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 18
BULLETIN 1

APRIL , 1988

SEMI-PARAMETRIC LIMITS
(THE MOST REASONABLE BOUNDS FOR CONFIDENCE BANDS)

INTRODUCTION

A PROBABILITY PAPER PLOT OF A CUMULATIVE DISTRIBUTION FUNCTION WITH A 90% CONFIDENCE BAND CONSISTS OF THREE DISTINCT PARTS.

THESE ARE

(1) THE CENTRAL LINE OF PREDICTION

(2) THE UPPER LIMIT, AT WHICH THERE IS ONLY A 5% CHANCE OF GOING HIGHER.

(3) THE LOWER LIMIT, AT WHICH THERE IS ONLY A 5% CHANCE OF GOING LOWER.

THESE PARTS ARE SHOWN GRAPHICALLY IN FIGURE 1 BELOW :

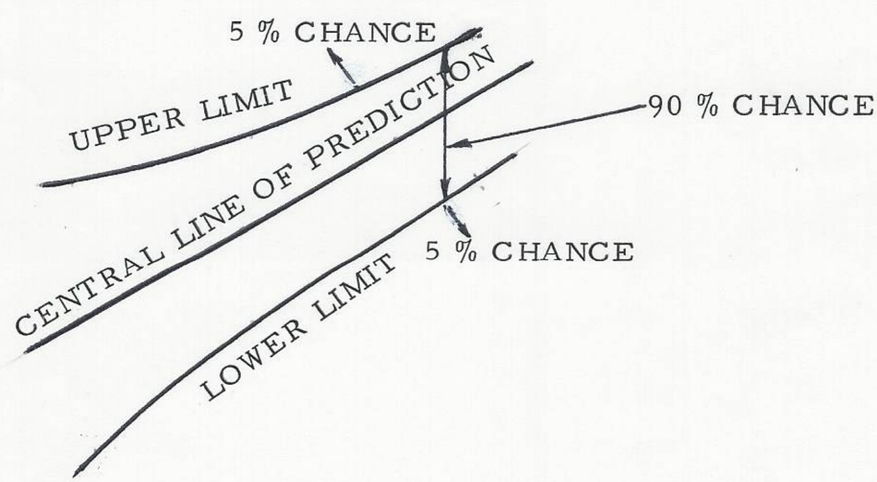

FIGURE 1

NOTICE THAT THE TOTAL REALM (100%) OF POSSIBILITIES IS THUS BROKEN UP INTO THREE PORTIONS:

(1) 5% ABOVE THE UPPER LIMIT

(2) 5% BELOW THE LOWER LIMIT

(3) 90% BETWEEN THE LOWER AND UPPER LIMITS

(THUS, WE CALL PORTION (3) THE 90% CONFIDENCE BAND FOR THE PREDICTED FUNCTION.)

CONVERTING THE CHANCES TO ODDS

QUESTION : WHAT DOES A 5% CHANCE OF BEING EXCEEDED MEAN ?

ANSWER : IT MEANS 1 CHANCE IN 20 OF BEING EXCEEDED, AS WELL AS 19 CHANCES IN 20 OF NOT BEING EXCEEDED.

THE RATIO 19:1, OR (19/1), IS CALLED THE ODDS AGAINST THE UPPER LIMIT BEING EXCEEDED.

LIKEWISE, WE SAY THAT THE ODDS AGAINST THE LOWER LIMIT BEING VIOLATED (BY GOING LOWER) IS 19:1, I.E., (19/1) .

THUS, IN CONSTRUCTING A 90% CONFIDENCE BAND WE USE (19/1) AS THE ODDS RATIO TO IDENTIFY BOTH THE UPPER AND LOWER CONFIDENCE LIMITS.

ANALYSIS OF BOUNDS FOR A SINGLE POINT

LET US TAKE ONE POINT ON THE CENTRAL LINE OF PREDICTION, AS IN FIGURE 2.

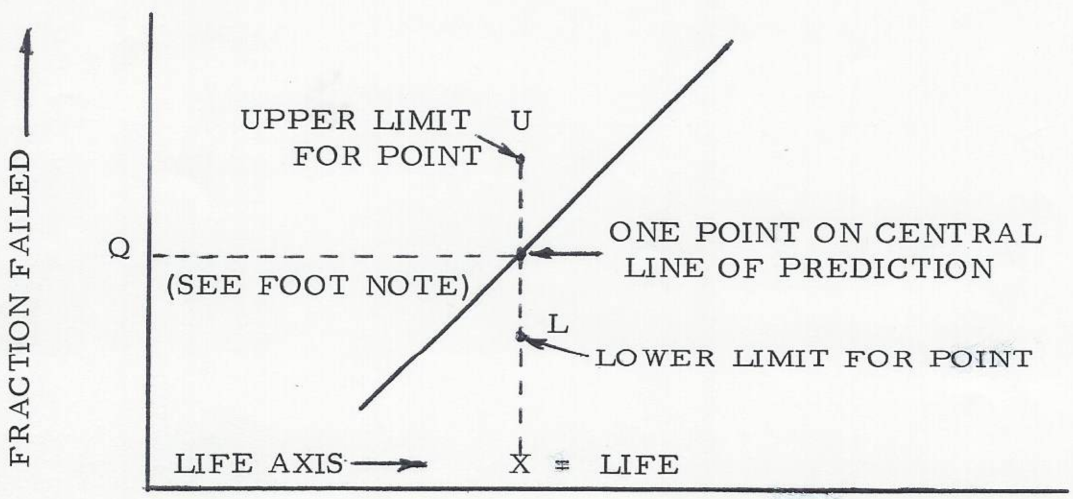

FIGURE 2

SINCE THE CENTRAL LINE OF PREDICTION IS A CDF (CUMULATIVE DISTRIBUTION FUNCTION) OF LIFE, IT FOLLOWS THAT EACH POINT ON IT HAS AN ORDINATE Q WHICH REPRESENTS THE FRACTION FAILED IN THE LIFE DISTRIBUTION REPRESENTED BY THE CDF. THE PROBLEM NOW IS TO LOCATE THE UPPER LIMIT U AND THE LOWER LIMIT L FOR SUCH A POINT (AT LEVEL Q ON THE CENTRAL LINE OF PREDICTION).

NOTE: Q IS ASSUMED TO BE .5 OR LESS.

SOLUTION BY THE USE OF ENTROPY

THE BASIC PROBLEM OF DEFINING CONFIDENCE BOUNDS FOR ANY POINT ON THE CENTRAL LINE OF PREDICTION IS HANDLED VERY EASILY BY USING THE CONCEPT OF ENTROPY.

WE DEFINE THE ENTROPY AT LEVEL Q TO BE

$$\text{ENTROPY} = \mathcal{E} = -\ln(1 - Q) \quad .$$

NEXT, WE MOVE ALONG THE ENTROPY SCALE TO THE UPPER LIMIT U BY MAKING USE OF THE UNIVERSAL LAW OF ODDS, WHICH SAYS

$$(\text{ODDS}) = (\text{ENTROPY RATIO})^{(\text{ODDS EXPONENT})}$$

OR

$$\text{ENTROPY RATIO} = (\text{ODDS})^{(1/\text{ODDS EXPONENT})} \qquad (1)$$

THREE TYPES OF ODDS EXPONENTS

THE ODDS EXPONENT WE USE DEPENDS ON THE AMOUNT OF UNCERTAINTY IN THE POPULATION PARAMETERS OF THE CDF. THERE MAY BE

(A) TOTAL UNCERTAINTY

(B) NO UNCERTAINTY

(C) A MEDIAN AMOUNT OF UNCERTAINTY

THE MOST REASONABLE PREDICTION WE CAN MAKE WHEN THE AMOUNT OF UNCERTAINTY IS UNKNOWN IS TO ASSUME (C), I.E., A MEDIAN AMOUNT OF UNCERTAINTY. IN THAT CASE WE TAKE THE MEDIAN ODDS EXPONENT DEFINED BY

$\pi/\sqrt{3}\ \sqrt{.5N\ (1+Q)}$, WHERE

(SEE APPENDIX FOR SITUATIONS WHERE $Q > .5$)

N = SAMPLE SIZE AT THE POINT ON THE CENTRAL LINE OF PREDICTION.

Q = FRACTION FAILED AT THE POINT ON THE CENTRAL LINE OF PREDICTION. ($Q \leq .5$)

THEN EQUATION (1) BECOMES

$$\text{ENTROPY RATIO} = (\text{ODDS})^{\sqrt{3}/\pi\sqrt{.5N\ (1\ +\ Q)}} = (\text{ODDS})^{[.78/\sqrt{N(1+Q)}]}$$

SINCE FOR 90% CONFIDENCE BOUNDARIES WE USE

ODDS = (19/1) , IT FOLLOWS THAT

$$\text{ENTROPY RATIO} = 19^{[.78/\sqrt{N\ (1\ +\ Q)}]} \qquad (2)$$

EVALUATION OF THE UPPER AND LOWER LIMITS

THE CENTRAL LINE OF PREDICTION AT LEVEL Q HAS

$$\text{ENTROPY} = -\ln(1 - Q) \quad .$$

WE MULTIPLY THIS CENTRAL ENTROPY BY THE ENTROPY RATIO IN (2) TO OBTAIN THE ENTROPY AT THE UPPER LIMIT U . THUS,

$$\text{UPPER ENTROPY} = -19^{[.78/\sqrt{N(1 + Q)}]} \ln(1 - Q)$$

OR

$$-\ln(1 - U) = -19^{[.78/\sqrt{N(1 + Q)}]} \ln(1 - Q)$$

OR

$$1 - U = (1 - Q)^{\{19^{[.78/\sqrt{N(1 + Q)}]}\}}$$

OR

$$U = 1 - (1 - Q)^{\{19^{[.78/\sqrt{N(1 + Q)}]}\}} \qquad (3)$$

IN A SIMILAR FASHION, THE LOWER LIMIT L IS OBTAINED BY DIVIDING THE CENTRAL ENTROPY BY THE SAME ENTROPY RATIO GIVEN BY (2) IN ORDER TO ARRIVE AT THE LOWER ENTROPY AT L . THIS YIELDS

$$L = 1 - (1 - Q)^{\{1/[19^{(.78/\sqrt{N(1 + Q)})}]\}} \qquad (4)$$

(3) AND (4) YIELD REASONABLE LIMITS WHEN THE UNCERTAINTY IS UNKNOWN.

A NUMERICAL EXAMPLE

SUPPOSE WE HAVE A POINT ON THE CENTRAL LINE OF PREDICTION SUCH THAT

Q = 30% FAILED (AT THE POINT) (I.E., Q = .3)

N = SAMPLE SIZE = 10 (AT THE POINT)

THEN $.78/\sqrt{N(1+Q)} = .78/\sqrt{10\,(1+.3)} = .21633$

ENTROPY RATIO $= 19^{.21633} = 1.8907$

THUS, FROM (3):

$U = 1 - (1 - .3)^{1.8907} = .4905$

$L = 1 - (1 - .3)^{1/1.8907} = .1719$

THUS , THE UPPER LIMIT FOR THE FRACTION FAILED IS 49.05% AND THE LOWER LIMIT FOR THE FRACTION FAILED IS 17.19%. SO, WE CONCLUDE WITH 90% CONFIDENCE THAT THE TRUE FRACTION FAILED IS SOMEWHERE BETWEEN 17.19% AND 49.05% WITH THE CENTRAL PREDICTED PERCENT FAILED AT 30% AT LIFE X .

CONCLUSION

IT CAN BE SEEN HOW CONVENIENT THE ENTROPY APPROACH IS WHEN WE KNOW THE UNIVERSAL LAW OF ODDS. IN THIS WAY WE DO NOT NEED ANY RANK TABLES.
IT IS A REAL BREAKTHROUGH !

APPENDIX

FOR SITUATION IN WHICH THE POINT ON THE CENTAL LINE OF PREDICTION HAS Q > .5 ARE HANDLED BY MODIFYING THE FORMULA FOR THE MEDIAN ODDS EXPONENT AS FOLLOWS :

INSTEAD OF TAKING $\pi/\sqrt{3}\ [\sqrt{.5N\ (1 + Q)}]$ AS THE FORMULA FOR THE MEDIAN ODDS EXPONENT WE USE THE MODIFIED FORMULA

MEDIAN ODDS EXPONENT $= \pi/\sqrt{3}\ (\sqrt{.5N\ (2 - Q)}$, WHERE $Q > .5$

HENCE , THE ENTROPY RATIO (2) BECOMES

$$\text{ENTROPY RATIO} = 19^{[.78/\sqrt{N\ (2 - Q)}]}$$

FORMULA (3) BECOMES

$$U = 1 - (1 - Q)\ \{19^{[.78/\sqrt{N\ (2 - Q)}]}\}$$

FORMULA (4) BECOMES

$$L = 1 - (1 - Q)\ \{1/[19^{(.78/\sqrt{N\ (2 - Q)})}]\}$$