The Annals of Statistics
1973, Vol. 1, No. 1, 166-169

AN ASYMPTOTIC UMP SIGN TEST IN THE PRESENCE OF TIES

BY J. KRAUTH

University of Düsseldorf

A modification of the one-sided sign test, which was proposed by Putter (1955), is shown to be an asymptotic UMP test. In contrast to the usual practice, zeros are not omitted but taken into account.

1. Introduction and results. Putter (1955) considered the sign test which amounts to "omitting the ties from the observations" for testing the hypothesis

$$H: P(Z_k > 0) = P(Z_k < 0),$$

against the alternative

$$A: P(Z_k > 0) > P(Z_k < 0).$$

He proved this test to be the UMP unbiased test based on n_+ and n_0.

In Putter's notation $Z_1, \cdots, Z_n$ are i.i.d. rv's, n_+ is the number of positive Z_k's, n_- of negative Z_k's, and n_0 of zeros among the Z_k's. Further denote

$$P(Z_k > 0 \mid H) = p_+, \qquad P(Z_k = 0 \mid H) = p_0, \qquad P(Z_k < 0 \mid H) = p_-,$$
$$P(Z_k > 0 \mid A) = q_+, \qquad P(Z_k = 0 \mid A) = q_0, \qquad P(Z_k < 0 \mid A) = q_-.$$

All these probabilities are assumed to be greater zero.

In this note we shall prove the following two theorems.

THEOREM 1. *An* UMP *test for testing H against A with a known constant $p_0 = q_0$ is given by*

$$n_+ + \tfrac{1}{2}n_0 > k_n(p_0). \tag{1.1}$$

The test (1.1) was proposed e.g. by Dixon and Mood (1946). Since in general the parameter p_0 is unknown, a usual practice is to take the cutoff point corresponding to $\mathfrak{B}(n, \frac{1}{2})$ instead of $k_n(p_0)$. Hemelrijk (1952) has shown that this modification of the test is uniformly inferior to the conditional test with regard to power. The following example shows that this is not true for the exact test, as we might have concluded from Theorem 1, of course.

EXAMPLE. Putter (1955) compares the power of the conditional test and the randomized test for $n = 10$, $q_0 = q_- = \frac{1}{4}$, $q_+ = \frac{1}{2}$. Both tests are considered on the .05 level. His values are .221 and .195, respectively. The power of the test (1.1) (for $p_0 = q_0$) is .230.

Received September 14, 1971; revised May 15, 1972.
AMS 1970 *subject classifications.* Primary 62G10.
Key words and phrases. Sign test, ties, optimality.

THEOREM 2. *An asymptotic* UMP *test for testing H against A under the additional restriction* $p_0 = q_0$ *is given by*

$$T_n = (2n_+ + n_0 - n)/(n - n_0)^{\frac{1}{2}} > k \,, \tag{1.2}$$

where the cutoff point k is corresponding to the $\mathfrak{N}(0, 1)$*—distribution.*

The test (1.2) was proposed by Putter (1955), who proved it to be asymptotically more efficient than the randomized test.

2. Outline of the proofs. Consider the distribution of

$$n_+ - n_- = 2n_+ + n_0 - n = 2(n_+ + \tfrac{1}{2}n_0) - n \,. \tag{2.1}$$

Under A,

$$P_A(x) = P(n_+ - n_- = x) = \sum_{n_+ - n_- = x} \frac{n!}{n_+!\, n_-!\, n_0!} q_+^{n_+} q_-^{n_-} q_0^{n_0}$$
$$= q_0^n (q_+/q_0)^x \sum_{i=0}^n \binom{n}{i}\binom{n-i}{i-x}(q_+ q_-/q_0^2)^{i-x} \,;$$

under H,

$$P_H(x) = P(n_+ - n_- = x) = p_0^n (p_+/p_0)^x \sum_{i=0}^n \binom{n}{i}\binom{n-i}{i-x}(p_+ p_-/p_0^2)^{i-x} \,,$$

$x = -n, -n+1, \cdots, n$. Thence

$$\frac{P_A(x)}{P_H(x)} = \frac{y^x A_n(z, x)}{y_0^x A_n(z_0, x)} \tag{2.2}$$

with

$$A_n(z, x) = \sum_{i=0}^n \binom{n}{i}\binom{n-i}{i-x} z^{i-x} \,, \tag{2.3}$$

$$y = q_+/q_0 \,, \qquad y_0 = p_+/p_0 = (1 - p_0)/2p_0 \,,$$
$$z = q_+ q_-/q_0^2 \,, \qquad z_0 = p_+ p_-/p_0^2 = (1 - p_0)^2/4p_0^2 \,.$$

$P_A(x)/P_H(x)$ is a strictly increasing function of x, as is proved in the lemma below. Therefore by the Neyman-Pearson lemma a UMP test for known $p_0 = q_0$ is given by (1.1). Considering $T_n \to_L \mathfrak{N}(0, 1)$ (Putter (1955), (4.3)) and (2.1) the test (1.2) is a most powerful asymptotic test independent of p_0, i.e. an asymptotic UMP test for $p_0 = q_0$. The following lemma proves $P_A(x)/P_H(x)$ to be a strictly increasing function of x, because $z < z_0$ and $(y/y_0)^x$ is strictly increasing in x for $y > y_0$.

LEMMA. *The inequality*

$$A_n(z, x)/A_n(z_0, x) > A_n(z, x-1)/A_n(z_0, x-1) \tag{2.4}$$

is valid for $z < z_0$, $x = -n+1, \cdots, n$.

PROOF. It suffices to prove

$$A_n(z, x)/A_n(z, x-1) > A_n(z_0, x)/A_n(z_0, x-1) \tag{2.5}$$

for $z < z_0$, $x = -n+1, \cdots, n$. We get (2.5) by showing that the derivative of

$$H(z, x) = A_n(z, x)/A_n(z, x-1) \tag{2.6}$$

with respect to z is negative for $z > 0$, $x = -n+1, \cdots, n$.

For $x \leqq 0$ we have

$$A_n(z, x) = \sum_{i=0}^{m} \binom{n}{i}\binom{n-i}{i-x} z^{i-x} \tag{2.7}$$

with $m = [(n + x)/2]$. Therefore it is enough to prove

$$\sum_{i=0}^{m} \sum_{j=0}^{m'} (i - j - 1)\binom{n}{i}\binom{n-i}{i-x}\binom{n}{j}\binom{n-j}{j-x+1} z^{i+j} < 0$$

with $m' = [(n + x - 1)/2]$ or

$$\sum_{i=0}^{m} \sum_{j=1}^{m'+1} (i - j)\binom{n}{i}\binom{n-i}{i-x}\binom{n}{j-1}\binom{n-j+1}{j-x} z^{i+j-1} < 0 \,. \tag{2.8}$$

We consider only the terms for which $i, j \in \{1, \cdots, m\}$, because the terms with $i = 0$ or $j = m + 1$ are negative anyhow. For $i = j$ the summands vanish. For the sum of two terms with

$$(i_1, j_1) = (s, t)\,, \qquad (i_2, j_2) = (t, s)\,, \qquad s, t \in \{1, \cdots, m\}\,, \tag{2.9}$$

we get

$$-(s - t)^2\binom{n}{s}\binom{n-s}{s-x}\binom{n}{t}\binom{n-t}{t-x} \times \frac{(n + x + 1)}{(n - 2t + x + 1)(n - 2s + x + 1)} z^{s+t-1}\,, \tag{2.10}$$

which is negative for all $s, t \in \{1, \cdots, m\}$. This completes the proof for $x \leqq 0$.

For $x \geqq 0$ we have

$$A_n(z, x) = \sum_{i=x}^{m} \binom{n}{i}\binom{n-i}{i-x} z^{i-x} \,. \tag{2.11}$$

We must prove

$$\sum_{i=x}^{m} \sum_{j=x}^{m'+1} (i - j)\binom{n}{i}\binom{n-i}{i-x}\binom{n}{j-1}\binom{n-j+1}{j-x} z^{i+j-1} < 0\,, \tag{2.12}$$

which follows in the same way as (2.8).

3. Generalization. Looking through the proofs one can easily see that the restriction $p_0 = q_0$ is unnecessarily strong. What we really needed was $y \geqq y_0$, $z < z_0$ or $y > y_0$, $z \leqq z_0$, respectively. Hence the results also hold, if the set of admissible alternatives is enlarged by those with either

$$p_0 < q_0 \quad \text{and} \quad q_+ - q_- \geqq (q_0 - p_0)/p_0 \tag{3.1}$$

or

$$p_0 > q_0 \quad \text{and} \quad (q_+ - q_-)^2/(q_+^{\frac{1}{2}} + q_-^{\frac{1}{2}})^2 \geqq (p_0 - q_0)/p_0 \tag{3.2}$$

or

$$p_0 > q_0 \quad \text{and} \quad (q_+ - q_-)^2/(q_+^{\frac{1}{2}} - q_-^{\frac{1}{2}})^2 \leqq (p_0 - q_0)/p_0 \,. \tag{3.3}$$

REFERENCES

[1] Dixon, W. J. and Mood, A. M. (1946). The statistical sign test. *J. Amer. Statist. Assoc.* **41** 557-566.

[2] HEMELRIJK, J. (1952). A theorem on the sign test when ties are present. *Nederl. Akad. Wetensch. Proc. Ser. A* **55** 322–326.

[3] PUTTER J. (1955). The treatment of ties in some nonparametric tests. *Ann. Math. Statist.* **26** 368–386.

UNIVERSITY OF DÜSSELDORF
HAROLDSTRASSE 19
4 DÜSSELDORF, GERMANY