193

# Decision-Theoretic Evaluation of Some Non-Parametric Methods*

by

Herman Rubin

Department of Statistics

Division of Mathematical Sciences

Mimeograph Series No. 193

July, 1969



---

* This research was supported in part by the Office of Naval Research under Contract N00014-67-A-0226-0008. Reproduction in whole or in part is permitted for any purpose of the United States Government.

Decision-Theoretic Evaluation of Some

Non-Parametric Methods*

by

Herman Rubin

Purdue University

## 1. Introduction

In this paper we shall evaluate for "large" and for "moderately large" samples the efficiency of some non-parametric methods, in particular, those of the Kolmogorov-Smirnov type, for both the one-sided and the two-sided testing problems. We also discuss, in general, some of the problems of moderately large samples.

In the two-sided case, the Kolmogorov-Smirnov (K-S) and Kuiper tests are asymptotically as efficient as the median for location parameters in the symmetric unimodal case. We have compared, numerically, the relative efficiency of the K-S test to the median for double-exponential and uniform alternatives for "reasonable" sample sizes, and we find a slow approach to 1.

For the one-sided case, the situation is different. Here the efficiency depends on the specific test used. The analysis of this situation indicates that a better test statistic than the one-sided K-S test is to use the difference of the positive and negative deviations.

## 2. Preliminaries

For both the one-sided and two-sided tests, we assume that the sample sizes are sufficiently large that the asymptotic distributions are suffi-

*This research was supported in part by the Office of Naval Research under Contract N00014-67-A-0226-0008. Reproduction in whole or in part is permitted for any purpose of the United States Government.

ciently good approximations. Also, for the two-sided case we assume that the sample is sufficiently large that the weight measure (loss function times prior probability measure) is approximately proportional to $\delta_\theta$ for rejection of the null hypothesis and to $|\theta|^q d\theta$ for acceptance - our calculations here are only carried out for $q = 0$. For the one-sided case we assume that the weight function is approximately proportional to $|\theta| d\theta$ for either kind of error - the author has not been able to find a practical example of any other kind.

We assume that the scale has been chosen so that the density at the median is $\frac{1}{2}$, and we define

$$X_+ = 2\sqrt{n} \sup (F_n(x) - F(x)),$$

$$X_- = 2\sqrt{n} \sup (F(x) - F_n(x)),$$

$$X_o = 2\sqrt{n} (\tfrac{1}{2} - F_n(0)),$$

$$\varphi = \sqrt{n}\, \theta,$$

where the null hypothesis is $\theta = 0$.

With this normalization, the median for small $\theta$ is approximately $X_o / \sqrt{n}$, and $X_o$ is approximately normal with mean $\varphi$ and variance 1. Also, under the null hypothesis, by the usual methods the joint c.d.f. of $X_+$ and $X_-$ is found to be [1] for $a, b \geq 0$,

$$\text{(1)} \qquad P(X_+ \leq a, X_- \leq b) = \sum_{-\infty}^{\infty} e^{-\frac{n^2}{2}(a+b)^2} - \sum_{-\infty}^{\infty} e^{-\frac{1}{2}(na+(n-1)b)^2}$$

Under the alternative $\varphi$,

$$\text{(2)} \qquad 2\sqrt{n}(F_n(x) - F(x)) = 2\sqrt{n}(G_n(x-\theta) - F(x-\theta) + F(x-\theta) - F(x))$$

$$= Y_n(x-\theta) - 2\varphi f(x-\lambda\theta)$$

$$\sim Y_n(x) - 2\varphi f(x),$$

and so

(3) $$2\sqrt{n}(F_n F^{-1}(t) - t) \sim X_n(t) - 2\varphi f F^{-1}(t).$$

In the double-exponential case $2fF^{-1}(t) = 1 - 2|t-\frac{1}{2}|$; in the uniform case $2fF^{-1}(t) = 1$; and in general, $2fF^{-1}(\frac{1}{2})=1$, $2fF^{-1}(t) \geq 0$ and unimodal at $\frac{1}{2}$. We will for certain purposes consider an "extreme" distribution with density $\frac{1}{2}$ "at θ" and "0" elsewhere. In the sequel we shall use h for $2fF^{-1}$.

For the double-exponential case, we can find the limiting distribution of the maximum and minimum of $X_+$ and $X_-$ given $X_o$ as

(4) $$P(X_+ \leq a,\ X_- \leq b | X_o=q) = \left( \sum_{-\infty}^{\infty} e^{-n^2(a+b)^2 - n(a+b)q} - \sum_{-\infty}^{\infty} e^{-(na+(n-1)b)^2 - (na+(n-1)b)q} \right)^2$$

independent of $\varphi$.

3. <u>The one-sided case</u>

In this case, the weight density is typically a multiple of $|\theta - \theta_o|$ where $\theta_o$ is that parameter value at which there is indifference as to which action to take. If we use a statistic with variance $\frac{\sigma^2}{n}$, we would obtain the Bayes risk as

(5) $$\rho = \frac{C}{n\sigma} \int\int_{\varphi x > 0} |\varphi| \frac{1}{\sqrt{2\pi}} \quad e^{-\frac{1}{2}\frac{(x+\varphi)^2}{\sigma^2}} dx\, d\varphi = \frac{1}{2}\frac{C\sigma^2}{n}$$

Thus we see that, for regular estimates, the efficiency is proportional to $1/\sigma^2$. We have chosen $\sigma^2 = 1$ for the median by our choice of units. Let us now consider what happens for the one-sided K-S test with uniform alternatives. Here the probability of error is

(6) $$\begin{cases} e^{-\frac{1}{2}(a+\varphi)^2} & \varphi > 0, \\ 1 - e^{-\frac{1}{2}(a+\varphi)^2} & -a < \varphi < 0, \\ 0 & \varphi < -a, \end{cases}$$

so, letting $\rho^* = \frac{n\rho}{C}$

(7) $$\rho^* = \frac{1}{2}a^2 + \int_{-a}^{\infty} \varphi e^{-\frac{1}{2}(a+\varphi)^2} d\varphi$$

$$= \frac{1}{2}a^2 + \int_{0}^{\infty} (x-a)\ e^{-\frac{1}{2}x^2} dx$$

$$= \frac{1}{2}a^2 - a\sqrt{\frac{\pi}{2}} + 1$$

This is minimized at $a = \sqrt{\frac{\pi}{2}}$, obtaining

(8) $$\hat{\rho}^* = 1 - \frac{\pi}{4} = .2146$$

for an efficiency of 2.33.

If we consider instead the statistic $X_+ - X_-$, we find from (1) that

(9) $$P(X_+ - X_- > \lambda) = \sum_{n=1}^{\infty} \frac{1}{4n^2 - 1}\ e^{-\frac{n^2\lambda^2}{2}},$$

yielding

(10) $$\rho^* = \int_{-\infty}^{\infty} |\lambda| \sum_{n=1}^{\infty} \frac{1}{4n^2 - 1}\ e^{-2n^2\lambda^2} d\lambda = 1 - \frac{\pi^2}{12} = .1776.$$

The efficiency of this test is 2.82, so that the test is 1.21 times as efficient as the usual one-sided K-S test.

For distributions with very large tails, the one-sided K-S test discriminates well against shifts in one direction, but poorly in the other. In the "extreme" case, no one-sided test has finite asymptotic risk. Even in the uniform case, the improved discrimination can be seen from the following table of error probabilities.

| $\lvert\varphi\rvert$ | K-S ($\varphi > 0$) | K-S ($\varphi < 0$) | DIF |
|---|---|---|---|
| 0.0 | .4559 | .5441 | .5000 |
| 0.2 | .3478 | .4258 | .3763 |
| 0.4 | .2549 | .3052 | .2623 |
| 0.6 | .1795 | .1922 | .1664 |
| 0.8 | .1215 | .0980 | .0931 |
| 1.0 | .0790 | .0276 | .0451 |
| 1.2 | .0498 | .0014 | .0187 |
| 1.4 | .0296 | 0 | .0066 |
| 1.6 | .0171 | 0 | .0020 |
| 1.8 | .0095 | 0 | .0005 |
| 2.0 | .0050 | 0 | .0001 |
| 2.2 | .0026 | 0 | .0000 |
| 2.4 | .0013 | 0 | .0000 |
| 2.6 | .0006 | 0 | .0000 |
| 2.8 | .0003 | 0 | .0000 |
| 3.0 | .0001 | 0 | .0000 |

Computations not yet completed indicate a similar result for the double exponential.

## 4. The two-sided case

Here the asymptotic Bayes risk efficiency is easy to obtain - the K-S and Kuiper tests each have asymptotic Bayes risk efficiency exactly equal to that of the median (see [2]). However, one must be careful of using the limiting argument here too quickly. For example, let $k$ test statistics be given, so that for a given sample size $N$ the $i$-th test should have a type I error of $\alpha_i$. If we now decide to reject if any test rejects at level $\min \alpha_i$, the resulting test has at least as good asymptotic relative efficiency as the best one!

We have computed the relative efficiency of the K-S test to that based on the median for various risks (expressed as multiples of the type I risk). The type II loss is here constant. A brief table is appended.

| Risk | Efficiency Uniform | Double-exponential |
|---|---|---|
| .5 | 1.961 | .798 |
| .1 | 1.675 | .847 |
| .01 | 1.540 | .880 |
| .001 | 1.469 | .898 |
| .0001 | 1.422 | .911 |
| .00001 | 1.388 | .921 |
| .000001 | 1.362 | .928 |
| .0000001 | 1.341 | .934 |

As we can see, there is a very slow approach of the efficiency to 1, as the range of sample sizes in this calculation exceeds $10^{14}$.

The author believes that the results will be similar for other distributions and other loss functions, and simulations studies will be made to investigate this.

Of course, one really should consider infinite dimensional families of alternatives. Preliminary considerations indicate that the convergence to the asymptotic relative efficiency is slower.

## REFERENCES

[1]. Doob, J. L. Heuristic approach to the Kolmogorov-Smirnov theorem. Ann. Math. Stat. **20** (1949) pp. 393-403.

[2]. Rubin, H. and Sethuraman, J. Bayes risk efficiency. Sankhyā, Series A, **27**, pp. 347-356.

Unclassified
Security Classification

## DOCUMENT CONTROL DATA - R&D

*(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)*

| | |
|---|---|
| 1. ORIGINATING ACTIVITY (Corporate author)<br>Purdue University | 2a. REPORT SECURITY CLASSIFICATION<br>Unclassified<br>2b GROUP |

3. REPORT TITLE

Decision-Theoretic Evaluation of Some Non-Parametric Methods

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)

Technical Report, July 1969

5. AUTHOR(S) (Last name, first name, initial)

Rubin, Herman

| | | |
|---|---|---|
| 6. REPORT DATE<br>July 1969 | 7a. TOTAL NO. OF PAGES<br>6 | 7b. NO. OF REFS<br>2 |

| | |
|---|---|
| 8a. CONTRACT OR GRANT NO.<br>N00014-67-A-0226-0008<br>b. PROJECT NO.<br>c.<br>d. | 9a. ORIGINATOR'S REPORT NUMBER(S)<br>Mimeograph Series No. 193<br>9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report) |

10. AVAILABILITY/LIMITATION NOTICES

Distribution of this document is unlimited.

| | |
|---|---|
| 11. SUPPLEMENTARY NOTES | 12. SPONSORING MILITARY ACTIVITY<br>Office of Naval Research<br>Washington, D.C. |

13 ABSTRACT

The Kolmogorov-Smirnov test is compared to the median for one-sided and two-sided alternatives. In the two-sided case, although the asymptotic Bayes risk efficiency is the same for symmetric unimodal distributions, the relative efficiency approaches one slowly. In the one-sided case, the efficiency depends heavily on the test. For uniform alternatives, the K-S test is much better than the median, and the test based on the difference of the positive and negative deviations is still better. This latter test appears to have good properties even when the one-sided K-S test does not.

DD FORM 1 JAN 64 1473

Unclassified
Security Classification

| 14. KEY WORDS | LINK A | | LINK B | | LINK C | |
|---|---|---|---|---|---|---|
| | ROLE | WT | ROLE | WT | ROLE | WT |
| Non-parametric | | | | | | |
| test | | | | | | |
| Kolmogorov-Smirnov | | | | | | |
| Bayes risk | | | | | | |
| efficiency | | | | | | |

## INSTRUCTIONS

1. ORIGINATING ACTIVITY: Enter the name and address of the contractor, subcontractor, grantee, Department of Defense activity or other organization *(corporate author)* issuing the report.

2*a*. REPORT SECURITY CLASSIFICATION: Enter the overall security classification of the report. Indicate whether "Restricted Data" is included. Marking is to be in accordance with appropriate security regulations.

2*b*. GROUP: Automatic downgrading is specified in DoD Directive 5200.10 and Armed Forces Industrial Manual. Enter the group number. Also, when applicable, show that optional markings have been used for Group 3 and Group 4 as authorized.

3. REPORT TITLE: Enter the complete report title in all capital letters. Titles in all cases should be unclassified. If a meaningful title cannot be selected without classification, show title classification in all capitals in parenthesis immediately following the title.

4. DESCRIPTIVE NOTES: If appropriate, enter the type of report, e.g., interim, progress, summary, annual, or final. Give the inclusive dates when a specific reporting period is covered.

5. AUTHOR(S): Enter the name(s) of author(s) as shown on or in the report. Enter last name, first name, middle initial. If military, show rank and branch of service. The name of the principal author is an absolute minimum requirement.

6. REPORT DATE: Enter the date of the report as day, month, year; or month, year. If more than one date appears on the report, use date of publication.

7*a*. TOTAL NUMBER OF PAGES: The total page count should follow normal pagination procedures, i.e., enter the number of pages containing information.

7*b*. NUMBER OF REFERENCES: Enter the total number of references cited in the report.

8*a*. CONTRACT OR GRANT NUMBER: If appropriate, enter the applicable number of the contract or grant under which the report was written.

8*b*, 8*c*, & 8*d*. PROJECT NUMBER: Enter the appropriate military department identification, such as project number, subproject number, system numbers, task number, etc.

9*a*. ORIGINATOR'S REPORT NUMBER(S): Enter the official report number by which the document will be identified and controlled by the originating activity. This number must be unique to this report.

9*b*. OTHER REPORT NUMBER(S): If the report has been assigned any other report numbers *(either by the originator or by the sponsor)*, also enter this number(s).

10. AVAILABILITY/LIMITATION NOTICES: Enter any limitations on further dissemination of the report, other than those imposed by security classification, using standard statements such as:

(1) "Qualified requesters may obtain copies of this report from DDC."

(2) "Foreign announcement and dissemination of this report by DDC is not authorized."

(3) "U. S. Government agencies may obtain copies of this report directly from DDC. Other qualified DDC users shall request through
________________."

(4) "U. S. military agencies may obtain copies of this report directly from DDC. Other qualified users shall request through
________________."

(5) "All distribution of this report is controlled. Qualified DDC users shall request through
________________."

If the report has been furnished to the Office of Technical Services, Department of Commerce, for sale to the public, indicate this fact and enter the price, if known.

11. SUPPLEMENTARY NOTES: Use for additional explanatory notes.

12. SPONSORING MILITARY ACTIVITY: Enter the name of the departmental project office or laboratory sponsoring *(paying for)* the research and development. Include address.

13. ABSTRACT: Enter an abstract giving a brief and factual summary of the document indicative of the report, even though it may also appear elsewhere in the body of the technical report. If additional space is required, a continuation sheet shall be attached.

It is highly desirable that the abstract of classified reports be unclassified. Each paragraph of the abstract shall end with an indication of the military security classification of the information in the paragraph, represented as *(TS)*, *(S)*, *(C)*, *or (U)*.

There is no limitation on the length of the abstract. However, the suggested length is from 150 to 225 words.

14. KEY WORDS: Key words are technically meaningful terms or short phrases that characterize a report and may be used as index entries for cataloging the report. Key words must be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location, may be used as key words but will be followed by an indication of technical context. The assignment of links, roles, and weights is optional.