Bayes' Rule :

Likelihood   Prior

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

Posterior

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

marginalization

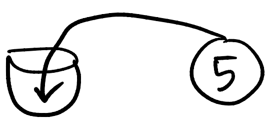
$$= \frac{P(B|A) \cdot P(A)}{\sum_{A} P(B|A) P(A)}$$

- Chain Rule

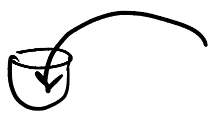
$$P(A,B,C) = P(A|B,C)\, P(B,C)$$

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

- Random Variables :



Variable X

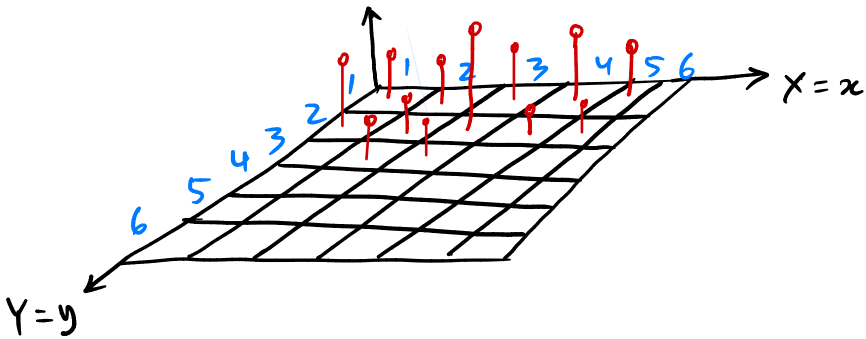


Random Variable X

$$\begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{bmatrix}$$

$$E[X] = \mu_x = \sum_{i} x_i f_X(x_i) \qquad Var(X) = E[(X-\mu_x)^2]$$

Joint Distribution



Conditional Distribution

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

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

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

- Independence :

  X and Y are independent when

  $$P(X, Y) = P(X)\,P(Y) \quad \text{for all values of } X, Y$$

- Variance and Covariance

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

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

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

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

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

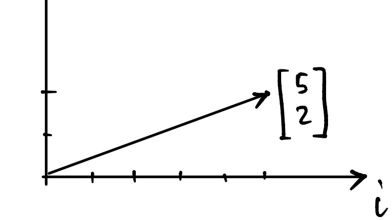
- Meaning of independence and correlation (or uncorrelation) in plain English?

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

- Will 2 independent RVs be always uncorrelated?

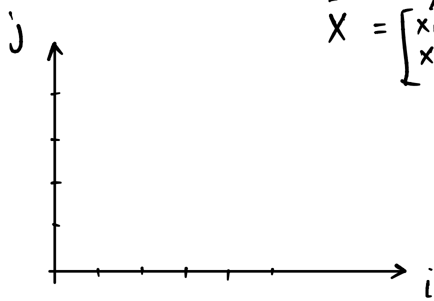
Vectors of Random Variable $\bar{X}$

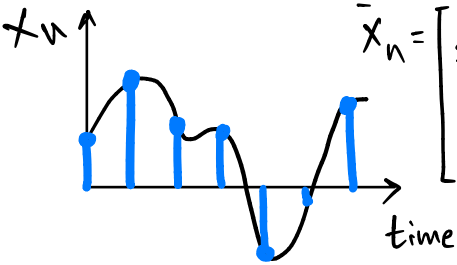
$$\bar{X} = \begin{bmatrix} 5 \\ 2 \end{bmatrix}$$



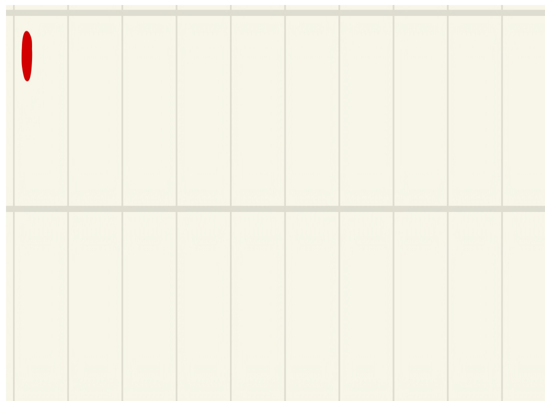
$$\bar{X} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$x_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \\ \vdots \\ 5 \end{bmatrix} \qquad x_2 = \begin{bmatrix} 1 \\ 2 \\ 3 \\ \vdots \\ 5 \end{bmatrix}$$





$$\bar{X}_n = \begin{bmatrix} 1 \\ 2.5 \\ 1.5 \\ \vdots \\ 2 \end{bmatrix}$$



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

$$E[\bar{X}] = \begin{bmatrix} \; \end{bmatrix}$$

- Vector version of variance → Covariance matrix

$$Cov(\bar{x}) = cov\left(\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}\right) = \text{Pairwise covariances}$$

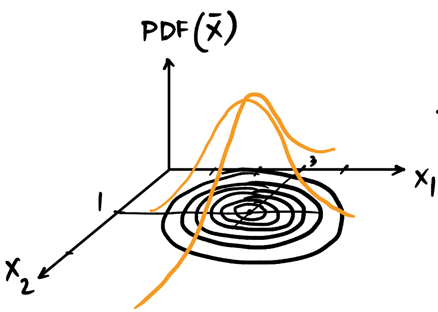
$$=$$

Covariance matrix $Cov(x) =$

$E[XX^T] =$

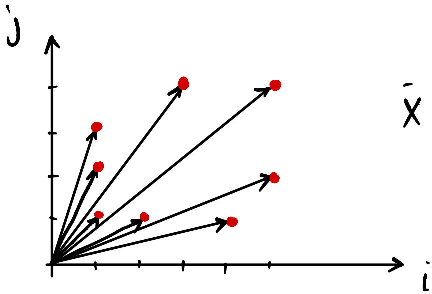
$cov(\bar{x}) = \Sigma_1 =$

- Visualize a 2D jointly Gaussian random variable. where x_1 and x_2 are uncorrelated and $N(3,1)$ and $N(1,1)$



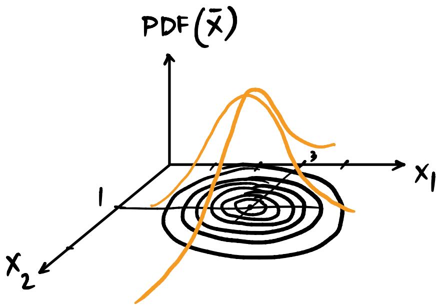
$f_X(x) =$

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

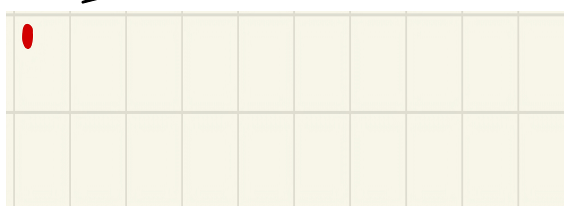


$$\bar{X} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \qquad \begin{bmatrix} 1 \\ 2 \\ 3 \\ \vdots \\ 5 \end{bmatrix} \qquad \begin{bmatrix} 1 \\ 2 \\ 3 \\ \vdots \\ 5 \end{bmatrix}$$

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



generate data

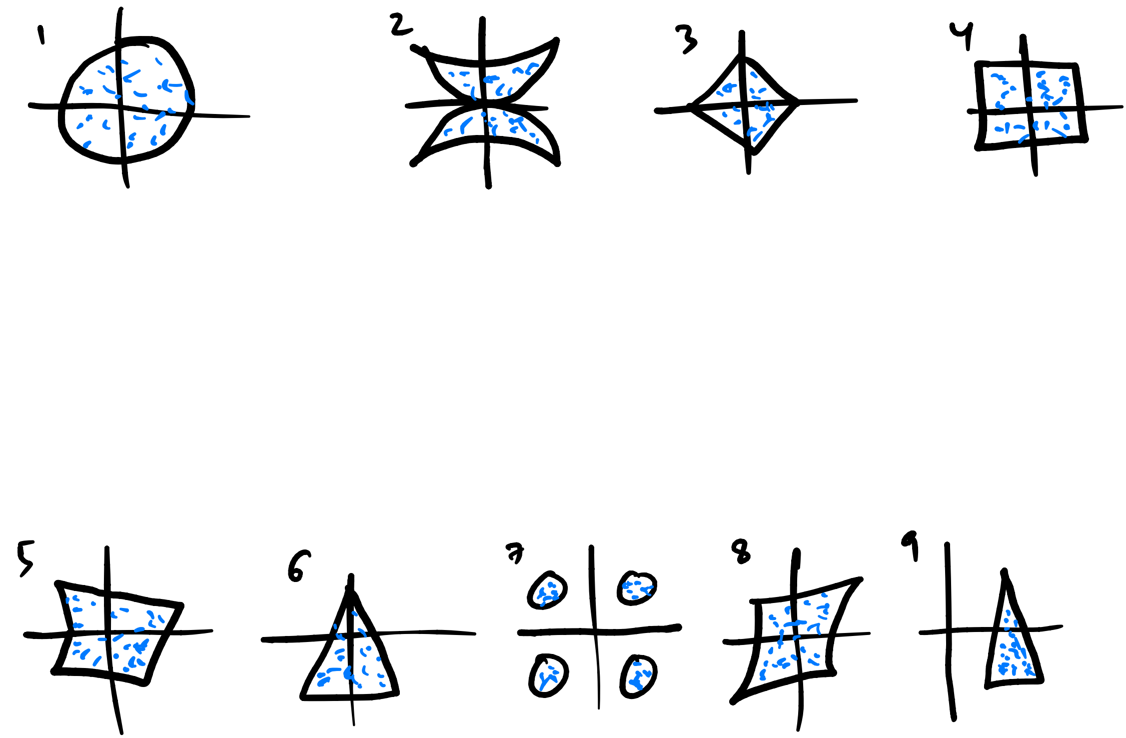


Is the data uncorrelated

■ How about these data. Label them as

- Correlated : C
- Uncorrelated : U

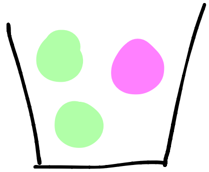
- Independent : I
- Dependent : D



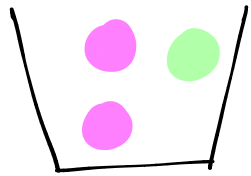
- Application : covariance of received mic. data.

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

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



A



B

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

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

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

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

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

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

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)