# 1   Pointwise vs. Uniform Consistency in Level

Let $X_i, i = 1, \ldots, n$ be i.i.d. data from some distribution $P \in \mathbf{P}$. Suppose one wishes to test the null hypothesis $H_0 : P \in \mathbf{P}_0 \subsetneq \mathbf{P}$. To this end, one may consider a test function $\phi_n = \phi_n(X_1, \ldots, X_n)$ such that controls the probability of a Type 1 error in some sense. Ideally, we would like the test to satisfy

$$E_P[\phi_n] \le \alpha \text{ for all } P \in \mathbf{P}_0 , \tag{1}$$

but many times this is too demanding a requirement. As a result, we may settle instead for tests such that

$$\limsup_{n \to \infty} E_P[\phi_n] \le \alpha \text{ for all } P \in \mathbf{P}_0 . \tag{2}$$

Tests satisfying (1) are said to be of level $\alpha$ for $P \in \mathbf{P}_0$, whereas tests satisfying (2) are said to be pointwise asymptotically of level $\alpha$ for $P \in \mathbf{P}_0$. The hope is that if (2) holds, then (1) holds approximately, at least for large enough $n$. But this is not true. All that (2) ensures is that for each $P \in \mathbf{P}_0$ and $\epsilon > 0$ there is an $N(P)$ such that for all $n > N(P)$

$$E_P[\phi_n] \le \alpha + \epsilon .$$

Importantly, the sample size required for the approximation to work, $N(P)$, may depend on $P$. As a result, it could be the case that for every sample size $n$ (even, e.g., for $n = 10^{10}$) there could be $P = P_n \in \mathbf{P}_0$ such that

$$E_P[\phi_n] \gg \alpha .$$

Consider the following concrete example of this phenomenon. Suppose $\mathbf{P} = \{P \text{ on } \mathbf{R} : 0 < \sigma^2(P) < \infty\}$ and $\mathbf{P}_0 = \{P \in \mathbf{P} : \mu(P) = 0\}$. Let $\phi_n$ be the $t$-test; that is, $\phi_n = I\{\sqrt{n}\bar{X}_n > \hat{\sigma}_n z_{1-\alpha}\}$, where $z_{1-\alpha}$ is the $1 - \alpha$ quantile of the standard normal distribution. We know that

$$E_P[\phi_n] \to \alpha \text{ for all } P \in \mathbf{P}_0 ,$$

but it turns out that the $t$-test suffers from the problem described above. In fact, we can show that for every $0 < c < 1$ and every sample size $n$ there

exists a $P = P_{n,c}$ such that

$$E_P[\phi_n] \geq c .$$

To see this, let $n$ and $c$ be given. Let $P$ be the distribution that puts mass $1 - p$ on $p$ and mass $p$ on $-(1 - p)$. We will specify $p$ in a minute, but first note that for such a distribution $P$ all of the $X_i$ are in fact equal to $p > 0$ with probability $(1 - p)^n$. For such a sequence of observations, $\hat{\sigma}_n = 0$ and $\sqrt{n}\bar{X}_n > 0$, so $\phi_n = 1$. The probability of rejection, $E_P[\phi_n]$, is therefore at least $(1 - p)^n$. Now all that remains is to choose $p$ so that $(1 - p)^n = c$; that is, $p = 1 - c^{1/n}$.

To rule this very disturbing possibility out, we need to ensure that the convergence in (2) is uniform for $P \in \mathbf{P}_0$; that is,

$$\limsup_{n \to \infty} \sup_{P \in \mathbf{P}_0} E_P[\phi_n] \leq \alpha . \tag{3}$$

Tests satisfying (3) are said to be uniformly asymptotically of level $\alpha$ for $P \in \mathbf{P}_0$. This requirement implies that for each $\epsilon > 0$ there is an $N$ (which does not depend on $P$) such that for all $n > N$

$$E_P[\phi_n] \leq \alpha + \epsilon .$$

In the case of the $t$-test, the above example shows us that this is not true for $\mathbf{P} = \{P \text{ on } \mathbf{R} : 0 < \sigma^2(P) < \infty\}$ and $\mathbf{P}_0 = \{P \in \mathbf{P} : \mu(P) = 0\}$.

It is possible that this shortcoming is due to the $t$-test $--$ perhaps there are other tests of the same null hypothesis that would behave more reasonably. Unfortunately, we can show that this is not the case, provided that $\mathbf{P}$ is "'sufficiently rich". Formally, we have the following result due to Bahadur and Savage (1956):

**Theorem 1.1** Let $\mathbf{P}$ be a class of distributions on $\mathbf{R}$ such that

(i) For every $P \in \mathbf{P}$, $\mu(P)$ exists and is finite;

(ii) For every $m \in \mathbf{R}$, there is $P \in \mathbf{P}$ such that $\mu(P) = m$;

(iii) $\mathbf{P}$ is convex in the sense that if $P_1$ and $P_2$ are in $\mathbf{P}$, then $\gamma P_1 + (1-\gamma)P_2$ is in $\mathbf{P}$ for $\gamma \in [0,1]$.

Let $X_i, i = 1, \ldots, n$ be i.i.d. with distribution $P \in \mathbf{P}$. Let $\phi_n$ be any test of the null hypothesis $H_0 : \mu(P) = 0$. Then,

(a) Any test of $H_0$ which has size $\alpha$ for $\mathbf{P}$ has power $\leq \alpha$ for any alternative $P \in \mathbf{P}$.

(b) Any test of $H_0$ which has power $\beta$ against some alternative $P \in \mathbf{P}$ has size $\geq \beta$.

The proof of this result will follow from the following lemma:

**Lemma 1.1** Let $X_i, i = 1, \ldots, n$ be i.i.d. with distribution $P \in \mathbf{P}$, where a $\mathbf{P}$ is the class of distributions on $\mathbf{R}$ satisfying (i) - (iii) in Theorem 2.1. Let $\phi_n$ be any test function. Define

$$\mathbf{P}_m = \{P \in \mathbf{P} : \mu(P) = m\} \ .$$

Then,

$$\inf_{P \in \mathbf{P}_m} E_P[\phi_n] \text{ and } \sup_{P \in \mathbf{P}_m} E_P[\phi_n]$$

are independent of $m$.

PROOF: We show first that $\sup_{P \in \mathbf{P}_m} E_P[\phi_n]$ does not depend on $m$. Let $m$ be given and choose $m' \neq m$. We wish to show that

$$\sup_{P \in \mathbf{P}_{m'}} E_P[\phi_n] = \sup_{P \in \mathbf{P}_m} E_P[\phi_n] \ .$$

To this end, choose $P_j, j \geq 1$ so that

$$\lim_{j \to \infty} E_{P_j}[\phi_n] = \sup_{P \in \mathbf{P}_m} E_P[\phi_n] \ .$$

Let $h_j$ be defined so that

$$m' = (1 - \frac{1}{j})m + \frac{1}{j}h_j \ .$$

3

Choose $H_j$ so that $\mu(H_j) = h_j$. Define

$$G_j = (1 - \frac{1}{j})P_j + \frac{1}{j}H_j \ .$$

Thus, $G_j \in \mathbf{P}_{m'}$. Note that with probability $(1 - \frac{1}{j})^n$, a sample of size $n$ from $G_j$ is simply a sample of size $n$ from $P_j$. Therefore,

$$\sup_{P \in \mathbf{P}_{m'}} E_P[\phi_n] \geq E_{G_j}[\phi_n] \geq (1 - \frac{1}{j})^n E_{P_j}[\phi_n] \ .$$

But $(1 - \frac{1}{j})^n \to 1$ and $E_{P_j}[\phi_n] \to \sup_{P \in \mathbf{P}_m} E_P[\phi_n]$ as $j \to \infty$. Therefore,

$$\sup_{P \in \mathbf{P}_{m'}} E_P[\phi_n] \geq \sup_{P \in \mathbf{P}_m} E_P[\phi_n] \ .$$

Interchanging the roles of $m$ and $m'$, we can establish the reverse inequality

$$\sup_{P \in \mathbf{P}_{m'}} E_P[\phi_n] \leq \sup_{P \in \mathbf{P}_m} E_P[\phi_n] \ .$$

We could replace $\phi_n$ with $1 - \phi_n$ to establish that $\inf_{P \in \mathbf{P}_m} E_P[\phi_n]$ does not depend on $m$. ■

PROOF OF THEOREM 2.1: (a) Let $\phi_n$ be a test of size $\alpha$ for $\mathbf{P}$. Let $P'$ be any alternative. Define $m = \mu(P')$. Then,

$$E_{P'}[\phi_n] \leq \sup_{P \in \mathbf{P}_m} E_P[\phi_n] = \sup_{P \in \mathbf{P}_0} E_P[\phi_n] = \alpha \ .$$

The proof of (b) is similar. ■

The class of distributions with finite second moment satisfies the requirements of the theorem, as does the class of distributions with infinitely many moments. Thus, the failure of the $t$-test is not special to the $t$-test; in this setting, there simply exist no "reasonable" tests. But this does not mean that all hope is lost. By restricting the class of distributions some, it is possible to construct reasonable tests about the mean. In fact, the $t$-test does satisfy (3) for certain large classes of distributions that are somewhat smaller than $\mathbf{P}_0$. See Chapter 11 of Lehmann and Romano (2005) for details.

4