STATISTICAL BULLETIN
Reliability & Variation Research

LEONARD G. JOHNSON
EDITOR

DETROIT RESEARCH INSTITUTE
P.O. Box 36504 • Grosse Pointe, MI 48236 • (313) 886-8435

WANG H. YEE
DIRECTOR

Volume 22
Bulletin 4

August , 1992
Page 1

LOGISTIC SPECIFICATION STATISTICS:

A COMPUTERIZED LOGISTIC THEORY FOR PERCENT OUT OF SPEC WHEN CONSIDERATION MUST BE TAKEN OF UPPER AND LOWER SPECIFICATION LIMITS, AS WELL AS OF OF NOMINAL VALUES AND THE SHIFTING OF DISTRIBUTION MEANS FROM THE NOMINAL VALUES.

INTRODUCTION

In this bulletin we are proposing a brand new theoretical outlook on unilateral and bilateral dimensional tolerances by using the **Logistic Distribution** as the one to which measured data are fitted. The graphical plots on **Semi-Log** paper will give us the evidence of a good fit to measured values, and how they are actually distributed. In other words, we will **tell the truth** with actual statistics on the measured values. We claim that this **Logistic Approach** of ours is every bit as valid as the **Classical Normal Distribution Theory** employed in **Statistical Quality Control.**

The final proof of the validilty of the entire theory will come out of the **Goodness of Fit** we obtain in plotting actual data with a **Logistic Cumulative Distribution Function** obtained by plotting measured values as abscissas on a **Linear Scale, and Odds Against Exceedance** as ordinates on a **Logarithmic Scale.** The computerization of this theory is so much more straightforward and simple, because **Table Look-Ups of Normal Curve Areas** are eliminated by having a direct and simple mathematical formula for **Cumulative Area in a Logistic Distribution Function.** It's about time we quit considering the normal distribution to be a sacred image from which it is an evil act to depart.

THE MATHEMATICS OF THE LOGISTIC DISTRIBUTION

DEFINITION: If a variable x has a *Logistic Distribution* with a *Mean* **M** and a *Standard Deviation* **s,** then the cumulative fraction F(x) of the variable accounted for at any value x is given by the formula

$$F(x) = \frac{1}{1 + \exp\{-(\pi/\sqrt{3})[(x - M)/s]\}}$$

If we standardize this distribution by defining Z = (x - M)/s, then this so-called *Z-Score* will have a cumulative distribution function whose formula is

$$F(Z) = \frac{1}{1 + \exp\{-(\pi/\sqrt{3})(Z)\}}$$

By taking the *DERIVATIVE* with respect to Z we obtain the *FREQUENCY CURVE (*or Probability Density Function) f(Z), defined by

$$f(Z) = \frac{\pi/\sqrt{3}}{\{\exp(\pi Z/2\sqrt{3}) + \exp(-\pi Z/2\sqrt{3})\}^2}$$

Now we can evaluate both the *Frequency Curve* (for heights) and the *Cumulative Distribution Function* (for Cumulative Area), by taking Z values from -3 to +3 at interval of .1, to obtain Table 1 and Figures 1 and 2.

TABLE 1

PROPERTIES OF THE LOGISTIC DISTRIBUTION

Z-SCORE	PDF HEIGHT	CUM. AREA
-3	.007797	.0043147
-2.9	.009331	.0051683
-2.8	.011163	.0061898
-2.7	.01335	.0074116
-2.6	.015958	.0088725
-2.5	.019064	.0106183
-2.4	.022759	.0127031
-2.3	.027147	.015191
-2.2	.032349	.0181572
-2.1	.038503	.0216899
-2	.045763	.0258917
-1.9	.054303	.0308818
-1.8	.064308	.0367973
-1.7	.075979	.0437947
-1.6	.089521	.0520508
-1.5	.105134	.0617629
-1.4	.122999	.0731472
-1.3	.143258	8.643669E-02
-1.2	.165988	.1018752
-1.1	.191169	.1197099
-1	.218647	.1401795
-.9	.248098	.1634991
-.8	.278993	.1898416
-.7	.310576	.2193154
-.6	.341862	.2519423
-.5	.371665	.2876346
-.4	.398659	.3261786
-.3	.42148	.367225
-.2	.438856	.4102916
-.1	.449741	.4547789
0	.45345	.5
.1	.449741	.5452211
.2	.438856	.5897084
.3	.42148	.6327749
.4	.398659	.6738214
.5	.371665	.7123654
.6	.341862	.7480576
.7	.310576	.7806845
.8	.278993	.8101584
.9	.248098	.8365009
1	.218647	.8598205
1.1	.191169	.8802901
1.2	.165988	.8981247
1.3	.143258	.9135632

TABLE 1 - Continued

PROPERTIES OF THE LOGISTIC DISTRIBUTION

1.4	.122999	.9268527
1.5	.105134	.938237
1.6	.089521	.9479491
1.7	.075979	.9562052
1.8	.064308	.9632027
1.9	.054303	.9691182
2	.045763	.9741082
2.1	.038503	.9783101
2.2	.032349	.9818427
2.3	.027147	.9848088
2.4	.022759	.9872968
2.5	.019064	.9893816
2.6	.015958	.9911274
2.7	.01335	.9925883
2.8	.011163	.9938102
2.9	.009331	.9948317
3	.007797	.9956852

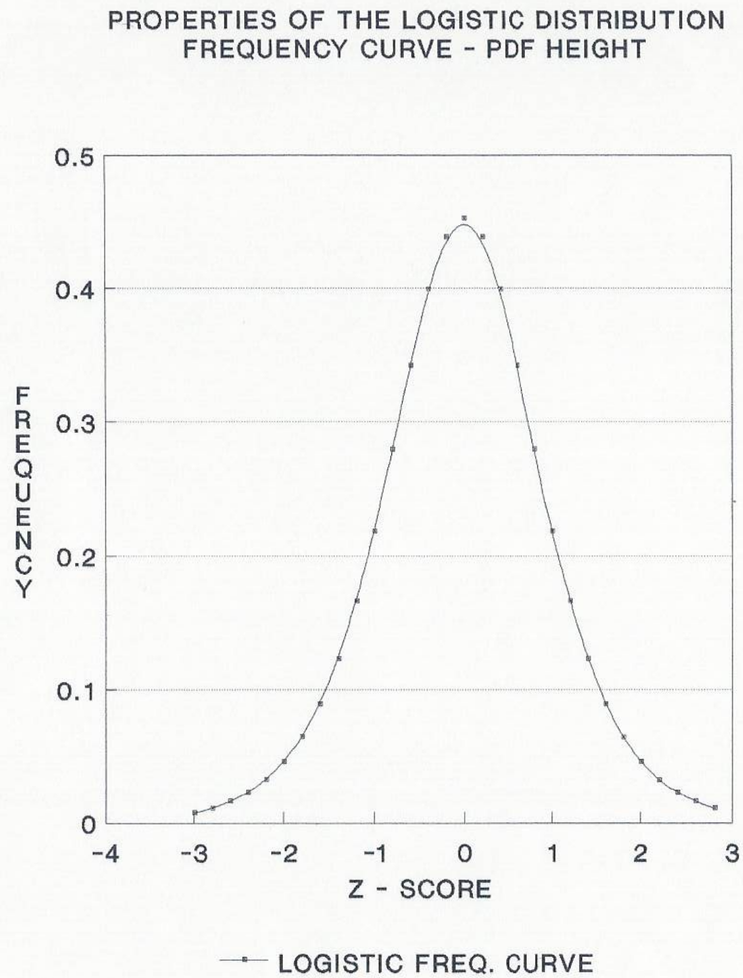

FIGURE 1

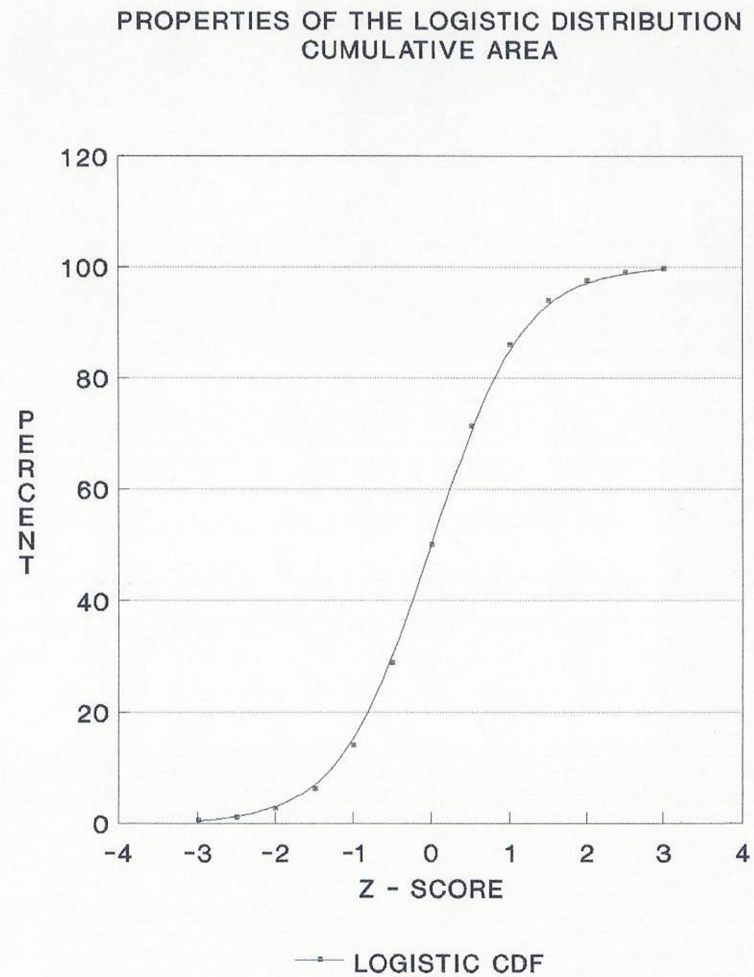

FIGURE 2

FIVE FACTOR SPEC ANALYSIS

The five factors of input into our specification compliance problem are

FACTOR #1 : The Lower Specification Limit (LSL)

FACTOR #2 : The Upper Specification Limit (USL)

FACTOR #3 : The Nominal Value (NV)

FACTOR #4 : The Logistic Slope (B)

FACTOR #5 : The Logistic Mean (M)

From these five factors we immediately calculate that the distribution mean is shifted by the amount $S = M - NV$.

Furthermore, the distribution SIGMA (i.e., Standard Deviation) is

$$s = SIGMA = (\pi/\sqrt{3})/(\text{Distribution Slope}) ,$$

and the Distribution Mean is where the **ODDS AGAINST EXCEEDANCE** are **EVEN** (i.e., 1 to 1) .

The Distribution Slope and Distribution Mean are determined from a **LEAST SQUARES FIT on SEMI-LOG paper** with measurements as abscissas on a **Linear Scale** and values of **F(x)/(1 - F(x)** as ordinates on a **Log Scale.**

NUMERICAL EXAMPLE

RAW DATA -----TABLE # 2

ACTUAL MEASURED VALUES	SPEC LIMITS
1.25	
1.27	Lower Spec = 1.265
1.31	
1.35	Nominal Value = 1.365
1.36	
1.39	Upper Spec = 1.465
1.43	
1.47	
1.49	
1.55	

NOTE: There are 10 measured values which have been put into numerical order from the smallest to the largest.

The raw data are enter as order statistic into the computer program "LOGIPAR" . The results come printed out as shown in TABLE 3 on Page 8. If a simulated experiment is desired in order to collect any number of random measurements which come from the predicted logistic distribution, we simply employ the parameters obtained (i.e., Logistic Slope and Logistic Mean) in a Logisitic Generating function. This simulation is carried out by the computer program "LOGISIM". For the example, such a set of 50 measurements is printed out in TABLE # 4 on page 9 .

LOGISTIC PARAMETER BY LEAST SQUARES - (LOGIPAR PROGRAM)

PROBLEM TITLE: LOGISTIC MEASUREMENTS

ORDER NO.	MEASUREMENTS
1	1.25
2	1.27
3	1.31
4	1.35
5	1.36
6	1.39
7	1.43
8	1.47
9	1.49
10	1.55

PROBLEM TITLE: LOGISTIC MEASUREMENTS

J= 1 AND X= 1.25 ODDS AGAINST EXCEEDENCE= 7.216495E-02 TO 1(AT X = 1.25)
J= 2 AND X= 1.27 ODDS AGAINST EXCEEDENCE= .1954023 TO 1(AT X = 1.27)
J= 3 AND X= 1.31 ODDS AGAINST EXCEEDENCE= .3506494 TO 1(AT X = 1.31)
J= 4 AND X= 1.35 ODDS AGAINST EXCEEDENCE= .5522389 TO 1(AT X = 1.35)
J= 5 AND X= 1.36 ODDS AGAINST EXCEEDENCE= .8245615 TO 1(AT X = 1.36)
J= 6 AND X= 1.39 ODDS AGAINST EXCEEDENCE= 1.212766 TO 1(AT X = 1.39)
J= 7 AND X= 1.43 ODDS AGAINST EXCEEDENCE= 1.810811 TO 1(AT X = 1.43)
J= 8 AND X= 1.47 ODDS AGAINST EXCEEDENCE= 2.851852 TO 1(AT X = 1.47)
J= 9 AND X= 1.49 ODDS AGAINST EXCEEDENCE= 5.117648 TO 1(AT X = 1.49)
J= 10 AND X= 1.55 ODDS AGAINST EXCEEDENCE= 13.85715 TO 1(AT X = 1.55)

LOGISTIC SLOPE = 15.84062
LOGISTIC MEAN = 1.387
GOODNESS OF FIT = .9898704
LOGISTIC SIGMA = .1145031

LOWER SPEC = 1.265
NOMINAL VALUE = 1.365
UPPER SPEC = 1.465

% BELOW LOWER SPEC = 12.64681 %
% BELOW NOMINAL VALUE = 41.3748 %
% ABOVE UPPER SPEC = 22.52082 %
DISTRIBUTION'S MEAN SHIFT FROM NOMINAL = 2.199996E-02
PERCENT WITHIN THE TWO SPECS = 64.83238 %
PERCENT OUTSIDE THE TWO SPECS = 35.16762 %

LOGISTIC SIMULATION PROGRAM

LOGISTIC SLOPE = 15.84062
LOGISTIC MEAN = 1.387

RANDOM VALUES
1 = 1.442
2 = 1.431
3 = 1.453
4 = 1.428
5 = 1.307
6 = 1.29
7 = 1.553
8 = 1.483
9 = 1.22
10 = 1.399
11 = 1.286
12 = 1.344
13 = 1.496
14 = 1.346
15 = 1.492
16 = 1.4
17 = 1.536
18 = 1.523
19 = 1.398
20 = 1.462
21 = 1.393
22 = 1.461
23 = 1.419
24 = 1.4
25 = 1.374
26 = 1.399
27 = 1.378
28 = 1.077
29 = 1.358
30 = 1.457
31 = 1.342
32 = 1.177
33 = 1.341
34 = 1.335
35 = 1.321
36 = 1.379
37 = 1.268
38 = 1.358
39 = 1.352
40 = 1.402
41 = 1.461
42 = 1.41
43 = 1.403
44 = 1.646
45 = 1.406
46 = 1.358
47 = 1.513
48 = 1.401
49 = 1.469
50 = 1.441 ENDED AFTER 50 RANDOM VALUES

THE RESULT OF SORTING THE RANDOM LIST OF 50

THE SORTING LIST IS :

```
# 1 = 1.213
# 2 = 1.215
# 3 = 1.222
# 4 = 1.234
# 5 = 1.25
# 6 = 1.267   <---- Lower Spec at 1.265
# 7 = 1.292
# 8 = 1.303
# 9 = 1.305
# 10 = 1.307
# 11 = 1.316
# 12 = 1.326
# 13 = 1.335
# 14 = 1.335
# 15 = 1.339
# 16 = 1.348
# 17 = 1.351
# 18 = 1.361  <---- Nominal Value at 1.365
# 19 = 1.37                                    \
# 20 = 1.371                                    > Mean Shift = +.022
# 21 = 1.378                                   /
# 22 = 1.388  <---- Distribution Mean at 1.387
# 23 = 1.392
# 24 = 1.396
# 25 = 1.399
# 26 = 1.399
# 27 = 1.404
# 28 = 1.404
# 29 = 1.411
# 30 = 1.434
# 31 = 1.442
# 32 = 1.452
# 33 = 1.457
# 34 = 1.46
# 35 = 1.465  <---- Upper Spec at 1.465
# 36 = 1.472
# 37 = 1.486
# 38 = 1.493
# 39 = 1.497
# 40 = 1.503
# 41 = 1.511
# 42 = 1.536
# 43 = 1.562
# 44 = 1.566
# 45 = 1.567
# 46 = 1.568
# 47 = 1.582
# 48 = 1.59
# 49 = 1.609
# 50 = 1.719
```

<u>CONCLUSION:</u> This list of 50 random values obtained by a simulated random process from the Predicted Logistic Population agree within possible sampling errors with the Predicted Percentages of TABLE # 3 on Page 8.

CONCLUSION

We conclude that our **LOGISTIC SPECIFICATION** statistical approach is a powerful tool which accurately predicts percentages inside and outside specification limits in a very simple and effective manner. The **GOODNESS OF FIT** tells us that it is totally valid in the example used to illustrate the procedure. Furthermore, in the **APPENDIX**, we have drawn the **SEMI-LOG PLOT of ODDS AGAINST EXCEEDANCE** versus each measured value. The good fitting **Linear Plot on Semi-Log Paper** proves that the approach is a valid one for the data set. This technique is so much handier than the classical **Normal Distribution** approach employed in **STATISTICAL QUALITY CONTROL** that it behooves all quality circles to switch to this straightforward and sensible approach.

APPENDIX
SEMI-LOG FOR THE EXAMPLE

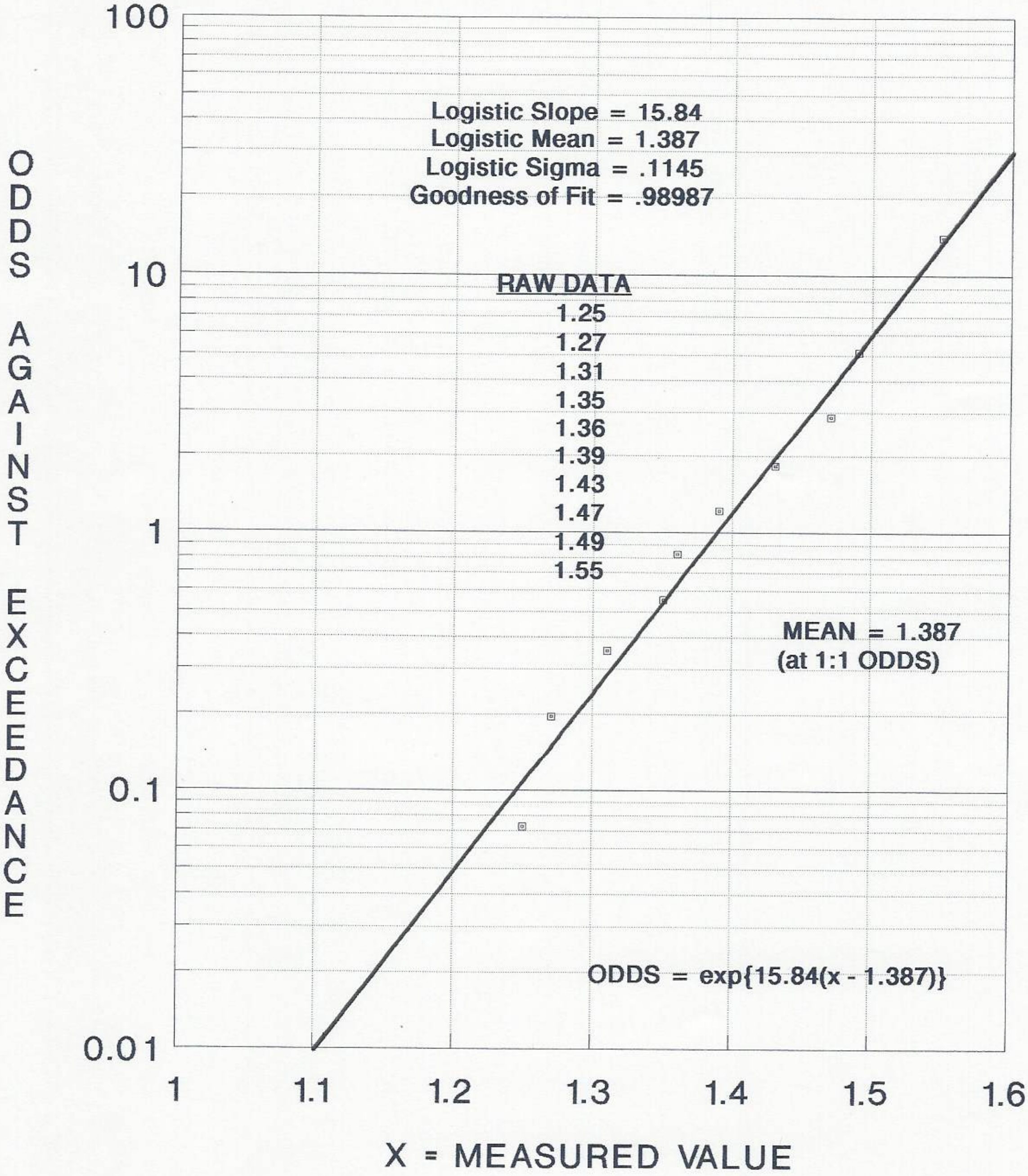

FIGURE 3