Multiparameter Hypothesis Testing and Acceptance Sampling

Roger L. Berger
The Florida State University
Department of Statistics
Tallahassee, FL 32306

The quality of a product might be determined by several parameters, each of which must meet certain standards before the product is acceptable. In this article, a method of determining whether all the parameters meet their respective standards is proposed. The method consists of testing each parameter individually and deciding that the product is acceptable only if each parameter passes its test. This simple method has some optimal properties including attaining exactly a prespecified consumer's risk and uniformly minimizing the producer's risk. These results are obtained from more general hypothesis-testing results concerning null hypotheses consisting of the unions of sets.

KEY WORDS: Consumer's risk; Producer's risk; Multiple inference.

1. INTRODUCTION

In many situations the quality of a product is determined by several parameters. The product is of acceptable quality to the consumer only if each of the parameters meets certain standards. For example, an upholstery fabric must meet standards for strength, colorfastness, and fire resistance. Based on some measurements on the product, the consumer must decide whether the product is acceptable, that is, all of the parameters meet the standards, or unacceptable, that is, one or more of the parameters do not meet the standards. In making this decision the consumer wishes to use a rule that controls the consumer's risk at a small level.

If there is only one parameter and only one kind of measurement, then a standard quality control text such as Burr (1976) or Duncan (1974) gives methods for making this decision. Different methods are given depending on whether the parameter is a mean, variance, or proportion of defectives and on whether the measurements are counts of defective units (sampling by attributes) or measurements on a continuous variable (sampling by measurements). But no text that the author has found deals with the situation in which there are multiple parameters of interest.

This problem will be formulated as a hypothesis-testing problem in which the null hypothesis states that one or more of the parameters do not meet their standards and the alternative hypothesis states that all of the parameters do meet their standards. Note that rejection of the null hypothesis does not correspond to rejection of the product. Rather, rejection of the null hypothesis corresponds to the decision that the product is acceptable. With this formulation the probability of a Type I error will be the consumer's risk. An $\alpha$-level test will be one for which the consumer's risk is less than or equal to $\alpha$.

The test proposed herein is so simple it must not be new. But the author has not been able to find the test described in hypothesis-testing or quality control literature. The test is the following. A hypothesis test is done on each parameter individually at level $\alpha$. The overall test rejects the null hypothesis and decides that all of the parameters meet their standards if and only if each individual test decides that the individual parameter meets its standard.

This test has several interesting properties. First, the individual tests are performed at level $\alpha$ and yet the overall test has level $\alpha$. Usually when doing simultaneous inference about many parameters (see, e.g., Miller 1966), inferences about individual parameters must be done with an error rate of less than $\alpha$ to achieve an overall error rate of $\alpha$. This, for example, is the basis of the Bonferroni method of simultaneous confidence intervals. Second, under very mild conditions, the size of this test is exactly $\alpha$. So the test is not being too conservative by requiring each of the individual tests to decide that the individual parameter meets its standard. Third, under more restrictive conditions a result of Lehmann (1952) can be used to
prove that this test is uniformly most powerful in a reasonable class of tests. In terms of risks this says that this test uniformly minimizes the producer's risk. These properties indicate that not only is the test extremely easy to implement, since it deals with only one parameter at a time, but it also seems to be a reasonably good test.

2. BASIC RESULTS

Let \( X = (X_1, \ldots, X_n) \) be a random vector of observations whose distribution is determined by a vector parameter \( \theta = (\theta_1, \ldots, \theta_k) \). Let \( \Theta \) denote the parameter space. Let \( \Theta_i, i = 1, \ldots, k, \) be subsets of \( \Theta \). Let \( \Theta_0 = \bigcup_{i=1}^k \Theta_i \). Let \( A' \) denote the complement of the set \( A \). Note that \( \Theta_0 = \bigcap_{i=1}^k \Theta_i \). The problem to be considered is that of testing \( H_0: \theta \in \Theta_0 \) versus \( H_1: \theta \in \Theta_0 \). In the example in the introduction \( \Theta_i \) is the hypothesis that \( \theta_i \) does not meet its standard. If \( \theta_i \) must be greater than \( c_i \) to meet its standard, then \( \Theta_i = \{ \theta: \theta_i \leq c_i \} \). If \( \Theta_i \) must be between \( c_i \) and \( d_i \) to meet its standard, then \( \Theta_i = \{ \theta: c_i \leq \theta_i \leq d_i \} \). With this formulation, \( H_0 \) is the hypothesis that at least one parameter does not meet its standard and \( H_1 \) is the hypothesis that every parameter meets its standard. Note that \( k \), the number of subsets, may be less than \( \ell \), the number of parameters. This will be the case if some of the parameters are nuisance parameters and do not have standards associated with them.

Let \( \alpha, 0 \leq \alpha \leq 1 \), be fixed. For \( i = 1, \ldots, k \), let \( \psi_i(X) \) be an \( \alpha \)-level test of \( H_{0i}: \theta \in \Theta_i \) versus \( H_{1i}: \theta \in \Theta_i \), that is, \( E_{\theta} \psi_i(X) \leq \alpha \) for all \( \theta \in \Theta_i \). Let \( \Psi \) be the test of \( H_0 \) versus \( H_1 \) that rejects \( H_0 \) if and only if every \( \psi_i \) rejects \( H_{0i} \).

Other authors such as Birnbaum (1954, 1955), Lehmann (1952), and Spjutvoll (1972) have considered testing hypotheses \( H_0 \) and \( H_1 \). But in all these papers, except the one result of Lehmann to be discussed in Section 3.2, the null hypothesis is of the form \( H_1 \).

Tsutakawa and Hewett (1978) propose the test \( \Psi \) for a problem comparing regression lines. Tsutakawa and Hewett's test is discussed in Example 4.1. Wilkinson (1951) has proposed a test like \( \Psi \) but in a very different situation. He assumes that the individual tests are \( \alpha \)-level tests for all of \( H_0 \), not just \( H_{0i} \). Wilkinson also assumes the individual tests are independent, which typically will not be the case in the problems considered herein.

The facts that \( \Psi \) is always an \( \alpha \)-level test and under mild conditions has size exactly \( \alpha \) are stated in Theorems 1 and 2. Theorem 2 is proved in the Appendix.

**Theorem 1.** \( \Psi \) is an \( \alpha \)-level test of \( H_0 \) versus \( H_1 \), that is, \( E_{\theta} \Psi(X) \leq \alpha \) for all \( \theta \in \Theta_0 \).

**Proof of Theorem 1.** Let \( R_i \) be the event that \( \psi_i \) rejects \( H_{0i} \). Then \( R = \bigcap_{i=1}^k R_i \) is the event that \( \Psi \) rejects \( H_0 \). Fix \( \theta \in \Theta_0 \). Then \( \theta \in \Theta_i \) for some \( i \) and

\[ E_{\theta} \Psi(X) = P_{\theta}(R) \leq P_{\theta}(R_i) = E_{\theta} \psi_i(X) \leq \alpha \]

since \( \theta \in \Theta_i \) and \( \psi_i \) is an \( \alpha \)-level test of \( H_{0i} \).

**Theorem 2.** Suppose \( \Theta_i \) and \( \psi_i \), \( i = 1, \ldots, k-1 \), satisfy (i), (ii), and (iii).

(i) The standard for \( \theta_i \) is one-sided, that is, \( \Theta_i = \{ \theta: \theta_i \leq c_i \} \).

(ii) The power of \( \psi_i \) depends only on \( \theta_i \) and \( \theta_{i+1}, \ldots, \theta_k \).

(iii) There exist numbers \( b_i \) (possibly infinite) such that, for any fixed values of \( \theta_{k+1}, \ldots, \theta_k \), \( \lim_{\theta \to b_i} E_{\theta} \psi_i(X) = 1 \).

Suppose \( \Theta_i \) and \( \psi_i \) satisfy (iv).

(iv) There exist values \( \theta^* \), \( \theta_{k+1}^*, \ldots, \theta^*_k \) such that, for any values of \( \theta_1, \ldots, \theta_{k-1} \), the vector \( \theta^* = (\theta_1, \ldots, \theta_{k-1}, \theta^*_k, \ldots, \theta^*_k) \) satisfies \( E_{\theta^*} \psi_i(X) = \alpha \).

Then \( \Psi \) has size exactly \( \alpha \), that is, \( \sup_{\theta \in \Theta_0} E_{\theta} \Psi(X) = \alpha \).

Theorem 1 puts no restrictions on the form of the hypotheses \( \Theta_i \). They may be one-sided, two-sided, or any other form. The test \( \Psi \) is easily constructed and implemented since the individual tests \( \psi_i \) test only one hypothesis \( \Theta_i \) at a time. This simplicity and the generality of the result in Theorem 1 may be useful in some problems. But the test \( \Psi \) might be quite conservative in that the maximum consumer's risk (size) might be much less than \( \alpha \). This might put a heavy burden on the producer.

Theorem 2, on the other hand, gives conditions under which the test \( \Psi \) is not too conservative in that the maximum consumer's risk (size) is exactly \( \alpha \). Theorem 2 requires \( k - 1 \) of the standards to be one-sided, that is, \( \theta_i \) must be greater than \( c_i \) to meet its standard (recall \( \Theta_i \) is the set of parameters for which \( \theta_i \) does not meet its standard). If some of the standards are one-sided but in the opposite direction, that is, \( \theta_i \) must be less than \( c_i \) to meet its standard, then Condition (i) is still satisfied. The model may be parameterized in terms of \( \phi_i = -\theta_i \) and \( \Theta_i \) may be written as \( \{ \theta: \phi_i \leq -c_i \} \), which is the form for Condition (i). Many standard one-sided tests satisfy Conditions (ii) and (iii) of Theorem 2. These include:

1. one-tailed \( z \) test if \( \theta_i \) is a normal mean and the variance is known,
2. one-tailed \( t \) test if \( \theta_i \) is a normal mean and the variance is a nuisance parameter,
3. one-tailed binomial test if \( \theta_i \) is the proportion of nondefective items,
4. one-tailed \( \chi^2 \) test if \( \theta_i \) is a normal variance and the mean is a nuisance parameter,
5. nonparametric one-tailed tests such as
Some values of the power function of the test $\Psi$, which decides both parameters meet their standards if and only if $\psi_1$ and $\psi_2$ both reject, are listed in Table 1. The values above or left of the dotted line are values of the consumer's risk for various parameter values. The values of the power function below and to the right of the dotted line are equal to one minus the producer's risk for various parameter values. Large values of the power function correspond to small producer's risks and vice versa.

The following characteristics of the risks, which can be noted in Table 1, are typical of the more general situation in which all of the tests $\psi_1, \ldots, \psi_4$ are chosen from the list (2.1). (In the following remarks, terms like "much better" or "much worse" must be interpreted relative to the associated standard deviations.)

**Remark 3.1.** Only the marginal distributions of $X$ and $\bar{Y}$, not the joint distribution of $(X, \bar{Y})$, were considered in constructing $\psi_1$ and $\psi_2$. In general, only the marginal distribution of $\psi_i$ need be considered to insure that $\psi_i$ is an $\alpha$-level test of $\Theta_i$. The joint distribution of the test statistics need not be considered, or even known, to construct the test $\Psi$. This contributes to the simplicity of $\Psi$.

**Remark 3.2.** The maximum consumer's risk of $\alpha$ occurs when $\theta_1$ is much better than its standard and $\theta_2$ just fails to meet its standard or vice versa. In general, as the proof of Theorem 2 shows (see Appendix), the maximum consumer's risk occurs when all but one of the parameters are much better than their standards and the one remaining parameter just fails to meet its standard.

**Remark 3.3.** If either $\theta_1$ or $\theta_2$ is much worse than its standard, then the consumer's risk is near zero. In general, if at least one of the parameters is much worse than its standard, then the consumer's risk is near zero.

**Remark 3.4.** The maximum producer's risk occurs when both $\theta_1$ and $\theta_2$ are just slightly better than their standards. In general, the maximum producer's risk occurs when all of the parameters just slightly exceed their standards. If each $\theta_i$ is only slightly better than its standard, the probability that $\psi_i$ rejects $\Theta_i$ is at most slightly larger than $\alpha$. The probability that all the $\psi_i$ reject is at most slightly larger than $\alpha$ and is likely to be much less than $\alpha$. If the tests $\psi_i$ are all independent, then the maximum producer's risk is $1 - \alpha^k$.

**Remark 3.5.** The producer's risk is near zero only if $\theta_1$ and $\theta_2$ are both much better than their standards. In general, the producer's risk will be near zero only if all of the parameters are much better than their standards.
These observations indicate that the test \( \Psi \) provides very good protection for the producer and presents the consumer with a burden of proof. Remark 3.3 indicates that even if only one parameter is far below its standard, the consumer's risk will be near zero. On the other hand, all the parameters must far exceed their standards for producer's risk to be near zero (Remark 3.5). But this problem can be overcome by the use of large samples. As the sample size increases, the amount by which \( \theta_j \) must exceed its standard for the producer's risk to be acceptably small decreases. Remark 3.4 indicates that the producer's risk will be large if each parameter barely exceeds its standard. Yet the result discussed in Section 3.2 indicates that the producer's risk is as small as can be expected.

A test that is based on the Bonferroni Inequality and that controls the producer's risk at \( \alpha \) can be constructed in this way. Let \( H_{01} : \theta_1 > c_1 \) and \( \theta_2 > c_2 \). Each of the hypotheses, \( H_{01} : \theta_j > c_j \), could be tested at level \( \alpha/2 \). \( H_{01} \) would be rejected if either \( H_{01} \) were rejected. Then the producer's risk is

\[
P(\text{reject } H_0 | H_0) \leq P(\text{reject } H_{01} | H_0) + P(\text{reject } H_{02} | H_0) \leq \alpha.
\]

In this normal example, for \( \alpha = .05 \), \( H_{01} \) would be rejected if \( X < c_1 - 1.96\sigma_1/\sqrt{n} \) and similarly for \( H_{02} \). This test is related to the test \( \Psi \) and, in fact, Theorem 2 shows that the maximum consumer's risk is exactly \( 1 - \alpha/2 \). The test that decides that \( \theta_1 \leq c_1 \) if \( X < c_1 - 1.96\sigma_1/\sqrt{n} \) and that \( \theta_2 > c_1 \) if \( X > c_1 - 1.96\sigma_1/\sqrt{n} \) is a test of \( H_{01} : \theta_1 \leq c_1 \) versus \( H_{12} : \theta_1 > c_1 \) with size \( 1 - \alpha/2 \). The same holds true for the test based on \( Y \). Thus, by Theorem 2, the maximum consumer's risk is \( 1 - \alpha/2 \). In general, if the \( k \) individual tests in the Bonferroni test satisfy the conditions of Theorem 2, then the maximum consumer's risk is exactly \( 1 - \alpha/k \). One would choose between the Bonferroni test and \( \Psi \) depending on whether one wished to control the producer's risk or the consumer's risk.

### 3.2. An Optimality Result

Despite facts like Remarks 3.4 and 3.5, which indicate that the producer's risk may be high even if all the parameters exceed their standards, Theorem 4.2 of Lehmann (1952) indicates that the test \( \Psi \) uniformly minimizes the producer's risk in a reasonable class of tests. Lehmann's result will now be informally described.

Suppose all the standards are one-sided, that is, \( \Theta_i = \{ \theta_i \leq c_i \} \). Suppose each test \( \psi_i \) is of the form "reject \( \Theta_i \) if \( Y_i > d_i \)" for an appropriate test statistic \( Y_i \). Suppose each \( \psi_i \) has size exactly \( \alpha \) and satisfies (ii) and (iii) of Theorem 2. The joint distribution of \( (Y_1, \ldots, Y_n) \) must satisfy a condition that roughly requires that as all the parameters \( \theta_1, \ldots, \theta_n \) increase, all the test statistics tend to increase. A rejection region is monotone if \( (y_1, \ldots, y_n) \) is in the rejection region and all the coordinates of \( (z_1, \ldots, z_n) \) satisfy \( z_i \geq y_i \) imply \( z_i \geq z_i \) is in the rejection region. Then Lehmann's result asserts that among all level-\( \alpha \) tests with monotone rejection regions that are based on these test statistics, \( \Psi \) uniformly minimizes the producer's risk.

The conditions of Lehmann's theorem will be satisfied, for example, if all the tests \( \psi_i \) are from the list (2.1) and are independent. But the conditions will also be satisfied in many cases when the test statistics are correlated. For example, if each test either \( \theta_i \) is a location parameter for \( \gamma_i \) or \( \theta_i \) is a scale parameter for \( \gamma_i \), and \( \gamma_i \) is positive, then the conditions of Lehmann's theorem will be met, regardless of the correlations between the test statistics. Most dispersion tests fall into the scale parameter framework.

### 4. EXAMPLES

Two examples of the use of the test \( \Psi \) are presented in this section. In the first example, a regression problem examined by Tsutakawa and Hewett (1978) is considered. The test they propose is the test \( \Psi \). The second example considers an acceptance sampling problem from the textile industry.

**Example 4.1.** This problem was considered by Tsutakawa and Hewett (1978). Let \( Y_{ij} = \alpha_i + \beta_i X_{ij} + \epsilon_{ij} \), \( i = 1, 2; j = 1, \ldots, n_i \), \( \alpha_i, \beta_i, X_{ij}, \) and \( \epsilon_{ij} \) are unknown parameters; the \( X_{ij} \) are known constants; and the \( \epsilon_{ij} \) are independent identically distributed normal random variables with mean zero and variance \( \sigma^2 \). It is desired to compare the two regression lines \( \alpha_1 + \beta_1 X \) and \( \alpha_2 + \beta_2 X \) on the finite interval \( I = [X_*, X^*] \).

Specifically, we desire a test for \( H_0 : \alpha_1 + \beta_1 X \leq \alpha_2 + \beta_2 X \) for some \( X \in I \). Let \( \alpha_i = \alpha_i + \beta_i X_* - \alpha_2 - \beta_2 X_* \) and \( \theta_2 = \alpha_1 + \beta_1 X_* - \alpha_2 - \beta_2 X_* \). Then \( \theta_2 = 0 \) can be written as \( H_0 : \theta_1 \leq 0 \) or \( \theta_2 \leq 0 \) versus \( H_1 : \theta_1 > 0 \) and \( \theta_2 > 0 \). Letting \( \Theta_1 = \{ \theta_1 \leq 0 \} \) and \( \Theta_2 = \{ \theta_2 \leq 0 \} \), \( H_0 \) and \( H_1 \) are seen to be of the form considered in Theorems 1 and 2. To derive the test \( \Psi \), it is necessary to derive the tests \( \psi_1 \) for \( \Theta_1 \) and \( \psi_2 \) for \( \Theta_2 \). Let \( a_1, b_1, a_2, b_2 \) denote the least squares estimates of \( \alpha_1, \beta_1, \alpha_2, \) and \( \beta_2 \). Then the test statistic \( T_i = (a_1 + b_1 X_* - a_2 - b_2 X_*)/S_{\hat{\sigma}} \) has a \( t \) distribution with \( v = n_1 + n_2 - 4 \) degrees of freedom if \( \theta_i = 0 \), where

\[
S_i^2 = \sum_{j=1}^{n_i} \sum_{j=1}^{n} (Y_{ij} - a_i - b_i X_{ij})^2 / (n_1 + n_2 - 4)
\]

and

\[
\sigma_{11}^2 = \sum_{i=1}^{2} \sum_{j=1}^{n_i} (X_{ij} - X_*)^2 / \left( \sum_{j=1}^{n_i} (X_{ij} - X_*)^2 \right).
\]
(See Tsutakawa and Hewett 1978 for details of the derivation.) Thus a size $\alpha$ test of $\Theta_1$ versus $\Theta_0$ rejects $\Theta_1$ if $T_1 > t_{1-\alpha}(v)$, the $1 - \alpha$ percentile of the $t$ distribution with $v = n_1 + n_2 - 4$ degrees of freedom. This test is $\psi_1$. The test $\psi_2$, a size $\alpha$ test of $\Theta_2$ versus $\Theta_1$, rejects $\Theta_2$ if $T_2 > t_{1-\alpha}(v)$, where $T_2 = (a_1 + b_2X^* - a_2 - b_2X^*)/S\sqrt{\sigma_{22}}$ and $a_{22}$ is the same as $a_{11}$ except $X^*$ replaces $X$. The test $\Psi$ then rejects $H_0$: $\theta_1 \leq 0$ or $\theta_2 \leq 0$ and accepts $H_1$: $\theta_1 > 0$ and $\theta_2 > 0$ only if $T_1 > t_{1-\alpha}(v)$ and $T_2 > t_{1-\alpha}(v)$. This is the test proposed by Tsutakawa and Hewett. The fact that $\Psi$ has size exactly $\alpha$ follows from Theorem 2 since both $\psi_1$ and $\psi_2$ are of the form of 2 in (2.1). The Tsutakawa and Hewett test is an example of how the results of Theorems 1 and 2 yield a size $\alpha$ test when $H_0$ can be conveniently written as the union of two or more sets.

Example 4.2. An example of the use of the test $\Psi$ in an acceptance sampling situation will be taken from the textile industry. Table 2 lists specifications for upholstery fabric from the American Society for Testing and Materials. The specifications give standards for nine parameters related to strength, colorfastness, flammability, and dimensional stability.

The first three parameters might be assumed to be normal means; mean breaking strength, mean tear strength and mean abrasion resistance. Each standard says the mean must be greater than some value. Each of these standards could be tested with a one-sided $t$ test. For example, $\psi_1$ would be the size $\alpha$ test of $H_{01}$: $\theta_1 \leq 50$ versus $H_{11}$: $\theta_1 > 50$.

The next five parameters, colorfastness and flammability, might be measured by binomial variables, each variable counting the number of units in a sample that pass the corresponding test. Each parameter would then be the proportion of units in the population (a particular manufacturer’s output) that achieve one of the standards. The usual one-sided binomial test could be used to test each of these five hypotheses. For example, suppose $\theta_4$ is the proportion of the manufacturer’s output that will attain a grade of class four or above in the water colorfastness test. Suppose that at least 90 percent of the output must attain this grade for the upholstery to be acceptable. Then $X_4$ would be the binomial variable that counts the number of units in a sample that attain a class four grade or better. $\psi_4$ would be the size $\alpha$ binomial test of $H_{04}$: $\theta_4 \leq .90$ versus $H_{04}$: $\theta_4 \leq .90$ versus $H_{14}$: $\theta_4 > .90$ based on $X_4$.

The last parameter $\theta_9$ might be assumed to be a normal mean, mean percentage change. The standard is two-sided. Thus $\psi_9$ must be a size $\alpha$ test of $H_{09}$: $\theta_9 \geq .02$ or $\theta_9 \leq -.05$ versus $H_{19}$: $-.05 < \theta_9 < .02$. If an upper bound for the variance can be assumed, say $\sigma^2 \leq \sigma^2_0$, then $\psi_9$ can be constructed in this way. Let $X_1, \ldots, X_m$ be independent identically distributed normal random variables with mean $\theta$ and variance $\sigma^2$. Assume $\sigma^2 \leq \sigma^2_0$. Let $\bar{X}$ and $S^2$ denote the sample mean and variance. Consider testing $H_0$: $\theta \geq d$ or $\theta \leq c$ versus $H_1$: $c < \theta < d$. Let $T = (\bar{X} - \theta_0)/(S/\sqrt{m})$, where $\theta_0 = (c + d)/2$. $T$ has a noncentral $t$ distribution with noncentrality parameter $\delta = \sqrt{m(\theta - \theta_0)/\sigma}$ and $m - 1$ degrees of freedom. Let $\delta_0 = \sqrt{m(d - c)/2d_0}$. Let $a$ be the number that satisfies $P(-a < T \leq a) = \alpha$, where $T_0$ has a noncentral $t$ distribution with noncentrality parameter $\delta_0$ and $m - 1$ degrees of freedom. Then the test that rejects $H_0$ if $-a < T < a$ is a size $\alpha$ test of $H_0$ versus $H_1$. This is true since $P(-a < T < a) = \alpha$ is a decreasing function of $|\delta|$ and $\delta_0$ is the smallest possible value of $|\delta|$ under $H_0$. In our example, $d = .02, c = -.05$. Assume $m = 9$ and $\sigma = .05$. Then $\delta_0 = 2.1$. If $\alpha = .10$ is used, then the value of $a$ satisfying $P(-a < T_2 < a) = .10$ is $a = .8300$. Thus, the test $\psi_9$ rejects $H_{09}$ if $-.8300 < T < .8300$. If $\alpha = .05$ is used, then $a = .5055$ and $\psi_9$ rejects $H_{09}$ if $-.5055 < T < .5055$. These values of $a$ were computed using the program MDTN of the IMSL library to compute noncentral $t$ probabilities. The values of $a$ could be approximated using tables of noncentral $t$ probabilities such as those of Locks, Alexander, and Byars (1963).

Since the first eight tests are from the list (2.1) and the ninth test has the probability of a Type I error, exactly $\alpha$ when $\theta_9 = .02$ and $\sigma = .05$, Theorem 2 guarantees that the test $\Psi$ obtained by combining these nine tests has size exactly $\alpha$. It would be very difficult to posit a realistic multivariate model for the observations on these variables. Some are discrete and some are continuous. Some are likely to be correlated. Yet it is relatively easy to construct a size $\alpha$ test for

Table 2. Standard Specification for Woven Upholstery Fabric—Plain, Tufted, or Flocked

<table>
<thead>
<tr>
<th>Test</th>
<th>Minimum Standard</th>
</tr>
</thead>
<tbody>
<tr>
<td>Breaking Strength</td>
<td>50 pounds</td>
</tr>
<tr>
<td>Tensile Tear Strength</td>
<td>6 pounds</td>
</tr>
<tr>
<td>Surface Abrasion (heavy duty)</td>
<td>15,000 cycles</td>
</tr>
<tr>
<td>Colorfastness to:</td>
<td></td>
</tr>
<tr>
<td>Water</td>
<td>class 1</td>
</tr>
<tr>
<td>Croaking</td>
<td>class 1</td>
</tr>
<tr>
<td>Dry</td>
<td>class 1</td>
</tr>
<tr>
<td>Wet</td>
<td>class 3</td>
</tr>
<tr>
<td>Light-Salt AATCC Fading Units</td>
<td>class 4</td>
</tr>
<tr>
<td>Flammability</td>
<td>Pass</td>
</tr>
<tr>
<td>Dimensional Change</td>
<td>5% shrinkage, .2% gain</td>
</tr>
</tbody>
</table>

each parameter individually. These nine tests can be combined into the overall test $\Psi$ to test the hypothesis $H_0$ with a maximum consumer's risk of exactly $z$. The upholstery will be deemed satisfactory only if each of the nine tests rejects its hypothesis at level $z$.

5. CONCLUSIONS

Acceptance sampling procedures for individual parameters are well known. This article proposes a way of combining these procedures in the situation in which the quality of a product is measured by standards on several parameters. Not only is the method easy to implement, but it controls the consumer's risk at exactly a preassigned level in typical situations when the standards are one-sided (either upper or lower bounds). Under slightly more restrictive conditions, this method also uniformly minimizes the producer's risk. This method can be used in hypothesis-testing problems other than acceptance sampling if the null hypothesis is a union of sets.

ACKNOWLEDGMENT

This research was supported by the U.S. Army Research Office Grant No. DAAG 29-79-C-0158.

The values in Table 1 were computed using a computer program written by Professor Charles E. McCulloch, Florida State University. The author gratefully acknowledges this assistance. The author also wishes to thank the referees and editor for comments that improved the presentation of this article.

APPENDIX

Proof of Theorem 2. Let $R_i$ and $R$ be defined as in the proof of Theorem 1. Let $\theta_i = (\theta_{1i}, \ldots, \theta_{ki-1}, \theta^*_{ki}, \theta^*_{k+1i}, \ldots, \theta^*_{pi})$, $i = 1, 2, \ldots$ be a sequence of parameter points satisfying $\theta_{ij} \rightarrow b_j$, $j = 1, \ldots, k - 1$, as $i \rightarrow \infty$. Then $\theta_i \epsilon \Theta$ for all $i$ and $P_{\theta_i}(R_i) = 1 - E_{\theta_i} \psi_i(X) = 1 - z$ for all $i$. Also, for $j = 1, \ldots, k - 1$, $\lim_{i \rightarrow \infty} P_{\theta_i}(R_j) = 1 - \lim_{i \rightarrow \infty} P_{\theta_i}(R_j) = 1 - 1 = 0$. Therefore,

$$\sup_{\theta \epsilon \Theta} E_{\theta} \psi(X) \geq \lim_{i \rightarrow \infty} E_{\theta_i} \psi(X)$$

$$\lim_{i \rightarrow \infty} P_{\theta_i} \left( \bigcap_{j=1}^{k} R_j \right) = \lim_{i \rightarrow \infty} \left( 1 - P_{\theta_i} \left( \bigcup_{j=1}^{k} R_j \right) \right) \geq 1 - \lim_{i \rightarrow \infty} \sum_{j=1}^{k} P_{\theta_i}(R_j)$$

$$= 1 - \lim_{i \rightarrow \infty} \left( (1 - z) + \sum_{j=1}^{k-1} P_{\theta_i}(R_j) \right)$$

$$= 1 - (1 - z) - 0 = z.$$

From Theorem 1, $\sup_{\theta \epsilon \Theta} E_{\theta} \psi(X) \leq z$. Thus the size of $\Psi$ is exactly $z$.

[Received October 1980. Revised February 1982.]

REFERENCES


LOCKS, M. O., ALEXANDER, M. J., and BYARS, B. J. (1963), New Tables of the Noncentral t Distribution, Canoga Park, Calif.: Rocketdyne, A Division of North American Aviation.


