

Chapter 3 Joint distributions

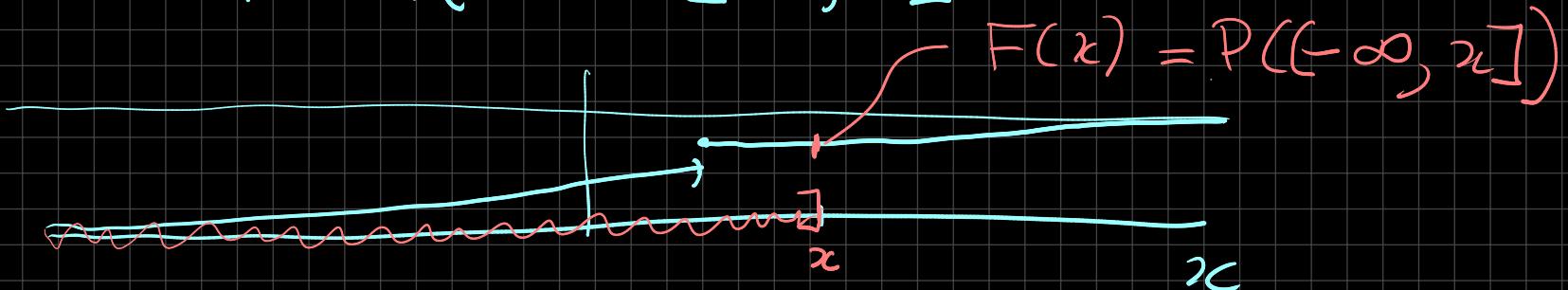
2 or more random variables

E.G. • Height • Weight

- Dose of a drug → time to recovery
- Gender, Age, Education, Income

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

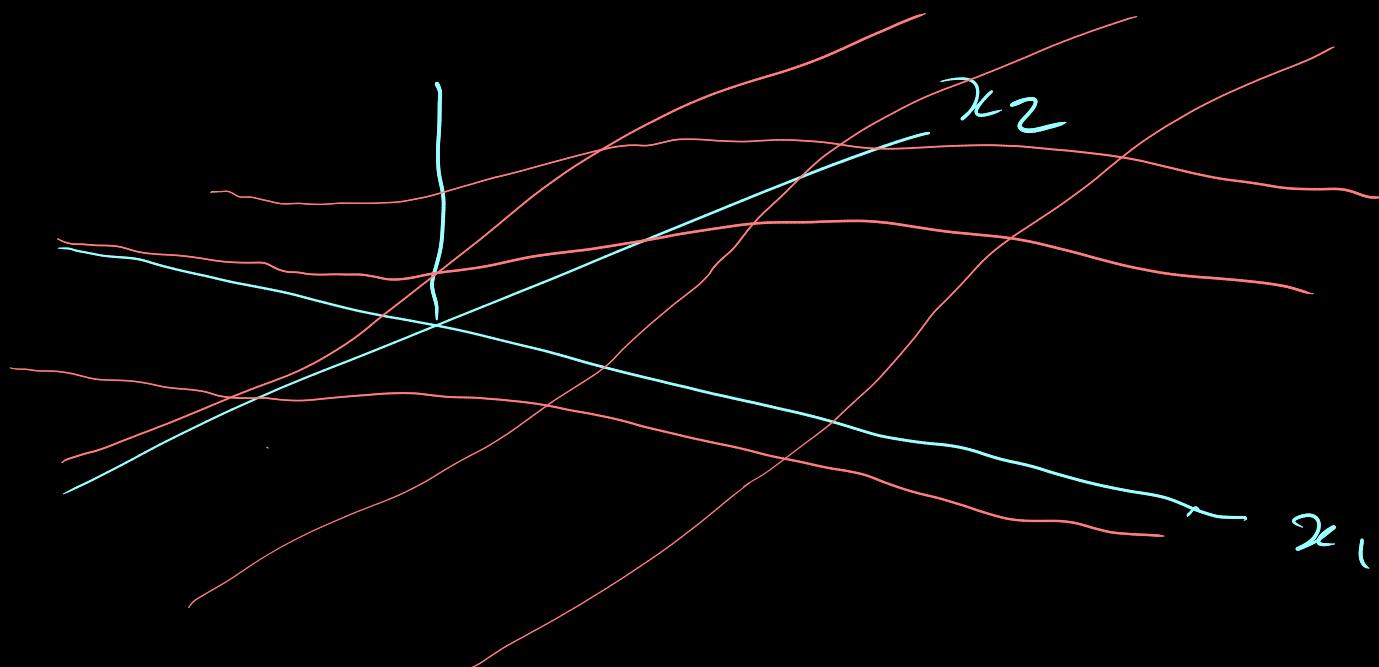
$$F: \mathbb{R} \rightarrow [0, 1]$$



CDF : Two R.V.s X_1, X_2

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

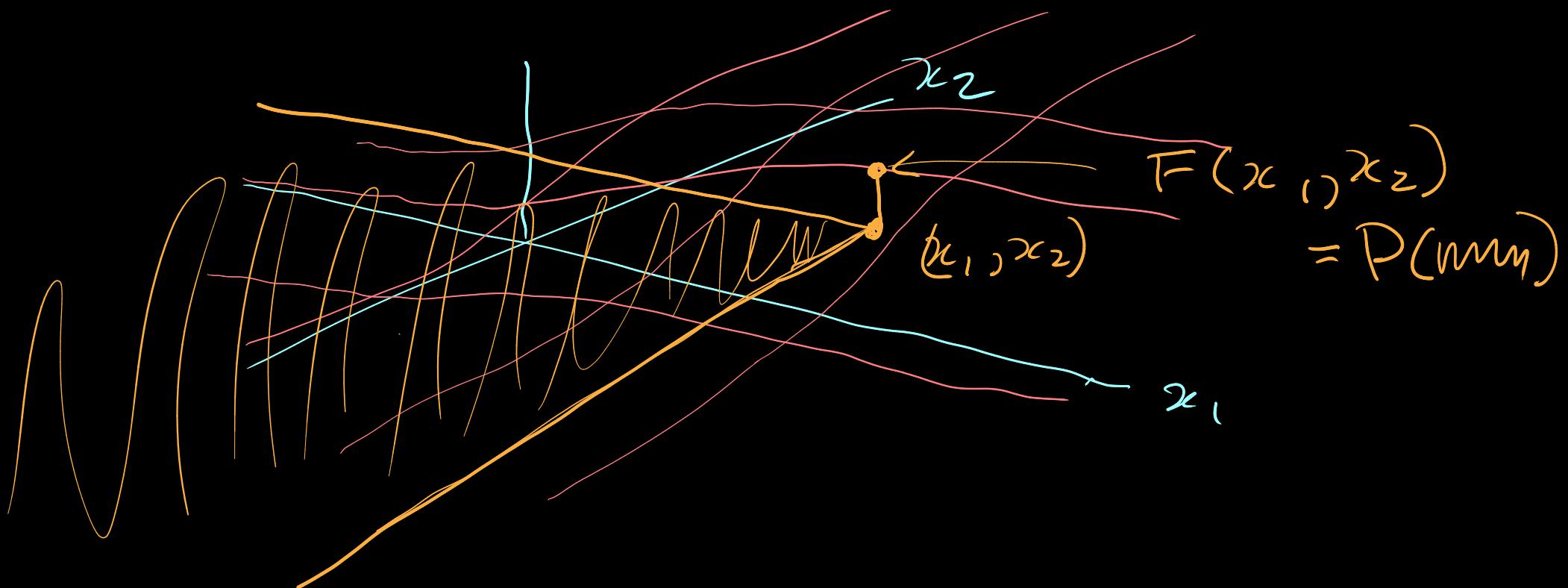
$$F: \mathbb{R}^2 \rightarrow [0, 1]$$



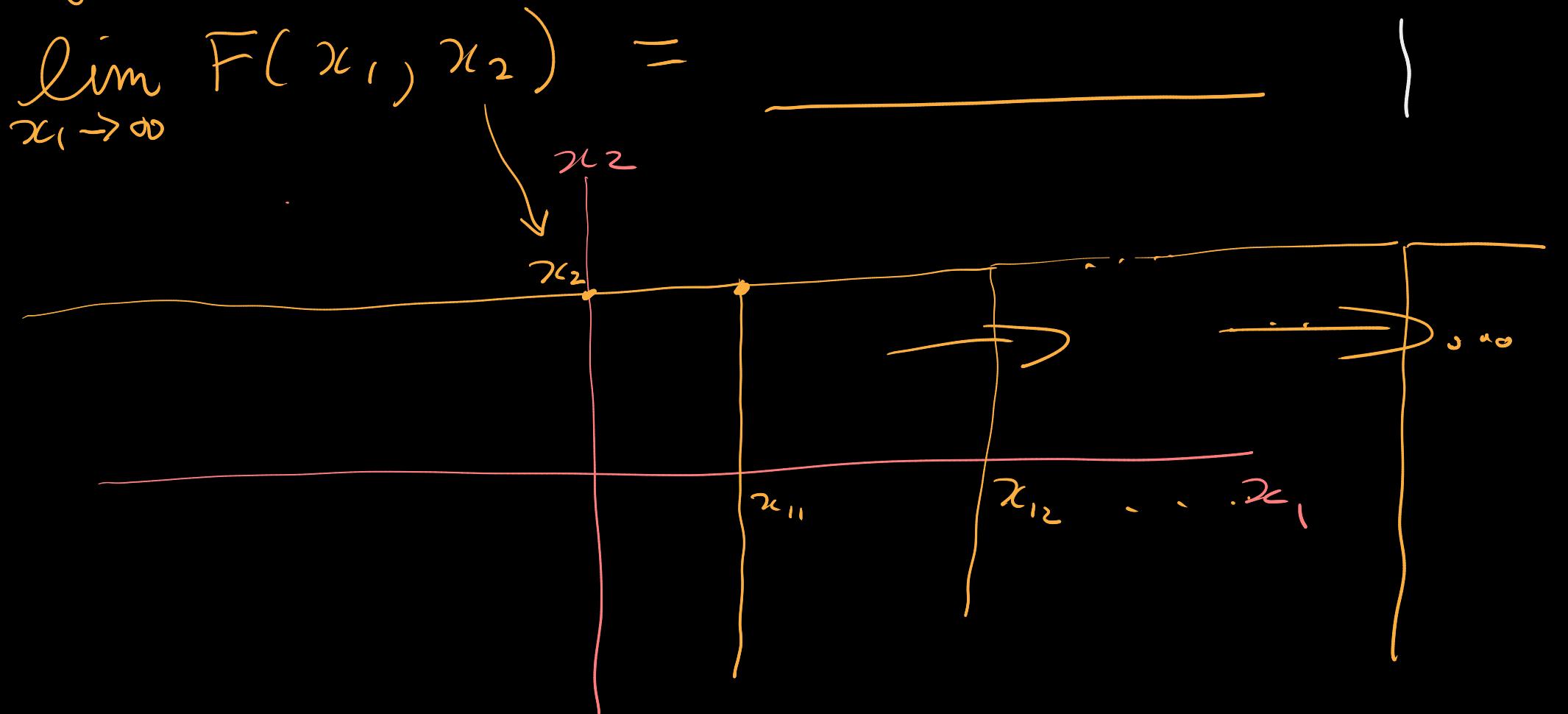
CDF : Two R.V.s X_1, X_2

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

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

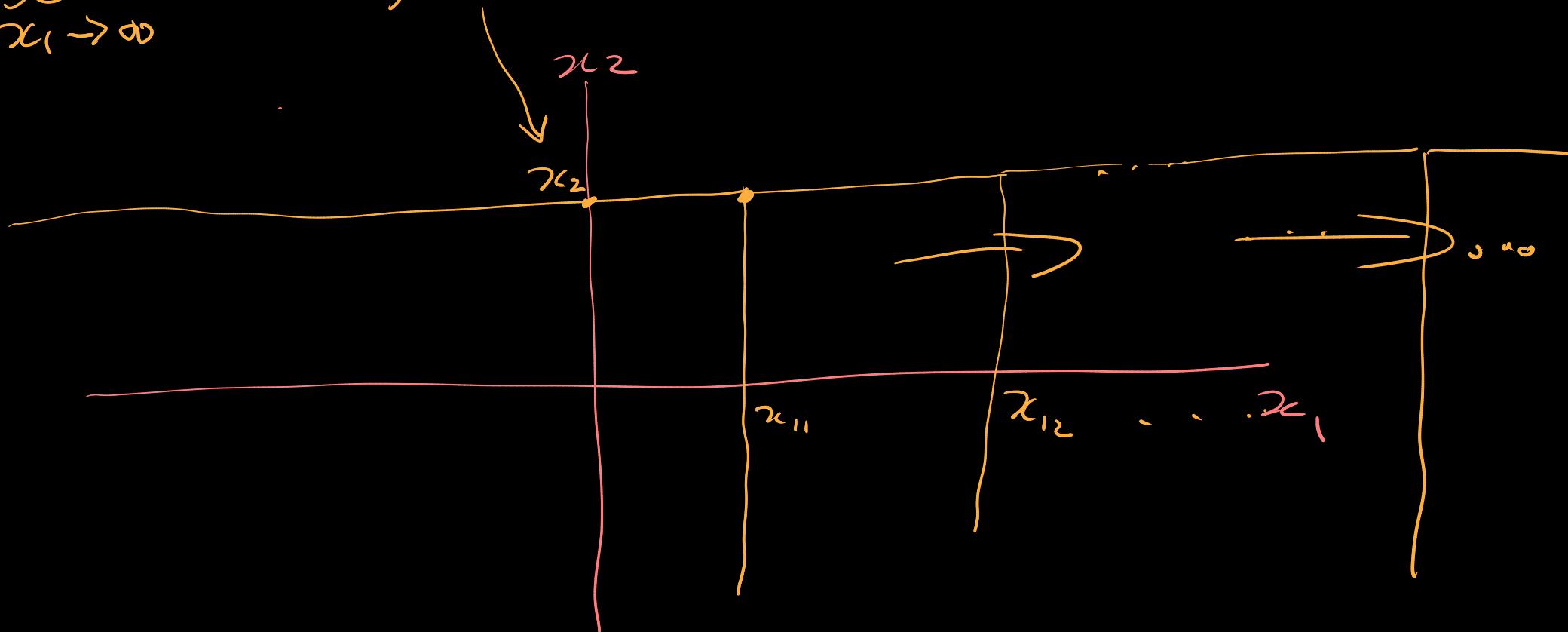


If you fix x_2 and let $x_1 \rightarrow \infty$



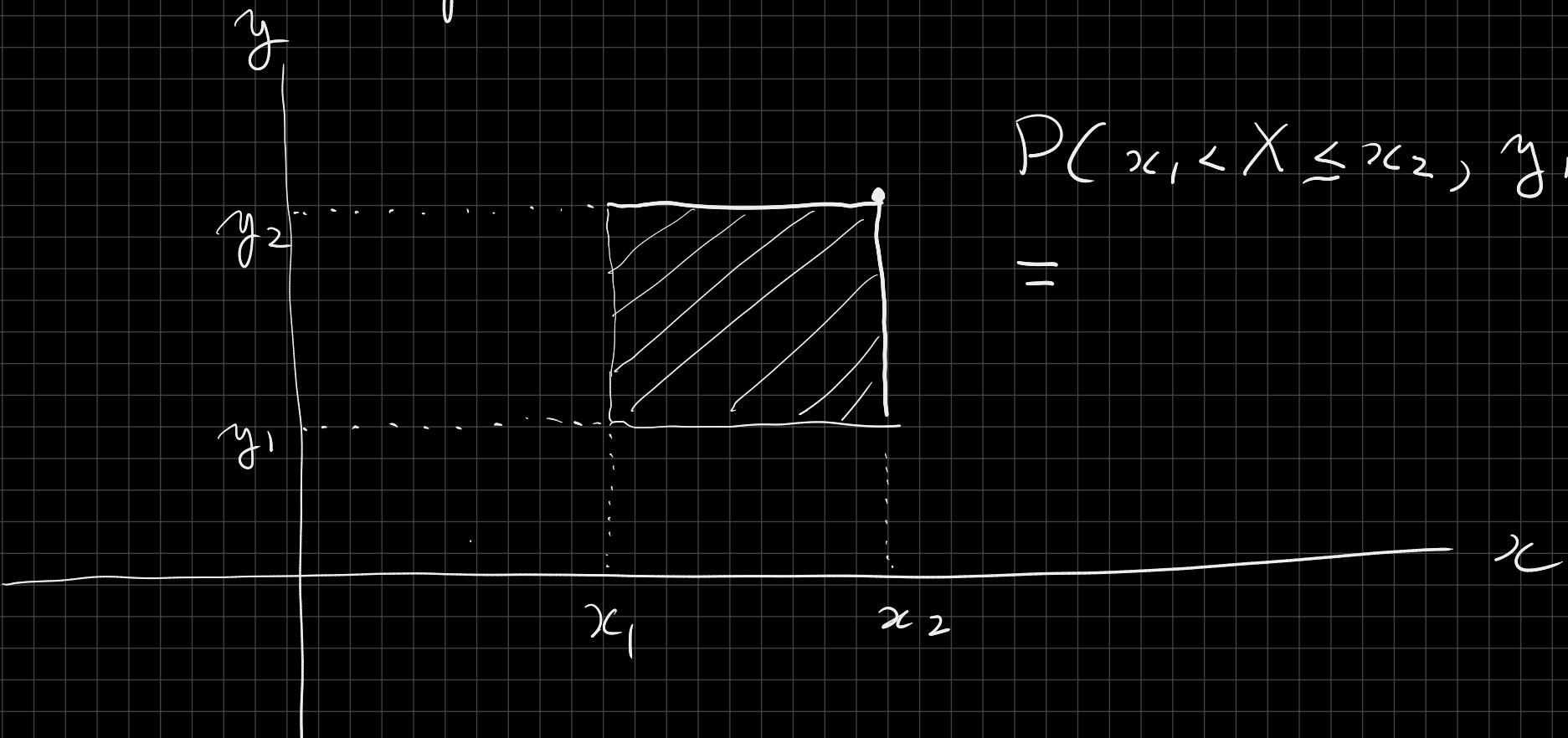
If you fix x_2 and let $x_1 \rightarrow \infty$

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



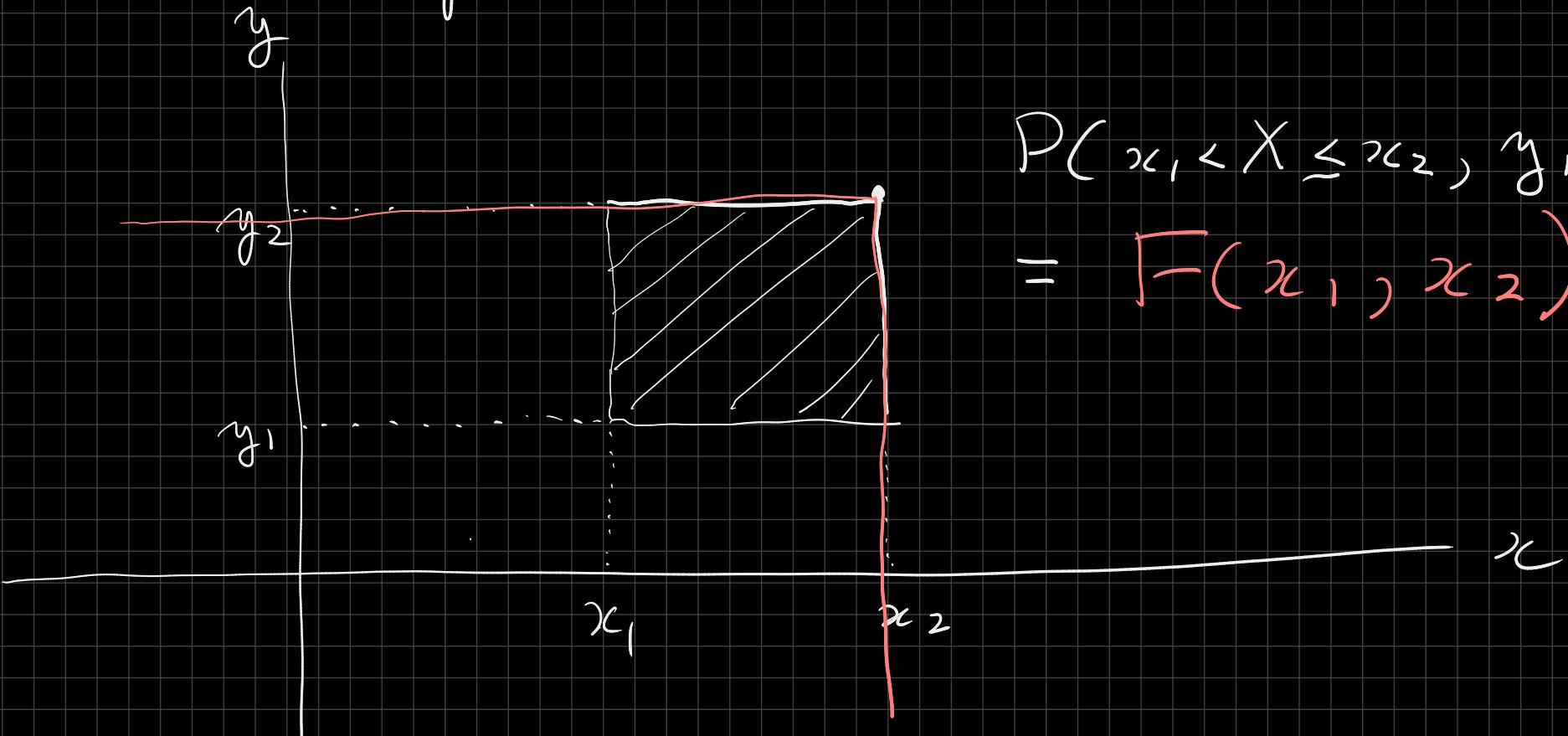
Similarly $\lim_{x_2 \rightarrow \infty} F(x_1, x_2) = F(x_1)$

Probability of a rectangle:



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

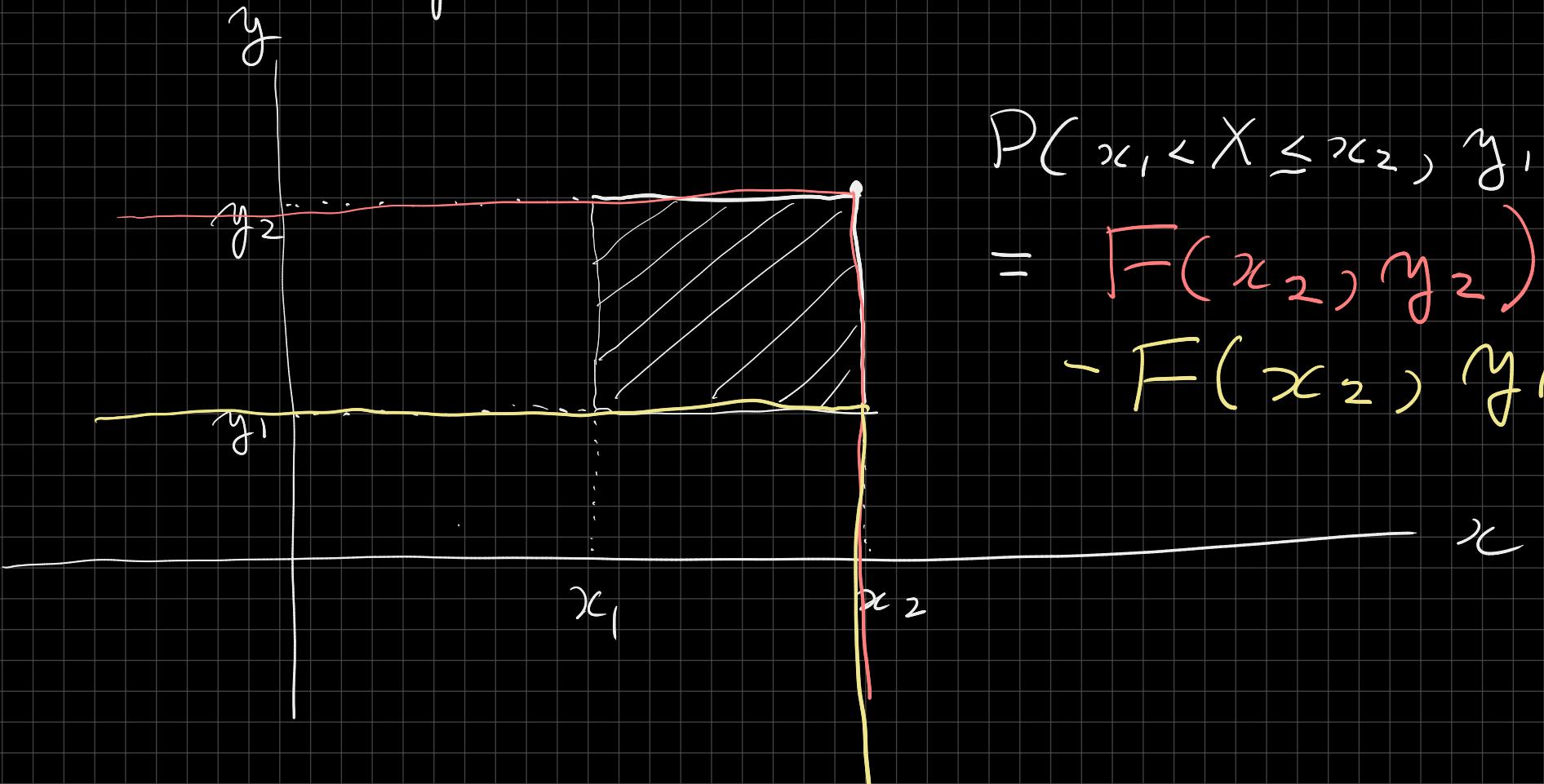
Probability of a rectangle:



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

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

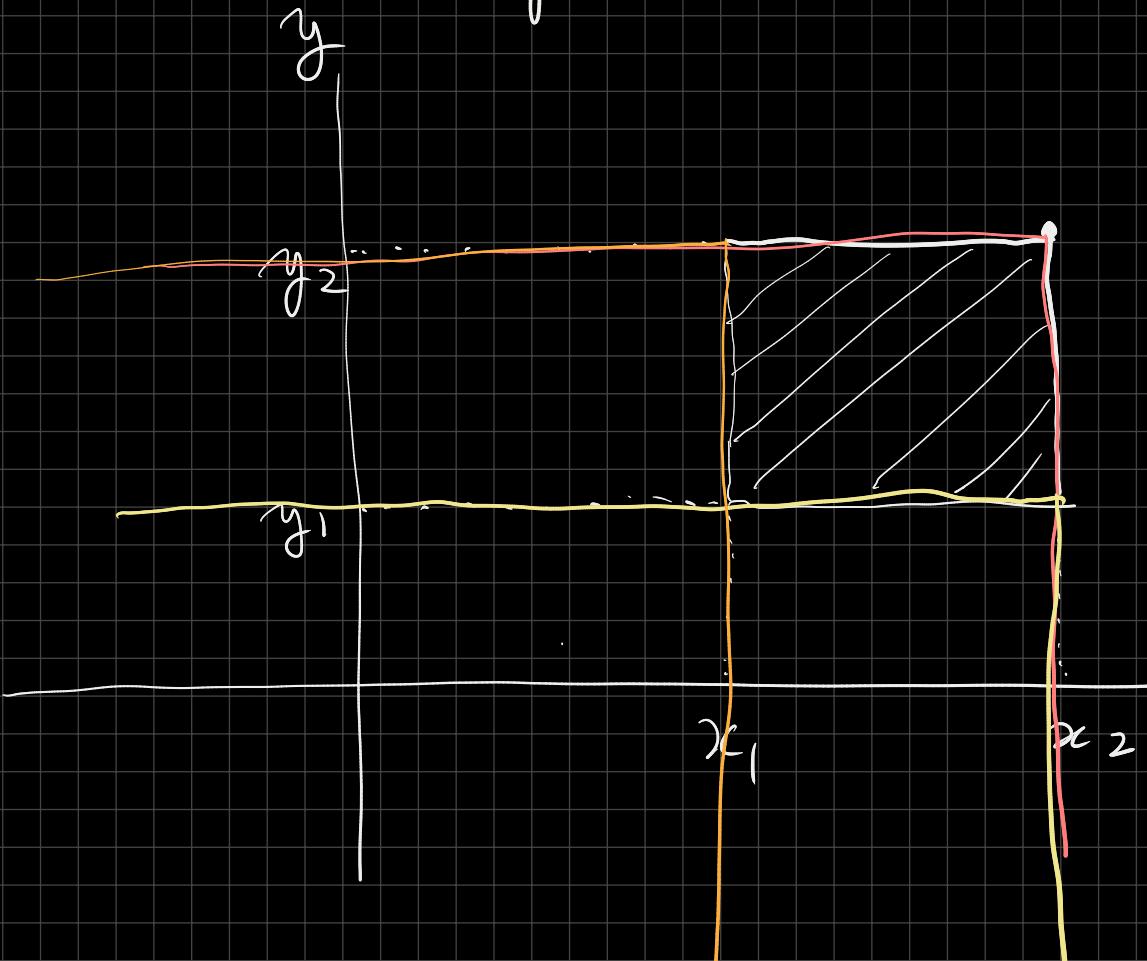
Probability of a rectangle:



$$\begin{aligned} P(x_1 < X \leq x_2, y_1 < Y \leq y_2) \\ &= F(x_2, y_2) \\ &\quad - F(x_2, y_1) \end{aligned}$$

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

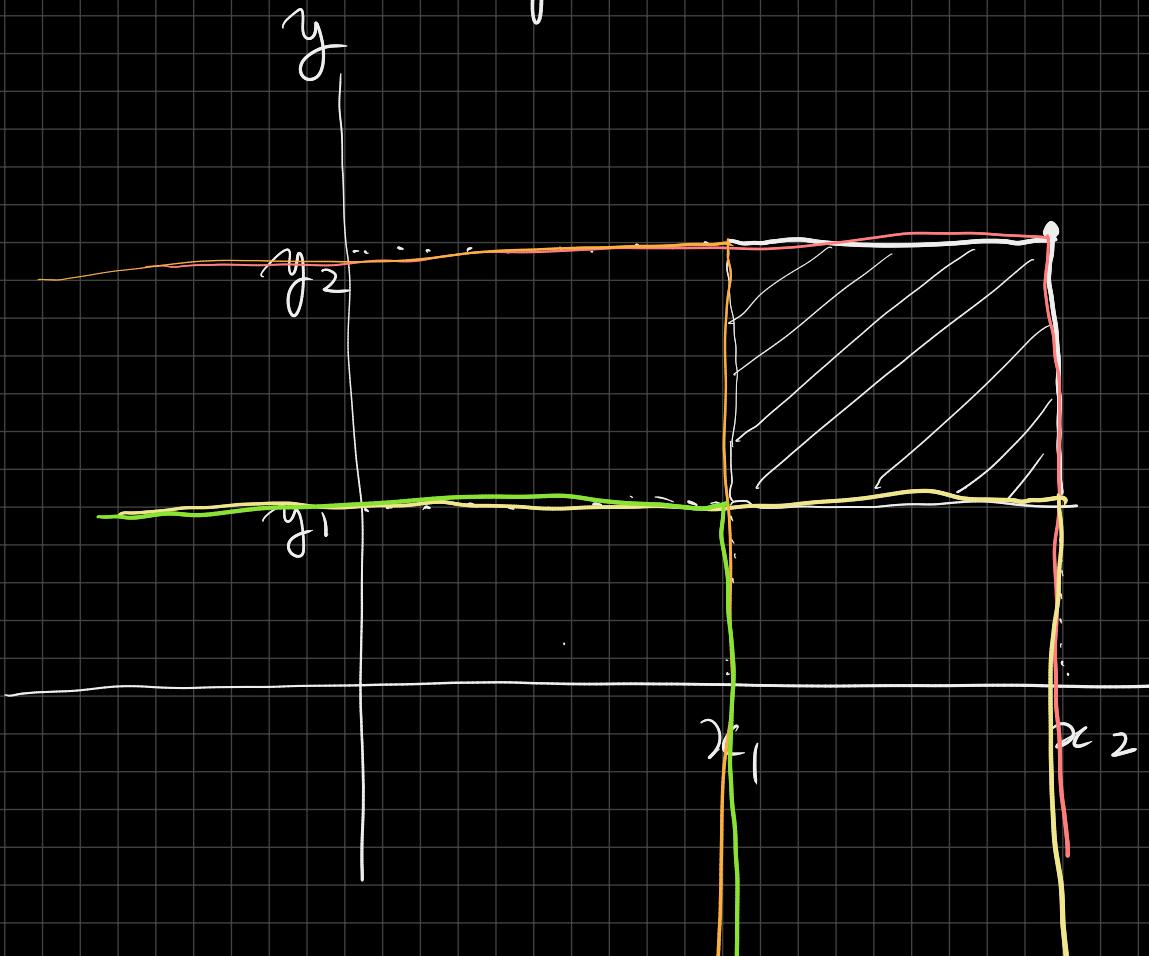
Probability of a rectangle:



$$\begin{aligned} P(x_1 < X \leq x_2, y_1 < Y \leq y_2) \\ &= F(x_2, y_2) \\ &\quad - F(x_2, y_1) \\ &\quad - F(x_1, y_2) \\ &\quad + F(x_1, y_1) \end{aligned}$$

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

Probability of a rectangle:



$$\begin{aligned} P(x_1 < X \leq x_2, y_1 < Y \leq y_2) \\ &= F(x_2, y_2) \\ &\quad - F(x_2, y_1) \\ &\quad - F(x_1, y_2) \\ &\quad + F(x_1, y_1) \end{aligned}$$

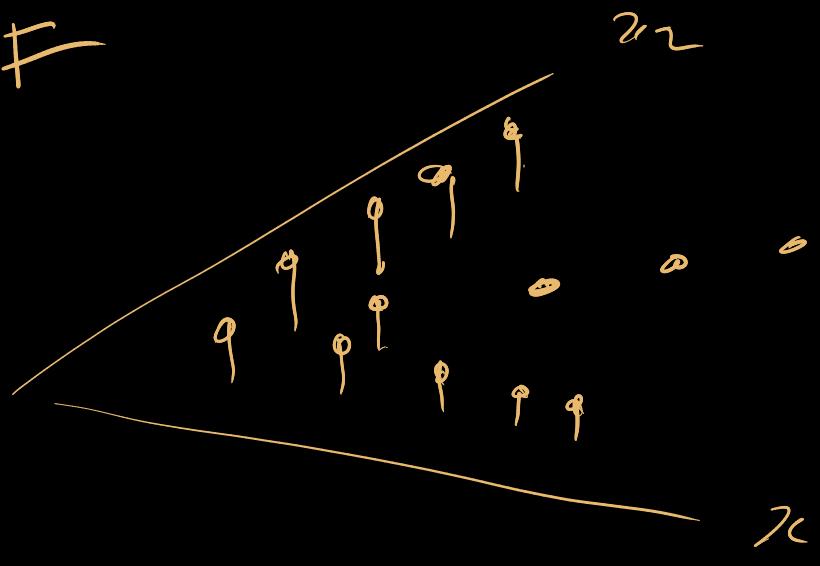
Recap :

$$\begin{aligned} P(x_1 < X \leq x_2, y_1 < Y \leq y_2) \\ = F(x_2, y_2) - F(x_2, y_1) - F(x_1, y_2) + F(x_1, y_1) \end{aligned}$$

Discrete random variables

$$\begin{aligned} \text{Joint PMF } P(X_1 = x_1, X_2 = x_2, X_3 = x_3) \\ = p(x_1, x_2, x_3) \end{aligned}$$

Joint PMF



if $\text{Supp}(X_1, X_2)$ is finite, we can use a table.
e.g. Toss a coin: $X = \# \text{ of Heads}$ + a die, $Y = \# \text{ rolled}$.

		Y						Joint PMF	
		1	2	3	4	5	6	marginal PMF for X	
X	0	$\frac{1}{2}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{2}$	$P(S_2)$
	1	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{2}$	
		$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{2}$	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 T T H H T T H H ($n=10$)
8 Hs and 2 Ts

$$\text{So } X = 8$$

But we could also record # of Ts, 2
and think of this as a joint distribution

for $X_1 = \# \text{ of Hs}$, and $X_2 = \# \text{ of Ts}$

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

Of course $X_1 + X_2 = n$

Here's the joint distribution for $n = \cancel{3}$ 2

		X_2 (Tails)		
		0	1	2
0		0	0	$\frac{1}{4}$
X_1 (Heads)	1	0	$\frac{1}{2}$	0
	2	$\frac{1}{4}$	0	0

Tosses	P	\underline{x}_1	\underline{x}_2
TT	$\frac{1}{4}$	0	2
TH	$\frac{1}{4}$	1	1
HT	$\frac{1}{4}$	1	1
TT	$\frac{1}{4}$	2	0

$(X_1, X_2) \sim \text{Multinomial}(\cancel{3}, P = (\frac{1}{2}, \frac{1}{2}))$

This generalized to any number of categories.

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

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

$$(X_1, X_2, X_3, X_4) \sim \text{Multinomial}(n, (p_1, p_2, p_3, p_4))$$

↑
proportions of B, H, B, G
in population

Or. $\bar{X} \sim \text{Multinomial}(n, P)$

PMF :

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

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

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

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

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

$$\text{Proof: } P(X_1=x_1) = \sum_{x_2} \binom{n}{x_1 x_2 x_3} P_1^{x_1} P_2^{x_2} P_3^{x_3}$$

$x_3 = n - x_1 - x_2$

Pulling out factors that don't depend on summands

$$= \frac{n!}{x_1!} P_1^{x_1} \sum \frac{1}{x_2! x_3!} P_2^{x_2} P_3^{x_3}$$

$x_2 + x_3 = n - x_1$

Greatest tricks in math { Multiply $\times 1$
What can we do here? add 0

= 0 0 0

Continuous Random Variables

\iint instead of \sum

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

$$= \int_{-\infty}^x \left\{ \int_{-\infty}^y f(u, v) dv \right\} du$$

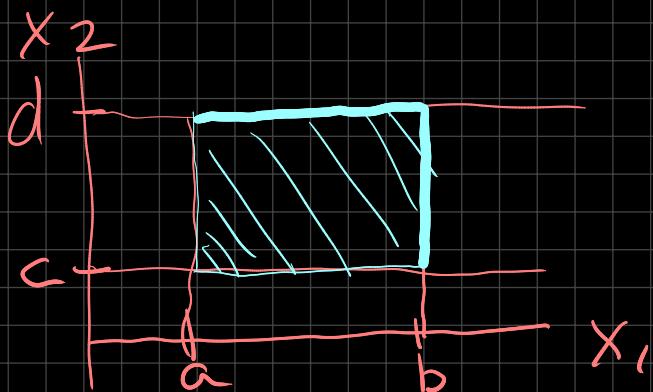
PDF from CDF

$$f(x_1, x_2) = \frac{\partial^2}{\partial x_1 \partial x_2} F(x_1, x_2)$$

Partial derivatives

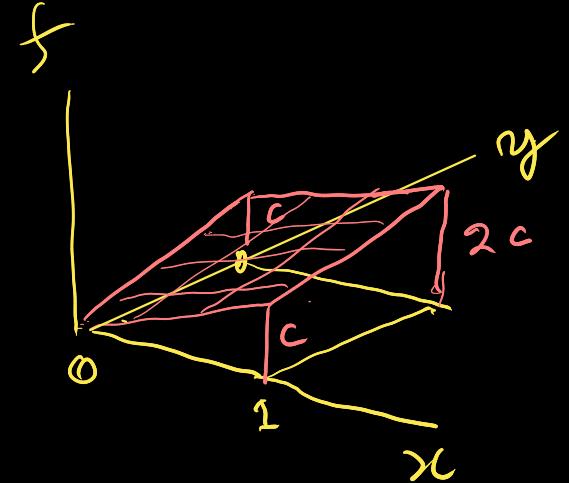
Probability of a rectangle (easier)

$$P(a < X_1 \leq b, c < X_2 \leq d) = \int_a^b \int_c^d f(x_1, x_2) dx_2 dx_1$$



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

1) Find c to make it a density



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

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

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

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

$$\text{So } c = 1$$

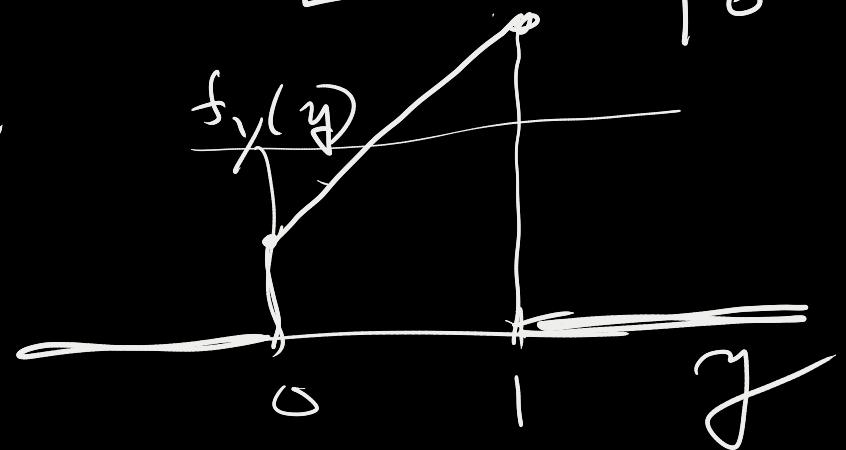
Marginal distribution of Y

Integrate out x :

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

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

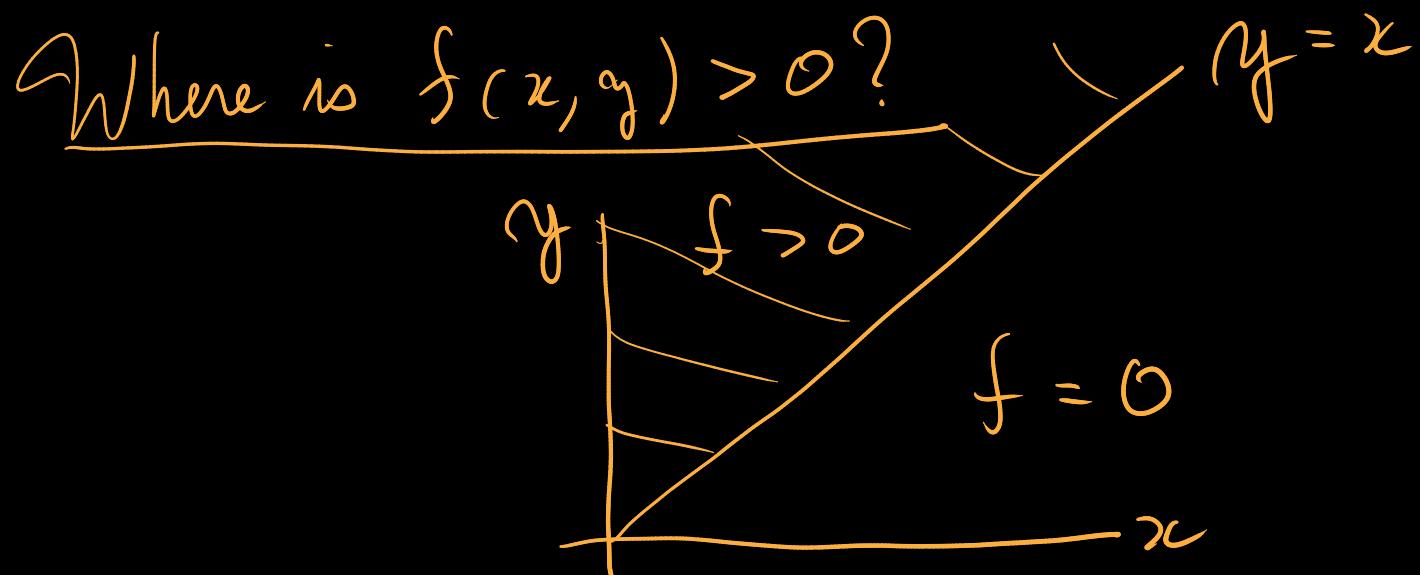
$$= \frac{1}{2} + y$$



$$\begin{aligned} S_2 \quad P(Y \leq \frac{1}{2}) &= \int_0^{\frac{1}{2}} f_Y(y) dy \\ &= \int_0^{\frac{1}{2}} \frac{1}{2} + y \ dy = \frac{3}{8} \end{aligned}$$

Harder example : p.79 Example D

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



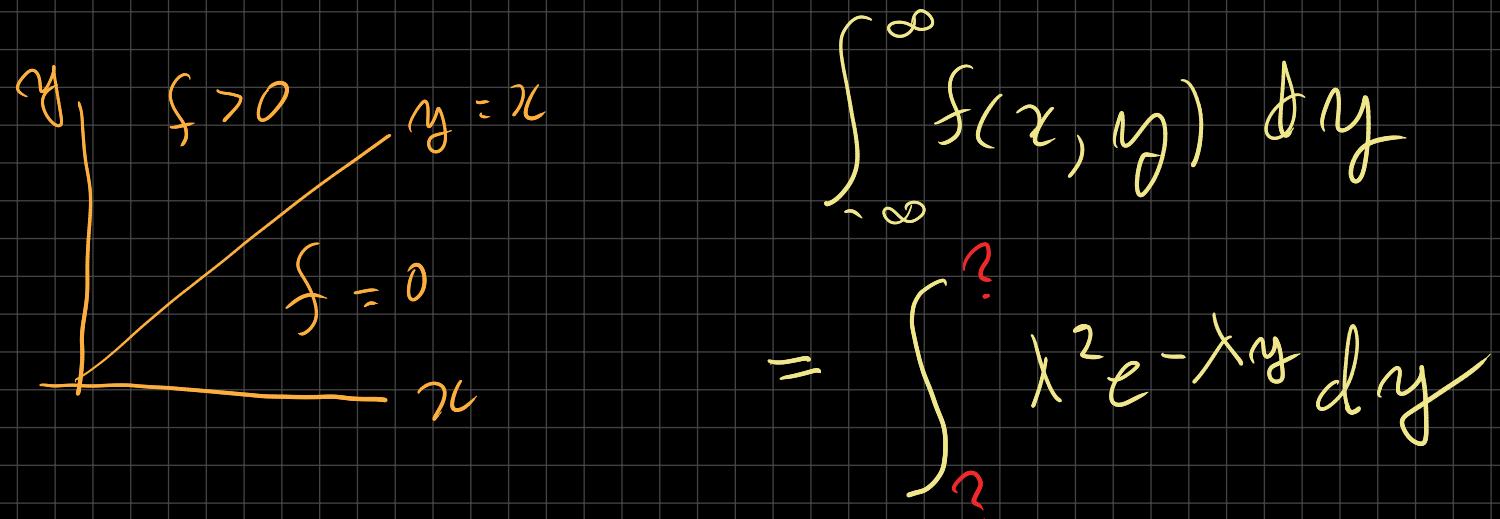
What does $f(x, y)$ look like?

Fix parameter $\lambda = 1$

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



Marginal distribution of X Integrate out Y



$$\int_{-\infty}^{\infty} f(x, y) dy = \int_{?}^{?} x^2 e^{-\lambda y} dy$$

$$= \int_x^{\infty} \lambda^2 e^{-\lambda y} dy = \left[\frac{\lambda^2 e^{-\lambda y}}{-\lambda} \right]_x^{\infty} = 0 - \left(-\lambda e^{-\lambda x} \right)$$

$$= \lambda e^{-\lambda x}, \quad y > 0$$

So X is exponential

$$\begin{aligned}
 f_Y(y) &= \int f(x, y) dx \\
 &= \int_0^y \lambda^2 e^{-\lambda x} dx \\
 &= \left[x \lambda^2 e^{-\lambda x} \right]_0^y \\
 &= y \lambda^2 e^{-\lambda y} = y^{2-1} \frac{\lambda^2 e^{-\lambda y}}{\Gamma(2)}
 \end{aligned}$$

So $\gamma \sim \text{Gamma}$
 with shape parameter
 $\alpha = 2$ and scale
 parameter λ

We will return to copulas on p. 78

and the bivariate normal on p. 81

after talking about independence section 3.4

3.4 Independent random variables

Def: Random variables X_1, \dots, X_n are independent iff joint CDF factors into product of marginal CDF:

$$F(x_1, x_2, \dots, x_n) = F_1(x_1) F_2(x_2) \cdots F_n(x_n)$$

$$\forall x_1, \dots, x_n.$$

Facts: ① If X_1, X_2 are independent RVs then all events involving only X_1 are independent of events involving only X_2 .

i.e. $P(X_1 \in A, X_2 \in B)$

$$= P(X_1 \in A) P(X_2 \in B)$$

(2) Events for any function g_1 of X_1 and g_2 of X_2 are independent:

$$P(g_1(X_1) \in A, g_2(X_2) \in B)$$

$$= P(g_1(X_1) \in A) \times P(g_2(X_2) \in B)$$

③ Discrete case : X_1, \dots, X_n are independent

iff $P(x_1, x_2, \dots, x_n)$
 $= P_1(x_1) \cdots P_n(x_n)$

④ Continuous case : X_1, \dots, X_n are independent

iff $f(x_1, \dots, x_n)$ can be expressed as
 $f_1(x_1) \cdots f_n(x_n)$

Back to p. 81 : Bivariate normal

Univariate $N(\mu, \sigma^2)$

Standard Normal $Z \sim N(0, 1) :$ $f_z(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}$

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

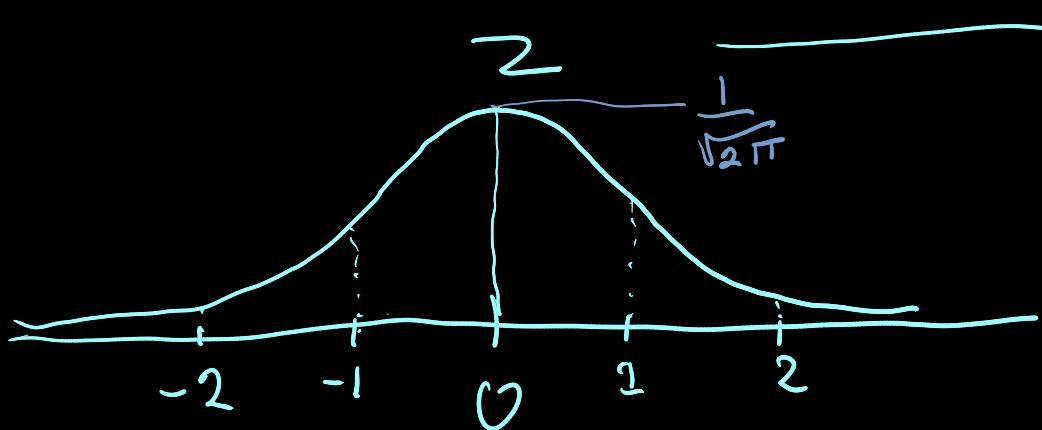
$$X = \mu + \sigma Z \quad Z = (X - \mu) / \sigma$$

$$f_x(x) = f_z((x - \mu) / \sigma) \left| \frac{dZ}{dx} \right| = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} \left(\frac{x-\mu}{\sigma} \right)^2} \times \frac{1}{\sigma}$$

$$= \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{1}{2} \left(\frac{x-\mu}{\sigma}\right)^2\right\}$$

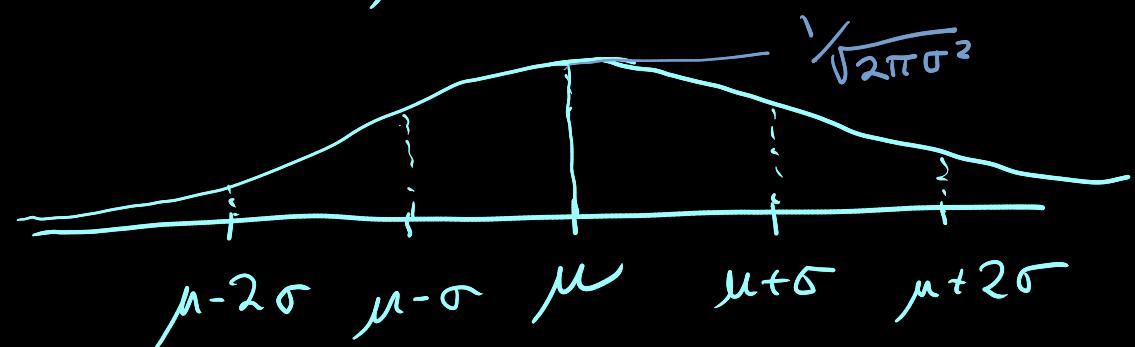
$$\text{OR} = \frac{1}{\sqrt{2\pi}\sigma^2} \exp\left\{-\frac{1}{2\sigma^2} (x-\mu)^2\right\}$$

Standard Normal



$$N(\mu, \sigma^2)$$

$$x = \mu + \sigma z$$



Multivariate normal in stages

Bivariate standard normal

2 independent normals

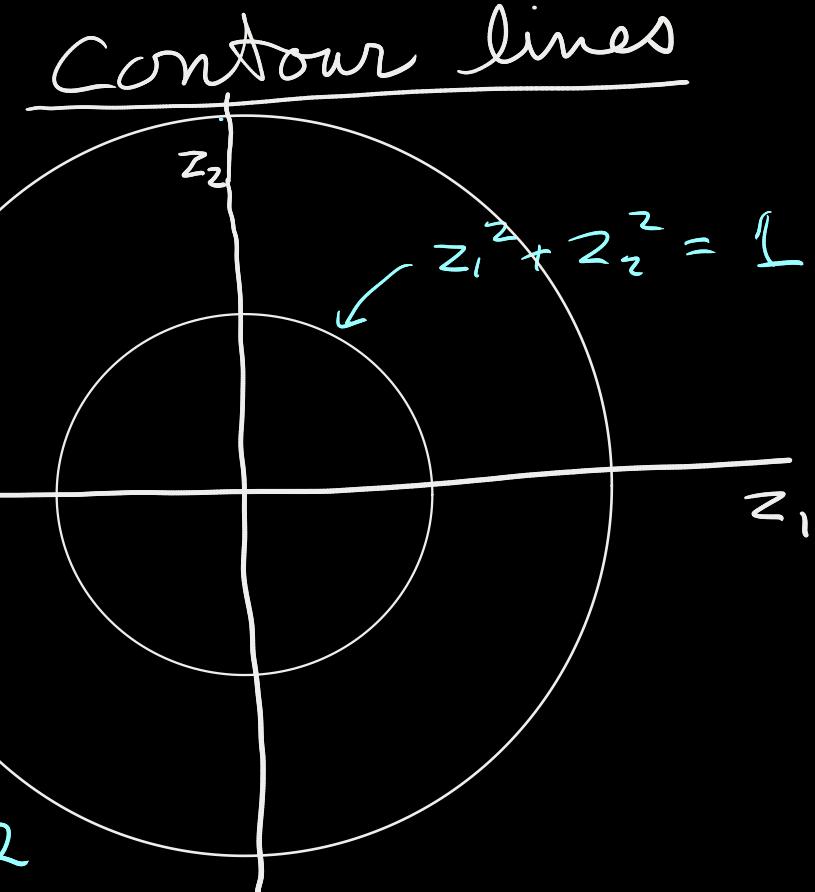
$$(z_1, z_2) \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right)$$

$$\begin{aligned} f(z_1, z_2) &= f(z_1) f(z_2) \\ &= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z_1^2} \times \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z_2^2} \end{aligned}$$

$$= \frac{1}{(2\pi)^{1/2}} \exp \left\{ -\frac{1}{2} (z_1^2 + z_2^2) \right\}$$



lines of
equal altitude
on a mountain
 $z_1^2 + z_2^2 = 2$



$\begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$ as a linear transformation of $\begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix}$

$$\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} + \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix}$$

$$\underline{X} = \underline{\mu} + A \underline{Z}$$

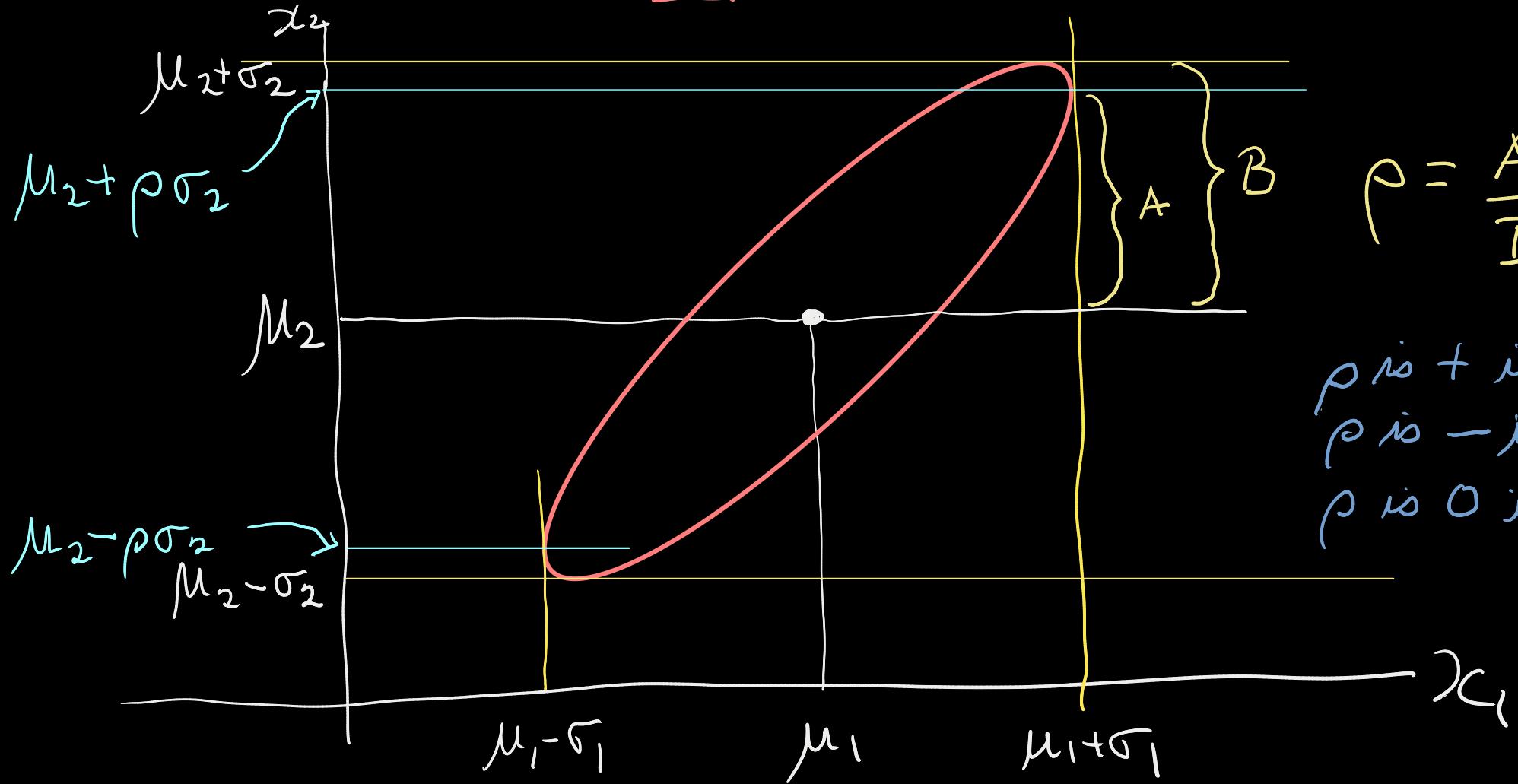
... takes us to the general bivariate normal
in the text book:

$$f(x_1, x_2) = \frac{1}{2\pi \sigma_x \sigma_y \sqrt{1-\rho^2}} \times$$

$$\exp \left\{ - \frac{1}{2(1-\rho^2)} \left[\left(\frac{x_1 - \mu_1}{\sigma_1} \right)^2 + \left(\frac{x_2 - \mu_2}{\sigma_2} \right)^2 - 2\rho \left(\frac{x_1 - \mu_1}{\sigma_1} \right) \left(\frac{x_2 - \mu_2}{\sigma_2} \right) \right] \right\}$$

Contours

$$\left[\dots \right] = 1$$



ρ is + if
 ρ is - if
 ρ is 0 if

Back to P 78 : Copula

CDF for X_1, X_2 on $[0, 1] \times [0, 1]$

so marginals for X_1, X_2 are $U(0, 1)$

$$C(x_1, x_2)$$

Note $C(x_1, \infty) = C(x_1, 1) = x_1$
if $0 < x_1 < 1$

Same for $C(\infty, x_1)$

