

THE CHARACTERIZATION OF DISTRIBUTIONS
OF CHEMISTRIES WITH THE SMA 12/60

by

Adeline V. Clarke

Technical Report No. 104
Department of Statistics NASA Contract

April 19, 1971

Research sponsored by National Aeronautics and Space Administration
Grant No. NGR 44-007-006

Reproduction in whole or in part is permitted
for any purpose of the United States Government.

This document has been approved for public release
and sale; its distribution is unlimited.

DEPARTMENT OF STATISTICS
Southern Methodist University

THE CHARACTERIZATION OF DISTRIBUTIONS
OF CHEMISTRIES WITH THE SMA 12/60

A Thesis Presented to the Faculty of the Graduate School

of

Southern Methodist University

in

Partial Fulfillment of the Requirements

for the degree of

Master of Science

with a

Major in Statistics

Adeline Vera Clarke
(B.S., Texas Woman's University, 1960)
M.T.(ASCP)

April 19, 1971

TABLE OF CONTENTS

Page	
Abstract.....	iv
Acknowledgments.....	v
I. Introduction.....	1
II. Systems of Frequency Curves.....	3
2.1 Pearson.....	3
2.2 Burr.....	16
III. Tolerance Limits.....	19
IV. Summary.....	32
Bibliography.....	35

CHAPTER I

INTRODUCTION

From (Burlington and May, 1970), I quote: "In nearly all fields there arise problems where precise measurements or observations are impossible to make, or where events are not exactly reproducible or predictable. The analysis of such situations leads to considerable uncertainty. Statistical methods lend themselves to the analysis of such situations and furnish a means of describing and indicating trends, of expectations, often with an associated degree of reliability. Statistics makes use of scientific methods of collecting, analyzing and interpreting data."¹

The value of modern statistical procedures is taking its right-

ful place in the clinical laboratory with a better interpretation of data accumulated.

Interests in this paper include exploiting statistical methods

in the laboratory and finding "normal values" for twelve chemistries with the sequential multiple analyzer, SMA 12/60, for the laboratory of

St. Paul Hospital, Dallas, Texas.

Two pathologists suggested criteria for placing patients in the

hospital into two groups, "normal" and "abnormal." A patient is included in the "normal group" if the systolic pressure is greater than or equal

to 100 and less than or equal to 160; the diastolic pressure is greater

than or equal to 40 and less than or equal to 90; temperature is less

than or equal to 100° F.; respiration is less than or equal to 20; WBC is

less than or equal to 14,000 and greater than or equal to 4,000; bands

¹S. B. Burlington and D. C. May, Handbook of Probability and Statistics with Tables, 2nd ed., (New York: McGraw-Hill Book Company, 1970), p. 1.

less than or equal to 10%; lymphs less than or equal to 45% for patients greater than 14 years; hemoglobin greater than or equal to 10.0gms, where the upper limit is 15.5 gms for men and 15.0 gms for women. For urines, patients are included in the "normal group" if there is no sign of sugar and if protein is less than 5.0 mgs % . There were some categories of patients such as those with open-heart surgery, diabetes, etc., which were excluded from the "normal group."

Patients' charts, chosen serially from the SMA 12/60 log book, were pulled in the hospital record room until 900 records were found. Of these, 31 could not be used due to incomplete records.

An analysis was done on 291 patients in the "normal group," 578 patients in the "abnormal group," and 869 patients in the "total group." Some chemistry values were censored because they were beyond the capacity of the recorder to register them.

In Chapter II, the curve-shaped characteristics by the Pearson system of curves are analyzed. Certain characteristics are mentioned of the Burr distribution which I hope to be able to expand later.

In Chapter III, "normal values" for each group are found by using a distribution-free or non-parametric procedure.

Chapter IV summarizes this work and gives ideas for its extension.

In the late 1800's Karl Pearson devised a system of frequency curves using the method of moments, a method which he named. William Palin Elderton's and Norman Lloyd Johnson's book, Systems of Frequency Curves, (1969) gives a practical approach for the application of Pearson's system of curves. It is a handy reference for formulae and methods. It also gives a fairly thorough explanation of the theory involved. Since I will be referring to this book extensively, it will be referred to as E & J.

In this part of Chapter II I write only of the three main types because the data used falls within Types I, IV, and VI. Tables 2.1 through 2.3 show values of computations of the central moments, the censored n , and many criteria needed and used for the fitting of the Pearson curves to the data used. By this method the roughness of the data can be removed. The average value of the variable of the observed data is the mean, the mode is the value of the variable which occurs most frequently, and the standard deviation shows the spread of a curve. The difference between the mean and the mode is a measure of skewness. Skewness is actually

Pearson

SYSTEMS OF FREQUENCY CURVES

CHAPTER II

Table 2.1

Statistics used to classify distribution curves according to the Pearson and Burr systems.

'Normal Group'

	Total Protein	Albumin	Calcium	Cholesterol	Glucose	BUN
Mean	.6977+01	.4155+01	.9189+01	.2075+03	.1018+03	.1338+02
Second	.3583-00	.4256-00	.2555-00	.1574+04	.2722+03	.3131+02
Third	.5543-01	-.8180-01	.1961-01	.2789+05	.4704+04	.3848+03
Fourth	.4419-00	.7864-00	.1882-00	.8390+07	.5114+06	.1168+05
Std Deviation	.5986-00	.6524-00	.5055-00	.3968+02	.1650+02	.5596+01
n	.2910+03	.2910+03	.2910+03	.2890+03	.2890+03	.2910+03
Kappa	.7484-01	.2823-01	-.5729-01	.9108-00	.2450-00	.2405+01
Delta	.1059-00	.3299-00	-.5161-01	.2700-01	.4557-00	.2254-00
Beta 1	.6676-01	.8679-01	.2304-01	.1994-00	.1097+01	.4825+01
Beta 2	.3441+01	.4341+01	.2883+01	.3385+01	.6901+01	.1192+02
Sqr(Beta 1)	.2584-00	.2946-00	.1518-00	.4466-00	.1047+01	.2197+01
Alpha 3	.2584-00	-.2946-00	.1518-00	.4466-00	.1047+01	.2197+01
Alpha 4	.3441+01	.4341+01	.2883+01	.3385+01	.6901+01	.1192+02
Skewness	.1066-00	-.8874-01	.8463-01	.2118-00	.2739-00	.7570-00

Table 2.1--Continued
'Normal Group'

	Uric Acid	Creatinine	Total Bilirubin	Alkaline Phosphatase	LDH	SGOT
Mean	.5480+01	.9236-00	.7474-00	.6200+02	.1391+03	.4790+02
Second	.3145+01	.5057-01	.7588-00	.1300+04	.3159+04	.6683+03
Third	.4320+01	.1220-01	.4737+01	.1463+06	.6267+06	.4019+05
Fourth	.3732+02	.1430-01	.3619+02	.3025+08	.2180+09	.4328+07
Std Deviation	.1773+01	.2249-00	.8711-00	.3606+02	.5620+02	.2585+02
n	.2890+03	.2910+03	.2910+03	.2830+03	.2900+03	.2870+03
Kappa	-.2040+01	.6481-00	-.1664+02	.4511+02	.1248+03	-.3392+01
Delta	-.3749-01	.2017-00	-.5218-00	.2665-01	.1240-01	-.2247-00
Beta 1	.6002-00	.1151+01	.5135+02	.9744+01	.1246+02	.5412+01
Beta 2	.3773+01	.5594+01	.6284+02	.1789+02	.2185+02	.9692+01
Sqr(Beta 1)	.7747-00	.1073+01	.7166+01	.3121+01	.3530+01	.2326+01
Alpha 3	.7747-00	.1073+01	.7166+01	.3121+01	.3530+01	.2326+01
Alpha 4	.3773+01	.5594+01	.6284+02	.1789+02	.2185+02	.9692+01
Skewness	.4187-00	.3823-00	.8214+02	.1482+01	.1722+01	.2113+01

Table 2.2

Statistics used to classify distribution curves according to the Pearson and Burr systems.

'Abnormal Group'

	Total Protein	Albumin	Calcium	Cholesterol	Glucose	BUN
Mean	.6912+01	.4070+01	.9161+01	.2147+03	.1164+03	.1487+02
Second	.4143-00	.5221-00	.3830-00	.2685+04	.1531+04	.8780+02
Third	-.9802-01	-.2107-00	-.2353-01	.1102+06	.1841+06	.3479+04
Fourth	.6990-00	.9947-00	.1007+01	.3882+08	.3370+08	.2175+06
Std Deviation	.6436-00	.7226-00	.6189-00	.5182+02	.3912+02	.9370+01
n	.5770+03	.5780+03	.5780+03	.5770+03	.5730+03	.5770+03
Kappa	.6114-01	.6976-00	.1135-02	.1942-00	-.4385+01	-.2296+02
Delta	.2460-00	.5444-01	.7803-00	.3444-00	-.3210-00	-.1026-00
Beta 1	.1351-00	.3120-00	.9852-02	.6271-00	.9453+01	.1788+02
Beta 2	.4073+01	.3649+01	.6863+01	.5384+01	.1439+02	.2821+02
Sqr(Beta 1)	.3676-00	.5586-00	.9926-01	.7919-00	.3075+01	.4228+01
Alpha 3	-.3676-00	-.5586-00	-.9926-01	.7919-00	.3075+01	.4228+01
Alpha 4	.4073+01	.3649+01	.6863+01	.5384+01	.1439+02	.2821+02
Skewness	-.1232-00	-.2519-00	-.1938-01	.2344-00	.4294+01	.2660+01

Table 2.2--Continued
'Abnormal Group'

	Uric Acid	Creatinine	Total Bilirubin	Alkaline Phosphatase	LDH	SGOT
Mean	.6101+01	.1102+01	.8193-00	.7150+02	.1567+03	.5598+02
Second	.3238+01	.7621-00	.5795-00	.1745+04	.6272+04	.1490+04
Third	.2151+01	.5746+01	.2601+01	.2068+06	.1434+07	.1621+06
Fourth	.3647+02	.5071+02	.1557+02	.4310+08	.5261+09	.2780+08
Std Deviation	.1800+01	.8730-00	.7612-00	.4177+02	.7920+02	.3860+02
n	.5680+03	.5780+03	.5740+03	.5720+03	.5710+03	.5640+03
Kappa	.1939-00	-.2198+02	-.1485+02	-.9969+01	-.4612+01	-.3812+01
Delta	.8428-01	-.6107-00	-.3562-00	-.1066-00	-.2594-00	-.3076-00
Beta 1	.1363-00	.7460+02	.3477+02	.8052+01	.8331+01	.7937+01
Beta 2	.3477+01	.8732+02	.4636+02	.1416+02	.1337+02	.1252+02
Sqr (Beta 1)	.3692-00	.8637+01	.5897+01	.2838+01	.2886+01	.2817+01
Alpha 3	.3692-00	.8637+01	.5897+01	.2838+01	.2886+01	.2817+01
Alpha 4	.3477+01	.8732+02	.4636+02	.1416+02	.1337+02	.1252+02
Skewness	.1580-00	.1950+02	.1025+02	.1803+01	.3000+01	.3661+01

Table 2.3

Statistics used to classify distribution curves according to the Pearson and Burr systems.

'Total Group'

	Total Protein	Albumin	Calcium	Cholesterol	Glucose	BUN
Mean	.6934+01	.4098+01	.9170+01	.2123+03	.1115+03	.1437+02
Second	.3965-00	.4914-00	.3405-00	.2326+04	.1156+04	.6936+02
Third	-.4930-01	-.1731-00	-.1183-01	.8799+05	.1360+06	.2497+04
Fourth	.6260-00	.9393-00	.7366-00	.2933+08	.2512+08	.1528+06
Std Deviation	.6297-00	.7010-00	.5836-00	.4823+02	.3401+02	.8328+01
n	.8680+03	.8690+03	.8690+03	.8660+03	.8620+03	.8680+03
Kappa	.1628-01	.1979-00	.4555-03	.1835-00	-.8355+01	.5560+02
Delta	.2644-00	.1484-00	.7156-00	.3559-00	-.1988-00	.4118-01
Beta 1	.3899-01	.2524-00	.3541-02	.6154-00	.1197+02	.1869+02
Beta 2	.3981+01	.3890+01	.6351+01	.5422+01	.1878+02	.3176+02
Sqr(Beta 1)	.1975-00	.5024-00	.5951-01	.7845-00	.3459+01	.4324+01
Alpha 3	-.1975-00	-.5024-00	-.5951-01	.7845-00	.3459+01	.4324+01
Alpha 4	.3981+01	.3890+01	.6351+01	.5422+01	.1878+02	.3176+02
Skewness	-.6458-01	-.1937-00	-.1224-01	.2291-00	.2871+01	.1997+01

Table 2.3--Continued

Total Group

	Uric Acid	Creatinine	Total Bilirubin	Alkaline Phosphatase	LDH	SGOT
Mean	.5891+01	.1042+01	.7951-00	.6836+02	.1508+03	.5326+02
Second	.3293+01	.5309-00	.6410-00	.1617+04	.5293+04	.1228+04
Third	.2904+01	.3910+01	.3311+01	.1895+06	.1198+07	.1254+06
Fourth	.3720+02	.3466+02	.2237+02	.3953+08	.4368+09	.2085+08
Std Deviation	.1815+01	.7286-00	.8006-00	.4022+02	.7275+02	.3504+02
n	.8570+03	.8690+03	.8650+03	.8550+03	.8610+03	.8510+03
Kappa	.1225+01	-.3283+02	-.1682+02	-.1605+02	-.6492+01	-.5276+01
Delta	.2381-01	-.5289-00	-.3824-00	-.6843-01	-.2081-00	-.2271-00
Beta 1	.2362-00	.1022+03	.4163+02	.8486+01	.9684+01	.8497+01
Beta 2	.3431+01	.1230+03	.5445+02	.1511+02	.1559+02	.1383+02
Sqr (Beta 1)	.4860-00	.1011+02	.6452+01	.2913+01	.3112+01	.2915+01
Alpha 3	.4860-00	.1011+02	.6452+01	.2913+01	.3112+01	.2915+01
Alpha 4	.3431+01	.1230+03	.5445+02	.1511+02	.1559+02	.1383+02
Skewness	.2319-00	.8740+02	.1371+02	.1688+01	.2666+01	.2671+01

the difference of the mean and mode divided by the standard deviation. For my calculations:

$$(2.1) \quad \text{Skewness} = \frac{\sqrt{B_1} (B_2 + 3)}{2(5B_2 - 6B_1 - 9)}$$

Frequency curves often start at zero, rise to a maximum, and then fall at either the same rate or a different rate.

Pearson uses a differential equation

$$(2.2) \quad \frac{dy}{dx} = \frac{y(x+a)}{b_0 + b_1x + b_2x^2 + \dots}$$

which he found to cover a range of widely different shaped curves. E & J explain thoroughly the theory involving Equation (2.2). E & J, p. 41, also give the equation for 'the criterion'.

$$(2.3) \quad K = \frac{B_1(B_2 + 3)^2}{4(2B_2 - 3B_1 - 6)(4B_2 - 3B_1)}$$

Looking again at Tables 2.1 through 2.3, one can classify

each chemistry by the value of $K = \text{KAPPA}$. E & J have a most valuable table, p. 45, which explains the limits or values of K for each type of curve along with other information. For Type I K is negative, for Type IV K is greater than 0 and less than 1, and for Type VI K is greater than 1.

$$y = y_0 (1 + x/a_2)^{-m_2} e^{-\sqrt{2} \tan^{-1}(x/a_2)}, -\infty < x < \infty. \quad (2.5)$$

Let us now consider Type IV and the equation

If m_1 is negative and m_2 is positive, the resulting curve is J-shaped.
If m_1 and m_2 are both positive, the resulting curve is bell-shaped.

where $m_1/a_1 = m_2/a_2$.

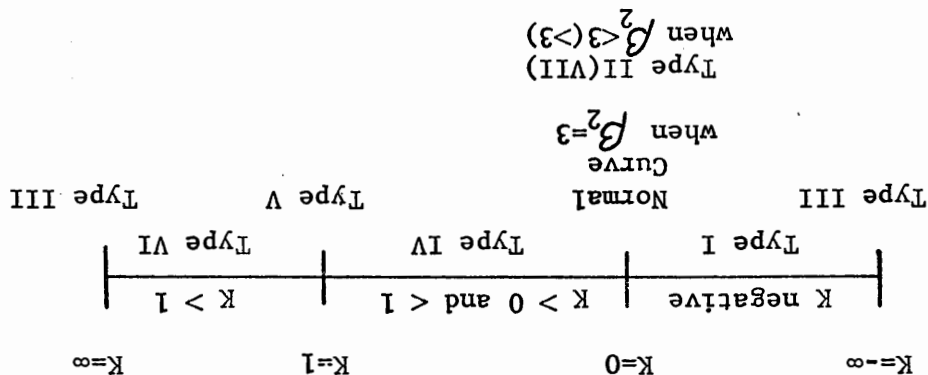
$$y = y_0 (1 + x/a_1)^{m_1} (1 - x/a_2)^{m_2}, -a_1 < x < a_2 \quad (2.4)$$

Let us consider Type I and the equation

see E & J.

are the transition curves. For more information of these curves, not overlap. The other curves shown besides Types I, IV, and VI how the types cover all possible values of the criterion and do Figure 1, which is taken directly from E & J, p. 49, shows

Figure 1: Kappa criterion.



This curve is always bell-shaped. μ_3 and ν have opposite signs.

Also let us consider Type VI and the equation

$$y = y_0 (x-a)^{q_2} x^{-q_1}, \quad a < x < \infty. \quad (2.6)$$

For this type q_1 is greater than q_2 . If q_2 is negative, the curve

is J-shaped.

Considering the three previous paragraphs, I have constructed a table, Table 2.4 which shows by my calculations the shape of the

curves of the data used.

Also Figures 2, 3, and 4 give the same conclusions.

Figure 2: A plot of (β_1, δ) for the twelve chemistries of the "normal group" which shows both the type of Pearson curve and whether the Burr can be used. (Scale not precise)

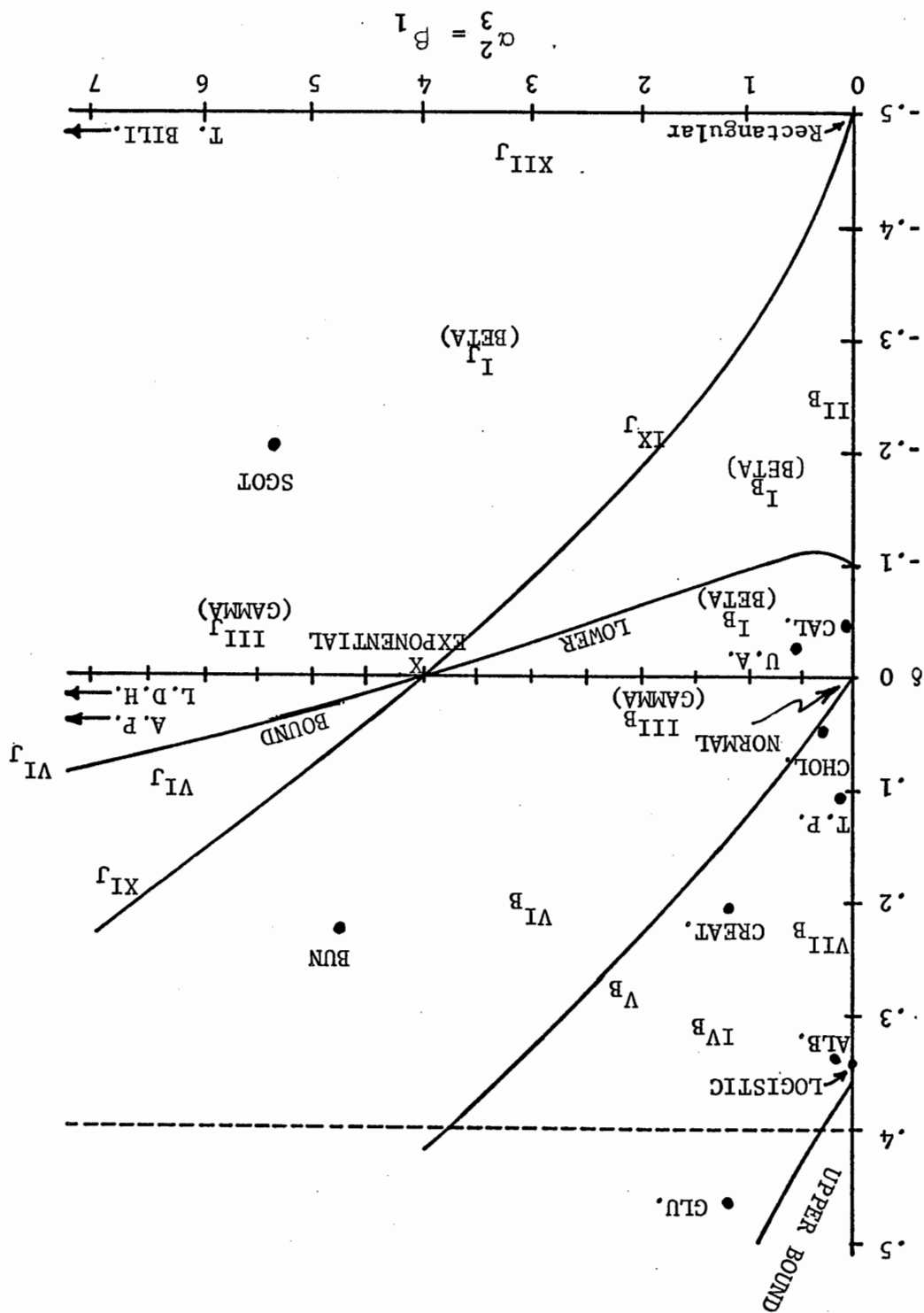
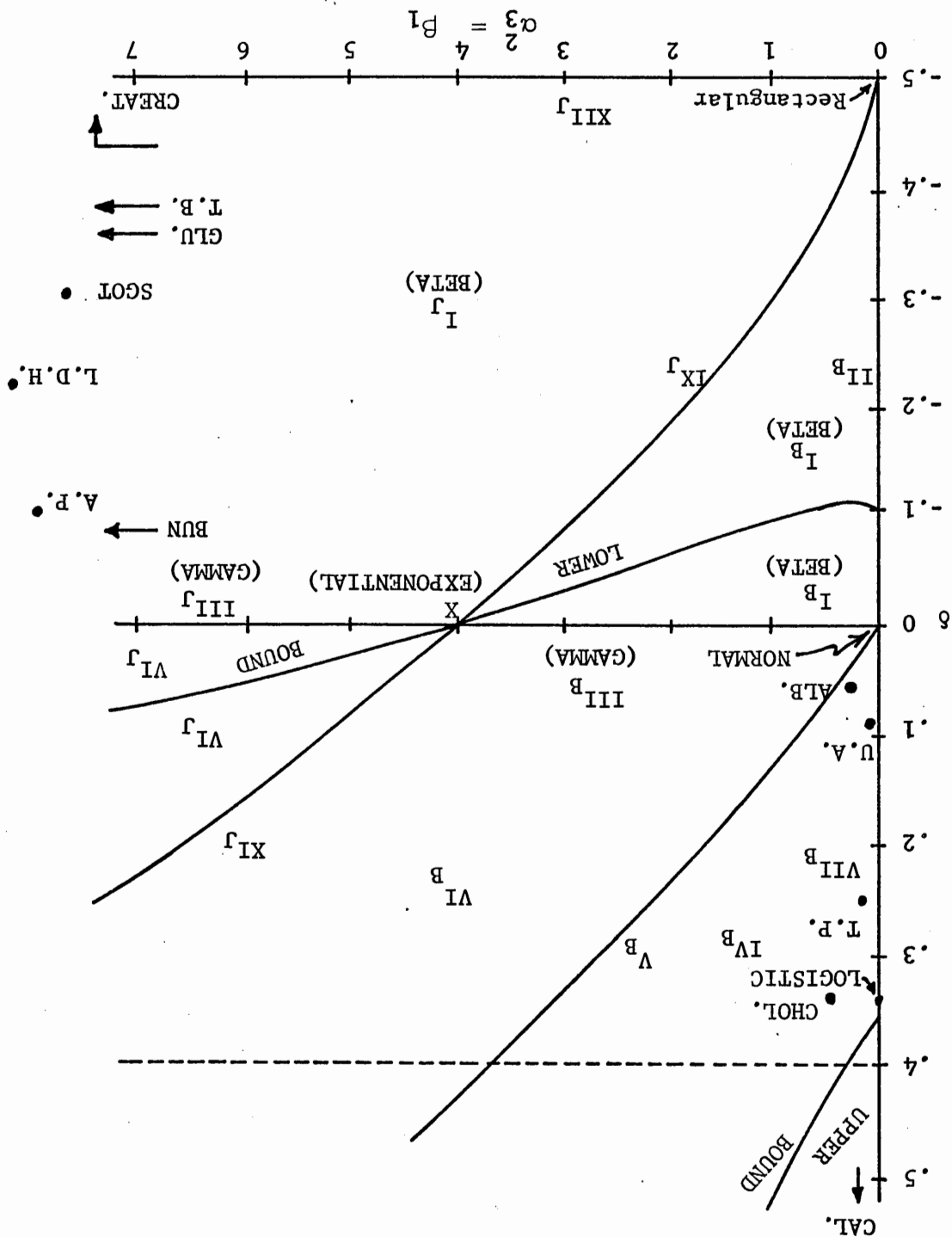
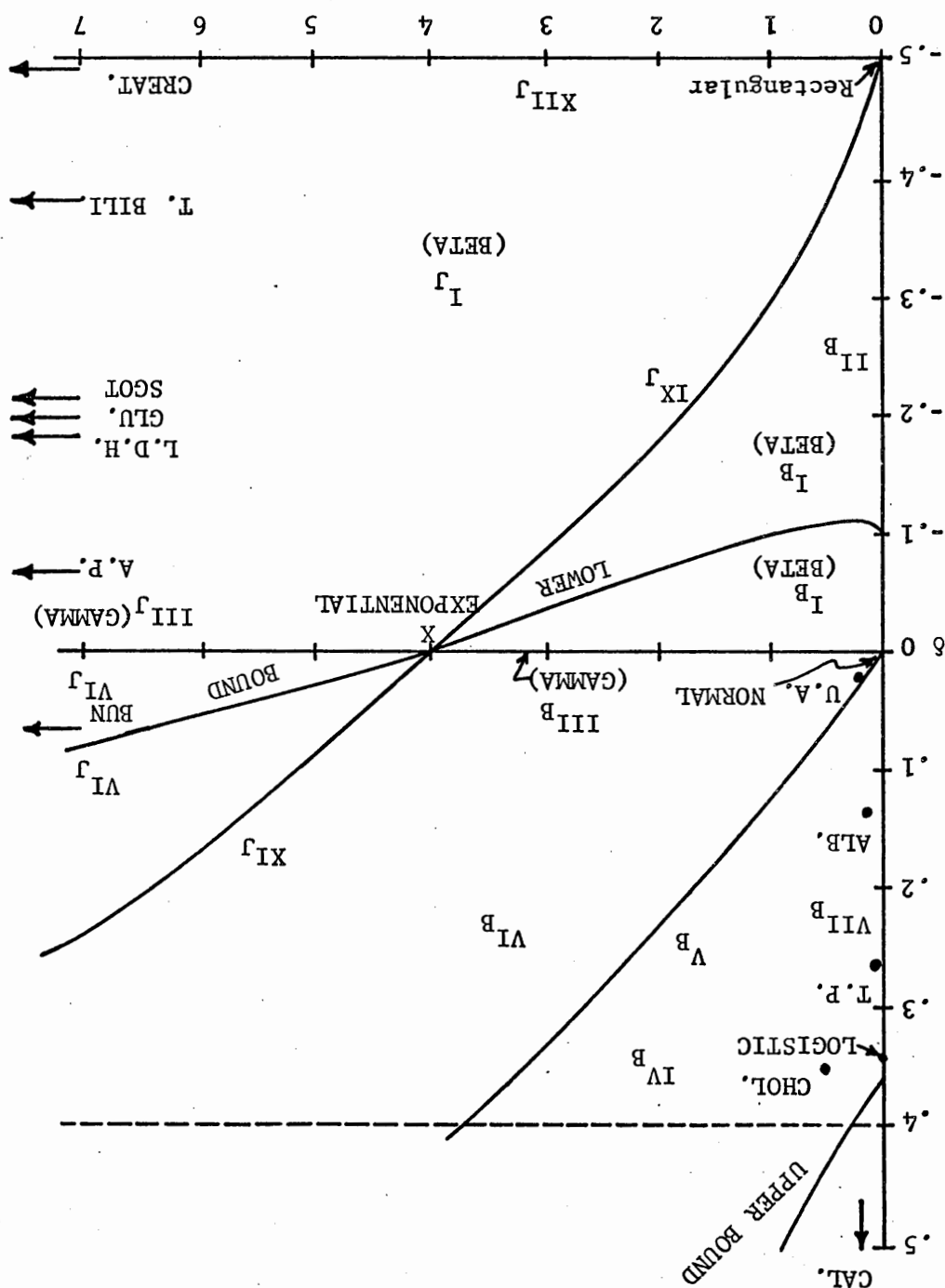


Figure 3: A plot of (β_1, δ) for the twelve chemistries of the "abnormal group" which shows both the type of Pearson curve and whether the Burr can be used. (Scale not precise)



$$I_B = \frac{3}{2} \alpha$$


15

Another means of approaching frequency curves was introduced by Irving W. Burr (1942) of Purdue University. Burr uses the equation

Table 2.4. A table of the type of Pearson curve for each of the twelve chemistries of the three groups.

Test	Normal Range	Value	Reference
Total Protein	6.0-8.3 g/dl	6.8	IV B
Albumin	3.5-5.0 g/dl	4.2	IV B
Calcium	8.8-10.0 mg/dl	9.2	IV B
Cholesterol	160-260 mg/dl	240	IV B
Glucose	70-110 mg/dl	95	IV B
BUN	8-20 mg/dl	12	VI B
Uric Acid	2.4-8.0 mg/dl	5.2	IV B
Creatinine	0.6-1.2 mg/dl	0.8	IV B
Total Bilirubin	0.1-1.2 mg/dl	0.2	IV B
Alkaline Phosphatase	44-146 U/L	110	VI B
LDH	100-250 U/L	180	VI B
SGOT	0-40 U/L	15	IV B

Considering the general characteristics of an empirical cumulative distribution, $F(x)$ is non-decreasing from $F(-\infty) = 0$ to $F(\infty) = 1$. Burr directly uses the cumulative distribution (2.7). He says that an advantage of the direct use of $F(x)$ is the fact that $f(x)$ is at least theoretically much simpler.

A differential equation which Burr uses to incorporate desirable properties of $F(x)$ is

$$\frac{dy}{dx} = y(1-y)g(x,y), \quad y = F(x) \quad (2.8)$$

where $g(x,y)$ is to be positive for $0 \leq y \leq 1$ and x in the range over which the solution is to be used. Burr explains that Equation (2.8) is similar to a differential equation which generates the Pearson system under certain conditions.

The Burr distribution (2.2) in the form of

$$F(z) = 1 - \left[1 + \left(\frac{\theta_2}{z + \theta_1} \right)^c \right]^{-k} \quad (2.9)$$

will be briefly discussed in the summary. Equation (2.9) increases the number of parameters by two, where θ_1 and θ_2 fix the origin and scale. Constants c and k are positive numbers.

Cecil C. Craig (1936) gives two characteristics for the Burr distribution which are valuable for graphing which then in turn would be a means of seeing if a set of data falls within the bounds

of the Burr distribution. One can plot α_3^2 against δ where $\alpha_3^2 = \beta_1$ in Pearson's notation and

$$\delta = \frac{(2\alpha_4 - 3\alpha_3^2 - 6)}{(\alpha_4 + 3)} \quad (2.10)$$

where

$$\alpha_1 = \frac{\mu_1}{\mu_1^{1/2} \mu_2} \quad (2.11)$$

which Burr and Cislak (1968) call the standardized central moments. Figures 2, 3, and 4 show that the Burr distribution covers portions of the three main types of Pearson's curves as well as some transitional types.

Looking at Figure 2, we see that eight chemistries are covered by the Burr distribution for the "normal group". Looking at Figures 3 and 4, we find only four chemistries (the same four) are covered for the "abnormal group" and the "total group".

TOLERANCE LIMITS

CHAPTER III

From (Burrington and May, 1970), I quote: "In recent times much effort has been placed on the development of statistical tests which are valid for situations in which little is known about the shape of the population distribution from which the samples are drawn. One class of such tests valid for samples from continuous population distributions of any shape, is called 'non-parametric tests' (or 'distribution-free tests'). They are based on order statistics."

For the analysis of the twelve chemistry tests of the SMA 12/60, the probability that at least a fraction β of a population of a random variable 'x' lies between A and B of a sample for each chemistry is at least α . A and B are the k th and $(n-k+1)$ th order statistics. The fraction β of the population values lying between A and B is $F = \int_A^B f(x)dx$. Since A and B are the k th and $(n-k+1)$ th order statistics in a sample size n , let us consider the distribution of $F(t)$ where t is the k th order statistic from a sample of size n . F has the density function

$$\frac{n!}{F^{k-1} [1-F]^{n-k}}, \quad 0 < F < 1.$$

Owen (1962, p. 321) uses Equation (3.1) to determine a sample size n if α and β are given.

²S. B. Burrington and D. C. May, Handbook of Probability and Statistics with Tables, 2nd ed., (New York: McGraw-Hill Book Company, 1970), p. 9.

$$n \beta^{n-1} \beta - (n-1) \beta^n \leq 1 - \alpha \quad (3.1)$$

Burington and May (1970) use the form

$$1 - \beta^n - n(1 - \beta) \beta^{n-1} = \alpha. \quad (3.2)$$

When n is large and β is close to 1, Equation(3.2) can be approxi-

mated by

$$e^{-n(1 - \beta)} [1 + n(1 - \beta)] = 1 - \alpha. \quad (3.3)$$

Ordinarily, values for α and β are near unity.

The three groups mentioned earlier with which I worked had

uncensored sample sizes of 291, 578, and 869. I found from Owen

(ibid.) that I could use combinations of α and β up to .98 and .98,

respectively.

We have been considering the equation,

$$P \left[\int_A^B f(x) dx \geq \beta \right] = \alpha. \quad (3.4)$$

Equation (3.4) written with order statistic notation is

$$P [F(x) > p] = \int_0^p \frac{n!}{(k-1)!(n-k)!} F^{k-1}(x) [1-F(x)]^{n-k} dF(x). \quad (3.5)$$

Tables 3.1 through 3.3 show the values of A and B. The values of the "normal group" will be the tentatively accepted values for the "normal values" for each chemistry by the method used. The values of the "abnormal group" and the "total group" can be used for comparisons for the time being.

$$\sum_{k=0}^x \binom{n}{k} p^k (1-p)^{n-k} \leq \frac{1}{2} (1 - \alpha).$$

n, a k was found so that (Harvard, 1955) and a two tailed test, and by interpolation for each The observed data were put in rank order. By using a book of tables The sum (3.6) is the form which I used for my computations. Bernoulli trials, each trial with a probability p of success. is exactly the probability of at least k "successes" throughout n

$$(3.6) \quad \sum_{x_1=k}^n \frac{n!}{x_1!(n-x_1)!} p^{x_1} (1-p)^{n-x_1}$$

sometimes easier to compute. The sum This integral (3.5) can be expressed in a form which is

Table 3.1

Tolerance limits, A and B, are the k th and $(n-k+1)$ th order statistics for given α and β .

'Normal Group'

$\alpha = .90$

β	n														
		291	291	291	289	289	291	289	291	289	291	291	283	290	287
.90		Total Protein gm %	Albumin gm %	Calcium mg %	Cholesterol mg %	Glucose mg %	BUN mg %	Uric Acid mg %	Creatinine mg %	Total Bilirubin mg %	Alkaline Phos phatase mU./ml	LDH mU./ml	SGOT mU./ml		
		6.1 7.95	3.15 5.2	8.48 9.95	150 270	83 128	7.2 20	3.2 8.75	.65 1.25	.3 1.3	30 105	90 213	26 92		
.95		5.85 8.15	2.8 5.48	8.25 10.15	138 290	78 147	6 27.5	2.65 9.8	.55 1.45	.25 1.75	25 167	77 284	23 127		
		5.65 8.4	2.1 5.7	8.02 10.5	117 330	55 160	4 39	1.85 10.8	.5 1.7	.2 8	18 235	71 456	20 174		
.98															

Table 3.1--Continued

'Normal Group'

$\alpha = .95$

β	n	291	291	291	289	289	291	289	291	291	283	290	287
		Total Protein gm %	Albumin gm %	Calcium mg %	Cholesterol mg %	Glucose mg %	BUN mg %	Uric Acid mg %	Creatinine mg %	Total Bilirubin mg %	Alkaline Phosphatase mU./ml	LDH mU./ml	SGOT mU./ml
.90	6.1	3.15	8.45	150	83	7	3.1	.65	.3	28	87	26	
	8	5.22	10	270	128	20.5	8.8	1.28	1.35	107	217	92	
.95	5.85	2.75	8.25	137	77	5.5	2.5	.55	.25	24	75	22	
	8.15	5.52	10.2	293	148	30.5	10.25	1.5	2.1	174	285	133	
.98	5.65	2.1	8.02	117	55	4	1.85	.5	.2	18	71	20	
	8.4	5.7	10.5	330	160	39	10.8	1.7	8	235	456	174	

Table 3.1--Continued

'Normal Group'

$\alpha = .98$

β	n												
		291	291	291	289	289	291	289	291	291	283	290	287
		Total Protein gm %											
		Albumin gm %											
		Calcium mg %											
		Cholesterol mg %											
		Glucose mg %											
		BUN mg %											
		Uric Acid mg %											
		Creatinine mg %											
		Total Bilirubin mg %											
		Alkaline Phosphatase mU./ml											
		LDH mU./ml											
		SGOT mU./ml											
.90	6.1 8	3.12 5.25	8.45 10	150 275	82 130	7 20.8	3 9.1	.62 1.3	.3 1.4	27 115	85 217	25 94	
.95	5.85 8.15	2.7 5.55	8.25 10.2	137 296	76 148	5 31.5	2.3 10.3	.55 1.55	.25 3.2	23 182	75 285	22 137	
.98	5.4 9.5	1.25 5.9	7.6 10.7	115 360	43 187	4 49	1.8 11.85	.5 2.05	.15 9.45	18 335	67 577	18 175	

Table 3.2

Tolerance limits, A and B, are the kth and (n-k+1)th order statistics for given α and β .

'Abnormal Group'

$$\alpha = .90$$

β	n												
		577	578	578	577	573	577	568	578	574	572	571	564
		Total Protein gm %											
		Albumin gm %											
		Calcium mg %											
		Cholesterol mg %											
		Glucose mg %											
		BUN mg %											
		Uric Acid mg %											
		Creatinine mg %											
		Total Bilirubin mg %											
		Alkaline Phosphatase mU./ml											
		LDH mU./ml											
		SGOT mU./ml											
.90	6 7.8	3 5.04	8.4 9.95	150 290	87 165	7.5 23.5	3.65 8.7	.7 1.4	.35 1.3	37 127	97 255	27 108	
.95	5.7 8	2.5 5.35	8 10.18	130 313	83 235	6.5 34.2	2.85 9.9	.65 1.9	.3 2	29 180	84 367	24 158	
.98	4.98 8.5	2.1 5.6	7.62 10.8	110 385	77 303	4 59	2.3 11.5	.55 4	.3 5.5	22 240	70 553	20 234	

Table 3.2--Continued
'Abnormal Group'

$\alpha = .95$

β	n												
		577	578	578	577	573	577	568	578	574	572	571	564
		Total Protein gm %											
		Albumin gm %											
		Calcium mg %											
		Cholesterol mg %											
		Glucose mg %											
		BUN mg %											
		Uric Acid mg %											
		Creatinine mg %											
		Total Bilirubin mg %											
		Alkaline Phosphatase mU./ml											
		LDH mU./ml											
		SGOT mU./ml											
.90	6	3	8.4	150	87	7.5	3.65	.7	.35	37	95	27	
	7.8	5.05	9.95	292	165	24	8.7	1.45	1.3	128	255	112	
.95	5.6	2.45	7.95	130	82	6.2	2.8	.65	.3	28	83	24	
	8	5.35	10.2	315	235	35	9.9	2	2.05	183	375	162	
.98	4.9	2.05	7.4	106	75	4	2.2	.5	.3	17	70	20	
	8.5	5.62	11	390	305	74.5	11.7	7.9	5.7	267	560	235	

Table 3.2--Continued
'Abnormal Group'

$\alpha = .98$

β	n	577	578	578	577	573	577	568	578	574	572	571	564
		Total Protein gm %	Albumin gm %	Calcium mg %	Cholesterol mg %	Glucose mg %	BUN mg %	Uric Acid mg %	Creatinine mg %	Total Bilirubin mg %	Alkaline Phosphatase mU./ml	LDH mU./ml	SGOT mU./ml
.90	6	2.85	8.3	147	87	7.5	3.55	.7	.35	37	95	27	
	7.8	5.1	9.95	293	168	25	8.75	1.45	1.35	132	262	119	
.95	5.48	2.38	7.9	126	82	6	2.7	.65	.3	28	82	24	
	8.05	5.4	10.2	317	235	35.8	9.95	2.05	2.1	188	410	170	
.98	4.9	2.05	7.35	96	74	3	2	.5	.3	17	70	19	
	8.6	5.65	11.05	390	310	78	11.7	9.05	6.82	277	574	237	

Table 3.3

Tolerance limits, A and B, are the kth and (n-k+1)th order statistics for given α and β .

'Total Group'

$$\alpha = .90$$

β	n	868	869	869	866	862	868	857	869	865	855	861	851
		Total Protein gm %											
		Albumin gm %											
		Calcium mg %											
		Cholesterol mg %											
		Glucose mg %											
		BUN mg %											
		Uric Acid mg %											
		Creatinine mg %											
		Total Bilirubin mg %											
		Alkaline Phosphatase mU./ml											
		LDH mU./ml											
		SGOT mU./ml											
.90		6.1 7.8	3.12 5.05	8.4 9.9	152 278	87 148	7.5 21.5	3.55 8.6	.7 1.35	.35 1.25	35 116	95 234	27 97
.95	8	5.82	2.6 5.35	8.2 10.15	137 308	82 184	6.5 30.5	2.88 9.7	.65 1.65	.3 1.85	27 165	84 327	24 144
.98		5.35 8.3	2.2 5.6	7.75 10.5	115 350	76 267	4.5 45.5	2.3 10.6	.55 2.4	.25 3.95	22 227	74 474	22 203

Table 3.3--Continued

'Total Group'

$\alpha = .95$

P	n	868	869	869	866	862	868	857	869	865	855	861	851
		Total Protein gm %	Albumin gm %	Calcium mg %	Cholesterol mg %	Glucose mg %	BUN mg %	Uric Acid mg %	Creatinine mg %	Total Bilirubin mg %	Alkaline Phosphatase mU./ml	LDH mU./ml	SGOT mU./ml
.90		6.1 7.8	3.1 5.1	8.4 9.95	150 280	86 150	7.5 21.5	3.5 8.6	.7 1.35	.35 1.28	35 178	95 235	27 98
.95		5.8 8.05	2.55 5.35	8.2 10.15	136 310	82 185	6.5 31	2.85 9.75	.65 1.7	.3 1.85	27 167	83 335	24 147
.98		5.15 8.4	2.12 5.6	7.7 10.6	115 360	75 277	4 49	2.25 10.8	.5 3.05	.25 5.5	22 235	71 480	20 220

Table 3.3--Continued

'Total Group'

$\alpha = .98$

B	n												
		868	869	869	866	862	868	857	869	865	855	861	851
		Total Protein gm %											
		Albumin gm %											
		Calcium mg %											
		Cholesterol mg %											
		Glucose mg %											
		BUN mg %											
		Uric Acid mg %											
		Creatinine mg %											
		Total Bilirubin mg %											
		Alkaline Phosphatase mU./ml											
		LDH mU./ml											
		SGOT mU./ml											
.90		6.05 7.8	3.1 5.1	8.4 9.95	150 280	86 150	7.5 22	3.45 8.7	.7 1.35	.35 1.3	34 123	95 240	27 100
.95		5.75 8.05	2.5 5.4	8.15 10.15	135 310	82 202	6 31.5	2.8 9.8	.6 1.7	.3 1.9	27 170	82 340	24 148
.98		5.1 8.5	2.1 5.62	7.68 10.65	114 360	74 294	4 52	2.2 11.3	.5 3.08	.25 5.7	20 240	70 523	20 225

Just a comment of the difference between a tolerance interval and a confidence interval: a tolerance interval involves a random variable 'x' while a confidence interval generally involves parametric values such as the mean, standard deviation, etc.. A sentence concerning a tolerance interval would state: The probability is α that at least β fraction of a population of a random variable 'x' would fall between two values A and B. For a confidence interval: The probability is α that the true value of a parameter, say μ , lies between two values A and B. A and B are calculated from the observed data in both instances.

CHAPTER IV

SUMMARY

Using criteria external to the twelve chemistries, patients were placed into two groups, "normal" and "abnormal". The analyses which were done on these two groups as well as on the combined group included classifications according to the Pearson system of curves, the Burr distribution, and the employment of a non-parametric procedure. Some chemistry values had to be censored due to the limited range of the SMA 12/60 chart.

By the method used, the data reveal coverage of Types I, IV and VI of the Pearson system of frequency curves. But the most interesting revelation is shown by the J-shaped curves. Looking at Table 2.4 and considering the "normal group", there are four J-shaped curves which show that most results cluster around smaller observed values.

Figure 2 shows that there are eight chemistries of the "normal group" which can be illustrated by the Burr distribution. There are four chemistries in each of the other two groups which can be described by the Burr distribution. This is shown in Figures 3 and 4. For the data used, these all happen to be bell-shaped curves.

Tolerance limits were found using order statistics and a non-parametric procedure.

Some laboratories use $\bar{X} \pm 2SD$ in order to find normal values. Table 4.1 gives values of $\bar{X} \pm 2SD$ and of the smallest and largest observed values of four chemistries of the "normal group", two of which are bell-shaped and two of which are J-shaped.

	Largest Observed Value	Smallest Observed Value	$\bar{X} + 2SD$	$\bar{X} - 2SD$
Total Protein	9.5	5.4	8.17	5.78
Cholesterol	360	115	287	128
LDH	577	67	252	26.7
SGOT	175	18	99.6	-3.8

Table 4.1. Comparison of some of the observed data with the traditional tolerance limits for the "normal group".

By comparing this table with Table 3.1, one can see that there is not a huge difference for the bell-shaped curves but there is a great discrepancy for the J-shaped curves. As can be seen, the lower value ($\bar{X} - 2SD$) for LDH is 26.7 when in fact the smallest observed value is 67. For SGOT, the lower value ($\bar{X} - 2SD$) is less than zero when the smallest observed value is 18.

For future projects, histograms could be plotted from the observed data, and frequency curves could be drawn by using the Pearson curves. A goodness of fit test could be done to see if

the actual frequency curve deviates significantly from the Pearson curve which was fitted to the data used.

Possibly in the future an easy means of finding tolerance limits for the Pearson curves will be found.

Percentage points of the Burr distribution, Equation (2.9), could be found by employing techniques of linear regression and by finding the maximum likelihood estimates of θ_1 , θ_2 , c , and k .

Then the Burr distribution could be plotted and confidence limits could be found for certain percentage points. When parameters are estimated, as in this case, then the tolerance limits are one-sided confidence limits of the percentage points.

The data could also be broken down into categories such as age, sex, etc..

A correlation analysis could be made to see if there are possible relations of one chemistry to another.

The data accumulated from the record room of St. Paul Hospital have been punched on cards and they will be kept on file for 5 years. At present the data are also on tapes.

This paper is only a sample of what could be done with the chemistries in a laboratory and of finding "normal values" of these chemistries. In the future I hope to be able to expand the area of investigation and improve on what I have begun. This has been a fascinating area for me to explore and hopefully this paper will help to stimulate an interest in continuing research in this field.

BIBLIOGRAPHY

- Burington, S. B. and May, D. C. Handbook of Probability and Statistics with Tables. 2nd ed. New York: McGraw-Hill Book Company, 1970.
- Burr, I. W. "Cumulative Frequency Functions," Annals of Mathematical Statistics, 13 (1942), 215-32.
- Burr, I. W. and Cislak, P. J. "On a General System of Distributions I. Its curve-shape Characteristics," Journal of the American Statistical Association, 62 (1968), 627-635.
- Craig, C. C. "A New Exposition and Chart for the Pearson System of Frequency Curves." Annals of Mathematical Statistics, 7 (1936), 16-28.
- Elderton, W. P. and Johnson, N. L. Systems of Frequency Curves. Cambridge: The University Press, 1969.
- Tables of the Cumulative Binomial Probability Distribution. Cambridge, Massachusetts: Harvard University Press, 1955.
- Owen, D. B. Handbook of Statistical Tables. Addison-Wesley, 1962.