

THE POWER OF CERTAIN NON-PARAMETRIC TESTS.*

by

WASSILY HOEFFDING

Special report to the Office of Naval
Research of work at Chapel Hill under
Project NR 042 031 for research in
probability and statistics.

Institute of Statistics
Mimeo Series No. 56
For Limited Distribution
August, 1952

THE POWER OF CERTAIN NON-PARAMETRIC TESTS.*

By Wassily Hoeffding

1. Introduction and summary.

Consider any non-parametric two-sample test which has a given significance level for testing the hypothesis H that all observed variates are independent with a common continuous distribution. Its power against the alternative A that the two samples come from two normal populations with common variance and different means is necessarily less than the power of the appropriate one-sided two-sample t -test of the same significance level, at least if the non-parametric test is "similar". This will be a fortiori the case if the non-parametric test depends only on the rank order, not on the values of the observations. It has, however, recently been found that there exist non-parametric tests, and even rank order tests, whose power against the alternative A is arbitrarily close to the power of the standard parametric test if the samples are sufficiently large. Analogous results apply to other tests of randomness, tests of independence, etc. This paper will give a sketch of these results and of their derivation. The two-sample case will serve to illustrate the essential points.

2. Tests based on permutations of observations.

Let \underline{X} denote the vector (X_1, \dots, X_N) , and let $t_N(\underline{X})$ be the standard t -statistic for testing whether the two samples (X_1, \dots, X_m) and (X_{m+1}, \dots, X_N) come from the same normal population. If $\underline{r} = (r_1, \dots, r_N)$ is any permutation of $(1, \dots, N)$, let $\underline{X}_{\underline{r}} = (X_{r_1}, \dots, X_{r_N})$. For any given \underline{X} let

$$t_N^{(1)}(\underline{X}) \leq t_N^{(2)}(\underline{X}) \leq \dots \leq t_N^{(M)}(\underline{X})$$

be the ordered values $t_N(\underline{X}_{\underline{r}})$ corresponding to the $M = N!$ permutations \underline{r} . Given

*Work done under the sponsorship of the Office of Naval Research.

a number α , $0 < \alpha < 1$, let k be defined by $k = M - \lfloor M\alpha \rfloor$, where $\lfloor M\alpha \rfloor$ denotes the largest integer $\leq M\alpha$. Let T be the test which rejects H if and only if

$$(1) \quad t_N(\underline{X}) > t_N^{(k)}(\underline{X}).$$

This test has significance level α for testing H . It differs only in a trivial way from a test which Lehmann and Stein [3] have shown to be most powerful similar of size α for testing H against the normal alternative A of the introduction, where the sign of the difference of the means is specified.

The standard one-sided t -test of size α rejects the null-hypothesis if and only if

$$(2) \quad t_N(\underline{X}) > \lambda_N,$$

where $\lim_{N \rightarrow \infty} \lambda_N = \lambda$,

$$\phi(-\lambda) = \alpha, \quad \phi(y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^y e^{-\frac{t^2}{2}} dt.$$

Suppose we have a sequence of distributions P_N of $\underline{X} = \underline{X}^{(N)}$, defined for every N , such that as $N \rightarrow \infty$,

$$(3) \quad t_N^{(k)}(\underline{X}) \rightarrow \lambda \text{ in probability}$$

and

$$(4) \quad \Pr \{ t_N(\underline{X}) \leq y \} \rightarrow H(y), \text{ a function continuous at } y = \lambda.$$

Then a comparison of (1) and (2) shows that the powers against the alternative P_N of the test T and the standard test tend to the same limit $1 - H(\lambda)$.

In the case of the normal alternative A , if σ^2 denotes the common variance and Δ the difference of the means, it can be shown by standard methods that the power of the t -test tends to a limit if and only if

$$(5) \quad \lim_{N \rightarrow \infty} \frac{\Delta^2}{\sigma^2} \frac{m(N-m)}{N} = \delta^2 (\leq \infty) \text{ exists,}$$

and then the limit of the power is $\phi(-\lambda + \delta)$. Here Δ and σ may depend on N .

The author has shown in [1] that (3) is satisfied for an extensive class of alternatives which includes the normal alternative A, where, if Δ/σ is fixed, m/N tends to some limit. We give here a brief outline of the method of proof.

I. For every vector $\underline{x} = (x_1, \dots, x_N)$ let $M_N^F(y, \underline{x})$ be the number of permutations \underline{r} for which $t_{N-\underline{r}}(\underline{x}) \leq y$. Then

$$\Pr \left\{ t_N^{(k)}(\underline{X}) \leq y \right\} = \Pr \left\{ F_N(y, \underline{X}) \geq \frac{k}{M} \right\}.$$

Since $\frac{k}{M} \rightarrow 1 - \alpha = \phi(\lambda)$, it easily follows that (3) is satisfied if, as $N \rightarrow \infty$,

$$(6) \quad F_N(y, \underline{X}) \rightarrow \phi(y) \text{ in probability for every } y.$$

II. Let \underline{R} be a random vector whose values are the $M = N!$ equally probable permutations \underline{r} . Let \underline{R}' have the same distribution as \underline{R} , and suppose that \underline{R} , \underline{R}' and \underline{X} are mutually independent. Then

$$\begin{aligned} EF_N(y, \underline{X}) &= \Pr \left\{ t_{N-\underline{R}}(\underline{X}) \leq y \right\}, \\ EF_N(y, \underline{X})^2 &= \Pr \left\{ t_{N-\underline{R}}(\underline{X}) \leq y, t_{N-\underline{R}'}(\underline{X}) \leq y \right\}. \end{aligned}$$

It follows that (6) is satisfied if $t_{N-\underline{R}}(\underline{X})$, $t_{N-\underline{R}'}(\underline{X})$ have the joint limiting distribution of two independent standard normal random variables,

$$(7) \quad \Pr \left\{ t_{N-\underline{R}}(\underline{X}) \leq y, t_{N-\underline{R}'}(\underline{X}) \leq y' \right\} \rightarrow \phi(y)\phi(y').$$

III. If the alternative A is true, the vector $(t_{N-\underline{R}}(\underline{X}), t_{N-\underline{R}'}(\underline{X}))$ can be shown to have the same limiting distribution (if any) as $(L_N(\underline{X}_{\underline{R}}), L_N(\underline{X}_{\underline{R}'}))$, where

$L_N(\underline{x})$ is a linear function of x_1, \dots, x_N . For any two fixed permutations $\underline{r}, \underline{r}'$ the vector $(L_N(X_{\underline{r}}), L_N(X_{\underline{r}'}))$ has a bivariate normal distribution whose correlation coefficient $\rho(\underline{r}, \underline{r}')$ and means $m_1(\underline{r})$ and $m_2(\underline{r}')$ are functions of \underline{r} and \underline{r}' , while the variances are independent of $\underline{r}, \underline{r}'$. It can be shown that as $N \rightarrow \infty$, $\rho(\underline{R}, \underline{R}')$ tends to 0 in probability. If $\Delta = 0$, then $m_1 = m_2 = 0$, and (7) is now easy to prove. In the general case it can be shown that $m_1(\underline{R})$ and $m_2(\underline{R}')$ have normal limiting distributions with zero means and are in the limit independent of the other random variables involved. This again implies (7).

A similar argument can be applied to more general, non-normal alternatives by using the central limit theorem for vectors. For details of the proof see [1].

3. Rank order tests.

Let R_i be the rank of X_i in the set (X_1, \dots, X_N) . If H is true, the random vector $\underline{R} = (R_1, \dots, R_N)$ takes on the $N!$ permutations \underline{r} of $(1, \dots, N)$ with equal probabilities. Let $Z_{N1} \leq Z_{N2} \leq \dots \leq Z_{NN}$ be the ordered values of N independent normal random variables with mean 0 and variance 1. Let

$$c_1(\underline{r}) = \sum_{i=m+1}^N EZ_{Ni} r_i,$$

and let c^* be the least value of $c_1(\underline{r})$ such that the inequality $c_1(\underline{r}) \geq c^*$ holds for not more than $\alpha N!$ permutations \underline{r} . Then the test which rejects H if and only if $c_1(\underline{R}) \geq c^*$ has significance level α for testing H . As shown in [2], this is the most powerful rank order test against the normal alternative A if Δ/σ is sufficiently small and the sample (X_{m+1}, \dots, X_N) comes from the population with the larger mean.

The author has shown in an unpublished paper that if this last-mentioned alternative is true and condition (5) is satisfied, the power of the c_1 -test

tends to the same limit as the power of the standard t-test, provided only that both sample sizes, m and $N-m$, tend to infinity. The proof proceeds as follows.

If X_1, \dots, X_N are independent normal with common variance σ^2 and means $EX_i = \Gamma, i = 1, \dots, m; EX_i = \Gamma + \Delta, i = m+1, \dots, N$, it is easy to show (see [2]) that the probability $P(\underline{r}, \delta_N)$ of $\underline{R} = \underline{r}$ can be written in the form

$$(8) \quad P(\underline{r}, \delta_N) = \frac{1}{N!} e^{-\delta_N^2/2} E e^{\delta_N \sum_{i=1}^N b_i Z_{Nr_i}},$$

where

$$\delta_N = \frac{\Delta}{\sigma} \sqrt{\frac{m(N-m)}{N}},$$

$$b_i = \frac{-N+m}{\sqrt{Nm(N-m)}}, \quad i = 1, \dots, m; \quad b_i = \frac{m}{\sqrt{Nm(N-m)}}, \quad i = m+1, \dots, N.$$

Assumption (5) implies that $\lim \delta_N = \delta$, and it suffices to consider the case where δ is finite.

Let

$$k_N(\underline{r}) = \sum_{i=1}^N b_i E Z_{Nr_i},$$

so that

$$k_N(\underline{r}) = \sqrt{\frac{N}{m(N-m)}} c_1(\underline{r}).$$

If the moment-generating function of $k_N(\underline{R}) - \delta_N$ when \underline{R} has the distribution (8) is denoted by $M_{N, \delta_N}(t)$, it is sufficient to show that

$$(9) \quad \lim_{N \rightarrow \infty} M_{N, \delta_N}(t) = e^{t^2/2}.$$

We have

$$\begin{aligned}
 (10) \quad M_{N, \delta_N}(t) &= \sum_{\underline{r}} e^{t \sqrt{k_N(\underline{r})} - \delta_N} P(\underline{r}, \delta_N) \\
 &= e^{-\delta_N^2/2 - t \delta_N} \frac{1}{N!} \sum_{\underline{r}} e^{(t + \delta_N) k_N(\underline{r})} E e^{\delta_N U_{N \underline{r}}},
 \end{aligned}$$

where

$$U_{N \underline{r}} = \sum_{i=1}^N b_i Z_{N r_i} - k_N(\underline{r}) = \sum_{i=1}^N b_i (Z_{N r_i} - E Z_{N r_i}).$$

The proof of (9) consists in showing (i) that (9) holds when $\delta_N = 0$ and (ii) that for "most" permutations \underline{r} , $U_{N \underline{r}}$ is close to 0 with high probability,

in such a way that the factor $E e^{\delta_N U_{N \underline{r}}}$ in (10) can be replaced by 1 as $N \rightarrow \infty$.

The proof of (i), that is,

$$(11) \quad M_{N, 0}(t) = \frac{1}{N!} \sum_{\underline{r}} e^{t k_N(\underline{r})} \longrightarrow e^{t^2/2},$$

is easy. From a well-known theorem of Wald and Wolfowitz and its extension by Noether, it follows that the moments of any order of $k_N(\underline{r})$ in this case converge to the corresponding moments of the standard normal distribution. With the help of Taylor's formula this easily implies that

$$\liminf_{N \rightarrow \infty} M_{N, 0}(t) \geq e^{t^2/2}.$$

On the other hand, since $E U_{N \underline{r}} = 0$, we have $E e^{\delta_N U_{N \underline{r}}} \geq 1$, and hence, from (10),

$$(12) \quad M_{N, \delta_N}(t) \geq e^{-\delta_N^2/2 - t\delta_N} M_{N,0}(t+\delta_N).$$

Letting $t = 0$ and replacing δ_N by t , we get

$$M_{N,0}(t) \leq e^{t^2/2}.$$

One part of (ii) is also easy to prove, for (11) and (12) imply

$$\liminf_{N \rightarrow \infty} M_{N, \delta_N}(t) \geq e^{t^2/2}.$$

The opposite inequality is less immediate. It can be proved by first showing that

$$\lim_{N \rightarrow \infty} \frac{1}{N!} \sum_r e^{\delta_N r} = 1$$

uniformly in any finite δ -interval, and then applying Höder's inequality to (10).

This method permits us to study the power of the c_1 -test only against the alternative with respect to which the test is most powerful. An alternative, more flexible method has been developed by M. Dwass in a paper which is not yet published.

4. Concluding remarks.

The results which have been outlined in this paper are of a rather restricted nature. Under ordinary assumptions the power of the standard test will converge to 1 as N tends to infinity. In this case we only obtain that the powers of the two non-parametric tests also tend to 1, and this property is undoubtedly shared by many other non-parametric tests. The results are non-trivial only if the limit of the power is less than 1, and this ordinarily means that the alternative is "close" to the null hypothesis when N is large. It would be desirable

to investigate more closely the performance characteristics of these tests when their power tends to 1.

REFERENCES.

- [1] W. Hoeffding, "The large-sample power of tests based on permutations of observations," Annals of Mathematical Statistics, Vol. 23 (1952), pp. 169 - 192.
- [2] W. Hoeffding, "'Optimum' non-parametric tests," Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, University of California Press, 1951, pp. 83-92.
- [3] E. L. Lehmann and C. Stein, "On the theory of some nonparametric hypotheses," Annals of Mathematical Statistics, Vol. 20 (1949), pp. 28-45.