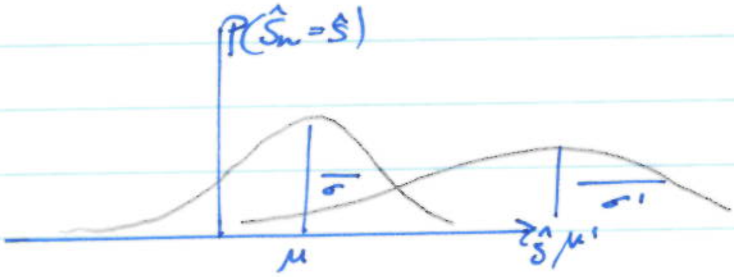
2.1. Gaussian sum

$\hat{S}_n = \sum_{i=1}^{n} X_i \qquad X_i \sim \mathcal{N}(\mu, \sigma^2) \quad \text{iid}$

- $\hat{S}_n$ is Gaussian distributed
- Mean: $E\hat{S}_n = E\sum_i X_i = \sum_i EX_i = n\mu$
- Variance: $\mathrm{var}(\hat{S}_n) = \sum_i \mathrm{var}\, X_i = n\sigma^2$

$\left.\right\}\ \hat{S}_n \sim \mathcal{N}(n\mu, n\sigma^2)$



mean $\to \infty$

variance $\to \infty$

- CLT: $\bar{S}_n = \dfrac{\hat{S}_n - n\mu}{\sqrt{n}\,\sigma} \sim \mathcal{N}(0,1)$

mean $= 0$

variance $=$ constant

$\bar{S}_n = \bar{s}$ about mean $\Rightarrow \hat{S}_n = \sqrt{n}\,\bar{s}$ about mean

look at $\hat{S}_n \sim \sqrt{n}$

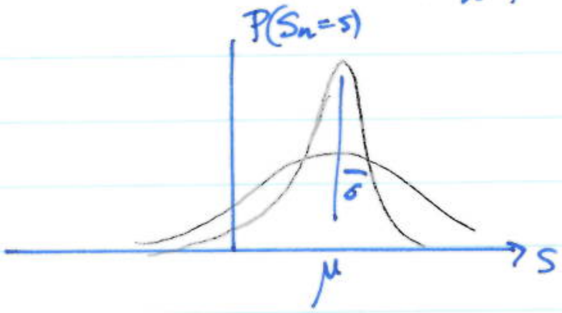
- Large deviations:

$S_n = \frac{1}{n}\sum_{i=1}^{n} X_i \qquad$ sample mean

mean = constant

variance $\to 0$

- Mean: $ES_n = \mu$
- Variance: $\mathrm{var}\, S_n = \mathrm{var}\left(\frac{\hat{S}_n}{n}\right) = \frac{1}{n^2}\mathrm{var}(\hat{S}_n) = \frac{\sigma^2}{n}$



$\sigma_n \to 0$

- $P(S_n = s) \to \delta(s - \mu) \qquad$ Law of large numbers
- $S_n \to \mu$ with probability 1 (or in probability)
- $P(|S_n - \mu| > \varepsilon) \to 0 \quad \forall \varepsilon > 0$
- Looks Gaussian around $s = \mu$ (really?)
- $S_n = s \Rightarrow \hat{S}_n = ns \qquad$ look at large values of sum

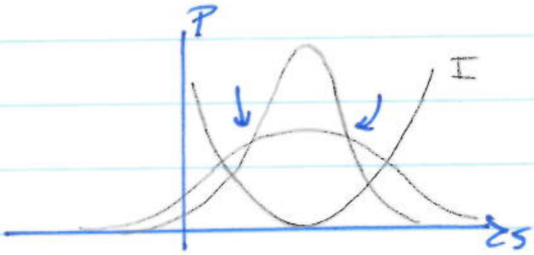
· Density :

$$P(S_n = s) = \underbrace{\sqrt{\frac{n}{2\pi\sigma^2}}}_{\text{subdominant}} \; e^{-n\frac{(s-\mu)^2}{2\sigma^2}}$$

n dominant : responsible for concentration

· Dominant part :

$$P(S_n = s) \approx e^{-n I(s)} \qquad I(s) = \frac{(s-\mu)^2}{2\sigma^2}$$



· Large deviation limit:

$$\lim_{n\to\infty} -\frac{1}{n} \ln P(S_n = s) = \frac{(s-\mu)^2}{2\sigma^2} - \underbrace{\lim \frac{1}{n} \ln \sqrt{\frac{n}{2\pi\sigma^2}}}_{\to 0}$$

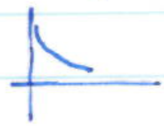
$$= I(s)$$

$$\Rightarrow \lim_{n\to\infty} -\frac{1}{n} \ln P(S_n = s) = I(s) \iff P(S_n = s) \approx e^{-n I(s)}$$

· More precise : $P(S_n = s) = e^{-n I(s) + o(n)}$

2.2. Exponential sample mean

$$S_n = \frac{1}{n}\sum_{i=1}^{n} X_i \qquad X_i \sim p(x) = \frac{1}{\mu} e^{-x/\mu}, \quad x \geq 0, \ \mu > 0$$



iid

$EX_i = \mu$

$P(S_n = s)$?

· Generating function method

$$G_{S_n}(k) = E[e^{ikS_n}]$$

$$= \mathcal{F} P_{S_n}$$

$$= E\left[e^{ik\frac{1}{n}\sum_i X_i}\right]$$

$$= E\left[e^{i\bar{k}\sum_i X_i}\right] \qquad \bar{k} = k/n$$

$$= E\left[\prod_{j=1}^{n} e^{i\bar{k}X_j}\right]$$

$$= \prod_{j=1}^{n} E\left[e^{i\bar{k}X_j}\right] \qquad \text{independence}$$

$$= G_X(\bar{k})^n \qquad \text{id. dist.}$$

Can also use GF

· Inverse Fourier:

$$P(S_n = s) = \mathcal{F}^{-1} G_{S_n}(k)$$

$$= \frac{1}{2\pi}\int_{-\infty}^{\infty} dk\ e^{-iks}\, G_X(k/n)^n$$

exactly or ...

· Saddle-point approximation:

$$P(S_n = s) \approx \int_{-\infty}^{\infty} dk\ e^{-inks}\, G_X(k)^n \qquad k \to nk$$

$$= \int_{-\infty}^{\infty} dk\ e^{-inks}\, e^{n \ln G_X(k)}$$

$$\approx e^{-nI(s)}$$

$$I(s) = k_s s - \lambda(k_s) \qquad \lambda(k) = \ln G(k)$$

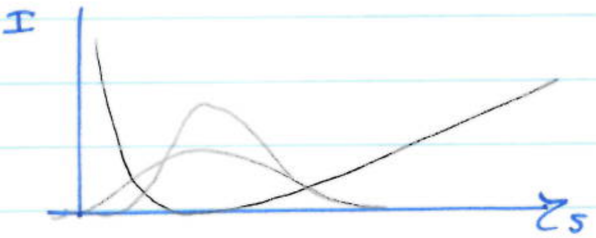
$$k_s: \quad \lambda'(k_s) = s$$

· Solution

$$\lambda(k) = -\ln\left(\left(\tfrac{1}{\mu} - k\right)\mu\right)$$
$$= -\ln\left(1 - k\mu\right)$$
$$= \ln \frac{1}{1-k\mu} \qquad k < \frac{1}{\mu}$$

$\lambda'(k) = s$ ... solve for k See exercises

$$I(s) = k_s s - \lambda(k_s) = \frac{s}{\mu} - 1 - \ln\frac{s}{\mu}$$



- $I(s^*) = 0 \quad s^* = \mu \quad \Rightarrow P(S_n = s) \to \delta(s-\mu)$ LLN
- $I(s) \sim \frac{s}{\mu} \quad s \gg \mu \quad \Rightarrow P(S_n = s) \approx e^{-ns/\mu}$ exp. tail
- Around mean: $I(s) \approx \frac{(s-\mu)^2}{2\sigma^2} \Rightarrow P(S_n = s) \approx e^{-n I(s)}$ Gaussian CLT

· Large deviation limit:

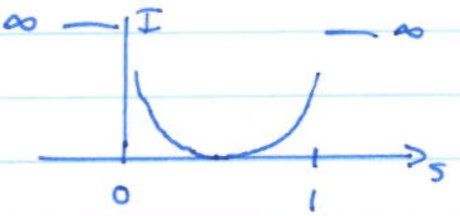
$$\lim_{n\to\infty} -\frac{1}{n}\ln P(S_n = s) = I(s)$$

· Rem: $P(S_n = s)$ for $s < 0$?
$I(s)$ " " ?

2.3. Bernoulli sample mean

$$S_n = \frac{1}{n}\sum_{i=1}^{n} X_i \qquad X_i \sim \mathrm{Bern}(p) \quad \text{i.e.} \quad \begin{aligned} P(X_i = 1) &= p \\ P(X_i = 0) &= 1-p \end{aligned}$$

· Rate function guess: $p = 1/2$



$n_1 + n_0 = n$

· Combinatorial calculation

$$S_n = s = \frac{1}{n}\sum_{i=1}^{n} X_i = 0\cdot\frac{n_0}{n} + \frac{n_1}{n}\cdot 1 = \frac{n_1}{n}$$

But

$$P(n_1 \cancel{=ns}) = \frac{n!}{n_0!\, n_1!}(1-p)^{n_0} p^{n_1}$$

So

$$P(S_n = s) = P(n_1 = ns)$$

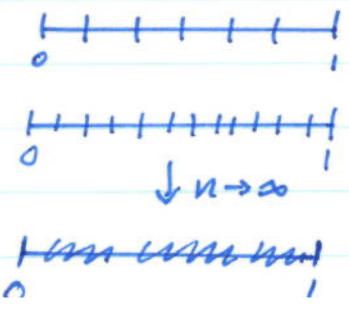
$$= \frac{n!}{(ns)!\,(n(1-s))!}(1-p)^{n(1-s)} p^{ns} \qquad \begin{aligned} n_0 &= n - n_1 \\ &= n - ns \end{aligned}$$

Use Stirling approximation: $n! \approx n^n e^{-n}$

$$P(S_n = s) \approx e^{-n I(s)}$$

$$I(s) = \cdots \qquad \text{See exercises}$$

· Rem: Values of S_n: $\left\{0, \frac{1}{n}, \frac{2}{n}, \frac{3}{n}, \dots, \frac{n-1}{n}, 1\right\}$



$P(S_n = s) \approx e^{-nI(s)}$ a density!

2.4. Large deviation principle – pdf case

Def: The pdf $P(S_n = s)$ satisfies the large deviation principle (LDP) if the limit

$$\lim_{n\to\infty} -\frac{1}{n} \ln P(S_n = s) = I(s)$$

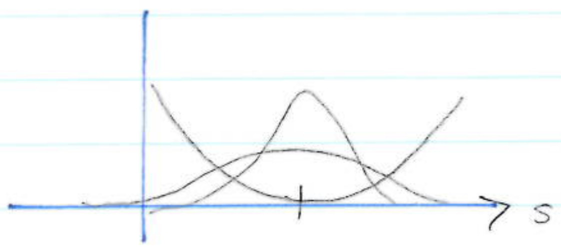
exists and yields a function $I(s)$ different from 0 and ∞.

$I(s)$: rate function

Interpretation:
$$\lim_{n\to\infty} -\frac{1}{n} \ln P(S_n = s) = I(s)$$
$$\Updownarrow$$
$$P(S_n = s) \approx e^{-nI(s)}$$
$$\Updownarrow$$
$$P(S_n = s) = e^{-nI(s) + o(n)}$$

- $P(S_n = s) \downarrow 0$ exp. with n $\forall$ s $\ni$ $I(s) > 0$
  $\not\downarrow 0$ " " " for s $\ni$ $I(s) = 0$

$\Rightarrow$ concentration on $s \ni I(s) = 0$



- Example: Gaussian
  Exponential
  ...

2.5. Large deviation principle – Probability case
S_n has LDP with r.f. $I(s)$

$$\Rightarrow \quad P(S_n \in B) = \int_B P(S_n = s)\, ds$$
$$\approx \int_B e^{-n I(s)}\, ds$$
$$\approx e^{-n \min_{s\in B} I(s)} \qquad \text{Laplace approximation}$$
$$= e^{-n I_B} \qquad I_B = \min_{s\in B} I(s)$$

$$\Rightarrow \quad \lim_{n\to\infty} -\frac{1}{n} \ln P(S_n \in B) = I_B \qquad \text{true } \forall B$$

Rem: Laplace principle

$$\sum_i e^{n a_i} \approx e^{n \max_i a_i} \qquad \lim \frac{1}{n} \ln \sum_i e^{n a_i} \to \max_i a_i$$

$$\int e^{n f(x)} dx \approx e^{n \max f(x)} \qquad \lim \frac{1}{n} \ln \int e^{n f(x)} dx = \max_x f(x)$$

Def: S_n satisfies the LDP if

$$\lim_{n\to\infty} -\frac{1}{n} \ln P(S_n \in B) = \min_{s\in B} \underbrace{I(s)}_{\text{rate fct}}$$

Rem: Upper + lower bounds
Reading: cf website

2.2. Why look at large deviations?

- System of N particles
- Energy $\sim N$
- Fluctuations $\sim \cancel{\frac{1}{\sqrt{N}}}$ $\sqrt{N}$

- System N particles, " $2N$ " $\Big\}$ Don't measure "twice" the properties

- Measurements $\rightarrow$ intensive quantities $\left\{\begin{array}{l}\text{specific heat} \\ \text{" energy} \\ \text{etc.}\end{array}\right.$
- Stability $\rightarrow$ fluctuations measured w/r these intensive quantities

- $h_N = \frac{H_N}{N} \rightarrow$ constant specific energy, mean "

- $var(h_N) \sim 1/N \downarrow 0$

- Concentration of fluctuations $\Rightarrow$ Large deviations