Notes on The Nonexistence of Certain Statistical Procedures in Nonparametric problems

Catherine Chen cyc2152

December 2023

1 Theorems

Let $\mathcal{F}$ be a given set of distribution functions $F, G, \ldots$ of a real variable, it must satisfy that:

1. For every $F \in \mathcal{F}$, $\mu_F = \int_{-\infty}^{\infty} z dF$ exists and is finite.

2. For every real m, there is an $F \in \mathcal{F}$ with $\mu_F = m$.

3. $\mathcal{F}$ is **convex**, that is, if F, $G \in \mathcal{F}$, π is a positive fraction, and $H = \pi F(1-\pi)G$ then $H \in \mathcal{F}$.

Let $X_1, X_2, \cdots \sim F$ denote an infinite sequence of independent RVs, i.e. $\Pr(X_i \leqq z) = F(z)$. Suppose that a (randomized, sequential) sampling procedure is given, i.e. a set of rules for observing $X_1, X_2, \cdots$ one by one up to a certain stage N s.t. at each stage the decision whether to continue depends (randomly) on the observed values in hand at that stage. The given procedure is assumed to be closed:

$$P_F(N < \infty) = 1, \tag{1}$$

for each $F \in \mathcal{F}$.

Denote the total outcome of the sampling procedure a random variable, V, i.e. $V = (X_1, X_2, \cdots, X_N)$. As in (1), for any event A defined on the sample space of V, $P_F(A)$ will denote the probability of A when F *obtains*, i.e., when each X_i is distributed according to F.

If φ is a real valued function of V, $E_F[\varphi]$ denotes the expected value of φ (if it exists) when F *obtains*.

For any real number m, let $\mathcal{F}_m$ denote the set of all $F \in \mathcal{F}$ with $\mu_F = m$.

1.1 Theorem I

For each bounded real valued function φ on the sample space of V, $\inf_{F \in \mathcal{F}_m} E_F[\varphi]$ and $\sup_{F \in \mathcal{F}_m} E_F[\varphi]$ are independent of m.

In other words, even if μ_F is known to equal one of two given values m_1 and m_2, the sample V cannot provide effective discrimination between the two hypothetical values.

Let H be the hypothesis that $\mu_F = 0$ (i.e., $F \in \mathcal{F}_0$); i.e. the **null hypothesis**. For any test t, let $\beta_F(t)$ denote the probability of rejecting H in using t when F *obtains*, i.e. $\beta_F(t)$ denotes the **power function** of t - the probability that the test correctly rejects the null hypothesis when the alternative hypothesis is true.

1. t is a **somewhere unbiased** level-α test if $\beta_F \leqq \alpha$ for $F \in \mathcal{F}_0$ and, for some $m \neq 0$, $\beta_F \geqq \alpha$ for $F \in \mathcal{F}_m$

2. t is a **similar** level-α test if $\beta_F = \alpha$ for each $F \in \mathcal{F}_0$

Note: level-α test, refers to the probability of rejecting the null hypothesis when the null hypothesis is true, i.e. the false positive rate, a.k.a. type I error.

Let $\varphi(V)$ denote the probability prescribed by t of rejecting H (the null hypothesis) on observation of V yields the following corollary.

1.2   Corollary 1

If t is a somewhere unbiased level-α test of H, or a similar level-α test of H, then $\beta_F(t) = \alpha$ for all $F \in \mathcal{F}$.

The corollary illustrates that all tests of the value of μ, assuming that it exists, are unsuccessful.

Let I denotes a confidence set for μ_F, i.e. I is a (randomized) function of V.

For any real m, let $C[m]$ denote the event that I

1.3   Corollary II

If $P_F\left(C\left[\mu_F\right]\right) \geqq 1 - \alpha$ for all $F \in \mathcal{F}$, then $P_F(C[m]) \geqq 1 - \alpha$ for all m and all $F \in \mathcal{F}$.

Proof. For each m, let $p_m(V) = P(C[m] \,|\, V) = P_F(C[m])$, $0 \leqq p \leqq 1$. Consider a fixed m, since $\mu_F = m$, for all $F \in \mathcal{F}_m$, we have that $P_F(C[\mu_F]) = P_F(C[m]) = E_F[p_m] \geqq 1 - \alpha$ for $F \in \mathcal{F}_m$. By Theorem I, *for each bounded real valued function φ on the sample space of V, $\inf_{F \in \mathcal{F}_m} E_F[\varphi]$ and $\sup_{F \in \mathcal{F}_m} E_F[\varphi]$ are independent of m.* Hence, $P_F(C[m]) = E_F[p_m] \geqq 1 - \alpha$ for all $F \in \mathcal{F}$. Proof is complete for arbitrary m. $\square$

1.4   Corollary III

Suppose that there exists at least one $F \in \mathcal{F}$ such that $P_F(I$ is a set bounded from below $) = 1$. Then $\inf_{p_e\mathcal{F}}\{P_F(C[\mu_F])\} = 0$.

Proof. For each $n = 1, 2, \cdots$

Let B_n denote the event that $I \in [-n, \infty)$ - i.e., I is bounded from below;
Let $q_n(V)$ denote the probability of B_n given V; $0 \leqq q_n \leqq q_{n+1} \leqq 1$.

Let $\bar{B}_n$ denote the complement of B_n, i.e. the event that $I \in (-\infty, -n]$;
Let $1 - q_n(V)$ denote the probability of $\bar{B}_n$ given V.

Let F be a distribution in $\mathcal{F}$ s.t. I is bounded from below w.p. 1 when F obtains. By the Monotone Convergence Theorem, we can swap the expectation and the limit: $E_F[\lim_n q_n] = \lim_n E_F[q_n] = \lim_n P_F(B_n) = P_F(I$ is bounded from below $) = 1 \implies \lim_n q_n(V) = 1$ except on a set of points V of P_F-measure zero.

For any $m < -n$, $p_m(V) = \Pr(m \in I \mid V) \leqq \Pr(\bar{B}_n \mid V) = 1 - q_n(V)$, thus, except on a P_F-null set,

$$\lim_{m \to -\infty} p_m(V) = 0. \tag{2}$$

Since $P_F(C[m]) = E_F[p_m]$ for all m, it follows from (2), by Lebesgue's theorem for boundedly convergent sequences (Bounded Convergence Theorem - corollary of Dominated Convergence Theorem), that

$$\lim_{m \to -\infty} P_F(C[m]) = \lim_{m \to -\infty} E_F[p_m] = E_F\left[\lim_{m \to -\infty} p_m\right] = 0. \tag{3}$$

By Corollary 2: $\inf_{G \in \mathcal{F}}\{P_G(C[\mu_G])\} = \inf_{G \in \mathcal{F}} \inf_m \{P_G(C[m])\}$. It follows from (3) that the common value of these infima is zero. $\square$