Probability

$(\Omega, \mathcal{F}, \mathbb{P})$ – Probability space

$X : \Omega \to \mathbb{R}$ – random variable

$\mathbb{E}[X] = \int X \, d\mathbb{P}$. [Integration]

Antiderivative

$$\int_a^b f(x)\,dx = F(b) - F(a)$$

$$F'(x) = f(x)$$

Riemann Integration

- Upper sums
- lower sums

} "area" under the curve

$f \in C[a,b]$

Theorem:- $f : [a,b] \to \mathbb{R}$ bounded is Riemann integrable

$\iff$

Set of discontinuities of f has lebesgue measure 0.

Theorem: (fundamental theorem of Integral Calculus)

$f \in \ldots$ $\exists\, F \in \ldots$ st $F' = f$

$$F(b) - F(a) = \int_a^b f(x)\,dx$$

Monotone Convergence Theorem on $(\Omega, \mathcal{B}, \mu)$

A B

If f, f_n non-negative measurable and $f_n(\omega) \nearrow f(\omega)$, $\forall \omega \in \Omega$, then

$$\int f d\mu = \sup_n \int f_n d\mu = \lim_{n\to\infty} \int f_n d\mu \ .$$

. N.S.C <... A and B

$$\boxed{f_n(\omega) = n \mathbb{1}_{[0,\frac{1}{n}]}}$$

Bounded

Domintated Convergence Theorem on $(\Omega, \mathcal{B}, \mu)$

[Uniform Integrability Convergence Theorem]

Let $\{f_n\}_{n=1}^{\infty}$ be a sequence of integrable functions, which is uniformly dominated by an integrable function : i.e. suppose there exists integrable g such that $|f_n(\omega)| \leq g(\omega) \ \ \forall \omega \in, \forall n$.

If $f_n(\omega) \to f(\omega) \ \ \forall \omega \in \Omega$, then

$$f \text{ is integrable and } \lim_{n\to\infty} \int f_n d\mu = \int f d\mu.$$

Fatou's Lemma

[No mass can arise at "∞"]

If $\{f_n : n = 1, 2, \cdots\}$ is any sequence of non-negative measurable functions, then

$$\int \liminf_{n\to\infty} f_n d\mu \leq \liminf_{n\to\infty} \int f_n d\mu \ .$$

Fubini's Theorem

$$\int_{\Omega_1}\int_{\Omega_2} f(\omega_1,\omega_2)\, d\omega_2\, d\omega_1 = \int_{\Omega_2}\int_{\Omega_1} f(\omega_1,\omega_2)\, d\omega_1\, d\omega_2$$

Probability space

Let $(\Omega, \mathcal{B}, \mu)$ denote the product of the σ-finite measure spaces $(\Omega_1, \mathcal{B}_1, \mu_1)$ and $(\Omega_2, \mathcal{B}_2, \mu_2)$. Let f be integrable on $(\Omega, \mathcal{B}, \mu)$. Then,

$\int |f(\omega_1,\omega_2)|\, d\mu_1(\omega_1)\, d\mu_2(\omega_2)$ & f is measurable w.r.t. $(\mu, \mathcal{B})$ — $\circledast$

Fubini's Theorem

In this class: $\mathbb{C} = \mathbb{R}$

(i) for μ_1-almost all x in Ω_1, the function

$$f^x : \Omega_2 \to \mathbb{C} \text{ given by } f^x(y) = f(x, y)$$

is $(\mathcal{B}_2, \mathcal{B}_{\mathbb{C}})$-measurable and in fact $f^x \in L^1(\Omega_2, \mathcal{B}_2, \mu_2)$;
(i)$'$ for μ_2-almost all y in Ω_2, the function

$$f_y : \Omega_1 \to \mathbb{C} \; f^y(x) = f(x, y)$$

is $(\mathcal{B}_1, \mathcal{B}_{\mathbb{C}})$-measurable and in fact $f_y \in L^1(\Omega_1, \mathcal{B}_1, \mu_1)$;

Fubini's Theorem

(ii) the μ_1-almost everywhere defined function

$$x \xrightarrow{g} \int f^x(y)d\mu_2(y)$$

is $(\mathcal{B}_1, \mathcal{B}_{\mathbb{C}})$-measurable and in fact it is integrable with respect to μ_1; $\therefore \int |g(x)| d\mu_1(x) < \infty$

(ii)′ the μ_2-almost everywhere defined function

$$y \xrightarrow{h} \int f_y(x)d\mu_1(x)$$

is $(\mathcal{B}_2, \mathcal{B}_{\mathbb{C}})$-measurable and in fact it is integrable with respect to μ_2; and $\therefore \int h(y) d\mu_2(y) < \infty$

Fubini's Theorem

(iii)

$$\overset{\mathrm{I}}{\boxed{\int_\Omega f d\mu}} = \overset{\mathrm{II}}{\boxed{\int_{\Omega_1}\left(\int_{\Omega_2} f(x,y)d\mu_2(y)\right)d\mu_1(x)}}$$

$$= \boxed{\int_{\Omega_2}\left(\int_{\Omega_1} f(x,y)d\mu_1(x)\right)d\mu_2(y).} \; \mathrm{III}$$

- worksheet due at 11:00 am

Till now :- Basics in (Measure Theoretic) Probability

- Integration - Expectation / variance
- limit theorems [Borel Cantelli / WLLN]
- Fatou's lemma, D.C.T., M.C.T., Fubini
- Moment generating functions / Chernoff bounds

Focus :

① Large Deviations Principle (LDP) .... (new)

② ✓ Strong law of large numbers ...

③ ✓ Central limit Theorem. $\rightsquigarrow$ [Stein's Method]

15. Large Deviation Principle (LDP)

Many questions in Probability can be formulated as a law of large numbers

Example 1:- $(\Omega, \mathcal{F}, \mathbb{P})$ - $A \in \mathcal{F}$ event

- independent trials (n) & $X_n = 1$ if A occurs, 0 otherwise

$$\bar{X}_n : \frac{X_1 + X_2 \dots + X_n}{n} = (\text{relative frequency of } A)_n$$

Application of Tschebychev:

(WLLN) $\mu = \mathbb{P}(A)$ $\quad \bar{X}_n \xrightarrow{P} \mu$ as $n \to \infty$

- Deviations from this typical behaviour

$\varepsilon > 0$ $\qquad \mathbb{P}(\bar{X}_n > \mu + \varepsilon) \leq \exp(-n(\dots))$

chernoff bounds

universal bound

$$\boxed{\exists\ \delta > 0 \quad E[e^{tX_1}] < \infty \quad t \in (-\delta, \delta)}$$

Question :- Is there a limit :

$$\frac{1}{n} \log \mathbb{P}(\bar{X}_n > \mu + \varepsilon) \longrightarrow ?$$

as $n \to \infty$

Example 2 : (Equilibrium Statistical Mechanics)

- [Model] in which each state has a certain energy
- Equilibrium - states with lower energy are likely to occur.

Model :- Gibbs measure.

E - Countable set of states ; $\beta > 0$; $H : E \to \mathbb{R}$ is the energy functions

$$\mathbb{P}_\beta(\{x\}) = \frac{e^{-\beta H(x)}}{\sum_{x \in E} e^{-\beta H(x)}} = e^{-\beta H(x) - \log Z_\beta}$$

where $\quad Z_\beta := \sum_{x \in E} e^{-\beta H(x)}$

$(*) \rightarrow$:. $\mathbb{P}_\beta(\{x\}) = e^{-\beta I(x)}$ [Assign higher probability to lower energy states]

where $I(x) = \dfrac{H(x) - \log Z_\beta}{\beta}$

Intuitively: Most probable state is the one where H takes its minimum.

LDP - are probability measures that behave like $(*)$ and $\beta \to \infty$ as $\epsilon \to 0$?

Example 3 :- $X_i \overset{d}{=} \text{Bernoulli}(p)$ $i \geq 1$

& each X_i's are independent.

- $E[X_i] = p$ $\quad$ $\text{var}[X_i] = p(1-p)$

$$S_n = \sum_{i=1}^{n} X_i \sim \text{Binomial}(n, p)$$

$$\mathbb{P}(S_n = k) = \binom{n}{k} p^k (1-p)^{n-k} \quad k = 0, 1, \ldots n$$

By Stirling's approximation

$$\mathbb{P}(S_n = k) = \binom{n}{k} p^k (1-p)^{n-k} = \frac{n!}{k!\,(n-k)!} p^k (1-p)^{n-k}$$

$$\cong \sqrt{\frac{n}{2\pi\, k(n-k)}} \cdot \exp\begin{pmatrix} n\log n - k\log k + k\log p \\ -(n-k)\log(n-k) \\ +(n-k)\log(1-p \end{pmatrix}$$

$$= \sqrt{\frac{n}{2\pi\, k(n-k)}} \exp\left\{ n \left(\frac{k}{n}\log p + \left(1-\frac{k}{n}\right)\log(1-p) - \frac{k}{n}\log\frac{k}{n} - \left(1-\frac{k}{n}\right)\log\left(1-\frac{k}{n}\right) \right)\right.$$

$\underline{k = [nx]}$

let $x \in [0,1]$

$$\mathbb{P}(S_n = [nx]) = \binom{n}{[nx]} p^{[nx]} (1-p)^{n-[nx]}$$

$$= \sqrt{\frac{1}{2\pi\, nx(1-x)}} \exp(-n\, I(x) + O(\log n))$$

$$= \exp(-n\, I(x) + O(\log n))$$

Entropy:

$$I(x) = x\log\left(\frac{x}{p}\right) + (1-x)\log\left(\frac{(1-x)}{1-p}\right) - (**)$$

$$\boxed{\therefore \frac{1}{n}\log \mathbb{P}(S_n = [nx]) \longrightarrow -I(x) \quad \text{as } n\to\infty}$$ — (LDP)

I given by $(**)$ $\quad I(p) = 0 \overset{(Ex)}{=} I'(p) \quad I''(p) = \frac{1}{p(1-p)} > 0$

Ex: p is the unique minima of $I(\cdot)$

[WLLN] $\frac{1}{n} S_n \longrightarrow p$ as $n \to \infty$ in Probability

$S_n = [nx]$ is

$x \neq p$, LDP $\equiv$ exponentially rare event.

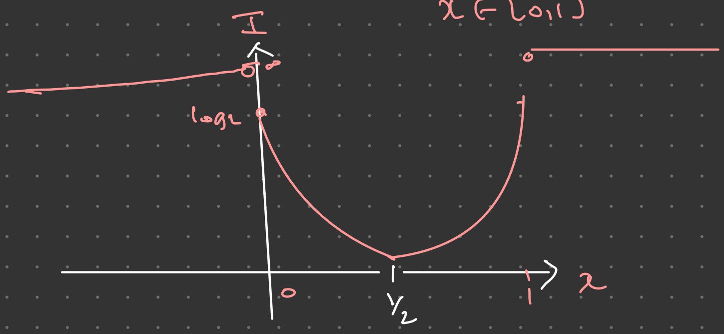
[Ex:-] Taylor expansion of $I(\cdot)$ around p

$$\mathbb{P}(S_n = [nx]) \approx \frac{1}{\sqrt{2\pi p(1-p)}} e^{-\frac{1}{2} I''(p) y^2}$$

where $x = p + \frac{y}{\sqrt{n}}$ (Central limit Theorem)

$p = \frac{1}{2}$ $\qquad I(x) = \log 2 + x \log x + (1-x) \log(1-x)$

$x \in [0,1]$



$\mathbb{P}(S_n = [nx]) = 0 \quad x < 0 \quad x > 1$

$\Rightarrow \quad \mathbb{P}(S_n = [nx]) \asymp e^{-n I(x)}$.

Theorem [CRAMER's Theorem] ~ [AMM - Dec 2011]
Cerf - Petit

Let $\{X_n\}_{n\geq 1}$ be a sequence of independent & identically distributed random variables and

$$\bar{X}_n = \frac{X_1 + \dots + X_n}{n}$$

Assume $E[e^{tX_1}] < \infty \quad \forall\, t \in \mathbb{R}$

- $\frac{1}{n} \log \mathbb{P}(\bar{X}_n \geq x) \xrightarrow[n\to\infty]{}$ a real number $\forall\, x \in \mathbb{R}$

- $\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}(\bar{X}_n \geq x) = \inf_{\lambda>0} \left(\log E[e^{\lambda X_1}] - \lambda x \right)$

Proof:-

$\forall\, x \in \mathbb{R} \qquad s(x) = \sup_{n\geq 1} \frac{1}{n} \log \mathbb{P}(\bar{X}_n \geq x)$

$\forall\, \lambda \in \mathbb{R} \qquad p(\lambda) = \log E[e^{\lambda X_1}]$

- Duality:- $\quad p(\lambda) = \sup_{u \in \mathbb{R}} (\lambda u + s(u))$

- Convergence: $\quad \frac{1}{n} \log \mathbb{P}(\bar{X}_n \geq x)$ converges to $s(x)$ as $n\to\infty \quad \forall\, x \in \mathbb{R}$