Chapter 3 Joint distributions

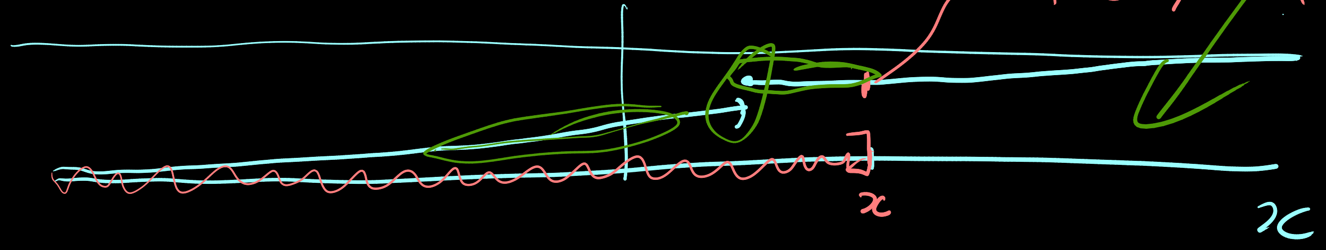
2 or more random variables

E.G.
- Height & Weight
- Dose of a drug & time to recovery
- Gender, Age, Education, Income

CDF One RV: $F(x) = P(X \leq x)$

$F : \mathbb{R} \to [0,1]$

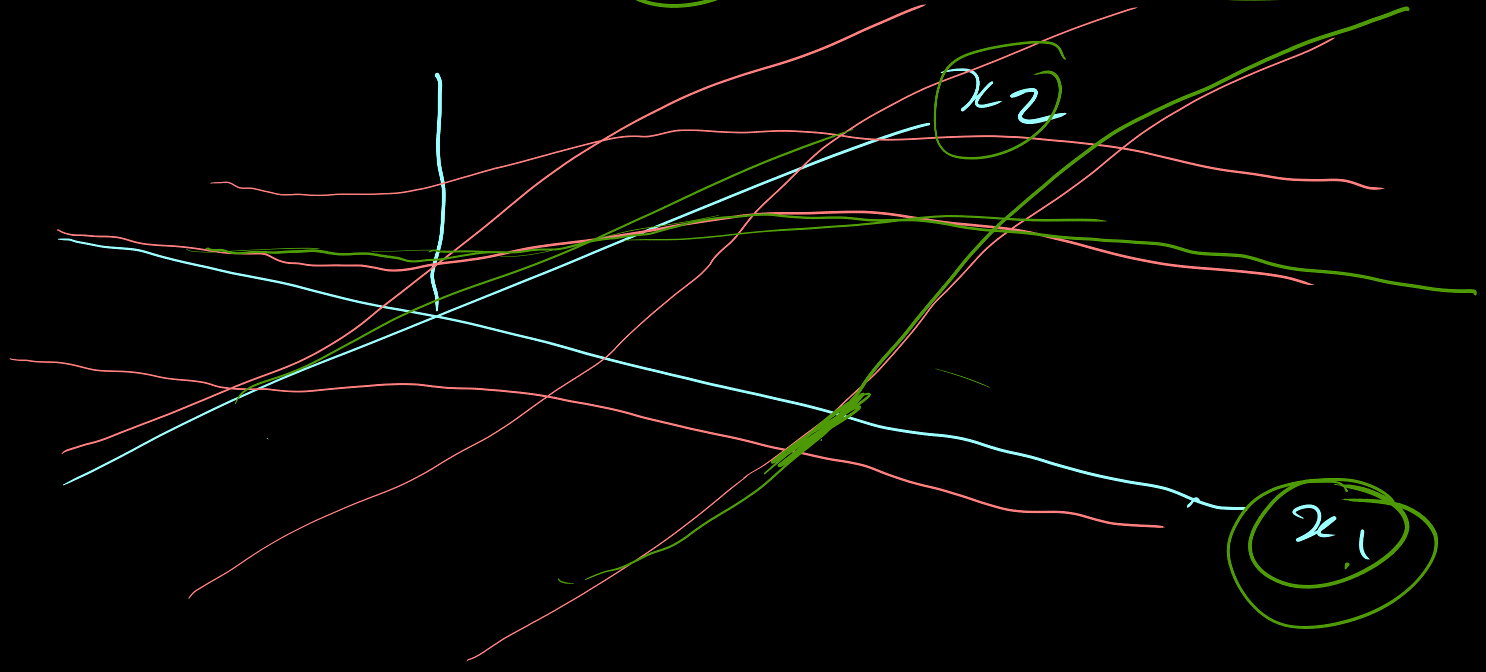
$F(x) = P((-\infty, x])$



CDF: Two R.V.s X_1, X_2

$$F(x_1, x_2) = P(X_1 \le \overset{x_1}{\cancel{x_2}} \text{ and } X_2 \le x_2)$$

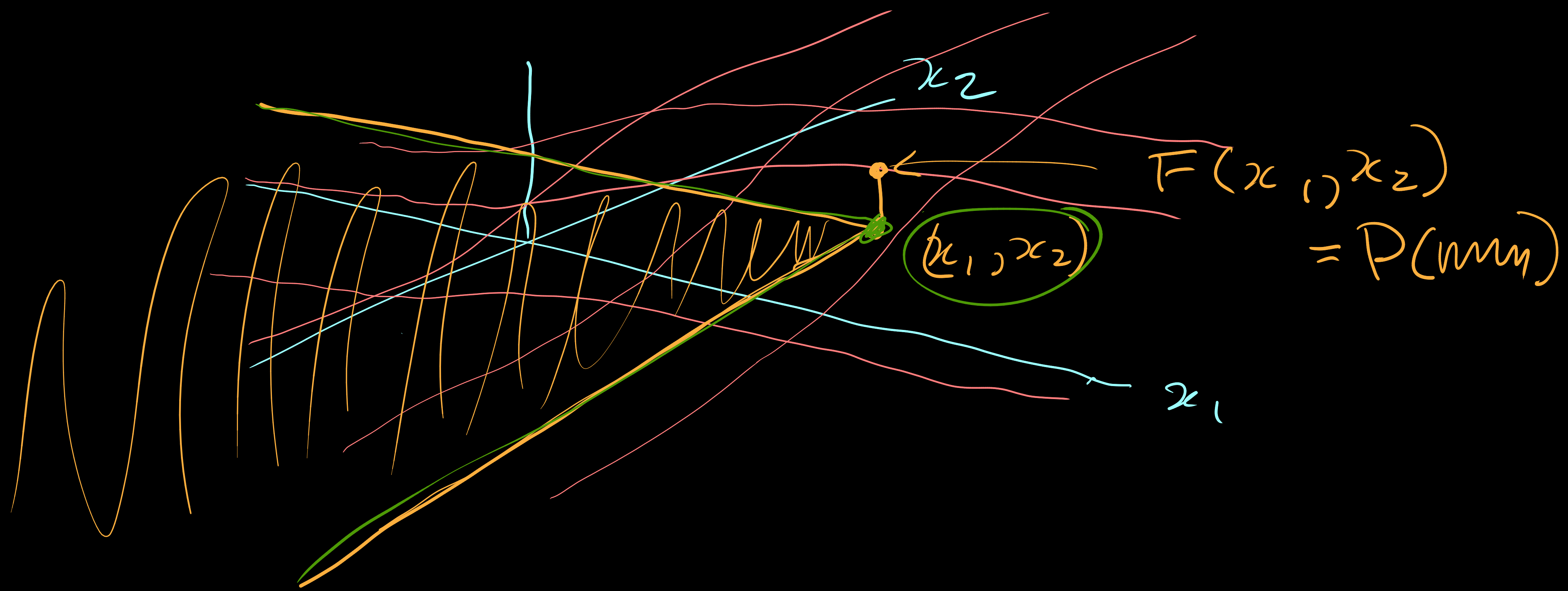
$$F : \mathbb{R}^2 \to [0, 1]$$



CDF : Two R.V.s X_1, X_2

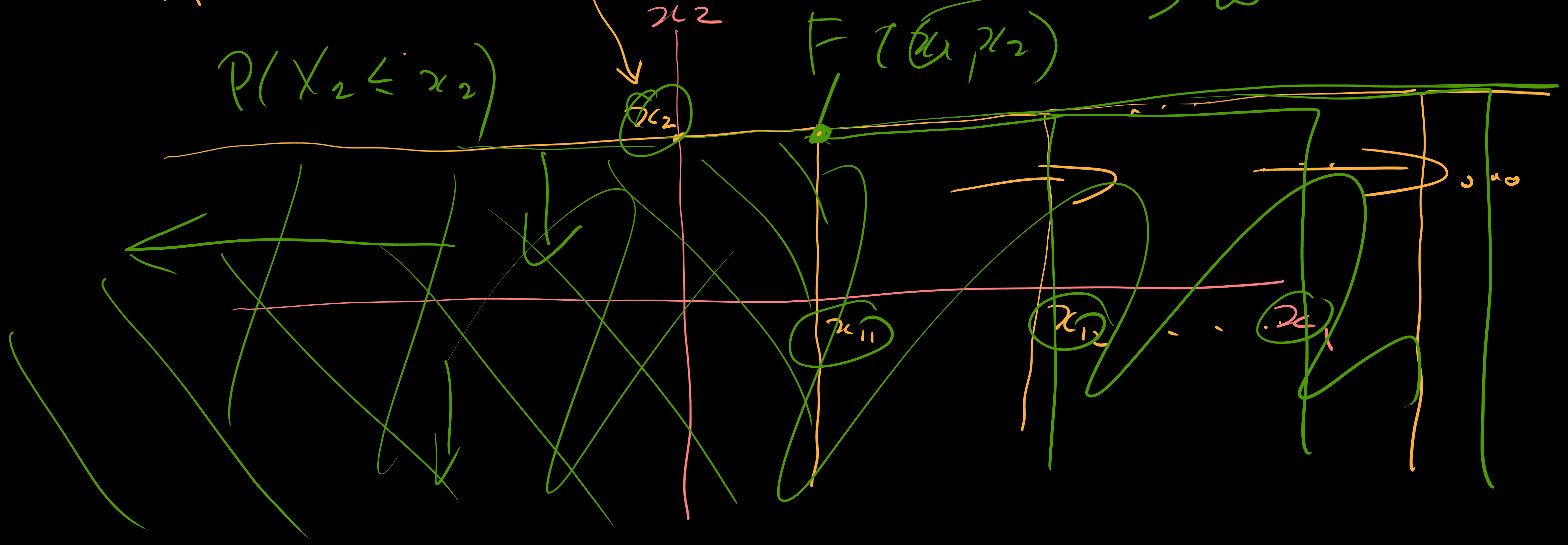
$$\boxed{F(x_1, x_2) = P(X_1 \le x_2 \text{ and } X_2 \le x_2)}$$

$$F : \mathbb{R}^2 \to [0, 1]$$



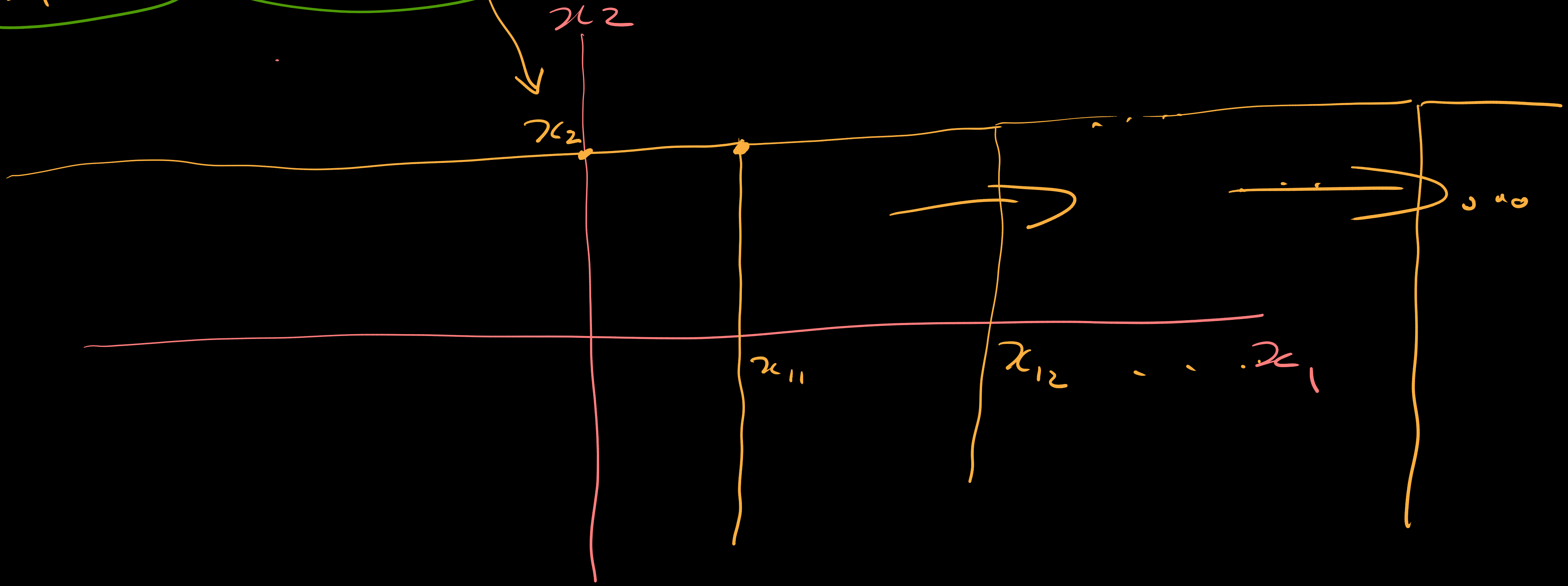
If you fix x_2 and let $x_1 \to \infty$

$$\lim_{x_1 \to \infty} F(x_1, x_2) = \underline{\hspace{6cm}}$$



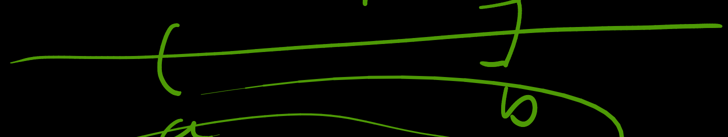
If you fix x_2 and let $x_1 \longrightarrow \infty$

$$\lim_{x_1 \to \infty} F(x, x_2) = F_2(x_2)$$

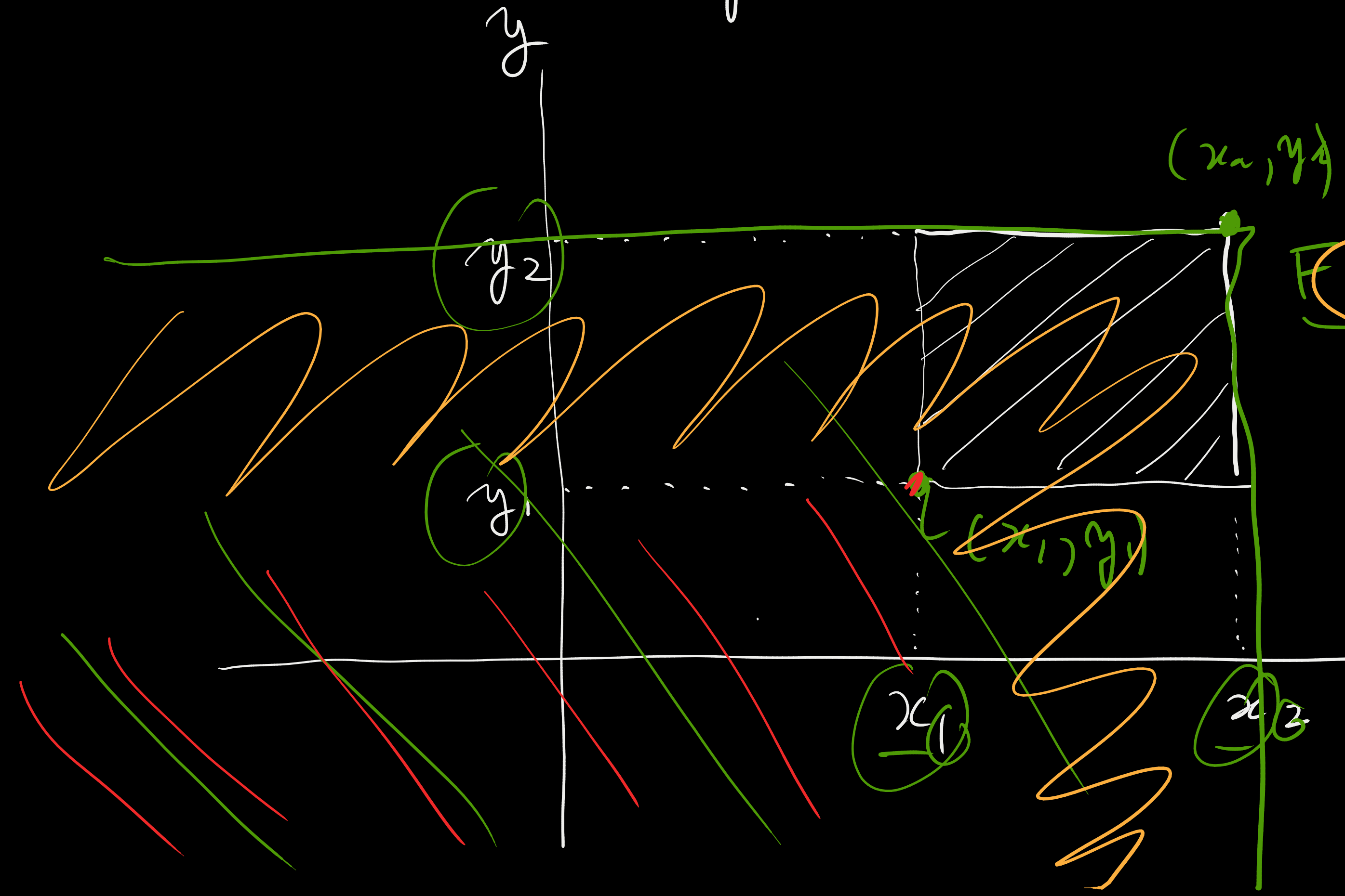


Similarly $\lim_{x_2 \to \infty} F(x_1, x_2) = F_1(x_1)$

Probability of a rectangle:



$F(b) - F(a)$

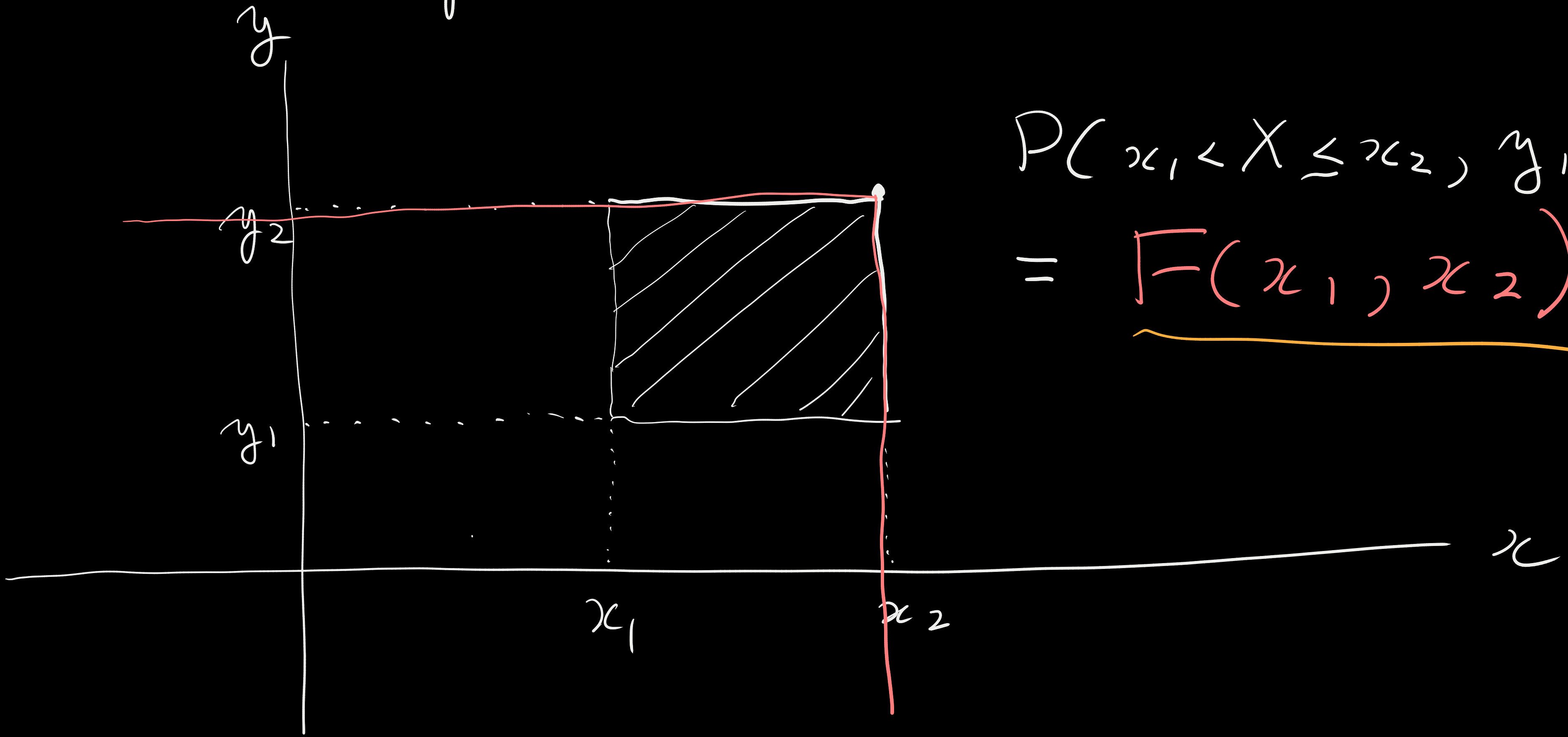


(x_2, y_2) $P(x_1 < X \le x_2,\ y_1 < Y \le y_2)$

$F(x_2, y_2) - F(x_1, y_1)$

Similarly $\lim\limits_{x_2 \to \infty} F(x_1, x_2) = F(x_1)$

Probability of a rectangle:

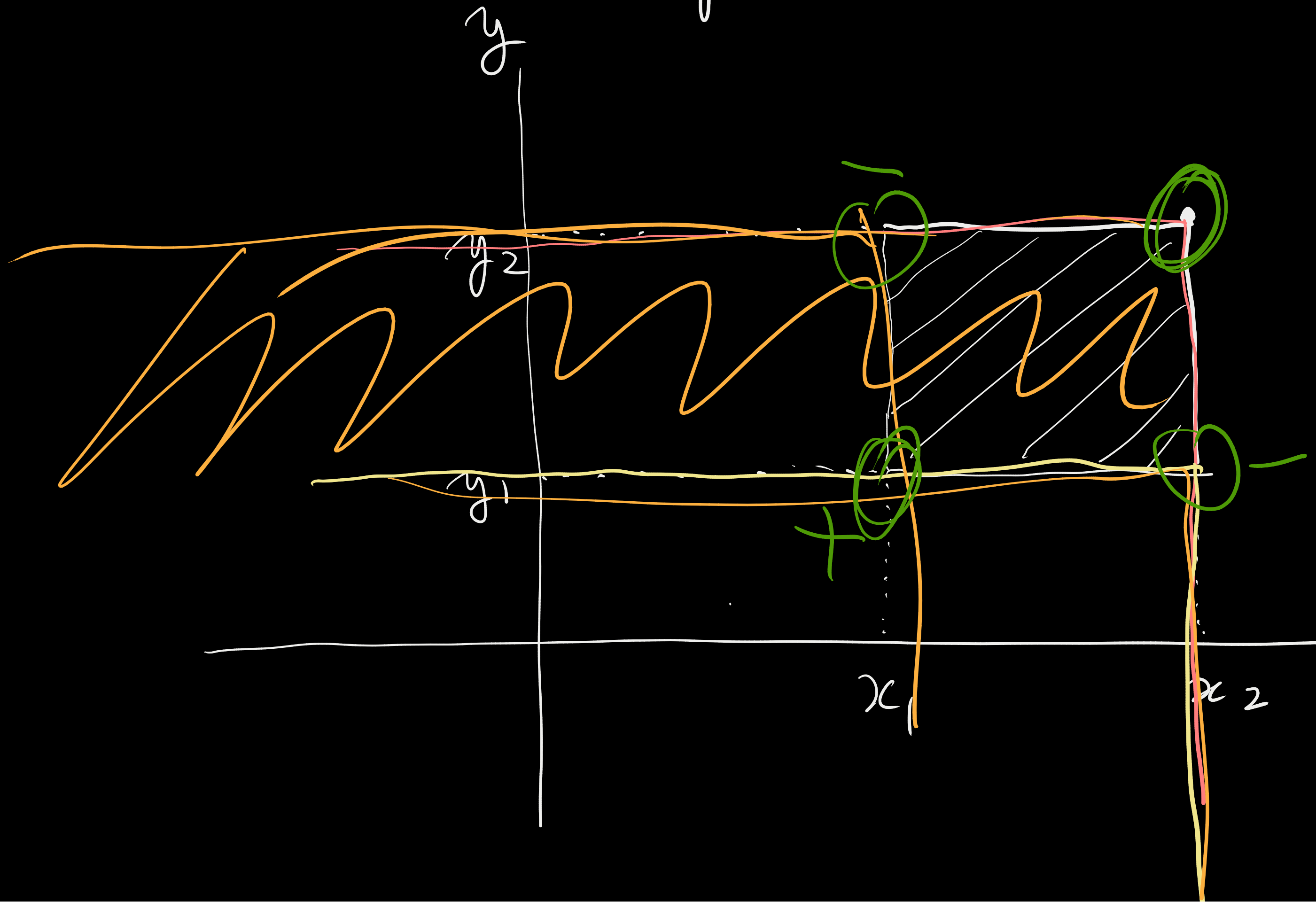


$P(x_1 < X \le x_2,\ y_1 < Y \le y_2)$

$= F(x_1, x_2)$

Similarly $\lim_{x_2 \to \infty} F(x_1, x_2) = F(x_1)$

Probability of a rectangle:



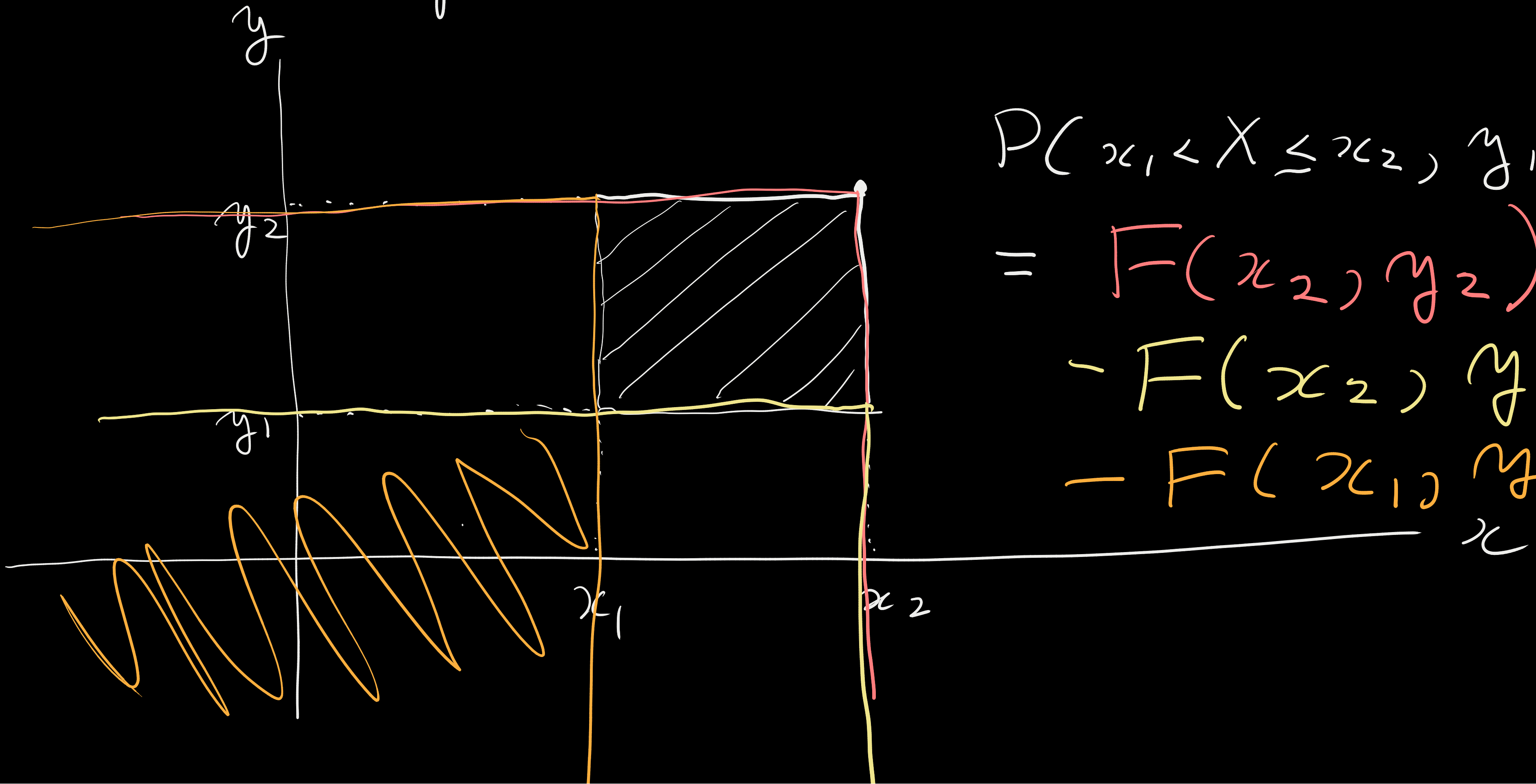
$P(x_1 < X \le x_2,\ y_1 < Y \le y_2)$

$= F(x_2, y_2)$

$- F(x_2, y_1)$

Similarly $\lim_{x_2 \to \infty} F(x_1, x_2) = F(x_1)$

Probability of a rectangle:

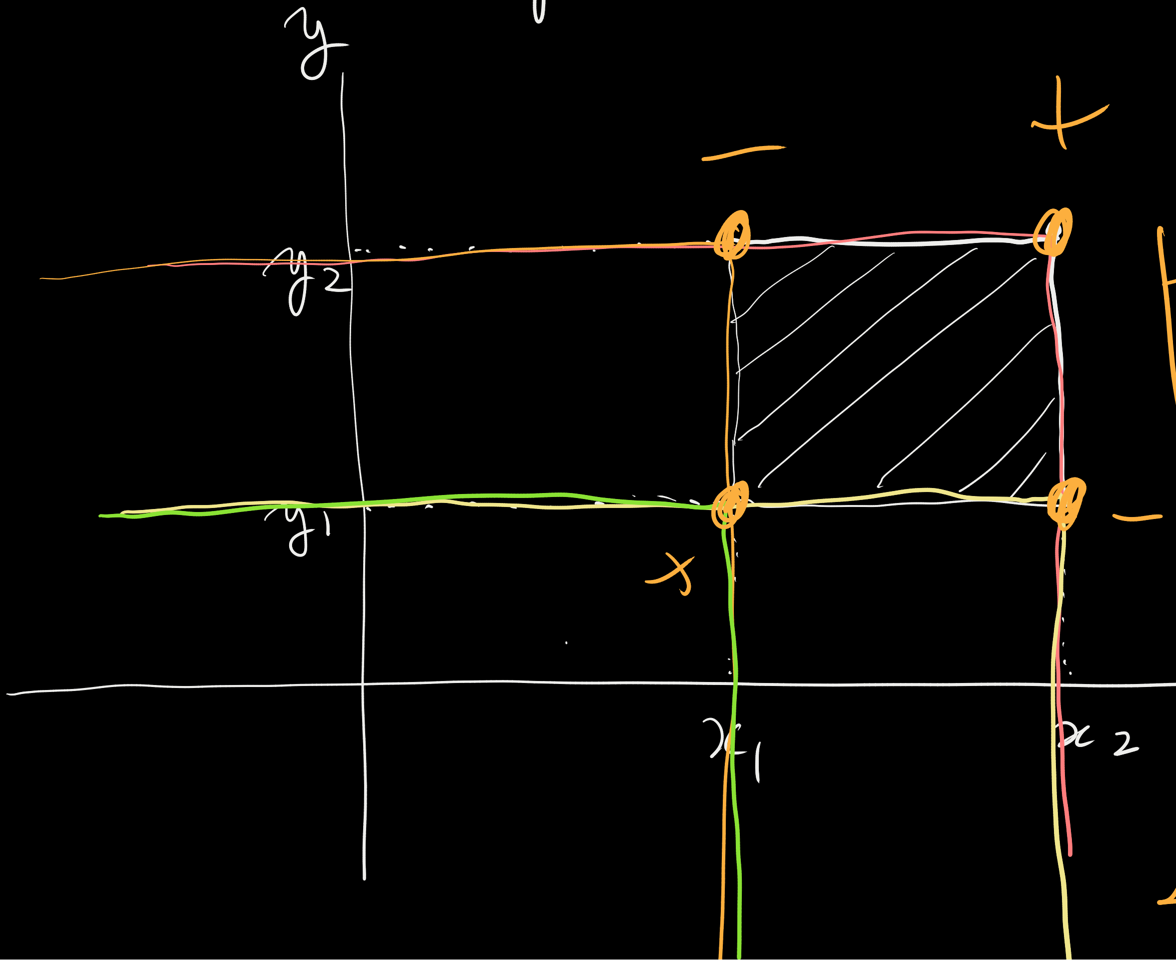


$$P(x_1 < X \le x_2,\ y_1 < Y \le y_2)$$

$$= F(x_2, y_2) - F(x_2, y_1) - F(x_1, y_2)$$

Similarly $\lim_{x_2 \to \infty} F(x_1, x_2) = F(x_1)$

Probability of a rectangle:



$$P(x_1 < X \leq x_2,\ y_1 < Y \leq y_2)$$

$$\boxed{\begin{aligned} = &\ F(x_2, y_2) \\ &- F(x_2, y_1) \\ &- F(x_1, y_2) \\ &+ F(x_1, y_1) \end{aligned}}$$

Recap :

$$P(x_1 < X \le x_2 , y_1 < Y \le y_2)$$

$$= F(x_2, y_2) - F(x_2, y_1) - F(x_1, y_2) + F(x_1, y_1)$$

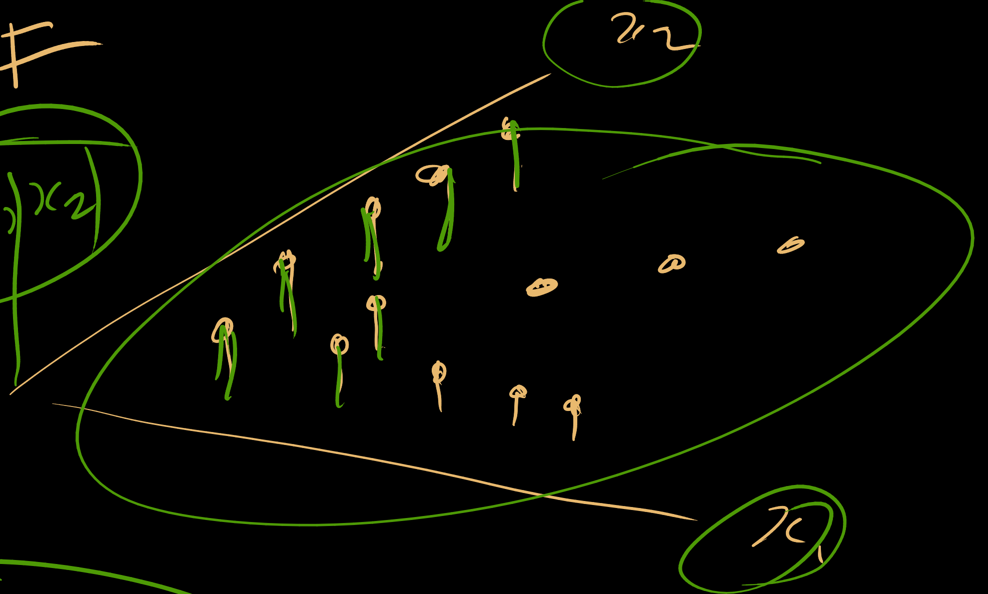
Discrete random variables

Joint PMF $\underline{P(X_1 = x_1, X_2 = x_2, X_3 = x_3)}$

$$= p(x_1, x_2, x_3)$$

Joint PMF

$p(x_1, x_2)$



If $\text{supp}(X_1, X_2)$ is finite, we can use a table.

e.g. Toss a coin: X = # of Heads & a die, Y = # rolled.

Joint PMF

$X \backslash Y$	1	2	3	4	5	6	
0	1/12	1/12	1/12	1/12	1/12	1/12	1/2
1	1/12	1/12	1/12	1/12	1/12	1/12	1/2
	1/6	1/6	1/6	1/6	1/6	1/6	1

marginal PMF for X

$P(\Omega)$

marginal for Y

In general: Joint PMF

$$p(x_1, x_2, x_3, x_4)$$

To get marginal PMF just sum over the variables you don't want:

$$p(x_1, x_3) = \sum_{\substack{\text{all } x_2\text{'s} \\ \text{all } x_4\text{'s}}} p(x_1, x_2, x_3, x_4)$$

Example: Multinomial distribution

Consider $X \sim \text{Binomial}(n, p)$ H T A

$P(H) = p$ $P(T) = 1 - p = q$

Toss n times   H H H T H H H T H H $(n = 10)$

8 Hs and 2Ts $(8, 2)$

X_1 X_2

So $X = 8$

But we could also record # of Ts, 2

and think of this as a joint distribution

for $X_1 = \#$ of Hs, and $X_2 = \#$ of Ts $X_1 + X_2 = n$

H T T H A T T H A

$n = 9$

$(3,\ 4,\ 2)$

$x_1 \quad x_2 \quad x_3$

$x_4 + x_T + x_A = 9$

$$\binom{9}{3\ 4\ 2} = \frac{9!}{3!\,4!\,2!}$$

1 2 1 3 4 0 1 4 2 1

$n = 10$

$x_1 \quad x_2 \quad x_3 \quad x_4 \quad x_5 \quad x_6$

4 2 1 2 0 1 $= 10$

Here $(X_1, X_2) = (8, 2)$

Of course $X_1 + X_2 = n$

Here's the joint distribution for $n = 2$

X_1 (Heads) \ X_2 (Tails)	0	1	2
0	0	0	1/4
1	0	1/2	0
2	1/4	0	0

H	T
T	T
H	T
T	H

$(X_1, X_2) \sim \text{Multinomial}(2, p = (1/2, 1/2))$

This generalized to any number of categories.

E.G. Take a sample of n students and record eye color: Black, Hazel, Blue, Gray

Counts: $X_1 \quad X_2 \quad X_3 \quad X_4$

$$(X_1, X_1, X_2, X_4) \sim \text{Multinomial}\left(n, (p_1, p_2, p_3, p_4)\right)$$

proportions of B, H, B, G in population

Or. $\underset{\sim}{X} \sim \text{Multinomial}(n, \underset{\sim}{p})$

PMF:

$$p(x_1, \dots, x_n) = \binom{n}{x_1\, x_2 \cdots x_n} p_1^{x_1} p_2^{x_2} \cdots p_n^{x_n} \quad \left(\frac{n!}{x_1!\, x_2! \cdots x_n!}\right)$$

$x_i = 0, 1, \dots$ if $x_i \geq 0$, $\sum x_i = n$,

$p_i > 0$, $\sum p_i = 1$

Let X_1, X_2, X_3 be Mulinomial $(n, (p_1, p_2, p_3))$

Then $X_1 \sim$ Binomial(n, p_1)

Proof: $P(X_1 = x_1) = \sum \binom{n}{x_1\ x_2\ x_3} p_1^{x_1} p_2^{x_2} p_3^{x_3}$

$x_1 + x_2 + x_3 = n$

$x_2 = 0 \ldots n - x_1$

$x_3 = n - x_1 - x_2$

$$= \frac{n!}{x_1!\,(n-x_1)!} p_1^{x_1} \sum_{x_2 + x_3 = n - x_1} \frac{(n-x_1)!}{x_2!\,x_3!} p_2^{x_2} p_3^{x_3}$$

pulling out factors that don't depend on summands

Greatest tricks in math $\begin{cases} \text{Multiply} \times 1 \\ \text{add } 0 \end{cases}$

What can we do here?

$$\frac{n!}{x_1!\,(n-x_1)!} p_1^{x_1} (1-p_1)^{n-x_1} = \ldots$$

$$\frac{n!}{x_1!\,x_2!\,x_3!} p_1^{x_1} p_2^{x_2} p_3^{x_3}$$

Continuous random variables

$\iint$ instead of $\sum_i$

(X, Y) have a bivariate continuous distribution

if $\quad P((X,Y) \in A) = \iint_A f(x,y)\,dx\,dy$

for some function f such that

1) $f \geq 0$ on $\mathbb{R}^2$

2) $\iint f(x,y)\,dx\,dy = 1$

CDF: $\quad F(x,y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f(u,v)\,dv\,du$

Simple example: $f(x,y) = c(x+y) \quad x,y \in (0,1)$

1) Find c to make f a density

$$\iint f(x,y)\,dx\,dy = \int_0^1 \int_0^1 (x+y)\,dy\,dx$$

$$= \int_0^1 \left[xy + \frac{y^2}{2} \right|_0^1 dx$$

$$= \int_0^1 \left(x + \frac{1}{2}\right) dx = \left[\frac{x^2}{2} + \frac{1}{2}x \right|_0^1$$

$$= \frac{1}{2} + \frac{1}{2} = 1$$

So $c = 1$

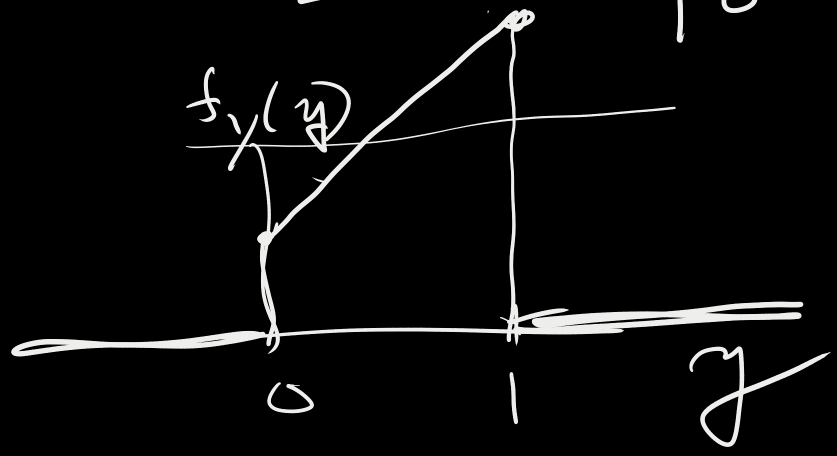
Marginal distribution of Y

Integrate out X :

$$f_Y(y) = \int f(x,y)\,dx$$

$$= \int_0^1 x + y \, dx = \left[\frac{x^2}{2} + xy\right]_0^1$$

$$= \frac{1}{2} + y$$



So $P(Y \le 1/2) = \int_0^{1/2} f_Y(y)\, dy$

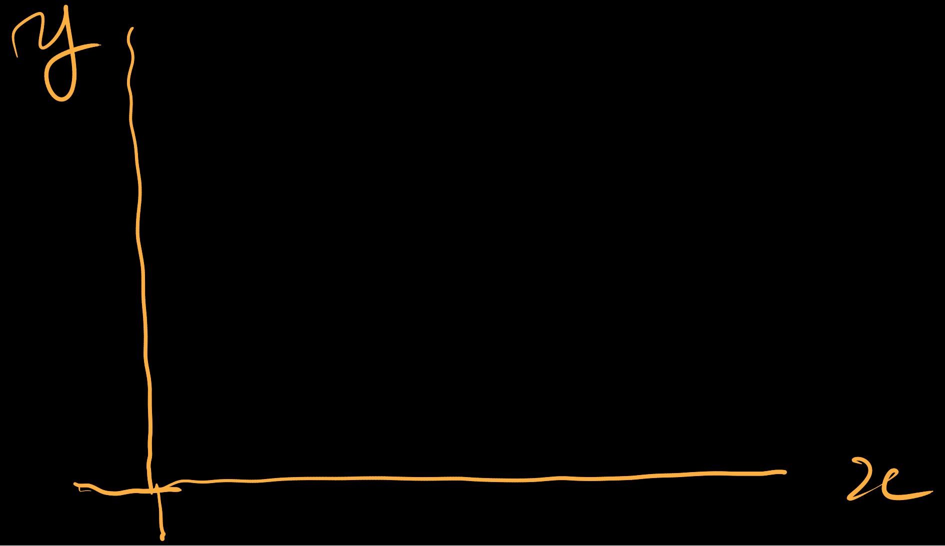
$\left[\frac{y}{2} + \frac{y^2}{2}\right|_0^{1/2} = \int_0^{1/2} \frac{1}{2} + y \, dy = \frac{3}{8}$

$\frac{1}{2}\left(1/2 + \frac{1}{4}\right)$

Harder example : p.79 Example D

$$f(x,y) = \begin{cases} \lambda^2 e^{-\lambda y} & 0 \le x \le y, \ \lambda > 0 \\ 0 & \text{otherwise} \end{cases}$$

Where is $f(x,y) > 0$?



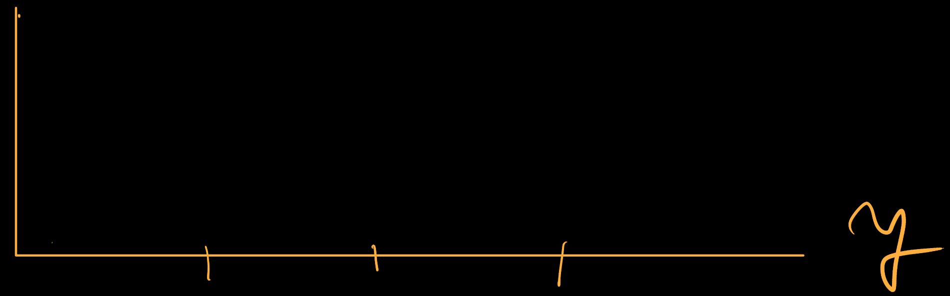
What does $f(x, y)$ look like?

Fix parameter $\lambda = 1$

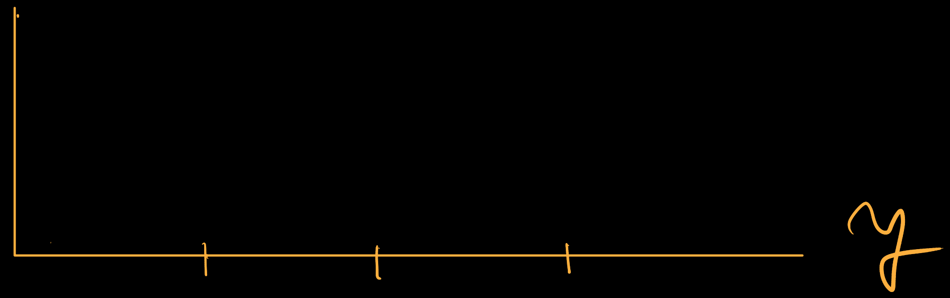
Consider different x's

e^{-y} for $y \geq x$

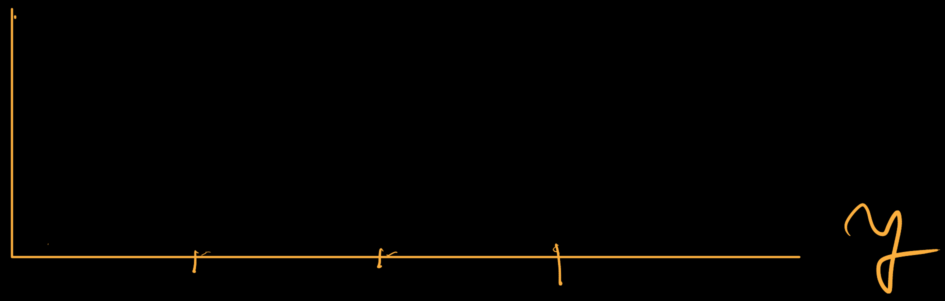
$x = 0$



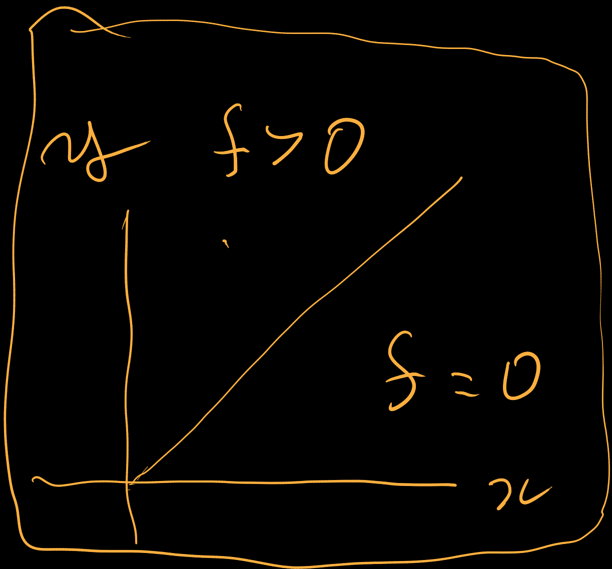
$x = 1$



$x = 2$



Marginal distn of X   Integrate out y



$$\int_{-\infty}^{\infty} f(x,y)\,dy$$

$$= \int_{?}^{?} \lambda^2 e^{-\lambda y}\,dy$$

$$= \int_x^\infty \lambda^2 e^{-\lambda y}\, dy = \left[\frac{\lambda^2 e^{-\lambda y}}{-\lambda} \right]_x^\infty$$

$$= 0 - \left[\frac{\lambda^2 e^{-\lambda x}}{-\lambda} \right]$$

$$= \lambda e^{-\lambda x} \text{ for } x \geq 0$$

So marginal for X is exponential(λ)

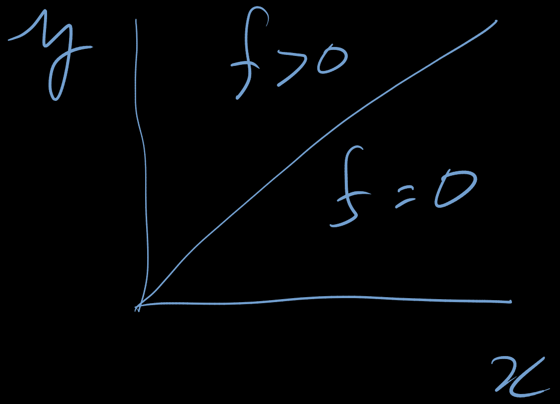
Marginal for Y

Integrate out x

$$f_Y(y) = \int_{-\infty}^{\infty} f(x,y)\,dx$$

for fixed y

$$= \int_{?}^{?} \lambda^2 e^{-\lambda y}\,dx$$



$$= \int_0^y \lambda^2 e^{-\lambda y} \, dx$$

$$= \left[x \lambda^2 e^{-\lambda y} \right]_0^y$$

$$= \lambda^2 y e^{-\lambda y}$$

Gamma$(2, \lambda)$ distribution