YZV 231E

15. 11. 2021

Probability Theory & Stats

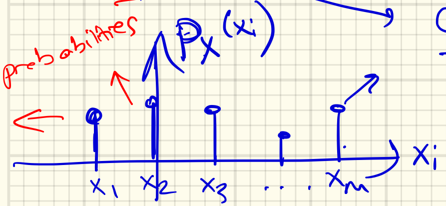
Gözde ÜNAL

Recap: R.v.'s : → Continuous r.v.'s.

A complete description of an r.v. X

Discrete r.v.s

pmfs : → CDFs

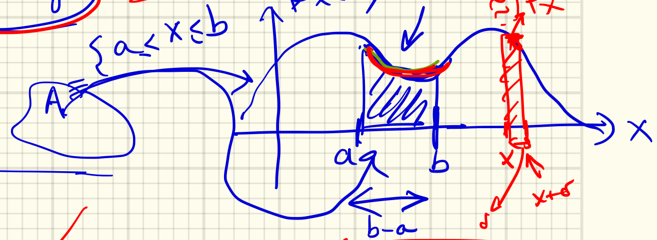


Continuous r.v.s.

pdf : probability / unit

$p_X(x)$

$\{a \le X \le b\}$



$\approx p_X(x)\cdot\delta$

Area ≈ Probability

$\boxed{p_X(x)\cdot\delta}$

δ is small

$$P(a \le X \le b) = \int_a^b p_X(x)\,dx$$

① $p_X(x) \ge 0$

prob. density

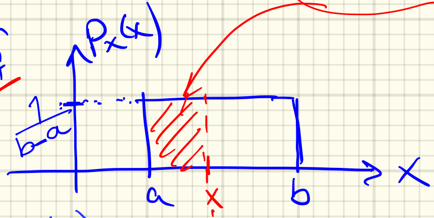
② $\int_{-\infty}^{\infty} p_X(x)\,dx = 1.$

pmf vs pdf

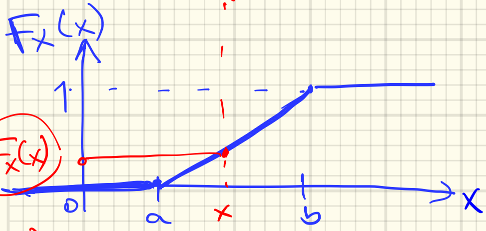
Cumulative Distribution Function (CDFs) : $F_X(x)$: unifies discrete & continuous r.v-s

$$\boxed{F_X(x) \triangleq P(X \leq x)} = \int_{-\infty}^{x} p_X(x)\,dx$$

eg. uniform pdf



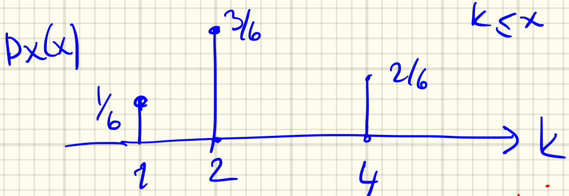
uniform cdf

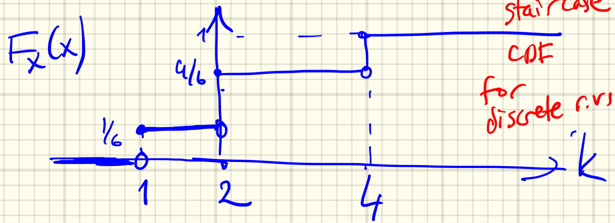


area under $p_X(x)$ for $(-\infty, x)$

Recall : for discrete r.v-s : same defn.

$$F_X(x) = P(X \leq x) = \sum_{k \leq x} p_X[k]$$





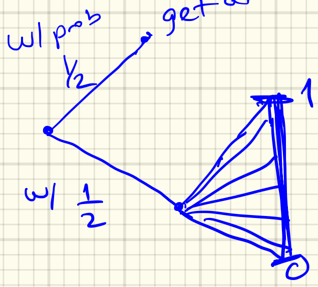
staircase CDF for discrete r.v.s.

$$\frac{d}{dx}F_X(x) = p_X(x)$$

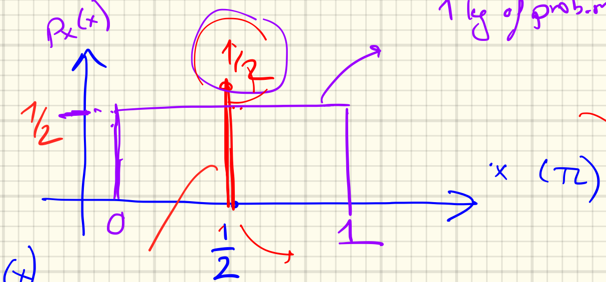
Mixed r.v.s : R.v.s that are a mixture of cont & discrete r.v.s

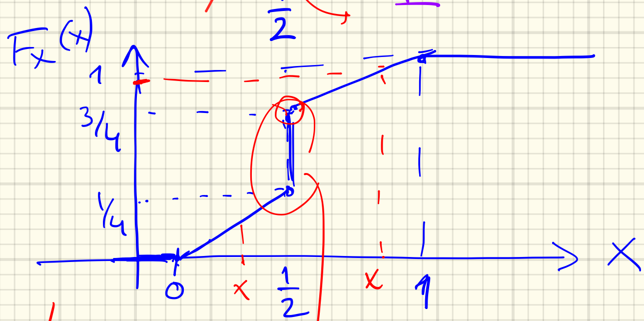
Ex: This is a mixed r.v. → combination of pdf & pmf

1 kg of prob. mass



you call your uniform $U[0,1]$ r.v. generator





$$F_X\left(\frac{1}{2}\right) = \frac{3}{4}$$

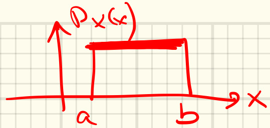
$$P\left(X \le \frac{1}{2}\right) \doteq$$ prob of getting a reward of $\frac{1}{2}$ TL or less

corresponds to discrete r.v.

$\boxed{\text{Generalized function}}$

Important pdfs:

1) Uniform pdf ✓ $X \sim U[a,b]$

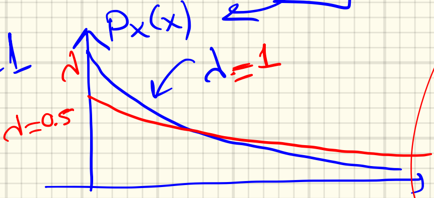


2) Exponential pdf: models lifetime of a product thing.

$X \sim \exp(\lambda)$; $p_X(x) = \begin{cases} \boxed{\lambda e^{-\lambda x}}, & x \geq 0 \\ 0, & o/w \end{cases}$

$p_X(x) \geq 0$

$\int_{-\infty}^{\infty} p_X(x)\,dx = 1$



check $\int_{-\infty}^{\infty} p_X(x)\,dx = 1$ ✓

$E(X) = \int_{-\infty}^{\infty} x \lambda e^{-\lambda x}\,dx$ exercise.

λ: event rate
failure rate.

e.g. $P(X \geq 100)$ = Prob that a device will last 100 days
$\equiv$ or if will fail after 100 days

Choose $\lambda = 0.01$ a parameter ... (exercise find a way to estimate λ) λ: failure rate
: death rate :
event rate

$= 1 - F_X(100) = \int_{100}^{\infty} \lambda e^{-\lambda x}\,dx = -e^{-\lambda x}\Big|_{100}^{\infty} = 0.367$

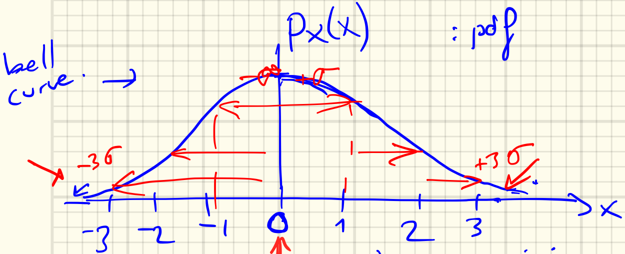
If $\boxed{\lambda = 0.001}$ $\rightarrow P(X \geq 100) = 0.904$

3) Gaussian (Normal) pdf (PDF) : ★★★

Standard Normal $\sim \mathcal{N}(0,1) \Rightarrow p_X(x) = \dfrac{1}{\sqrt{2\pi}} \boxed{e^{-x^2/2}}$, $\quad -\infty < x < \infty$,

↳ normalizing const.

↳ to make integral $= 1$.



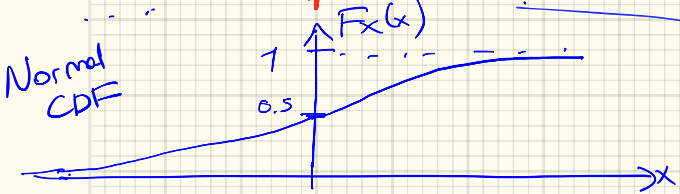
$E[X] = 0$,

$Var(X) = 1$. → calculate this IBP.

$= \sigma^2 = 1$

Std Dev $= \sqrt{Var} = \sigma$.

Normal CDF



$P\left(\underset{2\sigma}{x-3\sigma} \le x \le \underset{2\sigma}{x+3\sigma}\right) = 0.997\ldots$

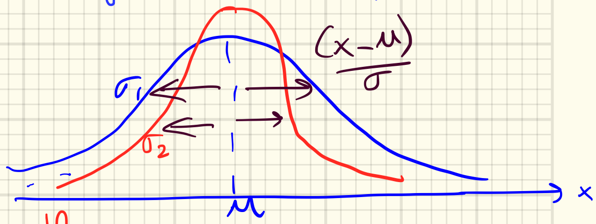
General Normal $\sim \mathcal{N}(\mu, \sigma^2)$

↳ $p_X(x) = \dfrac{1}{\sqrt{2\pi}\,\sigma} e^{-(x-\mu)^2/2\sigma^2}$

$\sigma \searrow$

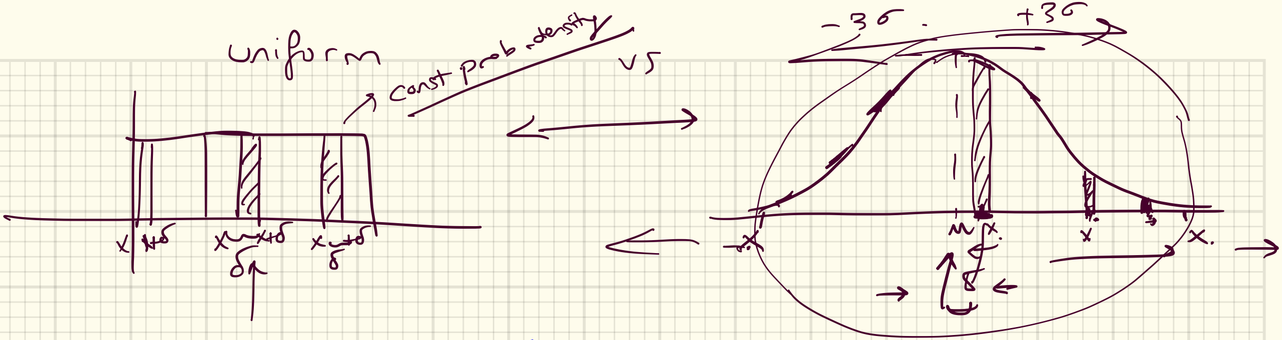
σ: small: narrow pdf

σ large: wide prob density.



$E[X] = \mu$

$Var(X) = \sigma^2$



Fact: $Y = aX + b$, (suppose X is normal).

$$E[Y] = \int y\, p_Y(y)\, dy$$

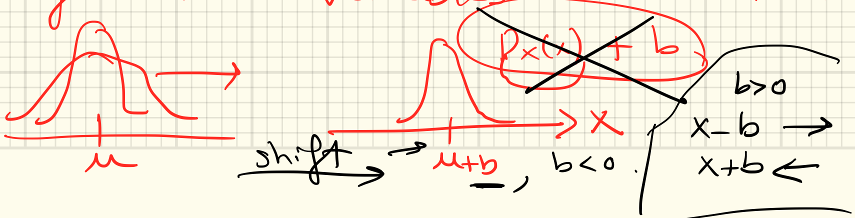
$$E[g(X)] = \int g(x)\, p_X(x)\, dx = \int (ax+b)\, p_X(x)\, dx = a\int_{-\infty}^{\infty} x\, p_X(x)\, dx + \int_{-\infty}^{\infty} b\, p_X(x)\, dx$$

$$E[Y] = a\,E[X] + b.$$

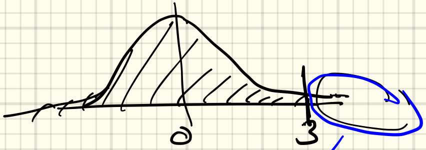
$$Var(Y) = a^2 \sigma^2$$

↳ $Y \sim N(a\mu + b,\ a^2\sigma^2)$: Fact: (We'll show later)

* Linear (Affine) functions of normal variables are also normal



Question: $P(X \leq 3) = ?$ $\quad X \sim N(0,1)$



$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{3} e^{-x^2/2}\, dx$$

: No closed form formula for this! →for the std normal.

Tail prob. $= 1 - P(X \leq 3) = P(X \geq 3)$

Solution: We tabulate it! Lookup tables.

	0.00	0.01	0.02	0.03	... 0.05 ...	0.09
0.0						
0.1						
0.2					0.5987	
⋮						
0.5						
0.60				0.7357		
⋮						

$P(X \leq 0.63) = 0.7357$

Calculating Normal Probabilities

ex: If $X \sim N(2, 16)$

$P(X \leq 3) = ?$

Use: If $X \sim N(\mu, \sigma^2)$

then $\boxed{\frac{X-\mu}{\sigma} \sim N(0,1)}$

This is called Standardization of an R.V.

$X \sim N(2,16) \rightarrow \mu = 2, \ \sigma = 4$

$\{X \leq 3\} \equiv \{\tilde{X} \leq \frac{1}{4}\}$ same event

$P(\ldots) = P(\ldots)$

$P(X \leq 3) = P\left(\frac{X-2}{4} \leq \frac{3-2}{4}\right) = \underset{\Phi}{\mathrm{CDF}}(0.25) = 0.5987$ from the table

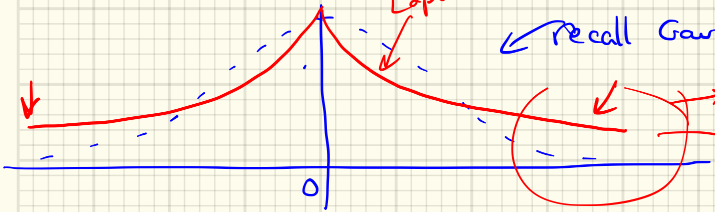
non-std Gaussian

$\tilde{X}$: std Gaussian $\sim N(0,1)$

④ Laplacian pdf: $p_x(x) \underset{\text{prop}}{\propto} e^{-|x|}$

Laplacian

recall Gaussian e^{-x^2}



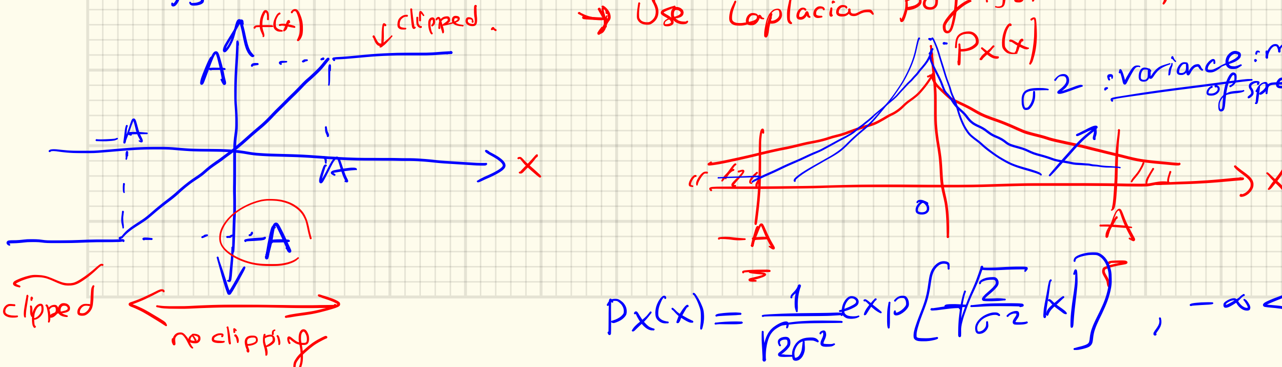
look at the tail probabilities → larger / heavier tails

$$p_x(x) = \frac{1}{\sqrt{2\sigma^2}} \exp\left(-\sqrt{\frac{2}{\sigma^2}}\,|x|\right), \quad \sigma^2 > 0, \quad \text{mean} \Rightarrow E[x] = 0$$

e.g. typical: used as a model for speech amplitudes.

Ex: (10.10) [Kay] Read. Set clipping levels A for speech signal. Speech sent over a xmission channel.

→ Use Laplacian pdf for the speech amplitude



σ^2: variance: measure of spread

$$p_x(x) = \frac{1}{\sqrt{2\sigma^2}} \exp\left[-\sqrt{\frac{2}{\sigma^2}}\,|x|\right], \quad -\infty < x < \infty$$

Design

→ Requirement : Xmit a speech signal w/o clipping 99 % of the time.

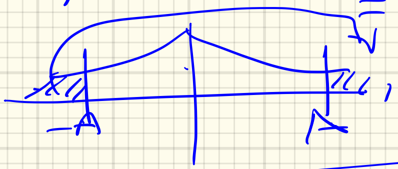
pytorch
random.
rand()
.randn(.)

Clipping occurs when $|x| > A$: Choose A so that

$$P_{clip} \leq 0.01.$$

$P_{clip} = P[X > A \text{ or } X < -A]$ = tail prob.

$= 2P[\cancel{X} > A]$ b/c Laplacian is symmetric around $x=0$



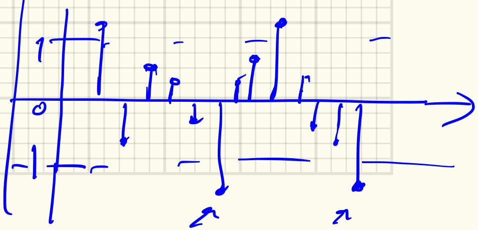
$$\Rightarrow P_{clip} = 2\int_A^\infty \frac{1}{\sqrt{2\sigma^2}} \exp\left(-\sqrt{\frac{2}{\sigma^2}}\,x\right) dx$$

Note: $\sigma^2 \sim$ speech power

A must be selected accordingly

$$P_{clip} = 2\left(-\frac{1}{2}\exp\left(-\sqrt{\frac{2}{\sigma^2}}\,x\right)\right)\Bigg|_A^\infty = \exp\left(-\sqrt{\frac{2}{\sigma^2}}A\right)$$

See Fig 10.32 $A = 1$, $\sigma^2 = 1$



want $P_{clip} \leq 0.01$

$$e^{-\sqrt{\frac{2}{\sigma^2}}A} \leq 0.01 \Longrightarrow -\sqrt{\frac{2}{\sigma^2}}A \leq \ln 0.01$$

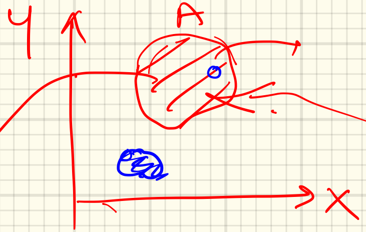
$$\rightarrow A \geq \sqrt{\frac{\sigma^2}{2}} \ln\left(\frac{1}{0.01}\right)$$

Generate uniform r.v. $U[0,1]$: $u_1, u_2, \ldots, u_m$: uniform r.v.

$x \to \underbrace{F_X^{-1}(u)}_{\text{eg Laplacian CDF}} \to x_1, \ldots, x_m$: are Laplacian r.v.

✓ ..

Multiple Joint PDFs : $\boxed{P_{X,Y}(x,y)}$



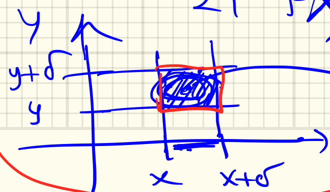
$$P((X,Y)\in A) = \iint_A \underbrace{P_{X,Y}(x,y)}\,dx\,dy$$

$P_{X,Y}(x,y)$ properties (joint pdf)

1) $\displaystyle\int_{-\infty}^{\infty}\int P_{X,Y}(x,y)\,dx\,dy = 1$

2) $P_{X,Y}(x,y) \geq 0$

height $P_{X,Y}(x,y)$ or probability per unit area



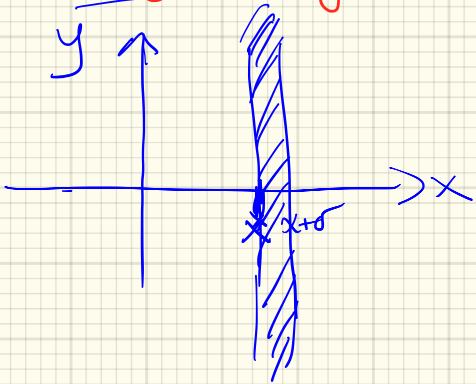
Similar to 1D intuition:

$$P(x \leq X \leq x+\delta,\ y \leq Y \leq y+\delta) \approx \boxed{P_{X,Y}(x,y)}\cdot\underbrace{\boxed{\delta^2}}_{\text{Area}} \implies$$

* Joint pdf (2D) is like Probability per Unit Area.
(in the nbhd of a given pt).

* From the Joint pdf to Marginal pdf.

$$\underbrace{P_X(x)}_{\text{marginal}} = \int_{-\infty\ (y)}^{\infty} P_{X,Y}(x,y)\,dy$$



$$P_Y(y) = \int_{-\infty}^{\infty} P_{X,Y}(x,y)\,dx$$

* X & Y are independent if $P_{X,Y}(x,y) = P_X(x) \cdot P_Y(y)$ $\forall x, y$

(same as in discrete r.v.s)

we factor out marginals from the joint.

* Expectation: $E[g(X,Y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y)\, P_{X,Y}(x,y)\,dx\,dy$

discrete $\left(\sum_x \sum_y g(x,y)\, P_{X,Y}(x,y)\right)$

2 r.v.s X & Y; their joint pdf:

Ex: $$p_{X,Y}(x,y) = \begin{cases} k\cdot(1-|2x-1|)(1-|2y-1|), & 0\le x\le 1,\ 0\le y\le 1 \\ 0, & \text{o/w} \end{cases}$$

What is k so that $p_{X,Y}(x,y)$ is a valid pdf.

$$1 = \int_0^1\int_0^1 k\cdot(1-|2x-1|)\cdot(1-|2y-1|)\,dx\,dy \Rightarrow$$

$$\frac{1}{k} = \int_0^1 (1-|2x-1|)\cdot dx \int_0^1 (1-|2y-1|)\,dy \Rightarrow k=4.$$

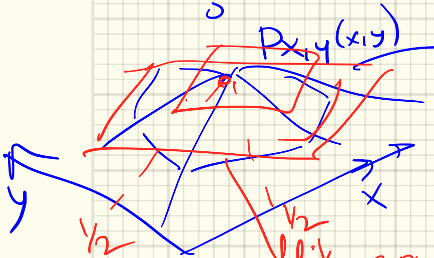
Q. Are X & Y independent? Yes!

$$p_{X,Y} = p_X \cdot p_Y$$

$p_X = 2.(1-|2x-1|)$, p_Y: $x \to y$

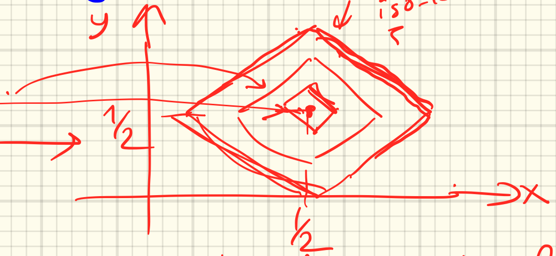
iso-contours

iso-level curves



like a pyramid in 3D.

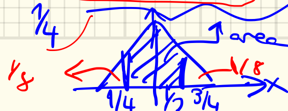
take contours of constant density



2D contour plot of the joint pdf.

$P(A) = ?$ $\quad A = \left\{ \frac{1}{4} \le X \le \frac{3}{4},\ \frac{1}{4} \le Y \le \frac{3}{4} \right\}$

$$\hookrightarrow P(A) = \int_{1/4}^{3/4}\int_{1/4}^{3/4} p_{X,Y}(x,y)\,dx\,dy = \int_{1/4}^{3/4} 2(1-|2x-1|)\,dx \cdot \int_{1/4}^{3/4} 2\cdot(1-|2y-1|)\,dy = \left(\frac{6}{8}\right)^2 = \frac{9}{16}$$



area $= 1-\left(\frac{2}{8}\right) = \frac{6}{8}$

$\rightarrow \; P_X(x) = \int_0^1 4\,\underbrace{(1-|2x-1|)}\,(1-|2y-1|)\cdot dy.$

$P_X(x) = 2\,(1-|2x-1|) \int_0^1 2\,(1-|2y-1|)\,dy$

$= 1$. → use your Calculus 101/102 skills to do this.

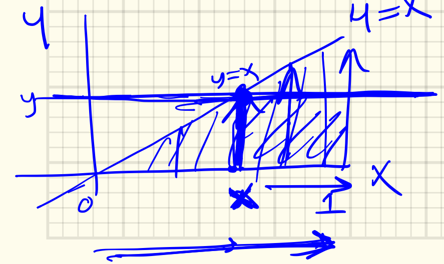
$$P_X(x) = \begin{cases} 2(1-|2x-1|), & 0 \le x \le 1 \\ 0, & o/w \end{cases}$$

→ valid pdf.

show → $$P_Y(y) = \begin{cases} 2(1-|2y-1|), & 0 \le y \le 1 \\ 0, & o/w \end{cases}$$

→ $P_{X,Y}(\;) = P_X(\;)\cdot P_Y(\;)$ ✓

$* \; P[Y \le X] = \int_0^1 \int_0^x p_{X,Y}(x,y)\,dx\,dy$



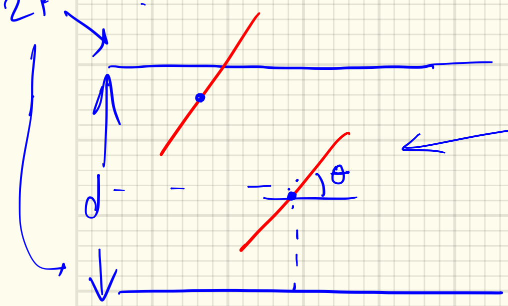
$= \iint 4\cdot(1-|2x-1|)\cdot(1-|2y-1|)\,dx\,dy$

$= \int_0^1 dy \int_{y=x}^1 dx \; P_{X,Y}(x,y) = \ldots$

Ex: Needle of Buffon: Throw needles at random over 2 sticks:

needle length $L < d$: distance between 2 sticks

$$P(\text{needle intersects one of the lines}) = ?$$



2 possibilities:

i) needle does not intersect a line.

ii) needle intersects one of the lines.

Our Prob. modeling Procedure:

1) Set up your sample space S;
2) Describe a prob. law on S
3) Identify the event of interest in S
4) Calculate its probability

1) model the location of the needle & the orientation of the needle.

1) $X \in [0, \frac{d}{2}]$: distance of the midpoint to nearest line.

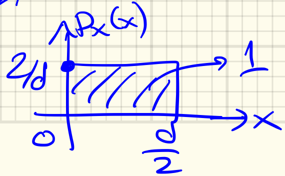
$\theta \in [0, \frac{\pi}{2}]$

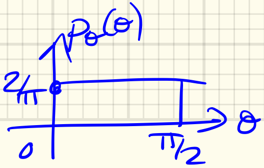
define S

2) Prob Model: X, θ: uniform, $\boxed{\text{independent}}$ intuitive simple model.

$$P_{X,\theta} = p_X(x) \cdot p_\theta(\theta), \quad 0 \le x \le \frac{d}{2}, \quad 0 \le \theta \le \frac{\pi}{2}$$

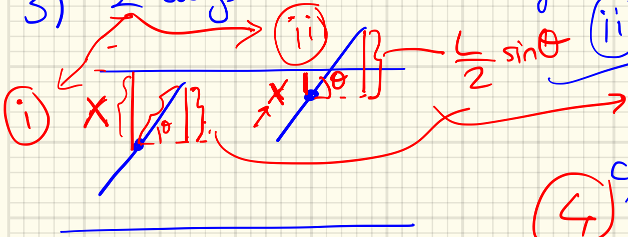
$$= \left(\frac{2}{d}\right)\left(\frac{2}{\pi}\right)$$





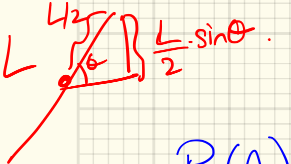
$$p_{X,\theta}(x,\theta) = \frac{4}{\pi d}.$$

3) 2 ways the needle can fall



(ii) ↳ $P\left(X \le \underbrace{\frac{L}{2}\sin\theta}\right)$

A = event defined ✓ (3)



(4) Calculate:

$$P\left(X \le \frac{L}{2}\sin\theta\right) = \iint_{\{X \le \frac{L}{2}\sin\theta\}} p_X(x)\cdot p_\Theta(\theta)\,dx\,d\theta$$

$$P(A) = \frac{4}{\pi d}\int_0^{\pi/2}\int_0^{\frac{L}{2}\sin\theta} dx\,d\theta = \frac{4}{\pi d}\int_0^{\pi/2}\frac{L}{2}\sin\theta\,d\theta$$

$$= \frac{\cancel{4}^{\,2}}{\pi d}\,\frac{L}{\cancel{2}}\cdot\left(\underbrace{-\cos\theta\Big|_0^{\pi/2}}_{1}\right)$$

$$\underbrace{P(\text{needle intersecting a line})} = \frac{2L}{\pi d}$$

Historical Note: They used this to calculate an approx value for π

→ Monte Carlo Method to evaluate integrals → later.

eg. They threw 1000 needles & calculated P(A) and knowing L & d,
$\pi = \frac{2L}{P(A)\,d}$