USE OF THE LARGEST NORMAL ORDER STATISTIC IN TESTING FOR INDUCIBLE ENZYME PRODUCTION

by

D. S. Robson

BU-815-M

May, 1983

Abstract

In circumstances where experimental error variance is inflated by the treatment as compared to a smaller variance among controls, perhaps due to uncontrolled variation in the intensity of the treatment application to the different experimental units, a loss of power may result if the within-treatment variance is utilized for purposes of testing the difference between treatment and control means. If the within-treatment distribution is highly skewed or even bimodal as compared to a normal distribution for the controls, then even the use of the treatment mean as a summary statistic is a dubious practice.

An extreme example of this nature arose in a consulting problem wherein the "treatment" data suggested that response was a sometimes thing, resulting in a seeming bimodal mixture of white noise responses and real, positive responses, while the control consistently produced only white noise responses. The largest of the n replicate observations on the treatment, Studentized by the control standard deviation, proved to be a better test statistic than the similarly Studentized treatment mean in this circumstance. A comparison of power against a bimodal mixture alternative to the null hypothesis revealed a dramatic difference in the power curves. If the control mean as well as the control standard deviation requires estimation then the order statistic test protocol calls for Dunnett's one-tailed critical values for comparing n treatments with a replicated control.

Paper No. BU-815-M in the Biometrics Unit, Cornell University.

USE OF THE LARGEST NORMAL ORDER STATISTIC IN TESTING FOR INDUCIBLE ENZYME PRODUCTION

by

D. S. Robson

BU-815-M May, 1983

Field isolates of a particular species of plant fungus were screened for presence or absence of the ability to produce a specific inducible enzyme used by fungi to detoxify the chemical defenses of host plants. For purposes of screening large numbers of isolates a quick assay for the enzyme was developed but proved to yield quite variable results in repeated trials on the same fungus isolate. In a technique study using isolates which had been pretested by a slow but accurate assay method it was found that the quick assay method reliably produced random normal deviates with constant variance when applied to control isolates known to lack the enzyme in question. When applied to isolates which were known producers of the enzyme, independently repeated trials of the quick assay generated data resembling control data but containing positive outliers; i.e., the data resembled that simulated by a "contaminated normal" distribution of the mixture type, as if the quick assay method frequently failed to induce the enzyme. Several quick assays of the same isolate ($n = 3$, 4 or 5) were thus often found to be needed to detect enzyme inducibility.

In order to circumvent the power-deflating effect of this outlier-caused large variance among replicate assays of the same isolate, only the variance among control replicates was used in constructing a statistical test for presence of the enzyme. The quick assay provided an estimate of the rate of enzyme activity; the response metameter was the slope (b) of a 4-point regression line, and in the absence of enzyme this slope was a normally distributed estimator of zero with a

---

Paper No. BU-815-M in the Biometrics Unit, Cornell University.

standard deviation s_c estimated with ν degrees of freedom from the empirical variance among control replicates. Since this H_o-distribution has a known mean, $\beta_0 = 0$, the sample mean of the controls was not needed for comparison with the slope of a test isolate ($\bar{b}_0$ was slightly but not significantly less than zero).

The statistical design and analysis adopted for the screening program consisted of performing several (n) replicate quick assays on each test isolate to obtain independent slope estimates $b_1, \cdots, b_n$ and then comparing the largest of these, $b_{[n]}/s_c = \max(b_1, \cdots, b_n)/s_c$, with the critical value of the largest Studentized normal order statistic, as tabulated by Pillai and Ramachandran (1954). The more conventional Studentized mean $\sqrt{n}\,\bar{b}/s_c$ when compared to the one-tailed critical value of the t_ν-distribution failed too frequently to detect enzyme in isolates which the prolonged assay had previously shown to be enzyme producers, while the Studentized order statistic test rarely failed to agree with the conclusions of the slow assay.

Comparison of power curves against "contaminated normal" mixtures provides some insights into the reason for the relatively better performance of the Studentized order statistic versus the Studentized mean test, and also reveals some anomalies of both procedures. Power comparisons are most easily implemented for the limiting case $\nu \to \infty$, which avoids the nuisance of estimation error in the standard deviation of the controls (s_c). Thus if we assume that for an enzyme-producing strain the distribution $F_b(x)$ of slope estimates is a mixture of two normals

$$H_a: \quad F_b(x) = p\Phi(x) + (1-p)\Phi\left(\frac{x-\beta}{\sigma}\right)$$

one with mean zero and unit standard deviation ($\sigma_c = 1$) and the other with mean $\beta > 0$ and a standard deviation of σ units, then the null hypothesis constitutes the boundary where $q = 1 - p = 0$. Letting $z_{1-\alpha}^{[n]} = \Phi^{-1}\left((1-\alpha)^{1/n}\right)$ where $\Phi(\cdot)$ denotes the standard normal distribution function, then the power of the statistical test

which rejects H_o when $b_{[n]} > z_{1-\alpha}^{[n]}$ is

$$P_{H_a}\left\{b_{[n]} > z_{1-\alpha}^{[n]}\right\} = 1 - \left[p(1-\alpha)^{1/n} + q\Phi\left(\frac{z_{1-\alpha}^{[n]} - \beta}{\sigma}\right)\right]^n$$

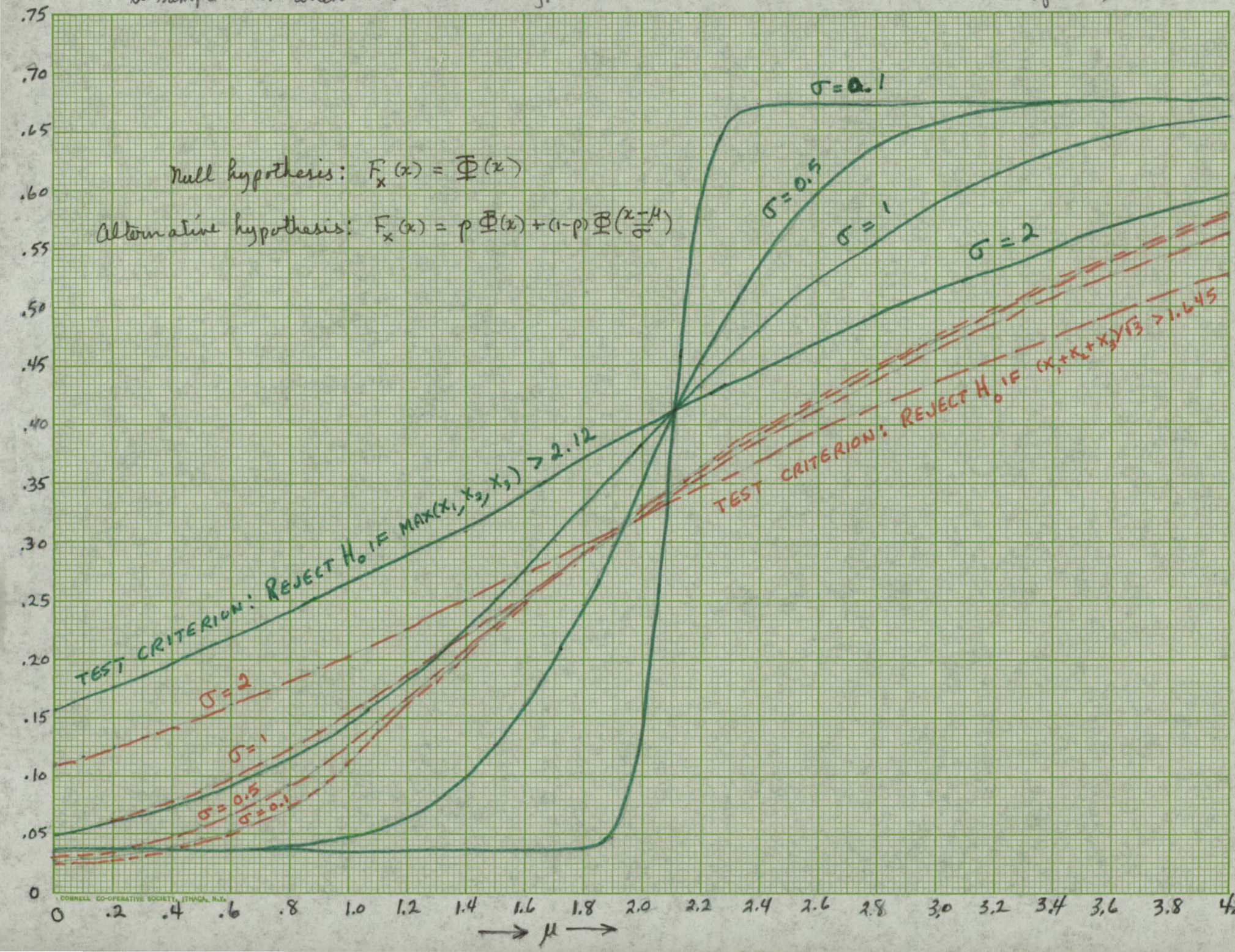
while the power of the mean test is

$$P_{H_a}\left\{\bar{b}\sqrt{n} > z_{1-\alpha}^{[1]}\right\} = 1 - \sum_{k=0}^{n}\binom{n}{k}p^k q^{n-k}\Phi\left((z_{1-\alpha}^{[1]} - \frac{n-k}{\sqrt{n}}\beta)/\sqrt{\frac{k}{n} + (1-\frac{k}{n})\sigma^2}\right).$$

Some graphs of these power curves are illustrated in Figure 1 for the case $n = 3$. Note that as $\beta \to +\infty$ the power of both tests against this mixture alternative approaches $1 - (1-\alpha)p^n$, irrespective of σ. A striking characteristic of the order statistic power curves is their common intersection at $\beta = z_{1-\alpha}^{[n]}$, where the power is $1 - [p(1-\alpha)^{1/n} + \frac{1}{2}q]^n$, irrespective of σ, so that as the ratio (σ) of "induced" standard deviation to control standard deviation approaches zero this sigmoidal power curve approaches a discrete step at $\beta = z_{1-\alpha}^{[n]}$, jumping directly from its lower limit $1 - [p(1-\alpha)^{1/n} + q\Phi(z_{1-\alpha}^{[n]}/\sigma)]^n$ to its upper limit $1 - p^n(1-\alpha)$. For β less than their respective critical values the behavior of both tests is somewhat anomalous for any value of σ other than $\sigma = 1$. Near $\beta = 0$ both tests are sensitive to this heteroscedasticity parameter σ, tending to reject H_o if σ is large and becoming biased if σ is small.

In the context of the fungus screening problem, where large β would be indicative of high enzyme levels, false negatives near $\beta = 0$ were of little concern since the screening objective was to identify isolates having the higher levels. The phenomenon of false positives at $\beta = 0$ due to $\sigma \gg 1$ was not expected to arise in this context because of an empirical relationship between mean and variance; thus, the anomalous features of the power curves of the Studentized largest order statistic could here be safely ignored while exploiting the better performance of this test statistic at higher enzyme levels.

Figure 1. Power curves of normal-theory one-tailed tests based on the largest order statistic versus tests based on the sample mean when the alternative hypothesis is a "contaminated" normal mixture ($p = .7$).



In other contexts where the control mean is not specified a priori but must be estimated by the sample mean, say $\bar{y}_c$ with estimated standard error $s_c/\sqrt{n_c}$, then it would become necessary to use the critical value of $(Y_{[n]} - \bar{y}_c)/s_c\sqrt{1 + \frac{1}{n_c}}$ as the Studentized largest order statistic of n equi-correlated normal deviates ($\rho = 1/(n_c + 1)$; see Dunnett, 1964). Under the mixture alternative hypothesis with μ and σ denoting the mean and standard deviation of the second normal component, the power of this test becomes

$$P_{H_a}\left\{\frac{Y_{[n]} - \bar{y}_c}{s_c\sqrt{1+\frac{1}{n_c}}} > t\right\} = 1 - E_{H_o}\left[p\Phi\left(\frac{\bar{y}_c - \mu_c + ts_c\sqrt{1+\frac{1}{n_c}}}{\sigma_c}\right) + q\Phi\left(\frac{\bar{y}_c - \mu + ts_c\sqrt{1+\frac{1}{n_c}}}{\sigma}\right)\right]^n$$

where t is the appropriate critical value and E_{H_o} denotes expectation with respect to the H_o-distribution of $(\bar{y}_c, s_c)$. As n_c gets large the graphs of this function would approach the corresponding curves in Figure 1.

References

K. C. S. Pillai and K. V. Ramachandran (1954), "On the distribution of the ratio of the i'th observation in an ordered sample from a normal population to an independent estimate of the standard deviation", Annals Math. Stat., Vol. 25, pp. 565-572.

C. W. Dunnett (1964), "New tables for multiple comparisons with a control." Biometrics 20, 482-491.