and $S_N = \sum_{i=1}^{N} X_i$. Then clearly

$$P_k := \text{Prob}[S_N = k] = p^k (1-p)^{N-k} \binom{N}{k}$$

whereas the CLT provides the Gaussian approximation

$$P_k^G = \frac{1}{\sqrt{2\pi\, p(1-p)}} \exp\left(- \frac{(k-pN)^2}{2p(1-p)N} \right)$$

For the "extreme event" $k = N$ this implies

$$\left.\begin{aligned} P_N &= p^N = \exp(-N \ln(1/p)) \\ P_N^G &= \frac{1}{\sqrt{2\pi\, p(1-p)}} \exp\left(- \frac{(1-p)^2 N^2}{2p(1-p)N}\right) \sim \exp\left(- \frac{(1-p)}{2p} N\right) \end{aligned}\right\}$$

Because in general $\frac{1-p}{2p} \neq \ln(1/p)$, the relative error of P_N^G is exponentially large in N.

$$\left.\begin{aligned} &\underline{p \to 0:} \quad \ln\left(\frac{1}{p}\right) \ll \frac{1-p}{2p}, \quad P_N^G \ll P_N \\ &\underline{p \to 1:} \quad \ln\left(\frac{1}{p}\right) > \frac{1-p}{2p}, \quad P_N^G \gg P_N \end{aligned}\right\}$$

$\Rightarrow$ CLT describes typical fluctuations within $O(N^{1/2})$ around the mean $\mu_1 N$, but fails for large deviations $O(N)$.

The large deviations theorem (LDT) provides an improved description of such large fluctuations.
It states that in general, for a sum

$$S_N = \sum_{i=1}^{N} X_i, \quad X_i \text{ i.i.d. RV}$$

$$\text{Prob}[S_N \simeq Ns] \sim e^{N g(s)}, \quad N \to \infty$$

where the limit $N \to \infty$ is taken at fixed
$s = S_N / N$

Here $g(s) \leq 0$ is called the rate function, with the property $g(\mu_1) = 0$.
To compute $g(s)$ we assume that the "partition function"

$$Z(\beta) = \int dx\, f(x)\, e^{-\beta x} = G_X(i\beta)$$

exists. Then clearly

$$\langle e^{-\beta S_N} \rangle = \langle e^{-\beta \sum_{i=1}^{N} X_i} \rangle = Z(\beta)^N$$

On the other hand $\mathrm{Prob}[S_N = S] \sim e^{N g(S/N)}$

$$\Rightarrow \langle e^{-\beta S_N} \rangle \sim \int ds\, e^{-\beta N s}\, e^{N g(s)} =$$

$$= \int ds\, e^{N(g(s) - \beta s)}$$

For $N \to \infty$ the integral is dominated by the maximum of $g(s) - \beta s$, $s^*(\beta)$:

$$\frac{d}{ds}(g(s) - \beta s) = 0 \Rightarrow g'(s^*) = \beta$$

$$\Rightarrow \langle e^{-\beta S_N} \rangle \sim e^{N(g(s^*) - \beta s^*)} \sim Z(\beta)^N$$

$$\Rightarrow \left.\begin{array}{ll} \text{(i)} & g(s) = \ln Z(\beta(s)) + \beta(s)\, s \\ \text{(ii)} & \beta(s) = g'(s) \end{array}\right\} \text{Legendre transform}$$

To compute $\beta(s)$ we take the derivative of (i):

$$\frac{d}{ds} g(s) = \frac{d \ln Z}{d\beta} \frac{d\beta}{ds} + \frac{d\beta}{ds} s + \beta \overset{!}{=} \beta$$

$$\Rightarrow \quad s(\beta) = - \frac{d \ln Z}{d\beta} \qquad \text{(iii)}$$

Application 1: Binomial distribution

$$Z(\beta) = 1 - p + p e^{-\beta}$$

$$\Rightarrow s(\beta) = - \frac{d}{d\beta} \ln \left(1 - p + p e^{-\beta}\right) = \frac{p e^{-\beta}}{1 - p + p e^{-\beta}}$$

$s \in [0,1]$ with $s(-\infty) = 1$, $s(\infty) = 0$, $s(0) = p$

The inversion yields

$$\beta(s) = \ln \left(\frac{s(1-p)}{p(1-s)} \right)$$

$$\Rightarrow \underline{g(s) = s \ln p + (1-s) \ln (1-p) - s \ln s - (1-s) \ln (1-s)}$$

as can also be obtained using the Stirling approximation.

<u>Special points:</u>

$$\left.\begin{aligned} g(p) &= 0 \quad \Rightarrow \quad \mu_1 = p \\ g(1) &= \ln p \\ g(0) &= \ln(1-p) \end{aligned}\right\}$$

in agreement with the exact binomial distribution to exponential accuracy.

Application 2: Limits of validity of the CLT

The CLT is recovered from the LDT by expanding to second order around μ_1:

$$g(s) \simeq -\frac{1}{2} g''(\mu_1)(s-\mu_1)^2 + \frac{1}{6} g'''(\mu_1)(s-\mu_1)^3 + \dots$$

$$\Rightarrow \text{Prob}[S_N \simeq S] = \exp(N g(S/N)) \cong$$

$$\cong \exp\left[\frac{1}{2} N g''(\mu_1)\left(\frac{S}{N} - \mu_1\right)^2\right] =$$

$$= \exp\left[\frac{1}{2} g''(\mu_1) \frac{(S - N\mu_1)^2}{N}\right] = e^{-\frac{(S-N\mu_1)^2}{2\sigma^2 N}}$$

with the identification $\sigma^2 = -1/g''(\mu_1)$

The higher order terms = the expansion quantify corrections to the CLT. In general we have

$$\frac{\mathrm{Prob}[S_N \simeq S]}{\mathrm{Gauss}[S_N \simeq S]} \simeq \exp\left[\frac{1}{6} N g'''(\mu_1)(s-\mu_1)^3\right]$$

$$= \exp\left[-\frac{1}{6} g'''(\mu_1)\frac{(S-N\mu_1)^3}{N^2}\right]$$

$\Rightarrow$ validity of the CLT requires

$$|S_N - N\mu_1|^3 \ll N^2 \Rightarrow \underline{|S_N - N\mu_1| \ll N^{2/3}}$$

If the distribution is symmetric we have

$g(s) = g(-s) \Rightarrow \mu_1 = 0,\ g'''(0) = 0$ and the estimate becomes

$$g''''(0)\ S^4/N^3 \ll 1 \Rightarrow \underline{|S_N| \ll N^{3/4}}$$

These estimates can be made rigorous by cumulant expansion, provided all cumulants exist.

What happens if $f(x)$ is a power law with exponent $\alpha > 2$? We assume

$$f(x) \simeq C x^{-(\alpha+1)}, \quad x \to \infty, \quad \alpha > 2.$$

Then it can be shown Problem

$$\underline{\mathrm{Prob}[S_N \simeq S] \simeq N C S^{-(\alpha+1)}}, \quad S \to \infty$$

This tail has to be matched to the CLT at some scale S^*:

$$\frac{1}{\sqrt{2\pi N\sigma^2}} e^{-S^2/2N\sigma^2} \simeq \frac{NC}{S^{\alpha+1}}$$

$$\Rightarrow \quad \frac{S^2}{2N\sigma^2} + \underbrace{\frac{1}{2}\ln(2\pi N\sigma^2)}_{\simeq \frac{1}{2}\ln N} \simeq (\alpha+1)\ln S - \underbrace{\ln(NC)}_{\simeq \ln N}$$

$$\Rightarrow \quad \frac{S^2}{2N\sigma^2} = (\alpha+1)\ln S - \frac{3}{2}\ln N = \ln\left(\frac{S^{\alpha+1}}{N^{3/2}}\right)$$

This suggests a solution $S^2/N \sim \ln N$

Ansatz: $S^* = a\sqrt{N \ln N}$

$$\Rightarrow \frac{S^2}{2N\sigma^2} = \frac{a^2 N \ln N}{2N\sigma^2} = \frac{a^2}{2\sigma^2}\ln N \overset{!}{=}$$

$$\overset{!}{=} \ln\left(N^{\frac{1}{2}(\alpha+1-3)}\right) + O(\ln\ln N)$$

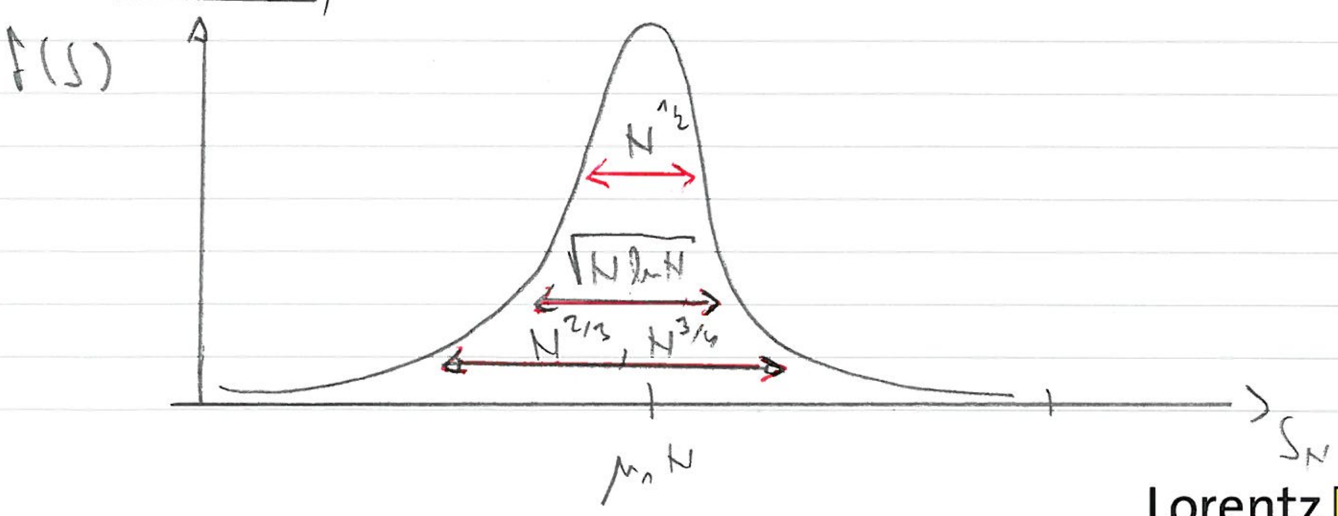
$$\Rightarrow \frac{a^2}{2\sigma^2} = \frac{1}{2}(\alpha-2) \Rightarrow \underline{a = \sigma\sqrt{\alpha-2}}$$

which is consistent for $\alpha > 2$. Thus in this case the CLT is limited to

$$\underline{|S_N - N\mu_1| \ll \sigma\sqrt{N \ln(N^{\alpha-2})}}$$

which is only slightly larger than $\sigma\sqrt{N}$.

Summary:



4° Extreme value theory

Consider N i.i.d. RV's $X_1, .., X_N$. What can we say about the RV

$$X_{max}(N) = \max_i \{X_1, .., X_N\} \; ?$$

More generally, we may rank order the RV's such that

$$X^{(1)} < X^{(2)} < \quad < X^{(N)}$$

and ask for the properties of the r'th order statistic $X^{(r)}$, where $X_{max} = X^{(N)}$, $X_{min} = X^{(1)}$.

In contrast to the sum S_N, X_{max} is a highly nonlinear function of the X_i.

As a physical motivation, consider the random energy model of spin glasses: (Derrida, 1981)

- N spins $\sigma_s = \pm 1$, $s = 1, .., N$
- Assign i.i.d. random energies E_i to each of the 2^N microstates

$\Rightarrow$ partition function $Z_N(\beta) = \sum_{i=1}^{2^N} e^{-\beta E_i}$ and

free energy $\mathcal{F}_N(\beta) = -\frac{1}{\beta} \ln Z_N(\beta)$

are also RV's. Consider two limits:

(i) $\underline{\beta \to 0}$: If $\beta E_i \ll 1$ for all i, we have

$$Z_N(\beta) \simeq \sum_{i=1}^{2^N} 1 - \beta E_i \simeq 2^N (1 - \beta \langle E \rangle)$$

according to the law of large numbers, and

$$\mathcal{F}_N(\beta) \cong -\frac{1}{\beta} \ln \left(2^N (1 - \beta \langle E \rangle) \right) \cong$$

$$= \langle E \rangle - T \underbrace{N \ln 2}_{\text{entropy } S} + \text{CLT-fluctuations}$$

(ii) $\underline{\beta \to \infty}$:

$$\underline{\mathcal{F}_N(\beta)} = -\frac{1}{\beta} \ln \Big(\underbrace{\sum_{i=1}^{N} e^{-\beta E_i}}_{\simeq e^{-\beta E_{min}}} \Big) = \underline{E_{min}}$$

$\Big\{$ $E_{min} = \min\{E_1, \dots, E_{2^N}\}$
ground state energy

a) Simple examples

Given N RV's with pdf f and distribution F, we have

$$\text{Prob}[\max(X_1, \ldots, X_N) < x] =$$

$$= \text{Prob}[X_1 < x \wedge X_2 < x \ldots \wedge X_N < x] = F(x)^N$$

$$\Rightarrow \quad F_{max}(x) = F(x)^N$$

The prob. density function is correspondingly

$$f_{max}(x) = \frac{d}{dx} F(x)^N = N f(x) F(x)^{N-1}$$

We consider some simple cases.

(i) Exponential: $F(x) = 1 - e^{-x}$, $f(x) = e^{-x}$

$$\Rightarrow \quad F_{max}(x) = (1 - e^{-x})^N$$

$$f_{max}(x) = N e^{-x} (1 - e^{-x})^{N-1} =$$

$$= N \exp(-x + (N-1) \ln (1 - e^{-x}))$$

How to scale x so that a nontrivial limit exists for $N \to \infty$?

The most probable value of X_{max} is obtained from

$$-1 + \frac{(N-1)\,e^{-x}}{1-e^{-x}} = 0 \;\Rightarrow\; N e^{-x} \simeq 1$$

$$\Rightarrow\; \underline{x = \ln N}$$

$\Rightarrow$ set $y = x - \ln N$, $x = y + \ln N$

$\Rightarrow$ $e^{-x} = \frac{1}{N} e^{-y}$

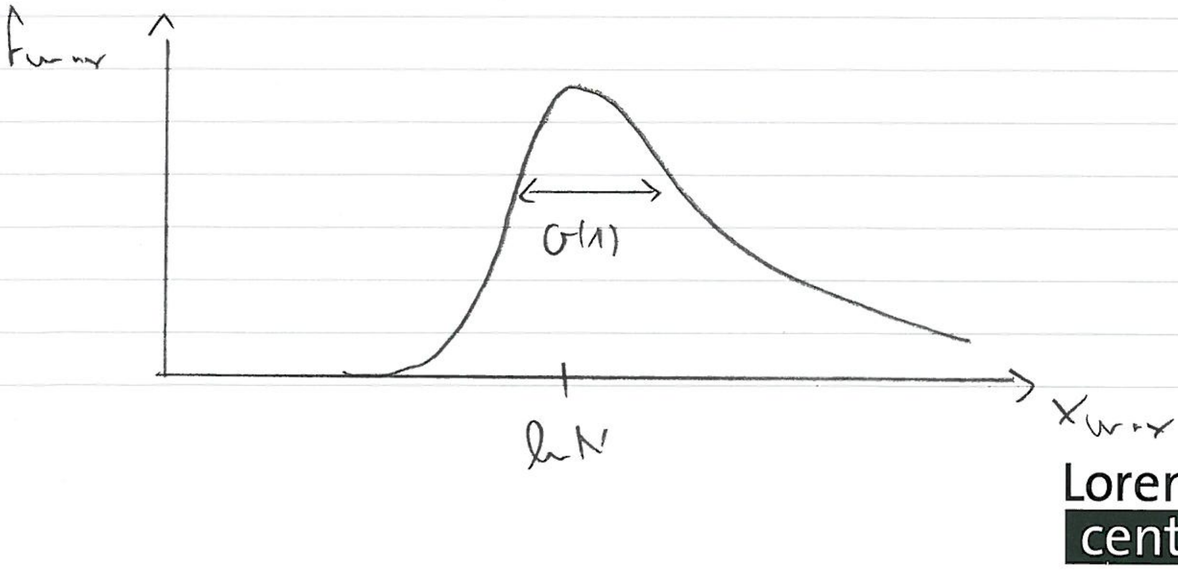
$\Rightarrow$ $F_{max} = \left(1 - \frac{1}{N} e^{-y}\right)^{N} \xrightarrow[N\to\infty]{} e^{-e^{-y}}$

$f_{max}(y) = \frac{d}{dy} F_{max}(y) = e^{-y-e^{-y}}$

<u>Gumbel distribution</u>

<u>Limits:</u> $f_{max}(y \to \infty) \simeq e^{-y}$

$f_{max}(y \to -\infty) \simeq e^{-e^{|y|}}$



Properties: $\mu_1 = \gamma \simeq 0.577215\ldots$ Euler constant
$\sigma^2 = \pi^2/6$

$\Rightarrow \langle X_{max} \rangle = \ln N + \gamma$

(ii) Uniform: $F(x) = x, \quad x \in [0,1]$

$\Rightarrow F_{max}(x) = x^N$
$f_{max}(x) = N x^{N-1}$

$\Rightarrow$ most likely value is $x = 1$.

Expected value:

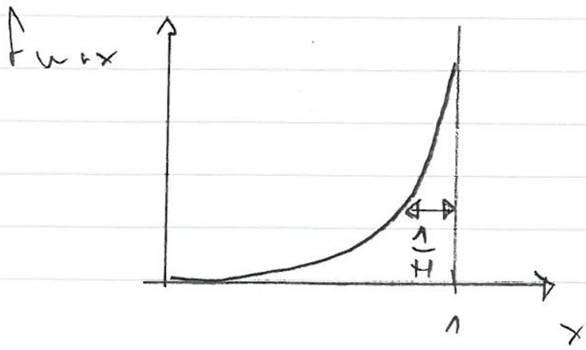
$$\langle X_{max} \rangle = \int_0^1 dx\, N x^N = \frac{N}{N+1} = 1 - \frac{1}{N+1}$$

Variance:

$$\langle X_{max}^2 \rangle = \int_0^1 dx\, N x^{N+1} = \frac{N}{N+2} = 1 - \frac{2}{N+2}$$

$$\Rightarrow \sigma^2 = \langle X_{max}^2 \rangle - \langle X_{max} \rangle^2 = \frac{N}{(N+2)(N+1)^2}$$

$$\simeq \frac{1}{N^2}$$



Scaling: $y = N(x - (1 - \frac{1}{N})) = 1 - N(1-x)$
$x = 1 - \frac{1}{N}(1-y)$

$$\Rightarrow F_{max} = x^N = \left(1 - \frac{1}{N}(1-y)\right)^N \xrightarrow[N\to\infty]{} e^{-(1-y)}$$

$$f_{max}(y) = e^{-(1-y)}, \quad y \in [-N, 1]$$

(iii) Pareto: $F(x) = 1 - x^{-\alpha}, \quad x \geq 1$

$$\Rightarrow F_{max}(x) = (1 - x^{-\alpha})^N$$
$$f_{max}(x) = \alpha x^{-(\alpha+1)} (1 - x^{-\alpha})^{N-1}$$

Most likely value $x \sim N^{1/\alpha}$

$\Rightarrow$ scale $y = x / N^{1/\alpha}$, $x = N^{1/\alpha} y$

$$\Rightarrow F_{max}(y) = \left(1 - (N^{1/\alpha} y)^{-\alpha}\right)^N =$$

$$= \left(1 - \frac{1}{N} y^{-\alpha}\right)^N \xrightarrow[N\to\infty]{} e^{-y^{-\alpha}}$$

$$f_{max}(y) = \alpha y^{-(\alpha+1)} e^{-y^{-\alpha}}$$