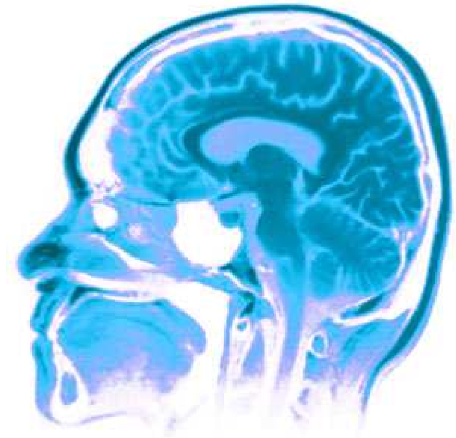
# CPSC340



## Distributions, expectations and Bernoulli random variables

Nando de Freitas
*September, 2012*
*University of British Columbia*

# Outline of the lecture

This lecture is a quick revision of statistics. The goal is for you to learn the following concepts:

- Expectations.
- Distributions.
- Connection between probability and expectation.
- Bernoulli random variables and their properties.

# Discrete random variables.

*A binary random variable (r.v.)* ***X*** *is a mapping from the sample space* ***Ω*** *to a discrete space:* ***E = {0,1}****. In math:*

$$X(\omega):\ \Omega \to E$$

*Where* $\omega$ *are the measurable sets of* ***Ω****.*

***<u>Example</u>***

*For a die, we may be interested in the events* ***E={even,odd}*** *. Here* $\Omega = \{1,2,3,4,5,6\}$ *and* $\boxed{X=\textbf{even}}$ *if* $\omega \in \{2,4,6\}$*, or* $X=\textbf{odd}$ *if* $\omega \in \{1,3,5\}$.

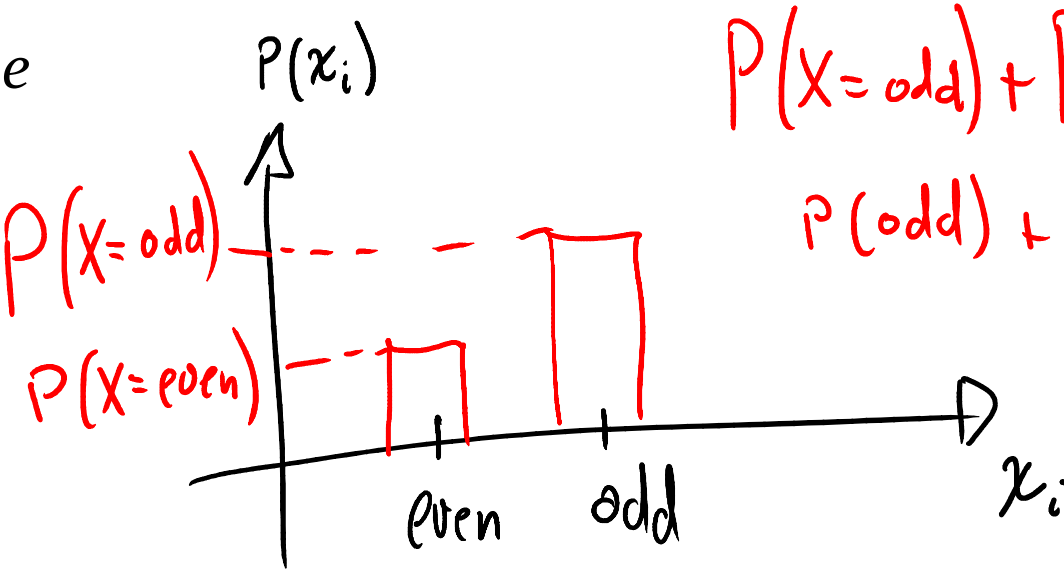
# Probability distributions

*An r.v. may be described with a* ***probability distribution***

$$p(x_i) = P(X = x_i)$$

***Example***

*For our die*

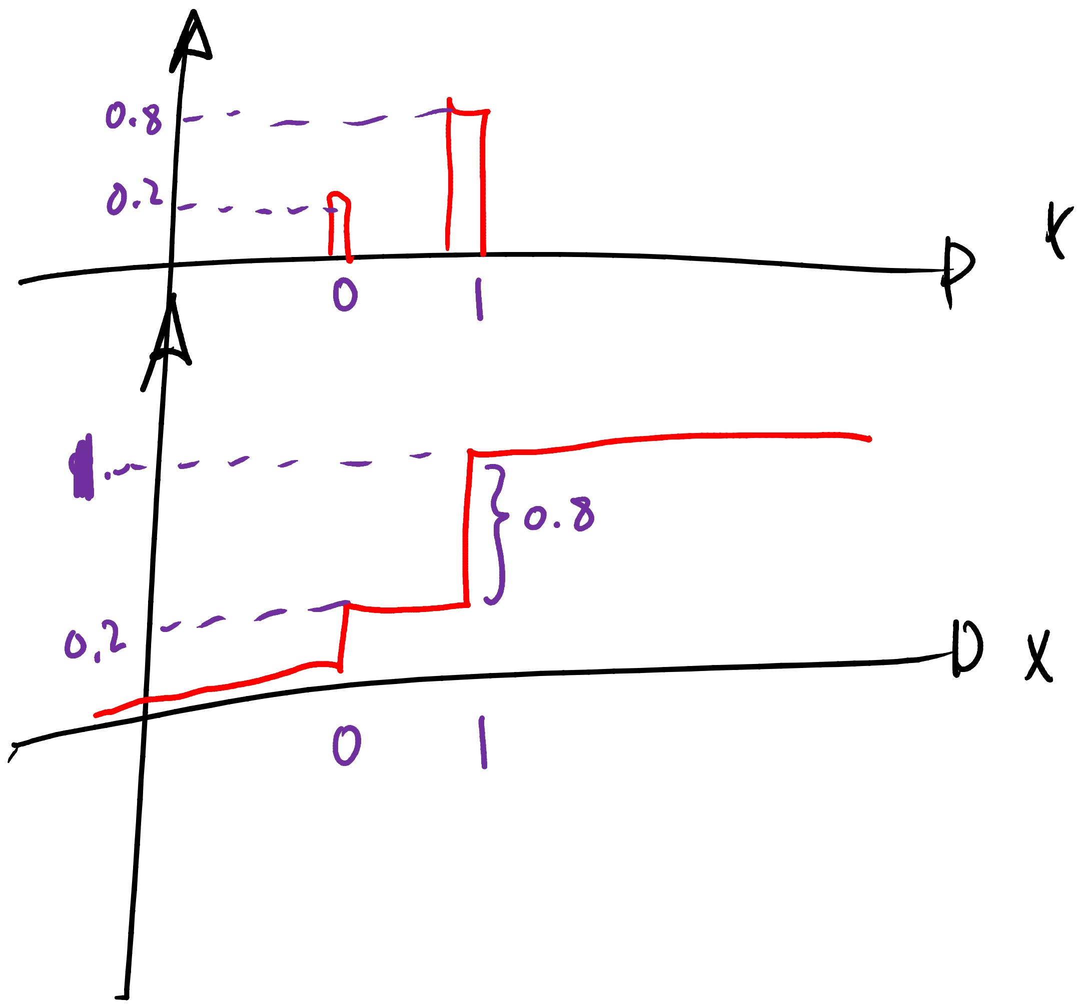



$P(X=\text{odd}) + P(X=\text{even}) = 1$

$p(\text{odd}) + p(\text{even}) = 1$

# The CDF

*The r.v. may also be described with a* ***cumulative distribution function***
***$F(x) = P(X \leq x)$***



# Expectation

$E = \{0, 1\}$

***Expectation*** *is a linear operator defined as follows:*

$$\mathbb{E}(X) = \sum_{x_i \in E} x_i \, p(x_i) = 1 \times P(x_i = 1) + 0\, P(x_i = 0)$$
$$= P(x_i = 1)$$

*Let us introduce the* ***indicator variable*** $\boxed{X = \mathbb{I}_A(w)}$

For a die

$\mathbb{I}_{Even}(w=5) = 0$

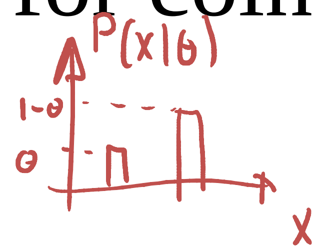
$\mathbb{I}_{Even}(w=4) = 1$

$$\boxed{X = \mathbb{I}_A(w)} = \begin{cases} \boxed{1} & \text{if } \boxed{w \in A} \\ \boxed{0} & \text{otherwise} \end{cases}$$

*Question: What is* $\mathbb{E}(\mathbb{I}_A(w))$ *equal to?*

$$\mathbb{E}\left[\mathbb{I}_A(w)\right] = \underbrace{1\, P(w \in A)}_{w \in A} + \underbrace{0 \cdot P(w \notin A)}_{w \notin A} = P(w \in A)$$
$$= P(A)$$

# Bernoulli: a model for coins




*A* ***Bernoulli r.v.*** ***X*** *takes values in {0,1}*

$$p(x|\theta) = \begin{cases} \theta & \text{if } x=1 \\ 1-\theta & \text{if } x=0 \end{cases}$$

$P(X=1) = \theta$

$P(X=0) = 1-\theta$

*Where* $\theta \in (0,1)$. *We can write this probability more succinctly as follows:*

$$\boxed{P(x|\theta) = \theta^{x}(1-\theta)^{1-x}} = \begin{cases} \theta & \text{if } x=1 \\ \theta^{0}(1-\theta)^{1} = 1-\theta & \text{if } x=0 \end{cases}$$

$$\boxed{P(x|\theta) = \theta^{\mathbb{I}_1(x)}(1-\theta)^{\mathbb{I}_0(x)}} = \begin{cases} \theta & \text{if } x=1 \\ 1-\theta & \text{if } x=0 \end{cases}$$

# Bernoulli expectation

*What is the expectation/mean of a Bernoulli variable?*

$$P(x|\theta) = \theta^x (1-\theta)^{1-x}$$

$$\mathbb{E}(x) = \sum_{x \in \{0,1\}} x\, p(x|\theta) = 1\, P(x=1|\theta) + 0\, P(x=0|\theta)$$

$$\mathbb{E}(x|\theta)$$

$$\mathbb{E}_\theta(x)$$

$$= 1\, P(x=1|\theta)$$

$$= \theta$$

# Bernoulli expectation

*What is the variance of a Bernoulli variable?*

$\mathrm{var}(x|\theta)$

$$= \mathbb{E}\left(\left(X - \mathbb{E}(x|\theta)\right)^2 \,\middle|\, \theta\right) = \mathbb{E}\left((X-\theta)^2 \,\middle|\, \theta\right)$$

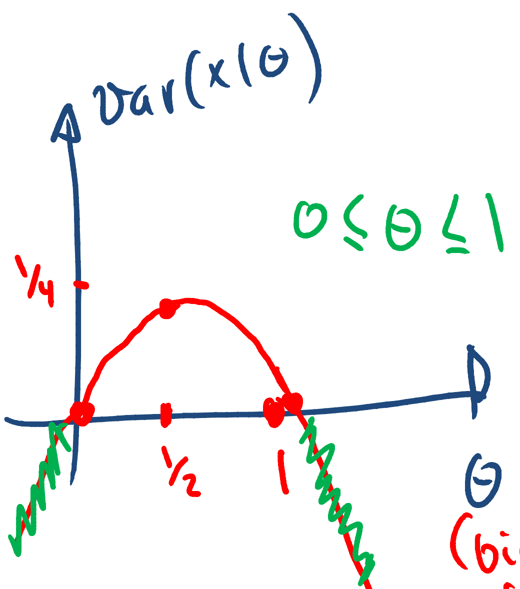
$$\boxed{\frac{d(\theta-\theta^2)}{d\theta} = 1 - 2\theta = 0 \qquad \theta = 1/2}$$

$$= (1-\theta)^2 P(x=1|\theta) + (0-\theta)^2 P(x=0|\theta)$$

$$= (1-\theta)^2\theta + \theta^2(1-\theta)$$

$$= \theta(1-\theta) = \theta - \theta^2$$

$$\mathrm{var}(x|\theta=1/2) = \frac{1}{2} - \frac{1}{4} = \frac{1}{4}$$




(bias of the coin)

# N independent tosses

*What is the distribution of N independent coin tosses?*

$$P(x_1, x_2 | \theta) = P(x_1 | \theta)\, P(x_2 | \theta)$$

$$P(x_{1:n} | \theta) = P(x_1\, x_2 \cdots x_n | \theta) = \underbrace{\prod_{i=1}^{n} P(x_i | \theta)}_{}$$

# Next lecture

In the next lecture, we introduce information theory and the principle of learning by maximizing likelihood.