Combinatorial Hypothesis Testing

April 13, 2023

0 Contents ❖

1 Introduction ❖

Suppose we observe an n-dimensional vector $\mathbf{X} = (X_1, ..., X_n)$. The null hypothesis H_0 is that the components of $\mathbf{X}$ are independent and identically distributed (i.i.d.) standard normal random variables. We denote the probability measure and expectation under H_0 by $\mathbb{P}_0$ and $\mathbb{E}_0$, respectively.

Combinatorics kicks in as we consider the alternative hypotheses: consider a class $\mathcal{C} = \{S_1, \ldots, S_N\}$ of N sets of indices such that $S_k \subset \{1, \ldots, n\}$ for all $k = 1, \ldots, N$. Under H_1, there exists an $S \in \mathcal{C}$ such that X_i has a distribution determined by whether i is in S:

1. In its simplest form, as discussed in [1–3], we consider

$$X_i \text{ has distribution } \begin{cases} \mathcal{N}(0,1), & \text{if } i \notin S \\ \mathcal{N}(\mu,1), & \text{if } i \in S \end{cases}$$

   where $\mu > 0$ is a positive parameter and components of $\mathbf{X}$ are independent.

2.

2 Moment Methods ❖

3 Extension ❖

3 References ❖

[1] Louigi Addario-Berry, Nicolas Broutin, Luc Devroye, and Gábor Lugosi. On combinatorial testing problems. 2010.

Last updated April 13, 2023

[2] Ery Arias-Castro, Emmanuel J Candes, and Arnaud Durand. Detection of an anomalous cluster in a network. *The Annals of Statistics*, pages 278–304, 2011.

[3] Ery Arias-Castro, Emmanuel J Candes, Hannes Helgason, and Ofer Zeitouni. Searching for a trail of evidence in a maze. 2008.