$Var(U) = Cov(U, U) = Cov(\underline{a}^T \underline{X}, \underline{a}^T \underline{X}) = \underline{a}^T \underline{\underline{\Sigma}}\, \underline{a}$

$= \sum_{i=1}^{n} a_i^2 \sigma_i^2$

Basics:

- $Var(X) = Cov(X, X)$
- $Cov(X, Y) = Cov(Y, X)$

Conditional Distributions:

- Discrete: $P(X=x \mid Y=y) = \dfrac{P(X=x \wedge Y=y)}{P(Y=y)}$
- Continuous: $f_{X|Y}(x|y) = \dfrac{f_{XY}(x,y)}{f_Y(y)}$

why? We want to predict
- Health
- elections
- wealth
- Future

Ex Discrete

X \ Y	0	1	2	
1	.15	.10	0	.25
2	.25	.30	.20	.75
	.4	.40	.20	

joint pmf

$P(X=x \mid Y=y)$

X \ Y	0	1	2
1	$.15/.40$	$\frac{.10}{.40}$	0
2	$.25/.40$	$\frac{.30}{.40}$	$\frac{.20}{.20}$

$E(X=x \mid Y=y)$

$\boxed{Ex}$ $E(X \mid Y=1) = \sum_{\text{all } x} x\, P(X=x \mid Y=1) = 1 \cdot \frac{1}{4} + 2 \frac{3}{4}$

$= \frac{7}{4}$

390	391
- Data Scraping	
- Modeling	

Today
1) Independence
2) Covariance
3) Cond. Dist.

Check X and Y are independent:

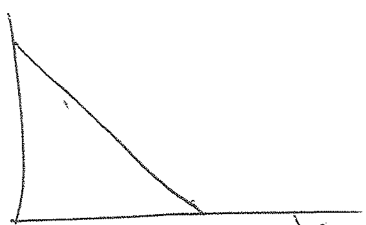
- Discrete $P(X=x, Y=y) = P_X(x)\,P_Y(y)$
- Continuous $F_{XY}(x,y) = F_X(x)\,F_Y(y)$

$$\Rightarrow f_{XY}(x,y) = f_X(x)\,f_Y(y) \quad *$$

- Check the support/joint contours

X & Y are not independent if the support is not "□" or "O"

Ex. $f_{XY}(x,y) = 60x^2y \quad 0<x<1,\ 0<y<1,\ x+y<1$



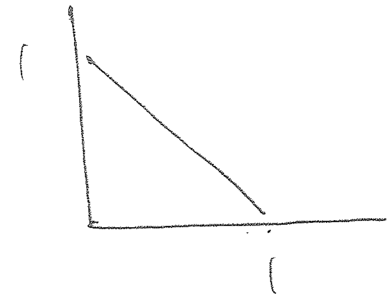
Ex. #1

X \ Y	0	1	2	
1	.15	.10	0	.25
2	.25	.30	.20	.75
	.40	.40	.20	

Check independence

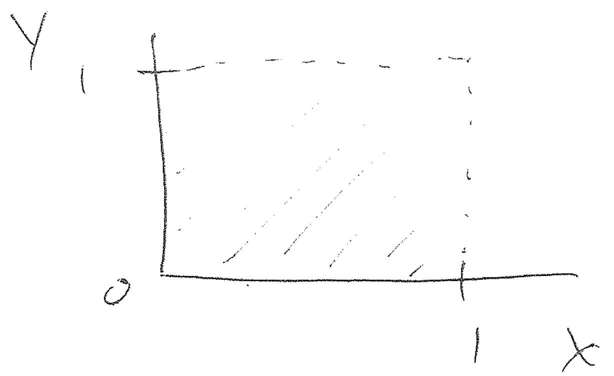
$$P(X=1, Y=2) = 0 \neq P(X=1)\,P(Y=2)$$

$\boxed{\text{Ex. \#2}}$ $f_{XY}(x,y) = 60x^2y, \; 0<x<1, \; 0<y<1, \; x+y<1$



Not rectangle. X and Y are dependent

$\boxed{\text{Ex \#3}}$ $f_{XY} = \begin{cases} x+y, & 0<x<1, \; 0<y<1 \\ 0 & \text{o.w.} \end{cases}$



$f_X(x) = \int_0^1 (x+y)\,dy = \cancel{y+\frac{1}{2}, \; 0<y<1}$ $x+\frac{1}{2}, \; 0<x<1$

$f_Y(y) = \int_0^1 (x+y)\,dx = y+\frac{1}{2}, \; 0<y<1$

$f_{XY} = (x+y) \neq f_X f_Y = (x+\frac{1}{2})(y+\frac{1}{2})$

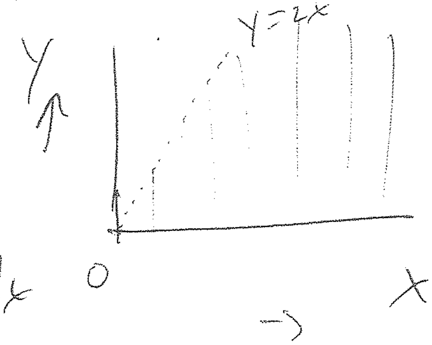
$\Rightarrow$ X and Y are dependent OR not independent

Ex #4 Suppose X and Y are independent

$X \sim \exp(\alpha)$, $Y \sim \exp(\beta)$ Find $P(2X > Y)$.

$$f_X(x) = \alpha e^{-\alpha x},\ x > 0 \quad ; \quad f_Y(y) = \beta e^{-\beta y},\ y > 0$$

$$f_{XY} = f_X f_Y$$



$$P(2X > Y) = \int_0^\infty \int_0^{2x} f_{XY}(x,y)\, dy\, dx$$

$$= \int_0^\infty \int_0^{2x} f_X(x) f_Y(y)\, dy\, dx = \int_0^\infty \alpha e^{-\alpha x} \left(\int_0^{2x} \beta e^{-\beta y}\, dy \right) dx$$

$$= \frac{2\beta}{\alpha + 2\beta}$$

$\boxed{\text{Ex \#5}}$ Let X and Y be independent R.v.'s

$X \sim \text{Geometric}(p = \frac{1}{3})$, $Y \sim \text{Poisson}(\lambda = 3)$

$$P_X(x) = \frac{1}{3}\left(\frac{2}{3}\right)^{x-1}, \; x = 1, 2, \ldots \qquad P_Y(y) = \frac{3^y e^{-3}}{y!}, \; y = 0, 1, 2, \ldots$$

a) $$P(X=Y) = \sum_{k=1}^{\infty} P_{XY}(X=k, Y=k) = \sum_{k=1}^{\infty} P_X(X=k)\, P_Y(Y=k)$$

$$= \sum_{k=1}^{\infty} \frac{1}{3}\left(\frac{2}{3}\right)^{k-1} \frac{3^k e^{-3}}{k!} = e^{-3} \sum_{k=1}^{\infty} \frac{2^{k-1}}{k!} \cdot \frac{2}{2} = \frac{e^{-3}}{2} \sum_{k=1}^{\infty} \frac{2^k}{k!}$$

starts at 1

$$= \frac{e^{-3}}{2}\left(\underbrace{\sum_{k=0}^{\infty} \frac{2^k}{k!}} - 1\right)$$

$$= \frac{e^{-3}}{2}\left(e^2 - 1\right) \approx .159$$

b) $$P(X = 2Y) = \sum_{k=1}^{\infty} P_{XY}(X=2k, Y=k) = \sum_{k=1}^{\infty} P_X(X=2k)\, P_Y(Y=k)$$

$$= \sum_{k=1}^{\infty} \frac{1}{3}\left(\frac{2}{3}\right)^{2k-1} \frac{3^k e^{-3}}{k!} = \sum_{k=1}^{\infty} \frac{1}{3}\left(\frac{2}{3}\right)^{2k} \frac{3}{2} \frac{3^k e^{-3}}{k!}$$

$$= \frac{e^{-3}}{2} \sum_{k=1}^{\infty} \frac{2^{2k}}{3^{2k}} \frac{3^k}{k!} = \frac{e^{-3}}{2} \sum_{k=1}^{\infty} \frac{4^k}{3^k} \frac{1}{k!} = \frac{e^{-3}}{2} \sum_{k=1}^{\infty} \frac{\left(\frac{4}{3}\right)^k}{k!}$$

$$= \frac{e^{-3}}{2}\left(e^{4/3} - 1\right)$$

Covariance

$$Cov(X,Y) = E\left[(X-\mu_X)(Y-\mu_Y)\right]$$
$$= E(XY) - [E(X)][E(Y)]$$
$$= E(XY) - \mu_X \mu_Y$$

* Measure of unstandardized linear association

Correlation

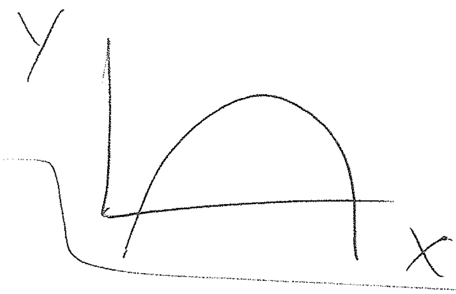
$$\rho_{XY} = \frac{Cov(X,Y)}{\sigma_X \sigma_Y}$$

* standardized measure of linear association

If X and Y are independent $\Rightarrow Cov(X,Y) = 0$

pf $E(XY) = EX\,EY$

The converse is not true.



Covariances of linear combs of r.v.s

- $Cov(aX + bY, cX + dY)$?

$$= ac\,Var(X) + bd\,Var(Y) + (ad + bc)\,Cov(X,Y) \quad *$$

- $Cov(a_1 X_1 + a_2 X_2,\ b_1 X_1 + b_2 X_2)$

$$\underline{X} = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}, \quad \underline{a} = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}, \quad \underline{b} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

$$\mathrm{Cov}(a_1 X_1 + a_2 X_2,\ b_1 X_1 + b_2 X_2) = \underline{a}^T \underline{\underline{\Sigma}}\, \underline{b} \quad *$$

$$\underline{X} = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix},\quad \underline{a} = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix},\quad \underline{b} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix},\quad \underline{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}$$

- symmetric $\underline{\underline{\Sigma}} = \underline{\underline{\Sigma}}^T$
- p.s.d.

$\underline{\underline{\Sigma}}$ = variance - covariance matrix

$$= \begin{bmatrix} \mathrm{Var}(X_1) & \mathrm{Cov}(X_1, X_2) \\ \mathrm{Cov}(X_1, X_2) & \mathrm{Var}(X_2) \end{bmatrix}$$

$$= E\left[(\underline{X} - \underline{\mu})(\underline{X} - \underline{\mu})^T\right] \quad \left(\text{Hint } \underline{X} - \underline{\mu} = \begin{bmatrix} X_1 - \mu_1 \\ X_2 - \mu_2 \end{bmatrix}\right)$$

$\boxed{\text{Ex}}$ a) Find in terms of σ_X^2, σ_Y^2, σ_{XY}

$$\mathrm{Cov}(2X + 3Y,\ X - 2Y) \qquad \underline{a} = \begin{bmatrix} 2 \\ 3 \end{bmatrix},\quad b = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$$

$$= \underline{a}^T \underline{\underline{\Sigma}}\, \underline{b} \qquad \underline{\underline{\Sigma}} = \begin{bmatrix} \sigma_X^2 & \sigma_{XY} \\ \sigma_{XY} & \sigma_Y^2 \end{bmatrix}$$

$$\sigma_X^2 = 1,\quad \sigma_Y^2 = 2,\quad \sigma_{XY} = 1$$

$$= \cancel{\begin{bmatrix} 2 \\ 3 \end{bmatrix}^T \Sigma}$$

$$= \begin{bmatrix} 2 & 3 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ -2 \end{bmatrix}$$

b) Suppose $X_1, \ldots, X_n$ are independent w/ ~~mean~~ variances $\sigma_1^2, \ldots, \sigma_n^2$. Let $U = \sum_{i=1}^{n} a_i X_i$. Find $\mathrm{Var}(U)$. $= \mathrm{Cov}(U, U)$

$$\underline{a} = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix},\quad \underline{a} = \underline{b},\quad \underline{\underline{\Sigma}} \text{ is diagonal} \Rightarrow \underline{\underline{\Sigma}} = \begin{bmatrix} \sigma_1^2 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & \cdots 0 & \sigma_n^2 \end{bmatrix}$$

Conditional Expectations

~~E(X~~ $E(Y|X) = \sum_{\text{all } y} y \, p(Y=y \mid X=x)$

$E(Y|X) = \int y f_Y(y|x)\, dy$