Bayes' Rule :

Likelihood Prior

$$\underbrace{P(A|B)}_{\text{Posterior}} = \frac{P(A,B)}{P(B)} = \frac{P(B|A) \cdot P(A)}{P(B)}$$

$$= \frac{P(B|A) \cdot P(A)}{\boxed{\sum_A P(B,A)}}$$

knife

Eve

marginalization

$$= \frac{P(B|A) \cdot P(A)}{\sum_A P(B|A)\, P(A)}$$

$\alpha \equiv P(\text{knife} \mid \text{Eve})$

$P_0 \equiv P(\text{Eve} \mid \text{knife})$

$\alpha \equiv P(\text{article} \mid \text{CNN})$

$P_0 \equiv P(\text{CNN} \mid \text{article})$ ✓

$P(\text{evidence} \mid \text{hypo})$

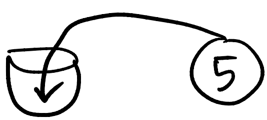
$P(\text{hypo} \mid \text{evidence})$

Chain Rule

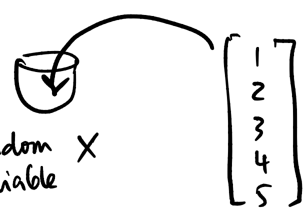
$$P(A,B,C) = P(A|B,C)\; P(B,C)$$

$$= P(A|B,C)\; P(B|C)\; P(C)$$

Random Variables :



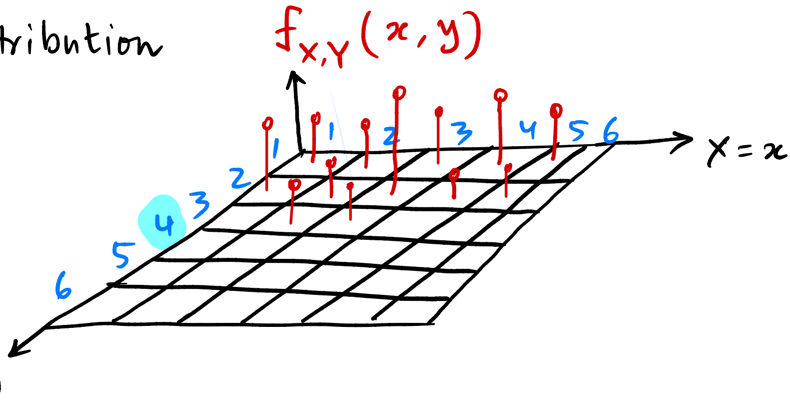
Variable X



Random Variable X

$$E[X] = \mu_X = \sum_i x_i\, f_X(x_i) \qquad Var(X) = E[(X-\mu_X)^2]$$

Joint Distribution



$$P\left(X=\begin{bmatrix}1\\2\\3\\\vdots\\6\end{bmatrix} \middle| Y=4\right)$$

Conditional Distribution

$$P(X|Y) = \frac{P(X\,Y)}{P(Y)}$$

$$P(X|Y=2) = \frac{P\left(X=\begin{bmatrix}1\\2\\3\\4\\5\\6\end{bmatrix}, Y=2\right)}{P(Y=2)}$$

$$P(Z) = \sum_M P(Z,M)$$

$$= \frac{P(X=[\;], Y=2)}{\sum_X P(X=[\;]\; Y=2)}$$

- Independence :

  X and Y are independent when

  $$P(X,Y) = P(X)P(Y)$$ for all values of X, Y

  If X, Y are indep

  $$P(X|Y) = P(X)$$

- Variance and Covariance

$$\begin{aligned}
Var(X) &= E[(X-\mu_x)^2] \\
&= E[X^2 + \mu_x^2 - 2X\mu_x] \\
&= E[X^2] + \mu_x^2 - 2\mu_x^2 \\
&= E[X^2] - E[X]^2
\end{aligned}$$

$$\begin{aligned}
Cov(X,Y) &= E[(X-\mu_X)(Y-\mu_Y)] \\
&= E[XY - X\mu_y - \mu_x Y + \mu_x \mu_Y] \\
&= E[XY] - E[X]\mu_Y - \mu_x E[Y] + \mu_x \mu_Y \\
&= E[XY] - E[X]E[Y]
\end{aligned}$$

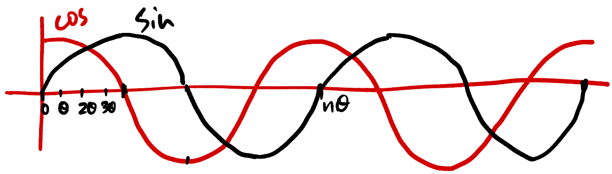
- When $Cov(X,Y) = 0$

  i.e., $E[XY] = E[X]E[Y]$ called uncorrelated

- Correlation coefficient $= \dfrac{\mathrm{cov}(X,Y)}{\sqrt{\mathrm{Var}(X)\,\mathrm{Var}(Y)}} \in [-1, 1]$

- Meaning of independence and correlation (or uncorrelation) in plain English?

- Two RVs X and Y are uncorrelated. Are they also independent?



$$\begin{bmatrix} \cos\theta \\ \cos 2\theta \\ \cos 3\theta \\ \vdots \\ \cos n\theta \end{bmatrix} \quad \begin{bmatrix} \sin\theta \\ \sin 2\theta \\ \sin 3\theta \\ \vdots \\ \sin n\theta \end{bmatrix}$$

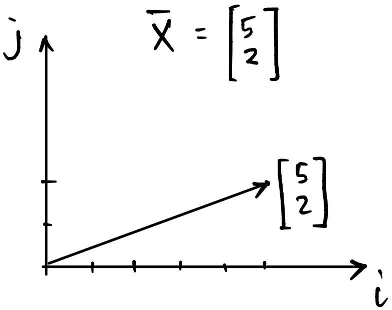
- Will 2 independent RVs be always uncorrelated?

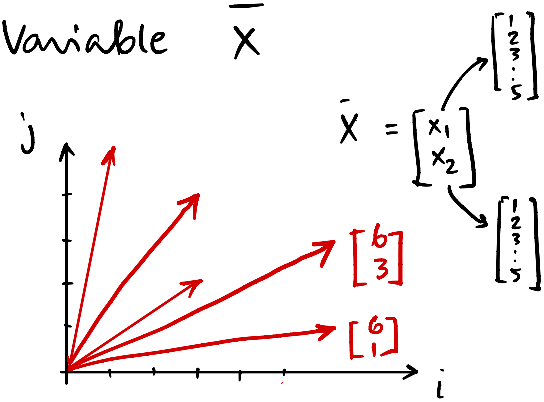
Yes.

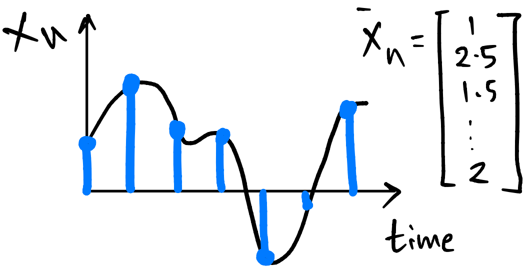
After class discussion

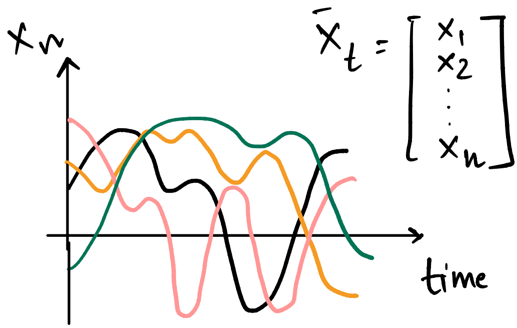
$$E[XY] = \sum_y \sum_x xy\, f_{XY}(x,y)$$
$$\sum \sum xy \boxed{f_X(x) f_Y(y)}$$
$$= \sum_y \sum_x x f_X(x)\, y f_Y(y)$$
$$= \sum y f_Y(y) \sum x f_X(x) = E[X]E[Y]$$

Vectors of Random Variable $\bar{X}$



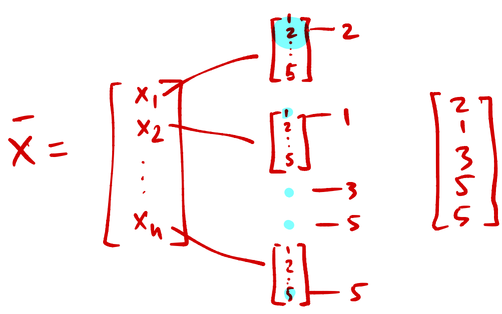






Expected value of Random Vector $E[\bar{X}]$

$$E[\bar{X}] = \begin{bmatrix} E[x_1] \\ E[x_2] \\ \vdots \\ E[x_n] \end{bmatrix}$$



Vector version of variance → Covariance matrix

$$Cov(\bar{x}) = Cov\left(\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}\right) = \text{Pairwise covariances} = Cov(x_i x_j) \quad \forall i,j$$

$$\begin{bmatrix} cov(x_1x_1) & cov(x_1x_2) & \cdots & cov(x_1x_n) \\ \vdots & & & \\ cov(x_nx_1) & \cdots & & cov(x_nx_n) \end{bmatrix}$$

Covariance matrix $Cov(x) = E[xx^T] - E[x]E[x^T]$

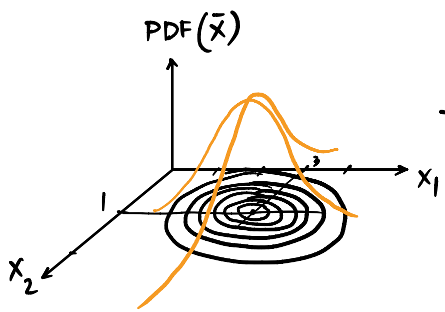
$$E[xx^T] = E\left[\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}[x_1\ x_2\ \cdots\ x_n]\right] = \begin{bmatrix} E[x_1x_1] & E[x_1x_2] & \cdots & E[x_1x_n] \\ E[x_2x_1] & E[x_2x_2] & \cdots & \\ & & & \\ E[x_nx_1] & & & E[x_nx_n] \end{bmatrix}$$

$- E[x_1]E[x_2^T]$

$$cov(\bar{x}) = \underline{\underline{\Sigma_1}} = \begin{bmatrix} \sigma_1^2 & cov(x_1x_2) & cov(x_1x_3) & \cdots & cov(x_1x_n) \\ cov(x_2x_1) & \sigma_2^2 & & & \\ cov(x_3x_1) & cov(x_3x_2) & \ddots & & \\ cov(x_nx_1) & \cdots & & & \sigma_n^2 \end{bmatrix}$$

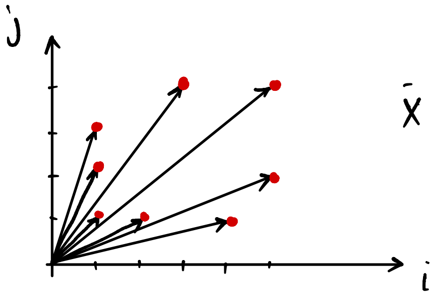
Visualize a 2D jointly Gaussian random variable.

where x_1 and x_2 are uncorrelated and $N(3,1)$ and $N(1,1)$



$$f_X(x) = \frac{1}{\sqrt{2\pi}\,|\Sigma_1|^{1/2}} e^{-\frac{(x-\mu_x)\Sigma_1^{-1}(x-\mu_x)^T}{2}}$$

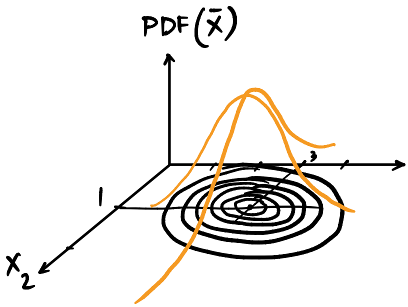
- Interpret 2D data points as a realization of rand. vectors.



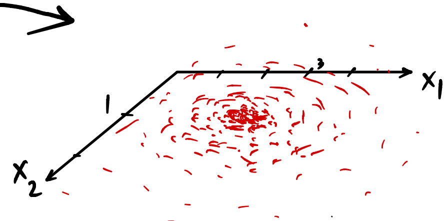
$$\bar{X} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \quad x_1 \rightarrow \begin{bmatrix} 1 \\ 2 \\ 3 \\ \vdots \\ 5 \end{bmatrix} \quad x_2 \rightarrow \begin{bmatrix} 1 \\ 2 \\ 3 \\ \vdots \\ 5 \end{bmatrix}$$

$$cov(\bar{x}) = \begin{bmatrix} \sigma_1^2 & cov(x_1,x_2) \\ cov(x_2,x_1) & \sigma_2^2 \end{bmatrix} \rightarrow \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix}$$

So data points have a model now $\rightarrow$ helps for analysis

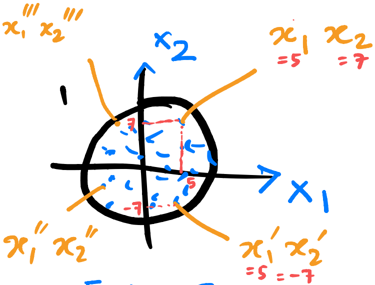


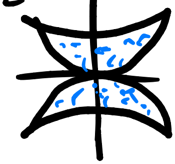
generate data $\rightarrow$

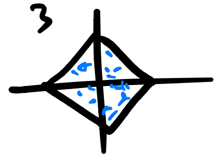


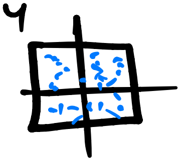
Is the data uncorrelated ↙

How about these data. Label them as
- { Correlated : C, Uncorrelated : U }
- { Independent : I, Dependent : D }



2


3


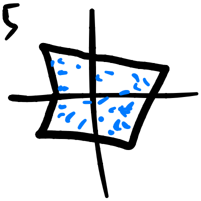
4


$$E[X_1X_2] = 0$$

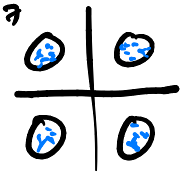
$$x_1x_2 + x_1'x_2' + x_1''x_2'' + \cdots\cdots = ?$$

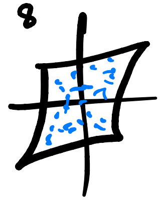
Yes, uncorrelated since they cancel out.

$P(x_2|x_1)$ is $\neq$ $P(x_2)$ $\Rightarrow$ Not independent

5


6


7


8


9


- Application : covariance of received mic. data.

  i.e., say $x = A\bar{s} + \bar{n}$

  $$\mathrm{Cov}(x) = E[xx^T]$$

$$P(A \mid Red) = \frac{P(A, Red)}{P(Red)} = \frac{P(Red \mid A)\, P(A)}{P(Red)}$$

$$= \frac{\frac{1}{3} \cdot \frac{1}{2}}{P(Red)} = \frac{\frac{1}{6}}{P(Red, A) + P(Red, B)}$$

$$= \frac{1}{6} \cdot \frac{1}{P(Red \mid A)\, P(A) + P(Red \mid B)\, P(B)}$$

$$= \frac{1}{6} \cdot \frac{1}{\frac{1}{3} \cdot \frac{1}{2} + \frac{2}{3} \cdot \frac{1}{2}}$$

$$= \frac{1}{6} \cdot \frac{1}{\frac{1}{6} + \frac{1}{3}} = \frac{1}{6} \cdot \frac{1}{\frac{3}{6}}$$

$$= \frac{1}{6} \cdot 2 = \frac{1}{3} \quad //$$

Likelihood is self evident (given hypothesis, you know the chance of the evidence)

↳ P(article | CNN)

Posterior requires you to know more about other hypothesis (given evidence, you need to understand how the evidence relates to other hypothesis)

↳ P(CNN | article)