UNIVERSITÄT DES SAARLANDES
FACHRICHTUNG MATHEMATIK
Dr. Marwa Banna

High Dimensional Probability
with Applications to Big Data Sciences

Summer term 2020

Assignment 2
Due: Monday, May 25, 2020

Problem 6 (Tails of Normal Distribution)**.** Let $g \sim \mathcal{N}(0,1)$. Prove that for all $t > 0$,

$$\left(\frac{1}{t} - \frac{1}{t^3}\right)\frac{1}{\sqrt{2\pi}}e^{-t^2/2} \leq \mathbb{P}(g \geq t) \leq \frac{1}{t}\frac{1}{\sqrt{2\pi}}e^{-t^2/2}.$$

Problem 7 (Exercises from lecture)**.**

1. Proposition 2.1: Prove that property $P5$ is equivalent to $P1$ - $P4$.
   Hint: Prove that $P3 \Rightarrow P5$ and $P5 \Rightarrow P1$.
2. Prove the general Hoeffding inequality in Theorem 2.3.
3. Prove Khintchine's inequality in Theorem 2.5.

Problem 8 (Chernoff's inequality)**.** Let X_i be independent Bernoulli random variables with parameters p_i, i.e. $\mathbb{P}(X_i = 1) = 1 - \mathbb{P}(X_i = 0) = p_i$. Set $S_N = \sum_{i=1}^{N} X_i$ and denote by $\mu = \mathbb{E}S_N$.

1. (Large deviations) Prove that for any $t > \mu$,
$$\mathbb{P}(S_N \geq t) \leq e^{-\mu}\left(\frac{e\mu}{t}\right)^t$$
   and for any $t < \mu$,
$$\mathbb{P}(S_N \leq t) \leq e^{-\mu}\left(\frac{e\mu}{t}\right)^t.$$
2. (Small deviations) Deduce that for $\delta \in (0,1)$
$$\mathbb{P}\big(|S_N - \mu| \geq \delta\mu\big) \leq 2e^{-c\mu\delta^2},$$
   where $c > 0$ is an absolute constant.
   Hint: Apply part 1 for $t = (1 \pm \delta)\mu$ and analyze the bounds for small δ.

Application to degree of random graphs

We give an application of Chernoff's inequality to random graphs. We consider the classical Erdös-Rényi model $G(n, p)$, which is constructed on a set of n vertices by connecting every pair of *distinct* vertices independently with probability p. In applications, this model appears as the simplest stochastic model for large, real-world, networks.

Problem 9. Consider an Erdös-Rényi random graph $G(n, p)$. For any $i = 1, \dots, n$, let X_i be the *degree* of the vertex v_i in the graph, i.e. the number of the edges incident to v_i.

1. Prove that $\mathbb{E}X_i = (n-1)p =: d$.

We say that a graph is regular if all vertices have equal degrees. We will show that relatively *dense graphs*, those where $d \gtrsim \log n$, are almost regular with high probability; i.e. the degrees of all vertices are approximately equal to d with high probability.

2. Assume that $d \geq C \log n$. Prove that if C is sufficiently large then with high probability (say, 0.9), all the vertices have degree between $0.9d$ and $1.1d$.
   Hint: Use Chernoff's inequality.

We consider now sparse graphs where $d = \mathcal{O}(\log n)$.

3. Prove that with high probability (say, 0.9), all the vertices have degree of order $\log n$.
   Hint: Use Chernoff's inequality.

We will see that the situation changes for very sparse graphs where $d = \mathcal{O}(1)$.

4. Prove that with high probability (say, 0.9), all the vertices have degree

$$\mathcal{O}\left(\frac{\log n}{\log(\log n)}\right).$$