

Probability & Random Processes

Irvin Avalos

irvin.l@berkeley.edu

CONTENTS

1	Probability Theory	2
1.1	Introduction	2
1.2	Discrete Random Variables	2
1.3	Continuous Random Variables	2
1.4	Order Statistics	2
1.5	Transforms and Sums	2
1.6	Concentration Inequalities and Convergences	2
2	Information Theory	3
2.1	Information Measures and Entropy	3
2.2	Source Coding Theorem	3
2.3	Huffman Coding	3
2.4	Channel Models	3
2.5	Channel Coding Theorem	3
3	Stochastic Processes	4
3.1	Discrete Time Markov Chains	4
3.2	Poisson Processes	4
3.3	Continuous Time Markov Chains	4
4	Random Graphs	5
4.1	Erdos-Reyni Model	5
4.2	Thresholds and Connectivity	6
5	Statistical Inference	7
5.1	Point Estimation	7

CHAPTER 1

PROBABILITY THEORY

1.1 INTRODUCTION

1.2 DISCRETE RANDOM VARIABLES

1.3 CONTINUOUS RANDOM VARIABLES

1.4 ORDER STATISTICS

1.5 TRANSFORMS AND SUMS

1.6 CONCENTRATION INEQUALITIES AND CONVERGENCES

CHAPTER 2

INFORMATION THEORY

2.1 INFORMATION MEASURES AND ENTROPY

2.2 SOURCE CODING THEOREM

2.3 HUFFMAN CODING

2.4 CHANNEL MODELS

2.5 CHANNEL CODING THEOREM

CHAPTER 3

STOCHASTIC PROCESSES

3.1 DISCRETE TIME MARKOV CHAINS

3.2 POISSON PROCESSES

3.3 CONTINUOUS TIME MARKOV CHAINS

CHAPTER 4

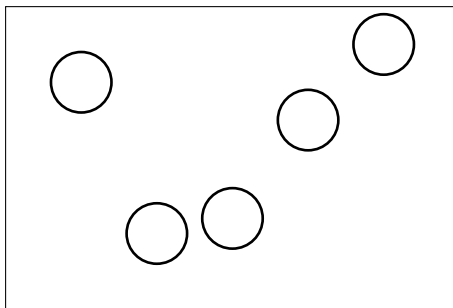
RANDOM GRAPHS

4.1 ERDOS-REYNI MODEL

Definition 4.1 (Erdos-Reyni random graph).

Given $n \in \mathbb{Z}^+$ and a probability $p \in [0, 1]$, the (E-R) random graph is denoted as $\mathcal{G}(n, p)$.

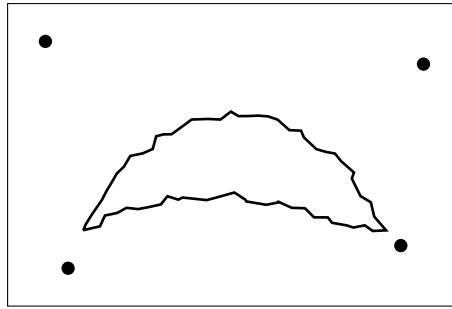
- The (RG) $\mathcal{G}(n, p)$ is undirected on all n vertices
- Each of the $\binom{n}{2}$ edges is formed independently from each other with probability p
- E-R stated a number of result that are based on "thresholds" of p which are needed to show the structural properties of the graph:
 1. p is a function of n , $p := p(n)$
 2. $\frac{1}{n^2}$ is the threshold for the first edges to appear in a (RG)
 3. $\frac{1}{n^{3/2}}$ is the threshold for the first trees to emerge
 4. $\frac{1}{n}$ is the threshold for the first cycles to appear in the graph
- If $p_n = 1/n$ then the "Giant component" emerges. Let $p_n = (1 - \varepsilon)/n$ for $0 < \varepsilon \ll 1$, then each "puddle" represents a region in the graph where all nodes are connected.



Subcritical: $p = \frac{1-\varepsilon}{n}$

Note that p_n is to the left of ε and that the size of the largest puddles are "small", $\rightarrow O(\log n)$ w.h.p.

- For $p_n = (1 + \varepsilon)/n$, we go to the right of $\varepsilon \rightarrow 1 + \varepsilon$.

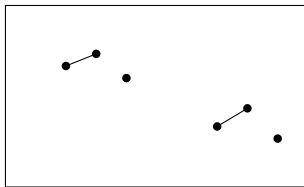


Supercritical: $p > \frac{1+\varepsilon}{n}$

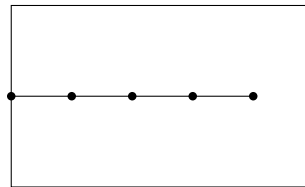
Here we have one large "island" where its size is relatively, $O(n)$ w.h.p.

4.2 THRESHOLDS AND CONNECTIVITY

When p_n is below $\log_e n \implies$ no connectivity, $P(\text{connectivity}) \longrightarrow 0$. Otherwise, if p_n is above $\log_e n \implies$ ensures connectivity, $P(\text{connectivity}) \longrightarrow 1$.



Disconnected: $p < \frac{\log_e n}{n}$



connected ($p > (1 + \varepsilon) \frac{\log n}{n}$)

CHAPTER 5

STATISTICAL INFERENCE

5.1 POINT ESTIMATION

- **Probability vs Statistics:** There exists two perspectives that surround the field of statistical inference, Bayesian vs Frequentist.

1. From the Bayesian perspective, unknown quantities are treated as RVs with known or assumed distributions and priors.
2. While under the Frequentist's point of view, unknowns are deterministic parameters that need to be estimated.

In essence, a person following the Bayesian perspective starts with a probabilistic model and assumes its distribution and shows through rigorous probabilistic analysis that measurements get smaller.

On the other hand, a Frequentist would make hidden assumptions about parameters present in the model for which we know nothing about.

- **Bayesian Inference:**