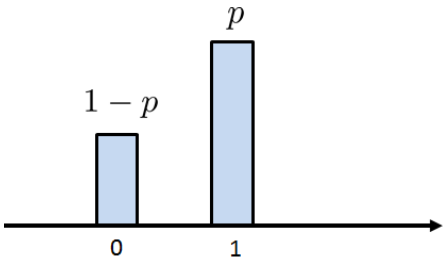
ECE 302: Lecture 3.6 Bernoulli Random Variables

Prof Stanley Chan

School of Electrical and Computer Engineering
Purdue University

The random variable of flipping a coin



- How can this ever be interesting?
- It can be used to model randomized algorithm.
- It can be used to model social networks.

Outline

- 3.1 Random variables
- 3.2 Probability mass functions (PMF)
- 3.3 Cumulative distribution functions (discrete case)
- 3.4 Expectation
- 3.5 Moments and variance
- 3.6 Bernoulli random variables
  - Definition of a Bernoulli random variable
  - Expectation, Variance
  - Application 1: Modeling randomized algorithms
  - Application 2: Modeling social networks
- 3.7 Binomial random variables
- 3.8 Geometric random variables
- 3.9 Poisson random variables

Bernoulli Random Variable

Definition

Let X be a **Bernoulli** random variable. Then, the PMF of X is

$$p_X(0) = 1 - p, \qquad p_X(1) = p,$$

where $0 < p < 1$ is called the Bernoulli parameter. We write

$$X \sim \mathrm{Bernoulli}(p)$$

to say that X is drawn from a Bernoulli distribution with a parameter p.

Example. Coin flip.

Mean, Variance

Proposition

If $X \sim \text{Bernoulli}(p)$, *then*

$$\mathbb{E}[X] = p, \qquad \mathbb{E}[X^2] = p, \qquad \text{Var}[X] = p(1-p).$$

Analysis of the variance

When will the variance of a Bernoulli random variable be maximized?

- When $p = 0$?
- When $p = 1$?
- When $p = \frac{1}{2}$?

Rademacher random variable

- Bernoulli: 0 and 1
- Rademacher: $+1$ and -1

PMF of Rademacher is

$$p_X(-1) = \frac{1}{2}, \quad \text{and} \quad p_X(+1) = \frac{1}{2}.$$

You can show that following result:

Proposition

If X is a Rademacher random variable, then $(X+1)/2 \sim \text{Bernoulli}(1/2)$. Conversely if $Y \sim \text{Bernoulli}(1/2)$ then $2Y - 1$ is Rademacher.

Application 1: Modeling randomized algorithms

Consider a large linear system:

$$\begin{bmatrix} y_1 \\ \vdots \\ y_N \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1N} \\ \vdots & \vdots & \ddots & \vdots \\ a_{M1} & a_{M2} & \dots & a_{MN} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_N \end{bmatrix}$$

where M and N are very large.

Example.

- Large network analysis (million-node network)
- Large-scale inverse problem (giga-pixel deconvolution)
- Large-scale routing (air traffic control)
- Large-scale decomposition (genome analysis)

Application 1: Modeling randomized algorithms

Brute-force computation:

$$y_i = \sum_{j=1}^{N} a_{ij} x_j.$$

Randomized computation: Let $I_j \sim \text{Bernoulli}(p_j)$. Then,

$$\widehat{y}_i = \sum_{j=1}^{N} a_{ij} x_j I_j / p_j,$$

- If $I_j = 0$, then you can skip that entry.

Application 1: Modeling randomized algorithms

- The compensation by p_j is to ensure $\mathbb{E}[\widehat{y}_i] = y_i$. How?

We can prove that with extremely high probability, the deviation between $\widehat{y}_i$ and y_i is very small.

Drineas, Kannan, Mahoney, Fast Monte Carlo algorithms for matrices, SIAM Journal on Computing 36(1), pp.132-157, 2006.

Application 2: Modeling social networks

Erdos-Renyi Graph

- A very famous (and simple) model for large networks.
- Is used nowadays to study network structures, e.g. Facebook, Google.

Graph **Matrix**

Bollobas, Random Graphs (2nd ed.). Cambridge University Press. 2001.

Application 2: Modeling social networks

Erdos-Renyi Graph says:

$$A_{ij} \sim \text{Bernoulli}(p), \tag{1}$$

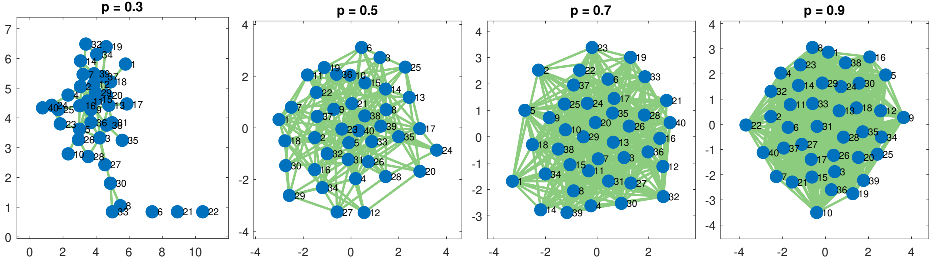


Figure: Erdős-Rényi graph with different parameters $p = 0.3, 0.5, 0.7, 0.9$.

Questions?