R.v's $X_1,\dots,X_n$ are "jointly absolutely continuous" if the random vector $X=(X_1,\dots,X_n)$ has an absolutely continuous distribution, i.e. $\exists$ integrable $f:\mathbb{R}^n\to\mathbb{R}$ such that

$$P\{X\in B\} = \int_B f(x)\,dx \qquad \forall B\in\mathcal{B}^n.$$

i.e.

$$P\{X_1\in B_1,\dots,X_n\in B_n\} = \int_{B_1\times\dots\times B_n} \underbrace{f(x_1,\dots,x_n)}_{\text{"joint pdf" of } X_1,\dots,X_n}\,dx_1\dots dx_n \qquad \forall B_i\in\mathcal{B}.$$

THM TFAE for $\forall$ random variables $X_1,\dots,X_n$:

(a) $X_1,\dots,X_n$ are (individually) absolutely continuous and independent;

(b) $X_1,\dots,X_n$ are (jointly) absolutely continuous and their joint pdf satisfies

$$f(x_1,\dots,x_n) = f_{X_1}(x_1)\cdots f_{X_n}(x_n) \quad \text{a.e. on } \mathbb{R}^n \qquad (*)$$

Proof for $n=2$

(a) $\Rightarrow$ (b). WTS:

$$P\{X\in B\} = \underbrace{\int_B f_{X_1}(x_1)\,f_{X_2}(x_2)\,dx_1\,dx_2}_{\overset{!!}{\mu(B)} \text{ is a prob. measure (check!)}} \qquad \forall B\in\mathcal{B}^2 \qquad (*)$$

Uniqueness of measures (p.36) $\Rightarrow$ enough to check $(*)$ for product sets

$$B = B_1\times B_2, \quad B_i\in\mathcal{B}$$

$$P\{X\in B\} = P\{X_1\in B_1, X_2\in B_2\} \overset{\text{indep}}{=} P\{X_1\in B_1\}\cdot P\{X_2\in B_2\} = \left(\int_{B_1} f_{X_1}(x_1)\,dx_1\right)\left(\int_{B_2} f_{X_2}(x_2)\,dx_2\right)$$

$$\overset{\text{Fubini}}{=} \int_{B_1\times B_2} f_{X_1}(x_1)\,f_{X_2}(x_2)\,dx_1\,dx_2. \qquad \square$$

(b) $\Rightarrow$ (a) Absolute continuity? $\forall B\in\mathcal{B}$:

$$P\{X_1\in B\} = P\{X\in B_1\times\mathbb{R}\} = \int_{B_1\times\mathbb{R}} \underbrace{f_X(x)}_{=\, f_{X_1}(x_1)\cdot f_{X_2}(x_2)}\,dx = \left(\int_{B_1} f_{X_1}(x_1)\,dx_1\right)\underbrace{\left(\int_{\mathbb{R}} f_{X_2}(x_2)\,dx_2\right)}_{=\,1}$$

$\Rightarrow$ abs. continuous, pdf f_1.

Independence?

$$P\{X_1\in B_1, X_2\in B_2\} \overset{\text{similarly}}{=} \left(\int_{B_1} f_{X_1}(x_1)\,dx_1\right)\left(\int_{B_2} f_{X_2}(x_2)\,dx_2\right) = P\{X_1\in B_1\}\cdot P\{X_2\in B_2\}$$

Ex Show that $(*)$ can be replaced by the weaker condition:

$\exists$ integrable functions $f_1,\dots,f_n:\mathbb{R}\to\mathbb{R}$ s.t. $f(x_1,\dots,x_n) = f_1(x_1)\cdots f_n(x_n)$ a.e. on $\mathbb{R}^n$.

HW

EXAMPLES:

(a) Standard normal distr (multivariate):

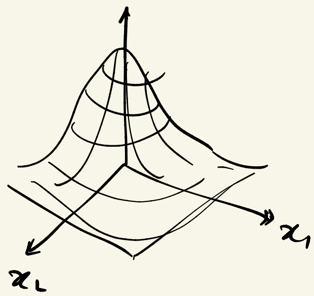
$$X=(X_1,\dots,X_n) \quad \text{where} \quad X_i \sim N(0,1)$$

Joint pdf:

$$f_X(x) = f_{X_1}(x_1)\cdots f_{X_n}(x_n) = \frac{1}{\sqrt{2\pi}} e^{-x_1^2/2} \cdots \frac{1}{\sqrt{2\pi}} e^{-x_n^2/2} = \boxed{\frac{1}{(2\pi)^{n/2}} e^{-\|x\|_2^2/2} \quad \forall x \in \mathbb{R}^n}$$

$x = (x_1,\dots,x_n)$

Def: $\boxed{X \sim N(0, I_n)}$



• Rotation invariance:

$\Downarrow$

$\boxed{\forall\, U \in O(n), \text{ the r.v.'s } X \text{ and } UX \text{ have the same distribution}}$

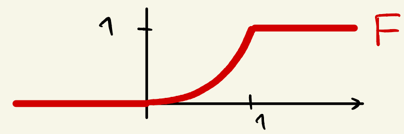
(b) If $X_1,\dots,X_n \sim \text{Unif}[0,1]$, then $X=(X_1,\dots,X_n) \sim \text{Unif}([0,1]^n)$

$\left[\; f(x) = f(x_1)\cdots f(x_n) = \mathbb{1}_{[0,1]}(x_1)\cdots \mathbb{1}_{[0,1]}(x_n) = \mathbb{1}_{[0,1]^n}(x) \;\right]$

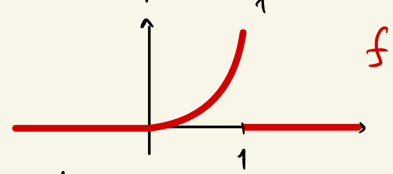
(c) (continued) $\mathbb{E}\,\overbrace{\max(X_1,\dots,X_n)}^{=X_{(n)}} = ?$

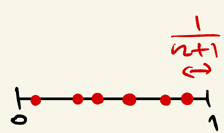
• cdf: $F(x) = \mathbb{P}\{X_{(n)} \le x\} = \mathbb{P}\{X_1 \le x,\dots, X_n \le x\} = \mathbb{P}\{X_1 \le x\}\cdots \mathbb{P}\{X_n \le x\}$

$$= \begin{cases} x^n, & 0 \le x \le 1 \\ 0, & x < 0 \\ 1, & x > 1 \end{cases}$$



• pdf: $f(x) = F'(x) = \begin{cases} n x^{n-1}, & 0 \le x \le 1 \\ 0, & \text{———} \end{cases}$





• $\mathbb{E} X_{(n)} = \int_{-\infty}^{\infty} x f(x)\,dx = \int_0^1 x \cdot n x^{n-1}\,dx = n \int_0^1 x^n\,dx = \boxed{\frac{n}{n+1}}$

(d) (continued) $\mathbb{E} X_{(n-1)} = ?$ ← second largest among $X_1,\dots,X_n$

• pdf: $\mathbb{P}\{X_{(n-1)} \le x\} = \mathbb{P}\{\text{at most one of } X_1,\dots,X_n \text{ is } > x\}$

$$= \mathbb{P}\{\text{all } X_i \le x\} + \sum_{i=1}^n \mathbb{P}\{X_i > x \text{ and } X_j \le x \ \forall j \ne i\}$$

$$= x^n + n\cdot(1-x)x^{n-1}, \quad 0 \le x \le 1.$$

Continue... $\mathbb{E} X_{(n-1)} = \boxed{\frac{n-1}{n+1}}$

Generally, $\boxed{\mathbb{E} X_{(k)} = \frac{k}{n-1}}$ (HW?)

Prop (Joint to marginal density) If $X_1, \dots, X_n$ have (jointly) continuous distr., then

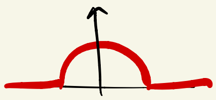
$$f_{X_n}(x_n) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x_1, \dots, x_n)\, dx_1 \cdots dx_{n-1}$$

Proof for $n=2$: $\quad f_X(x) = \int_{-\infty}^{\infty} f(x,y)\, dy$?

$\forall B \in \mathcal{B}$: $\quad \mathbb{P}\{X \in B\} = \mathbb{P}\{(X,Y) \in B \times \mathbb{R}\} = \int_{B \times \mathbb{R}} f(x,y)\, dx\, dy = \int_B \underbrace{\left(\int_{\mathbb{R}} f(x,y)\, dy \right)}\, dx$

Uniqueness of density (from Radon-Nikodym Thm) $\Rightarrow$ pdf of X is this. $\square$

Ex $(X,Y) \sim \mathrm{Unif}(\text{disk}) \Rightarrow$ pdf of X:



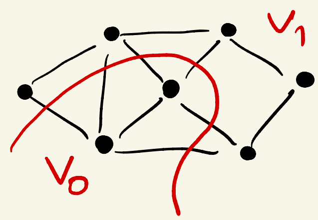
APPLICATIONS

① Finding a large cut in a graph [Alon-Spencer's Book]

Thm $\forall$ graph $G=(V,E)$, the vertices can be partitioned $V = V_0 \sqcup V_1$ s.t.

$$|E(V_0, V_1)| \geq \frac{|E|}{2}$$

Proof Cut the graph at random: put each vertex into V_0 or V_1 independently at random, with prob. $1/2$



$\delta_v \sim Ber(1/2)$ iid; $V_0 = \{v \in V : \delta_v = 0\}$
$V_1 = \{v \in V : \delta_v = 1\}$

$$|E(V_0, V_1)| = \sum_{(u,v)\in E} \mathbb{1}_{\{u \in V_0 \text{ and } v \in V_1, \text{ or vice versa}\}}$$

$$= \sum_{(u,v)\in E} \mathbb{1}\{\delta_u \delta_v = 0\}$$

$$\Rightarrow \mathbb{E}|E(V_0,V_1)| = \sum_{(u,v)\in E} \underbrace{\mathbb{P}\{\delta_u \delta_v = 0\}}_{= 1/2} = \frac{|E|}{2}.$$

$\Rightarrow \exists$ realization s.t. $E(V_0, V_1) \geq \frac{|E|}{2}$ $\square$.

$\left(\mathbb{E}X \geq c \Rightarrow \mathbb{P}\{X \geq c\} > 0 \right)$ HW?

Remark | MaxCut is NP-hard; [Goemans-Williamson]: 0.878-approximation (SDP)
Cover?

② Balancing Vectors (AS)

THM $\forall$ unit vectors $v_1, \dots, v_n \in \mathbb{R}^d$ $\exists$ signs "$\pm$" such that

$$\| \pm v_1 \pm v_2 \pm \dots \pm v_n \|_2 \le \sqrt{n}$$

Remark: optimal for the orthonormal Basis.

Proof Choose the signs at random, independently:

$$P\{\varepsilon_i = 1\} = P\{\varepsilon_i = -1\} = 1/2$$

(Rademacher distribution)

$$X := \left\| \sum_{i=1}^{n} \varepsilon_i v_i \right\|_2$$

$$\mathbb{E}[X^2] = \mathbb{E} \left\langle \sum_{i=1}^{n} \varepsilon_i v_i, \sum_{j=1}^{n} \varepsilon_j v_j \right\rangle = \mathbb{E} \sum_{i,j=1}^{n} \varepsilon_i \varepsilon_j \langle v_i, v_j \rangle = \sum_{i,j=1}^{n} \mathbb{E}[\varepsilon_i \varepsilon_j] \langle v_i, v_j \rangle =$$

- If $i = j$, $\mathbb{E}[\varepsilon_i \varepsilon_j] = \mathbb{E}\, 1 = 1$

  If $i \neq j$, $\mathbb{E}[\varepsilon_i \varepsilon_j] = \mathbb{E}[\varepsilon_i]\, \mathbb{E}[\varepsilon_j] = 0 \cdot 0 = 0$

$$= \sum_{i=1}^{n} \langle v_i, v_i \rangle = \sum_{i=1}^{n} \underbrace{\|v_i\|_2^2}_{=1} = n.$$

$\Rightarrow \exists$ realization $X^2 \le n$. $\square$

Question: if $n \gg d$, is there a Better Bound?