AD715127

THEMIS SIGNAL ANALYSIS STATISTICS RESEARCH PROGRAM

INVESTIGATION OF SYSTEM READINESS WHEN SOME DEFECT TYPES UNKNOWN

by

Edwin R. Huber and John E. Walsh

D D C
DEC 1971
D

Technical Report No. 82
Department of Statistics THEMIS Contract

Reproduction in whole or in part is permitted
for any purpose of the United States Government.

This document has been approved for public release
and sale; its distribution is unlimited.

Reproduced by
NATIONAL TECHNICAL
INFORMATION SERVICE
Springfield, Va. 22151

Department of Statistics
Southern Methodist University
Dallas, Texas 75222

INVESTIGATION OF SYSTEM READINESS WHEN SOME DEFECT TYPES UNKNOWN

Edwin R. Huber
U. S. Naval Torpedo Station, Keyport

John E. Walsh*
Southern Methodist University**

ABSTRACT

A system consists of subsystems and performs satisfactorily when defects in subsystems do not cause its failure. For each subsystem, defect types are identified by their nature and by their level of probabilistic influence on system failure (finite number of levels). The subsystems and possible defect types are so defined that, for satisfactory system performance, a defect type can occur at most once in a subsystem. Also, for this conditional case, probabilities for a defect type are not influenced by occurrence of other defect types. Moreover, the defect types are independent and have small probabilities with respect to occurrence. System ability is represented by the Readiness Index (RI), which is the probability of no defect that causes system failure. Statistical investigation of the RI is complicated by possible existence of defect types which have not yet been identified. Suitable data are available for each combination of subsystem and level of probabilistic influence on system failure. For every combination, the number of defect types occurring is observed over some repetitions. Unbiased estimation, also approximate tests and confidence intervals, are developed (some results are conservative and/or apply to at least moderately large RI values).

---

*Based on work performed at the Quality Evaluation Laboratory, U. S Naval Torpedo Station, Keyport, Washington.
**Research partially supported by ONR Contract N00014-68-A-0515 and by Mobil Research and Development Corporation. Based on some methods developed under NASA Grant NGR 44-007-028.

INTRODUCTION AND DISCUSSION

Considered is the ability of a system (for example, a torpedo) to perform satisfactorily over an operation. This system is composed of subsystems and performs satisfactorily (does not fail) when occurrence of defects in subsystems does not cause its failure.

Within subsystems, types of defects are identified by their nature and also by the probability that the defect does not cause system failure. Only a finite number of levels are considered to occur for the probability that a defect does not cause system failure. The number of levels and the values for these probabilities are known and can change with the subsytem.

The ability of a system is respresented by its Readiness Index (RI), which is the probability that no defect type occurs whose influence causes system failure. Methods are developed for investigating the RI from observational data. Observations are obtained for every possible combination of subsystem and probability level (for not causing system failure). For a given subsystem, an observation furnishes the observed number of defect types that have the specified probability level.

A complication in investigation of the RI is that there may be defect types whose existence has not yet been identified. Moreover, the number of undiscovered defect types, and their correspondence with the possible probability levels, are unknown for each subsystem.

The results are based on some assumptions that are to hold for the conditional case of no system failure. These are:

(a) Any given defect type can occur at most once in a subsystem.

(b) For any defect type, the probability that it does not cause system failure is not influenced by occurrence of other defect types in its subsystem or of defect types in other subsystems.

(c) With respect to occurrence, all defect types (over all subsystems) are statistically independent.

(d) No probabilities for occurrence of defect types are of more than moderate size and almost all of these probabilities are small.

(e) The observational data are statistically independent and data for the same subsystem and probability level combination constitute a random sample. Also, these data are obtained under conditions that correspond to the case of a system that has not failed.

(f) Consideration of only a finite (almost always small) number of probability levels (for not causing system failure) for each subsystem yields acceptable accuracy for the RI and its investigation.

In a number of cases, the subsystems and defect types can be defined so that assumptions (a) - (e) are met to a reasonable approximation. For example, this seems to be the case for many situations involving torpedoes. Now, consider assumption (f). Often, there are limitations on the accuracy to which the probability of not causing system failure can be determined for a defect of a given nature. Use of a small set of representative values, each of which corresponds to an interval of values, is about as good as can be done under these circumstances. Of course, use of enough levels (say, equally spaced) should provide sufficient accuracy.

However, too many levels may not be warranted and also can introduce difficulties in the collection of enough data for use of some of the approximate results that are developed. With these assumptions, especially (b), the RI becomes the probability that the system does not fail due to defects.

The principal results consist of an unbiased estimate for the RI, some conservative one-sided confidence intervals and significance tests for the RI, some approximate one-sided intervals and tests for the RI, and some two-sided intervals and tests for the RI (conservative and approximate). In some cases, the RI is assumed to be at least moderately large or the expected number of defects that cause failure is assumed to be small. Here, a conservative interval has a confidence coefficient at least equal to a determined value that is appropriate for intervals. A conservative test has a significance level that is at most equal to a determined value that is suitable for tests.

Notation and some basic expressions are given in the next section. The following section contains the unbiased estimate for the RI, including an expression for the variance of this estimate and an unbiased estimate of this variance. Inequalities and approximations that are used in developing the intervals and tests are stated in the next following section. The final three sections contain the material on conservative one-sided intervals and tests, approximate one-sided intervals and tests, and two-sided intervals and tests, respectively. Additional material, associated principally with the results for intervals and tests, is given in two appendices.

NOTATION AND BASIC EXPRESSIONS

Most of the notation used is introduced here.

p_{ij} = conditional probability that if defect type j occurs in subsystem i, (i=1,...,n), failure of the system does not happen because of the occurrence of this defect. The defect types for which $p_{ij} < 1$ are designated by j = 1,...,m(i) and only these types receive consideration. The value of n is known but the value of m(i) is unknown.

$p_i(u)$ = u-th of a set of U(i) possible values that are considered to occur for the p_{ij} that are less than unity. U(i) and all the $p_i(u)$ have known values.

d_{ij} = probability that defect type j occurs in subsystem i.

$d_{ij}(u)$ = d_{ij} when j is such that $p_{ij} = p_i(u)$, and equals zero otherwise. (Any defect type corresponds to exactly one value of u.)

$$d_i(u) = \sum_{j=1}^{m(i)} d_{ij}(u)$$

$y_{iv}(u)$ = observed number of defect types with conditional probability level $p_i(u)$ that occur for the v-th observation on the combination of this probability level and system i, where v = 1,...,V(i,u) ≥ 1.

$$\hat{d}_i(u) = \sum_{v=1}^{V(i,u)} y_{iv}(u)/V(i,u)$$

$$s^2 = \sum_{i=1}^{n} \sum_{u=1}^{U(i)} \{[1 - p_i(u)]^2/V(i,u)[V(i,u)-1]\} \sum_{v=1}^{V(i,u)} [y_{iv}(u) - \hat{d}_i(u)]^2$$

$$S(\epsilon)^2 = \epsilon \sum_{i=1}^{n} \sum_{u=1}^{U(i)} [1-p_i(u)]^2 \hat{d}_i(u)/V(i,u), \qquad (0 < \epsilon \leq 1)$$

$x_{ij}(u,v)$ = random variable that equals 1 if defect type j occurs for the v-th observation on the combination of probability level $p_i(u)$ and subsystem i, where $p_i(u) = p_{ij}$, and equals 0 if defect type j does not occur for this observation, $v = 1,\ldots,V(i,u)$.

K_α = deviate of standardized normal distribution (zero mean, unit variance) that is exceeded with probability α.

$M = \sum_{i=1}^{n} m(i)$. The value of M is unknown.

R = the RI = probability that the system does not fail due to occurrence of any of the defect types

$$= \prod_{i=1}^{n} \prod_{j=1}^{m(i)} [1-d_{ij}(1-p_{ij})].$$

The expression for R follows from assumptions (a) - (c) and from the consideration that the probability of no system failure due to defect type j of system i equals

(prob. type j does not occur)+(prob. type j occurs but does not cause failure)

$$= (1 - d_{ij}) + (d_{ij}p_{ij}) = 1 - d_{ij}(1-p_{ij}).$$

The random variable $x_{ij}(u,v)$ is introduced for use in derivations and need not be observed. In all cases, the observed data are the $y_{iv}(u)$. Often, in practice, all the defects that occur for a given subsystem are observed at the same time during the subsystem operation, and are later subdivided to obtain the $y_{iv}(u)$ for different u and this value of v. Such a procedure

is permissible since assumption (c) guarantees independence, with respect to occurrence, of all defect types. Of course, V(i,u) has the same value for all u when the data are obtained in this manner.

UNBIASED ESTIMATE

An exactly unbiased estimate of R is given by

$$\hat{R} = \prod_{i=1}^{n} \prod_{u=1}^{U(i)} \left[\frac{1}{V(i,u)} \sum_{v=1}^{V(i,u)} p_i(u)^{y_{iv}(u)} \right].$$

The unbiased nature of this estimate follows from the relation

$$Ep_i(u)^{y_{iv}(u)} = E\left[\frac{1}{V(i,u)} \sum_{w=1}^{V(i,u)} p_i(u)^{y_{iw}(u)} \right],$$

which holds for all v on the basis of assumption (e), and the relation

$$\prod_{j=1}^{m(i)} [1-x_{ij}(u,v)(1-p_{ij})] = p_i(u)^{y_{iv}(u)}.$$

Combined with assumptions (a) - (c) and (e), these relations imply that

$$E\hat{R} = \prod_{i=1}^{n} \prod_{u=1}^{U(i)} Ep_i(u)^{y_{iv}(u)}$$

$$= \prod_{i=1}^{n} \prod_{u=1}^{U(i)} \prod_{j=1}^{m(i)} E[1-x_{ij}(u,v)(1-p_{ij})] = R.$$

The variance of $\hat{R}$, on the basis of assumptions (a) - (c), (e) and use of ref. 1, equals

$$\prod_{i=1}^{n} \prod_{u=1}^{U(i)} \left(\mathrm{var}[p_i(u)^{y_{iv}(u)}]/V(i,u) + \{E[p_i(u)^{y_{iv}(u)}]\}^2 \right)$$

$$- \prod_{i=1}^{n} \prod_{u=1}^{U(i)} \{E[p_i(u)^{y_{iv}(u)}]\}^2 .$$

When $V(i,u) \geq 2$ for all i and u, this variance is unbiasedly estimated by

$$\prod_{i=1}^{n} \prod_{u=1}^{U(i)} \left[\frac{1}{V(i,u)} \sum_{v=1}^{V(i,u)} p_i(u)^{y_{iv}(u)}\right]^2$$

$$- \prod_{i=1}^{n} \prod_{u=1}^{U(i)} \left\{\frac{V(i,u)}{V(i,u)-1}\left[\frac{1}{V(i,u)} \sum_{v=1}^{V(i,u)} p_i(u)^{y_{iv}(u)}\right]^2 \right.$$

$$\left. - \frac{V(i,u)^{-1}}{V(i,u)-1} \sum_{v=1}^{V(i,u)} p_i(u)^{2y_{iv}(u)}\right\},$$

since

$$E\left[\frac{1}{V(i,u)} \sum_{v=1}^{V(i,u)} p_i(u)^{y_{iv}(u)}\right]^2 = \mathrm{var}[p_i(u)^{y_{iv}(u)}]/V(i,u)$$

$$+ \{E[p_i(u)^{y_{iv}(u)}]\}^2$$

and $\{E[p_i(u)^{y_{iv}(u)}]\}^2$ equals

$$E\left\{\frac{V(i,u)}{V(i,u)-1}\left[\sum_{v=1}^{V(i,u)} p_i(u)^{y_{iv}(u)}\right]^2 - \frac{V(i,u)^{-1}}{V(i,u)-1} \sum_{v=1}^{V(i,u)} p_i(u)^{2y_{iv}(u)}\right\},$$

which follows from material in ref. 1.

INEQUALITIES AND APPROXIMATIONS

First, suppose that the largest of the $d_{ij}(1-p_{ij})$ does not exceed .08 and that their arithmetic average does not exceed .02. This assumption is somewhat (but not much) more stringent than assumption (d). The value of $d_{ij}(1-p_{ij})$ is the probability that defect j occurs in system i and causes system failure. Consider failure and nonfailure as the outcomes of binomial events for each of the M combinations for i and j. From ref. 2, the Poisson approximation is applicable to these binomial events and

$$P(\text{no failures}) = R \doteq \exp\left[-\sum_{i=1}^{n}\sum_{j=1}^{m(i)} d_{ij}(1-p_{ij})\right]. \qquad (1)$$

Thus, with this somewhat stronger assumption, approximate confidence intervals and significance tests for R are directly obtainable from intervals and tests for

$$\sum_{i-1}^{n}\sum_{j=1}^{m(i)} d_{ij}(1-p_{ij}) = \sum_{i=1}^{n}\sum_{u=1}^{U(i)} d_i(u)\,[1-p_i(u)]. \qquad (2)$$

Moreover, this approximate expression for R also applies when a mild form of m-dependence occurs for the data (so that assumption (e) is violated).

Second, suppose that (2) is less than unity. Then (see ref. 3, for case of independence), the sharp inequalities

$$1 - \sum_{i=1}^{n}\sum_{j=1}^{m(i)} d_{ij}(1-p_{ij}) \le R \le \left[1 - \frac{1}{M}\sum_{i=1}^{n}\sum_{j=1}^{m(i)} d_{ij}(1-p_{ij})\right]^{M} \qquad (3)$$

hold. Also, by expansion, the upper bound is at most equal to

$$1 - \sum_{i=1}^{n} \sum_{u=1}^{U(i)} d_i(u)\,[1-p_i(u)] + \tfrac{1}{2}\left[\sum_{i=1}^{n} \sum_{u=1}^{U(i)} d_i(u)\,[1-p_i(u)]\right]^2, \qquad (4)$$

which is not a sharp upper limit but one that requires no knowledge of M. When the value of (2) is at most .2, the sharp upper and lower bounds are nearly equal to each other and to (4). Then

$$R \doteq \left[1-\tfrac{1}{2}\sum_{i=1}^{n} \sum_{u=1}^{U(i)} d_i(u)[\,1-p_i(u)\,]\right]^2, \qquad (5)$$

where this expression for R approximately minimizes its maximum deviation from the upper and lower bounds. Conservative intervals and tests for R can be obtained by use of the sharp lower bound and by use of (4). These are based on the assumption that (2) has a value less than unity. Approximate intervals and tests can be obtained from (5) for cases where the value of (2) is believed to be at most .2.

Sometimes, more information is available about R than is available about the $d_{ij}(1-p_{ij})$ or about (2). If $R > e^{-1} \doteq .368$, the value of (2) is less than unity and both of the inequalities (3) hold for R (see Appendix A). That is, both of the inequalities (3) and (4) are usable when the RI is of at least moderate size. If $R \geq .8$, the approximation (5) is also usable.

A fundamental statistic used for the confidence intervals and tests is

$$\sum_{i=1}^{n} \sum_{u=1}^{U(i)} \hat{d}(u)\,[1-p_i(u)], \qquad (6)$$

which is an unbiased estimate of (2). By expressing the $y_{iv}(u)$ in terms of the $x_{ij}(u,v)$, the variance of (6) is easily seen to be

$$\sum_{i=1}^{n} \sum_{j=1}^{m(i)} \sum_{u=1}^{U(i)} (1-p_{ij})^2 d_{ij}(u)[1-d_{ij}(u)]/V(i,u). \qquad (7)$$

The distribution of (6) should be approximately normal when (7) is not too small, which should often be the case. In deriving results, the distribution of (6) is considered to be approximately normal for cases where the confidence coefficients are not too near unity (say, at most .995) and the significance levels are not too small (say, at least .005).

On the basis of assumption (e), it is easily seen that s^2 is an unbiased estimate of (7). A conservative estimate of (7), with a larger expectation but smaller variation (can be much smaller) than s^2, is provided by $S(1)^2$.

$$ES(1)^2 = \epsilon^{-1}ES(\epsilon)^2 = \sum_{i=1}^{n} \sum_{j=1}^{m(i)} \sum_{u=1}^{U(i)} (1-p_{ij})^2 d_{ij}(u)/V(i,u),$$

which, according to assumption (d), should at least roughly equal (7).

Use of $S(\epsilon)^2$ with an appropriate value for ϵ should provide a satisfactory approximate estimate for (7). In many cases, the value of (7) can be assumed to be at least $ES(.96)^2$. Then use of $S(.98)^2$ as the estimate of (7) would seem suitable. The value of $ES(.98)^2$ would differ from (7) by at most two percent, which implies that one percent is about the maximum change that would be appropriate in the corresponding expression for the observed standard deviation that is used in intervals

and tests. A one percent change in the statistics, through adjustment in the probability level used, does not cause an important change in the confidence coefficient and significance level values that are considered. Similarly, the value of (7) nearly always should be at least $ES(.92)^2$ when assumption (d) holds. Then, use of $S(.96)^2$ as the estimate would call for a change of at most about two percent in the observed standard deviation. A two percent change in this statistic can be accomplished by a moderately unimportant change in a confidence coefficient or significance level value.

In the intervals and tests presented, $S(\epsilon)^2$ is used as the estimate of (7), so that $S(\epsilon)$ is the observed standard deviation. The value for ϵ is appropriately chosen (nearly always, so that $.96 \leq \epsilon \leq 1$). The variance of $S(\epsilon)^2$ is obtained in Appendix B. Approximate estimation of the variance of $S(\epsilon)^2$ is also considered in Appendix B.

CONSERVATIVE ONE-SIDED INTERVALS AND TESTS

The results of this and the following two sections are based on the assumption that the value of (2) is less than unity or that $R > .368$. Also, the distribution of (6) is assumed to be acceptably near normality for the confidence coefficient and significance level values that are considered.

A conservative one-sided interval with random lower endpoint is provided (approximately) by the relation

$$P\left[1 - \sum_{i=1}^{n} \sum_{u=1}^{U(i)} \hat{d}_i(u)\,[1 - p_i(u)] - S(\epsilon) K_\alpha \leq R\right] \geq 1 - \alpha .$$

This follows from the lower bound of (3) and occurrence of a distribution that is approximately standardized normal for the quantity

$$\left\{1 - \sum_{i=1}^{n} \sum_{u=1}^{U(i)} \hat{d}_i(u)[1 - p_i(u)]\right\} - \left\{1 - \sum_{i=1}^{n} \sum_{u=1}^{U(i)} d_i(u)[1 - p_i(u)]\right\}$$

divided by $S(\epsilon)$. Often, the true confidence coefficient value will be definitely greater than $1 - \alpha$ when the value of (2) is not substantially less than unity, since R will be substantially greater than its sharp lower bound. However, the confidence coefficient should often be near $1 - \alpha$ when (2) is at most .2.

A conservative one-sided interval with random upper endpoint is provided (approximately) by

$$P\left\{R \le 1 - \sum_{i=1}^{n} \sum_{u=1}^{U(i)} \hat{d}_i(u)[1 - p_i(u)] + S(\epsilon)K_\alpha + \tfrac{1}{2}\left[\sum_{i=1}^{n} \sum_{u=1}^{U(i)} \hat{d}_i(u)[1 - p_i(u)] - S(\epsilon)K_\alpha\right]^2\right\} \ge 1 - \alpha.$$

This follows from the approximate normality for (6), the upper bound (4) for R, and the fact that $1 - Z + (1/2)Z^2$ is a strictly monotonically decreasing function of Z for $0 \le Z < 1$. Here, the true confidence coefficient should be near $1 - \alpha$ when the value of (1) is at most .4.

Direct use of these intervals provides conservative one-sided significance tests. In all cases, the null hypothesis asserts that $R = R_0$, where R_0 is a specified value.

First, consider emphasis of the alternative hypothesis $R > R_0$. For this one-sided test, $R = R_0$ is rejected in favor of $R > R_0$ if and only if

$$R_0 < 1 \sum_{i=1}^{n} \sum_{u=1}^{U(i)} \hat{d}_i(u)[1 - p_i(u)] - S(\epsilon)K_\alpha \ .$$

The significance level of this test is (approximately) at most α and the value used for R_0 is at least .368. Often, the true significance level is substantially less than α. However, it is frequently near α when R_0 is at least .8.

Now, consider emphasis of $R < R_0$. For this one-sided test, $R = R_0$ is rejected in favor of $R < R_0$ if and only if

$$R_0 > 1 - \sum_{i=1}^{n} \sum_{u=1}^{U(i)} \hat{d}_i(u)[1 - p_i(u)] + S(\epsilon)K_\alpha + \tfrac{1}{2}\left[\sum_{i=1}^{n} \sum_{u=1}^{U(i)} \hat{d}_i(u)[1 - p_i(u)] - S(\epsilon)K_\alpha\right]^2 \ .$$

The significance level of this test is (approximately) at most α. The true significance level should often be near α when $R_0 \geq .6$.

APPROXIMATE ONE-SIDED INTERVALS AND TESTS

Results based on the Poisson approximation are considered first. Here, by assumption, the largest of the $d_{ij}(1-p_{ij})$ does not exceed .08 and their arithmetic average does not exceed .02. Also, the error in approximating R through (1) is assumed to be small compared to variation in the statistic (interval endpoint, or test statistic) involved.

An approximate one-sided interval with random lower endpoint is provided by

$$P\left\{\exp\left[-\sum_{i=1}^{n} \sum_{u=1}^{U(i)} \hat{d}_i(u)[1 - p_i(u)] - S(\epsilon)K_\alpha\right] \leq R\right\} \doteq 1 - \alpha. \quad (8)$$

This follows from (1), the approximate normality that is assumed for (6), and the fact that e^{-Z} is a strictly monotonically decreasing function Z.

An approximate one-sided interval with random upper endpoint is furnished by

$$P\left\{R \leq \exp\left[-\sum_{i=1}^{n}\sum_{u=1}^{U(i)} \hat{d}_i(u)[1-p_i(u)] + S(\epsilon)K_\alpha\right]\right\} \doteq 1 - \alpha. \qquad (9)$$

This too follows from (1) and the approximate normality assumed for (6).

Now, consider the case where the approximation (5) is used. The error in using this approximation is assumed to be small compared to the variation in the statistic involved when the value of (2) is at most .2, which is the situation that is assumed to occur. Also, the additional notation that L_α equals

$$\max\left\{0,\ 1 - \tfrac{1}{2}\sum_{i=1}^{n}\sum_{u=1}^{U(i)} \hat{d}_i(u)[1 - p_i(u)] - [S(\epsilon)/2]K_\alpha\right\}.$$

is introduced.

An approximate one-sided interval with random lower endpoint is provided by

$$P(L_\alpha^* \leq R) \doteq 1 - \alpha.$$

This follows from (5), the approximate normality assumed for (6), the small probability of a negative value for

$$1 - \tfrac{1}{2}\sum_{i=1}^{n}\sum_{u=1}^{U(i)} \hat{d}_i(u)[1 - p_i(u)] - [S(\epsilon)/2]K_\alpha,$$

and the relation

$$P(L_\alpha \leq \sqrt{R}) = P(L_\alpha^a \leq R),$$

which is generally valid.

Likewise, an approximate one-sided interval with random upper endpoint is furnished by

$$P(R \leq L_{1-\alpha}^a) \doteq 1 - \alpha,$$

and has the same kind of basis.

The null hypothesis is $R = R_0$ and direct use of the intervals provides corresponding one-sided tests. Use of the Poisson approximation is considered first.

For the one-sided test that emphasizes $R > R_0$, the null hypothesis is rejected in favor of $R > R_0$ if and only if

$$R_0 < \exp\left[-\sum_{i=1}^{n}\sum_{u=1}^{U(i)} \hat{d}_i(u)\,[1-p_i(u)] - S(\epsilon)K_\alpha\right].$$

The significance level of this test is approximately α when the assumptions for (8) are satisfied for the case of $R = R_0$. These assumptions become more readily acceptable as the value used for R_0 increases (for null uses).

For the one-sided test that emphasizes $R < R_0$, the null hypothesis is rejected in favor of $R < R_0$ if and only if

$$R_0 > \exp\left[-\sum_{i=1}^{n}\sum_{u=1}^{U(i)} \hat{d}_i(u)\,[1-p_i(u)] + S(\epsilon)K_\alpha\right].$$

This test has a significance level of approximately α when the assumptions for (9) hold with $R = R_0$. Here too the assumptions are more readily

acceptable as R_0 increases (for null uses).

Finally, consider the one-sided tests that are based on (5). Here, the values used for R_0 are at least .8.

For the one-sided test that emphasizes $R > R_0$, the null hypothesis is rejected in favor of $R > R_0$ if and only if $R_0 < L_\alpha^{\mathbf{a}}$. For the test where $R < R_0$ is emphasized, $R = R_0$ is rejected in favor of $R < R_0$ if and only if $R_0 > L_{1-\alpha}^{\mathbf{a}}$. Each test has a significance level that approximately equals α.

TWO-SIDED INTERVALS AND TESTS

The two-sided confidence intervals and significance tests are obtained directly from the one-sided intervals and tests presented in the preceding two sections. Consideration of their development is limited to intervals, since the two-sided tests are obtained from the two-sided intervals.

Specifically, for intervals, let

$$P[R_1'(\alpha_1) \leq R], \qquad P[R \leq R_2'(\alpha_2)] \tag{9}$$

define one-sided intervals, where (approximately) the confidence coefficient for the first interval is either $1 - \alpha_1$ or at least $1 - \alpha_1$ (depending on whether the interval is approximate or conservative, respectively), and for the second interval is either $1 - \alpha_2$ or at least $1 - \alpha_2$. In all cases, $R_2'(\alpha_2) > R_1'(\alpha_1)$. Then by considering the complements of these intervals,

$$P[R < R_1'(\alpha_1)], \qquad P[R_2'(\alpha_2) < R]$$

define one-sided intervals, where (approximately) the confidence coefficient for the first interval is either α_1 or at most α_1, and for the second interval is either α_2 or at most α_2.

Thus, a two-sided interval and its confidence coefficient properties are provided by

$$P[R_1'(\alpha_1) \leq R \leq R_2'(\alpha_2)] = 1 - P[R < R_1'(\alpha_1)] - P[R_2'(\alpha_2) < R].$$

If both the intervals of (9) are conservative, the confidence coefficient is (approximately) at least $1 - \alpha_1 - \alpha_2$, with $P[R < R_1'(\alpha_1)]$ at most α_1 and $P[R_2'(\alpha_2) < R]$ at most α_2. When the first interval of (9) is conservative and the second approximate, the confidence coefficient is (approximately) at least $1 - \alpha_1 - \alpha_2$, with $P[R < R_1'(\alpha_1)]$ at most α_1 and $P[R_2'(\alpha_2 < R]$ approximately α_2. If the first interval of (9) is approximate and the second conservative, the confidence coefficient is (approximately) at least $1 - \alpha_1 - \alpha_2$, with $P[R < R_1'(\alpha_1)]$ approximately α_1 and $P[R_2'(\alpha_2) < R]$ at most α_2. When both intervals of (9) are approximate, the confidence coefficient is approximately $1 - \alpha_1 - \alpha_2$, with $P[R < R_1'(\alpha_1)]$ approximately α_1 and $P[R_2'(\alpha_2) < R]$ approximately α_2.

The assumptions for both of the intervals of (9) should be satisfied. Also, when both intervals of (9) are approximate, it is desirable that they both have the same basis for the approximation of R. Then, considerations similar to those in ref. 4 indicate that close approximation to the normality assumption is not so important, especially when intervals with $\alpha_1 = \alpha_2$ are used.

The null hypothesis for tests is still $R = R_0$. In all cases, the alternative hypothesis is $R \neq R_0$. Specifically, $R = R_0$ is rejected in favor of $R \neq R_0$ if and only if either $R_0 < R_1'(\alpha_1)$ or $R_0 > R_2'(\alpha_2')$. The significance level is

$$P[R_0 < R_1'(\alpha_1) \,|\, R_0 = R] + P[R_2'(\alpha_2) < R_0 \,|\, R_0 = R],$$

and its properties are determined from the properties of these two

probabilities. As an example, suppose that both intervals of (9) are conservative. Then, $P[R_0 < R_1'(\alpha_1) \,|\, R_0 = R]$ is at most α_1 and $P[R_2'(\alpha_2) < R_0 \,|\, R_0 = R]$ is at most α_2. As another example, suppose that the first interval of (9) is conservative and the second is approximate. Then, $P[R_0 < R_1'(\alpha_1) \,|\, R_0 = R]$ is at most α_1 and $P[R_2'(\alpha_2) < R_0 \,|\, R_0 = R]$ is approximately α_2.

In determining null properties of tests, it is only necessary that the assumptions for the intervals (9) are satisfied when the null hypotheses holds. Thus, for assumptions expressed in terms of R, the null value R_0 can be used for R in deciding whether the assumptions hold (as was done for some one-sided tests in the preceding two sections).

REFERENCES

1. Leo A. Goodman, "The variance of the product of K random variables," Journal of the American Statistical Association, Vol. 57, 1962, pp. 54-60.

2. John E. Walsh, Handbook of Nonparametric Statistics, D. Von Nostrand, Princeton, N. J., 1962, p. 191.

3. John E. Walsh and Grace J. Kelleher, "Approximations to large probabilities of all successes for general case and some operations research applications, "Journal of the Operations Research Society of Japan, Vol. 11, 1969, pp. 114-118.

4. John E. Walsh and Grace J. Kelleher, "Robust variables control charts based on sample means," Proceedings of International Quality Control Conference, 1969-Tokyo, 1969, pp. 507-509.

APPENDIX A

Here, it is shown that the value of (2) is less than unity when $R > e^{-1}$. From the sharp upper bound in (3)

$$R \leq \left[1 - \frac{1}{M} \sum_{i=1}^{n} \sum_{j=1}^{m(i)} d_{ij} (1-p_{ij})\right]^M$$

so that (2) is at most $M(1 - R^{1/M})$ which can be expressed as

$$M\{1 - [1 - (1 - R)]^{1/M}\}$$

$$= M[1 - 1 + (1/M)(1 - R) + (1/2)(1/M)(1 - 1/M)(1 - R)^2$$

$$+ (1/6)(1/M)(1 - 1/M)(2 - 1/M)(1 - R)^3 + \cdots]$$

$$\leq (1 - R) + (1/2)(1 - R)^2 + (1/3)(1 - R)^3 + \ldots = -\log_e R.$$

Thus, $R > e^{-1}$ implies that (2) is less than unity.

APPENDIX B

Development of the variance of $S(\epsilon)^2$ and of an estimate for this variance are considered here.

The same considerations that yielded (7), the variance of (6), show that the variance of $S(\epsilon)^2$ is

$$\epsilon^2 \sum_{i=1}^{n} \sum_{j=1}^{m(i)} \sum_{u=1}^{U(i)} (1 - p_{ij})^4 d_{ij}(u)[1 - d_{ij}(u)]/V(i,u)^3.$$

A conservative estimate for the variance is provided by

$$\epsilon^2 \sum_{i=1}^{n} \sum_{u=1}^{U(i)} [1 - p_i(u)]^4 \hat{d}_i(u)/V(i,u)^3,$$

whose expected value is

$$\epsilon^2 \sum_{i=1}^{n} \sum_{j=1}^{m(i)} \sum_{u=1}^{U(i)} (1 - p_{ij})^4 d_{ij}(u)/V(i,u)^3.$$

Using the same basis as that for estimating (7),

$$\epsilon^3 \sum_{i=1}^{n} \sum_{u-1}^{U(i)} [1 - p_i(u)]^4 \hat{d}_i(u)/V(i,u)^3$$

is used as the estimate for the variance of $S(\epsilon)^{\alpha}$.

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)	2a. REPORT SECURITY CLASSIFICATION
SOUTHERN METHODIST UNIVERSITY	UNCLASSIFIED
	2b. GROUP: UNCLASSIFIED

3. REPORT TITLE

"Investigation of system readiness when some defect types unknown"

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

Technical Report

5. AUTHOR(S) (First name, middle initial, last name)

Edwin R. Huber

John E. Walsh

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
September 28, 1970	22	4

8a. CONTRACT OR GRANT NO	9a. ORIGINATOR'S REPORT NUMBER(S)
N00014-68-A-0515	
b. PROJECT NO: NR 042-260	82
c.	9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)
d.	

10. DISTRIBUTION STATEMENT

This document has been approved for public release and sale; its distribution is unlimited. Reproduction in whole or in part is permitted for any purpose of the United States Government.

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY
	Office of Naval Research

13. ABSTRACT

A system consists of subsystems and performs satisfactorily when defects in subsystems do not cause its failure. For each subsystem, defect types are identified by their nature and by their level of probabilistic influence on system failure (finite number of levels). The subsystems and possible defect types are so defined that, for satisfactory system performance, a defect type can occur at most once in a subsystem. Also, for this conditional case, probabilities for a defect type are not influenced by occurence of other defect types. Moreover, the defect types are independent and have small probabilities with respect to occurrence. System ability is represented by the Readiness Index (RI), which is the probability of no defect that causes system failure. Statistical investigation of the RI is complicated by possible existence of defect types which have not yet been identified. Suitable data are available for each combination of subsystem and level of probabilistic influence on system failure. For every combination, the number of defect types occurring is observed over some repetitions. Unbiased estimation, also approximate tests and confidence intervals, are developed (some results are conservative and/or apply to at least moderately large RI values).

DD FORM 1 NOV 68 1473 (PAGE 1)

S/N 0101-807-6811

UNCLASSIFIED

Security Classification