Week #4

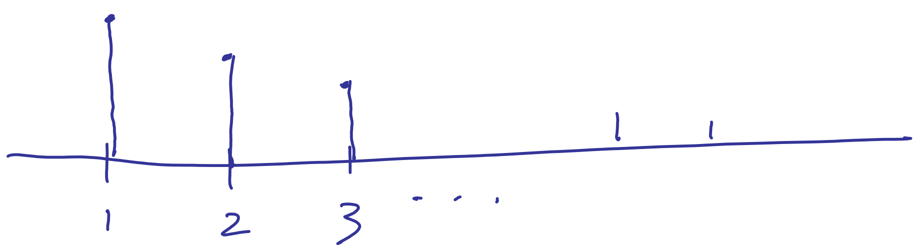
iid Bernoulli(p) rv's

$X_1, X_2, \dots$

Y = "time" until the first 1

$\sim$ geometric(p)

$$P(Y=k) = q^{k-1}p \quad , \quad k=1,2,\dots$$



$$f(x) = q^{x-1}p \quad , \quad x=1,2,\dots$$
$$= 0 \quad , \quad \text{ow}$$

$$P(Y\in B) = \sum_{y\in B} f(y)$$

$$G(s) = \underbrace{\sum_{\text{all } y} s^y f(y)}_{E(s^Y)} = \sum_{k=1}^{\infty} s^k q^{k-1} p$$

$$= \frac{p}{q} \sum_{k=1}^{\infty} (qs)^k$$

$$= \frac{p}{q} \frac{qs}{1-qs} = \frac{ps}{1-qs} \quad , \quad |s| < \frac{1}{q}$$

$\underline{Var(Y)?}$ $\left(= E(Y-\underset{E(Y)}{\underset{\uparrow}{\mu}})^2 = E(Y^2) - [E(Y)]^2 \right)$

$$\boxed{\begin{array}{l} \underline{\text{Lemma}}\ \text{(i)}\ E(aX+b) = a\,E(X) + b \\ \quad \text{(ii)}\ Var(aX+b) = a^2\,Var(X) \end{array}}$$

Look at $W = Y - 1$. Then

$$E(W) = E(Y) - 1 \quad , \quad Var(W) = Var(Y)$$

The pgf of W is

$$G_W(s) = E(s^W) = E(s^{Y-1})$$

$$= \frac{1}{s} E(s^Y)$$

$$= \frac{1}{s} G_Y(s)$$

$$\Rightarrow \quad G_W(s) = \frac{p}{1-qs}, \quad |s| < \frac{1}{q}$$

$$\Rightarrow \quad G_W^{(1)}(s) = \frac{pq}{(1-qs)^2} \overset{\text{to get } E(W)}{\underset{s=1}{=}} \frac{pq}{(1-q)^2} = \frac{q}{p}$$

$$\left\{ E(Y) = 1 + E(W) = 1 + \frac{q}{p} = \frac{1}{p} \right\}$$

$$G^{(2)}(s) = \frac{2pq^2}{(1-qs)^3} \underset{s=1}{=} \frac{2q^2}{p^2}$$

to get $E(W(W-1))$

So

$$\left. \begin{aligned} E(W) &= q/p \\ E(W^2) - E(W) &= 2q^2/p^2 \end{aligned} \right\} \rightarrow E(W^2) = \frac{2q^2}{p^2} + \frac{q}{p}$$

$$\Rightarrow Var(Y) = Var(W) = \frac{2q^2}{p^2} + \frac{q}{p} - \frac{q^2}{p^2}$$

$$= \frac{q^2}{p^2} + \frac{q}{p} = \frac{q^2 + pq}{p^2} = \frac{q}{p^2}$$

Note The mgf is

$$m(t) = E(e^{tY})$$

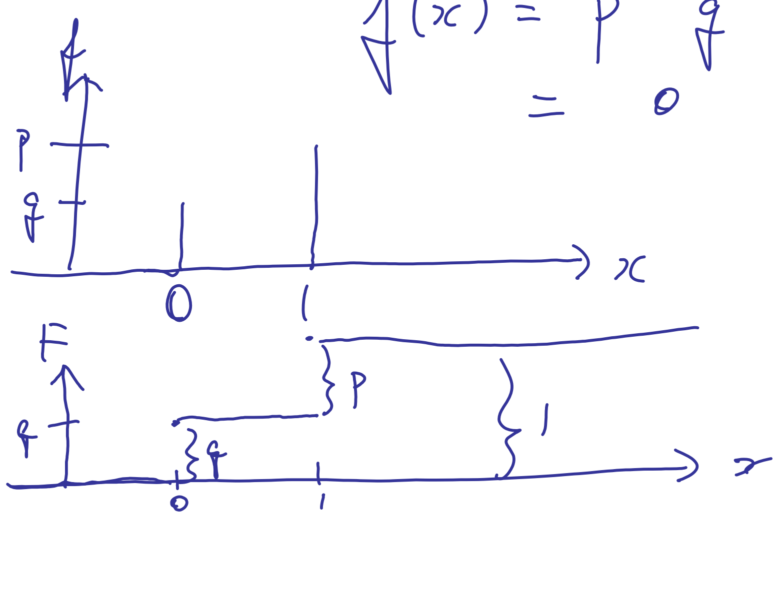
- $m(0) = 1$
- $m^{(k)}(0) = E(Y^k)$
- $m(t) = G(e^t)$ for counting rv's

pgf, mgf, $f(x)$ are representatives of the probability dist'n. Are there others? Yes, quite a few useful f'ns.

Dist'n Function $F(x) = P(X \le x)$, $\forall x$

eg $X \sim$ Bernoulli(p)

$$f(x) = p^x q^{1-x}, \quad x = 0, 1$$
$$= 0, \quad ow$$



Properties of all df's

- F is increasing (nondecreasing)
- $\lim_{x \to \infty} F(x) = 1$, $\lim_{x \to -\infty} F(x) = 0$
- F is right cts

$X \geq 0$ & $E(X) = 0 \implies P(X = 0) = 1$

We need more tools.

Proposition $X \leq Y \implies E(X) \leq E(Y)$

Proof $X \leq Y \implies Y - X \geq 0$

$$\implies E(Y - X) \geq 0$$

$$\implies E(Y) - E(X) \geq 0$$

$$\implies E(Y) \geq E(X)$$

qed

Application (Boole's Inequality)

events $A_1, A_2, \ldots$

$$P\left(\bigcup_k A_k\right) \le \sum_k P(A_k)$$

<u>Proof:</u> $I\left(\bigcup_k A_k\right) \le I(A_1) + I(A_2) + \cdots$

$\{$ ie $I\left(\bigcup_k A_k\right)(s) \le I(A_1)(s) + I(A_2)(s) + \cdots, \forall s \in S \}$

$$\Rightarrow E\left[I\left(\bigcup_k A_k\right)\right] \le E\left[\sum_k I(A_k)\right] = \sum_k E\{I(A_k)\}$$

$$\Rightarrow P\left(\bigcup_k A_k\right) \le \sum_k P(A_k)$$

Markov's Inequality

If $c > 0$ is a constant then

$$P(|X| \geq c) \le \frac{E(|X|)}{c}$$

Aside $P(|X-\mu| \geq k\sigma) \leq \frac{1}{k^2}$ $\boxed{\text{Chebyshev's Inequality}}$

Follows from Markov as

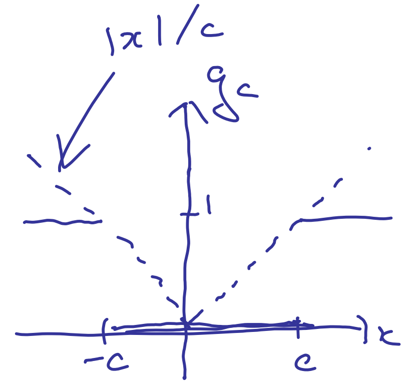
$$P(|X-\mu| \geq k\sigma) = P((X-\mu)^2 \geq k^2\sigma^2)$$

$$\leq \frac{E(X-\mu)^2}{k^2\sigma^2} = \frac{\sigma^2}{k^2\sigma^2} = \frac{1}{k^2}$$

Proof of Markov

$$P(|X| \geq c) = E[I(|X| \geq c)]$$

$$= E[g_c(X)],$$



where

$$g_c(x) = 1, \quad |x| \geq c$$
$$= 0, \quad |x| < c,$$

But

$$g_c(x) \leq \frac{|x|}{c}, \quad \forall x$$

+ so

$$g_c(X) \leq \frac{|X|}{c} \Rightarrow E[g_c(X)] \leq \frac{E(|X|)}{c}$$

$$\Rightarrow \quad P(|X| \geq c) \leq \frac{E(|X|)}{c}$$

~~qed~~

<u>Problem</u> $X \geq 0$ & $E(X) = 0 \Rightarrow P(X=0) = 1$

<u>Sol'n</u> $0 \leq P(X>0) = P\left(\bigcup_{m=1}^{\infty} \{X \geq \frac{1}{m}\}\right)$

$$\leq \sum_{m=1}^{\infty} P\left(X \geq \frac{1}{m}\right) \qquad \text{(Boole)}$$

$$\leq \sum_{m=1}^{\infty} \frac{E(X)}{\left(\frac{1}{m}\right)} \qquad \text{(Markov)}$$

$$= 0$$

$$\Rightarrow \quad P(X>0) = 0 \quad \Rightarrow P(X=0) = 1$$

<u>Consequence</u> Any rv X with $Var(X) = 0$.

This says $E[(X-\mu)^2] = 0$

$$\Rightarrow (X-\mu)^2 = 0 \quad , \underbrace{\text{wp1}}_{\text{with prob } 1}$$

$$\Rightarrow X \overset{\text{wp1}}{=} \mu$$

ie X is constant wp1.

Back to the cdf / df F

X ; $F(x) = P(X \le x)$, $-\infty < x < \infty$,

$\lim_{x \downarrow a} F(x) = F(a)$ — F is right cts at a.

$\lim_{x \uparrow a} F(x) = P(X < a)$ — clearly?

will show

Now $P(X \le a) - P(X < a) = P(X = a)$ = size of jump in F at a

If F is cts at a then $P(X=a) = 0$.

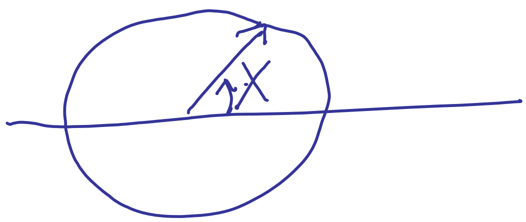
<u>Note</u> $F(b) - F(a) = P(X \le b) - P(X \le a)$
$(a < b)$
$= P(a < X \le b)$

$\because \{X \le a\} \cup \{a < X \le b\} = \{X \le b\}$

$$= \underbrace{P(X \le a)}_{F(a)} + P(a < X \le b) = \underbrace{P(X \le b)}_{F(b)}$$

$$\Longrightarrow \boxed{P(a < X \le b) = F(b) - F(a)}$$

eg



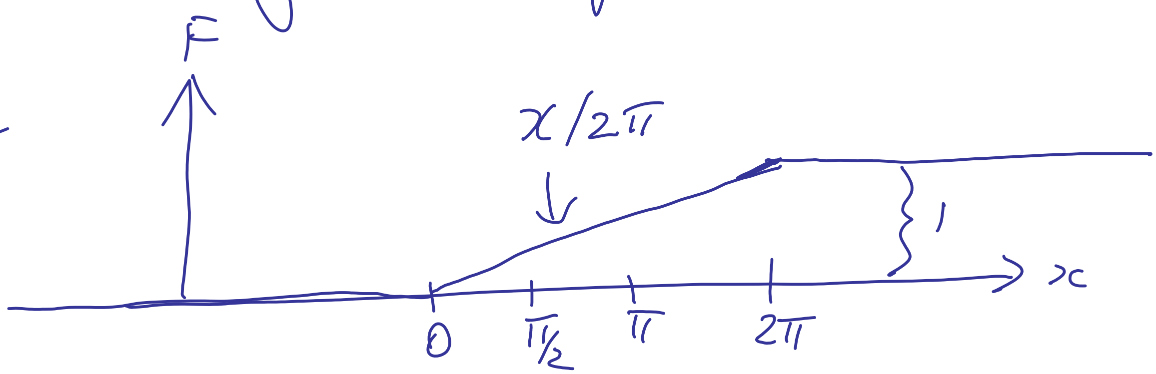
Spin needle
No preference
Observe $X \in [0, 2\pi)$

It makes sense for

$$\left.\begin{array}{l} P(0 < X \le \pi/2) \\ P(0 < X < \pi/2) \end{array}\right\} = \frac{1}{4}$$

which only holds if $P(X = \pi/2) = 0$

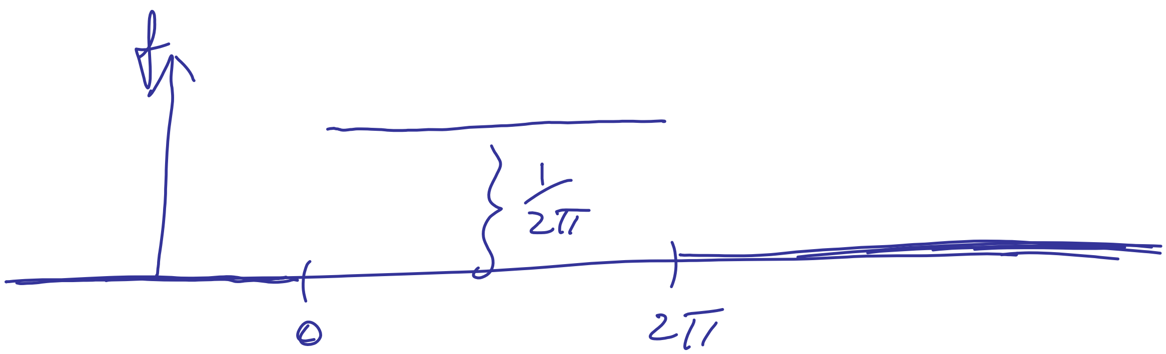
df



$$P\left(\frac{\pi}{2} < X \le \pi\right) = F(\pi) - F\left(\frac{\pi}{2}\right)$$
$$= \frac{1}{2} - \frac{1}{4} = \frac{1}{4}$$

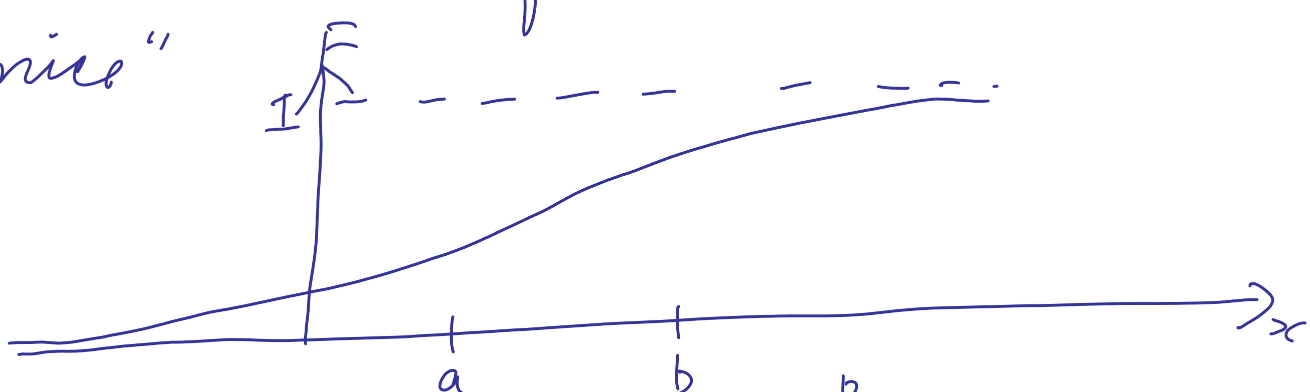
This is the df of a uniform $([0, 2\pi))$ r.v.

Look at $\boxed{F'(x) = f(x)}$



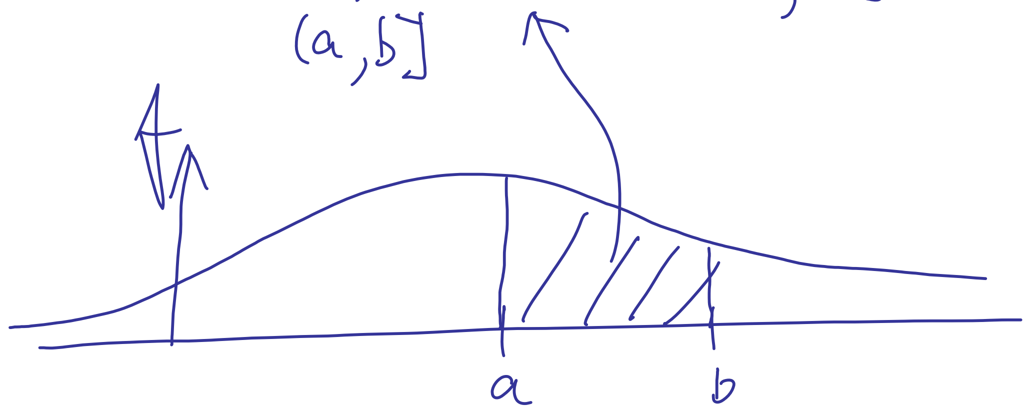
areas under the graph yield probabilities

More generally, if F is a cts df with derivative $f(x) = F'(x)$ which is "nice"



$$P(a < X \le b) = F(b) - F(a) = \int_a^b F'(x)\,dx$$

$$= \int_{(a,b]} f(x)\,dx = \text{area over } (a,b] \text{ \& under } f$$



<u>Note</u>

$$\left.\begin{array}{l} f(x) \geq 0 \\ \underbrace{\text{area under } f}_{\int_{-\infty}^{\infty} f(x)\,dx} = 1 \end{array}\right\} \text{pdf}$$

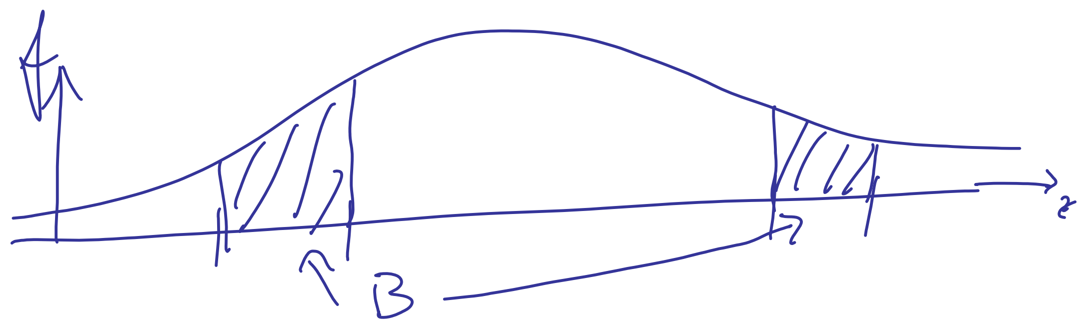
eg Here are 2 pdf's

$$\left.\begin{array}{ll} f(x) = 1, & 0 < x < 1 \\ \quad = 0, & \text{ow} \end{array}\right\} \text{uniform}(0,1) \text{ pdf}$$

$$\left.\begin{array}{ll} f(x) = e^{-x}, & x > 0 \\ \quad = 0, & \text{ow} \end{array}\right\} \text{standard exponential pdf}$$

So for a cts rv X with pdf f

$$P(X \in B) = \int_B f(x)\,dx$$



Fact The df F determines the probability dist'n (ie if we know F then we can calculate any probability)