Corrections entered
June 30
6 copies
given to Miss Kellington
for transmittal to NOTS
on July 1, 1952

NATIONAL BUREAU OF STANDARDS REPORT

1694

SIMULTANEOUS TESTS OF RANDOMNESS AND OTHER HYPOTHESES

By

I. Richard Savage

U. S. DEPARTMENT OF COMMERCE
NATIONAL BUREAU OF STANDARDS

U. S. DEPARTMENT OF COMMERCE
Charles Sawyer, *Secretary*

NATIONAL BUREAU OF STANDARDS
A. V. Astin, *Acting Director*

THE NATIONAL BUREAU OF STANDARDS

The scope of activities of the National Bureau of Standards is suggested in the following listing of the divisions and sections engaged in technical work. In general, each section is engaged in specialized research, development, and engineering in the field indicated by its title. A brief description of the activities, and of the resultant reports and publications, appears on the inside of the back cover of this report.

1. **Electricity.** Resistance Measurements. Inductance and Capacitance. Electrical Instruments. Magnetic Measurements. Electrochemistry.

2. **Optics and Metrology.** Photometry and Colorimetry. Optical Instruments. Photographic Technology. Length. Gage.

3. **Heat and Power.** Temperature Measurements. Thermodynamics. Cryogenics. Engines and Lubrication. Engine Fuels.

4. **Atomic and Radiation Physics.** Spectroscopy. Radiometry. Mass Spectrometry. Physical Electronics. Electron Physics. Atomic Physics. Neutron Measurements. Nuclear Physics. Radioactivity. X-Rays. Betatron. Nucleonic Instrumentation. Radiological Equipment. Atomic Energy Commission Instruments Branch.

5. **Chemistry.** Organic Coatings. Surface Chemistry. Organic Chemistry. Analytical Chemistry. Inorganic Chemistry. Electrodeposition. Gas Chemistry. Physical Chemistry. Thermochemistry. Spectrochemistry. Pure Substances.

6. **Mechanics.** Sound. Mechanical Instruments. Aerodynamics. Engineering Mechanics. Hydraulics. Mass. Capacity, Density, and Fluid Meters.

7. **Organic and Fibrous Materials.** Rubber. Textiles. Paper. Leather. Testing and Specifications. Organic Plastics. Dental Research.

8. **Metallurgy.** Thermal Metallurgy. Chemical Metallurgy. Mechanical Metallurgy. Corrosion.

9. **Mineral Products.** Porcelain and Pottery. Glass. Refractories. Enameled Metals. Building Stone. Concreting Materials. Constitution and Microstructure. Chemistry of Mineral Products.

10. **Building Technology.** Structural Engineering. Fire Protection. Heating and Air Conditioning. Exterior and Interior Coverings. Codes and Specifications.

11. **Applied Mathematics.** Numerical Analysis. Computation. Statistical Engineering. Machine Development.

12. **Electronics.** Engineering Electronics. Electron Tubes. Electronic Computers. Electronic Instrumentation.

13. **Ordnance Development.** Mechanical Research and Development. Electromechanical Fuzes. Technical Services. Missile Fuzing Research. Missile Fuzing Development. Projectile Fuzes. Ordnance Components. Ordnance Tests. Ordnance Research.

14. **Radio Propagation.** Upper Atmosphere Research. Ionospheric Research. Regular Propagation Services. Frequency Utilization Research. Tropospheric Propagation Research. High Frequency Standards. Microwave Standards.

15. **Missile Development.** Missile Engineering. Missile Dynamics. Missile Intelligence. Missile Instrumentation. Technical Services. Combustion.

NATIONAL BUREAU OF STANDARDS REPORT

NBS PROJECT | **NBS REPORT**

1103-21-5119 | 28 May 1952 | 1694

SIMULTANEOUS TESTS OF RANDOMNESS AND OTHER HYPOTHESES

by

I. Richard Savage

Research supported in part by the
U. S. Naval Ordnance Test Station, Inyokern

FOREWORD

It is shown in this report that rank order statistics and symmetric statistics are independently distributed under the assumption of random sampling. Techniques are given that are particularly useful in testing certain hypotheses concerning stochastic processes.

This research was carried out in the Statistical Engineering Laboratory, and was sponsored in part by the U. S. Naval Ordnance Test Station, Inyokern, under Contract NP45-507/L4-3 with the National Bureau of Standards. The Statistical Engineering Laboratory is Section 11.3 of the National Applied Mathematics Laboratories (Division 11 of the National Bureau of Standards), and is concerned with the development and application of modern statistical methods in the physical sciences and engineering.

J. H. Curtiss
Chief, National Applied
Mathematics Laboratories

A. V. Astin
Acting Director
National Bureau of Standards

SIMULTANEOUS TESTS OF RANDOMNESS AND OTHER HYPOTHESES

By

I. Richard Savage

Summary: In this report it is shown that under the hypothesis of randomness rank order statistics and symmetric statistics are independent. This fact is of use in testing hypotheses as is shown by examples.

Introduction: A common problem in statistical inference is "Do the observations $x_1, \ldots x_n$ come from a population centered at the origin?" One solution to this problem is to use the ordinary "t" test of this hypothesis. A more careful observer might notice that there is no assumption of normality which is needed in order to justify the use of the "t" test. The observer might then interpret "centered at the origin" as "median at the origin" and use the sign test. Finally the observer may notice that there is no assumption that the observations were made on independently and identicaly distributed random variables. The objective of this report is to show that many of the tests of randomness (identical and independent random variables) are independent of the common tests of hypotheses. More precisely it will be shown that if one has random samples then many of the tests

of randomness will be independent of ordinary tests of hypothesis. Or under the null hypothesis of random sampling tests of randomness are independent of ordinary tests of hypotheses. Using this fact the statistician will be able to test both the "random" and "parametric" parts of a hypothesis with known significance levels.

As a further example one might have $x(t)$ an observation on a Fundamental Random Process [Mann(1951)]*. One knows that if the null hypothesis is true then the quantities $x(i\delta) - x[(i-1)\delta]$ $(i=1,\ldots n)$ form a random sample from a normal distribution with mean zero and variance proportional to δ. To test this hypothesis one must test for both randomness and normality. The test for randomness would depend on the alternatives of interest, perhaps one of the run tests [Levene(1952)] would be found suitable. The test of normality could be performed using the classical chi square goodness of fit test with one parameter estimated. If this were the test program this report will show that the run test and the chi square test are independent under the null hypothesis.

Assumptions and Notation: We shall be concerned with real valued random variables x_i or x_{ij} (depending if there is one or several samples, x_i being the i^{th} observation in a

---

* Names followed by dates in square brackets refere to items in the bibliography at the end of this report.

sample, and x_{ij} being the j^{th} observation in the i^{th} sample). We shall assume that all of the observations come from distributions that have continuous cummulative distribution functions. Letting P(A) stand for the probability of the event A(P(A|B) will stand for the conditional probability of the event A given that the event B has occurred), the assumption of continuity implies:

(1) $P(x_{ij}=a) = 0$

(2) $P(x_{ij}=x_{i'j'},$ and either $i \neq i'$ or $j \neq j') = 0$.

A statistic is any measureable function of the observations (x_{ij}), and will be generally denoted $t(x_{ij})$. In some of the applications the statistics used in this report will actually be vector valued, and consequently all of the proofs are given in the case where $t(x_{ij})$ is a vector. In writing that a statistic is equal to a certain value or less than or equal to a certain value this should be interpreted in terms of the components of the statistics. In defining certain statistics subsequently we will not include any values if the event in (2) (this event will be denoted by E) above occurs. Since the event E has probability zero, any statistic which is not defined explicity when this event occurs may be defined arbitrarily, without effect on the distribution of the statistic.

The rank of x_{ij}(in the i^{th} sample) is a statistic and equals the number of observations in the i^{th} sample

whose values are less than or equal to that of x_{ij}.

A rank order statistic is a statistic which is a function only of the ranks (within samples) of the observations, and will be denoted by R.

A symmetric statistic is a symmetric function of the observations within samples, and will be denoted by S.

The "t" statistic for two samples is an example of a symmetric statistic. If we had several samples and from each of them formed a run statistic [Levene(1952)], then any function of the several run statistics would be a rank order statistic. In particular for one sample the total number of runs up and down is a rank order statistic.

Lemmas: In the following lemmas we shall need the sets A_j, defined as follows: (x_i) belongs to A_j if

$$x_{j_1} < x_{j_2} < \dots < x_{j_n} \text{ where } (j_1, j_2, \dots, j_n)$$

is a permutation of the first n integers. Then it is clear that there are n! sets A_j, and that they are all disjoint. Further the sum of the sets A_j and the set corresponding to the event E[(see (2) above] are disjoint and together form a n-dimensional Euclidean space E_n.

Lemma 1: $P[R=r \mid (x_i) \text{ belongs to } A_j]$ is one or zero depending

on if there is at least one point (x'_i) in A_j such that $R(x'_i) = r$ or not.

<u>Proof:</u> If (x_i) and (x'_i) are two points in A_j then it is clear that x_i and x'_i have the same ranks for each $i = 1, 2, \ldots, n$. But a rank order statistic is a function of the ranks only, and therefore any rank order statistic can assume only one value in the set A_j. From this the lemma follows.

Lemma II: Under the assumption that (x_i) is a random sample from a continuous distribution:

$$P[(x_i) \text{ belongs to } A_j] = 1/n!$$

<u>Proof:</u> Let $F(x)$ be the distribution function, then:

$$P[(x_i) \text{ belongs to } A_j] = \int \cdots \int_{A_j} \prod_{i=1}^{n} dF(x_i)$$

$$[\text{Let } F(x_i) = y_i]$$

$$= \int \cdots \int_{0 < y_{j_1} < y_{j_2} < \cdots < y_{j_n} < 1} \prod_{i=1}^{n} dy_i$$

$$= 1/n! \quad \text{Q. E. D.}$$

Lemma III: Under the assumption of sampling from a distribution with a continuous distribution function:

$$P[S \leq s \mid (x_i) \text{ belongs to } A_j] = P(S \leq s).$$

<u>Proof</u>: Let s_j be the set of points which are in A_j and such that $S \leq s$. Then it is clear that:

$$P(s_j) = P(s_{j'})$$

Hence using Lemma II we have that:

$$P[S \leq s \mid (x_i) \text{ belongs to } A_j]$$

$$= \frac{P[S \leq s \text{ and } (x_i) \text{ belongs to } A_j]}{P[(x_i) \text{ belongs to } A_j]}$$

$$= P(s_j)\, n!$$

$$= \frac{1}{n!} \sum_j P(s_j) n!$$

$$= P(S \leq s) \qquad \text{Q. E. D.}$$

<u>Fundamental Result</u>: In this section we state and prove a theorem which is immediately useful for the problem about the Fundamental Random Process mentioned in the introduction.

Theorem I: In samples of fixed size of random variables (x_i) which are independently and identically distributed any rank order statistic and any symmetric statistic are independently distributed.

Proof: Clearly any particular rank order statistic R can take on at most n! values and therefore has a discrete distribution. And thus we perform the following analysis to show the independence of R and S:

$P(R = r \text{ and } S \leq s)$

$= P[R = r \text{ and } S \leq s \text{ and } (x_i) \text{ belongs to } E]$

$+ P[R = r \text{ and } S \leq s \text{ and } (x_i) \text{ does not belong to } E]$

$= P[R = r \text{ and } S \leq s \text{ and } (x_i) \text{ belongs to the sum of the } A_j\text{'s}]$

$= \sum_j P[R = r \text{ and } S \leq s \text{ and } (x_i) \text{ belongs } A_j]$

$= \sum_j P[S \leq s | R = r \text{ and } (x_i) \text{ belongs to } A_j]$

$P[R = r \text{ and } (x_i) \text{ belongs to } A_j]$

(Now by Lemma I)

$= \sum_j P[S \leq s | (x_i) \text{ belongs to } A_j] \cdot$

$P[R = r \text{ and } (x_i) \text{ belongs to } A_j]$

(Now by Lemma III)

$= P(S \leq s) \sum_j P[R = r \text{ and } (x_i) \text{ belongs to } A_j]$

$= P(S \leq s)P(R = r)$ Q. E. D.

<u>Applications of Theorem I</u>: (1) We wish to test the hypothesis that observations $x_1, \ldots x_n$ (made in that order) come from a distribution with median zero. The possible alternatives are that the median is greater than zero, or that the median is shifting toward larger values as the successive observations are made. To test this hypothesis it would be appropriate to use the sign test statistic [MacStewart (1941)], and the [Mann (1945)] statistic for trend. It is clear that large values of either of these statistics should be used to reject the null hypothesis. The sign test statistic is a symmetric statistic and the Mann statistic is a rank order statistic so clearly Theorem I is applicable.

In order to make a test of significance at the α level, using these two statistics, there are available two standard techniques. We may use the χ^2 technique for combining tests of significance. An extensive analysis of this method can be found in [Wallis (1942)] see also [Fisher (1925)]. The other technique is to choose levels of significance α_1 and α_2 for the two tests based on the two statistics. Then the probability of rejecting the null hypothesis (when true) by at least one of the test is $1 - (1-\alpha_1)(1-\alpha_2) = \alpha_1 + \alpha_2 - \alpha_1\alpha_2$. Since α_1 and α_2 are at our disposal we can usually pick a desired combination of them so that $\alpha_1 + \alpha_2 - \alpha_1\alpha_2 = \alpha$. These techniques for combining independent tests can

of course be used in all of the remaining examples. When there are more than two independent tests the χ^2 method is still applicable, and in the other technique we pick levels of significance $\alpha_1, \alpha_2, \ldots, \alpha_e$ (say there are e independent tests). Then the probability that the null hypothesis will be rejected (when true) is $1 - \prod_{i=1}^{e}(1-\alpha_i)$ and we can usually choose the α_i such that $\alpha = 1 - \prod_{i=1}^{e}(1-\alpha_i)$.

(2) We wish to test the hypothesis that x(t) is an observation from a fundamental random process, and the alternatives are that $x(i\delta) - x[(i-1)\delta]$ $(i=1,\ldots,n)$ are either not normally distributed or are autocorrelated. The alternative of not being normally distributed may be tested by using a χ^2 goodness-of-fit test on the variables $x(i\delta) - x[(i-1)\delta]$, after estimating the common variance of the random variables $x(i\delta) - x[(i-1)\delta]$. This is a symmetric statistic. The autocorrelation may be tested by using the rank order autocorrelation coefficients [see Noether (1950)]. These are clearly rank order statistics. Thus Theorem I assures us that under the null hypothesis the χ^2 statistic is independent of the rank order autocorrelation coefficients.

<u>Several Samples:</u> In this section we shall state and prove a theorem analagous to Theorem I, applicable to several samples.

Theorem II: Assume that the set of random variables $(x_{ij})(j=1,\ldots,n_i, i=1,\ldots,N)$ are independently distributed and identically distributed for fixed i (ith sample) then any rank order statistic of the N samples is independent of any symmetric statistic of the N samples.

Proof: The sample space is again the Euclidean $n(n = \sum_{i=1}^{N} n_i)$ space E_n. E_n can be written as the product of N Euclidean spaces of dimension n_i that is

$$E = \prod_{i=1}^{N} E_{n_i} .$$

Further each E_{n_i} can be decomposed into sets E^i, A_j^i, $(i=1,\ldots,N, j=1,\ldots,n_i!)$ the sets E^i and A_j^i, being analogous to the sets E and A_j of the lemmas. Thus

$$E_n = \prod_{i=1}^{N}\left(\sum_{i=1} A_j^i + E^i\right) ,$$

which can be written as;

$$\sum_{k=1}^{\prod n_i!} B_k + E .$$

Here E is the union of all of those sets in E_n, as expressed in the above product which contain at least one E^i. That is E is the set of those points in E_n with at least two coordinates equal within a sample. B_k are the $\prod_{i=1}^{N} n_i!$ disjoint sets of the form $\prod_{i=1}^{N} A_{j_i}^{\ i}$, that is no ties within samples.

It is clear by the continuity assumptions that:

$$P(E) = 0.$$

Then:

$$P(S \leq s \text{ and } R = r)$$

$$= \sum_k P[S \leq s \text{ and } R = r \text{ and } (x_{ij}) \text{ belongs to } B_k]$$

$$+ P[S \leq s \text{ and } R = r \text{ and } (x_{ij}) \text{ belongs to } E]$$

$$= \sum_k P[S \leq s \text{ and } R = r \text{ and } (x_{ij}) \text{ belongs to } B_k]$$

$$= \sum_k P[S \leq s \mid (x_{ij}) \text{ belongs to } B_k \text{ and } R = r]$$

$$P[(x_{ij}) \text{ belongs to } B_k \text{ and } R = r]$$

$$= P(S \leq s)\ P(R = r) \quad \text{Q. E. D.}$$

Applications of Theorem: (1) Assume we have an observation on a stochastic process and wish to test the hypothesis that the increments are independent and identically distributed. The alternatives we wish to guard against are such things as autocorrelation and that $x(t)$ has some marked change at some time t_m[say $x(t)$ is displacement in the horizontal of a shell and t_m is the approximate time at which the shell stops rising and starts descending]. Let $y_i = x(t_m+i\delta) - x[t_m+(i-1)\delta]$ and $i = -(n_1-1),\ldots,n_2$. Then the null hypothesis implies that $y'_i = y_i[i=-(n_1-1,\ldots,0]$ and $y_i(i=1,\ldots,n_2)$ from two samples of identical and independent

observations. We might use the [Kolmogoroff (1941)] goodness-of-fit test to see that the observations from both samples come from the same distribution. Let the statistic used be called S. We can test autocorrelation by computing statistics R' and R which are the rank order autocorrelation coefficients from the first and second samples respectively. The vector (R',R) corresponds to the quantity R of the theorem, and since S is symmetric we have under the null hypothesis that (R',R) and S are independent. Also of course under the null hypothesis R' and R are independent.

(2) n litters of animals (n_i in the ith litter) are to be used for an experiment. Preliminary to the experiment we wish to test that average weights of the litters are the same. The null hypothesis is that all of the weights are independent, normal with same mean and variance. As alternatives we shall be concerned if there is a birth order effect within litters or that the litters have different average weights. The F test of analysis of variance will test the homogeneous nature of the means, and the totaled number of runs up and down in the ith litter R_i will be a test of birth order. It is clear that the F is a symmetric and the vector (R_i) is a rank order statistic and therefore the theorem is applicable.

Power: It is hard to describe the power functions of tests based on several statistics, when the several statistics do not remain independent. When the tests are independent, even under the alternative hypotheses it is possible to find the power function for each test separately and then of course it is possible to find the power function in most cases (it is trivial if we combine the test results by the second method discussed above).

If the random variables do remain independent and identically distributed, then even though the parametric portion of the null hypothesis may be false, the statistics used to test the "random" and "parametric" parts of the null hypothesis do remain independent.

Generalizations: The results of this report may be extended to the case where the observations instead of being real-valued are simply points in any abstract space. In this case the definition of symmetric statistic is not affected. However, rank order statistics must be redefined. One way of doing this is to form a real valued functional of the observations, and define the rank order statistics in terms of the values of this functional. If this is done in such a manner that the set E still remain of measure zero then the theorems of the report will be applicable to this more general system.

BIBLIOGRAPHY

R. A. Fisher

1925 Statistical methods for research workers.

Oliver and Boyd, Edinburgh,

Howard Levene

1952 On the power function of tests of randomness based on runs up and down.

Annals of Mathematical Statistics 23, 34-56.

W. MacStewart

1941 A note on the power function of the sign test.

Annals of Mathematical Statistics 12, 236-239.

Henry B. Mann

1945 On a test for randomness based on signs of differences.

Annals of Mathematical Statistics 16, 193-199.

1951 Introduction to the theory of stochastic processes depending on a continuous parameter.

National Bureau of Standards Report, 1293.

Gottfried E. Noether

1950 Asymptotic properties of the Wald-Wolfowitz tests of randomness.

Annals of Mathematical Statistics 21, 231-246.

W. Allen Wallis

1942 Compounding probabilities from independent significance tests.

Econometrika 10, 229-248.

THE NATIONAL BUREAU OF STANDARDS

Functions and Activities

The functions of the National Bureau of Standards are set forth in the Act of Congress, March 3, 1901, as amended by Congress in Public Law 619, 1950. These include the development and maintenance of the national standards of measurement and the provision of means and methods for making measurements consistent with these standards; the determination of physical constants and properties of materials; the development of methods and instruments for testing materials, devices, and structures; advisory services to Government Agencies on scientific and technical problems; invention and development of devices to serve special needs of the Government; and the development of standard practices, codes, and specifications. The work includes basic and applied research, development, engineering, instrumentation, testing, evaluation, calibration services and various consultation and information services. A major portion of the Bureau's work is performed for other Government Agencies, particularly the Department of Defense and the Atomic Energy Commission. The scope of activities is suggested by the listing of divisions and sections on the inside of the front cover.

Reports and Publications

The results of the Bureau's work take the form of either actual equipment and devices or published papers and reports. Reports are issued to the sponsoring agency of a particular project or program. Published papers appear either in the Bureau's own series of publications or in the journals of professional and scientific societies. The Bureau itself publishes three monthly periodicals, available from the Government Printing Office: The Journal of Research, which presents complete papers reporting technical investigations; the Technical News Bulletin, which presents summary and preliminary reports on work in progress; and Basic Radio Propagation Predictions, which provides data for determining the best frequencies to use for radio communications throughout the world. There are also five series of nonperiodical publications: The Applied Mathematics Series, Circulars, Handbooks, Building Materials and Structures Reports, and Miscellaneous Publications.

Information on the Bureau's publications can be found in NBS Circular 460, Publications of the National Bureau of Standards ($1.00). Information on calibration services and fees can be found in NBS Circular 483, Testing by the National Bureau of Standards (25 cents). Both are available from the Government Printing Office. Inquiries regarding the Bureau's reports and publications should be addressed to the Office of Scientific Publications, National Bureau of Standards, Washington 25, D. C.

NBS