TD Large deviations (1)

1°) All the X_i are $\mathcal{G}(\mu,\sigma)$, iid, $S_n = \sum_{i=1}^{n} X_i$ is $\mathcal{G}(n\mu, \sigma\sqrt{n})$, $s = \frac{1}{n} S_n$

$$p_{S_n}(S) = \frac{1}{\sqrt{2\pi\sigma^2 n}} \exp\left[-\frac{(S-n\mu)^2}{2\sigma^2 n}\right]$$ pdf of S_n

$$\Rightarrow \quad p_n(s) = \frac{1}{\sqrt{2\pi\sigma^2/n}} \exp\left[-\frac{(s-\mu)^2}{2\sigma^2/n}\right]$$

p_{S_n} admits a large deviation form: $p_{S_n}(S = ns) = \frac{1}{\sqrt{2\pi\sigma^2 n}} e^{-n\phi(s)}$

$$\boxed{\phi(s) = \frac{1}{2}\left(\frac{s-\mu}{\sigma}\right)^2}$$

$$I_n(s) \equiv \frac{1}{n}\log p_n(s) = \phi(s) - \underbrace{\frac{1}{2n}\log(2\pi\sigma^2/n)}_{\xrightarrow[n\to\infty]{} 0}$$

NB $\frac{1}{2}\log 2\pi \simeq 0{,}91$ compatible with $n=1$ plot on right panel

$\boxed{\text{Method 2}}$ Sanov theorem. Formulated for discrete variables.

Discretize... assume ok for continuous here:

Minimize $D(q \| p(x))$ with $\begin{cases} \int q(x)\,dx = 1 \\ \int x\, q(x)\,dx = s \text{ ; not } \mu \end{cases}$; $p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$

Call q^* the corresponding optimum.

Then $P[S_n = ns] \doteq \exp\left(-n\, D(q^* \| p)\right)$

Minimize $\int_{\mathbb{R}} q(x)\log\frac{q(x)}{p(x)} - \lambda\int q - \mu\int x\, q(x)\,dx$

$$\frac{\delta}{\delta q(x)} = 0 = \log\frac{q(x)}{p(x)} + 1 - \lambda - \mu x$$

$$\text{ie} \quad q(x) = p(x)\, e^{1-\lambda-\mu x}$$

Hence $q(x)$ is gaussian. We know its mean s; its variance is the same as that of $p(x)$.

$$\Rightarrow \quad q^*(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-s)^2}{2\sigma^2}}$$

$$D(q^*\|p) = \int q^*(x)\left[\frac{(x-\mu)^2}{2\sigma^2} - \frac{(x-s)^2}{2\sigma^2}\right] = \frac{1}{2\sigma^2}\int q^*(x)\left[(s-\mu)(2x-s-\mu)\right] = \frac{1}{2\sigma^2}(s-\mu)(s-\mu)$$

$$\phi(s) = +D(q^*\|p) = +\frac{1}{2\sigma^2}(s-\mu)^2 \quad ✓$$

$\boxed{\text{Method 3}}$ Gärtner-Ellis theorem

We have to compute the Legendre transform of the cumulant generating function

$$K(t) = \log\langle e^{tX}\rangle = \frac{1}{2}t^2\sigma^2 + t\mu$$

$$\phi(s) = \sup_t\left(ts - K(t)\right)$$

$$\frac{\partial}{\partial t}\left[ts - K(t)\right] = s - t\sigma^2 + -\mu \quad \Rightarrow \quad t^* = \frac{s-\mu}{\sigma^2}$$

$$\phi(s) = t^*s - K(t^*) = s\frac{s-\mu}{\sigma^2} - \left[\frac{1}{2}\sigma^{-2}(s-\mu)^2 + \frac{s-\mu}{\sigma^2}\mu\right]$$

$$= \frac{1}{2}\frac{(s-\mu)^2}{\sigma^2} \quad ✓$$

2°) Bernoulli, $X=0$ proba $1-\alpha$
$X=1$ " α

Sanov route: q is a Bernoulli, with mean s $\Rightarrow$ proba $1-s$ for $X=0$
proba s for $X=1$

there is no minimization left!

$$D[q\|p] = (1-s)\log\frac{1-s}{1-\alpha} + s\log\frac{s}{\alpha} = \phi(s)$$

immediately

the calculation could have been made from the binomial: if the X_i are Bernoulli mean α
then their sum $S_n = \sum_{i=1}^n X_i$ is a binomial $\mathcal{B}(x, n\alpha)$

$$P\left[\sum_{i=1}^n X_i = ns\right] = \binom{n}{ns}\alpha^{ns}(1-\alpha)^{n-ns}$$

3°) Coin tossed $n=100$ times. Outcome of one toss is Bernoulli

⊛ p is unfair: $\alpha = 0{,}9$ (for getting T). We take q fair with proba $0{,}5$

$$D[q\|p] = 0{,}5\log_2\frac{0{,}5}{0{,}1} + 0{,}5\log_2\frac{0{,}5}{0{,}9} \simeq 0{,}73$$

$$\Pr[50H/50T] \doteq 2^{-73} \simeq 10^{-22}$$

Large deviations | 2

(*) p is fair, $r = 0,5$; q is unfair with proba $0,9$ for Tail

$$D[q\|p] = 0,9 \log_2 \frac{0,9}{0,5} + 0,1 \log_2 \frac{0,1}{0,5} \simeq 0,53$$

$$\Pr[90T/10H] \doteq 2^{-53} \simeq 10^{-16}$$

These 2 proba are different : more likely to get a large dev with the fair coin, that the converse for the biased coin.
this illustrates the fact that the Kullback-Leibler "distance" is non symmetric.

$$D[p\|q] \neq D[q\|p]$$

4°) $S_n = \sum_{i=1}^{n} \eta_i \quad ; \quad p(\eta) = \frac{1}{2} e^{-|\eta|}$.

$\boxed{\text{Sanov}}$ written in the continuum :

Minimize $D(q\|p) = \int_{\mathbb{R}} q(\eta) \log \frac{q(\eta)}{p(\eta)} d\eta$; constraint $\int \eta\, q(\eta)\, d\eta = s$

$\int q(\eta)\, d\eta = 1$

$$\Rightarrow \quad \log \frac{q(\eta)}{p(\eta)} + 1 +\!- \lambda +\!- \mu \eta = 0$$

$$\Rightarrow \quad q(\eta) = \frac{1}{Z(\mu)} e^{-|\eta| + \mu\eta}$$

"partition function" $\boxed{Z(\mu) = \int_{\mathbb{R}} d\eta\, e^{-|\eta| + \mu\eta}}$

$\boxed{-1 < \mu < 1}$ for normalization

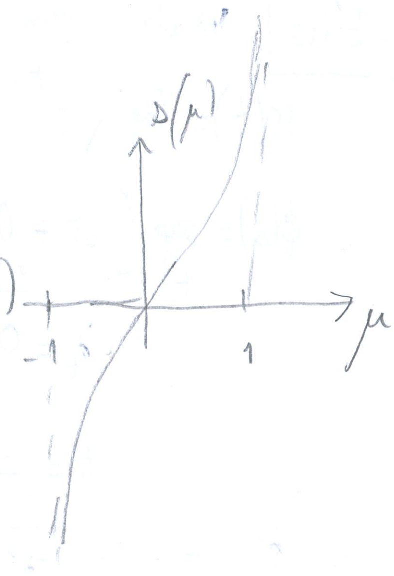
where the Lagrange multiplier μ is s.t.

$$\int \eta\, q(\eta)\, d\eta = s = \frac{1}{Z} \partial_\mu Z = \frac{\partial \log Z}{\partial \mu}$$

$$Z(\mu) = \int_{-\infty}^{0} e^{\eta + \mu\eta} d\eta + \int_{0}^{\infty} e^{-\eta + \mu\eta}$$

$$= \frac{1}{1+\mu} + \frac{1}{1-\mu} = \frac{2}{1-\mu^2}$$

$$\frac{\partial \log Z}{\partial \mu} = -\partial_\mu \log(1-\mu^2) = \frac{2\mu}{1-\mu^2} = s(\mu)$$



Hence all possible s values on $\mathbb{R}$ can be reached by some μ s.t. $\frac{2\mu}{1-\mu^2} = s$

$$\frac{2\mu}{1-\mu^2} = s \iff (1-\mu^2)s = 2\mu \iff s\mu^2 + 2\mu - s \overset{!}{=} 0$$

$$\mu = \frac{-2 \pm \sqrt{4+4s^2}}{2s} = \frac{-1 \pm \sqrt{1+s^2}}{s}$$

We need the root in $[-1,+1]$: $s>0$: $\mu = \frac{-1+\sqrt{1+s^2}}{s}$

$s<0$: $\mu = \frac{-1+\sqrt{1+s^2}}{s}$ also!

the other root is not in $[-1,+1]$: $\left(\frac{-1+\sqrt{\ }}{s}\right)^2 = \frac{1+1+s^2+2\sqrt{\ }}{s^2} > 1$ always

All in all:

$$q(\eta) = \frac{1-\mu^2}{2} e^{-|\eta|+\mu\eta} \qquad 1-\mu^2 = \frac{2\mu}{s} = \frac{(-1+\sqrt{1+s^2})2}{s^2}$$

$$= p(\eta)\,\frac{(-1+\sqrt{1+s^2})2}{s^2} e^{\mu\eta}$$

$$D(q\|p) = \int q(\eta) \log\left(\frac{q}{p}\right) d\eta = \int d\eta\, \underbrace{q(\eta)\,[\mu\eta}_{\hookrightarrow \mu\langle\eta\rangle_q = \mu s} + \log\left(\frac{(-1+\sqrt{1+s^2})2}{s^2}\right)]$$

$$= \phi(s)$$

$$\phi(s) = s\,\mu(s) + \log\left(\frac{(-1+\sqrt{1+s^2})2}{s^2}\right)$$

$$\boxed{\phi(s) = -1+\sqrt{1+s^2} + \log\left(\frac{(-1+\sqrt{1+s^2})2}{s^2}\right)} \quad ; \quad \phi(s) = \phi(-s)$$

from symmetry of $p(\eta)$

$\boxed{\text{Gärtner - Ellis}}$ we start from the cumulant generating function:

$$\kappa(t) = \log\langle e^{t\eta}\rangle = -\log(1-t^2)$$

; same as calculation of partition function $Z(\mu = t)$ above

$$\phi(s) = \sup_t [\underbrace{st + \log(1-t^2)}]$$

$$\partial_t = 0 = s + -\frac{2t}{1-t^2} \;;$$

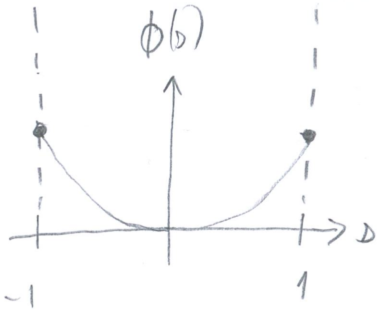
again, same as above with $\mu = t$

note $-1 \leq t \leq 1$ again

$$t = \frac{-1+\sqrt{1+s^2}}{s}$$

$$\Rightarrow \boxed{\phi(s) = -1+\sqrt{1+s^2} + \log\left(\frac{2(-1+\sqrt{1+s^2})}{s^2}\right)}$$

Large deviations | 3



$$\phi(s) \begin{cases} \xrightarrow{s\to 0} & \dfrac{s^2}{4} + (\ldots)s^4 \\ \xrightarrow{s\to 1} & -1+\sqrt{2}+\log\left(2(\sqrt{2}-1)\right) \simeq 0{,}226 \end{cases}$$

$$P(S_n) \doteq e^{-n\,\phi\left(\frac{S_n}{n}\right)} \sim e^{-n\left(\frac{S_n}{n}\right)^2\frac{1}{4}} \quad \text{for } |S_n| \ll n$$

$$\sim e^{-\frac{S_n^2}{4n}}$$

and we recover the central limit theorem :
$$\langle \eta \rangle = 0$$
$$\langle \eta^2 \rangle = \frac{1}{2}\int \eta^2 e^{-|\eta|} = 2$$
$$V(\underbrace{\sum_{i=1}^{n} \eta_i}_{S_n}) = n\,V(\eta) = 2n$$

5c) System with hamiltonian $\mathcal{H}(\mathcal{C})$ for a given configuration $\mathcal{C}$ (microscopic)

$$Z(\beta) = \sum_{\mathcal{C}} e^{-\beta\mathcal{H}(\mathcal{C})} \quad \text{is the partition function}$$

$$\langle e^{tE} \rangle = \langle e^{t\mathcal{H}} \rangle = \frac{1}{Z(\beta)}\sum_{\mathcal{C}} e^{t\mathcal{H}-\beta\mathcal{H}} = \frac{Z(\beta-t)}{Z(\beta)}$$

the cumulant generating function $\mathcal{K}(t)$ follows :

$$\mathcal{K}(t) = \log\langle e^{tE}\rangle = \log\frac{Z(\beta-t)}{Z(\beta)} = \beta F(\beta) - (\beta-t)F(\beta-t)$$

$$\langle E^2\rangle - \langle E\rangle^2 = \left.\frac{d^2\mathcal{K}}{dt^2}\right|_{t=0} = \left.\frac{\partial^2}{\partial t^2}\log Z(\beta-t)\right|_{t=0} = \frac{\partial^2}{\partial\beta^2}\log Z(\beta)$$

$$= \frac{\partial}{\partial\beta}\Big(\underbrace{\frac{1}{Z}\frac{\partial Z}{\partial\beta}}_{-\langle E\rangle}\Big) \quad ; \quad \left.\frac{\partial\langle E\rangle}{\partial T}\right|_V = c_V$$

$$= kT^2\left.\frac{\partial\langle E\rangle}{\partial T}\right|_V$$

$$= kT^2\, c_V$$