Probabilities and statistics

Lecturer: Prof. Dr. Sara van de Geer
Prof. Dr. Martin Larsson

Serie 2

March 2nd, 2014

Q1. Let $G = (V, K)$ be an arbitrary finite and undirected graph with vertices V and edges K, i.e., V is a finite set and $K \subseteq \{\{x, y\} \in V : x \neq y\}$. The MAX-CUT problem is to find a subset $A \subseteq V$ such that the number of edges connecting A and A^c (K_A) is as large as possible, i.e., $K_A = \{\{x, y\} \in K : x \in A, y \in A^c\}$. We want to show that there exists $A \subseteq V$ so that $|K_A| \geq \frac{1}{2}|K|$.

(a) Choose $A \subseteq V$ to be random set uniformly in 2^V. Calculate $\mathbb{P}(e \in K_A)$, i.e. $\mathbb{P}(\{A : e \in K_A\})$

(b) Using the linearity of the expectation show that

$$\mathbb{E}[|K_A|] = \frac{1}{2}|K|.$$

(c) Show that there exists an A so that $|K_A| \geq \frac{1}{2}|K|$.

Q2. **(a)** Take $p \in [0, 1]$ and $n \in \mathbb{N}\backslash\{0\}$. We say that $X \sim Bin(n, p)$ if the distribution of X is

$$\mathbb{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k}, \qquad k \in \{0, 1, ..., n\}.$$

Show that this is indeed a probability distribution using 2 different methods:

i. Calculating $\sum_k \mathbb{P}(X = k)$.

ii. Representing this probability in terms of the box model with replacement.

Calculate the expected value of X using 2 different methods (the one listed above).

(b) Take $K, n \in \mathbb{N}$ and $N \in \mathbb{N}\backslash\{0\}$ with $K, n \leq N$. We say that a random variable $X \sim Hyp(N, k, n)$ if its distribution is given by

$$\mathbb{P}(X = k) = \frac{\binom{K}{k}\binom{N-K}{n-k}}{\binom{N}{n}} \qquad k \in \{\max\{0, n + K - N\}, .., \min\{n, K\}\}$$

Show that this is indeed a probability distribution using 2 different methods:

i. Calculating $\sum_k \mathbb{P}(X = k)$.
Hint: Calculate $(1 + x)^n$ in two different ways and identify the terms.

ii. Representing this probability in the box model without replacement.

Calculate the expectation using both methods.

Q3. The voting problem Assume you have n votes in an election with two candidate (all people vote for one and only one of them) and the winning candidate have k more votes than the loser. If the votes were counted in a random way (the uniform measure in all possible ways of ordering the votes). What is the probability that there was never a moment, except the beginning, where the loser candidate has the same number or more number of votes than

the winning one.
Hint: Define $(S_l)_{0\leq n\leq N} := \sum_{i=1}^{l} X_i$ where

$$X_i := \begin{cases} 1 & \text{the vote was for the winner,} \\ -1 & \text{the vote was for the loser.} \end{cases}$$

Note that the event we are looking for is $A := \bigcap_{l=1}^{n}\{w \in \Omega : S_l(\omega) > 0\}$, calculate $|A|$ and $|\Omega|$.

Have a nice week ☾☼☽!!.