1}}$$

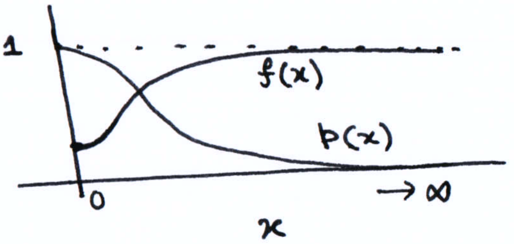
Is there a simple limiting formula for $P_{max}^{(N)}(x)$ if we shift + rescale properly? (OR $F^{(N)}(x)$)

Generalized CLT	Extreme value
• $\text{Prob}\left(\dfrac{\sum_i^N x_i - b_N}{a_N} = y\right) \xrightarrow[N\to\infty]{} p^*(y)$	• $\text{Prob}\left(\dfrac{\max\{x_1 \cdots x_N\} - b_N}{a_N} = y\right) \xrightarrow[N\to\infty]{} p^*(y)$
• Analyze using characteristic fn $G^{(N)}(k) = \left(g(k)\right)^N$	• Analyze using cumulative prob. $F^{(N)}(x) = \left(f(x)\right)^N$
• tails of $p(x)$ decides $p^*(y)$ Either Gaussian OR Lévy stable distributions by parameter α.	• tails of $p(x)$ decides $p^*(y)$ A single parameter family of distributions [See soon]

Specific cases:

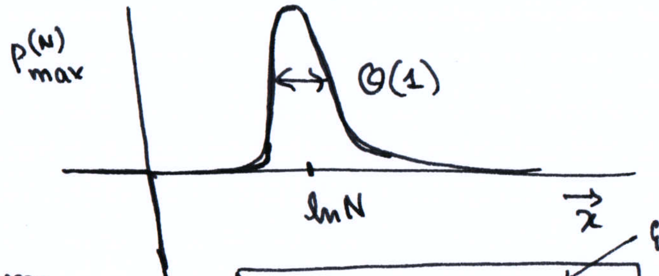
Case 1 $p(x) = e^{-x}$ in domain $x \in [0, \infty]$

$$\Rightarrow f(x) = 1 - e^{-x}$$



$$\Rightarrow F^{(N)}(x) = \left(f(x)\right)^N = \left(1 - e^{-x}\right)^N$$

$$\Rightarrow p^{(N)}_{max}(x) = N \cdot f'(x) \cdot \left(f(x)\right)^{N-1} = N\, e^{-x + (N-1)\ln(1-e^{-x})}$$



Then, we expect that the fluctuation around maximum will have limiting distribution.

$$\boxed{\begin{aligned} \langle x \rangle &\simeq \ln N + \gamma_E + \cdots \\ \langle x^2 \rangle_c &\simeq \frac{\pi^2}{6} + \cdots \end{aligned}}$$

Euler Gamma.

$$y = x - \ln N$$

up to a numerical (normalization) factor

$$\Rightarrow p^{(N)}_{max}(y) = p^{(N)}_{max}(x = y + \ln N) \xrightarrow[N \to \infty]{} e^{-y - e^{-y}} = p^{*}_{max}(y)$$

[Gumbel distribution]

$$\left[\text{Equivalently} \quad F^{*}(y) = \int_0^y dx\, p^{*}_{max}(x) = e^{-e^{-y}} \right]$$

Remark: why do we care about limiting distribution?

1. Universality. Many different distributions have same large N asymptotics.

2. For large N, they give a good enough simple description.

$$\text{e.g.} \quad \tilde{p}^{(N)}_{max}(x) = p^{*}_{max}(x - \ln N) = N e^{-x - N e^{-x}}$$

fits well the original distr. $p^{(N)}_{max}(x)$.

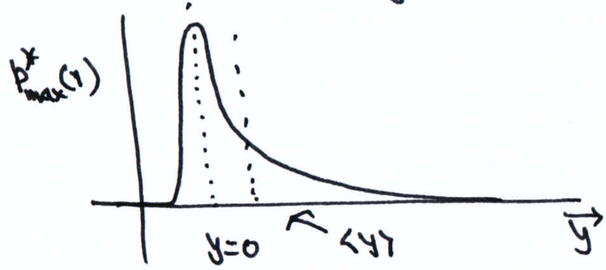
Plot $p^{(N)}_{max}$ and $\tilde{p}^{(N)}_{max}$ to compare

~~Remarks: for the distribution~~

Remark: for the Gumbel distribution $p^*_{max}(y) = \frac{1}{Z_1} e^{-y-e^{-y}}$

mean : $\langle y \rangle \simeq 1.2602$

although the distribution has maximum at $y=0$.



(important)

Remark: The limiting distribution $p^*(y)$ is NOT unique, it depends on the "shift" and "rescaling" chosen for "coarse-graining".

let's say

$$P^{(N)}_{max}\left(\frac{x - \ln N - b_1}{a_1} = y_1\right) \xrightarrow[N\to\infty]{} p^*_1(y_1)$$

$$P^{(N)}_{max}\left(\frac{x - \ln N - b_2}{a_2} = y_2\right) \xrightarrow[N\to\infty]{} p^*_2(y_2)$$

using the relation

$$y_2 = \frac{y_1 a_1 - b_2 - b_1}{a_2^2}$$

we get

$$p^*_1(y_1) = \frac{a_1}{a_2}\, p^*_2\left(\frac{y_1 a_1 - b_2 - b_1}{a_2^2}\right)$$

This means, to get a unique $p^*(y)$ we need to fix a, b first.
Of course, different limiting distributions are easily connected.

[This point will be important for RG analysis]

(functional) RG flow for extreme value statistics

<u>Ref:</u> Eric Bertin & Géza Györgyi, 2010, J stat mech.

$s=0$, $n=2^0$: $\quad x_1 \quad x_2 \quad x_3 \quad x_4 \quad x_5 \quad x_6 \quad x_7 \quad x_8 \qquad p(x)$

$s=1$, $n=2^1$: $\quad y_1^{(1)} = \dfrac{x_{max}^{(1)} - b(1)}{a(1)} \quad y_2^{(1)} \quad y_3^{(1)} \quad y_4^{(1)} \qquad \tilde{p}^{(1)}(y)$

$s=2$, $n=2^2$: $\quad y_1^{(2)} = \dfrac{x_{max}^{(2)} - b(2)}{a(2)} \quad y_2^{(2)} \qquad \tilde{p}^{(2)}(y)$

$s=3$, $n=2^3$: $\quad y_1^{(3)} = \dfrac{x_{max}^{(3)} - b(3)}{a(3)} \qquad \tilde{p}^{(3)}(y)$

$$\tilde{p}^{(3)}(y) \xrightarrow{s \to \infty} \tilde{p}^{*}(y)$$

<u>Note</u>: ① Here $x_{max}^{(s)} = \max\{x_1, \dots, x_{2^s}\}$

Subtle point. ② note the difference of coarse-graining done here

$$y_1^{(s)} = \frac{x_{max}^{(s)} - b(s)}{a(s)}$$

compared to what we did for CLT. In that case, we would have

$$\hat{y}^{(s)} = \frac{\max\{y_1^{(s-1)}, y_2^{(s-1)}\} - \hat{b}(s)}{\hat{a}(s)}$$

Convince yourself that they are equivalent with $b(s), a(s)$ related to $\hat{b}(n), \hat{a}(n)$ for $n \le s$.

If $P^{(s)}_{max}(x)$ is the probability of $\max\{x_1, \cdots, x_{2^s}\} = x$

then
$$\tilde{p}^{(s)}(y) = a(s)\, P^{(s)}\left(x = a(s)y + b(s)\right)$$
$\uparrow$ from normalization.

$$\left[\ \frac{x - b(s)}{a(s)} = y\ \right]$$

We shall do calculation using cumulative probability. (just as we use characteristic fn for CLT)
Easy to see that their cumulative probability

$$f^{(s)}(y) = F^{(s)}\left(a(s)y + b(s)\right)$$

On the otherhand, we know that

$$F^{(s)}(x) = \left[f^{(0)}(x)\right]^{2^s}$$

$$\left[\text{we denote, } f^{(0)}(x) = \int^x dy\, p(x)\right]$$

Means,
$$\boxed{f^{(s)}(y) = \left[f^{(0)}\left(a(s)y + b(s)\right)\right]^{2^s}}$$

Let's denote $f^{(0)}(x) \equiv f(x)$

This, when demanded that there is a unique asymptotic ~~start~~ $f^*(y)$ exist, gives the RG flow and also determines $f^*(y)$.

How? see next page.

[Remind yourself that $\tilde{p}^*(y) = \partial_y f^*(y)$]

RG algebra: a typical convention is to define $\underline{\underline{n = e^s}}$ (rather than 2^s).

Step 1: Define

$$g^{(s)}(y) = -\log\left(-\log f^{(s)}(y)\right)$$

$$\left[\text{means } f^{(s)}(y) = e^{-e^{-g^{(s)}(y)}}\right]$$

Then,

$$f^{(s)}(y) = \left[f^{(0)}\left(a(s)y + b(s)\right)\right]^{e^s}$$

gives

$$\boxed{g^{(s)}(y) = g^{(0)}\left(a(s)y + b(s)\right) - s}$$

Moreover $a(0) = 1$, $b(0) = 0$ [see the RG flow chart earlier]

subtle point

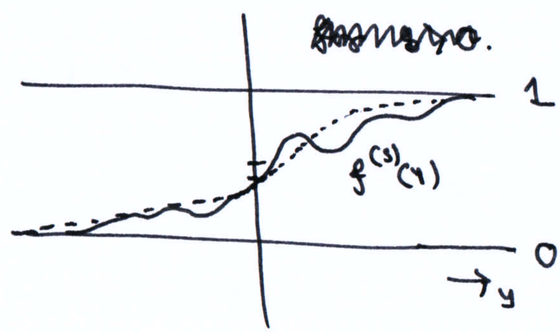
Step 2: ~~[illegible]~~
there are two free parameters, one for shift, one for rescaling. [Recall the remark on page 5]
we fix this by two suitable conditions

<u>For $s>0$:</u>

$$f^{(s)}(0) = 1/e$$

and

$$\partial_y f^{(s)}(0) = 1/e$$



$$\Rightarrow \boxed{g^{(s)}(0) = 0 \text{ and } \partial_y g^{(s)}(0) = 1}$$

$$\Rightarrow \boxed{g^{(0)}\left(b(s)\right) = s} \quad \text{and} \quad \boxed{g'^{(0)}\left(b(s)\right) = \frac{1}{a(s)}}$$

~~show that this gives a(s) and b(s) starting with a(0)=1, b(0)=0~~

Step 3: RG flow of $a(s)$ and $b(s)$.

$$g^{(0)}(b) = s \Rightarrow \dot{b}(s)\cdot g'^{(0)}(b(s)) = 1 \Rightarrow \boxed{\dot{b}(s) = \frac{1}{g'^{(0)}(b(s))}}$$

and $a(s)\, g'^{(0)}(b(s)) = 1$ means $\boxed{\dot{b}(s) = a(s)}$

$$\dot{a}(s) = -\frac{g''^{(0)}(b(s))}{\left(g'^{(0)}(b(s))\right)^2}\cdot \dot{b}(s)$$

$$\Rightarrow \boxed{\frac{\dot{a}(s)}{a(s)} = -\frac{g''^{(0)}(b(s))}{\left(g'^{(0)}(b(s))\right)^2} = \gamma(s)} \quad \text{(say)}$$

Step 4: RG flow of the distribution function.

$$g^{(s)}(y) = g^{(0)}\big(a(s)y + b(s)\big) - s$$

$$\Rightarrow \frac{d}{ds} g^{(s)}(y) = g'^{(0)}\big(a(s)y + b(s)\big)\big(\dot{a}(s)y + \dot{b}(s)\big) - 1$$

and

$$\frac{d}{dy} g^{(s)}(y) = g'^{(0)}\big(a(s)y + b(s)\big)\cdot a(s)$$

Together,

$$\frac{d}{ds} g^{(s)}(y) = \frac{d}{dy} g^{(s)}(y)\cdot \underbrace{\left(\frac{\dot{a}(s)}{a(s)}y + \frac{\dot{b}(s)}{a(s)}\right)}_{\gamma(s)y+1} - 1$$

$$\Rightarrow \boxed{\frac{d}{ds} g^{(s)}(y) = \big(1 + \gamma(s)\, y\big)\cdot \frac{d}{dy} g^{(s)}(y) - 1}$$

Step 5: Fixed point

$$\frac{d}{ds} g^*(y) = 0 = \left(1+\gamma^* y\right)\frac{d}{dy} g^*(y) - 1$$

$$\Rightarrow \quad \frac{d}{dy} g^*(y) = \frac{1}{1+\gamma^* \cdot y}$$

$$\Rightarrow \quad g^*(y) = \frac{1}{\gamma^*} \log\left(1+\gamma^* y\right)$$

[using $g^*(0)=0$ condition we chose]

Gives the limiting cumulative distribution

$$\boxed{f^*(y) = e^{-\left(1+\gamma^* \cdot y\right)^{-\frac{1}{\gamma^*}}}}$$

and

$$p^*(y) = \frac{d}{dy} f^*(y) = \frac{e^{-(1+\gamma^* y)^{-1/\gamma^*}}}{\left(1+\gamma^* y\right)^{1+\frac{1}{\gamma^*}}}$$

Remark: how do we get γ^*?

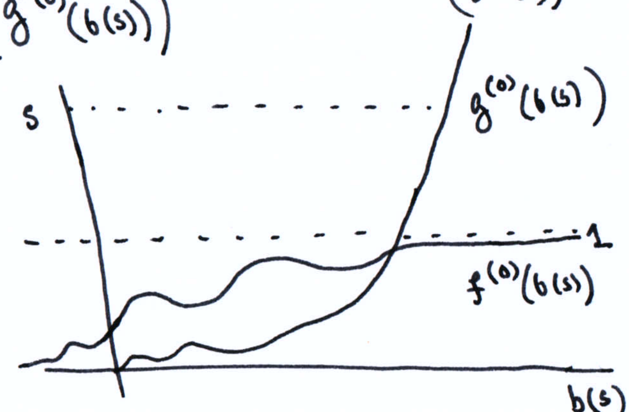
For example: $\gamma^* = -\lim\limits_{s\to\infty} \dfrac{g''^{(0)}(b(s))}{\left(g'^{(0)}(b(s))\right)^2} = -\lim\limits_{y\to\infty} \dfrac{g''^{(0)}(y)}{\left(g'^{(0)}(y)\right)^2}$

given the condition

$$g^{(0)}(b(s)) = s$$

and

$$f^{(0)}(y) = e^{-e^{-g^{(0)}(y)}} \quad \Rightarrow$$



Explicit examples

Example 1 :

$$p(x) = e^{-x} \quad \text{for} \quad x \in [0, \infty]$$

$$\Rightarrow f^{(0)}(x) = 1 - e^{-x}$$

$$\Rightarrow g^{(0)}(x) = -\log\left(-\log\left(1 - e^{-x}\right)\right)$$

$$\simeq x \quad \text{for large } x.$$

means

$$\gamma^* = -\lim_{x\to\infty} \frac{g''^{(0)}(x)}{\left(g'^{(0)}(x)\right)^2} = 0 \Rightarrow \boxed{\gamma^* = 0}$$

$$\boxed{\Rightarrow f^*(y) = \lim_{\gamma^*\to 0} e^{-(1+\gamma^* y)^{-1/\gamma^*}} = e^{-e^{-y}}}$$

and $b(s)$ for large s comes from

$$g^{(0)}\left(b(s)\right) = s \Rightarrow b^*(s) = s = \log n \leftarrow [n = e^s]$$

$$\Rightarrow \boxed{b^*(s) = \log n}$$

and

$$a(s) = \dot{b}(s) \Rightarrow \boxed{a^*(s) = 1}$$

This agrees with our exact analysis on page 4, that

$$P_{max}\left(\frac{x - \log n}{1} = y\right) \longrightarrow e^{-y - e^{-y}}$$

and corresponding $f^*(y) = e^{-e^{-y}}$

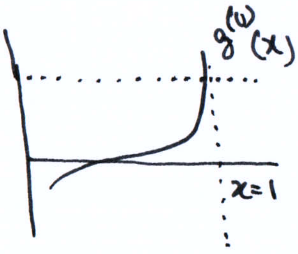
Gumbel distribution.

Example 2: uniform distribution

$$p(x) = 1 \quad \text{for} \quad x \in [0,1] \quad (\text{bounded domain})$$

$$\Rightarrow f^{(0)}(x) = x$$

Gives $$g^{(0)}(x) = -\log(-\log x) = -\log\log\frac{1}{x}$$



leads to

$$\gamma^* = -\lim_{x \to 1} \frac{g''^{(0)}(x)}{g'^{(0)}(x)^2} = -1$$

also

$$g^{(0)}(b(s)) = s \Rightarrow \quad b(s) = e^{-e^{-s}} = e^{-1/n} \simeq 1 - \frac{1}{n}$$

and $$a(s) = \dot{b}(s) = \frac{1}{n} e^{-1/n} \simeq \frac{1}{n}$$

And the limiting distribution

$$f^*(y) = e^{-(1-y)}$$

This means,

$$P_{max}\left(\frac{x-1}{1/n} = y\right) \longrightarrow e^{-(1-y)}$$

Example 3: Pareto distribution (power-law tail)

$$p(x) \simeq \frac{\alpha}{x^{1+\alpha}}$$

for x large and $x \in [0,\infty]$

$$\Rightarrow f^{(0)}(x) = 1 - \frac{1}{x^{\alpha}}$$

Gives $g^{(0)}(x) \simeq \alpha \log x$ for large x.

Gives $\gamma^* = \frac{1}{\alpha}$, $b(s) = e^{s/\alpha} = n^{1/\alpha}$

$$a(s) = \frac{1}{\alpha} n^{1/\alpha}$$

limiting distribution: $f^*(y) = e^{-\left(1+\frac{y}{\alpha}\right)^{-1/\alpha}}$

means,

$$P_{max}\left(\frac{x_{max} - n^{1/\alpha}}{\frac{1}{\alpha} n^{1/\alpha}} = y\right) = e^{-\left(1+\frac{y}{\alpha}\right)^{-1/\alpha}}$$

Remark: The above means

$$x_{max} \simeq n^{1/\alpha}.$$

Our earlier CLT/Levy analysis showed that, for $\alpha < 1$

$$\sum_{i=1}^{n} x_i \simeq n^{1/\alpha}$$

This means, for $\alpha < 1$, sum of random variable is of same order as x_{max}, thus the sum is dominated by single events!

Example 4. Gaussian

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{-\frac{x^2}{2\sigma^2}} \quad \text{for } -\infty < x < \infty$$

$$f^{(0)}(x) = 1 - \sqrt{\frac{\sigma^2}{2\pi}} \cdot \frac{e^{-x^2/2\sigma^2}}{x} \quad \text{for large } x.$$

$$\Rightarrow g^{(0)}(x) \simeq -\log\left(+ \sqrt{\frac{\sigma^2}{2\pi}}\, \frac{e^{-x^2/2\sigma^2}}{x} \right)$$

$$\simeq \frac{x^2}{2\sigma^2} + \log x + \frac{1}{2}\log\frac{2\pi}{\sigma^2} \quad \text{for } x \text{ large.}$$

Gives

$$\gamma^* \simeq -\frac{g''^{(0)}(x)}{\left(g'^{(0)}(x)\right)^2} \simeq -\frac{1}{x^2} \longrightarrow 0 \quad \text{for } x \to \infty$$

$$\Rightarrow \boxed{\gamma^* = 0} \Rightarrow f^*(y) = e^{-e^{-y}}$$

$$p^*(y) = e^{-y-e^{-y}}$$

Gumbel distr.

And $b(s) \simeq \sqrt{2\sigma^2 s} \Rightarrow b^* = \sqrt{2\sigma^2 \ln n}$

and $a^* = \sqrt{\frac{\sigma^2}{2s}} \simeq \sqrt{\frac{\sigma^2}{2\ln n}}$

All these mean

$$P_{max}\left(\frac{x_{max} - \sqrt{2\sigma^2 \ln n}}{\sqrt{\frac{\sigma^2}{2\ln n}}} = y \right) \longrightarrow e^{-y-e^{-y}}$$

Remark: for x_{min} this sign becomes (+).

Mathematical statement about the limit theorem:

Ref. ~~[illegible]~~ Fisher & Tippet, 1928

Gnedenko, 1943, Ann Math 44, 423.

Suppose there is a sequence of numbers a_n, b_n with $a_n > 0$ such that

$$y^{(n)} = \frac{x^{(n)}_{max} - b_n}{a_n}$$

has a non-degenerate distribution for $n \to \infty$, ie the limit

$$\lim_{n\to\infty} \left[f^{(0)} (a_n y + b_n) \right]^n = f^*(y)$$

exist and non-degenerate. Then, $f^*(y)$ belongs to the one-parameter family of function

$$f^*(y) = e^{-(1+\gamma^* \cdot y)^{-\frac{1}{\gamma^*}}} \quad \text{with } (1+\gamma^* \cdot y) > 0$$

and $\gamma^* \in$ Real.

Three well known ~~[illegible]~~ class.

① $\gamma^* > 0$. Fréchet distribution: when x is on an unbounded domain with $f^{(0)}(x)$ having a power-law tail.

Ex: $p(x) \approx \frac{1}{x^{1+\alpha}}$

② $\gamma^* = 0$. Gumbel distr.: x unbounded, but $f^{(0)}(x)$ decay faster that power-law.

Ex: $p(x) = e^{-x}$

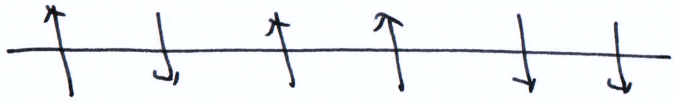
~~x is bounded domain with smooth p(x).~~

③ $\gamma^* < 0$. Weibull distr.: x in a bounded domain, with power-law behavior at boundary for $f^{(0)}(x)$.

Ex: $p(x) = 1$

A physics example: Random energy model.

Derrida, 1981.



L-spins.

total $n = 2^L$ configurations

Energy of each configuration is chosen Randomly from distr.

$$p(E) = \frac{1}{\sqrt{2\pi L}} e^{-\frac{E^2}{2L}}$$

because energy E of each config is extensive.

Our extreme value analysis (Example 4) shows that

note the sign.

$$p_{min}\left(\frac{E_{min} + \sqrt{2L\ln n}}{\sqrt{\frac{L}{2\ln n}}} = y\right) \to e^{-y-e^{-y}}$$

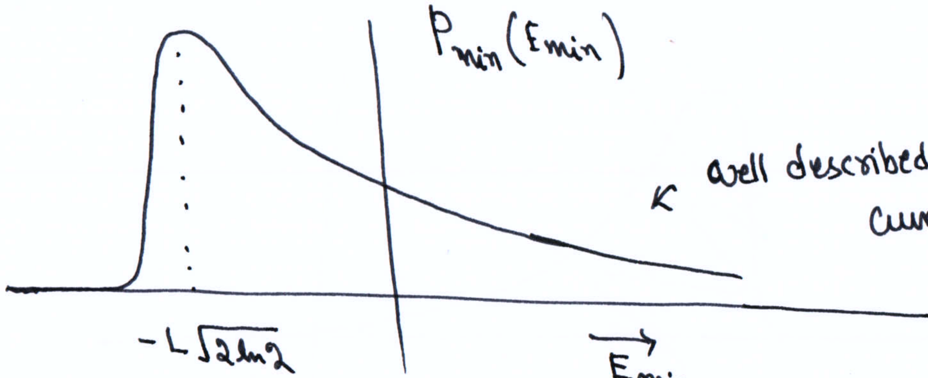
Equivalently

$$p_{min}\left(E_{min} = -\sqrt{2L\ln n} + y\cdot\sqrt{\frac{L}{2\ln n}}\right)$$

$$= p_{min}\left(E_{min} = -L\sqrt{2\ln 2} + y\sqrt{\frac{1}{2\ln 2}}\right)$$

$$\simeq e^{-y-e^{-y}}.$$

Means



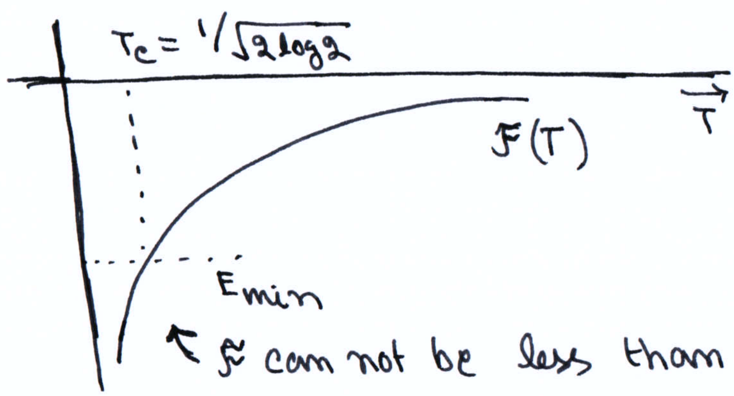
Gives average $\boxed{\langle E_{min}\rangle \simeq -L\sqrt{2\ln 2}}$

Free energy (Exact)

$$\tilde{F}(\beta) = -\frac{1}{\beta}\log\left\langle \underbrace{\sum_{i=1}^{2^L} e^{-\beta E_i}}_{\to\, e^{L\left(\frac{\beta^2}{2}+\log 2\right)}} \right\rangle$$

$$\Rightarrow \tilde{F}(\beta) = -L\frac{\log 2}{\beta} - L\frac{\beta}{2}$$

$$= L\left(-T\log 2 - \frac{1}{2T}\right) \quad \left(\beta = \frac{1}{T} \quad k_B = 1\right)$$



$\tilde{F}$ can not be less than Emin!

System freezes at Tc: you can see the by calculating entropy

$$S = -\frac{\partial}{\partial T}\tilde{F}(T) = L\left(\log 2 - \frac{1}{2T^2}\right) = 0 \quad \text{at } T = T_c.$$

~~Q: What is the $\tilde{F}$ below Tc? Good free energy is Quenched~~ $\cancel{F = -\frac{1}{\beta}\langle \log \sum e^{-\beta E_i}\rangle}$.
~~Gives $F \to E_{min}$ at $T \to 0$.~~

Remark: zero entropy means that system freezes into it's ground state E_{min}.

Remark: what is the $\mathcal{F}$ below T_c?

A well behaved free energy is Quenched $\mathcal{F}^{(Q)}$.

$$\mathcal{F}^{(Q)} = -\frac{1}{\beta}\left\langle \log \sum_{i=1}^{2^L} e^{-\beta E_i} \right\rangle$$

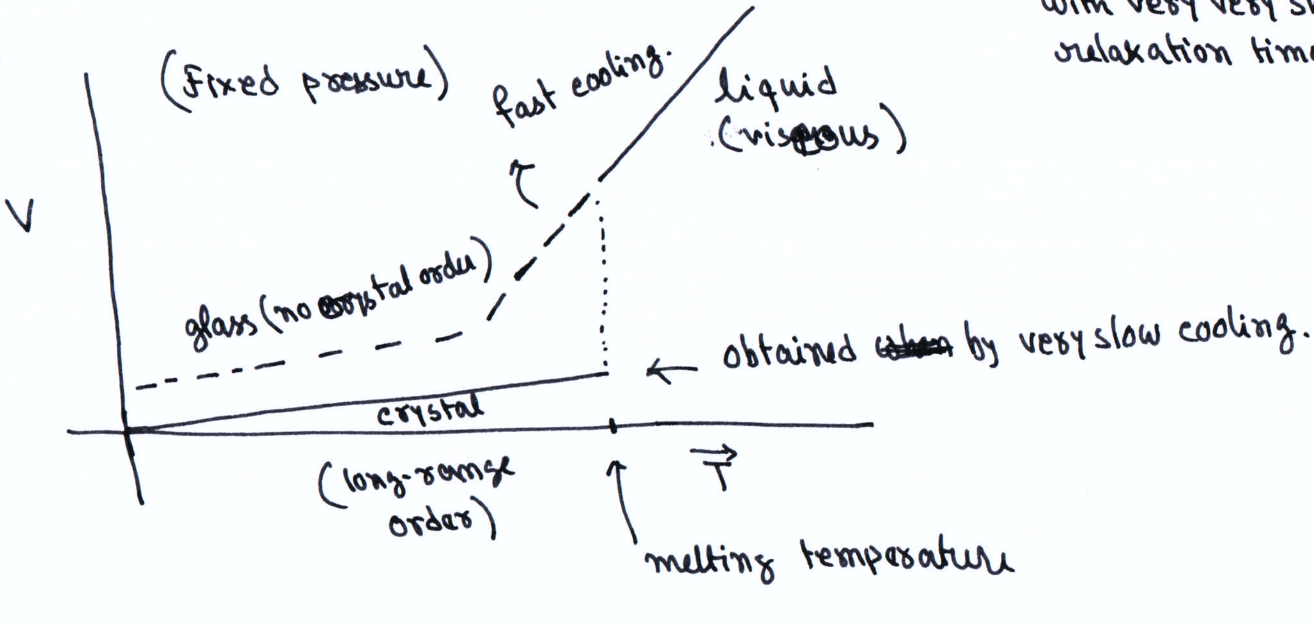
(note, in conventional $\mathcal{F}$, average is taken inside log. This is known a Annealed average)

for $T \to 0$ (or $\beta \to \infty$)

$$\boxed{\mathcal{F}^{(Q)} \simeq E_{min}}$$ (shown earlier).

This Annealed/Quenched average is a crucial concept in theory of disordered systems, e.g. glasses.

What are glasses: (loosely speaking) amorphous solid with very very slow relaxation time.



Easy read: Debenedetti & Stillinger, Nature, 410 (2001).