INSE6300	Jia Yuan Yu
6: Hypothesis Testing	
Concordia	March 29, 2016

We have seen different estimators for the unknown parameter of an unknown probability distribution that belongs to a known set.

In hypothesis testing, we have observations of a random variable, whose distrubition is unknown, up to a set of distributions $\mathcal{P} = \{P_\theta : \theta \in \Omega\}$. A hypothesis is a subset H of $\mathcal{P}$. There are two decisions: accept or reject the hypothesis.

First, we review how the Automobile game models supply chain management.

1 Example

Suppose that the observation X takes values in a finite set M. Consider $\mathcal{P} = \{P_0, P_1\}$, and the hypothesis $H = \{P_0\}$. The complement of the hypothesis is called the alternative and denotes $K = H^c$.

Suppose that each observation X take values in $\mathbb{R}$. Let d_0 and d_1 denote accepting or rejecting the hypothesis H respectively. A nonrandomized decision rule is a mapping $\delta : M \to \{d_0, d_1\}$. As in the study of control charts, there are two types of errors associated with the two decisions and two hypotheses. One objective of quality assurance or performance guarantees is to find decision rules that tradeoff the two types of errors.

For a given $\alpha \geq 0$, the objective is to minimize

$$\mathbb{P}(\delta(X) = d_0) \quad \text{for all } \mathbb{P} \in H^c,$$

subject to

$$\mathbb{P}(\delta(X) = d_1) \leq \alpha \quad \text{for all } \mathbb{P} \in H.$$

In other words, we want to minimize false-alarms, subject to a constraint on missed detections.

Observe that a nonrandomized decision rule δ is a subset S_δ of M. Since H^c contains a single element, we can find the optimal decision rule by finding the subset A:

$$\begin{aligned} \min_A \quad & \sum_{x \in A} P_0(X = x) \\ \text{subject to} \quad & \sum_{x \in A^c} P_1(X = x) \leq \alpha. \end{aligned}$$

Observe that minimizing $\mathbb{P}(\delta(X) = d_0)$ for $\mathbb{P} \in H^c$ is equivalent to maximizing $\mathbb{P}(\delta(X) = d_1)$ for $\mathbb{P} \in H^c$. Hence, the optimal decision rule can be found by solving:

$$\begin{aligned} \max_S \quad & \sum_{x \in S} P_1(X = x) \\ \text{subject to} \quad & \sum_{x \in S} P_0(X = x) \leq \alpha. \end{aligned}$$

The set S corresponds to rejecting the hypothesis.

One method to solve the above optimization is to rank all $x \in M$ according to

$$\frac{P_1(x)}{P_0(x)},$$

and adding elements to S until the threshold α is reached.

Remark 1 (Randomized rules). Mapping $x \in M$ to a probability distribution $\phi(x)$, then flip a coin with probability $\phi(x)$ to determine accept or reject.

2 References

- Lehmann and Romano's Testing Statistical Hypotheses.
- Robert W. Keener's "Theoretical Statistics: Topics for a Core Course."