Statistical Bulletin
Reliability & Variation Research

DETROIT RESEARCH INSTITUTE
21900 GREENFIELD ROAD • OAK PARK, MICHIGAN 48237 • (313) 968-1818

LEONARD G. JOHNSON
EDITOR

WANG H. YEE
DIRECTOR

Volume 6 Bulletin 1 April, 1976

THE DISTINCTION BETWEEN MIXTURE PROBLEMS AND SYSTEMS PROBLEMS

Published 8 times a year by the Detroit Research Institute. Subscription Rate: One year $32.00. Quantity rates on request.
Address subscription orders and editorial correspondence to:
Detroit Research Insitute, 21900 Greenfield Road, Oak Park, Michigan 48237

Introduction

In our educational experience during the past fifteen years at Detroit Research Institute we have noticed that there is quite often a lack of a clear understanding concerning the distinction between two types of problems in statistical reliability. These two types of problems are

I Mixture Problems

II Systems Problems with items in series which produce different modes of failure

Because of this lack of a clear understanding regarding the natures of Mixture Problems and Systems Problems, we have decided to use this issue of the DRI Statistical Bulletin as the medium for clarifying the distinction between Mixture Problems and Systems Problems.

The Meaning Of A Mixture Problem

A Mixture Problem arises whenever a data set under investigation is heterogeneous and has more than one homogeneous sub-population within the total data set. Examples of such mixed situations are

(1) Mixed data from two or more vendors

(2) Mixed data from two or more working shifts

(3) Mixed data from different manufacturing plants

(4) Mixed data from different seasons of production (e.g., summer items and winter items) which are different due to differences in temperature and humudity

(5) Mixed data from different climates of use

(6) Mixed data from different geographical locations

(7) Mixed data from customers of different temperament or ethnic origin, which makes for a difference in severity of use or even product abuse.

It should be understood that in a mixture problem we are dealing with the same failure mode, but that this same failure mode shows a better longevity distribution in some parts of the mixture than in others.

The Analysis Of Mixture Problems

As an example of a mixture problem consider the following data set on the hours to leakage of sixteen oil seals:
(NOTE: Median Ranks have been listed in the third column for a sample of 16.)

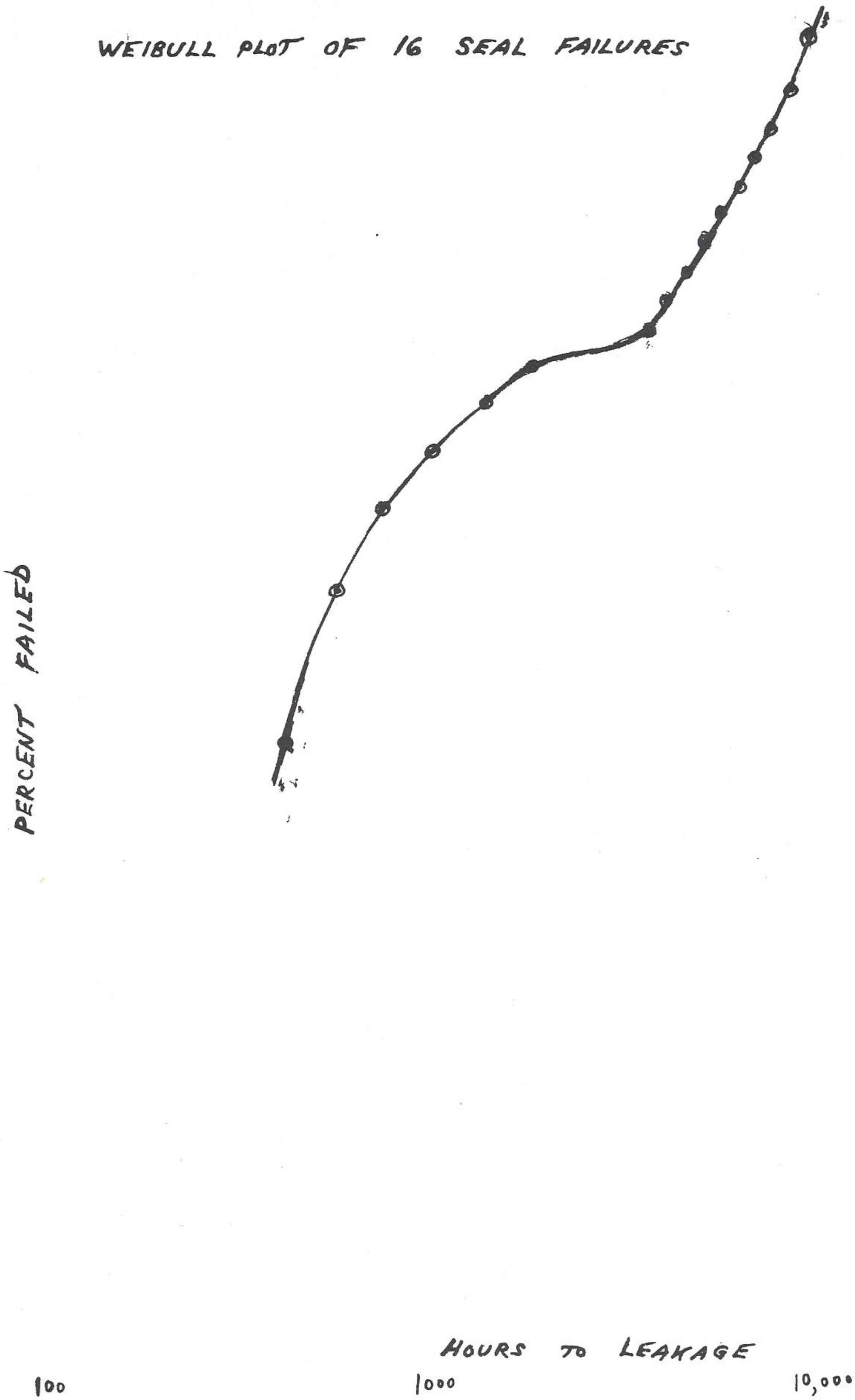
TABLE 1

Leakage Failure No.	Hours to Leakage	Percent of Population Failed (Median Rank)
1	459 hrs.	4.24 %
2	620	10.34 %
3	836	16.44 %
4	1129	22.54 %
5	1525	28.65 %
6	2058	34.75 %
7	4212	40.85 %
8	4675	46.95 %
9	5190	53.05 %
10	5760	59.15 %
11	6394	65.25 %
12	7097	71.35 %
13	7878	77.46 %
14	8745	83.56 %
15	9707	89.66 %
16	10,774 hrs.	95.76 %

Plotting the hours to leakage as abscissas and the Median Ranks as ordinates on Weibull Paper we obtain the Weibull plot shown in FIGURE 1.

FIGURE 1

WEIBULL PLOT OF 16 SEAL FAILURES



The Interpretation Of Figure 1

Obviously the Weibull plot shown in FIGURE 1 is not a straight line. As a matter of fact, it is an S-shaped curve with an inflection point at $37\frac{1}{2}$ % on the ordinate scale. There are 6 points to the left of the inflection point and 10 points to the right of the inflection point. Interestingly enough, it so happens that $37\frac{1}{2}$% of the total sample size of 16 is exactly 6. We therefore separate the total sample of 16 into two samples, one sample consisting of the first six data values, and the other sample consisting of the last 10 data values. Then we have the following pair of samples:

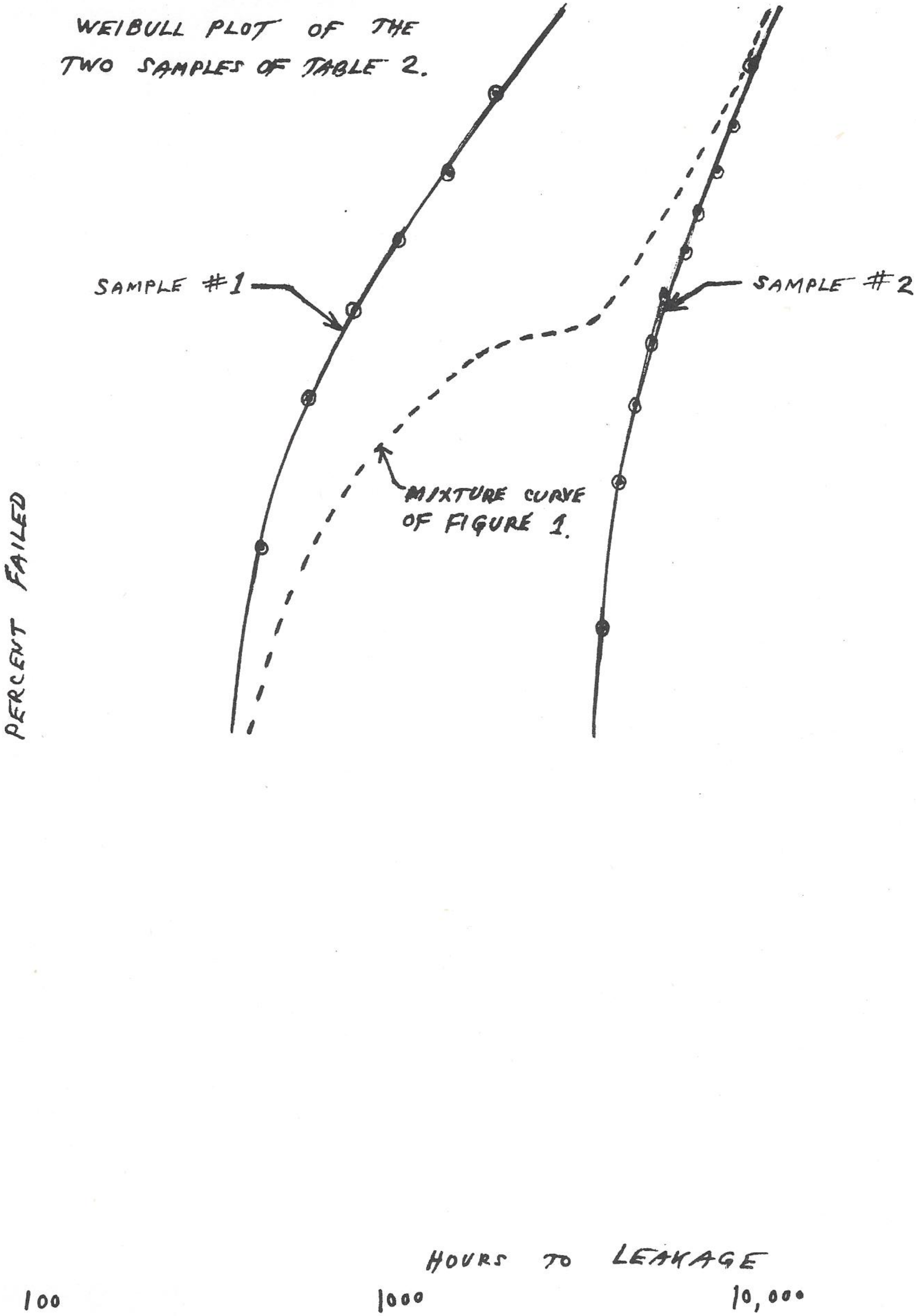
TABLE 2

SAMPLE # 1 (N_1 = 6)		SAMPLE # 2 (N_2 = 10)	
	MED. RANK		MED. RANK
459 hrs.	10.91 %	4212 hrs.	6.70 %
620	26.55 %	4675	16.32 %
836	42.18 %	5190	25.94 %
1129	57.82 %	5760	35.57 %
1525	73.45 %	6394	45.19 %
2058 hrs.	89.09 %	7097	54.81 %
		7878	64.43 %
		8745	74.06 %
		9707	83.68 %
		10,774 hrs.	93.30 %

These two samples are plotted as solid lines on Weibull paper in FIGURE 2.

FIGURE 2

WEIBULL PLOT OF THE TWO SAMPLES OF TABLE 2.



Conclusions From Figure 2

Upon examining FIGURE 2 we notice that

(1) The Mixture Curve of FIGURE 1 lies between the two sub-populations represented by Sample # 1 and Sample # 2.

(2) The ordinate of the inflection point on the dotted Mixture Curve represents the percentage of the mixture belonging to Part # 1 (represented by Sample # 1).

This is always the case when the data represent a mixture of two non-intersecting sub-populations. *

(3) The mathematical formula for the Mixture CDF (cumulative distribution function of failures) is

$$F_{Mixture}(X) = Q_1 F_1(X) + Q_2 F_2(X)$$

Two-Part Mixture Formula

where

$F_1(X)$ = CDF of Part (1) of the mixture

$F_2(X)$ = CDF of Part (2) of the mixture

Q_1 = Fraction of mixture coming from Part (1)

Q_2 = Fraction of mixture coming from Part (2)

$$(Q_1 + Q_2 = 1)$$

---

* If two sub-populations intersect at ordinate Y_0, then the mixture curve will pass through the intersection point and will have two inflection points, one at ordinate $Q_1 Y_0$ and the other at ordinate $Q_2 + Q_1 Y_0$, where Q_1 is the fraction of the mixture coming from sub-population (1) of smaller slope, and $Q_2 = 1 - Q_1$ is the fraction of the mixture from sub-population (2) of steeper slope.

The Meaning Of A Systems Problem

A systems failure problem is a situation in which an assembly (i.e., system) can fail in more than one way. Any one of several components in series making up the system can fail, thus causing an assembly failure. Each such individual component failure is called a <u>failure</u> <u>mode</u>. The basic mathematical relation which governs system reliability in these situations is the so-called <u>PRODUCT RULE</u>, which says that the SYSTEM RELIABILITY to a target is the product of all the individual component reliabilities to the same target. Thus, if a system consists of component (1) and component (2) in series, the product rule is written

$$R_{System}(X) = R_1(X)\, R_2(X)$$

where $R_1(X)$ = Reliability of component (1) to target time X

$R_2(X)$ = Reliability of component (2) to target time X

$R_{System}(X)$ = Reliability of the system to target time X

From this mathematical relation it follows that the formula for the system CDF (cumulative distribution function of failures) is

$$F_{System}(X) = 1 - \left[1 - F_1(X)\right] \left[1 \quad F_2(X)\right]$$

where $F_1(X)$ = CDF of component (1) = $1 - R_1(X)$

$F_2(X)$ = CDF of component (2) = $1 - R_2(X)$

$F_{System}(X)$ = CDF of the system = $1 - R_{System}(X)$

Analysis Of Systems Problems

As an example of a systems problem consider the following data on failures of an assembly which can fail in two ways (i.e., Type (1) or Type (2) failures):

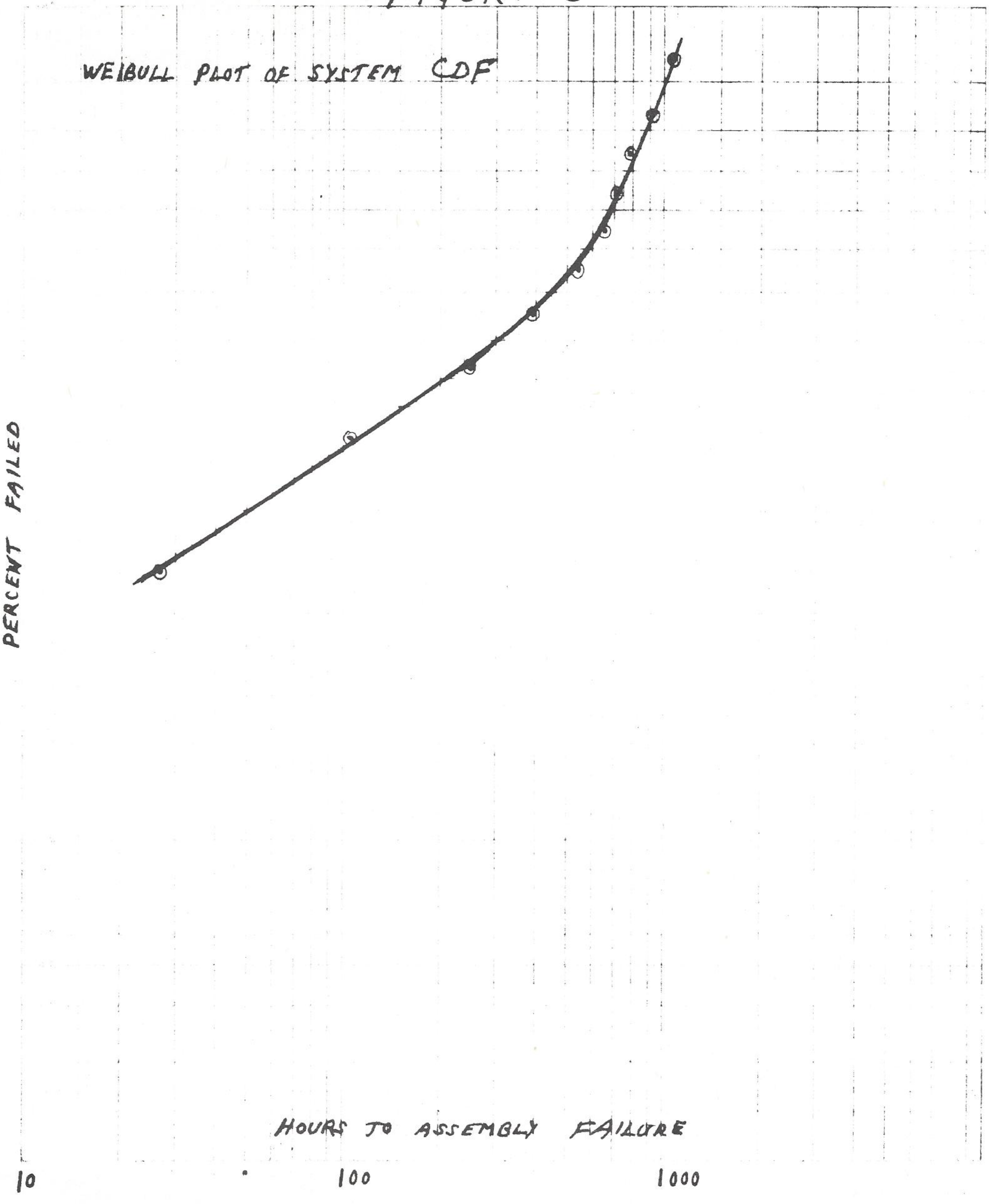
TABLE 3 (N = 10)

Assembly Failure No.	Hours to Failure	Failure Type	Median Rank
1	27	(1)	6.70 %
2	106	(1)	16.32 %
3	249	(1)	25.94 %
4	392	(2)	35.57 %
5	540	(2)	45.19 %
6	652	(2)	54.81 %
7	701	(1)	64.43 %
8	791	(2)	74.06 %
9	916	(2)	83.68 %
10	1075	(2)	93.30 %

Plotting all of these 10 assembly failures (regardless of type of failure) on Weibull paper yields FIGURE 3 .

DRI STATISTICAL BULLETIN
Vol. 6, Bulletin 1
April, 1976

FIGURE 3



Analyzing The Two Modes Of Failure Separately

In order to analyze the Type (1) failures we must treat all Type (2) failures as suspended items. Thus, for the data table on Type (1) failures we have the following mean order numbers (by the New Increment Technique) and median ranks for the Type (1) failures in the list given in TABLE 3:

TABLE 4 (Type (1) Failures)

	Mean Order No.	Median Rank
27 hrs. (Failed)	# 1 in 10	6.70 %
106 hrs. (Failed)	# 2 in 10	16.32 %
249 hrs. (Failed)	# 3 in 10	25.94 %
392 hrs. (Suspended)		
540 hrs. (Suspended)		
652 hrs. (Suspended)		
701 hrs. (Failed)	# 4.6 in 10	41.35 %
791 hrs. (Suspended)		
916 hrs. (Suspended)		
1075 hrs. (Suspended)		

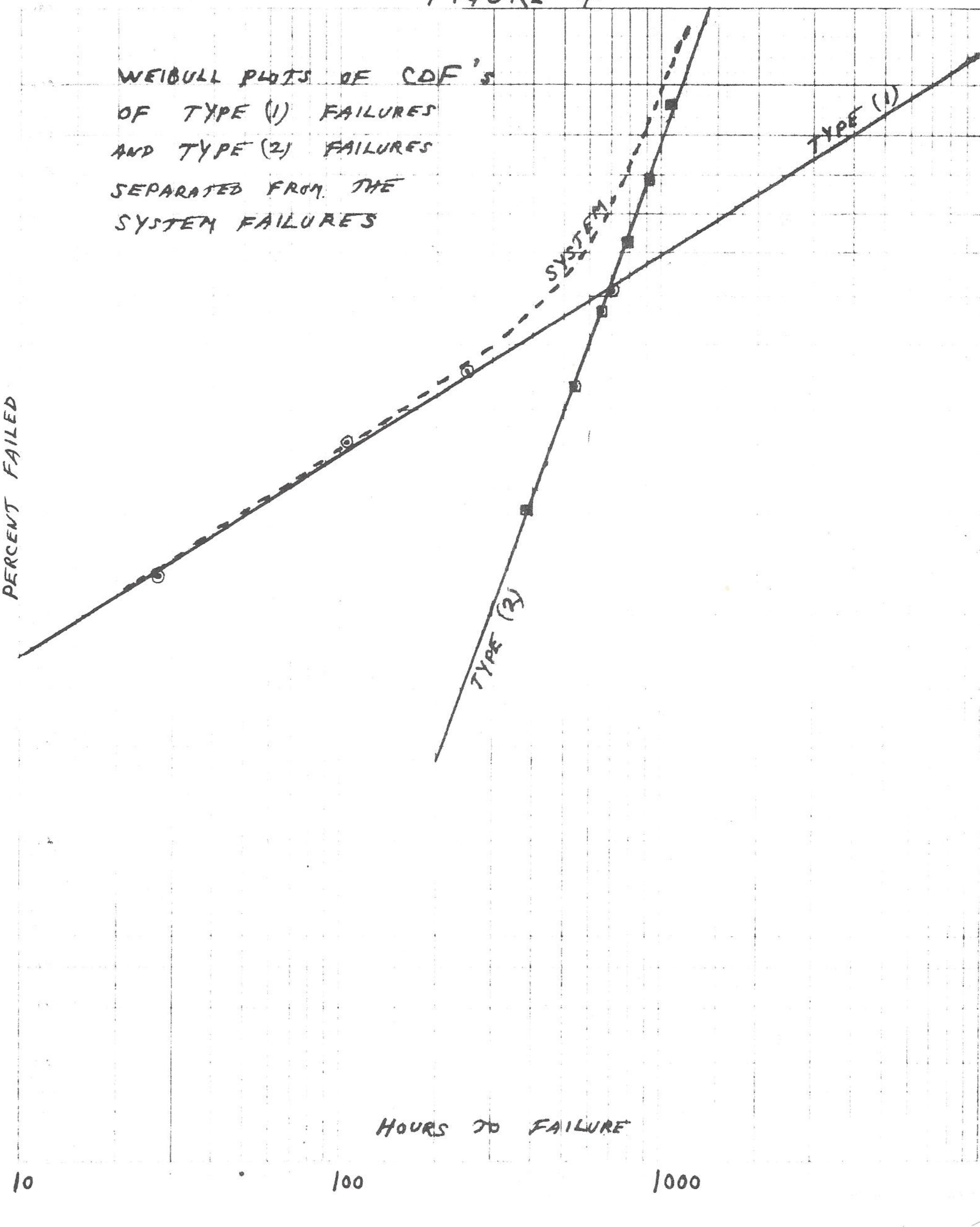
In order to analyze the Type (2) failures we treat all Type (1) failures as suspended items, as shown in TABLE 5 below: (Mean Order Numbers for the Type (2) failures have been calculated by using the New Increment Technique):

TABLE 5 (Type (2) Failures)

	Mean Order No.	Median Rank
27 hrs. (Susp.)		
106 hrs. (Susp.)		
249 hrs. (Susp.)		
392 hrs. (Failed)	# 1.375 in 10	10.34 %
540 hrs. (Failed)	# 2.75 in 10	23.56 %
652 hrs. (Failed)	# 4.125 in 10	36.78 %
701 hrs. (Susp.)		
791 hrs. (Failed)	# 5.84375 in 10	53.31 %
916 hrs. (Failed)	# 7.5625 in 10	69.83 %
1075 hrs. (Failed)	# 9.28125 in 10	86.36 %

Plotting the data of TABLES 4 and 5 on Weibull paper we obtain the two solid lines of FIGURE 4 as the separate CDF graphs of failure types (1) and (2) respectively. The dotted curve is the same systems CDF we plotted in FIGURE 3.

FIGURE 4


WEIBULL PLOTS OF CDF'S
OF TYPE (1) FAILURES
AND TYPE (2) FAILURES
SEPARATED FROM THE
SYSTEM FAILURES
SYSTEM
TYPE (1)
TYPE (2)
PERCENT FAILED
HOURS TO FAILURE
10
100
1000

Conclusions From FIGURE 4

From FIGURE 4 we can say that

(1) The SYSTEM CDF lies to the left of the CDF's of Modes (1) and (2).

(2) The SYSTEM CDF initially follows the failure mode of lower slope, and later on (above the intersection point of the two modes) it moves in the direction of the failure mode of steeper slope.

From these illustrations it should be clear that MIXTURE PROBLEMS and SYSTEMS PROBLEMS are of entirely different natures.