

# NATIONAL BUREAU OF STANDARDS REPORT

1626

HOW TO DECIDE OBJECTIVELY WHETHER  
AN OUTLYING OBSERVATION SHOULD BE REJECTED

by

Frank Proschan



U. S. DEPARTMENT OF COMMERCE  
NATIONAL BUREAU OF STANDARDS



## 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. Enamelled 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.

3011-60 0002

25 April 1952

1626

HOW TO DECIDE OBJECTIVELY WHETHER  
AN OUTLYING OBSERVATION SHOULD BE REJECTED

by

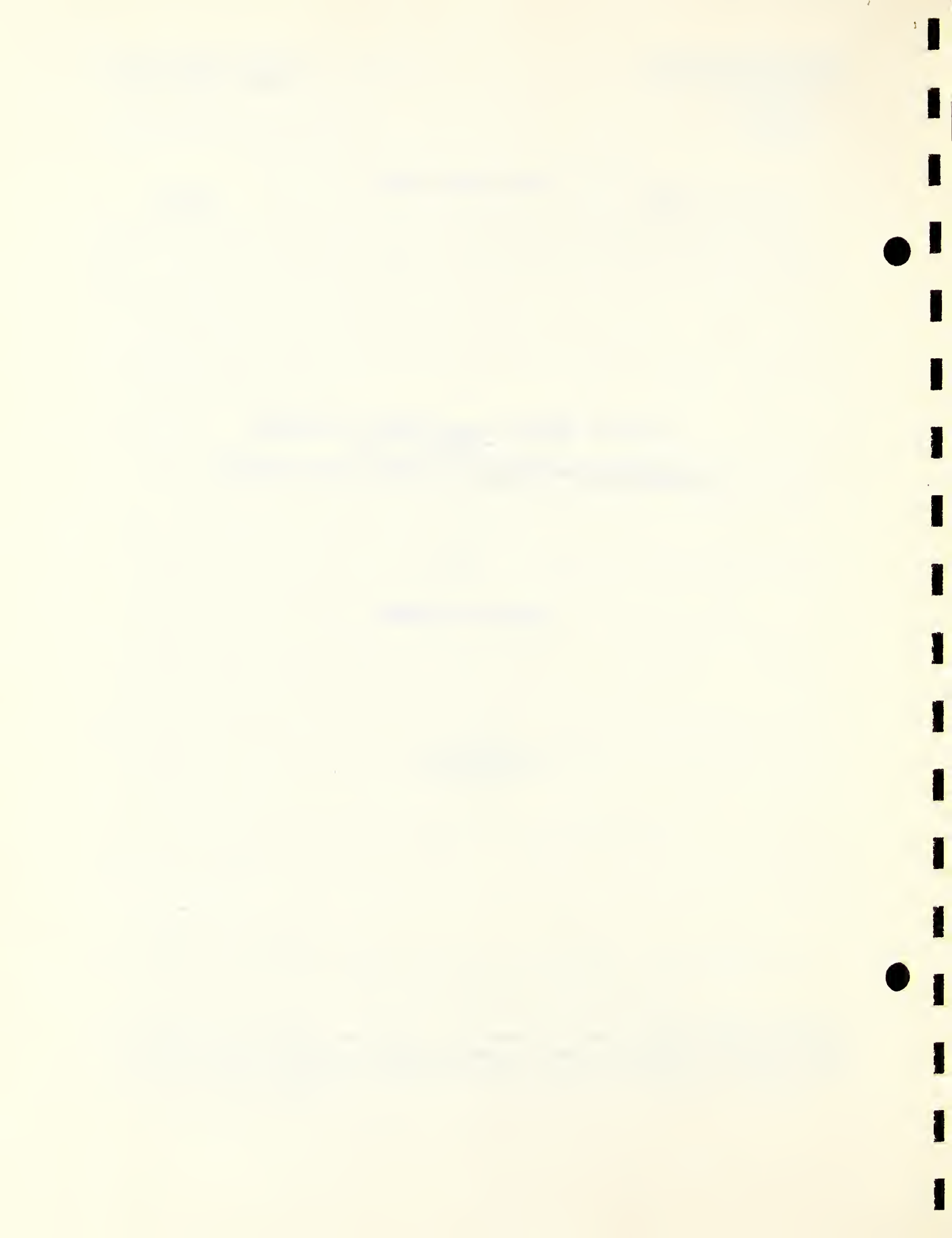
Frank Proschan



This report is issued for  
in any form, either in who  
from the Office of the Dir

Approved for public release by the  
Director of the National Institute of  
Standards and Technology (NIST)  
on October 9, 2015

reprinting, or reproduction  
ession in writing is obtained  
ington 25, D. C.



## FORWORD

A perennial problem vexing the experimenter is that of rejection of suspected data. For one hundred years attempts at the solution of this problem have been advanced, most of them to be themselves rejected as suspect. Fortunately modern statistical theory has proposed useful, reliable methods for objectively rejecting deviant values. However the solution is far from complete at present.

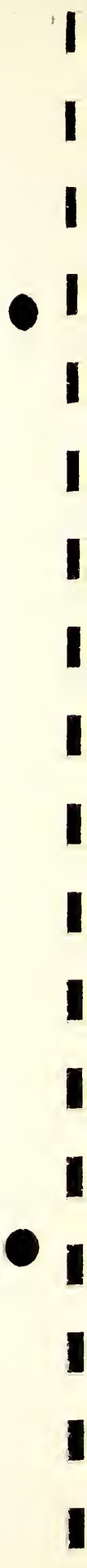
This report describes for the experimenter two of the modern statistical tests available for possible rejection of outlying observations. These two methods have been selected because they apply in a majority of the actually occurring situations, and because they are so easy to apply. The report was originally motivated by Mr. Proschan's consultative work in the Ordnance Development Division, National Bureau of Standards. However it is applicable more generally to the scientific and engineering work at the National Bureau of Standards.

It should be borne in mind that this report, although intended for practical every-day use in our laboratories, is not the final word on the subject of rejection of suspected observations.

J. H. Curtiss  
Chief, National Applied  
Mathematics Laboratories

A. V. Astin  
Acting Director  
National Bureau of Standards

[The text in this block is extremely faint and illegible, appearing as a series of horizontal lines across the page.]



HOW TO DECIDE OBJECTIVELY WHETHER  
AN OUTLYING OBSERVATION SHOULD BE REJECTED

by

Frank Proschan

1.1 PROBLEM: Here is a very common problem facing experimenters at the National Bureau of Standards. The typical scientist, call him X. Perry Menter, makes a number (say five) of repeated measurements of some unknown quantity. One of the values (say the largest) is so far removed from the other four that he suspects that it may be in error. However, Perry has no specific knowledge that a mistake actually did occur. Let us assume, too, that he has no previous data from which to estimate the precision of measurement. How can he decide, from the values themselves, whether the suspected value is in error or not?

The answer seems clear: He should consider the suspected value as in error when it seems too far from the other four values. But how can he judge when it is "too far from the other four values"?

1.2 LOGICAL APPROACH: Here is a simple, logical, objective criterion. Suppose Perry could somehow make millions of sets of five observations each. Suppose, too, that he could guarantee that none of these observations had any mistakes. Call a typical set  $x_1, x_2, x_3, x_4, x_5$ , where the  $x$ 's are arranged in order of size, so that  $x_1 \leq x_2 \leq x_3 \leq x_4 \leq x_5$ . Now a logical

REPORT PAGE

[The text in this section is extremely faint and illegible. It appears to be a multi-paragraph report or document.]



measure of the distance between the largest value and the other four values is

$$r_1 = \frac{x_5 - x_4}{x_5 - x_1} \quad (1)$$

i.e., the proportion of the total range, that the distance between the suspected value and its adjacent value, is.

Now Perry records with what frequency, among his millions of sets of five values each, different values of  $r_1$  occur. He finds that a value of  $r_1$  larger than .780 occurs one time in one hundred. He then reasons this way:

"I have found that among sets of five observations each (containing no mistakes) a value of  $r_1$  larger than .780 is quite unlikely (occurs only once in one hundred). If now, in my future experiments I get a set of five observations for which  $r_1$  is larger than .780 I will conclude that my largest observation is in error."

1.3 CONFIDENCE IN THE TEST: This seems reasonable. But what confidence can Perry have in such a procedure? How often will he consider as mistaken a perfectly good observation? How often will he consider acceptable an incorrect observation?

Clearly from the way he derived the test, he will classify a perfectly good largest observation as mistaken once among one hundred sets of five each, on the average. But there is no general answer to the question of how often he will let

1950

1951

1952

1953

1954

1955

1956

1957

1958

1959

1960

1961

1962

1963

1964

1965

1966

1967

1968

1969

1970

1971

1972

1973

1974

1975

1976

1977

1978

1979

1980

1981

1982

1983

1984

1985

1986

1987

1988

1989

1990

1991

1992

1993

1994

1995

1996

1997

1998

1999

2000

2001

2002

2003

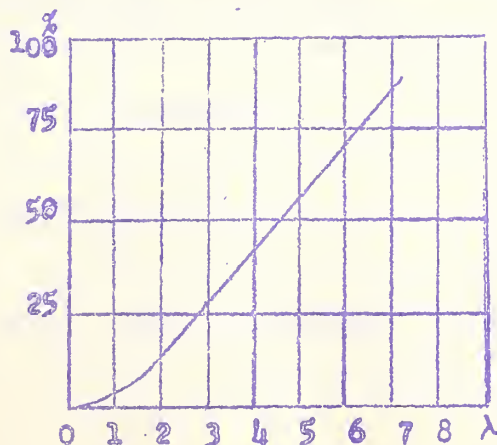
2004

2005

pass a mistaken observation. This depends on how "mistaken" the mistaken observation is. If a very large error were made, his test would tend to reject the observation almost certainly. If a very small error were made, his test would tend to reject the observation with a small probability.

The following chart [1] gives some idea of the performance of  $r_1$  in detecting mistaken observations. It is based on a sampling experiment in which samples of five from a normal population with mean  $\mu$ , standard deviation  $\sigma$  were contaminated with values drawn from a normal population with mean  $\mu + \lambda \sigma$  and standard deviation  $\sigma$ . The ordinate shows the percent discovery of contaminants (the proportion of the time the contaminating population provides an extreme value and the test discovers this value) while the abscissa shows  $\lambda$ , the magnitude of the shift (error) of the contaminator in standard deviations.

Performance of r-test



(From Dixon's article [1])

The first part of the document discusses the importance of maintaining accurate records of all transactions. It emphasizes that every entry should be supported by a valid receipt or invoice. The text also mentions the need for regular audits to ensure the integrity of the financial data.

In the second section, the author details the various methods used for data collection and analysis. This includes the use of statistical software and manual calculations. The document highlights the challenges of handling large volumes of data and the importance of using appropriate sampling techniques.

The third part of the document focuses on the results of the study. It presents a series of tables and graphs that illustrate the trends and patterns in the data. The author concludes that the findings are significant and provide valuable insights into the subject matter.

	1	2	3	4
A				
B				
C				

We said above that once in every 100 sets of values (on the average) Perry would consider as mistaken a perfectly good observation. If he were to reject this observation, and then compute the mean and standard deviation of the remaining values, these would be biased estimates. Also when a good observation is rejected, any further statistical tests of significance will become less reliable. This is the price that he must pay for improving the data in the cases where a mistaken observation is removed.

1.4 MATHEMATICAL DERIVATION: Of course, .780, the value of  $r_1$  exceeded by chance 1 percent of the time (called the 1 percent level of significance of  $r_1$ ), is not determined by actually making millions of sets of five observations each. Rather it may be calculated mathematically [2] with even greater accuracy than if millions of sets of five observations had been used. The basic assumption is that repeated measurements would follow the normal distribution.

Smallest Observation Suspected. What if Perry suspects the smallest observation in a set (of five, say)?

In this case, he computes

$$r_s = \frac{x_2 - x_1}{x_5 - x_1} \quad (2)$$

He compares it with .780, as before. Exactly the same reasoning applies.

[The text in this section is extremely faint and illegible. It appears to be a multi-paragraph document, possibly a letter or a report, with several lines of text per paragraph. The content is not discernible.]

1.5 Table I. Let us now define  $r$  as either  $r_1$  or  $r_2$  depending on whether the largest or smallest of a set is being tested. Table I at the end of the paper gives  $R$ , the value of  $r$  at various significance levels,  $\alpha$ , for sample size  $n$  from 3 to 30.

Thus for example, for  $n = 8$  and  $\alpha = .05$ , the table gives  $R = .468$ . This means that in 100 sets of 8 observations each free of mistakes, five values of  $r_1$  will be larger than .468 and 5 values of  $r_2$  will be larger than .468 (on the average).

Why are various significance levels given? The reason is that no one significance level is appropriate to all problems. For example consider these two cases:

- (a) Additional observations are not possible.
- (b) Additional observations are possible.

In case (a) for many problems it might be appropriate to compute  $r$  and test it at the 1 percent level of significance. If the observed value of  $r$  is larger than the table value for  $\alpha = 1$  percent, it might then be a good idea to exclude the responsible observation.

In case (b), for many situations a reasonable procedure might be to test  $r$  at the 5 percent level. If the sample value of  $r$  is significant at the 5 percent level, one or more additional observations would be taken. If the observation originally suspected remained outlying, it would be tested again, using the combined set of observations. This time, however, the  $r$  test would be performed at the 1 percent level

The first part of the document discusses the importance of maintaining accurate records of all transactions. It emphasizes that every entry should be supported by a valid receipt or invoice. This ensures transparency and allows for easy verification of the data.

In the second section, the author details the various methods used to collect and analyze the data. This includes both manual and automated processes. The goal is to ensure that the data is as accurate and comprehensive as possible.

The third part of the document focuses on the results of the analysis. It shows that there is a clear trend in the data, which is consistent with the initial hypothesis. This finding is significant as it provides strong evidence for the proposed model.

Finally, the document concludes with a summary of the findings and a list of recommendations for future research. It suggests that further studies should be conducted to explore the underlying causes of the observed trends.



of significance. If the outlier were significantly deviant at the 1 percent level, it would be rejected. It should be noted that among many sets tested this way, the proportion of perfectly good largest values rejected this way will be less than 1 percent. This is because the observation has a "second chance" before it is finally rejected.

1.6 SUMMARY: A set of  $n$  observations is made. No previous data are available from which to estimate the variability of a measurement. What is a rational procedure for testing whether the largest (or smallest) of the set is too deviant to be explained by the ordinary errors of measurement?

Consider the  $n$  observations in order of size,

$$x_1 \leq x_2 \leq \dots \leq x_n$$

Compute

$$r = \frac{x_n - x_{n-1}}{x_n - x_1} \quad (\text{if } x_n \text{ is suspected}) \quad (3)$$

or

$$r = \frac{x_2 - x_1}{x_n - x_1} \quad (\text{if } x_1 \text{ is suspected}).$$

Table I may be used to determine how likely it is to get as large a value of  $r$  as actually obtained, simply by chance. A procedure that might be appropriate for many problems is as follows:

(a) No additional observations possible: In this case, compare the computed  $r$  with the Table I value at the 1 percent level. If the computed value of  $r$  is larger than the

The first part of the paper is devoted to the study of the asymptotic behavior of the eigenvalues of the Laplacian on a Riemannian manifold with boundary. The second part is devoted to the study of the asymptotic behavior of the eigenvalues of the Laplacian on a Riemannian manifold with boundary.

The third part of the paper is devoted to the study of the asymptotic behavior of the eigenvalues of the Laplacian on a Riemannian manifold with boundary. The fourth part is devoted to the study of the asymptotic behavior of the eigenvalues of the Laplacian on a Riemannian manifold with boundary.

$$\lambda_1 \sim \frac{1}{4} \mu_1^2$$

$$\lambda_2 \sim \frac{1}{4} \mu_2^2$$

$$\lambda_3 \sim \frac{1}{4} \mu_3^2$$

The fifth part of the paper is devoted to the study of the asymptotic behavior of the eigenvalues of the Laplacian on a Riemannian manifold with boundary. The sixth part is devoted to the study of the asymptotic behavior of the eigenvalues of the Laplacian on a Riemannian manifold with boundary.

The seventh part of the paper is devoted to the study of the asymptotic behavior of the eigenvalues of the Laplacian on a Riemannian manifold with boundary. The eighth part is devoted to the study of the asymptotic behavior of the eigenvalues of the Laplacian on a Riemannian manifold with boundary.

Table I value, exclude the deviant observation. Otherwise, do not.

(b) Additional observations possible: In this case, compare the computed  $r$  with the table value of  $r$  at the 5 percent level. If the computed value of  $r$  is larger than the table value, take one or more additional observations (depending on convenience, cost of observations, etc.). Otherwise accept the suspected value without taking additional observations.

In in the enlarged set (containing all the original and the additional observation(s)), the previously suspected value remains outlying, compute  $r$  for the enlarged set. This time compare it with the Table I value at the 1 percent level. If the computed value exceeds the table value, exclude the outlier; otherwise do not.

1.7 EXAMPLES: 1. Anna List, chemist, determines five values for the iron content of an unknown solution by a new process: 7.42, 7.48, 7.39, 7.61, 7.44 (in percent). She suspects 7.61 as being mistaken, since it is so much larger than the other values. However she cannot determine additional values because no more samples are available. How can she decide whether to exclude 7.61 or not?

Since no previous data are available from which to compute the precision of measurement, the  $r$  test is appropriate. The first step is to arrange the five values in order of size:

Faint, illegible text covering the page, possibly bleed-through from the reverse side. The text is too light to transcribe accurately.

7.39, 7.42, 7.44, 7.48, 7.61

Then

$$r = \frac{x_5 - x_4}{x_5 - x_1} = \frac{7.61 - 7.48}{7.61 - 7.39} = \frac{.13}{.22} = .591$$

Since this is less than .780, the .01 point of  $r$  for  $n = 5$ , Anna List retains the suspected value, 7.61.

2. N. G. Neer fires projectiles from a gun under constant conditions. Here are the distances in feet (in order of size):

6801	7683
7424	7720
7502	7799
7544	

He suspects 6801 as being inconsistent with the other values. What shall he do?

He computes

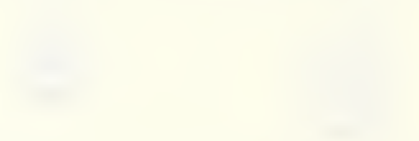
$$r = \frac{x_2 - x_1}{x_7 - x_1} = \frac{7424 - 6801}{7799 - 6801} = \frac{623}{998} = .624$$

This value lies between the .01 and the .05 points of  $r$  for  $n = 7$ . Hence N. G. Neer fires an additional round and gets a new value of 7603.

Since 6801 remains outlying in the enlarged set of eight, he computes  $r$  for this set of eight. Obviously  $r$  remains .624. This time it is larger than the .01 point of  $r$  for  $n = 8$ . Hence N. G. Neer rejects 6801 and uses only the remaining seven values.

$$x^2 - 3x - 10 = \left(x - \frac{3}{2}\right)^2 - \frac{49}{4}$$

The graph of the parabola  $y = x^2 - 3x - 10$  is shown below. The vertex is at  $\left(\frac{3}{2}, -\frac{49}{4}\right)$ . The x-intercepts are at  $x = -2$  and  $x = 5$ . The y-intercept is at  $y = -10$ .



The graph of the parabola  $y = x^2 - 3x - 10$  is shown below. The vertex is at  $\left(\frac{3}{2}, -\frac{49}{4}\right)$ . The x-intercepts are at  $x = -2$  and  $x = 5$ . The y-intercept is at  $y = -10$ .

$$x^2 - 3x - 10 = \left(x - \frac{3}{2}\right)^2 - \frac{49}{4}$$

The graph of the parabola  $y = x^2 - 3x - 10$  is shown below. The vertex is at  $\left(\frac{3}{2}, -\frac{49}{4}\right)$ . The x-intercepts are at  $x = -2$  and  $x = 5$ . The y-intercept is at  $y = -10$ .

The graph of the parabola  $y = x^2 - 3x - 10$  is shown below. The vertex is at  $\left(\frac{3}{2}, -\frac{49}{4}\right)$ . The x-intercepts are at  $x = -2$  and  $x = 5$ . The y-intercept is at  $y = -10$ .

The graph of the parabola  $y = x^2 - 3x - 10$  is shown below. The vertex is at  $\left(\frac{3}{2}, -\frac{49}{4}\right)$ . The x-intercepts are at  $x = -2$  and  $x = 5$ . The y-intercept is at  $y = -10$ .

## 2 Estimate of Measurement Variability Available.

In a great many laboratory situations, past data are available for estimating the uncertainty of a measurement. Thus if a laboratory makes routine analysis in triplicate of incoming specimens, there will be sets of triplicate values from which to estimate the precision of measurement. It is clear that where such information is available, it should be used, in deciding whether an outlier is mistaken or not. This will make the decision more reliable than if only the one set containing the suspected value is used.

2.1 u Test. The test ratio used now is

$$u = \frac{x_n - \bar{x}}{s_d} \quad (\text{if } x_n \text{ is the suspected value}) \quad (4)$$

or

$$u = \frac{\bar{x} - x_1}{s_d} \quad (\text{if } x_1 \text{ is the suspected value})$$

where

$s_d$  = standard deviation of an individual measurement, based on  $d$  degrees of freedom.

2.2 Calculating  $s_d$ . To determine  $s_d$  from a single set of measurements we would first calculate the sum of the squares of the deviations of the observations from their mean. Then we would divide by one less than the number of observations. This would give us  $s_d^2$ . Thus

$$s_d^2 = \frac{\sum_{i=1}^d (x_i - \bar{x})^2}{d-1} \quad (5)$$

... ..  
 ... ..  
 ... ..  
 ... ..  
 ... ..  
 ... ..  
 ... ..  
 ... ..  
 ... ..  
 ... ..  
 ... ..

... ..

... ..

... ..  $\frac{1}{2} \frac{d^2 y}{dx^2} + \dots = \dots$

... ..  $\frac{1}{2} \frac{d^2 y}{dx^2} + \dots = \dots$

... ..  
 ... ..

... ..

... ..  
 ... ..  
 ... ..

$$\frac{1}{2} \frac{d^2 y}{dx^2} + \dots = \dots$$



On the other hand, suppose a number of sets of observations were available:

<u>SETS</u>	<u>OBSERVATIONS</u>	<u>MEAN</u>
1	$x_{11}, x_{12}, \dots, x_{1n_1}$	$\bar{x}_1$
2	$x_{21}, x_{22}, \dots, x_{2n_2}$	$\bar{x}_2$
...	.....	...
k	$x_{k1}, x_{k2}, \dots, x_{kn_k}$	$\bar{x}_k$

Now we could calculate  $s_d^2$  from

$$s_d^2 = \frac{\sum_{i=1}^{n_1} (x_{1i} - \bar{x}_1)^2 + \sum_{i=1}^{n_2} (x_{2i} - \bar{x}_2)^2 + \dots + \sum_{i=1}^{n_k} (x_{ki} - \bar{x}_k)^2}{(n_1-1) + (n_2-1) + \dots + (n_k-1)} \quad (6)$$

$\delta$ , the number of degrees of freedom for estimating the uncertainty of measurement, is  $(n_1-1) + (n_2-1) + \dots + (n_k-1)$ .

2.3 EXAMPLE: An example will make the whole procedure clear:

Dr. Kem Mist has just run a set of three determinations on the yield of an ore-refining process. His values are:

39.35, 39.30, 39.00, with mean,  $\bar{x} = 39.22$ .

He wishes to test whether the "39.00" is unusually deviant. He has past data to estimate the precision of measurement:

OPERATIONS

$$\frac{1}{x^2} = x^{-2} \Rightarrow -2x^{-3} = -\frac{2}{x^3}$$

$$\frac{1}{x^3} = x^{-3} \Rightarrow -3x^{-4} = -\frac{3}{x^4}$$

$$\dots$$

$$\frac{1}{x^n} = x^{-n} \Rightarrow -nx^{-n-1} = -\frac{n}{x^{n+1}}$$

$$\text{more } \frac{d}{dx} x^{-n} = -nx^{-n-1}$$

$$\frac{d}{dx} \left( \frac{1}{x^2} + \frac{1}{x^3} + \dots + \frac{1}{x^n} \right) = -\frac{2}{x^3} - \frac{3}{x^4} - \dots - \frac{n}{x^{n+1}}$$

... (faint text describing the derivation of the power rule for negative exponents)

... (faint text describing the derivation of the power rule for negative exponents)

... (faint text describing the derivation of the power rule for negative exponents)

... (faint text describing the derivation of the power rule for negative exponents)

Set	Analysis		
	<u>1</u>	<u>2</u>	<u>3</u>
1	36.51	36.57	36.70
2	30.27	30.35	30.19
3	35.00	35.53	35.36
4	43.51	43.65	43.65
5	51.06	51.17	51.00
6	48.03	38.19	48.31
7	39.27	39.51	39.36
8	33.46	33.21	33.28

Although we could calculate  $s_d^2$  from (6), it is generally more convenient (especially with a computing machine available) to use (7):

$$s_d^2 = \left[ \sum_{i=1}^{n_1} x_{1i}^2 - \frac{\left( \sum_{i=1}^{n_1} x_{1i} \right)^2}{n_1} + \sum_{i=1}^{n_2} x_{2i}^2 - \frac{\left( \sum_{i=1}^{n_2} x_{2i} \right)^2}{n_2} + \dots + \sum_{i=1}^{n_k} x_{ki}^2 - \frac{\left( \sum_{i=1}^{n_k} x_{ki} \right)^2}{n_k} \right] / [(n_1-1) + (n_2-1) + \dots + (n_k-1)] \quad (7)$$

It can easily be shown that (7) is algebraically equivalent to (6); thus substituting values into (7) yields:

$$s_d^2 = \left[ 36.51^2 + 36.57^2 + 36.70^2 - \frac{(36.51+36.57+36.70)^2}{3} + 30.27^2 + 30.35^2 + 30.19^2 - \frac{(30.27+30.35+30.19)^2}{3} + \dots + 33.46^2 + 33.21^2 + 33.28^2 - \frac{(33.46+33.21+33.28)^2}{3} \right] / [(3-1) + (3-1) + \dots + (3-1)] = \frac{3082}{16} = 0.19263$$

	1	2	3
1975	1975	1975	1
1976	1976	1976	1
1977	1977	1977	1
1978	1978	1978	1
1979	1979	1979	1
1980	1980	1980	1
1981	1981	1981	1
1982	1982	1982	1
1983	1983	1983	1

... (faint text) ...

$$\frac{1}{2} \left( \frac{1}{2} \frac{d^2}{dt^2} + \dots \right) = \dots$$

$$\dots = \frac{1}{2} \frac{d^2}{dt^2} \dots$$

... (faint text) ...

$$\dots = \dots$$

$$\dots = \dots$$

$$\dots = \dots$$

Hence  $s_d = .139$

The u-ratio defined by (4) gives:

$$u = \frac{\bar{x} - x_1}{s_d} = \frac{39.22 - 39.00}{.139} = 1.58$$

He now uses Table II which gives the 5 percent and 1 percent levels of u for various values of n and d. n is the size of the sample which contains the suspected value, while d is the number of degrees of freedom on which  $s_d$  is based. In the present case n = 3 and d = 16.

The observed value of u, 1.58, is considerably less than the table value of u at the 5 percent level, 1.90. Hence he concludes that the suspected value 39.00 is not significantly outlying. In other words the deviation of 39.00 from the mean of the set of three measurements is easily explainable in terms of the precision of the measurement process. Hence, 39.00 is accepted into the fold of good measurements.

2.4 TABLE 2. When past data are available, the u-ratio may be computed and Table 2 used just as the r-ratio and Table I were used for the case where no past data were available. The procedure outlined in paragraph 1.5 for the two cases (a) and (b) may be followed just as before (using u and Table 2 instead of r and Table I).

III. Cautions and Comments. a. Obviously, if the experimenter knows that a mistake has occurred he should reject the observation. The tests of this paper are used only if he doesn't know that a mistake has occurred.

1971 - 1972

Annual Report of the Board of Directors

$$C = \frac{a + b \cdot x}{c + d \cdot x}$$

The following table shows the results of the regression analysis for the period 1971-1972. The dependent variable is the ratio of the number of employees to the number of full-time employees. The independent variable is the number of full-time employees. The regression equation is  $C = \frac{a + b \cdot x}{c + d \cdot x}$ . The results are as follows:

Dependent Variable: Ratio of employees to full-time employees  
Independent Variable: Number of full-time employees  
Regression Equation:  $C = \frac{a + b \cdot x}{c + d \cdot x}$   
Coefficients:  $a = 0.12$ ,  $b = 0.0001$ ,  $c = 0.0001$ ,  $d = 0.0001$   
R-squared: 0.95  
F-statistic: 100.0  
t-statistic: 10.0  
p-value: 0.0001

The results indicate a strong positive correlation between the number of full-time employees and the ratio of employees to full-time employees. The regression equation is highly significant, as indicated by the high R-squared value and the low p-value.

b. If the experimenter uses this technique for a certain routine type of measurement, he should apply it, implicitly or explicitly, every time he makes that type of measurement. After several explicit applications of this technique, he will probably be able to perform the  $r$ (or  $u$ ) test in all but the most doubtful cases without actually explicitly doing the arithmetic, since he will have the critical value of  $r$ (or  $u$ ) in mind. He should not, however, reject outliers by the  $r$  test in some cases, and accept others just as badly deviating, simply because he did not apply the test in these latter cases.

c. Both the  $r$  and  $u$  tests are based on the assumption that repeated measurements of the same unknown follow the normal frequency distribution. If, in actual practice, the distribution of repeated measurements is markedly different from the normal curve, then the use of these tests will lead to different risks than originally intended.

d. The use of the .01 and .05 points is arbitrary. The individual experimenter should use whatever levels of significance are most appropriate.

e. Suppose the type of measurement is such that the suspected value is practically always the smallest in the set (for example, a chemical analysis where some of the material may be washed out), or practically always the largest. Then as stated above, 1 percent of the time a perfectly good

The first part of the report deals with the general situation of the country and the progress of the work done during the year. It is followed by a detailed account of the various projects and schemes which have been undertaken during the year. The report concludes with a summary of the work done and a statement of the progress made.

The second part of the report deals with the financial position of the organization. It gives a detailed account of the income and expenditure for the year and shows how the funds have been applied to the various projects and schemes. It also shows the balance of the funds at the end of the year.

The third part of the report deals with the personnel of the organization. It gives a list of the staff and shows the work done by each of them during the year. It also shows the progress made in the recruitment of new staff and the training of existing staff.



observation will be rejected in case (a) of paragraph 1.5. Suppose however the type of measurement is such that the suspected value may be either the largest or the smallest. In this case about 2 percent of the time a perfectly good observation will be rejected. The appropriate tabular point should be selected with this in mind.

f. Other tests for rejection of suspected values are available [1]. However, the r and u tests have been selected because of their ease of application.

g. The whole question of rejection of suspected values is a difficult one, and has given rise to quite a bit of discussion. [See references.] If the scientist or engineer at the National Bureau of Standards is in doubt about any assumption, procedure, or conclusion involved in a specific practical problem, we suggest that he consult the Statistical Engineering Laboratory (11.3).

1. The first part of the document is a letter from the Secretary of the State to the Governor, dated 18th March 1871. It contains a report on the progress of the work done during the year 1870. The letter is signed by the Secretary, and is addressed to the Governor.

2. The second part of the document is a report on the work done during the year 1870. It is written by the Secretary, and is addressed to the Governor. The report contains a detailed account of the work done, and is signed by the Secretary.

3. The third part of the document is a report on the work done during the year 1870. It is written by the Secretary, and is addressed to the Governor. The report contains a detailed account of the work done, and is signed by the Secretary.

4. The fourth part of the document is a report on the work done during the year 1870. It is written by the Secretary, and is addressed to the Governor. The report contains a detailed account of the work done, and is signed by the Secretary.

5. The fifth part of the document is a report on the work done during the year 1870. It is written by the Secretary, and is addressed to the Governor. The report contains a detailed account of the work done, and is signed by the Secretary.

## REFERENCES

- [1] Dixon, W. J., Analysis of extreme values, Annals of Math. Stat. Vol. 21, No. 4, Dec. 1950.
- [2] Dixon, W. J., Ratios involving extreme values, Annals of Math. Stat. Vol. 22, No. 1, 1951.
- [3] Grubbs, F. E., Sample criteria for testing outlying observations, Annals of Math. Stat. Vol. 21, No. 1, March 1950.
- [4] Nair, K. R., The distribution of the extreme deviate from the sample mean and its studentized form, Biometrika, Vol. 35, Parts I and II, May 1948.
- [5] Pearson, E. S. and Sekar, C. C., The efficiency of statistical tools and a criterion for the rejection of outlying observations, Biometrika, Vol. XXVIII, Parts III and IV, Dec. 1936.
- [6] Rider, P. R., Criteria for rejection of observations, Washington University Studies, -New Series, Science and Technology - No. 8, Oct. 1933.

MEMORANDUM

TO : SAC, [illegible]

FROM : [illegible]

SUBJECT: [illegible]

[illegible]

[illegible]

[illegible]

TABLE I

$$\Pr(r > R) = \alpha$$

$n \backslash \alpha$	.005	.01	.02	.05	.10	.20	.30	.40	.50	.60	.70	.80	.90	.95
3	.994	.988	.976	.941	.886	.781	.684	.591	.500	.409	.316	.219	.114	.059
4	.926	.889	.846	.765	.679	.560	.471	.394	.324	.257	.193	.130	.065	.033
5	.821	.780	.729	.642	.557	.451	.373	.308	.250	.196	.146	.097	.048	.023
6	.740	.698	.644	.560	.482	.386	.318	.261	.210	.164	.121	.079	.038	.018
7	.680	.637	.586	.507	.434	.344	.281	.230	.184	.143	.105	.068	.032	.016
8	.634	.590	.543	.468	.399	.314	.255	.208	.166	.128	.094	.060	.029	.014
9	.598	.555	.510	.437	.370	.290	.234	.191	.152	.118	.086	.055	.026	.013
10	.568	.527	.483	.412	.349	.273	.219	.178	.142	.110	.080	.051	.025	.012
11	.542	.502	.460	.392	.332	.259	.208	.168	.133	.103	.074	.048	.023	.011
12	.522	.482	.441	.376	.318	.247	.197	.160	.126	.097	.070	.045	.022	.011
13	.503	.465	.425	.361	.305	.237	.188	.153	.120	.092	.067	.043	.021	.010
14	.488	.450	.411	.349	.294	.228	.181	.147	.115	.088	.064	.041	.020	.010
15	.475	.438	.399	.338	.285	.220	.175	.141	.111	.085	.062	.040	.019	.010
16	.463	.426	.388	.329	.277	.213	.169	.136	.107	.082	.060	.039	.019	.009
17	.452	.416	.379	.320	.269	.207	.165	.132	.104	.080	.058	.038	.018	.009
18	.442	.407	.370	.313	.263	.202	.160	.128	.101	.078	.056	.036	.018	.009
19	.433	.398	.363	.306	.258	.197	.157	.125	.098	.076	.055	.036	.017	.008
20	.425	.391	.356	.300	.252	.193	.153	.122	.096	.074	.053	.035	.017	.008
21	.418	.384	.350	.295	.247	.189	.150	.119	.094	.072	.052	.034	.016	.008
22	.411	.378	.344	.290	.242	.185	.147	.117	.092	.071	.051	.033	.016	.008
23	.404	.372	.338	.285	.238	.182	.144	.115	.090	.069	.050	.033	.016	.008
24	.399	.367	.333	.281	.234	.179	.142	.113	.089	.068	.049	.032	.016	.008
25	.393	.362	.329	.277	.230	.176	.139	.111	.088	.067	.048	.032	.015	.008
26	.388	.357	.324	.273	.227	.173	.137	.109	.086	.066	.047	.031	.015	.007
27	.384	.353	.320	.269	.224	.171	.135	.108	.085	.065	.047	.031	.015	.007
28	.380	.349	.316	.266	.220	.168	.133	.106	.084	.064	.046	.030	.015	.007
29	.376	.345	.312	.263	.218	.166	.131	.105	.083	.063	.046	.030	.014	.007
30	.372	.341	.309	.260	.215	.164	.130	.103	.082	.062	.045	.029	.014	.007



TABLE II

Upper per cent points of the studentized extreme deviate  
 $(x_n - \bar{x})/s_d$  or  $(\bar{x} - x_1)/s_d$

n	$\alpha = .05$							$\alpha = .01$						
	3	4	5	6	7	8	9	3	4	5	6	7	8	9
10	2.02	2.29	2.49	2.63	2.75	2.85	2.93	2.76	3.05	3.25	3.39	3.50	3.59	3.67
11	1.99	2.26	2.44	2.58	2.70	2.79	2.87	2.71	3.00	3.19	3.33	3.44	3.53	3.61
12	1.97	2.22	2.40	2.54	2.65	2.75	2.83	2.67	2.95	3.14	3.28	3.39	3.48	3.55
13	1.95	2.20	2.38	2.51	2.62	2.71	2.79	2.63	2.91	3.10	3.24	3.34	3.43	3.51
14	1.93	2.18	2.35	2.48	2.59	2.68	2.76	2.60	2.87	3.06	3.20	3.30	3.39	3.47
15	1.92	2.16	2.33	2.46	2.56	2.65	2.73	2.57	2.84	3.02	3.16	3.27	3.35	3.43
16	1.90	2.14	2.31	2.44	2.54	2.63	2.70	2.55	2.81	3.00	3.13	3.24	3.32	3.39
17	1.89	2.13	2.30	2.42	2.52	2.61	2.68	2.52	2.79	2.97	3.10	3.21	3.29	3.36
18	1.88	2.12	2.28	2.41	2.51	2.59	2.66	2.50	2.77	2.95	3.08	3.18	3.27	3.34
19	1.87	2.11	2.27	2.39	2.49	2.58	2.65	2.49	2.75	2.92	3.06	3.16	3.24	3.31
20	1.87	2.10	2.26	2.38	2.48	2.56	2.63	2.47	2.73	2.91	3.04	3.14	3.22	3.29
24	1.84	2.07	2.23	2.35	2.44	2.52	2.59	2.43	2.68	2.85	2.97	3.07	3.15	3.22
30	1.82	2.04	2.20	2.31	2.40	2.48	2.55	2.38	2.62	2.79	2.91	3.01	3.08	3.15
40	1.80	2.02	2.17	2.28	2.37	2.44	2.51	2.34	2.57	2.73	2.85	2.94	3.02	3.08
60	1.78	1.99	2.14	2.25	2.33	2.41	2.47	2.30	2.52	2.68	2.79	2.88	2.95	3.01
120	1.76	1.97	2.11	2.21	2.30	2.37	2.43	2.25	2.48	2.62	2.73	2.82	2.89	2.95
$\infty$	1.74	1.94	2.08	2.18	2.27	2.33	2.39	2.22	2.43	2.57	2.68	2.76	2.83	2.88

1. Introduction

2. Methodology

3. Results

4. Discussion

5. Conclusion

6. References



# THE NATIONAL BUREAU OF STANDARDS

## *Functions and Activities*

The National Bureau of Standards is the principal agency of the Federal Government for fundamental and applied research in physics, mathematics, chemistry, and engineering. Its activities range from the determination of physical constants and properties of materials, the development and maintenance of the national standards of measurement in the physical sciences, and the development of methods and instruments of measurement, to the development of special devices for the military and civilian agencies of the Government. The work includes basic and applied research, development, engineering, instrumentation, testing, evaluation, calibration services, and various scientific and technical advisory services. A major portion of the NBS work is performed for other government agencies, particularly the Department of Defense and the Atomic Energy Commission. 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. The scope of activities is suggested in 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.

