Statistical Bulletin
Reliability & Variation Research

DETROIT RESEARCH INSTITUTE
P.O. BOX 36504 • GROSSE POINTE, MICHIGAN 48236 • (313) 886-7976

LEONARD G. JOHNSON
EDITOR

WANG H. YEE
DIRECTOR
Volume 17
Bulletin 4

August , 1987

THE CULMINATED ENTROPY METHOD OF ANALYZING DESIGNED EXPERIMENTS

INTRODUCTION

The present day emphasis on cost reductions and quality improvements for purposes of competitive advantage and efficiency have made it imperative that we use scientific methods of empirical experimentation on factors which might contribute to improved performance of manufactured products. Consequently, such techniques as complete and fractional factorial designs in experimentation have become much more widely used in industry than ever before.

Once such experiments are designed and run we are faced with the problem of analyzing the resulting data on the response variable at different levels of the input factors.

In the present bulletin we shall introduce the Culminated Entropy Method of Analysis of response data at different levels of a factor. This amounts to a comparison problem between two average entropies. It will be seen that this technique is much more effective than the conventional analysis of variance in demonstrating the significance of effects.

EXPERIMENTAL DESIGN EXAMPLE

(3 FACTORS AT 2 LEVELS EACH)

(COMPLETE FACTORIAL)

The numbers inside the squares are response values , such as hours to failure, for different combinations of levels of factors A , B , and C.

	A_1		A_2	
	C_1	C_2	C_1	C_2
B_1	90	200	155	275
B_2	115	150	200	240

EVALUATING THE SIGNIFICANCE OF FACTOR A

Factor A has two levels : A_1 and A_2 thus, two data sets are included in the experiment for evaluating the effect of changing levels of Factor A . These two data sets are the following (after putting the responses into numerical order for each level of Factor A) :

DATA SET # 1 (FACTOR A AT LEVEL 1)	DATA SET # 2 (FACTOR A AT LEVEL 2)
90	155
115	200
150	240
200	275

These two data sets are two samples (each of size 4) which can now be compared by the Culminated Entropy Technique. The first thing to do is to plot data set #1 on some type of probability paper, say Weibull paper in case the responses are life test results, i.e., times to failure . Using Median Rank plotting (Benard's Formula) we obtain the results shown on the next page .

ANALYSIS DETAILS FOR FACTOR A

SAMPLE # 1 (N = 4)

Order No. = j	Response Value	Med. Rank = j - .3/N + .4
1	90	.1591
2	115	.3864
3	150	.6136
4	200	.8409

Plotting response values as abscissas on Weibull paper together with median ranks as ordinates we obtain Figure 1 below :

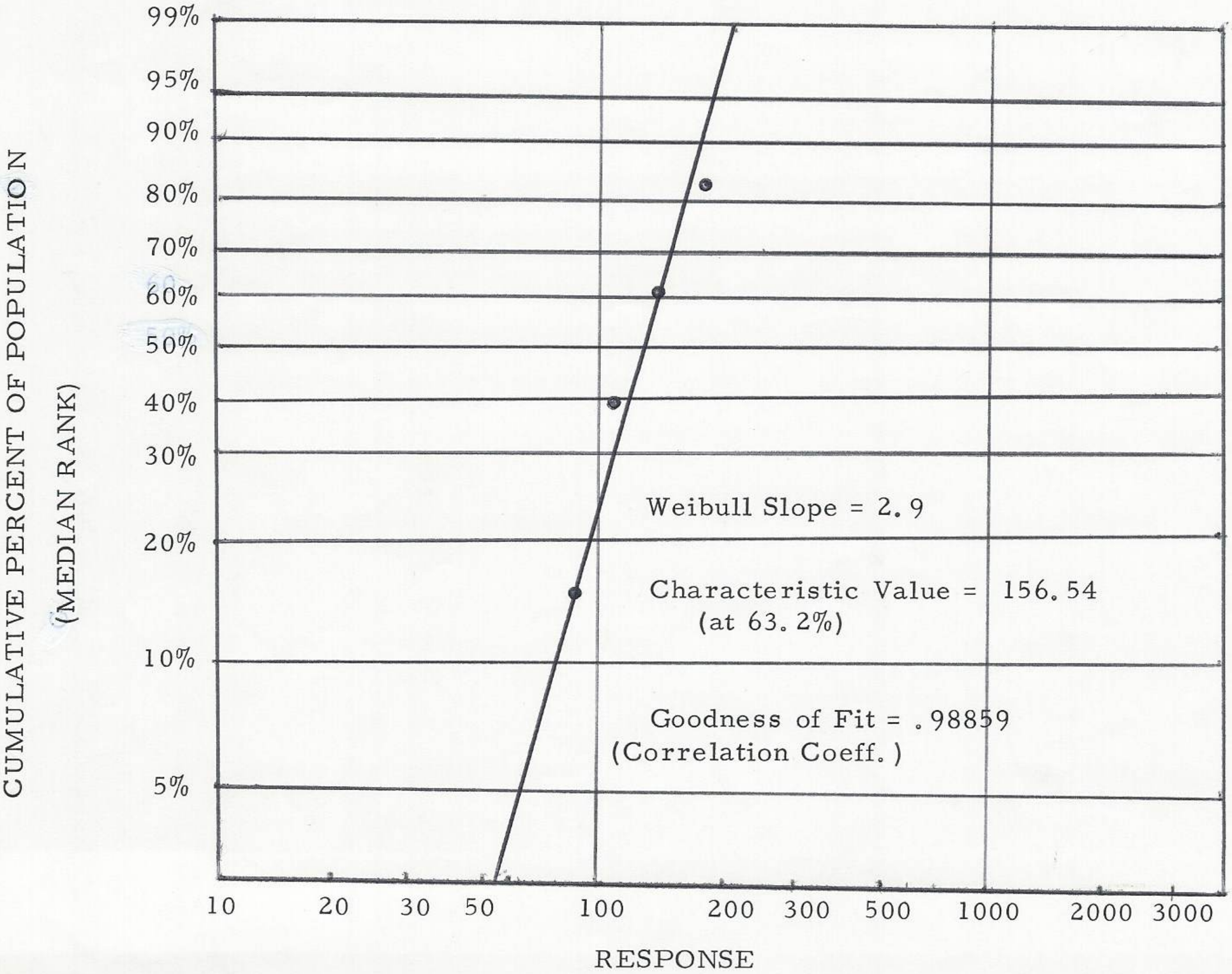

FIGURE 1

We take the Weibull line of Level 1 as our standard reference. Then for any response x the Entropy in the Weibull line of Level 1 is $(x/\theta_1)^{b_1}$,

where b_1 = Weibull slope of Level 1 = 2.9

and θ_1 = Characteristic Value of Level 1 = 156.54 (at 63.6%)

so , the 4 responses of Level 2 , i.e., 150, 200, 240, 275 have (in Level 1's Weibull line) an Entropy total (Culminated Entropy) of

$$(150/156.54)^{2.9} + (200/156.54)^{2.9} + (240/156.54)^{2.9} + (275/156.54)^{2.9} = 11.49615 .$$

If Level 2 had been in the same population as Level 1 (i.e., no different than Level 1) then the Culminated Entropy for Level 2 in Level 1's Weibull line would have been only 4 (same as the sample size). Here, however , the Culminated Entropy is way up to 11.49615 .

The standard deviation of an Entropy total of N items is $\sqrt{N}$. Hence, in this case , the standard deviation of the Entropy total (Culminated Entropy) is $\sqrt{4} = 2$.

To compare Level 2 with Level 1 we calculate the so-called t-score to Coincidence defined as

$$\frac{\mathcal{E}_2 - \mathcal{E}_1}{\sigma_1 + \sigma_2}$$

where $\mathcal{E}_1$ = Culminated Entropy for Level 1

$\mathcal{E}_2$ = Culminated Entropy for Level 2

σ_1 = Standard Deviation of Culminated Entropy (Level 1)

σ_2 = Standard Deviation of Culminated Entropy (Level 2)

In the example we have

$$\mathcal{E}_1 = 4 \qquad \sigma_1 = \sqrt{4} = 2$$

$$\mathcal{E}_2 = 11.49615 \qquad \sigma_2 = \sqrt{4} = 2$$

Hence, the t-score to Coincidence is 11.49615 - 4/(2+2) = 7.49615/4 = 1.87404

The confidence of a real difference between Level 1 and 2 is then

$$C = \frac{1}{1 + e^{-\frac{\pi}{\sqrt{3}} k\hat{t}}}$$

where

$$k = \sqrt{1 + \frac{\sqrt{N_1 N_2}}{\frac{1}{2}(N_1 + N_2)}} = \sqrt{1 + \frac{\sqrt{16}}{\frac{1}{2}(4 + 4)}} = \sqrt{2} .$$

(N_1 and N_2 are the sample sizes of Level 1 and 2)

Thus,

$$C = 1/(1 + e^{-1.8138\sqrt{2}(1.87404)}) = .99189$$

So, the A effect is significant enough to yield over 99% confidence of a real change from Level 1 to Level 2.

CONCLUSION

From the example we see how the confidence of a real change in response at Level 2 (compared to Level 1) is readily calculated for any factor simply by using the Culminated Entropies of Level 2's responses in the distribution function estimated for Level 1. This is much more convenient and more easily interpretable than the Analysis of Variance. Furthermore, the Analysis of Variance requires Normal distributions. Entropy for any response can be calculated in any distribution having a cumulative distributin function F(x) simply by evaluating the general formula for Entropy at response x, which is $\mathcal{E}(x) = \ln(1/1 - F(x))$.

In the case of a Weibull:

$$F(x) = 1 - e^{-(x/\theta)^b}$$

so,

$$\mathcal{E}(x) = \ln \frac{1}{e^{-(x/\theta)^b}} = \ln(e^{(x/\theta)^b}) = (x/\theta)^b .$$

NOTE: If any response x_i in a life test is a suspended item (unfailed), then the Culminated Entropy for x_i is $\mathcal{E}(x_i) = (x_i/\theta)^b + 1$ ← (always add 1 for each suspended item)