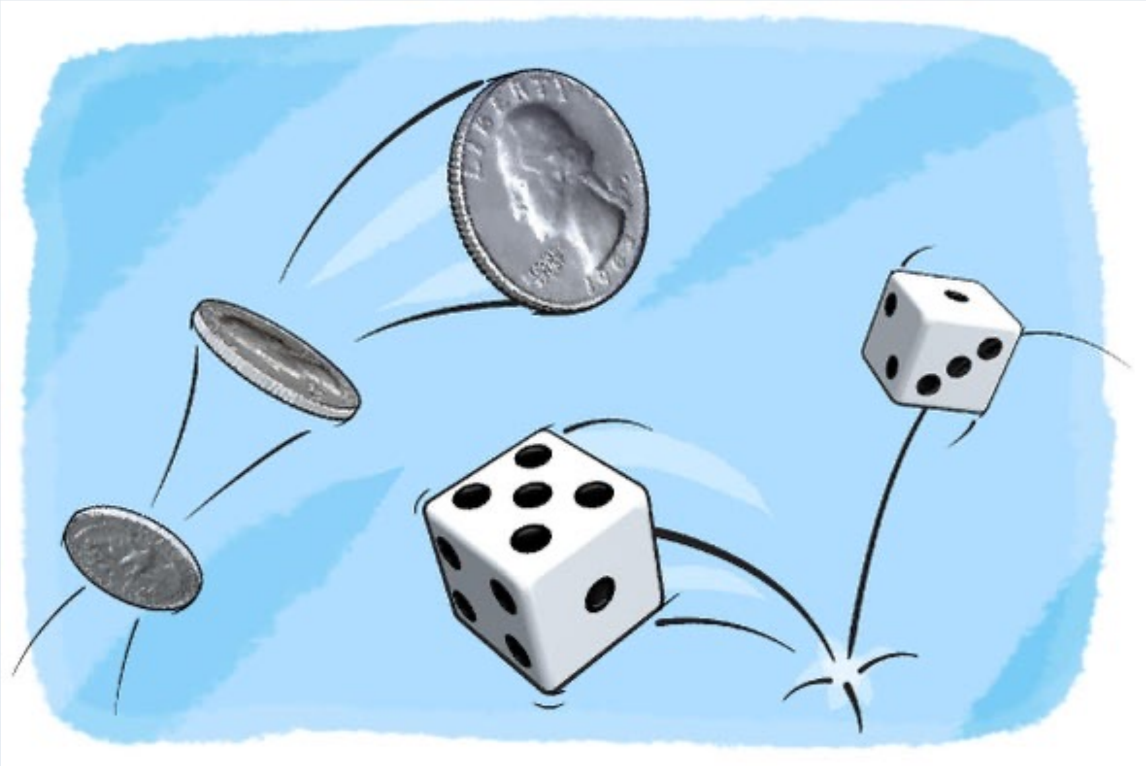
Basic Probability

TUAN NGUYEN

Uncertainty



Variables

Deterministic Variable

Code

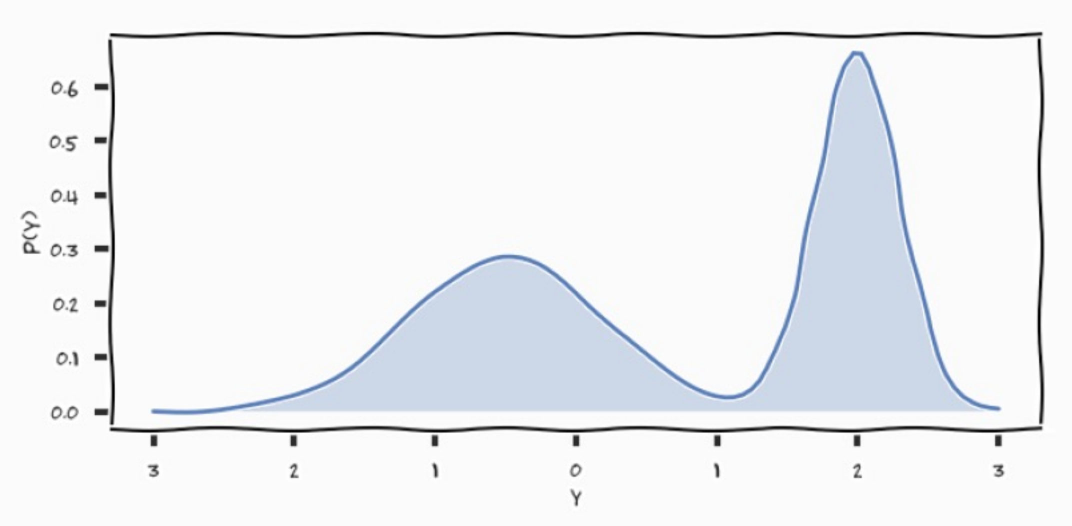
```
int x = 3
float y = 3.14
```

Stochastic Variable

$$x \sim p(x)$$

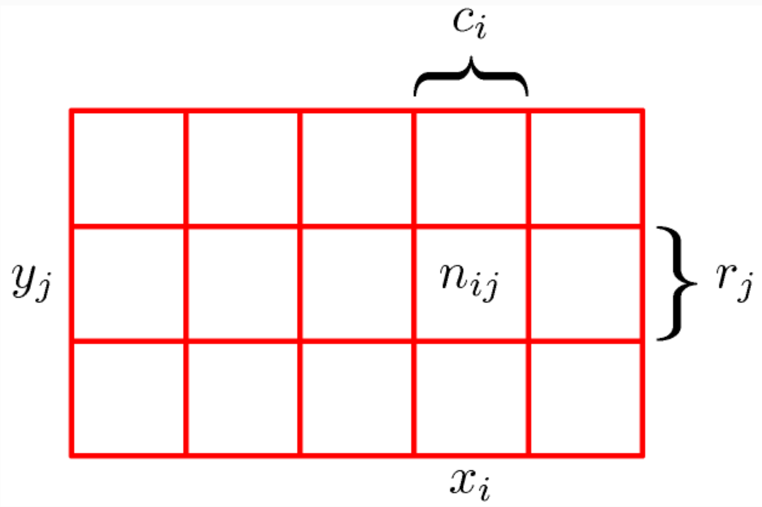
$$y \sim \mathcal{N}(0, I)$$

Probability Theory



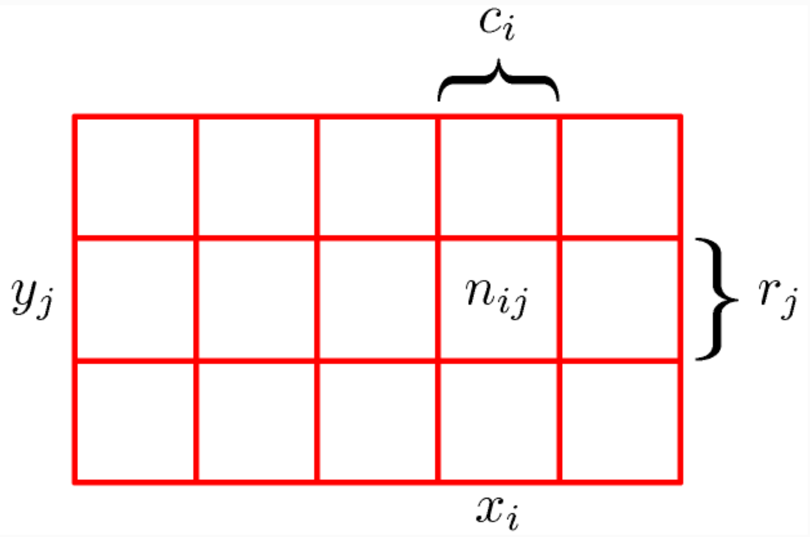
- Probability theory is a framework for manipulating uncertainty
- Random variable, is a stochastic variable that follows a distribution

Probability Theory



$$\{X = x_i, Y = y_j\} = n_{ij}$$

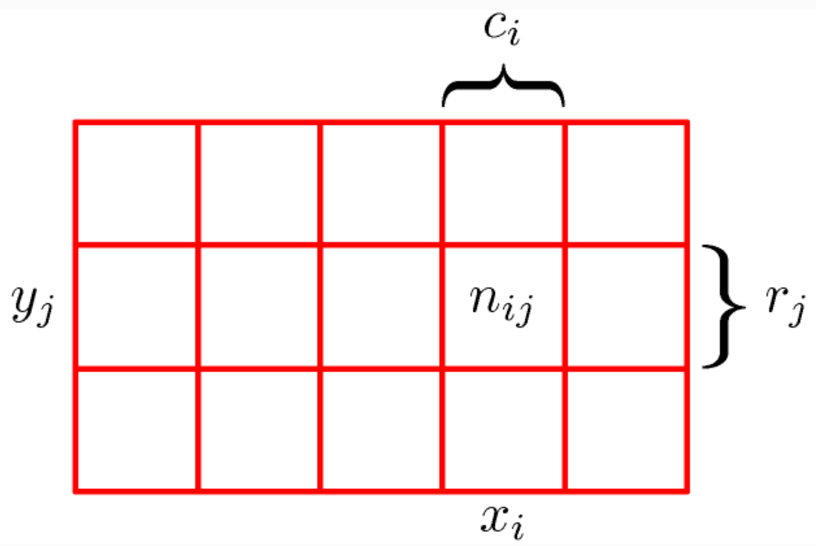
Probability Theory



Joint Probability

$$p(X = x_i, Y = y_j) = \frac{n_{ij}}{\sum_{kl} n_{kl}} = \frac{n_{ij}}{N}$$

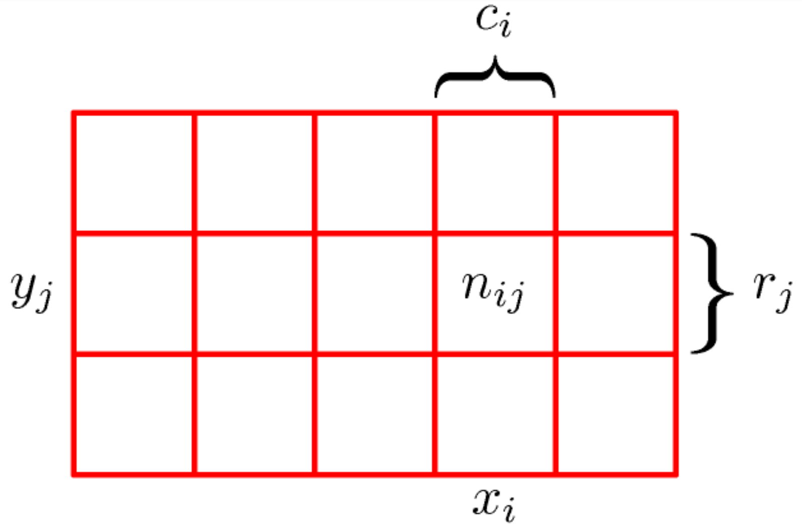
Probability Theory



Marginal Probability

$$p(X = x_i) = \frac{\sum_j n_{ij}}{N} = \frac{c_i}{N}$$

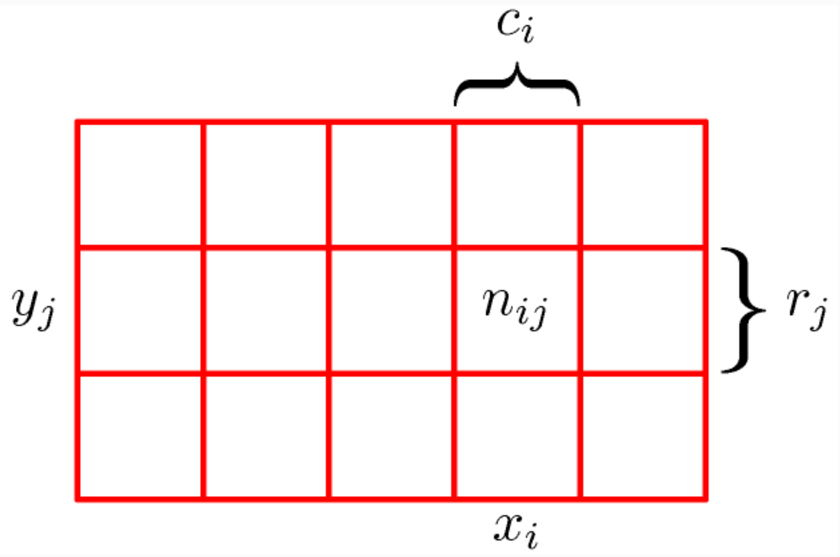
Probability Theory



Sum rule

$$p(X = x_i) = \frac{\sum_j n_{ij}}{N} = \sum_j \frac{n_{ij}}{N}$$

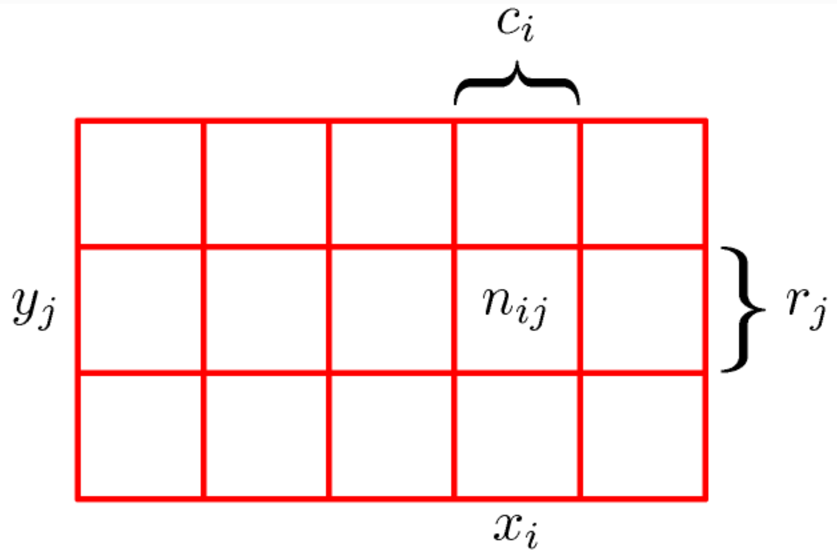
Probability Theory



Conditional

$$p(Y = y_j | X = x_i) = \frac{n_{ij}}{c_i}$$

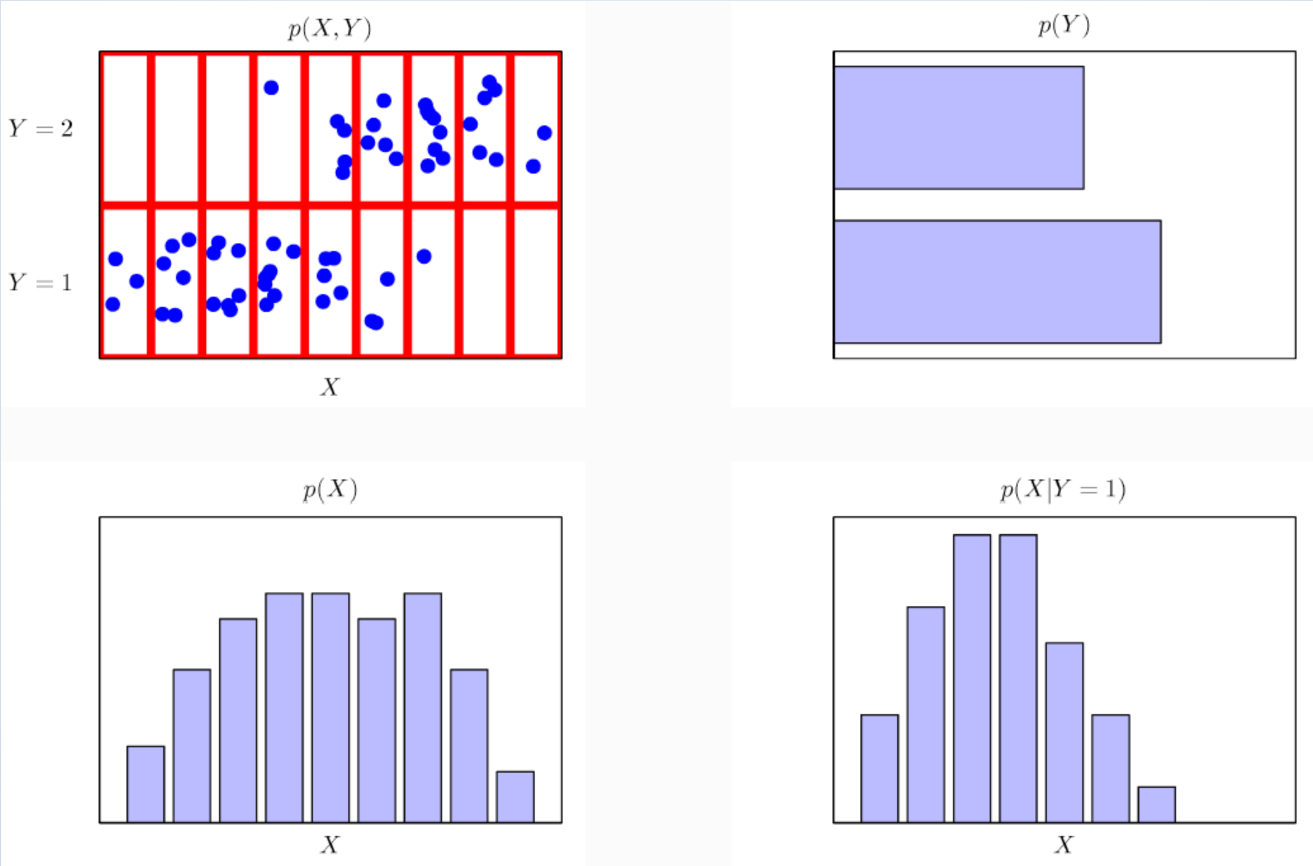
Probability Theory



Product rule

$$p(X = x_i, Y = y_j) = \frac{n_{ij}}{N} = \frac{n_{ij}}{c_i} \cdot \frac{c_i}{N} = p(Y = y_j | X = x_i) p(X = x_i)$$

Probability Theory



Rules of Probability

Sum Rule

$$p(X) = \sum_{Y} p(X, Y)$$

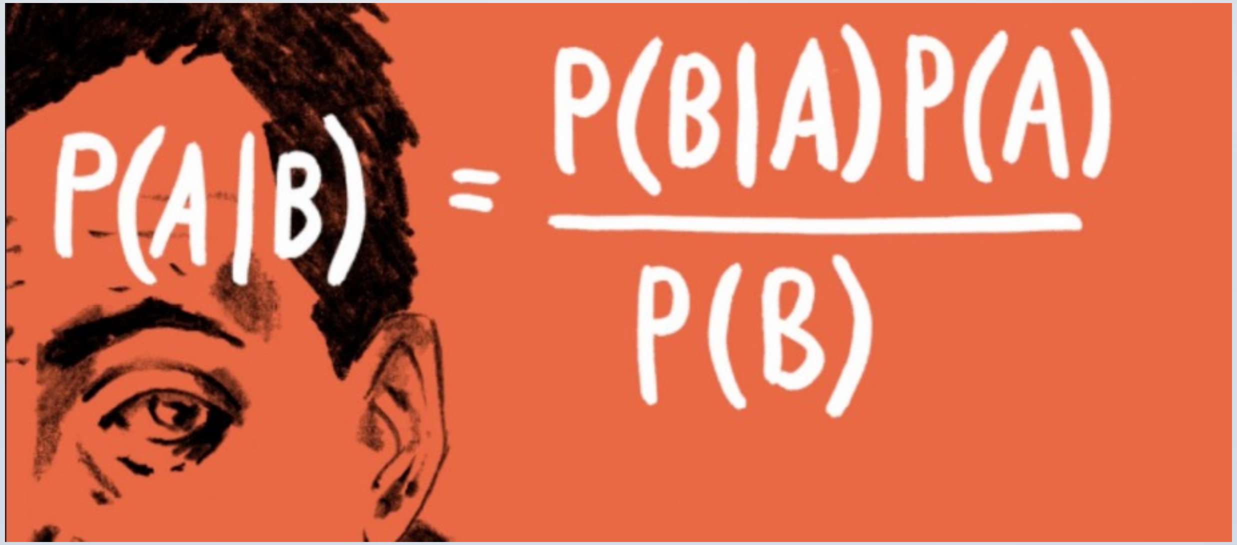
Product Rule

$$p(X, Y) = p(Y|X)p(X)$$

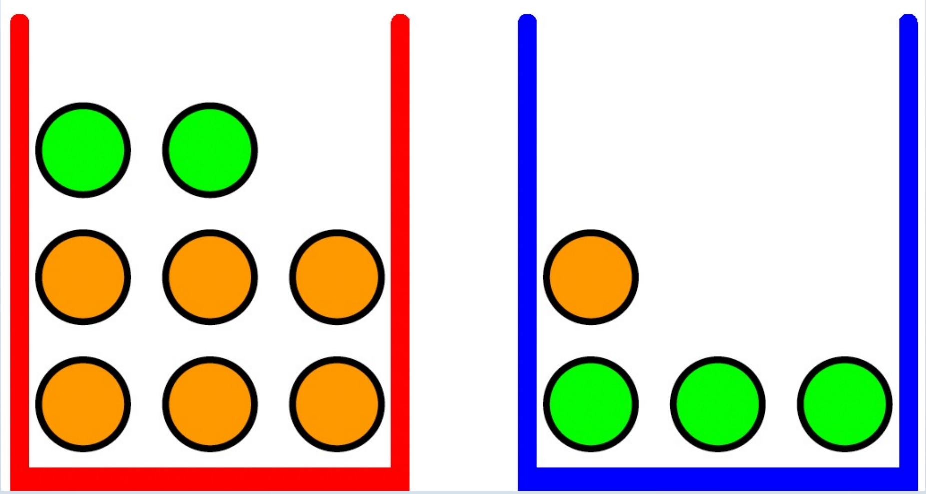
Baye’s Rule

$$
\begin{aligned}
p(X, Y) &= p(Y|X)p(X) \\
p(X, Y) &= p(X|Y)p(Y) \\
p(X|Y)p(Y) &= p(Y|X)p(X) \\
p(X|Y) &= \frac{p(Y|X)p(X)}{p(Y)} \\
&= \frac{p(Y|X)p(X)}{\sum_X p(Y|X)p(X)}
\end{aligned}
$$

Baye’s Rule



Exercise



Random variable B: the box will be chosen, $p(B = r) = 0.4$, $p(B = b) = 0.6$

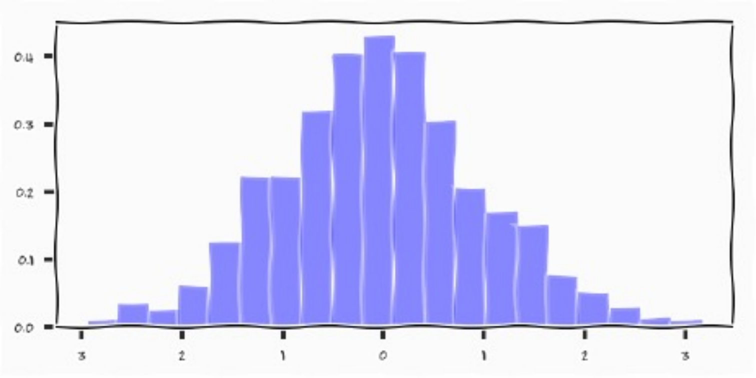
Question:

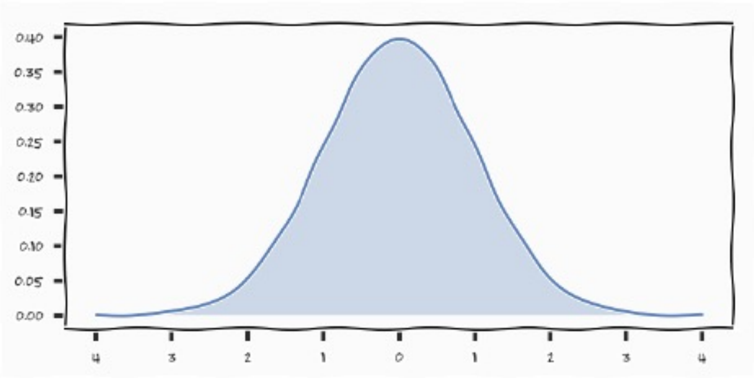
- What is the overall probability that the selection procedure will pick an apple?
- Given that we have chosen an orange, what is the probability that the box we chose was the blue one?

Exercise

A basketball team is to play two games in a tournament. The probability of winning the first game is .10. If the first game is won, the probability of winning the second game is .15. If the first game is lost, the probability of winning the second game is .25. What is the probability the first game was won if the second game is lost?

Discrete vs Continuous



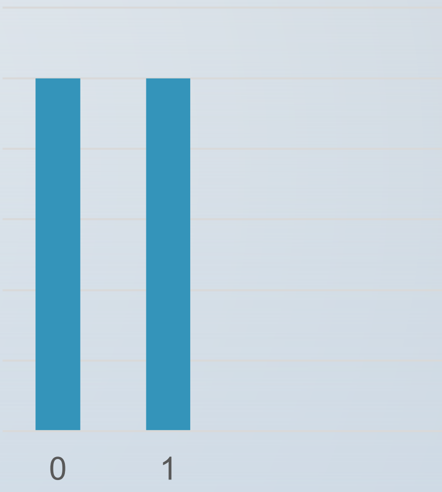


$$\Sigma \qquad\qquad \int$$

Discrete Probability

The probability mass function for X, the number of heads that appear in two tosses of an fair coin

x	P(x)
0	0.5
1	0.5



Bernoulli Distribution

The Bernoulli distribution is a discrete distribution having two possible outcomes labelled by n=0 and n=1 in which

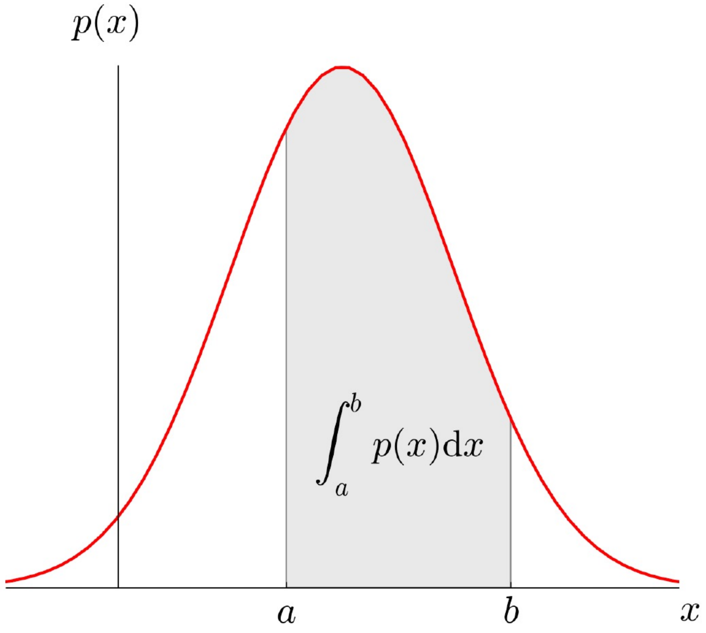
$$P(X = x) = \begin{cases} p, & x = 1 \\ 1 - p, & x = 0 \end{cases}$$

$$P(X = x) = p^x(1 - p)^{1-x}$$

Proof normalization.

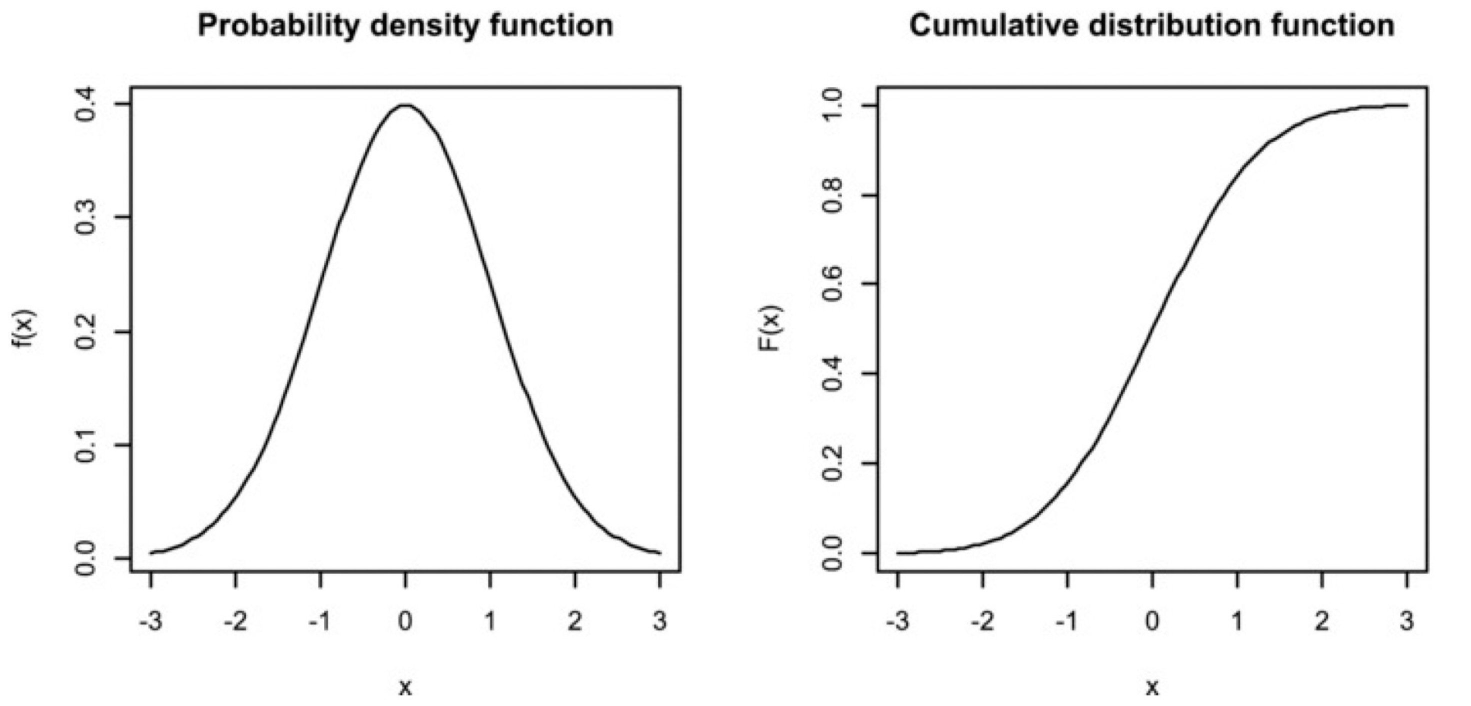
Probability density function (PDF)

$$P(a \leq X \leq b) = \int_a^b f(x)dx \geq 0$$



Cumulative Distribution Function (CDF)

$$F_X(x) = P(X \leq x)$$



Expectation

The expectation

$$\mu = \mathbb{E}[x]$$

Discrete random variable

$$\mathbb{E}[x] = \sum_{x} xP(X = x)$$

Continuous random variable

$$\mathbb{E}[x] = \int_{-\infty}^{+\infty} xf(x)dx$$

Exercise

Repair costs for a particular machine are represented by the following probability distribution. Calculate the expectation of reparing?

x	\$50	\$200	\$350
P(X=x)	0.3	0.2	0.5

Variance

The variance

$$Var(X) = \mathbb{E}[(X - \mathbb{E}[X])^2] = \mathbb{E}[X^2] - \mathbb{E}[X]^2$$

Discrete random variable

$$Var(X) = \sum_x (x - \mu)^2 P(X = x)$$

Continuous random variable

$$Var(X) = \int_{-\infty}^{+\infty} (x - \mu)^2 f(x) dx$$

Exercise

Calculate mean and variance of Bernoulli distribution.