0.1 Basic permutation test

Let X be data taking values in a sample space $\mathcal{X}$. Let G be a finite set of transformations $g : \mathcal{X} \to \mathcal{X}$, such that G is a group with respect to the operation of composition of transformations. This means that G satisfies the following three properties: G contains an identity element (the map $x \mapsto x$); every element of G has an inverse in G; for all $a_1, a_2 \in G$, $a_1 \circ a_2 \in G$. This assumption of a group structure for G is fundamental throughout the paper, since it ensures that $Gg = G$ for all $g \in G$, i.e. that the set G is permutation invariant.

Considering a general group of transformations rather than only permutations is useful, since in many practical situations the group consists of, for example, rotations Langsrud2005, Solari2014 or maps that multiply part of the data by -1 Pesarin2010. We write $g(X)$ as gX. Consider any test statistic $T : \mathcal{X} \to R$. Throughout this paper, we are concerned with testing the following null hypothesis of permutation invariance.

Let H_p be any null hypothesis which implies that the joint distribution of the test statistics $T(gX)$, $g \in G$, is invariant under all transformations in G of X. That is, writing $G = \{a_1, \ldots, a_G\}$, under H_p

$$(T(a_1X), \ldots, T(a_GX))\, d{=}\, (T(a_1gX), \ldots, T(a_GgX))$$

for all $g \in G$.

Note that eq:null$_h ypothesisholdsinparticularwhenforall$g $\in G$

$$X d{=} gX.$$

Composite null hypotheses are usually not of the form H_p, but for specific scenarios, properties of tests of such hypotheses can be established using results in this paper.

The most basic permutation test rejects H_p when $T(X) > T^{(k)}(X)$, where

$$T^{(1)}(X) \leq \cdots \leq T^{(G)}(X)$$

are the sorted test statistics $T(gX)$, $g \in G$, and $k = \lfloor (1-\alpha)G \rfloor$ with $\alpha \in [0, 1)$. As is known and stated in the following theorem, this test has level at most α.

Under H_p, $P\{T(X) > T^{(k)}(X)\} \leq \alpha$.

We now give two proofs: a conditioning-based approach and an approach without conditioning. Both approaches are more or less known. The conditioning-based proof is similar to that in Pesarin2015, but the setting is more general. For each $x \in \mathcal{X}$, define O_x to be the orbit of x, which is the set $\{gx : g \in G\} \subseteq \mathcal{X}$.

[Proof of Theorem 1] Let $A = \{x \in \mathcal{X} : T(x) > T^{(k)}(x)\}$ be the set of elements of the sample space that lead to rejection. Suppose H_p holds. By the group structure, $Gg = G$ for all $g \in G$. Consequently, $T^{(k)}(gX) = T^{(k)}(X)$ for all $g \in G$. Thus, $\#g \in G : gX \in A = \#g \in G : T(gX) > T^{(k)}(gX)\} = \#g \in G : T(gX) > T^{(k)}(X)\} \leq \alpha G$.

Endow the space of orbits with the σ-algebra that it inherits from the σ-algebra on $\mathcal{X}$. Analogously to the proof of Theorem 15.2.2 in Lehmann2005, we obtain

$$P(X \in A | O_X) = \frac{1}{G} \# g \in G : gX \in A.$$

By the argument above, this is bounded by α. Hence,

$$P(X \in A) = E\{P(X \in A|O_X)\} \leq \alpha$$

as was to be shown.

We now state a different proof without conditioning. A similar proof can be found in Hoeffding1952 and Lehmann2005 (p. 634).

[Alternative proof of Theorem 1] By the group structure, $Gg = G$ for all $g \in G$. Hence, $T^{(k)}(gX) = T^{(k)}(X)$ for all $g \in G$. Let h have the uniform distribution on G. Then, under H_p, the rejection probability is

$$P\{T(X) > T^{(k)}(X)\} = P\{T(hX) > T^{(k)}(hX)\} = P\{T(hX) > T^{(k)}(X)\}.$$

The first equality follows from the null hypothesis, and the second equality holds since $T^{(k)}(X) = T^{(k)}(hX)$. Since h is uniform on G, the above probability equals

$$E\left[(G)^{-1} \cdot \#g \in G : T(gX) > T^{(k)}(X)\}\right] \leq \alpha,$$

as was to be shown.

0.2 Permutation p values

Permutation p values are p values based on permutations of the data. Here we will discuss permutation p values based on the full permutation group. p values based on random permutations are considered in Section **??**.

It is essential to note that there is often no unique null distribution of $T(X)$, since H_p often does not specify a unique null distribution of the data. Correspondingly, $T^{(k)}(X)$ should not be seen as the $(1-\alpha)$-quantile of the null distribution.

When a test statistic t is a function (which is not random) of the data and has a unique distribution under a hypothesis H, then a p value in the strict sense, $P_H(t \geq t_{obs})$, is defined where t_{obs} is the observed value of t. Since under H_p $T(X)$ does not always have a unique null distribution, often there exists no p value in the strict sense based on this test statistic. However, under Condition 1 the statistic

$$D = \#\{g \in G : T(gX) \geq T(X)\}$$

does have a unique null distribution. Thus, a p value in the strict sense based on $-D$ is then defined. Denoting by d the observed value