Chapter 3 Joint distributions

2 or more random variables

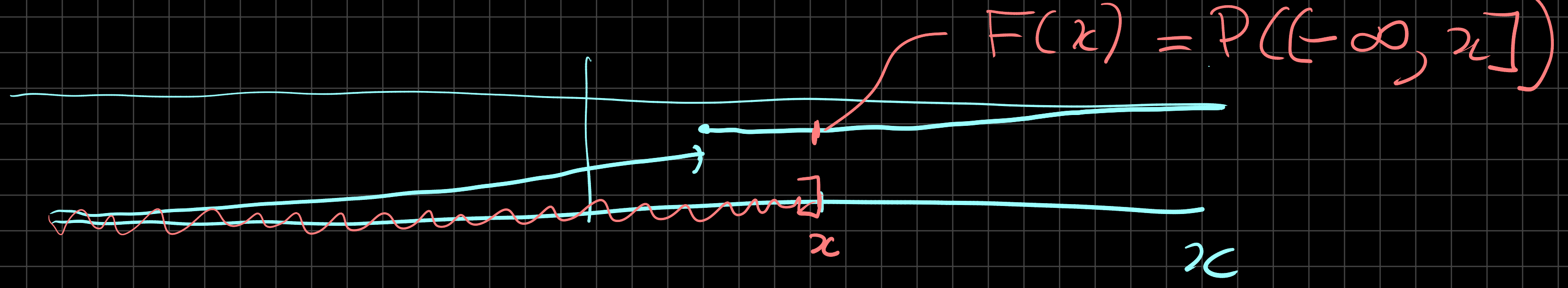
E.G. • Height • Weight

• Dose of a drug • time to recovery

• Gender, Age, Education, Income

<u>CDF</u> One RV: $F(x) = P(X \le x)$

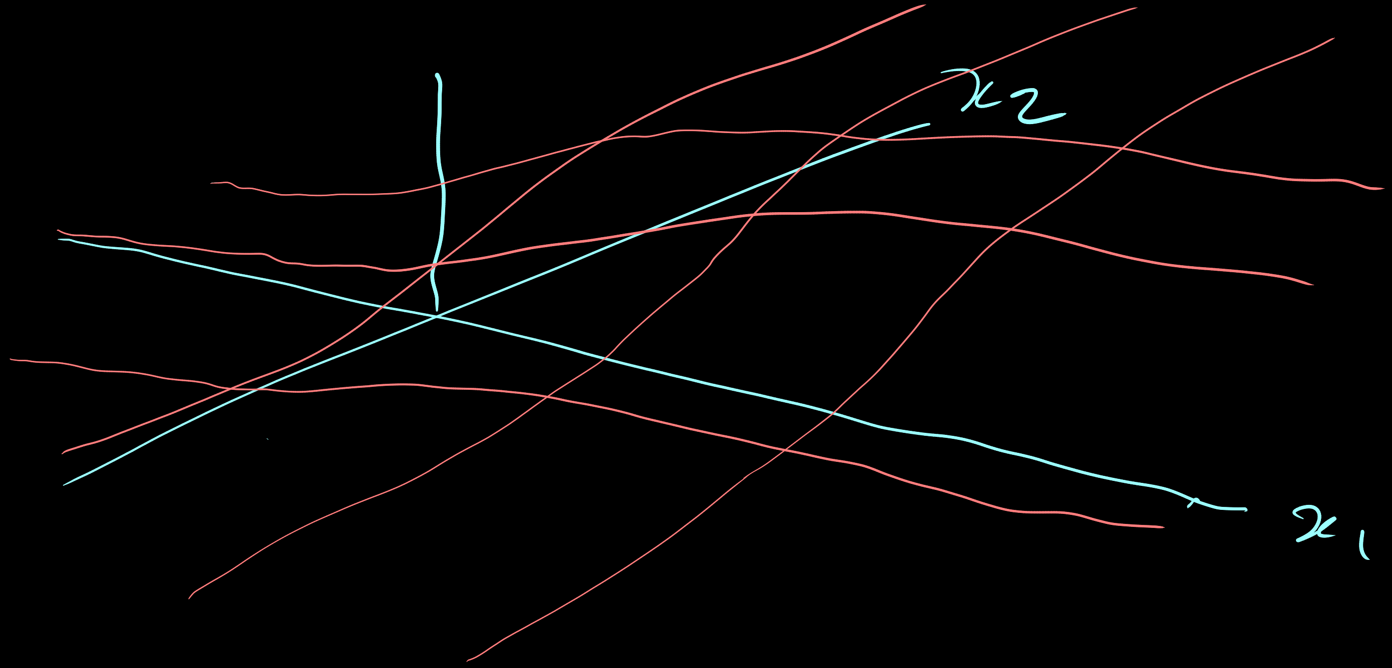
$$F : \mathbb{R} \to [0,1]$$



CDF: Two R.V.s X_1, X_2

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

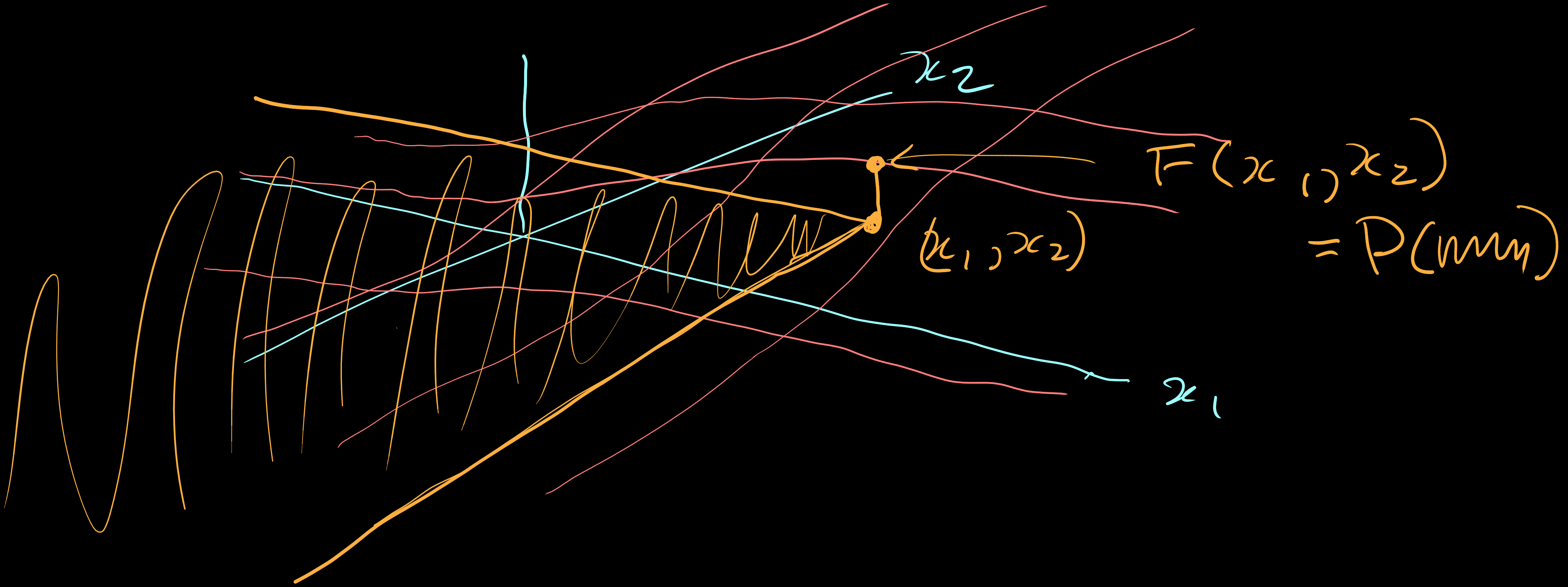
$$F : \mathbb{R}^2 \to [0, 1]$$



CDF: Two R.V.s X_1, X_2

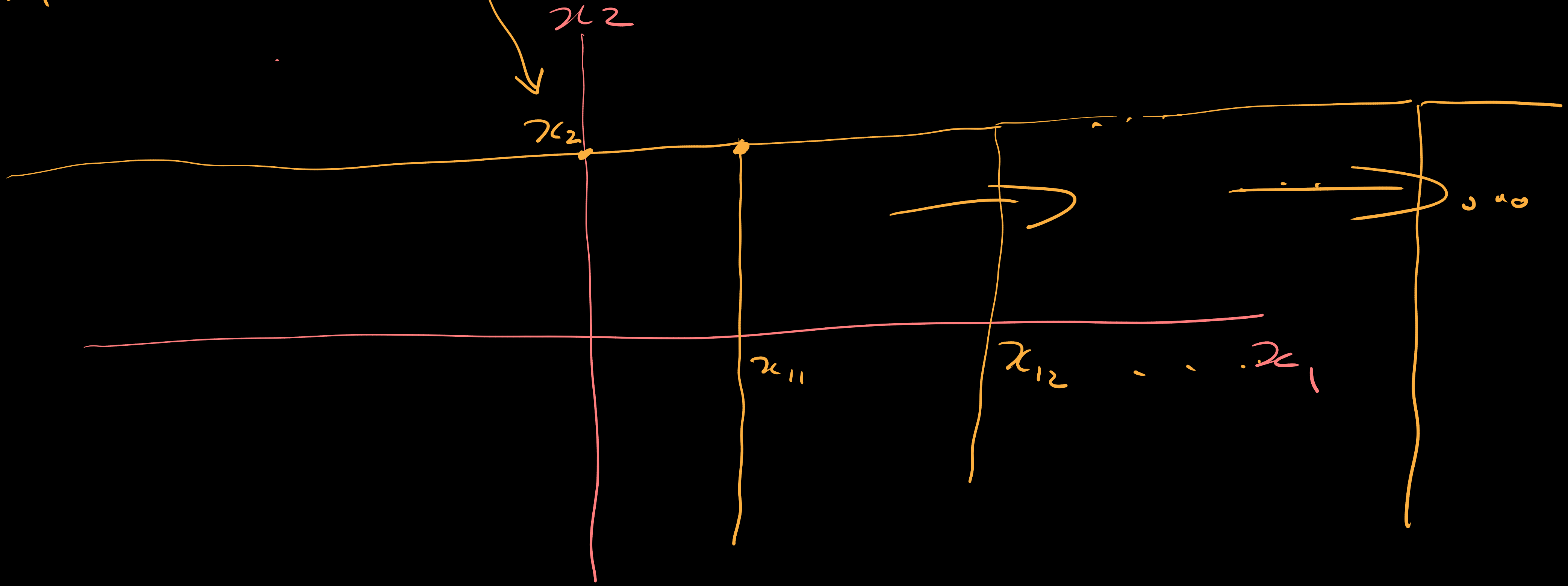
$$F(x_1, x_2) = P(X_1 \le x_2 \text{ and } X_2 \le x_2)$$

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



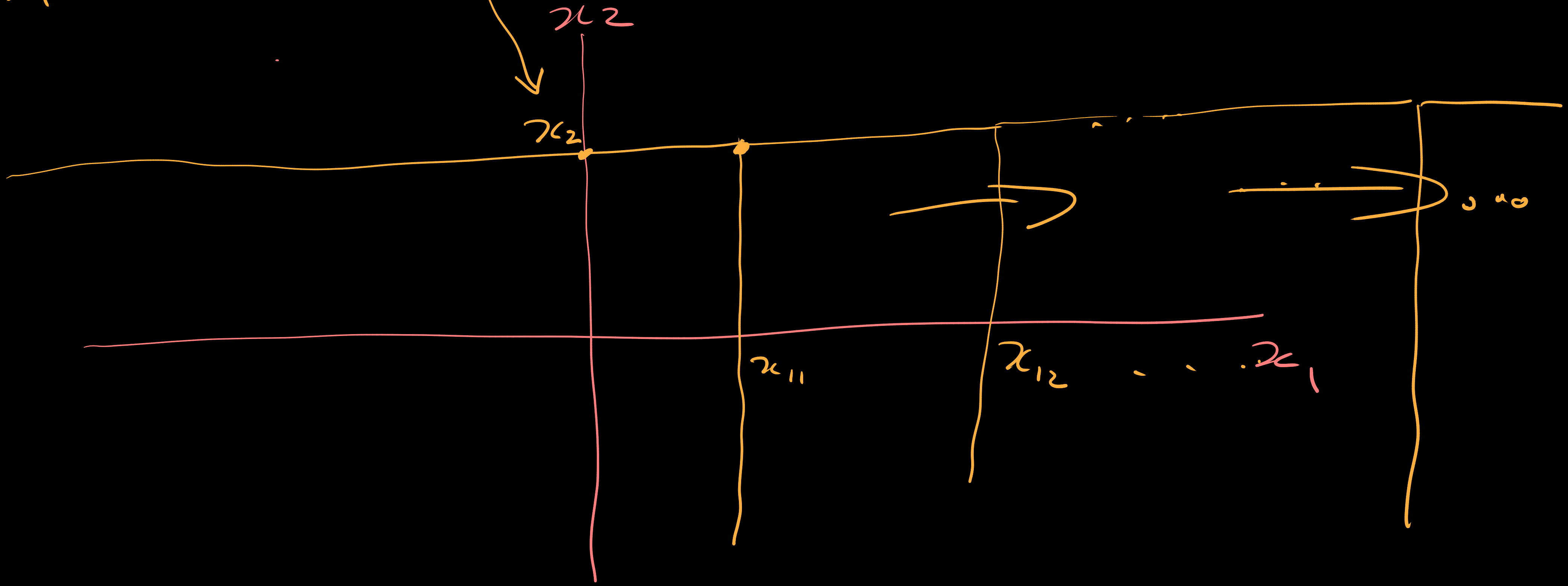
If you fix x_2 and let $x_1 \to \infty$

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



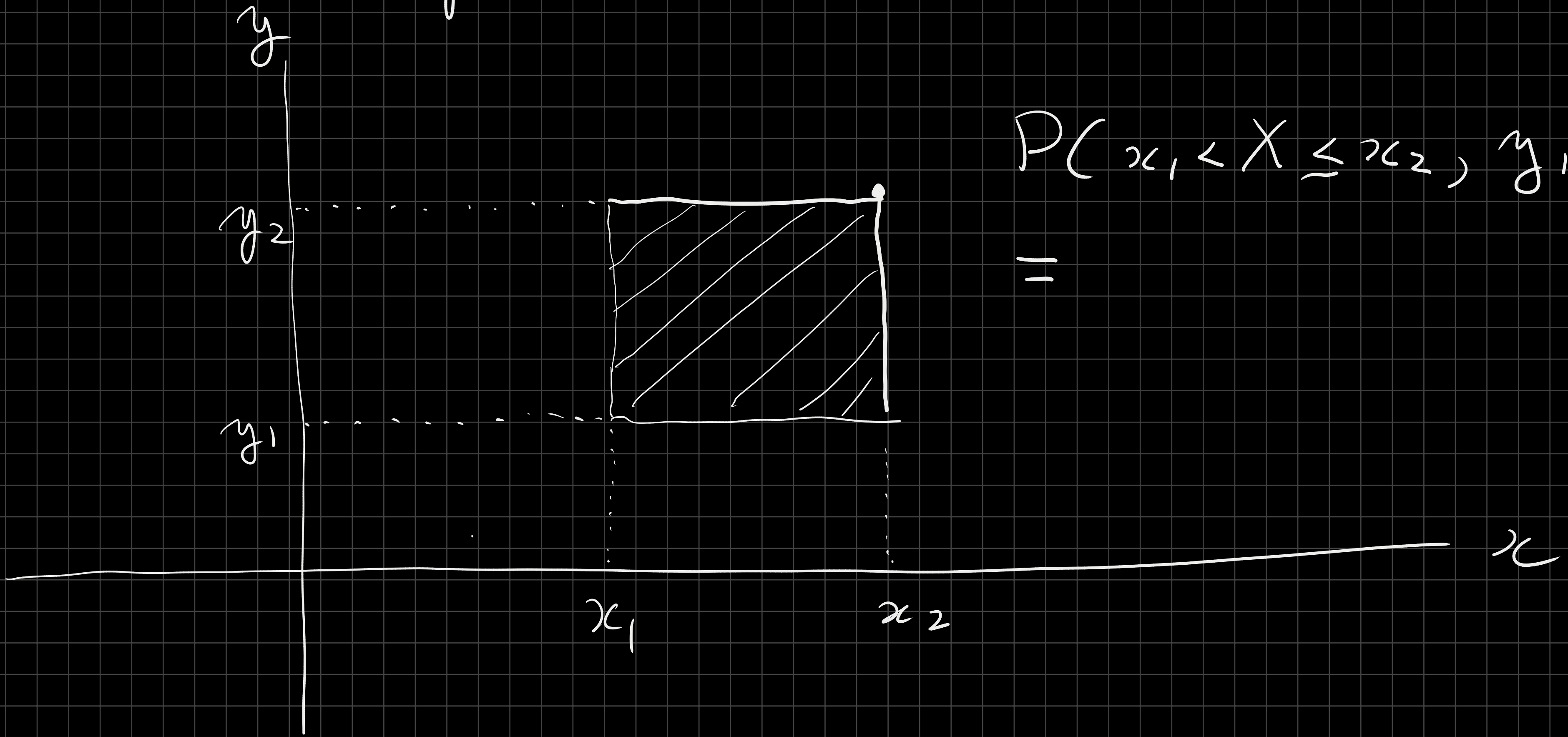
If you fix x_2 and let $x_1 \longrightarrow \infty$

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



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

Probability of a rectangle:

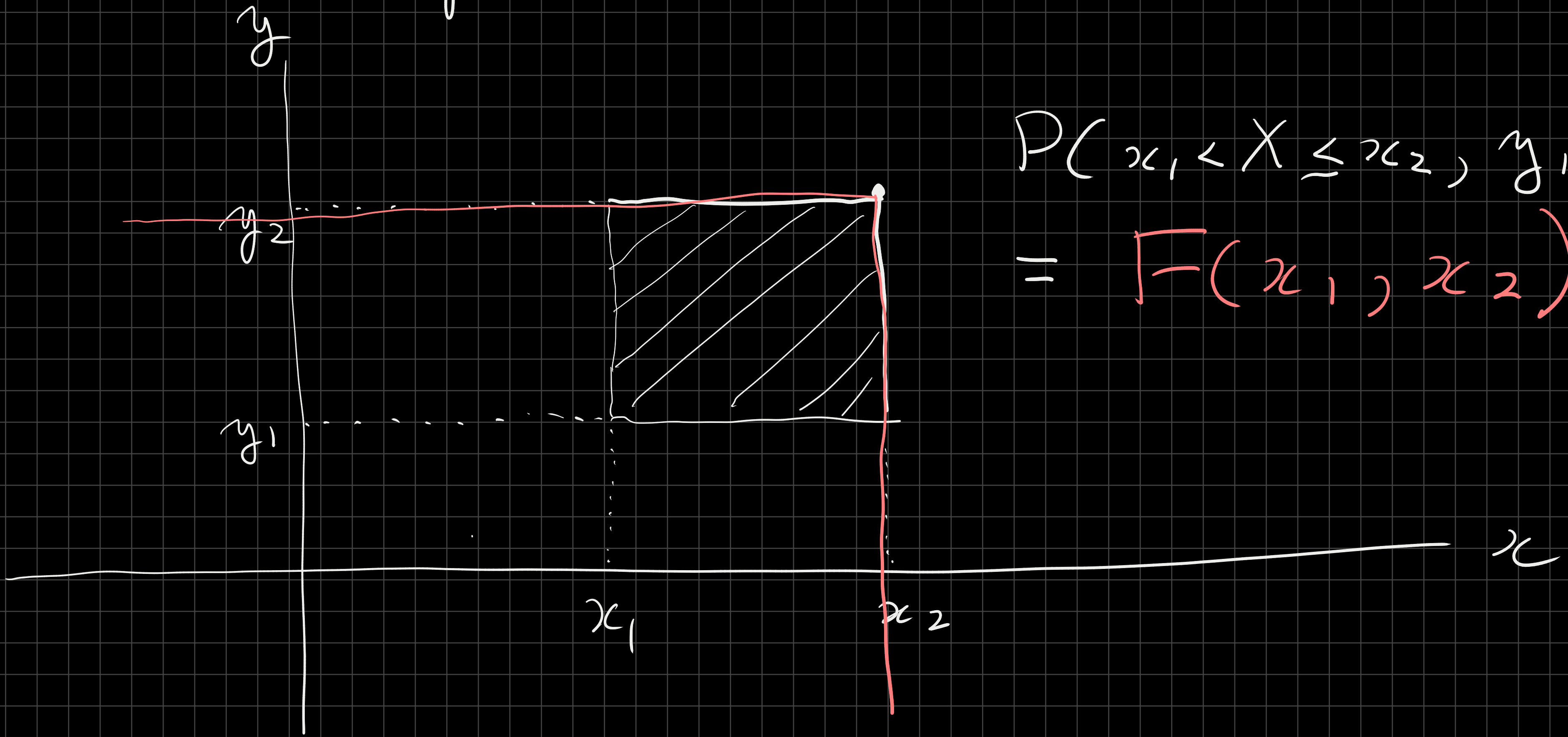


$$P(x_1 < X \leq x_2,\ y_1 < Y \leq y_2)$$

$$=$$

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

Probability of a rectangle:

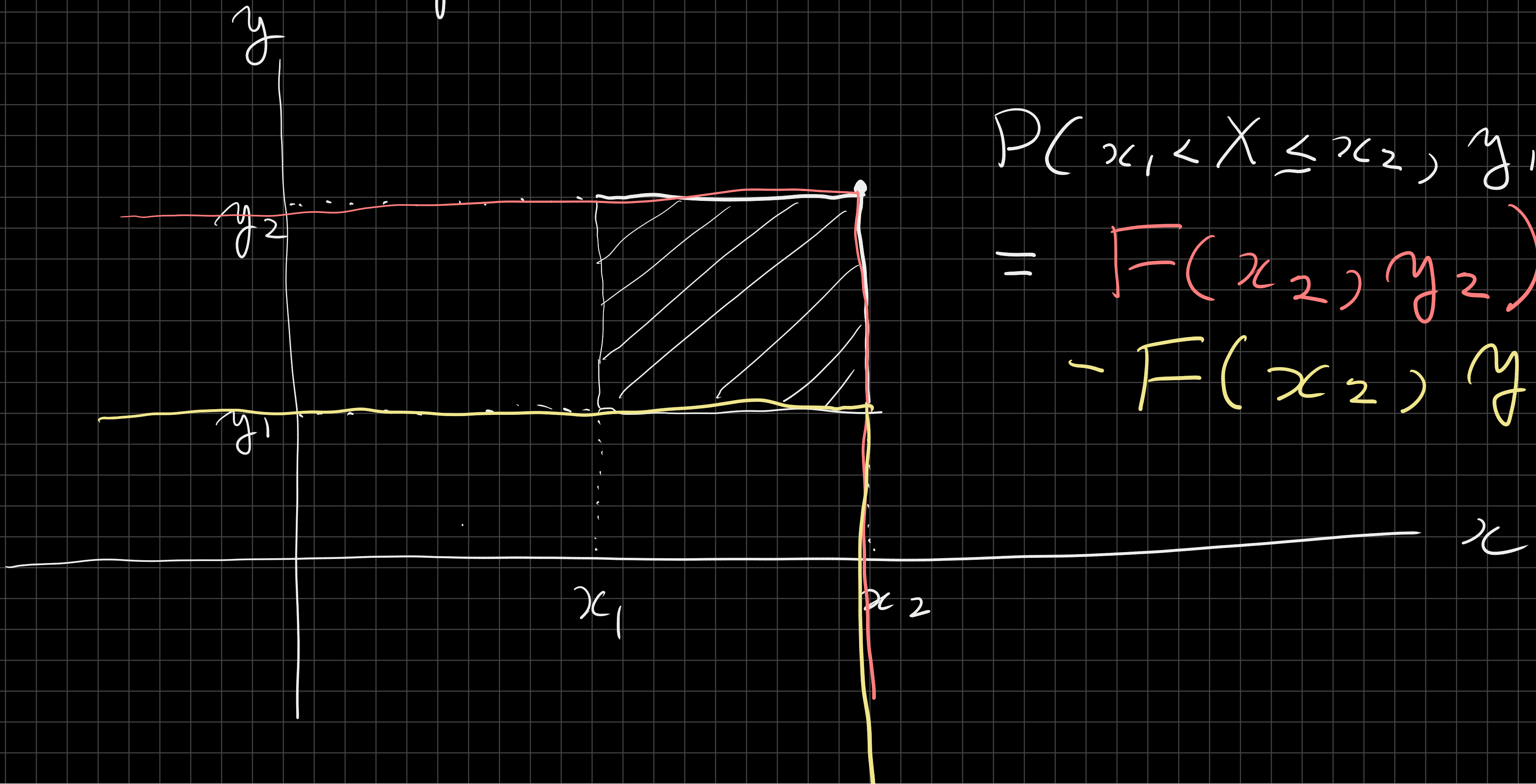


$$P(x_1 < X \leq x_2, y_1 < Y \leq y_2)$$

$$= F(x_1, x_2)$$

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

Probability of a rectangle:



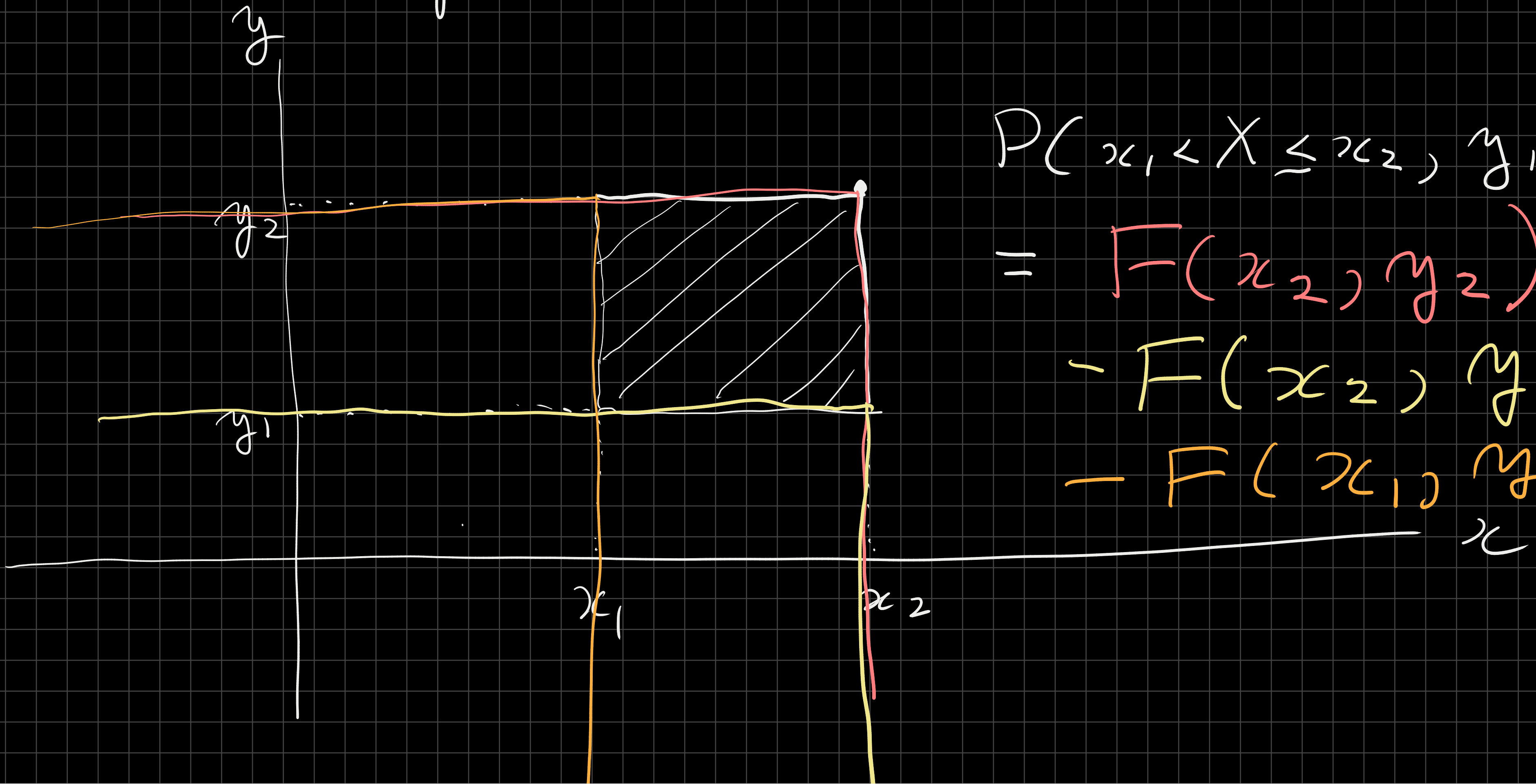
$$P(x_1 < X \leq x_2,\ y_1 < Y \leq y_2)$$

$$= F(x_2, y_2)$$

$$- F(x_2, y_1)$$

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

Probability of a rectangle:

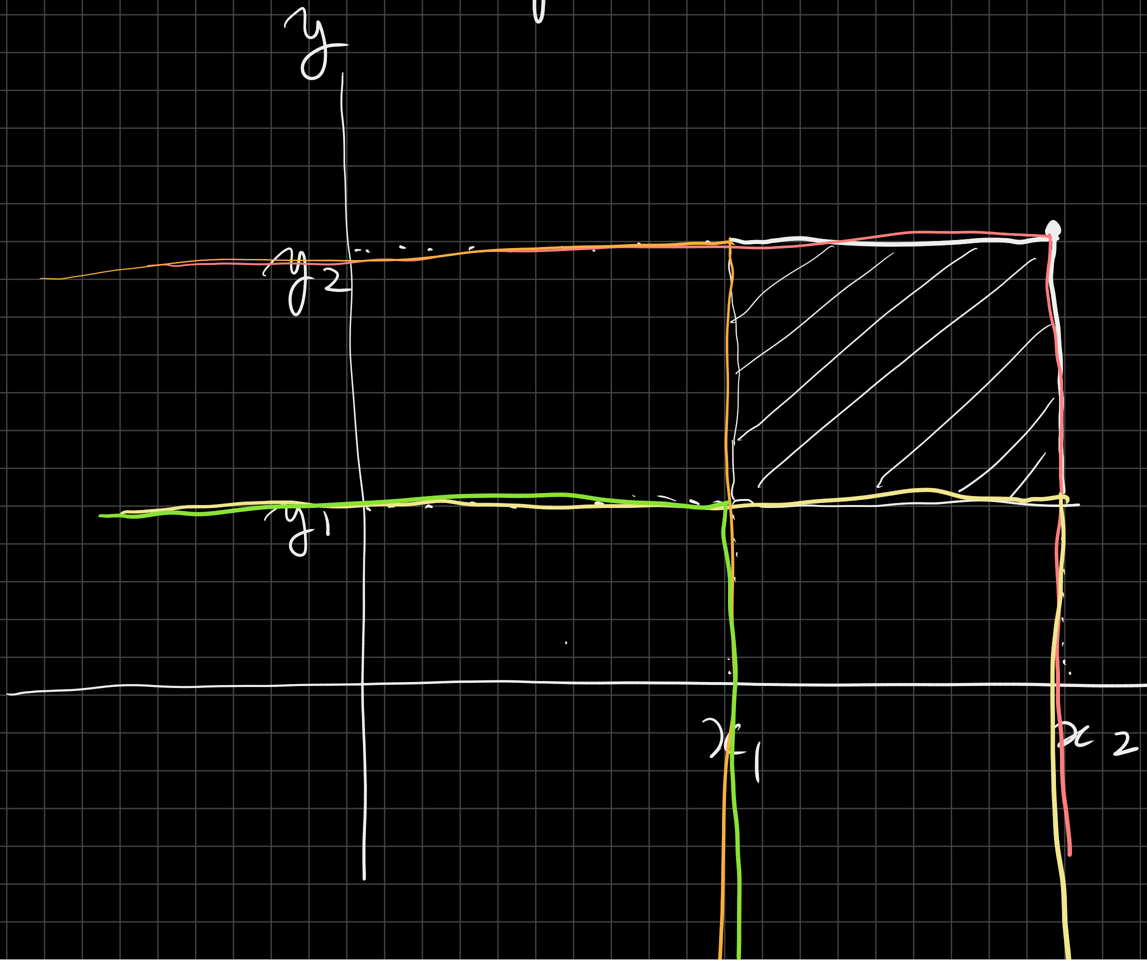


$$P(x_1 < X \leq x_2, y_1 < Y \leq y_2)$$

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

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

Probability of a rectangle:



$$P(x_1 < X \le x_2,\ y_1 < Y \le y_2)$$
$$= F(x_2, y_2)$$
$$- F(x_2, y_1)$$
$$- F(x_1, y_2)$$
$$+ F(x_1, y_1)$$

Recap:

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

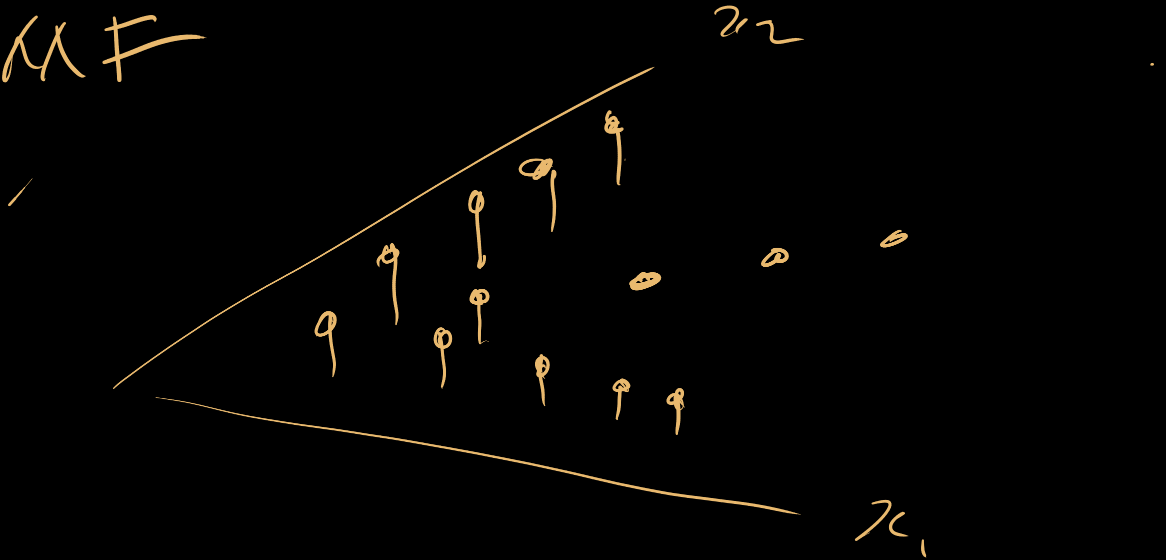
$$= F(x_2, y_2) - F(x_2, y_1) - F(x_1, y_2) + F(x_1, y_1)$$

Discrete random variables

Joint PMF $P(X_1 = x_1, X_2 = x_2, X_3 = x_3)$

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

Joint PMF



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.

$X \backslash Y$	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

Joint PMF

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 $X \sim \text{Binomial}(n, p)$

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

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

8 Hs and 2 Ts

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

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

Of course $X_1 + X_2 = n$

Here's the joint distribution for $n = 3$

X_2 (Tails)

X_1 (Heads)

0	0	1/4
0	1/2	0
1/4	0	0

$(X_1, X_2) \sim \text{Multinomial}\left(3, p = \left(\tfrac{1}{2}, \tfrac{1}{2}\right)\right)$

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 X_2 X_3 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:

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

$$\text{if } x_i \geq 0, \ \sum x_i = n,$$

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

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

Then $X_1 \sim \text{Binomial}(n, p_1)$

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

$$= \underbrace{\frac{n!}{x_1!} p_1^{x_1}}_{\text{pulling out factors that don't depend on summands}} \sum_{x_2 + x_3 = n - x_1} \frac{1}{x_2!\, x_3!} p_2^{x_2} p_3^{x_3}$$

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

What can we do here?

$= \ldots$

Continuous random variables

$\iint$ instead of Σ

(X, Y) have a bivariate continuous distribution if $P((X,Y) \in A) = \iint_A f(x,y)\,dx\,dy$

for some function f such that

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

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

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

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 $c = 1$

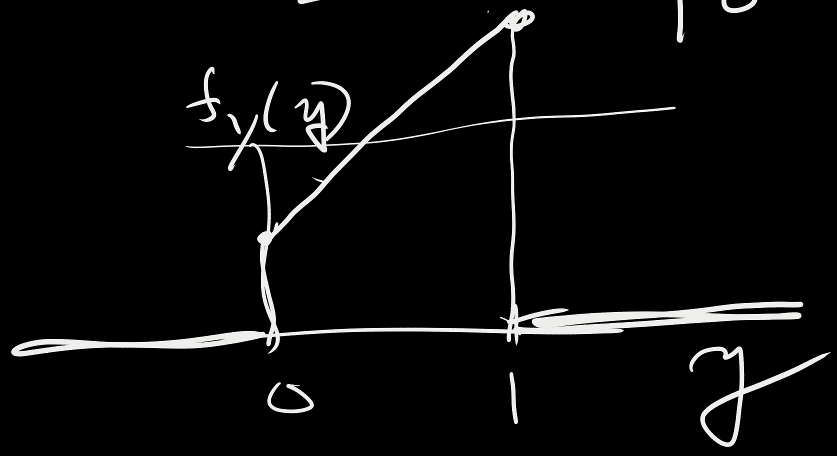
Marginal distribution of Y

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

$$= \int_0^{1/2} \frac{1}{2} + y \, dy = \frac{3}{8}$$