Prob Stats L11

Covariance & Correlation

Co-Variance

March 16, 2023

Expectation for Joint R.V.

X, Y

- joint pdf $f_{X,Y}(x,y) = f(x,y)$

- function on R.V.s
  $Z = g(X,Y)$

$Z = g(X,Y) =$
ex: $3X + 2XY^2$

$$E[Z] = E[g(X,Y)] = \int_{x=-\infty}^{\infty} \int_{y=-\infty}^{\infty} g(x,y) \cdot f(x,y)\, dy\, dx$$

$$E[X] = \int_{x=-\infty}^{\infty} x\, f_X(x)\, dx$$

Revisit Linearity of Expectation

$g(x,y) = rX + sY$

$$E[rX+sY] = \sum_i \sum_j (r a_i + s b_j) \cdot \Pr(X=a_i, Y=b_j)$$

$$= r \sum_i \sum_j a_i \cdot \Pr(X=a_i, Y=b_j) + s \sum_i \sum_j b_j \cdot \Pr(X=a_i, Y=b_j)$$

$$= r \sum_i a_i \Big(\underbrace{\sum_j \Pr(X=a_i, Y=b_j)}_{\text{marginalizing over } Y} \Big) + s \sum_j b_j \Big(\underbrace{\sum_i \Pr(X=a_i, Y=b_j)}_{\text{marg... } X} \Big)$$

$$= r \underbrace{\sum_i a_i \Pr(X=a_i)}_{E[X]} + s \underbrace{\sum_j b_j \Pr(Y=b_j)}_{E[Y]}$$

$$= r E[X] + s E[Y]$$

Covariance

2 RVs: X, Y

measures how X, Y co-vary

$$\mathrm{Cov}(X,Y) = E\left[(X - E[X]) \cdot (Y - E[Y])\right]$$

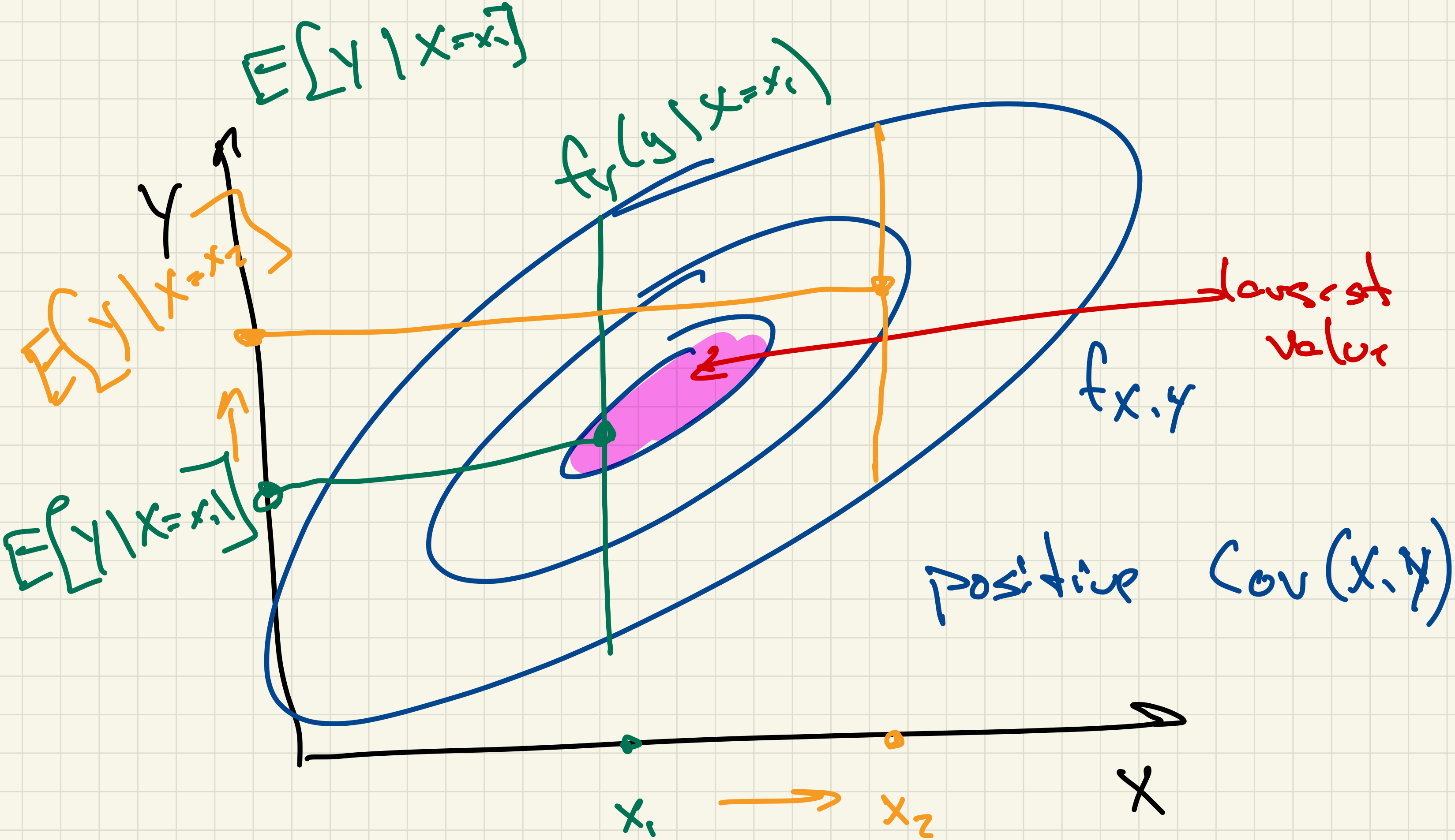
$$\mathrm{Cov}(X,X) = \mathrm{Var}(X)$$

- if $\mathrm{Cov}(X,Y) > 0$

  then as X gets larger

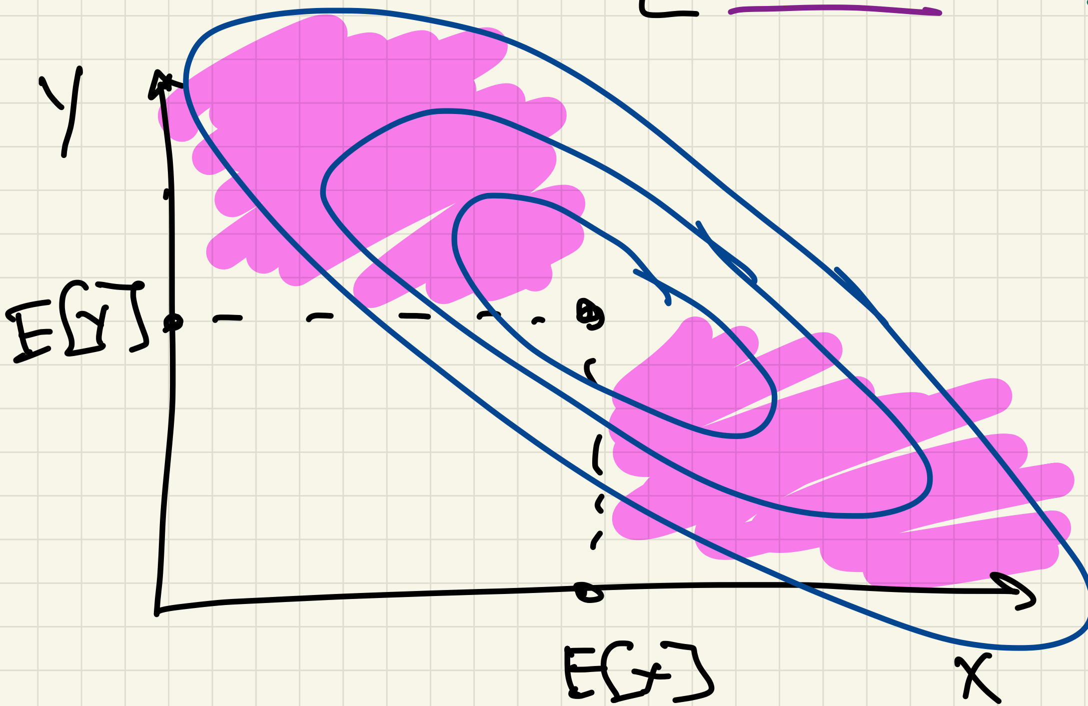
  Y tends to also

$$\mathrm{Var}(X) = E\left[(X - E[X])^2\right]$$

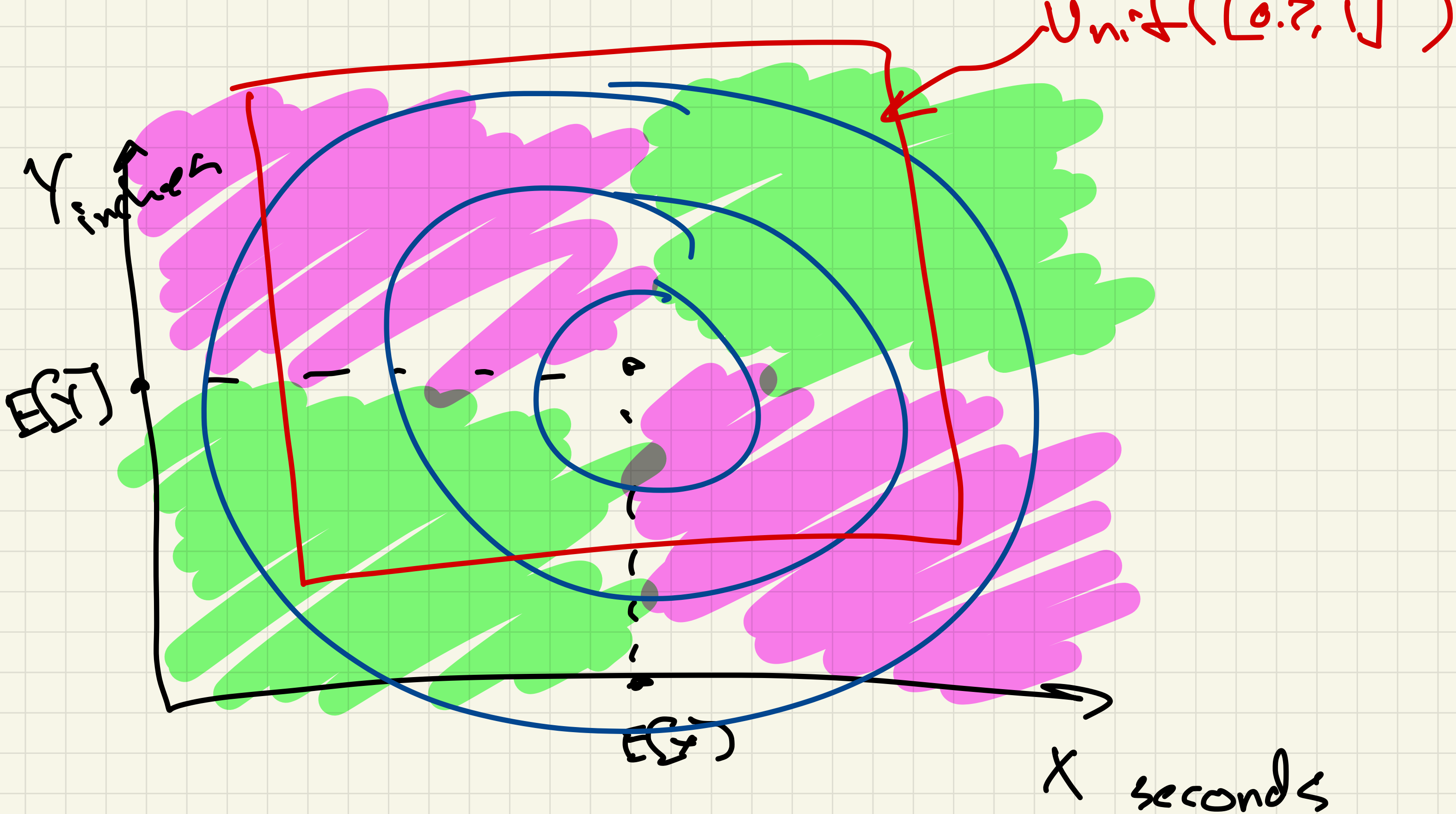

E[Y | X=x₂]
E[Y | X=x₂]
f_Y(y | X=x₁)
Y
E[Y | X=x₁]
densest value
f_X,Y
positive Cov(X,Y)
X
x₁
x₂

what if $Cov(X, Y) < 0$

$$Cov(X, Y) = E\left[\underline{(X - E[X])} \cdot \underline{(Y - E[Y])}\right]$$



What if $Cov(X, Y) = 0$



If X, Y Independent

then $Cov(X, Y) = 0$

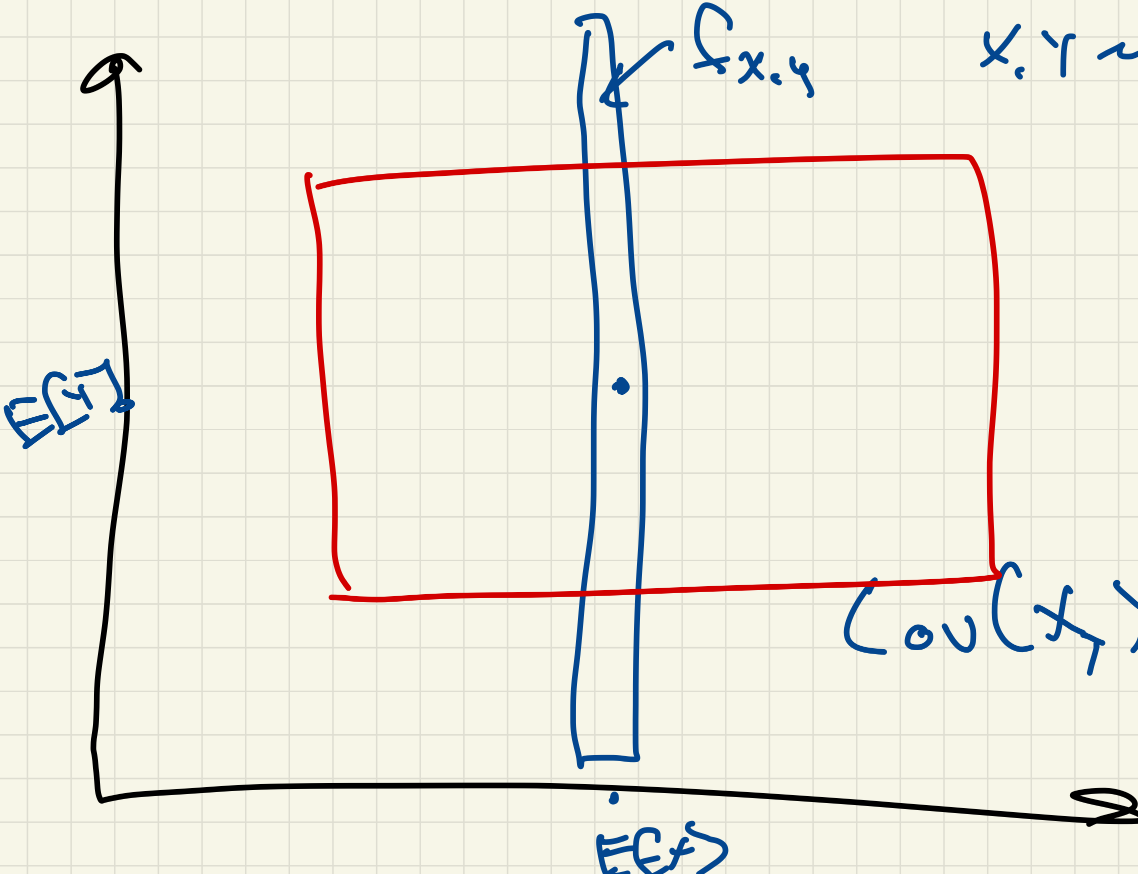
but

if $Cov(X, Y) = 0$

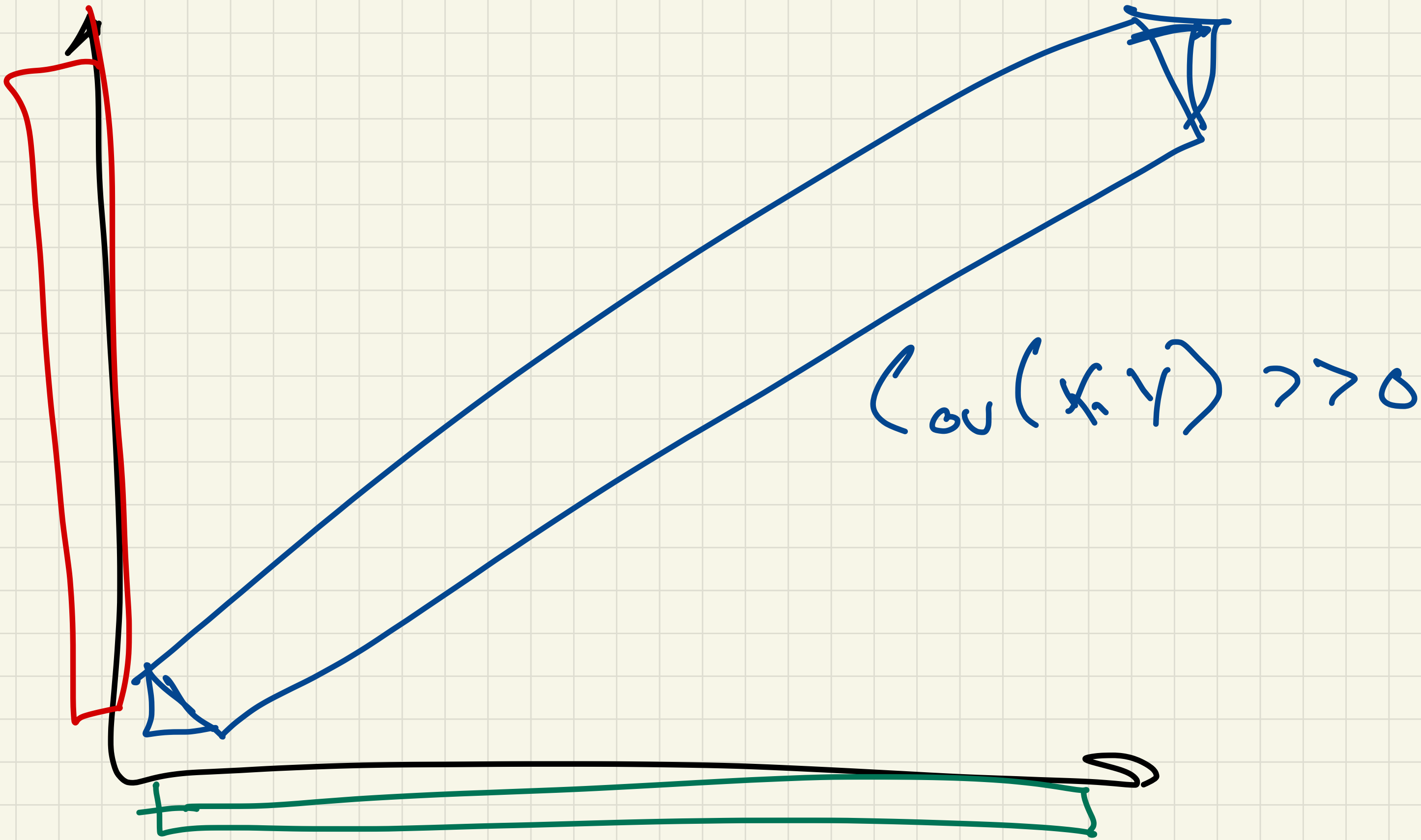
it might not mean

that X, Y are Independent.

$X, Y \sim \mathrm{Unif}\left(x = 0.5 \pm, \; y \in [0,1]\right)$



$\mathrm{Cov}(X,Y) = 0$


Cov(X, Y) >> 0

Correlation $\rho(X, Y)$

because we don't know how big is big w/ $Cov(X,Y)$

$$\rho(X,Y) = \frac{Cov(X,Y)}{\sqrt{Var(X) \cdot Var(Y)}}$$

m · s

m^2 $\quad$ s^2

$\sqrt{m^2 \cdot s^2}$

X meters

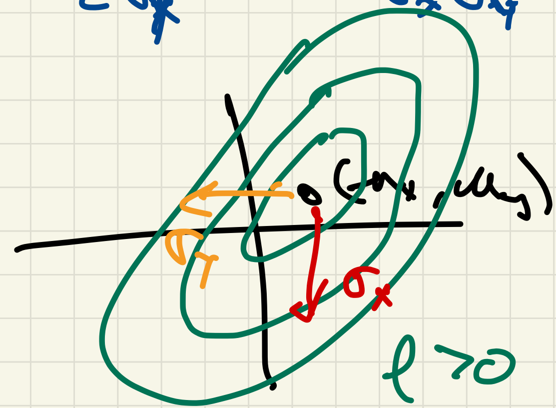
Y seconds

$$\rho(X,Y) \in [-1, 1]$$

Gaussian = multiVariate Normal   d=2

$$f(x,y) = \boxed{\text{normalization}} \cdot \exp(\ldots)$$

$$\exp\left(-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_x)^2}{2\sigma_x^2} + \frac{(y-\mu_y)^2}{2\sigma_y^2} + \frac{2\rho(x-\mu_x)(y-\mu_y)}{\sigma_x\sigma_y}\right]\right)$$



$\mu_x = E[X]$ $\mu_y = E[Y]$ $\rho = \rho(X,Y)$ $\sigma_x^2 = Var(X)$ $\sigma_y^2 = Var(Y)$

Review for Quiz 8

Continuous Joint RVs

pdf $f_{X,Y}(x,y) = f(x,y)$

1. $f(x,y) \geq 0$
2. $\displaystyle\int_{x=-\infty}^{\infty} \int_{y=-\infty}^{\infty} f(x,y)\, dy\, dx = 1$
3. $\Pr(a \leq X \leq b,\ c \leq Y \leq d) =$

$$= \int_{x=a}^{b} \int_{y=c}^{d} f(x,y)\, dy\, dx$$

Review part 2

Covariance

$$Cov(X,Y) = E[(X-E[X])\cdot(Y-E[Y])]$$
$$= E[XY] - E[X]E[Y]$$

if $Cov(X,Y) > 0$

then if X increases

Y expects to increase

Correlation

$$\rho(X,Y) = \frac{Cov(X,Y)}{\sqrt{Var(X)\cdot Var(Y)}}$$

$$\rho(X,Y) \in [-1,1]$$

Review

Independence

if X, Y independent

then $\Longrightarrow$ $Cov(X, Y) = 0$

$\rho(X, Y) = 0$

but if $Cor(X, Y) = 0$

maybe not independent