STATISTICS IN BALLISTICS

Applications of Mathematical Statistics to Dispersion in Exterior Ballistics

Lewis M. Campbell

PREFACE

This work is about applications of mathematical statistics to ordnance, especially to the dispersion of rounds in the pattern of fall of shot. Although mathematics is used here, this is not a book of mathematics. No mathematical proofs are offered.

This book is intended for engineers, scientists, and other technical workers engaged in designing, developing, testing or maintaining ordnance systems. Those systems usually involve the fire or delivery of an ordnance warhead onto a target. Such a delivery process always involves an error, to a greater or lesser extent, in the placement of the warhead. The dispersion of the shot, rounds, or warheads is the subject of this work.

References are given at the end of each chapter and at the end of the book to more extensive treatments of the several statistical methods used here.

The writer wishes to acknowledge the many insights gained from co-workers over 29 years at the Naval Ordnance Laboratory in White Oak, Maryland. Dr. Russell Glock was very helpful in explaining several applications of statistics. Dr. Bordelon's warnings about the pitfalls and subtleties of probability and statistics often come to mind. Dr. Cohen's guidance was helpful, and Dr. van Tuyl's encouragement to investigate computer graphics made this book possible. Dr. Ellingson's kindly scholarship is missed. He was a master algebraist; I have not met his equal at that art. Jim Bob McQuitty was one of the very best at applied mathematics. He too, is missed. Col. E. H. Harrison, U.S. Army, Rtd. and Senior Technical Editor of the American Rifleman magazine recognized and encouraged my efforts.

Many librarians have contributed their skilled services to this work. All the staff of the Technical Library at the Naval Ordnance Laboratory, White Oak, Maryland were helpful over many years. Mr. Wes Price at the Dahlgren Laboratory has provided his assistance. Mr. Neil Young of the

Imperial War Museum, London and Mr. Russell Lee of the Smithsonian Air and Space Museum shared their understanding of a term used in aircraft bombing. Shirley and Linda at the Smoot Library in King George, Virginia have cheerfully aided my searches.

Doug Miller has kindly reviewed the equations herein.

Of course, I owe all to the patience of my wife Judith.

Editor's Note

The author, Lewis Michael Campbell, passed away in 2019. This work was collected and edited by his son, Steven Campbell. The material was composed of images, spreadsheets, and text that utilized software that is no longer commonly available. Any errors or omissions are almost certainly the fault of the editor.

CONTENTS

INTRODUCTION

This book is intended for weapons officers, artillery officers, fire controllers, strike planners, and damage assessment analysts. In addition, ordnance program managers and others will benefit from study of the methods presented.

Application of mathematical statistics is the subject in this work. This involves a review of the normal probability distribution. That review has been kept short, as most readers of this book will have studied at least an elementary course in statistics, and many will have completed more advanced work.

Discussion begins with a graph illustrating a typical pattern of fall of shot, or impact pattern of a group of rounds of any kind. Variance, standard deviation and circular error probable (CEP) are mentioned as being usual measures of scatter. Sample mean and median are discussed in their role as common measures of central tendency. The use of small arms as models of larger guns leads to a comparison of large versus small gun designs. Similarity of chemistry of propellants is observed. Some current gun development efforts are mentioned. Conventional terms used in discussions of graphs are noted.

Normal, bivariate normal, and circular normal probability density distributions are introduced and discussed. The Rayleigh density distribution is presented as a development of the circular normal. The concept of the circular error probable (CEP) is introduced. Notes and references are given.

Use of the term CPE versus CEP is discussed. Circular error probable is derived and shown on a contour plot. The ratio of radius from center of impact to circular error probable (R/CEP) is developed .Graphs and tables of the relationship between probability and (R/CEP) are presented.

Some current ordnance developments are discussed. "Accurate" versus "Very accurate" versus "Precision" weapons are defined. Examples of applications of the tables are given. Effects of launch damage and long-term storage are considered. Non-circular patterns of dispersion are mentioned.

Calculation of the CEP from test data is discussed. An example of the calculation is given, and a table is presented to be used in correcting a bias found in the ordinary method for calculation of standard deviation.

Another characteristic of patterns of impact of rounds is the central tendency, or tendency of the points to cluster around a particular point or line. The median is introduced as an alternative to the mean for description of central tendency. An example from small arm sight adjustment is presented. The procedure is applicable to adjustment of sights for any direct-fire weapon.

Regarding the question of mean versus median in practical applications of statistics, the work of Tokishige Hojo is reviewed. A numerical study shows the accuracy of his calculations. The median is shown to be satisfactory for many if not all applications.

Nonparametric statistics is considered. The work of S. S. Wilks is applied to an example from manufacturing production quality control. The larger sample sizes needed in the nonparametric approach are balanced against freedom from having to first identify the underlying probability distribution.

An application of nonparametric statistics is made to the extreme spread of rounds upon a target. The relationship among confidence, sample size and population coverage is shown.

Sturges' rule for plotting of histograms is discussed. A table of sample size versus quantity of bins recommended is given. The importance of Sturges' rule as a symmetry test and in studying the probability distribution underlying a sample is emphasized. The probable derivation of Sturges' rule from the binomial distribution is given. This derivation is from the late Mr. G. J. Bradley, by kind permission of Mr. J. R. King. A sketch of the details of the computation of the table is given.

Extreme-value probability theory and plotting is applied to the extreme spreads from small arms targets. Comparison of the histogram to a graph of the extreme-value probability density function indicates that an extreme-value plot might be applicable. The plot of extremes is made, and examples are given of the estimates which may be made.

An application of statistical analysis to a psychovisual experiment is made. The experiment compares mean position of a small sample with the center estimated by eye. The result indicates that the human eye is quite accurate in its estimation of position.

CHAPTER 1

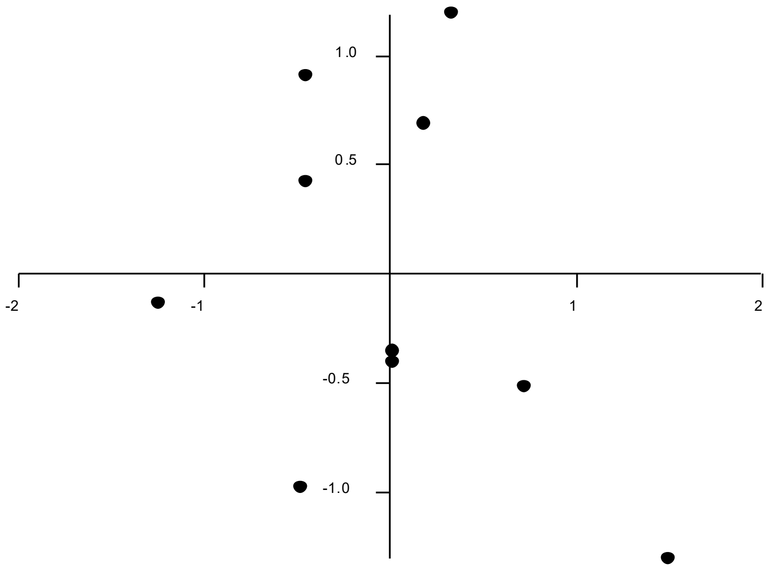
AN EXAMPLE OF DISPERSION OF ROUNDS IN EXTERIOR BALLISTICS

Figure 1.1 shows a typical pattern of fall of shot, or impact pattern of a salvo or group of rounds. This pattern is similar to those found on targets fired on small arms firing ranges. The ten rounds are scattered randomly across the target. In the case of small arms, the scale markings might represent one-inch divisions on the target. In the case of aircraft bombs, naval gunfire or ground forces artillery, the scale divisions would likely be meters or yards. The statistical methods used to describe the dispersion of rounds is the same in all these cases.

An obvious characteristic of the pattern is the manner in which the individual rounds are scattered across and along the target plane. The measure of amount of scatter is the variance or its square root, the standard deviation. A measure of scatter which is more often used in military and naval parlance is the circular error probable or CEP. The CEP is discussed more fully in chapters following.

A larger number of rounds, or sample, would show that the group of rounds has a tendency to cluster toward the center. There are several measures commonly used to define the center of the group, with the mean and median being used most often. The mean is simply computed, being the average distance of the positions of the rounds as measured along a given axis. The mean is the measure of central tendency used in the normal probability distribution, which will be familiar to the readers of this work. A brief review of that distribution will be given in the following chapter. Those readers whose memory needs no refreshing may simply pass over the discussion there.

FIGURE 1.1 TYPICAL PATTERN OF FALL OF SHOT.



In this work, the writer often has used small arms fire, or the pattern of bullet holes upon a plane target as an example for statistical analysis. In such a case, the small arm is a convenient model for any larger size gun. Gun designs of all sizes use the same materials except that wood is used only in some small arms today. The chemistry of propellants used in large guns is the same as that for small arms, with two exceptions. The current developments of liquid propellants and electrothermal-chemical (ETC) propellant systems have no counterparts in small arms. But the primary difference between the trajectory of small arms and that of large guns is the much greater range of the large gun projectile. For accurate fire at long range, the gunner must correct for both the curvature of the earth's surface and the rotation of the earth (the Coriolis correction). Neither correction is made for small arms fire. Other than the differences noted, the small arm is an excellent model for the study of fire from larger guns.

A note on graphs. In the discussions on graphs, the abscissa may be referred to as the "x" axis, or the lateral direction. The ordinate may be called the "y" axis or vertical axis if a vertical target is considered or the range axis if surface bombardment is the subject of discussion.

CHAPTER 2

THE NORMAL PROBABILITY DISTRIBUTION

Introduction

Figure 1.1 shows how bullet holes or round impacts appear on a plane target or plane surface around a target or aimpoint. The following paragraphs address application of the normal probability distribution to describe the dispersion of rounds.

NORMAL (GAUSSIAN) PROBABILITY DISTRIBUTION

It is established that dispersion of rounds of gunfire at moderate ranges is well described by the normal or Gaussian probability distribution. This distribution is often written as follows:

$$p(x) = (1/s(2\_)^{1/2}) * (e^{-((x-m)^{\wedge}2)}/2s^2 \qquad \textbf{EQ.2.1}$$

where:

p(x) = normal probability density function

x = a random variate. x may take on any value between minus to plus infinity.

s = standard deviation. A measure of the magnitude of dispersion of the variable. The standard deviation is always a positive number.

m = mean of the distribution. The mean may be any number.

e = 2.718..._ = 3.14159... (of course)

^ means: 'raised to the power of'

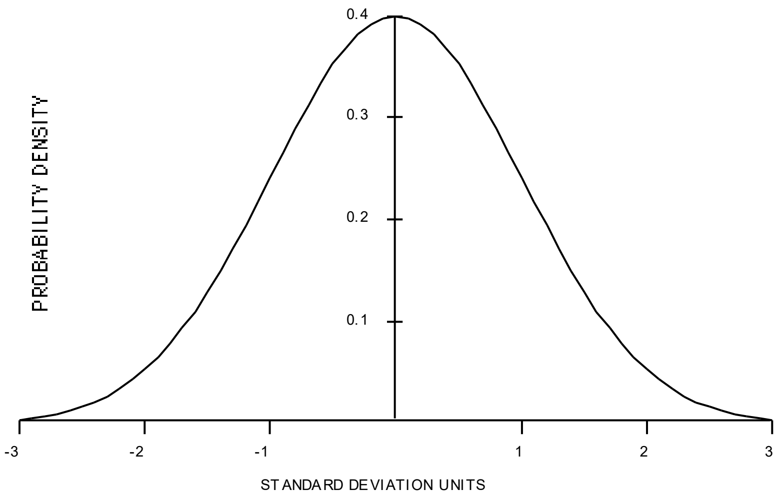
The function p(x) is called a density function by analogy to mechanics or mathematical dynamics, in which the density of a given body is defined by a mathematical function, and the mass of the body is computed by integrating the density function throughout the volume of the body.

The so-called 'standard' normal probability density function is described and tabulated in many references. The standard normal probability density function is computed by setting the mean to zero and the standard deviation equal to unity. Figure 2.1 shows a graph of the standard function. The peak of the function is equal to 0.3989... .

Equation 2.1 describes the dispersion of the rounds along the abscissa in Figure 1.1. A separate probability density function will describe the dispersion of the rounds in the ordinate direction in Figure 1.1. It would be the same as EQ.2.1, but with y replacing x, and perhaps other values for standard deviation and mean.

In the world in which naval and military operations take place, every observation is influenced by many random factors or errors. These random errors introduce uncertainty into our our measurements and predictions. Hence, designers of sensor or detection systems speak of 'noise in the system'. That is to say, our every observation or measurement is in fact a function, to a greater or lesser extent, of multiple random variates. In the simplest multivariate case, the random part of the observation depends upon only two variables, for example, our x and y above. If also the variables are normally distributed, the

FIGURE 2.1 STANDARD NORMAL PROBABILITY DENSITY



resulting probability distribution is often termed bivariate normal. In many practical cases, the dispersion in one direction is statistically independent of the dispersion in other directions. That is to say, given knowledge of the value of the horizontal position of a certain round, one cannot predict the vertical position of that or any other round. Conversely, one cannot predict the horizontal position from knowledge of the vertical position. If, in addition to the circumstance of independent variables, one has the standard deviation of the horizontal variate (our x) equal to the standard deviation of the vertical variate (our y), the resulting probability density function is called circular normal. An equation for the circular normal probability density function having zero means is as follows:

$$p(x,y) = (1/2ps^2)*(e^{-((x^2+ y^2)/2s^2)}) \qquad \textbf{EQ.2.2}$$

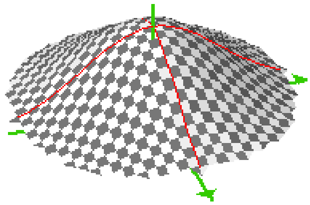
In EQ.2.2, s is the standard deviation of either x or y, as the two are equal.

Other terms have the same meanings as before. The means of both x and y are zero, and do not appear in EQ.2.2.

Figure 2.2 shows how a plot of EQ.2.2 might look if viewed from above and to one side. The plot is made for values of x and y between minus 2 and plus 2 and with standard deviation of 1.0. In the figure, the function p(x,y) is multiplied by 10 to make the shape of the function more apparent.

FIGURE 2.2 BIVARIATE (CIRCULAR) NORMAL PROBABILITY DENSITY

WITH X AND Y INDEPENDENT, MEANS ZERO, AND STANDARD DEVIATIONS EQUAL TO 1.0.



As the range of both x and y is limited, Figure 2.2 does not show the manner in which the function approaches the x,y plane as x and y become large and p(x,y) approaches zero. However, the plot does give the reader a general idea of the function and its symmetry. The plot appears much like a small mound or hillock situated upon a plain. The function has a low peak in the center equal to 1/2p, approximately 0.159. The probability that a point lies in any given area of the x,y plane is found by computing the volume that lies above the given area and below the p(x,y) probability density surface.

This circular normal probability density function may be expressed in polar coordinates as:

$$p(R) = (R/s^2)*(e^{-((R\wedge 2)/2s\wedge 2)}) \qquad \textbf{EQ.2.3}$$

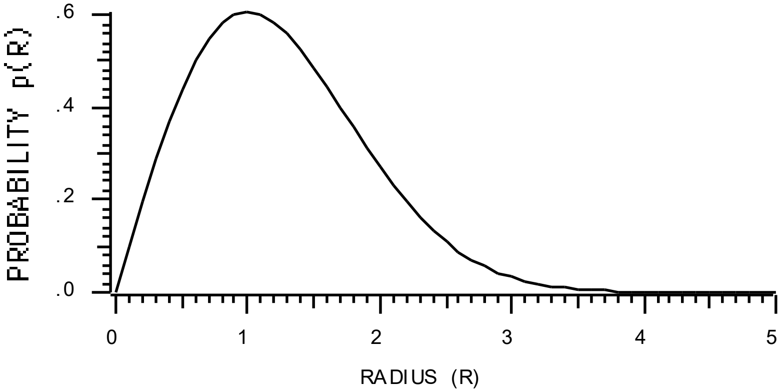
where $R^2 = x^2 + y^2$; mean of x = mean of y = 0,

and ^ means 'raised to the power of', as before.

R is the radius vector from the origin of the coordinate system and s is the common standard deviation as in p(x,y).

The polar form of the circular normal probability function is sometimes known as the Rayleigh distribution, and has a number of applications. For example, it is used to describe the characteristics of reverberation in active sonar transmission. Figure 2.3 shows a graph of how the probability varies as a function of radius R.

FIGURE 2.3 RAYLEIGH PROBABILITY DENSITY FUNCTION
p(R) = R(EXP(-(R^2)/2))



The probability P associated with a given value of R is found by integrating p(R) (EQ. 2.3) from zero to R. The result is:

$$P = 1 - e^{-(R^2)/2s^2} \qquad \textbf{EQ.2.4}$$

where P is the probability that a point chosen at random will lie within a given distance R from the origin of the polar coordinate system. Or consider the game of darts. Suppose that an unbiased dart is thrown at the center of a dart board. EQ. 2.4 gives the probability that the dart will strike within R units of the center. By virtue of the manner in which all probabilities are defined, P may be zero or unity or any number in between. Equation 2.4 incorporates the coordinate system and probability distribution function to which the measure of dispersion known as the circular error probable is applied. The circular error probable is discussed in the next chapter.

NOTES AND REFERENCES FOR CHAPTER 2

The normal or Gaussian probability distribution is described in most introductory statistics and probability texts. Or see

Burington, R. S. and May, D. C., Jr., “Handbook of Probability and Statistics with Tables,” Handbook Publishers, Inc., Sandusky, OH, 1958.

The application of the normal distribution to dispersion of rounds is well established. See:

Crow, E. L., Davis, F. A., Maxfield, M. W., “Statistics Manual,” Dover, NY, 1960.

Grubbs, F. E., “Statistical Measures of Accuracy for Riflemen and Missile Engineers,” 1964.

Gnedenko, B. V. and Kinchin, A. Ya., “An Elementary Introduction to the Theory of Probability,” Dover, NY, 1962.

Herrman, E. E., “Exterior Ballistics,” U.S Naval Institute, Annapolis, MD, 1935.

For application of the Rayleigh probability density to reverberation, see:

Urick, R. J., “Principles of Underwater Sound,” 3rd ed., McGraw, 1983.

CHAPTER 3

THE CIRCULAR ERROR PROBABLE

Introduction

In the analysis of ordnance systems, a commonly-used measure of dispersion of the fall of bombs or the strike of projectiles is the "circular error probable." This measure, often abbreviated "CEP," is defined as the radius of the circle centered on the mean point of impact which would contain half of all impact points, given a large number of impacts. This chapter presents a graph of the relationship between CEP and the underlying probability, and also presents tables of the CEP versus probability, and of the inverse relationship.

A note on the terms used

The CEP is applicable to description of the dispersion of a wide variety of munitions. It has been applied to gun projectiles, bombs, rockets, guided missiles, and guided or homing bombs as well as small arms bullets. To avoid the confusion of constantly shifting terms, the writer will use the term "round" to refer to any one of the above named munitions. The word "round" is well established, having come into naval parlance with the introduction of cannon on board ship. Round shot referred to the spherical ball fired in early gun designs.

In some earlier work, the term "circular probable error" (CPE) is used. CPE has the same definition and meaning as CEP, and may be used interchangeably. For clarity, only the term CEP is used here.

The last chapter developed the circular normal probability distribution in polar coordinates, leading to EQ.2.4, repeated here:

$$P = 1 - e^{-(R^2)/2s^2} \qquad \textbf{EQ.3.1}$$

The origin of the coordinate system is understood to be at the center of impact of the rounds. The radius vector R may take on any value greater than or equal to zero. Let some value of the radius vector R, say R_0, be chosen. Then the probability P that a point chosen at random on the plane will lie at a distance equal to or less than R_0 is given by EQ.3.1.

In the study of gunnery, EQ.3.1 allows the introduction of a measure of dispersion of the fall of shot, or rounds. For a particular gun, rocket, bombing or missile system, a ready measure of dispersion is given by the radius R which will enclose half of the rounds delivered under a given set of conditions. To find this value of R, set P equal to 0.5 in EQ.3.1 and solve for R. The result is:

$$CEP = s(2LN(2))^{1/2} = (1.1774)s \qquad \textbf{EQ.3.2}$$

LN( ) = "natural logarithm of ( )"

That particular value of R is called the 'circular error probable', abbreviated CEP. Substituting EQ.3.2 back into EQ.3.1, we have:

$$P = 1 - e\text{-}(((R/CEP)\hat{}2)*Ln(2)) \qquad \textbf{EQ.3.3}$$

By expressing the probability P in terms of the ratio of radius R to CEP, the equation is generalized to every size of pattern of rounds.

In what follows, the discussion is in terms of the ratio of the radius R to the CEP. By use of the ratio R/CEP, the graphs and tables are made general and apply to any system, no matter how large or small its CEP. This will be made clearer in the examples to be discussed.

FIGURE 3.1. VALUES OF PROBABILITY (P) FOR (R/CEP) EQUAL TO 1, 2, AND 3.

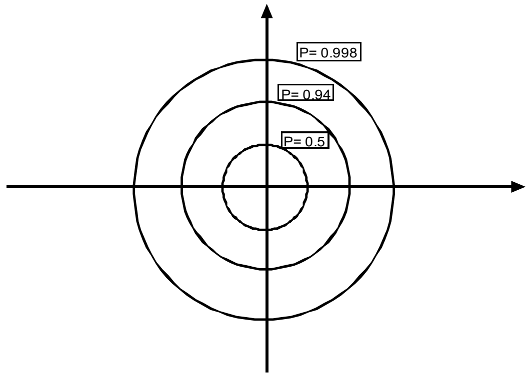


Figure 3.1 shows a contour plot, looking down upon the plane. The rounds must be imagined to fall upon this perfect plane, with the center of impact of the group, salvo or stick of bombs at the intersection of the axes shown. Half of the rounds should lie within the circle labeled P = 0.5, given a large number of rounds delivered under constant conditions. The radius to the P = 0.5 circle is the CEP for that particular ordnance system and conditions. For example, a small caliber machine gun fired at a range of two hundred yards might have a CEP of about eight to twelve inches or more, depending upon a great many factors. A battalion concentration of artillery fire would likely have a CEP of dozens of yards or meters, as would a salvo or pattern of rounds of naval gunfire. The CEP of unguided bombs or rockets will be even larger. The introduction of guidance and homing systems into bombs and missiles will greatly reduce the CEP, but can never completely eliminate it. Also in Figure 3.1, about 94 percent of the rounds will fall within a radius of twice the CEP. And finally, only about two rounds out of a thousand are likely to fall beyond three times the CEP from the center of impact of the pattern of rounds. These relations are summarized in Table 3.1. If committed to memory, Table 3.1 is helpful in discussions and evaluations of competing ordnance systems. For example, a rough estimate of the likely effectiveness of a particular weapon against a given target may be made by comparing twice the CEP to the dimensions of the target. Twice the CEP will enclose 15/16, or nearly all of the rounds. If twice the CEP is comparable to or less than the dimensions of the target, a properly-aimed strike should be effective.

FIGURE 3.2. PROBABILITY VS. RATIO OF RADIUS TO CIRCULAR ERROR PROBABLE (R/CEP)

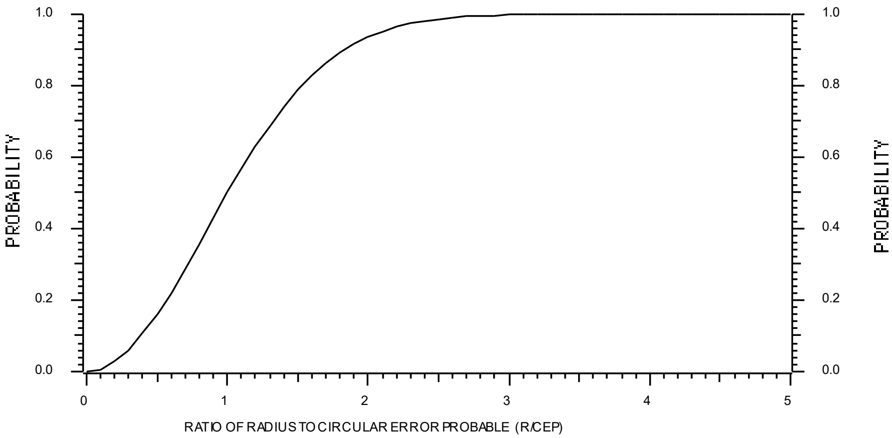


TABLE 3.2 RATIO OF RADIUS TO CIRCULAR ERROR PROBABLE (R/CEP) FOR GIVEN PROBABILITY (P)

P	R/CEP		P	R/CEP
0.02	0.170723009		0.52	1.029025602
0.04	0.24268022		0.54	1.058439528
0.06	0.298776402		0.56	1.088312718
0.08	0.346834591		0.58	1.118721935
0.1	0.389875741		0.6	1.149751319
0.12	0.42944682		0.62	1.181494256
0.14	0.466466971		0.64	1.214055678
0.16	0.501536406		0.66	1.247554948
0.18	0.535074		0.68	1.282129553
0.2	0.567387077		0.7	1.317939905
0.22	0.598710256		0.72	1.355175733
0.24	0.629228636		0.74	1.39406473
0.26	0.659092425		0.76	1.434884556
0.28	0.688426603		0.78	1.477979895
0.3	0.717337558		0.8	1.523787418
0.32	0.745917789		0.82	1.572873545
0.34	0.774249359		0.84	1.625993908
0.36	0.802406499		0.86	1.684191577
0.38	0.830457633		0.88	1.748969322
0.4	0.858467002		0.9	1.822615729
0.42	0.886496021		0.92	1.908888732
0.44	0.914604432		0.94	2.014669623
0.46	0.94285136		0.96	2.154960833
0.48	0.971296284		0.98	2.375680153
0.5	1		0.99	2.577567883

Figure 3.2 shows a graph of EQ.3.3. It may be seen that the probability P is 0.5 at R/CEP equal to 1.0. Figure 3.2 illustrates the slow rate of change of the probability at larger values of R/CEP.

FIGURE 3.3. PROBABILITY (P) VS. RATIO OF RADIUS TO CIRCULAR ERROR PROBABLE (R/CEP).

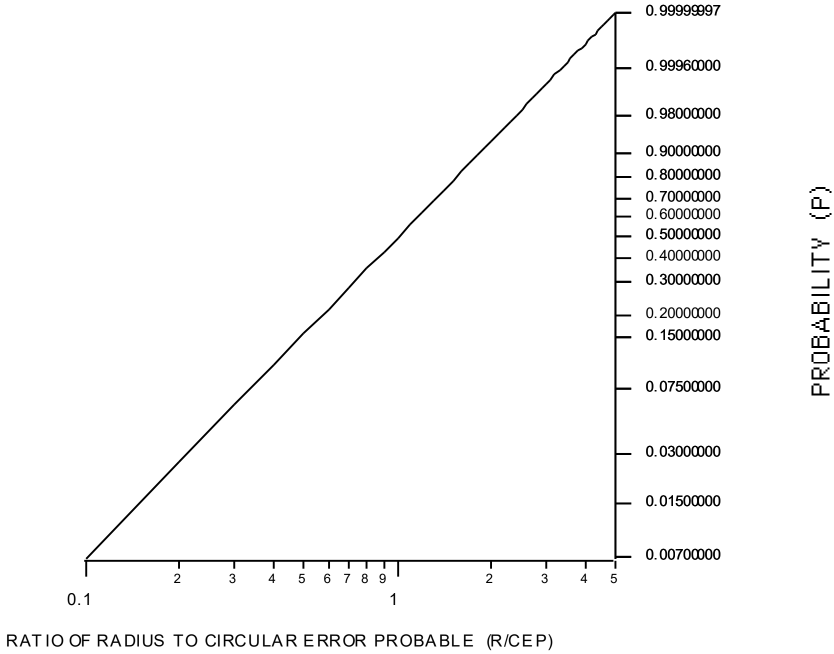


TABLE 3.3 PROBABILITY (P) FOR RATIO OF RADIUS TO CIRCULAR ERROR PROBABLE (R/CEP)

R/CEP	P		R/CEP	P
0	0		2.6	0.990773495
0.1	0.006907505		2.7	0.99361014
0.2	0.027345053		2.8	0.995635597
0.3	0.060477251		2.9	0.997060065
0.4	0.104974929		3	0.998046875
0.5	0.159103585		3.1	0.998720319
0.6	0.22083542		3.2	0.9991731
0.7	0.287974902		3.3	0.999473033
0.8	0.358287051		3.4	0.999668798
0.9	0.429618142		3.5	0.999794703
1	0.5		3.6	0.999874498
1.1	0.567731384		3.7	0.999924334
1.2	0.631432696		3.8	0.999955009
1.3	0.690073075		3.9	0.999973616
1.4	0.742971543		4	0.999984741
1.5	0.789775896		4.1	0.999991297
1.6	0.830424459		4.2	0.999995104
1.7	0.86509647		4.3	0.999997284
1.8	0.894156836		4.4	0.999998514
1.9	0.918100412		4.5	0.999999198
2	0.9375		4.6	0.999999573
2.1	0.952961039		4.7	0.999999776
2.2	0.965084777		4.8	0.999999884
2.3	0.974440561		4.9	0.999999941
2.4	0.98154699		5	0.99999997
2.5	0.986860994		5.1	0.999999985

In Figure 3.3, the probability is plotted as a straight line against nonlinear axes. This graph is useful for estimating probabilities. It also shows the slow change of probability for large R/CEP. Figure 3.3 is plotted from EQ.3.4:

$$LG[-LN(1-P)] = 2LG(R/CEP) + LG[LN(2)] \qquad \textbf{EQ.3.4}$$

where LG[ ] = logarithm to the base 10 or common logarithm
and LN( ) = natural logarithm.

TABLE 3.1 SHORT TABLE OF PROBABILITY FOR SELECTED VALUES OF (R/CEP).

R/CEP	PROBABILITY	FRACTION	COMMENT
1	0.5	1/2	EXACT
2	0.9375	15/16	EXACT
3	0.998+	> 998/1000	APPROX.

EQ.3.4 is of course the point-slope intercept form of the equation of a straight line:

$$y = mx + b$$

with $y = LG[-LN(1-P)]$, $m = 2$, $x = LG(R/CEP)$, and $b = LG[LN(2)]$.

EQ.3.4 is derived from EQ.3.3.

For the analysis of data from tests or in planning the tests themselves in a research, development or munitions surveillance program, more precise values of probability or R/CEP may be needed. Tables 3.2 and 3.3 provide information which may be of value. Table 3.2 gives R/CEP for selected values of probability P from 0.2 to 0.99. Table 3.3 treats the inverse problem, giving the values of probability P for selected values of R/CEP going from 0.05 to 5.1.

In the next chapter we consider applications of these tables.

NOTES AND REFERENCES FOR CHAPTER 3

For a derivation of CEP see pp. 99-101 of:

Burington, R. S. and May, D. C., Jr., “Handbook of Probability and Statistics with Tables,” Handbook Publishers, Inc., Sandusky, OH, 1958.

Another derivation of the CEP is found on p. 29 of:

Crow, E. L., Davis, F. A., and Maxfield, M. W., “Statistics Manual,” Dover, NY, 1960.

CHAPTER 4

APPLICATIONS OF TABLES OF THE CEP

Current Ordnance Development

As this is being written (*editor's note: approximately 1997*), among the more recent ordnance developments are the aircraft-delivered joint standoff weapon (JSOW) and the joint direct attack munition (JDAM). The JSOW is the more accurate of the two, with a CEP of 33 feet (nominal ten meters). The JSOW combines global positioning system (GPS) and inertial navigation system (INS) elements in an integrated guidance system. The JDAM is a nominal 2,000 pound bomb with a GPS - INS unit installed in the tail fin assembly. The JDAM achieves an accuracy of 15 to 20 meters CEP.

Based upon Mr. Canan's paper (see reference at end of chapter) , we may classify these weapons according to accuracy (probable terminal error) as follows:

"Accurate" munition: 15 to 20 meters CEP (example: JDAM)

"Very accurate" munition: nominal 10 meters CEP (example: JSOW)

"Precision" guided munition (PGM) : 3 meters CEP or less

The above definitions are somewhat different from those in technical usage. In general, a precision device has a small dispersion, or random error. An accurate device must have a small systematic, or aiming error, in addition to having a small random error.

The classifications as given above of weapons systems by their terminal accuracy or error is necessarily somewhat subjective and perhaps arguable. But the several classes seem satisfactory for practical operational use in the near future.

It seems clear that the circular error probable will be used widely in designing, developing, testing and operational planning for these weapons. The tables and graphs presented here should prove useful in all those areas.

The following discussion is based entirely upon artificially-generated or "school problems." However, the problems will serve to illustrate the use of the tables.

1. Suppose a Joint Standoff Weapon test is being planned. This weapon has a CEP of (nominal) 10 meters. A number of JSOW are to be launched against a test target.

What is the radius of the circle around the mean point of impact which should enclose 99 percent of the impacts?

In Table 3.2, at probability P of 0.99 the corresponding R/CEP is 2.58. So

R = CEP(2.58) = 10(2.58) = 25.8 meters.

Thus, 99 percent of the JSOW units should fall within 25.8 meters of the mean point of impact.

2. A Joint Direct Attack Munition (JDAM) has struck 32 meters from the mean point of impact of the group of JDAM launched against the target. The JDAM is considered an "accurate" munition with a CEP of 15 to 20 meters.

Is it likely that this particular JDAM is defective?

The R/CEP is 32/15 = 2.13 to 32/20 = 1.6. We use the case of 2.13 to enter Table 3.3. At R/CEP of 2.1, P equals .953. Hence there is a 95 percent chance that a properly-operating JDAM will fall as far as 32 meters from the mean point of impact. This performance is within the range of what we might reasonably expect. Hence, we cannot state that the munition is defective.

3. The JDAM strike considered in case 2 above is being reviewed. The production lot sample acceptance test data show that the lot actually achieved a CEP of 20 meters.

How does this additional information affect our conclusion?

As R/CEP = 32/20 =1.6, enter Table 3.3 at 1.6 to find P equal to 0.83. Thus there is a 17 percent chance that a JDAM from that particular production lot might fall as far as 32 meters (or more) from the mean point of impact. The 32-meter radius of this particular impact is not especially large, given the CEP of 20 meters from the lot test data.

Some Cautions

At this juncture, some warnings and cautions are in order. First, the calculations of radius made above are estimates of the impact radius of the munition and do not take into account the lethal radius of the munition's warhead. Hence the tables cannot be used alone in safety or effectiveness studies. Secondly, munitions are usually stored, often for a very long time before being employed. While in the stockpile, the performance of the munition does not improve. We may anticipate a gradual decrease in performance as munitions age. Thus, the dispersion of the munitions may increase. The analyst must be aware of this possibility and of the effect it may have on system effectiveness and personnel safety. Thirdly, the calculations assume that environmental influences do not change. This clearly is not true. The weather, for example, has much influence upon operations and is continuously changing. Fourth, it can occur that a munition is slightly damaged during the launching process. For example, an air-launched weapon may, on rare occasions, strike the fuselage or other part of the delivering aircraft. When this does occur, the stabilizing fins or control surfaces of the weapon may be slightly distorted. This bend or distortion may cause the munition to fly in an erratic manner. That particular munition cannot and will not obey the calculations and predictions made for it.

Another question concerns the shape of the pattern of fall of shot, or impact points of the rounds. In many practical cases, the pattern is roughly circular, and the development presented here is applicable. In other cases, such as the pattern of gunfire at long range against a surface target, the pattern tends to become elongated in range. This is especially serious for Marines calling in gun fire support, as the number of "short rounds" tends to increase as the range increases. Since gun fire support is often over and beyond the position of friendly troops, a short round is a serious matter.

A more obvious non-circular pattern is that of a long "stick" of bombs. In that case, the pattern is not even roughly circular, and other methods will have to be used to characterize the dispersion.

Tables 3.2 and 3.3 present more significant figures than are likely to be needed for the simple estimates made above. However, the tables were left in the form given because it is not possible to predict the many uses to which the tables may be put. It may be that the additional precision of the tables will find application in some other work.

NOTES AND REFERENCES FOR CHAPTER 4

The information on JSOW and JDAM is from the Navy League publication Sea Power:

Canan, J. W., "Smart and Smarter," Sea Power, Vol. 38, No. 4, Navy League of the United States, Arlington, VA, April 1995.

"JSOW Moves Forward," Sea Power, Navy League of the United States, Arlington, VA, April 1997.

For a method for analyzing severely non-circular patterns, see pp 112-115 of:

Burington, R. S. and May, D. C., Jr., "Handbook of Probability and Statistics with Tables," Handbook Publishers, Inc., Sandusky, OH, 1958.

CHAPTER 5

CALCULATING THE CEP FROM TEST DATA

Introduction

In previous chapters, calculations have been performed using given values of circular error probable (CEP). This chapter discusses the steps needed to obtain the desired value of CEP.

Test Data

For most weapons and especially explosive ordnance, the likely value of CEP to be obtained is a characteristic of first importance. Early in the research or development program, ballistics or flight and guidance tests should give a few preliminary samples of data. These early data will give some indication of the CEP to be expected. Later in the development or Low-Rate-Initial-Production (LRIP) program, larger sample sizes will be made available. It may be possible to pool some of the data to obtain a larger sample size, but this must be done with care, since engineering changes made throughout the development program may cause some units to perform in a manner unlike the majority of the population. Those units belong to a different population, and their data should not be included with production design units. The purpose of combining data is of course to increase the sample size and thereby increase confidence in the estimates made.

In early stages of programs, sample sizes are small, and care must be taken to wring as much information from the available data as possible. In what follows, we show a method by which the estimate of CEP may be made more accurate.

First consider the first three rounds of table 6.1, duplicated below:

SHOT NO.	HORIZ. (X)	VERT. (Y)
1	-1.72	2.84
2	-1.37	2.98
3	0.15	-0.02

Now suppose that the numbers represent measured impact points from a test of a Precision Guided Munition, with distances measured in meters. Let us determine the CEP. First we calculate the standard deviation of the horizontal, or 'x' values. We use this equation to calculate the standard deviation of the sample:

$$s_X = \text{STD. DEV.}(X) = [(S(X_i\text{-MEAN}(X))^2/(N-1)]^{1/2} \qquad \textbf{EQ.5.1}$$

The S or sum is taken over the number of data points, and in this case the index number i goes from 1 to 3. The result is:

$\text{STD. DEV.}(X) = 0.994 = s_X$

A similar calculation for the Y data gives:

$\text{STD. DEV.}(Y) = 1.670.= s_Y$

The usual method of calculating the standard deviation is slightly biased. The result is that the calculated standard deviation is somewhat smaller than the theoretically expected value. Frank Grubbs has shown how the bias may be corrected (see references at end of chapter). Table 5.1 gives the correction factors to be used for sample sizes of from 2 through 20. For our sample size of 3, the correction factor is 1.1284. After multiplying each standard deviation by 1.1284, we have the corrected value of:

$s_{X(\text{corrected})} = 1.122$

and

$s_{Y(corrected)} = 1.88$

In theory, the two standard deviations should be equal, but we may expect influences from many causes, and the small sample size allows considerable variation from test to test. Let us combine the two sample standard deviations s_X and s_Y to give an estimate of the common standard deviation for the calculation of the CEP. Since we can add the variances directly, let us simply average, or take the mean of the squares of the two standard deviations, and then take the square root of that mean:

$$[((s_X)^2 + (s_Y)^2)/2]^{1/2} = 1.55 \text{ meters} = s_{(estimated)}.$$

Now we may estimate the CEP from:

$$CEP = [(2 * LN(2))^{1/2}] * (s_{(estimated)}) = (1.1774) * (1.55) = 1.83 \text{ meters.}$$

So the estimate of CEP is CEP = 1.83 meters.

This estimate may now be used in making the calculations and estimates similar to those described in earlier chapters.

TABLE 5.1 CORRECTION FACTORS FOR STANDARD DEVIATIONS CALCULATED FROM SMALL SAMPLES.

SMPL SIZE (N)	STD.DEV. CORR.
2	1.2533
3	1.1284
4	1.0854
5	1.0638
6	1.0509
7	1.0424
8	1.0362
9	1.0317
10	1.0281
11	1.0253
12	1.023
13	1.021
14	1.0194
15	1.018
16	1.0168
17	1.0157
18	1.0148
19	1.014
20	1.0132

The corrections given in Table 5.1 differ from those given by Grubbs, because of a change that has occurred in the method used in calculating the standard deviation. Grubbs used the method common at that time of taking the square root of the average of the squared deviations from the mean. In computing the average, the sum is divided by the sample size, N. Today when computing sample standard deviations, we divide by (N-1) which is a simple, if slightly inaccurate, method of correcting for bias. As Table 5.1 shows, the correction needed is largest at sample size N of 2, where the bias is slightly more than 25 percent low. The magnitude of correction needed decreases rapidly as sample size increases, becoming less than five percent for sample sizes of seven or more.

Table 5.1 is computed by taking the reciprocal of equation (15) on p. 23 of Grubbs' book. This becomes:

$$\text{correction factor} = [(N/2)^{1/2}] * [G(N-1)/2] / [G(N/2)] \qquad \textbf{EQ.5.2}$$

where G is the gamma function.

NOTES AND REFERENCES FOR CHAPTER 5

The gamma function is discussed by Burington and also in most texts on special functions.

Burington, R. S. and May, D. C., Jr. , "Handbook of Probability and Statistics with Tables," Handbook Publishers, Inc., Sandusky, OH, 1958.

The correction for the standard deviation is on p. 23 of:

Grubbs, F. E., "Statistical Measures of Accuracy for Riflemen and Missile Engineers," 1964.

For a discussion of the gamma function, see chapter 2 of:

Rainville, E.D., "Special Functions," MacMillan, NY, 1960.

CHAPTER 6

THE MEDIAN AND ITS APPLICATION TO THE SPOTTING OF ROUNDS

Introduction

In studying a sample of data, one often looks for a single number to characterize the entire set or group of numbers. An "average" value of some sort is wanted. The mean is the most used of the possible averages which might be chosen. The mean has the great virtue of simplicity of calculation, being the sum of the data values divided by the total number of such values. In many probability distributions, including the normal, the mean describes the center of the distribution. But the mean is not the only measure of central tendency available. The median is an excellent indicator of center, in some respects, superior to the mean.

As used here, a 'spot' is a correction applied to weapon or battery orders to bring the rounds onto the target. 'Spotting' refers to the process of observing the fall of shot and estimating the necessary corrections. For ground forces artillery, 'adjustment of fire' means essentially the same as spotting. A similar procedure is used to adjust gun sights on small arms and other direct-fire weapons.

The Median

For a theoretical probability distribution, the median is defined to be that location such that a point chosen at random has equal probability of being greater than or less than the given median location. That is, a randomly-chosen point has a probability of 0.5 of being less than the median and a probability of 0.5 of being greater than the median.

The median of a sample is defined to be a value which is greater in size than half of the elements of the sample and less than the other half. By virtue of its definition, the median is independent of the parameters of any given probability distribution. Hence it is fair to consider the median a non-parametric measure of central tendency.

Now consider the following sample of five elements:

1.5
3.1
5.2
7.8
9.6

The median is 5.2. When dealing with small samples having an odd number of elements, one quickly picks the center element as the median (5.2 in the example above). For a sample having an even number of elements, it is customary to compute the median as the mean of the two centermost elements. For example, given the sample:

1.5
3.1
5.2
7.8
9.6
11.3

the median is taken to be 1/2(5.2+7.8) = 6.5. The median computed in this manner is truly correct only in the special case in which the two centermost elements have the same value.

A practical rule for calculating the median is as follows: sort the measured values in order of size. Then number each value according to its position, one for the first, two for the second, and so on. In general, the median point is $(1/2)*(n+1)$, where n is the sample size, or total number of elements in the sample.

TABLE 6.1 SMALL ARM PROJECTILE IMPACTS ON VERTICAL TARGET.

SHOT NO.	HORIZ. (X)	VERT. (Y)
1	-1.72	2.84
2	-1.37	2.98
3	0.15	-0.02
4	0.37	2.68
5	0.55	5.85
6	0.91	4.76
7	1.48	1.45
8	1.98	3.04
9	2.22	2.81
10	2.53	1.98

NOTES:

(1) DEVIATIONS DEFINED AS POSITIVE UPWARD AND TO THE RIGHT.

(2) SHOTS NUMBERED SEQUENTIALLY BY HORIZONTAL POSITION, LEFT TO RIGHT.

(3) DEVIATIONS MEASURED IN INCHES.

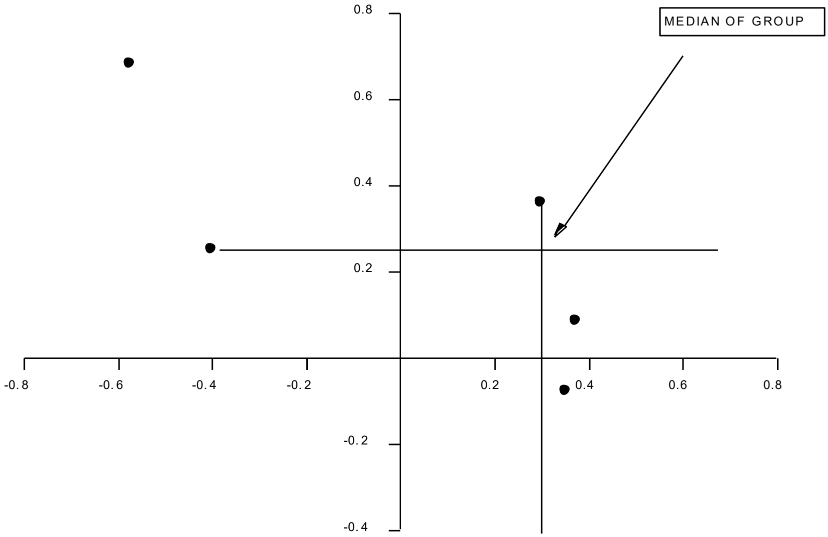
In the 10-shot sample of Table 6.1, the horizontal (X) data are sorted in order of size. The median point lies at (1/2)*(10+1) = 5.5. Of course, no sample point lies at that position, so the value of the (horizontal) mid-point between shots 5 and 6 is computed as follows: (1/2)*(0.55+0.91) = 0.73. A sample having an odd number of elements is easier to deal with. Consider a nine-shot salvo of naval gunfire. The median shot is (1/2)*(9+1) = 5. The fifth shot, counting along the range direction from either end of the pattern, is the median in range, and it may be used as an estimate of the center of impact. This estimate is particularly important, as it may be used in the adjustment of fire; that is, the adjustment of the laying of the battery to give the maximum number of hits upon the target.

Perhaps the most important quality of the median is that it is not disturbed by changes in the variance or in the extreme values of the sample. An increase or decrease in the pattern size will not disturb the median. Also, an extreme round that falls somewhat closer to or farther from the center of impact will not change the median. By way of contrast, the mean point of impact is affected by any change in the position of impact of any round.

Figure 6.1 illustrates the use of the median in the adjustment of a gunsight. The discussion here is in terms of small arms. Suppose that a rifle shooter has fired a group of five shots upon a vertical target at a fixed range. The shots are distributed upon the target as in Figure 6.1. The aim-point is at the crossing of the numerically labeled axes. We find the median of the group as follows:

The ordinal number of the median shot is (N+1)/2 = (5+1)/2 = 3. Thus we count from the top shot down (or bottom shot up) to the third bullet hole and strike a horizontal line across the target, as shown. Similarly, after counting from the leftmost shot rightward to the third shot, we strike a vertical line down to intersect the horizontal line which was struck previously. This intersection is the median of the group, as labeled. It may happen that a particular round is the median in both the horizontal direction and in the vertical direction. In that case, the position of the median of the group is at the location of that particular round.

FIGURE 6.1. USE OF MEDIAN ROUND POSITION
FOR ADJUSTMENT OF SIGHTS.



The adjustment of the rifle sights is done as follows:

The rifle is placed in a simple holding fixture and moved or shimmed until it is aimed exactly at the aim point (the intersection of the numbered axes in Figure 6.1). The rifle is securely clamped in place, being careful that the aim is not disturbed. Then the sights are adjusted so as to aim directly at the median of the group as labeled in the figure. Now the sights have been adjusted to point directly at the median position of the group, and the rifle is said to be "sighted in." All that remains is to fire a group of shots at the same range to verify the sight setting.

The procedure is somewhat simpler if the rifle has an optical or telescopic sight having a crosshair reticle. In that case, the shooter may aim at the aim point used to fire the group, and clamp the rifle in place. Then the horizontal cross hair is adjusted to cut, as nearly as possible, the third bullet hole counting from the top or bottom. A similar procedure is performed with the vertical cross hair. Then the reticle is secured and a check target is fired to verify that the sight setting is correct.

Essentially the same procedure will work for automatic cannon or machine guns on armored vehicles or aircraft, as well as tank cannon or other direct-fire weapons, including those using laser, infrared or low-light-level sighting systems.

NOTES AND REFERENCES FOR CHAPTER 6

The paper by Campbell contains many printing errors. The reader should review the Naval Engineers Journal, July 1983, pp. 153-156 for corrections.

Campbell, L. M., "Applications of Hybrid Statistics and the Median," Naval Engineers Journal, Washington, D.C., January, 1983.

The practical use of the median for the spotting of rounds was pointed out by Herrman:

Herrman, E. E., "Exterior Ballistics," U.S Naval Institute, Annapolis, MD, 1935.

For a more general approach to the calculation of the sample median, see:

Jackson, Dunham, "Note on the Median of a Set of Numbers," Bulletin of the American Mathematical Society, Vol. 27, pp. 160-164, 1921.

A summary of Dunham Jackson's approach is found in:

Whittaker, E., and Robinson, G., "The Calculus of Observations," Dover, pp 197-199, NY,1967.

The writer has used the median for the adjustment of sights on small arms. The procedure is effective and efficient.

CHAPTER 7

MEDIAN VERSUS MEAN IN SMALL SAMPLES FROM A NORMAL DISTRIBUTION

Introduction

Practical applications of statistics often involve gathering a sample of data and “reducing” the sample to a few numbers. Typically the average or mean is computed in order to have a measure of the center of the sample. The population from which the sample is drawn is sometimes called the parent population. Occasionally, the researcher is sampling from a physical process with a known probability distribution. For example, in the case of gunfire at a vertical target at moderate range, the rounds usually will be normally distributed in both the vertical and in the horizontal directions. In addition, the sample mean is also normally distributed with sample mean equal to the parent population mean and a standard deviation of:

$$\text{standard deviation of sample mean} = s_{mean} = s/(N)^{1/2}$$

where s is the standard deviation of the parent population and N is the sample size, that is, the number of measured values available for study.

Theory shows that the mean of the sample median is also equal to the mean of the parent population. Thus we expect the sample median to provide an unbiased estimate of the center of the sample.

The theory also indicates that the dispersion or variance of the median is greater than that of the mean. If we take the standard deviation (the square root of the variance) as our measure of dispersion, then the ratio of the standard deviation of the median to the standard deviation of the mean is generally greater than unity. Indeed, as the sample size N becomes larger and larger, the ratio

$$[s_{median}/s_{mean}] \text{ approaches } (p/2)^{1/2} = 1.2533...$$

Thus, at worst case in sampling from a normally distributed variate, the sample median might have a standard deviation approximately 25 percent greater than the standard deviation of the mean. For much research work, the question is moot, since the researcher often (if not usually) begins without knowledge of the statistical distribution being dealt with and is fortunate to have any well-behaved measure of central tendency.

With small samples, the ratio of standard deviations is not so great. This ratio (of standard deviations of sample median to sample mean) for small samples from a normal population) was studied by Tokishige Hojo.

TABLE 7.1 RATIO OF STANDARD DEVIATION OF MEDIAN TO STANDARD DEVIATION OF MEAN, COMPARED TO THEORETICAL VALUES, SAMPLES FROM NORMAL DISTRIBUTION.

N	THEORY	NBS RATIO	NBS % ERROR	RAN RATIO	RAN % ERROR
2	1	1	0	1	0
3	1.1602	1.14817	-1.037	1.17021	0.863
4	1.0922	1.08555	-0.609	1.09549	0.301
5	1.1976	1.19318	-0.369	1.20289	0.442
6	1.1351	1.13468	-0.037	1.14681	1.032
7	1.2137	1.21161	-0.172	1.22587	1.003
8	1.16	1.1696	0.828	1.16387	0.334
9	1.2226	1.22123	-0.112	1.22799	0.441
10	1.1768	1.17508	-0.146	1.17488	-0.163
11	1.2286	1.22356	-0.41	1.24058	0.975
12	1.1898	1.19115	0.113	1.18779	-0.169

The values of the ratio of standard deviation of sample median to standard deviation of sample mean as calculated by Hojo for sample size N are given in the second column of Table 7.1, headed “THEORY.” These are the theoretical values which are used as reference in subsequent calculations. The remaining columns give calculated values of the ratio using both the NBS and the RANNUM routines for the generation of Gaussian-distributed psuedo-random numbers. (See end of chapter for references.) The percent errors were computed as follows:

percent error ={[(CALCULATED VALUE) - (THEORY)]/(THEORY)}*(100).

An inspection of the table shows that the ratio of standard deviations (s_{median}/s_{mean}) is always greater for the odd sample sizes when compared to the adjacent even sample sizes. Figure 7.1 graphs the first two columns of Table 7.1. The greater magnitude of the ratio for odd sample sizes is apparent from the graph. By convention, the even sample size median is computed by averaging the two centermost values. This tends to “smooth” the median for the even sample sizes, thus reducing the variance and standard deviation as compared to the odd sample sizes, in which the center value is picked for the median, without benefit of smoothing. To take account of this difference, Hojo generated two separate equations, one for the odd-sample ratio and the other for the even-sample ratio. As stated above, the conventional method for computing the median of an even sample size is simply to compute the mean of the two centermost values. Hence, for a sample size of two, the computations of median and mean are the same, the ratio is unity and the percent errors are zero, as shown in the table. Figure 7.2 shows the percent errors for two different Gaussian random number generators used to compute the ratio of standard deviations. The theoretical values computed by Hojo are used as reference. As the graph shows, the errors are small, barely exceeding one percent.

Computation of the Table

Computations leading to Table 7.1 were as follows: for each sample size N from 2 through 12, 10,000 samples were generated using the NBS algorithm with mean of zero and standard deviation of unity. The ratios of standard deviation of median to standard deviation of mean were computed for each sample size.

The ratios were compared to Hojo's theoretical values and the percent errors computed. This process was repeated using the RANNUM routine to generate the Gaussian-distributed random numbers. The average error for the NBS samples is minus 0.177 percent, while that for the RANNUM-generated samples is +0.475 percent. These errors are quite small, and provide verification of the accuracy of Hojo's work, which was calculated directly from mathematical principles.

Practical Considerations

The above discussion means that the observer spotting gunfire or other ordnance delivery in the field may use the median as a quick estimate of center of impact of the rounds, with little loss of accuracy.

FIGURE 7.1. RATIO OF STD. DEV. OF MEDIAN TO STD. DEV. OF MEAN,
FOR SMALL SAMPLES FROM NORMAL DISTRIBUTIONS.

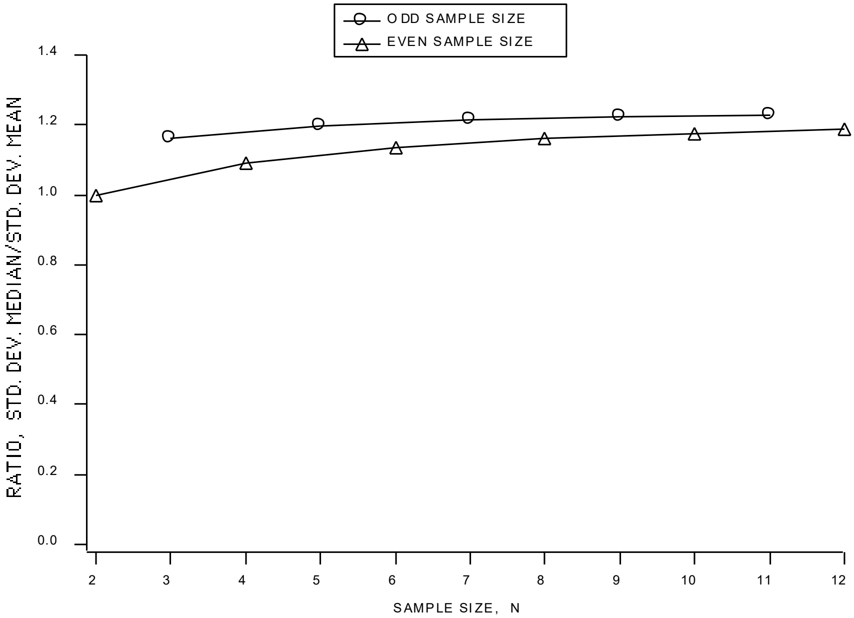
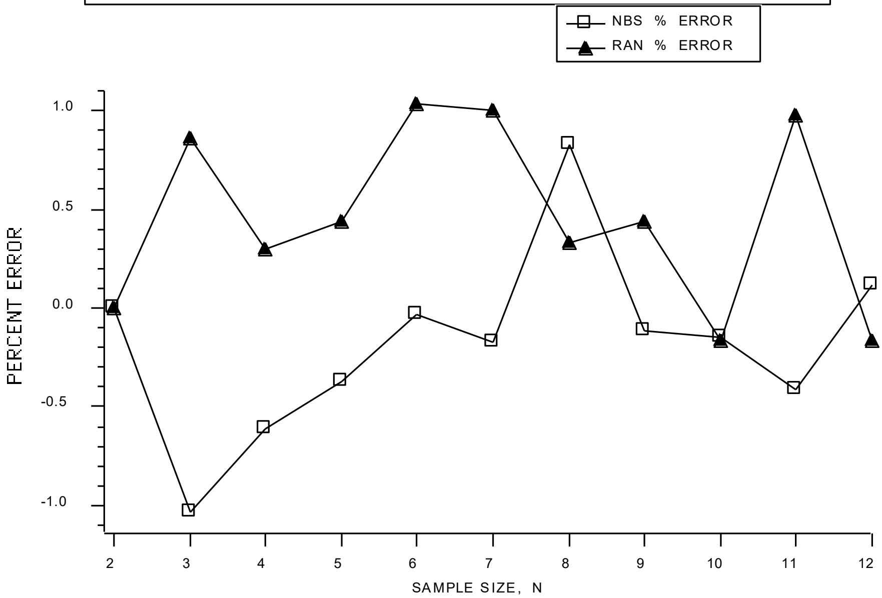


FIGURE 7.2. PERCENT ERROR VS. SAMPLE SIZE FOR RATIO OF STD. DEV. OF MEDIAN TO STD. DEV. OF MEAN, COMPARED TO THEORETICAL VALUES, SAMPLES FROM NORMAL DISTRIBUTIONS.



NOTES AND REFERENCES FOR CHAPTER 7

The NBS algorithm for generating Gaussian-distributed random numbers is from pp. 952-953 of:

Abramowitz, M., and Stegun, I. A., "Handbook of Mathematical Functions," National Bureau of Standards, 1964.

The mean and standard deviation of sample mean is given on p. 345 of:

Cramér, H., "Mathematical Methods of Statistics," Princeton University Press, 1963.

Hojo, T., "Distribution of the Median, Quartiles and Interquartile Distance in Samples from a Normal Population," Biometrika, Vol. 23, pp. 315-363, (1931).

The difference between the median of odd and even samples was emphasized by F. L. Weaver:

Weaver, F. L., Naval Engineers Journal, pp. 153-156, July 1983.

The RANNUM routine for generating Gaussian-random numbers was resident on the computer system at the Naval Ordnance Laboratory, White Oak, MD. That laboratory is now closed, and the fate of the software is unknown to the writer.

CHAPTER 8

NONPARAMETRIC STATISTICS AND APPLICATIONS

Introduction

Classical statistics considers various probability distributions and their parameters such as the mean, variance, and higher-order moments. In contrast, nonparametric statistics ignores those parameters, but yields a remarkable amount of information nevertheless.

Quality of Manufacture

S.S. Wilks considered the application of mathematical statistics to the practical problem of controlling quality of a manufactured product. For example, a given quality characteristic might be measured by a variable "X" where "X" might be the "blowing time" in seconds for a particular type of electrical fuse or the "breaking strength" of a sample of parachute cord. In testing the breaking strength of a sample of cord, for example, the results will reveal a range of values of breaking strengths. As a rule, high values of breaking strength are acceptable, and attention is given only to the minimum values. Wilks refers to this situation as the problem of "one tolerance limit." In particular, an answer is desired to the question: "With what confidence 'C' may it be predicted that a given fraction 'R_c' of the population of breaking strengths will be greater than the measured minimum value 'X', from a sample of size 'n'?"

Wilks shows that the confidence "C" and the population fraction "R_c" are related to the sample size "n" by the definite integral equation:

$$1n \; \_ \; x^{(n-1)}dx = C$$ **EQ.8.1**

R_c

(*editor's note: this equation was unfortunately corrupted*) where x is a “dummy” variable of integration.

This equation integrates simply to:

$$1 - (R_c)^n = C \qquad \textbf{EQ.8.2}$$

The relationship described by EQ. 8.2 is graphed in Figure 8.1. The figure shows that increases in sample size beyond about 300 does not return much increase in confidence.

By transposing terms, taking logarithms to the base 10 and multiplying through by minus one, we have:

$$-LG(1-C) = -LG(R_c)[n] \qquad \textbf{EQ.8.3}$$

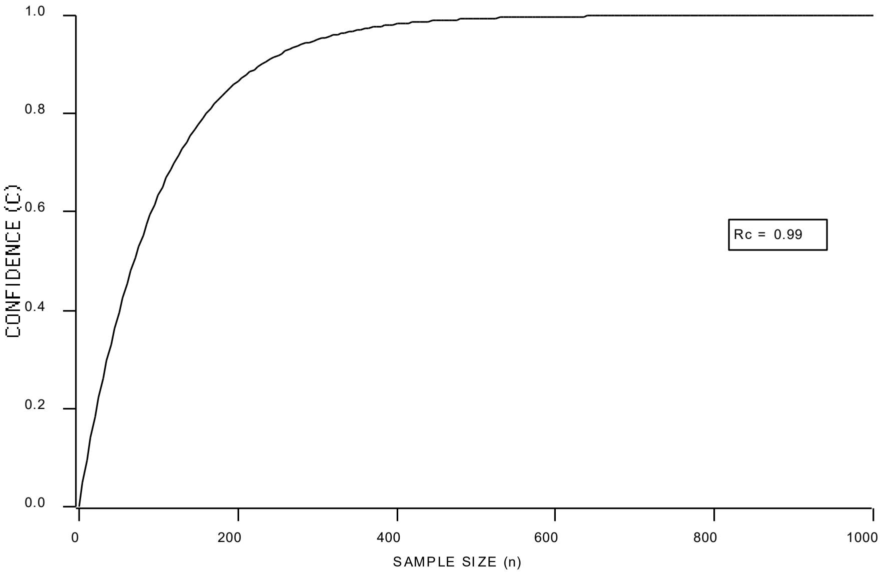
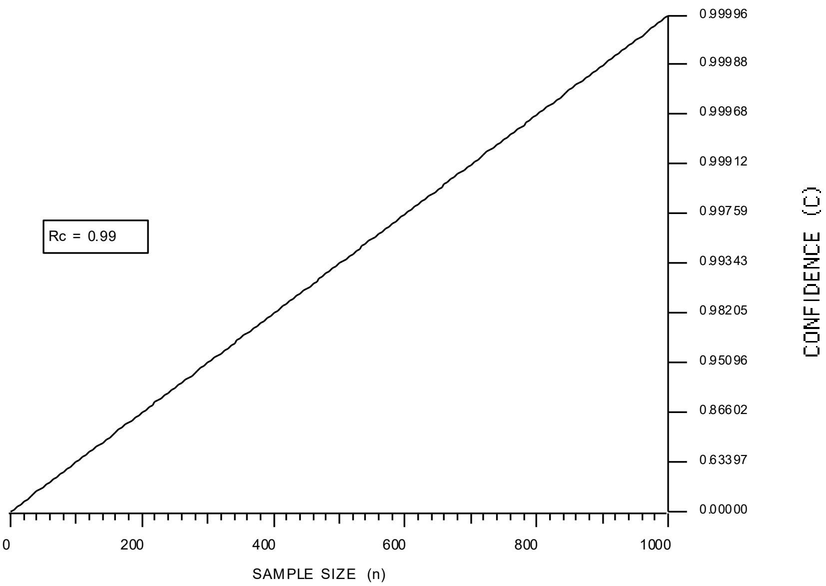


FIGURE 8.1. CONFIDENCE (C) VS. SAMPLE SIZE (n).

FIGURE 8.2. CONFIDENCE (C) VS. SAMPLE SIZE (n)



The multiplication by minus one is merely a convenience to allow the graph of EQ.8.3 to plot in the first quadrant. This linear equation is shown graphed in Figure 8.2, for a fixed fraction "R_c" equal to 0.99. Again, the larger sample sizes do not give much increase in confidence.

This is as good a place as any for a consideration of the choice of a particular numerical value for "R_c." If a relatively small value is chosen for R_c, then a substantial part of the population may have values which lie outside the range of the sample. Thus the sample range cannot represent, even approximately, the population. On the other hand, a fraction R_c which is too near unity will need an inordinately large sample size to give a confidence of 0.9 or greater, which is desired. In the work here, a fraction R_c equal to 0.99 is chosen as a compromise to give a reasonable confidence with sample sizes which are not too high.

Table 8.1 gives the sample sizes needed to yield confidence levels of 0.6, 0.8, 0.9, 0.95, and 0.99. The values of "n" have been rounded up to yield more convenient sample sizes. An example in the use of the table follows.

Suppose that a sample from a production lot of parachute cord has been received from the manufacturer and must be tested to verify that the breaking strength is high enough. The minimum breaking strength is determined by the design load of the parachute. The test is performed by choosing a sample of "n" cords and subjecting each cord to a gradually increasing force or load, until the cord breaks. The test records provide the sample distribution of ultimate strengths. In this case, we want to know if 99 percent (R_c =0.99) of all cords which might be produced by this manufacturer will have breaking strengths of at least the minimum obtained in the test, with a confidence of 95 percent ($C = 0.95$). Referring to Table 8.1, at C equal to 0.95, n is 300. Hence a sample size of 300 cords must be tested. The above application rests upon the assumption that the manufacturing process remains "in control" and that the remaining variability of the product's breaking strength may be considered to be random.

Returning to Table 8.1, It is seen that the sample sizes required are fairly large.

TABLE 8.1 SAMPLE SIZE n NEEDED TO GIVE CONFIDENCE C THAT A FRACTION R_c OF THE POPULATION WILL EXCEED A GIVEN REQUIREMENT

C	n
0.6	92
0.8	161
0.9	230
0.95	300
0.99	460

This is the price that is paid for the use of nonparametric statistics. At the outset, a cost analysis or estimate must be performed. If measurements must be repeated on many lots or batches of product, then one should study the product so as to fit a particular probability distribution to the variability of the data. In this way, a more efficient test may be devised which will have a smaller sample size and thus lower cost. On the other hand, if only a small number of lots are to be tested or if an answer is needed immediately and time is not available for theoretical studies, the use of nonparametric statistics may yield the needed information with minimum cost and delay.

NOTES AND REFERENCES FOR CHAPTER 8

The test described by Campbell on p. 71 of the reference below is not correct. The load must be gradually applied to the ultimate strength of the material, in order to discover the minimum breaking strength in the sample.

Campbell, L. M., "Some Applications of Extreme-Value and Non-Parametric Statistics to Naval Engineering," Naval Engineers Journal, Vol. 89, No. 6, pp. 67-74, Washington D. C., December, 1977.

Wilks, S. S., "Determination of Sample Sizes for Setting Tolerance Limits," Annals Math. Stat., Vol. 12, pp. 91-96, (1941).

Wilks, S. S., "Statistical Prediction with Special Reference to the Problem of Tolerance Limits," Annals Math. Stat., Vol. 13, pp. 400-409, (1942).

It has been said before, but bears repeating: any work by S. S. Wilks is worthy of study.

CHAPTER 9

NONPARAMETRIC STATISTICS CONTINUED
AN APPLICATION TO DISPERSION IN EXTERIOR BALLISTICS

Introduction

In development and manufacture of small arms ammunition, the dispersion of rounds when fired at fixed range upon a vertical target is a good indication of quality of the ammunition. Many different measures of dispersion have been proposed and used. The most popular measure among rifle shooters is the "extreme spread." This is the maximum distance between all possible pairs of bullet holes on the target.

FIGURE 9.1. BULLET IMPACTS ON VERTICAL TARGET.

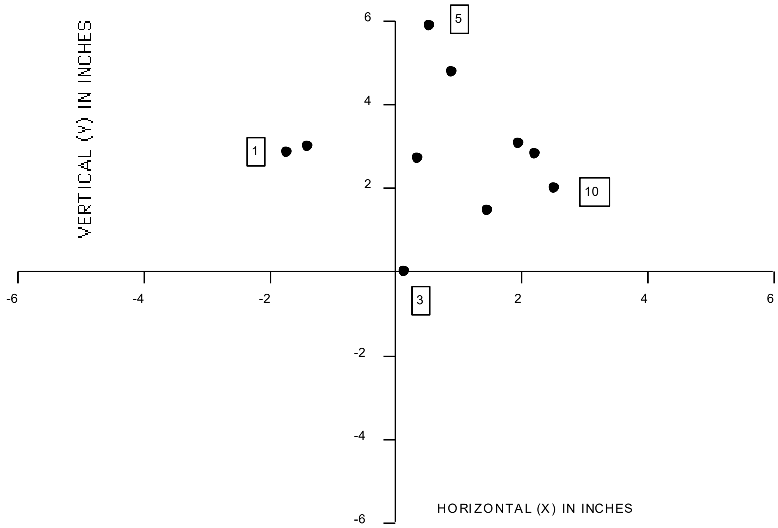


Figure 9.1 shows how a typical target from a small arms range might appear. This particular target was made by a shooter firing a .45 caliber revolver at 25 yards range. Table 6.1 gives the measured data from which the figure is plotted.

In Figure 9.1, the shots are numbered from left to right across the target. This is not the sequence in which the shots were fired, but were numbered to aid in the discussion. It is often possible to identify, simply by inspection, the pair of bullet holes which have the largest distance between, or extreme spread. In some cases, it is necessary to measure several candidates to identify the largest. In Figure 9.1, the spread between shots 3 and 5 appears to be the largest. On the actual target, one could simply measure the distance with a metal scale or tape measure. In a large shooting match, the measurement process might be automated with on-target sensors to report the coordinates of

each round striking the target. In our case, we have the measured horizontal and vertical deviations in Table 6.1. Subtracting the x- and y-coordinates of round 3 from the corresponding coordinates of round 5, we may compute:

EXTREME SPREAD = $[(X_5-X_3)^2 + (Y_5-Y_3)^2]^{1/2}$,or about 5.88 inches.

As noted by Grubbs, a group of "N" projectile hits upon a target generates "n" possible spreads, each spread being the linear distance between each possible pair of points. That is, "N" points are chosen two at a time. This is known formally as the combinations of "N" things, two at a time, and can be expressed as follows:

$$COMB(N;2) = (N!)/(2!)(N-2)! = n = N(N-1)/2 \qquad \textbf{EQ.9.1}$$

where $N! = (N)*(N-1)*(N-2)*...*(2)*(1)$.

Substituting for "n" in EQ.7.2 yields:

$$C = 1- (R_c)^{N*(N-1)/2} \qquad \textbf{EQ.9.2}$$

Using EQ.9.2, Table 9.1 is computed, which gives the confidence "C" that the extreme spread measured across a group of "N" projectile hits will cover 99 percent ($R_c = 0.99$) of all possible groups fired under similar conditions. Notice that the confidence for small sample sizes is quite low, being only 36 percent for the commonly-used 10-shot group, but reaching 99 percent for a group of 31 shots.

TABLE 9.1 CONFIDENCE C THAT EXTREME SPREAD OF GROUP OF N ROUNDS WILL COVER 99 PERCENT ($R_c = 0.99$) OF ALL GROUPS OF N ROUNDS FIRED UNDER SIMILAR CONDITIONS.

N	C		N	C
1	0		21	0.87883118
2	0.01		22	0.90188623
3	0.029701		23	0.921349
4	0.05851985		24	0.93758145
5	0.09561792		25	0.95095911
6	0.13994165		26	0.96185495
7	0.19027213		27	0.97062666
8	0.24528071		28	0.97760745
9	0.30358678		29	0.98309991
10	0.36381451		30	0.98737272
11	0.42464525		31	0.9906596
12	0.48486288		32	0.99315999
13	0.54339025		33	0.99504113
14	0.59931535		34	0.99644087
15	0.65190689		35	0.99747105
16	0.70061961		36	0.99822101
17	0.74509024		37	0.99876109
18	0.78512555		38	0.99914583
19	0.82068432		39	0.99941699
20	0.85185501		40	0.99960604

Some Examples in the Use of Table 9.1

Although these examples are taken from small-arms ordnance, the techniques are believed to be applicable to any projectile-throwing or missile-delivering ordnance system.

Example 1

We wish to measure, with a confidence of 0.90, the ballistic dispersion of a particular rifle firing a certain production lot of ammunition. Referring to Table 9.1, a confidence of 0.90 corresponds most closely to a sample size of 22 shots. Hence, the test procedure is as follows:

A group of 22 shots is fired upon a target and the resulting extreme spread is measured. Then 90 percent ($C = 0.90$) of all groups fired under similar conditions should measure less than or equal to the measured value of extreme spread.

Example 2

The accuracy of a particular production lot of ammunition must be determined, with a high confidence (say 99 percent) of being correct. From Table 9.1, a confidence of 0.99 requires a sample size or group of 31 shots. After having fired the 31-shot group and having measured the extreme spread, we may expect that 99 percent of all groups ($C = 0.99$) fired under similar conditions will have extreme spreads of less than or equal to the measured value.

Example 3

The dispersion of a particular lot of ammunition must be determined, using existing records. The records reveal that a 15-shot group was fired, yielding an extreme spread of 4 inches at a range of 100 yards. From Table 9.1, at $N = 15$, the confidence is 0.65. Thus we have our answer of 4 inches for the extreme spread, but at the relatively low confidence level of 65 percent. Our confidence in the measurement is limited by the relatively small 15-shot sample size.

The reader will notice that these last three examples are concerned with the upper tolerance limit (the maximum dispersion), whereas the example of the previous chapter dealt with the lower tolerance limit (minimum breaking strength). The extension of the method is based upon Wilks' statement that the problem of an upper tolerance limit is entirely similar to that of a lower tolerance limit.

A comparison of the confidence levels predicted by Table 9.1 with actual practice is illuminating. George L. Jacobsen, former Assistant Superintendent of Frankford Arsenal, states that for accuracy testing of small arms ammunition, a sample of at least 30 rounds is needed. This corresponds to a confidence of 0.987, which is not unreasonable for accuracy research. A second indication of the essential correctness of Table 9.1 is provided by the experience of workers at the U.S. Army Ballistics Research Laboratory during World War Two. In tests of the dispersion of 0.50 caliber aircraft machine guns, many groups of 20 shots each were fired. Table 9.1 indicates a confidence of 0.85 that the extreme spread of each group fired would cover 99 percent ($R_c = 0.99$) of the corresponding population of spreads. The 20-shot group is a reasonable compromise between confidence needed and time available for the test.

NOTES AND REFERENCES FOR CHAPTER 9

Campbell, L. M., "Some Applications of Extreme-Value and Non-Parametric Statistics to Naval Engineering," Naval Engineers Journal, Vol. 89, No. 6, pp. 67-74, Washington D. C., December, 1977.

Grubbs, F. E., "Statistical Measures of Accuracy for Riflemen and Missile Engineers," 1964.

Jacobsen, G. L., "Factors in Accuracy," The NRA Handloader's Guide, p. 145, National Rifle Association of America, Washington, D. C., 1969.

"Ballisticians in War and Peace. A History of the United States Army Ballistics Research Laboratories 1914-1956," Vol. 1, p. 55, United States Army Ballistics Research Laboratories, Aberdeen Proving Ground, MD, November, 1975.

Wilks, S. S., "Determination of Sample Sizes for Setting Tolerance Limits," Annals Math. Stat., Vol. 12, pp. 91-96, 1941.

Wilks, S. S., "Statistical Prediction with Special Reference to the Problem of Tolerance Limits," Annals Math. Stat., Vol. 13, pp. 400-409, (1942).

CHAPTER 10

STURGES' RULE FOR DATA PLOTTING.

Introduction

In the study of observed and measured data, one often assigns a number to a particular physical quantity. One of the graphical tools developed for presenting and clarifying masses of numerical data is the histogram. Most elementary texts on statistics introduce the histogram and give procedures for plotting. We may think of the process as sorting objects by size into various "bins." In the case of statistical analysis, usually we are sorting numerical values rather than physical objects. Upper and lower numerical limits are determined for each bin, and the data values which lie between the limits are placed in the bin. The result is a tally or count of the total number of data values in each bin, and is plotted as tally versus bin location.

A note on the terms used

The discussion here uses the term "bin" to describe the numerical interval into which the data values are sorted. The term "bin" seems to have been introduced with applications of statistics to signal processing and especially to the detection of very small (low-power) signals. Earlier writers used the word "cell" to mean the same thing as bin. And still earlier in this century, writers referred to the "class interval" when constructing a histogram. As observed by E. Bright Wilson (see references at end of chapter), a simple numerical measurement is in a sense a sorting of data into classes. Thus the earlier writers were on good ground when using the term "class interval." In this work, the writer shall conform to current word usage of bin.

Example

The graph of Figure 12.2 includes a plot of the mean horizontal positions of 32 three-shot groups. A useful method for gaining insight into the phenomenon under study is to plot the data as a histogram. This is done in Figure 10.1, using six bins. Inspection of Figure 10.1 reveals that the data is not distributed symmetrically, as one might expect in sampling from a normal distribution. The "X-mean" values are mostly negative, 21 as against 11 positive values. But we might equally as well have chosen three bins, as in Figure 10.2. Now here, the negative bias is still apparent, but the finer structure of the distribution of the sample values has been lost. For example, the single value lying at the top of the range in Figure 10.1 is not visible in the histogram of Figure 10.2 with only three bins. A plot using 12 bins, as in Figure 10.3, shows more detail but has two empty bins which is somewhat bothersome, indicating that the data are perhaps being "stretched" too far.

Very well, six bins seems about right, but how do we know what number of bins to choose? A quick way is to refer to Table 10.1, which gives a recommended number of bins for a given sample size.

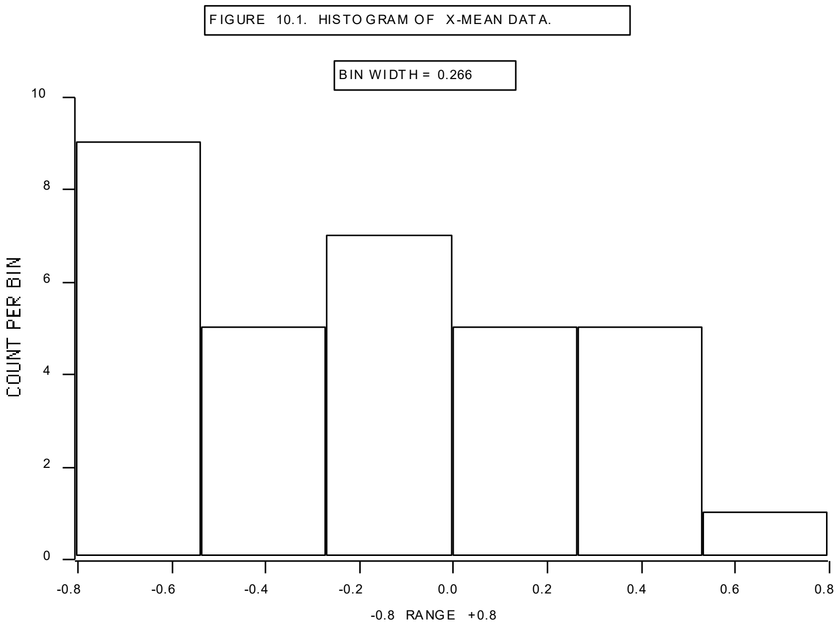


FIGURE 10.1. HISTOGRAM OF X-MEAN DATA.
BIN WIDTH = 0.266
COUNT PER BIN
10
8
6
4
2
0
-0.8
-0.6
-0.4
-0.2
0.0
0.2
0.4
0.6
0.8
-0.8 RANGE +0.8

FIGURE 10.2. HISTOGRAM OF X-MEANS (3 BINS).

BIN WIDTH = 0.533

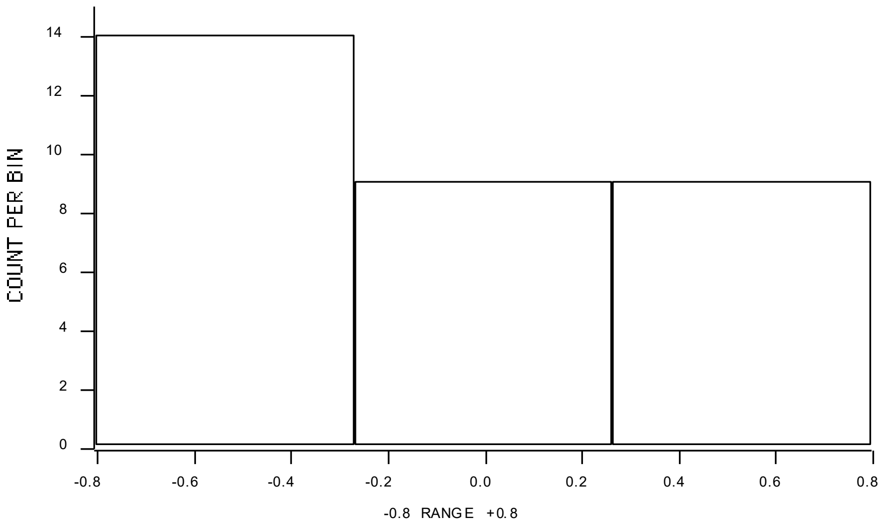


FIGURE 10.3. HISTOGRAM OF X-MEANS (12 BINS).

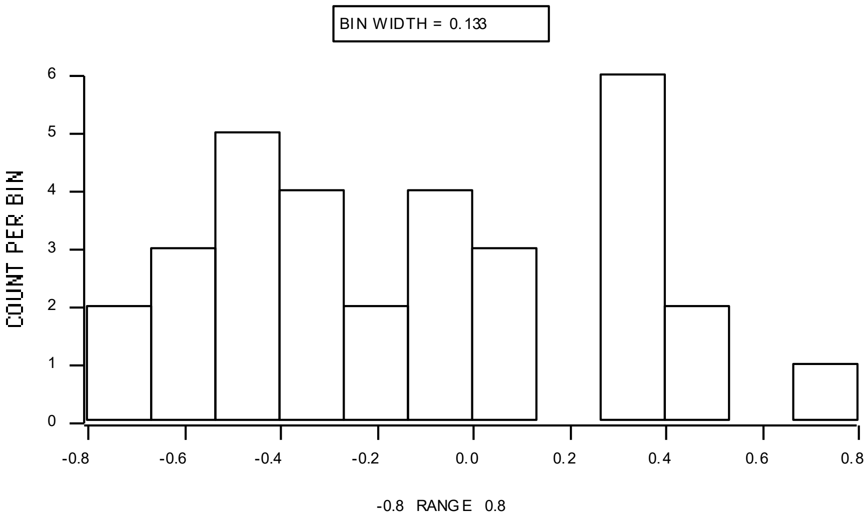


TABLE 10.1 RECOMMENDED QUANTITY OF BINS (k) FOR SAMPLE SIZE (N).

N	k
24-45	6
46-91	7
92-181	8
182-362	9
363-725	10
726-1448	11
1449-2896	12

Experience has shown that histograms with less than five or six bins are of limited usefulness, so the table is not extended to lower values. At the higher end of the table, 12 bins will suffice for nearly 2900 datum points, which should accommodate most applications.

Development of the Table

The following is a discussion of the development of Table 10.1. A knowledge of the development is not necessary for use of the table, but gives some insight into how and why the histogram is so very useful.

Herbert A. Sturges (reference at end of chapter) considered a practical problem which arises when plotting a histogram of numerical values from a statistical sample of size N having range R, the range being defined as the largest value minus the smallest value. Sturges wished to estimate the optimum class interval for the histogram. (Today the class interval is known as the cell width or bin width.) Sturges gave the following equation for estimating the class interval:

$$C = R/[1 + 3.322LG_{10}(N)] \qquad \textbf{EQ.10.1}$$

Where C is the optimal class interval (cell width or bin width) R is the range (largest value minus smallest value) N is the sample size (number of items in the sample) LG_{10} is the logarithm to the base 10 (common logarithm).

Sturges did not provide an explanation of how the equation was developed except to hint that it is based on a series of binomial coefficients. The following exposition of the likely development of Sturges' rule is from G.J. Bradley.

Given the binomial expansion of $(q+p)^m$: Where p is the probability of success in a single trial.

q is 1-p (= probability of failure)

m is the number of trials

For m going from 0 to 4	
0	1
1	$q+p$
2	$q^2+2qp+p^2$
3	$q^3+3q^2p+3qp^2+p^3$
4	$q^4+4q^3p+6q^2p^2+4qp^3+p^4$

If now we set down only the numerical coefficient of each term above, we have (Pascal's triangle):

m numerical coefficients										
					1					
				1		1				
			1		2		1			
		1		3		3		1		
4		1		4		6		4		1

Note that the number of coefficients is always m+1 and that the sum of the coefficients is 2^m.

A useful application of Pascal's triangle is to give the frequency distribution of the combinations expected in m repeated trials of an event. For example, m = 4 corresponds to the outcome of tossing four coins. A simple enumeration of the possible outcomes of the toss will show:

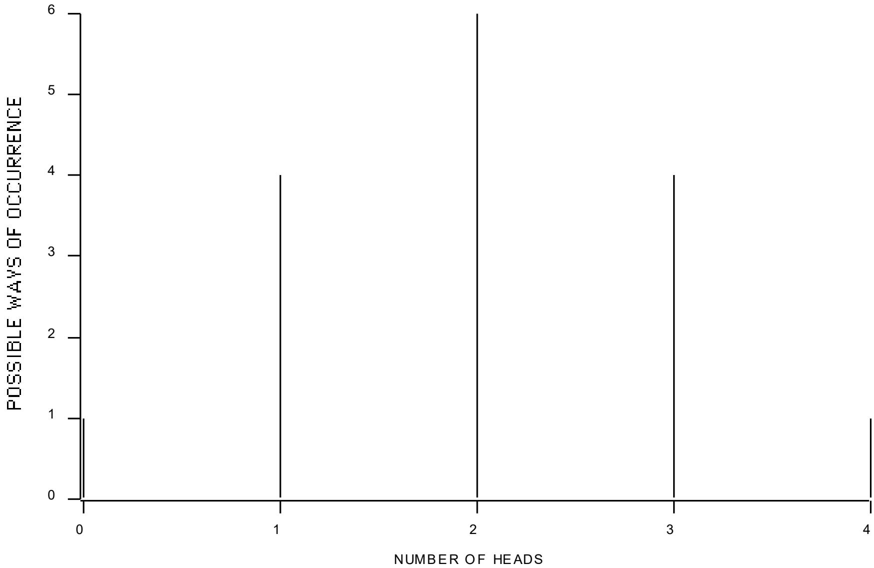
number of heads	number of ways to obtain
0	1
1	4
2	6
3	4
4	1

This result is given by the m = 4 row of Pascal's triangle. As previously stated, there are (m+1) terms having $(1+1)^m$ or 2^m possible combinations.

A graph of the m = 4 line from Pascal's triangle is given in Figure 10.4. This may be thought of as the possible results of tossing four coins. The coins are assumed to be fair; that is, the probability of heads is 0.5 as is the probability of tails. As Figure 10.4 shows, 0,1,2,3, or 4 heads can occur in 1,4,6,4,1 possible ways, respectively.

Now if a random sample of 16 elements (sample size = 16) is drawn from a symmetrical unimodal parent population, we should expect that the histogram of the sample should have the general form or shape as shown in Figure 10.4. The maximum of the histogram should occur near the central value of the population, and the dispersion of the sample will be determined by the dispersion of the parent population.

FIGURE 10.4. POSSIBLE WAYS OF OCCURRENCE OF HEADS IN TOSS OF FOUR COINS.



This suggests that Pascal's triangle (the binomial coefficients) might be used as a guide in plotting histograms. More specifically, consider a sample of size N which is to be plotted as a histogram with k bins. Each one of the N items is to be placed into one of the bins. By analogy with Figure 10.4, the number of bins k corresponds to the several classes of events (in Figure 10.4, the number of heads), which is (m + 1). Similarly, the sample size N corresponds to the total number of possible combinations, which is 2^m. That is $2^m = N$.

Using logarithms to the base 10

$$mLG_{10}(2) = LG_{10}(N)$$

so

$$m = [LG_{10}(N)]/LG_{10}(2)$$

Now let $m+1 = k$ where k is the number of terms. That is, k corresponds to the number of possible outcomes of the toss. Thus if $m = 4$, $k = 5$, and the five possible outcomes are 0,1,2,3, or 4 heads. Then

$$k = 1+ [LG_{10}(N)]/LG_{10}(2)$$

or

$$k = 1 + [LG_{10}(N)]/(0.301030)$$

and

$$k = 1 + 3.32193(LG_{10}(N))$$

Now if a statistical sample of size N with range R is plotted as a histogram with k equally sized classes (cells or bins), then the class interval (cell width or bin width) C is

$$C = R/k$$

By substitution in this equation

$$C = R/\{1 + 3.322[LG_{10}(N)]\}$$

which is Sturges' rule.

Histogram plotting via Sturges' rule is an excellent analytical tool in itself. This writer has used Sturges' rule plotting in studying the extreme-value distribution of projectile dispersion. But it seems that an important property of Sturges' rule plotting is not fully appreciated: Sturges'rule plotting is a symmetry test. The symmetry (or asymmetry) of a probability distribution is a property of fundamental importance. It is a prime characteristic to look for when attempting to identify the underlying population distribution from a study of a sample drawn from that population. Since Sturges' rule is based upon the perfect symmetry of the binomial coefficients, it provides a "symmetry yardstick" or benchmark against which to measure the symmetry (or asymmetry) of the sample.

Although Sturges' rule plotting is useful, it is not very sensitive. That is, it is not applicable to small samples. Experience has shown that it is necessary to plot at least six cells in order to begin to see asymmetry in the sample. However, Sturges' rule is not rigid. It is only a guide to rational plotting of data.

Computation of Table 10.1

Table 10.1 is computed by taking the nearest integer to

$$N = 2^{(m-.5)}$$

for the lower limit and

$$N = 2^{(m+.5)}$$

for the upper limit. For example, if k = 6 (six bins), m =5 and

$2^{(4.5)} = 22.6$ (lower limit = 23)

and

$2^{(5.5)} = 45.25$ (upper limit = 45).

For k = 7, the lower limit is computed by adding 1 to the upper limit for k = 6, which gives 46. The upper limit for k = 7 is computed as before, yielding 91. The other bin limits are developed in a similar manner. A similar table (Table 22-5) may be found in "The Quality Control Handbook." The use of either table will result in more satisfactory histogram plots.

NOTES AND REFERENCES FOR CHAPTER 10

The table referred to by Mr. Bradley is in section 22-4 of:

Juran, J. M., “Quality Control Handbook,” 3rd Ed., McGraw, 1974.

James R. King’s work is the best available for practical statistical analysis. See chapter 3 of the following book for Sturges' rule and tips on data presentation.

King, J. R. , “Probability Charts for Decision Making,” Industrial Press, Inc., NY, 1971.

A brief discussion of Sturges’ rule and histogram plotting is on pp.139-140 in the following book:

National Association of Relay Manufacturers, “Engineers’ Relay Handbook,” Hayden Book Company, Inc., NY, 1966.

Sturges’ paper is found on pp. 65-66 of:

Sturges, H. A., “The Choice of a Class Interval,” American Statistical Association Journal, March 1926.

The exposition of the development of Sturges’ rule is contained in a letter from G. J. Bradley to James R. King, letter dated September 7, 1978. Mr.Bradley is now deceased. The development in this chapter is published with the kind permission of Mr. King.

The concept of measurement as a classification is from Sec. 7.7 of:

Wilson, E. B., Jr., “An Introduction to Scientific Research,” McGraw-Hill, 1952.

CHAPTER 11

EXTREME-VALUE STATISTICS APPLIED TO THE EXTREME SPREAD IN BALLISTIC DISPERSION.

Introduction

In exterior ballistics, the dispersion of rounds about the average trajectory is a characteristic of great importance. Frank Grubbs considers the various measures of dispersion in his monograph (see references at end of chapter). Grubbs discusses eight different measures of dispersion. One measure of dispersion much used by shooters is the extreme spread, defined as the maximum distance between all possible pairs of bullet holes. As discussed in Chapter 9, in selecting pairs of points from a group of N points, we are selecting combinations of two things from a group of N things. In this case, the total number of such combinations is given by EQ. 9.1, repeated here:

$$COMB(N;2) = (N!)/(2!)(N-2)! = n = N(N-1)/2 \qquad \textbf{EQ.11.1}$$

where $N! = (N)*(N-1)*(N-2)*...*(2)*(1)$.

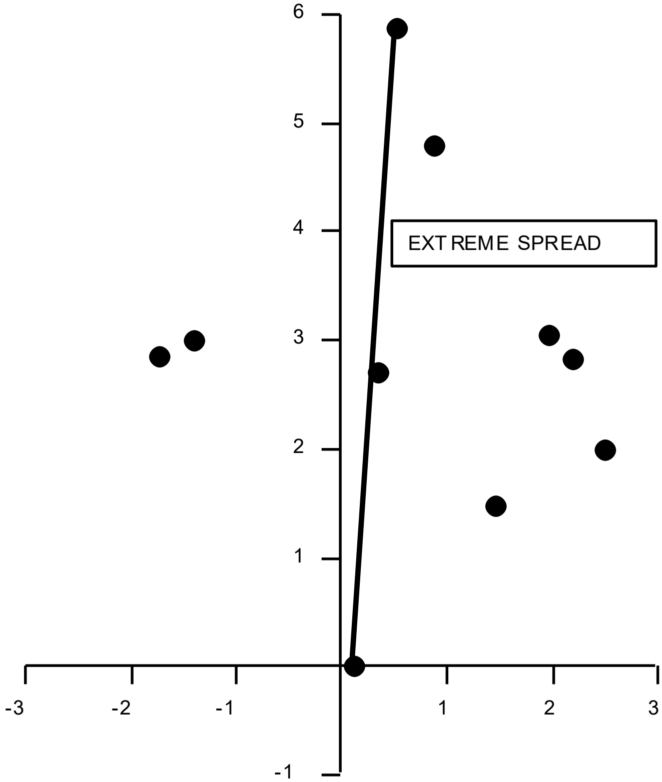
A group of ten rounds would give a total of 45 extreme spreads. By convention, the maximum of the extreme spreads would be selected as a measure of dispersion of the group of rounds.

Figure 11.1 is plotted from the data of Table 6.1 and illustrates the extreme spread of a ten-round group of shots. The aiming point is at the crossing of the axes. The center of the group is about three inches above and one-half inch to the right of the aiming point. The straight line labeled as ‘Extreme Spread’ shows the maximum distance between any two impact points, bullet holes in this case. On a particular small arms target, the extreme spread may be apparent. Occasionally, two or more spreads may be near in magnitude, and several measurements must be made to determine which is the largest.

In naval gunfire support, we would consider the extreme spread of the fall of shot. Artillery fire controllers might look to the extreme spread of the pattern of impact points of rounds. Evaluators of air or missile strikes would view the extreme spread of bomb, rocket, or missile impacts. In all these cases, the interest in and the mathematics needed to deal with the phenomena are the same.

In studying the statistics of ballistic dispersion, the writer fired upon 44 targets while shooting upon a small-arms range. Table 11.1 gives the data of interest.

FIGURE 11.1 BULLET IMPACTS ON VERTICAL TARGET
SHOWING EXTREME SPREAD.



DIMENSIONS IN INCHES.
AIMPOINT AT CROSSING OF AXES

TABLE 11.1 MEASURED EXTREME SPREADS FROM FORTY FOUR TARGETS

E.S.	E.S.	E.S.	E.S.
3.81	5.75	6.84	8.12
3.91	5.84	6.88	9.12
4.06	5.88	7	9.84
4.09	5.97	7.09	10.41
4.34	5.97	7.12	10.44
4.88	6.12	7.25	10.59
5	6.22	7.28	10.81
5	6.41	7.28	12.06
5.12	6.44	7.38	12.09
5.12	6.56	7.5	12.31
5.25	6.84	8	16.56

NOTES:

(1) FIREARM USED: 22 CALIBER REVOLVER.

(2) AMMUNITION: COMMERCIAL 22 LONG RIFLE CARTRIDGES, STANDARD VELOCITY.

(3) TEN ROUNDS FIRED ON EACH TARGET.

(4) RANGE: 25 YARDS.

(5) EXTREME SPREAD IN INCHES MEASURED ON TARGET PLANE.

A Search for the Probability Distribution of the Extreme Spread

As shown in Figure 11.1, the extreme spread is the largest of all the possible spreads which may be measured upon a given target or group of impact points of rounds. The extreme spread is one of several available measures of dispersion. Since the extreme spread is determined by the impact positions of the rounds, which positions are themselves random variates, the extreme spreads are also random variates.

In analyzing test or field data, the analyst is well advised to identify the probability distribution of the random portion of the measured data from the phenomenon under study. A knowledge of the underlying probability distribution will sharpen the analysis, reveal answers to questions of interest, suggest fruitful areas for investigation and also enable one to make rational predictions of performance of systems under examination. With identification of the probability distribution as our goal, let us return to study our data.

Histogram Plot

As a first step in studying the measured data, the 44 extreme spreads were plotted in a histogram. Sturges' rule was consulted for the number of bins for the histogram. As Table 10.1 indicates, six bins is recommended for a sample size of 44. Since the rule is a guide and not a law, seven bins were used, as seven seemed to better reveal the structure of the distribution of data.

Still not satisfied with the plot, the common logarithms of the data were plotted in a histogram, which is shown in Figure 11.2. The most striking characteristics of this histogram are the strong central tendency as shown by the tall bin near the center; and the right-hand skew.

Both characteristics suggest that either a log-normal or an extreme-value plot might best fit the data. This may be made more clear by comparing Figure 11.2, the histogram plot, with Figure 11.3, which is a graph of the extreme-value probability density function. A discussion of the extreme-value probability function is given later in this chapter. For now, the important thing is to see the general similarity in shape between the histogram of Figure 11.2 and the probability density function of Figure 11.3.The histogram displays the skewed distribution of values characteristic of extreme-value phenomena. The right-hand skew of the histogram indicates that the extremes are of large values. A skew of the histogram to the left would indicate that the smaller values are increasingly more rare.

Figure 11.3 is a graph of the extreme-value probability density function. This function describes the probability distribution of the extremes of many different kinds of random variates. The extremes referred to are the largest values of the variate which are observed in a group of samples of that variate. As an example, the floods of rivers and streams illustrate the extremes of a physical phenomenon, as the floods are the largest of the flows. The flow of a particular river is observed each day, which results in a sample size of 365 observations in a typical year. The largest flow in a given year is chosen to represent the flood for that year. This largest flow is of course the extreme of the given sample which was recorded for a particular year. Gumbel proposed that the extreme-value probability distribution be used to describe flood flows. Clearly the flood flows are random variates and certainly are examples of the extremes of phenomena.

In the following, the mathematical development of the extreme-value distribution is sketched. A brief description of the construction of the extreme-value probability plot is given. For details, the reader is urged to consult the books and papers by J. R. King and E. J. Gumbel. The data of table 11.1 are plotted in Figure 11.4 to test the possibility that the extreme spread in ballistic dispersion obeys the extreme-value probability distribution.

FIGURE 11.2 FREQUENCY OF OCCURRENCE VS.
COMMON LOGARITHMS OF EXTREME SPREADS.

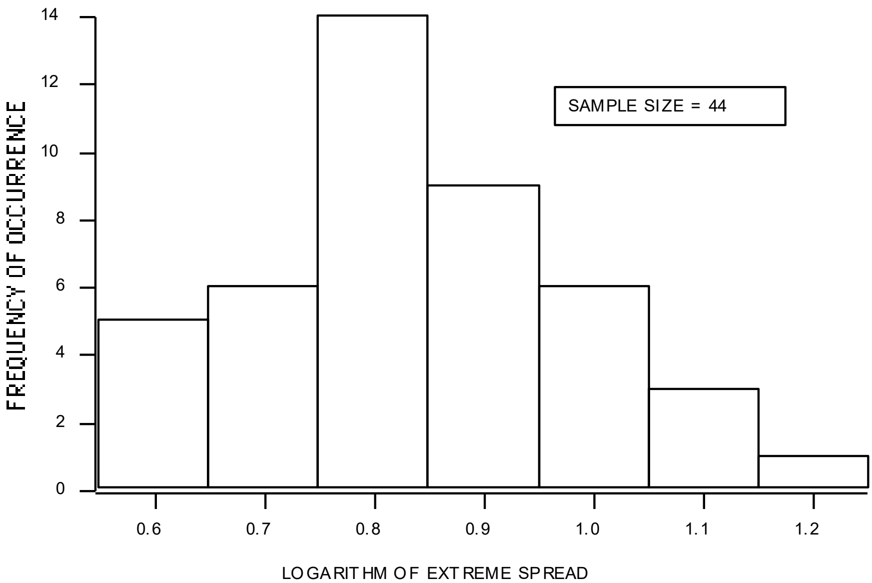
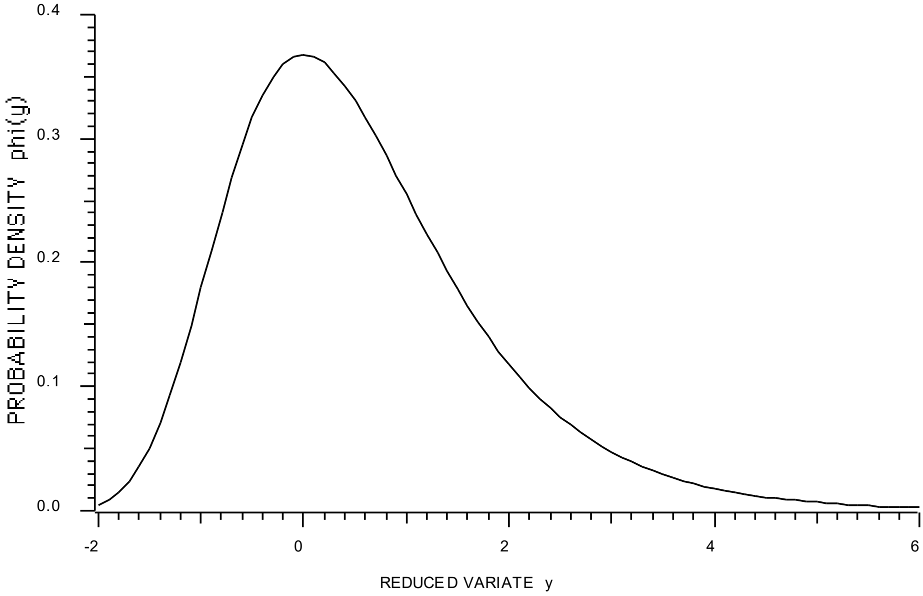


FIGURE 11.3 EXTREME-VALUE PROBABILITY DENSITY FUNCTION



The equation used by Gumbel to describe the probability density function of the extremes is:

phi(y) or f(y) = EXP[-y-EXP(-y)] **EQ. 11.1**

where -_<y<+_.

The reduced variate "y" is given by:

y = a(x-u) **EQ.11.2**

where a is a scale parameter and u is a location parameter, both computed from the data, as discussed later. The parameter u is the most probable value of the variate. The reciprocal of a is the slope of the theoretical straight line on an extreme-value probability plot. For the theoretical extreme-value probability density function of EQ.11.1, the most probable value is the maximum, which occurs at y = 0. Hence the theoretically most probable value of x is u , for which the probability is 1/e = 0.367879...

Extreme-Value Plot

The data were plotted on both log-normal and extreme-value probability graphs. The best fit was given by the extreme-value plot, as shown in Figure 11.4.

The abscissa in Figure 11.4 is the cumulative extreme-value probability distribution, given by integrating EQ.11.1 from minus infinity to y, and is:

F(y) = EXP(-EXP(-y)) **EQ.11.3**

where y is given by EQ.11.2, as before.

Extreme-value probability plots are sometimes confusing to those unfamiliar with them. In early applications, the variate was termed "x," and was plotted as the ordinate. The reduced variate "y" was plotted as the abscissa. This is the reverse of our usual convention with cartesian-coordinate graphing, and takes some getting used to. The reduced variate y sometimes is plotted beneath the cumulative probability scale as an additional abscissa scale. This y scale is used primarily in constructing the graphs.

Figure 11.4 , in which the logarithm of the variate fits the extreme-value distribution, is an example of Gumbel Type II extreme-value behavior. The theoretical line in Figure 11.4 is given by:

$x = (1/a)*(y) + u$ **EQ.11.4**

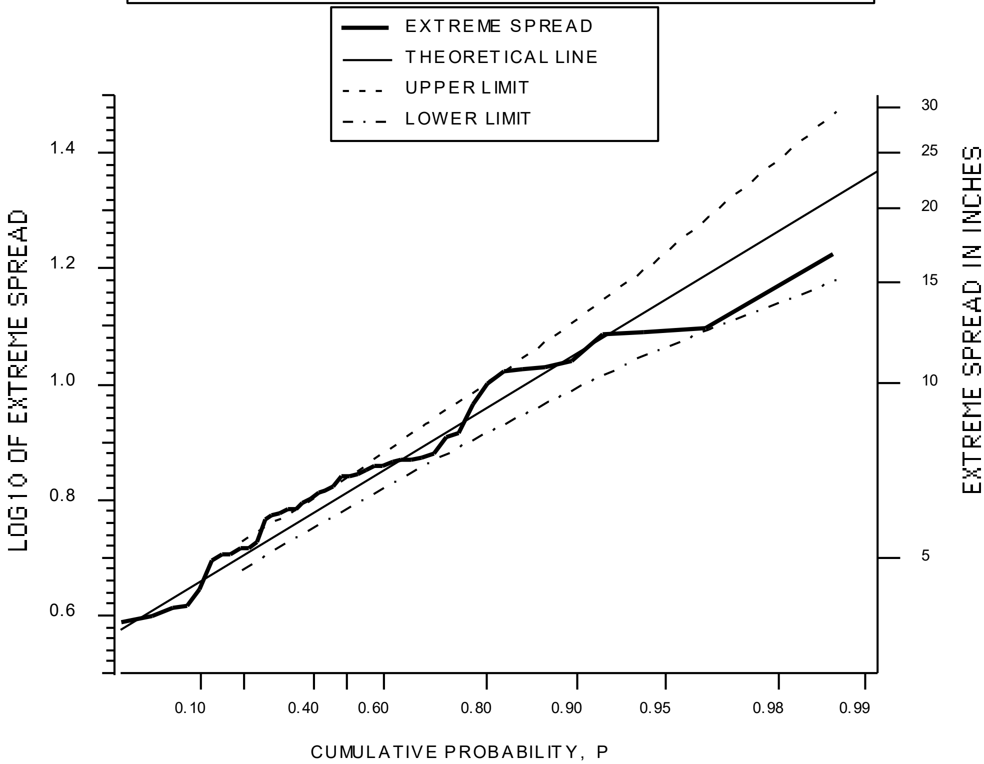
where

$(1/a) = (\text{STD. DEV.})/s_N$ **EQ.11.5**

and

$u = (\text{MEAN}) - (y_N)/a$ **EQ. 11.6**

FIGURE 11.4 EXTREME VALUE PLOT OF EXTREME SPREADS



The standard deviation and mean in EQ.11.5 and EQ.11.6 above are computed from the data. The factors s_N and y_N are given in a table by Gumbel and Carlson. For our sample size of 44, s_N equals 1.1499, and yN equals 0.5458

Common logarithms of magnitudes of the extreme spreads were plotted using the left-hand ordinate scale. The data values are ranked in order of increasing size, and the plotting position on the probability abscissa computed from

$p_i = (i - 0.44)/(n + 0.12)$ **EQ. 11.7**

where i is the index number of rank of a particular value.Thus the smallest of the extremes has rank 1, the next larger has rank 2, and so on to the largest of the extremes, which has rank n, where n is the sample size. In our case, n equals 44. This equation for computing the probability plotting position is from page 51 of J. R. Kings's book "Frugal Sampling Schemes."

The fit is good, with most points within the one-sigma limits shown.

Evaluation Dispersion from the Extreme-Value Plot

The computed theoretical line from EQ.11.4 is shown in Figure 11.4. The theoretical line gives the probability corresponding to a given extreme spread. For example, reading the right-hand scale at 15 inches, the corresponding probability is about 0.95. Hence the probability is about 0.95 that a 10-round group fired under similar conditions will show an extreme spread of 15 inches or less. Similarly, the probability is 0.99 that no 10-round group will exceed 23 inches extreme spread, under similar firing conditions. The median probability of 0.50 could be used as a measure of ammunition production quality when lot test samples are fired under controlled conditions. For the data considered here, the median value of extreme spread corresponding to the cumulative probability of 0.50 is about 6 inches. Many other numerical estimates may be made from the graph. For example, an extreme spread of 5 or less inches has a probability of occurrence of about 0.12. As previously noted, an extreme spread of 15 inches corresponds to a cumulative probability of 0.95. The probability that the extreme spread in a given test will lie between 5 and 15 inches is the difference between these probabilities, 0.95 minus 0.12, or 0.83. This prediction will be correct so long as ammunition quality and test conditions remain the same.

Evaluating the Goodness of Fit of the Plot to the Data

The upper and lower one-sigma limits are used in Figure 11.4 to verify the goodness of fit. Values for the curves are computed from a table of factors, also given by Gumbel and Carlson. Gumbel considers the data values which lie along probabilities in the interval $0.15 < F < 0.85$. Data values along that interval are normally distributed about their mean, which is the straight line in the graph. For the normal distribution, plus and minus one sigma encloses 0.68 of the area under the normal curve, which is the probability that a variate will lie within one standard deviation of the mean. The probability that a given data point might lie outside the one-sigma curves is thus about 0.32, or roughly one chance in three. The most extreme of the extremes are not normally distributed, and Gumbel makes special provisions to calculate the control curves for the three largest data values. If, after the plot is made, the analyst sees that the great majority of data points lie within one-sigma limits, the "fit" may be said to be "good." This is a judgment call by the analyst. Formulas and residuals may be used to refine the decision.

Additional Procedures for Plotting and Analysis

The development of extreme-value and many other types of probability plotting has been greatly extended by James R. King in the references given at the end of this chapter.

NOTES AND REFERENCES FOR CHAPTER 11

Grubbs, F. E., Statistical Measures of Accuracy for Riflemen and Missile Engineers, 1964.

The tables used for plotting Figure 11.4 are found in:

Gumbel, E. J., and Carlson, P. G. , "Extreme Values in Aeronautics," Journal of the Aeronautical Sciences, Vol. 21, No. 6, pp. 389-398, June, 1954.

See also:

Gumbel, E. J., "Statistics of Extremes," Columbia University Press, NY 1958.

The most efficient methods for probability plotting are given by King in:

King, J. R., "Probability Charts for Decision Making," Industrial Press, Inc., NY, 1971.

Chapters 11, 12, and 13 cover extreme-value plotting.

See also:

King, J. R., "Frugal Sampling Schemes," Technical and Engineering Aids for Management, Tamworth, NH, 1980.

Appendix C discusses extreme- value plotting. Anyone performing statistical analysis should have this book.

CHAPTER 12

AN INVESTIGATION OF HUMAN VISION IN THE ESTIMATION OF POSITION

Introduction

This investigation was prompted by a remark made by a senior U.S. Navy aviator. This pilot had much experience in making air searches of large ocean areas. He was quite familiar with all the varied types of electronic equipment for detecting ships and submarines, but insisted that the instrument which he called the "MARK I eyeball" was not fully appreciated. He said that the information to be gained by simply flying over the ocean and looking around is truly remarkable.

Human Factors

Much improvement in operations of systems of any kind may be obtained by study of the interface between the human operator and the equipment being operated. This area comes under the general heading of "human factors." The writer is not an expert in this field, but made the investigation discussed here to illustrate the application of statistical methods to a psychovisual study.

The Experiment

This chapter discusses a simple experiment which was performed to investigate the ability of human vision to estimate the center of impact of a group of rounds. This is an important procedure in adjusting the sights of a small arm or other direct-fire weapon. Typically, on a rifle range, the rifle coach will have the shooter fire two or three rounds at the target to obtain an estimate of where the rifle is shooting. The shooter is told to aim at a particular point on the target, usually the bottom ("six-o'clock hold") of the bullseye. The rifle coach observes the strike of the rounds on the target and then recommends the necessary sight adjustments to bring subsequent shots to the center of the target. It would appear that the coach is in some way estimating the center of the group, or center of impact of the rounds. An experiment was developed to investigate the accuracy and dispersion of these visual estimates.

A Sketch of the Experiment

One hundred and ninety-two numbers were generated by a psuedo-random number generator. The numbers have a normal (Gaussian) distribution with zero mean and standard deviation of 1.0. These numbers were grouped into pairs, thus defining the x-y coordinates of 96 points. The points were divided into groups of three and each group plotted on a one-tenth-inch grid. Each number was rounded to one decimal place to fit the grid. This process yielded 32 graphs, similar to Figure 12.1.

An "eyeball estimate" of the center of each group of three shots was made and marked on each graph, as shown on Figure 12.1. The x-y coordinates of each estimate of center of each group is given by Table12.1.

FIGURE 12.1. ESTIMATE OF CENTER OF IMPACT OF THREE-ROUND GROUP.

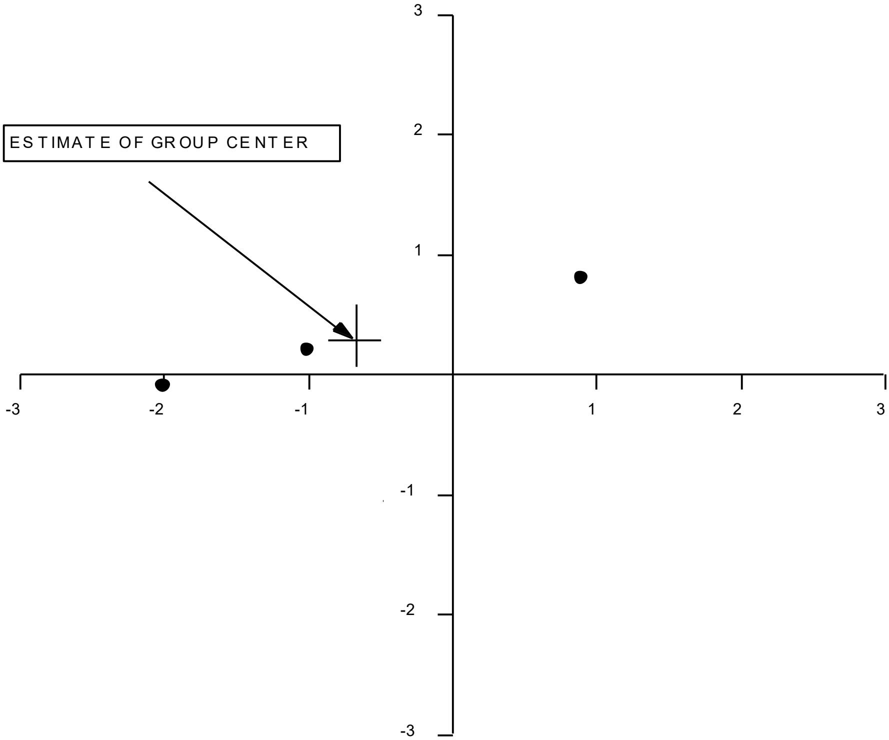


TABLE 12.1 COORDINATES OF ESTIMATES OF CENTERS OF GROUP

TARGET NO.	X	Y		TARGET NO.	X	Y
1	-0.7	0.3		17	0.4	-0.6
2	-0.3	-0.1		18	0.1	-0.3
3	-0.4	0.2		19	-0.5	-0.2
4	-0.3	-0.7		20	-0.3	-0.6
5	-0.1	-0.7		21	0.2	0.7
6	-0.6	0.2		22	0.3	-0.3
7	0.2	-0.1		23	0.1	-0.6
8	-0.3	0.7		24	-0.6	0.1
9	-0.6	-0.4		25	0.5	0.6
10	0.1	0.2		26	-0.1	-0.3
11	0.6	-0.2		27	-0.3	0.1
12	-0.5	-0.1		28	0.3	0.3
13	0.2	-0.4		29	-0.6	-0.1
14	0.4	0.5		30	-0.6	-0.1
15	-0.7	-0.1		31	0.6	-0.2
16	-0.1	-0.3		32	0.1	0.4

In addition, the mean horizontal (X) position and the mean vertical (Y) position was calculated for each of the 32 groups of three shots. Figure 12.2 shows the mean (X) and "eyeball estimate" (X) plotted versus group number, the groups being numbered from zero to 31. Figure 12.3 shows a similar plot for the mean (Y) and eye (Y) estimates. As may be seen from the plots, there is little to choose between the mean and the eye estimates. That is, for this small sample size, the human eye does very well at estimating the center of impact or group center.

A measure of performance to be considered is the dispersion of the calculated mean versus the dispersion of the estimates group centers. The comparison is made in Table 12.2.

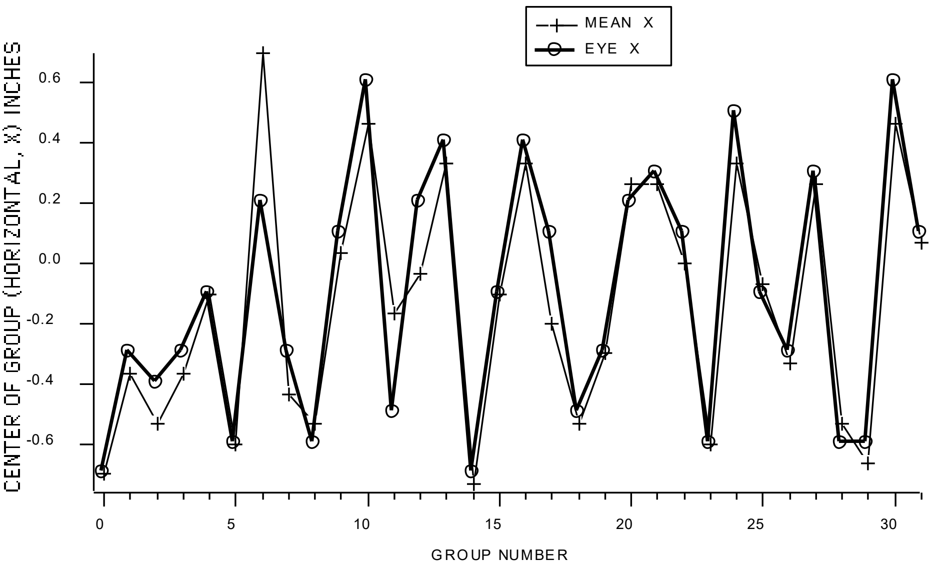
TABLE 12.2 STANDARD DEVIATIONS OF GROUP MEANS AND VISUALLY-ESTIMATED GROUP CENTERS FOR THIRTY-TWO GROUPS OF THREE ROUNDS EACH.

GROUP	MEAN(X)	EYE(X)	MEAN(Y)	EYE(Y)
0	-0.7	-0.7	0.3	0.3
1	-0.36667	-0.3	-0.03333	-0.1
2	-0.53333	-0.4	0.5	0.2
3	-0.36667	-0.3	-0.43333	-0.7
4	-0.1	-0.1	-0.63333	-0.7
5	-0.6	-0.6	0.166667	0.2
6	0.7	0.2	-0.6	-0.1
7	-0.43333	-0.3	0.666667	0.7
8	-0.53333	-0.6	-0.36667	-0.4
9	0.033333	0.1	0.266667	0.2
10	0.466667	0.6	-0.23333	-0.2
11	-0.16667	-0.5	-0.06667	-0.1
12	-0.03333	0.2	-0.33333	-0.4
13	0.333333	0.4	0.466667	0.5
14	-0.73333	-0.7	0	-0.1
15	-0.1	-0.1	-0.36667	-0.3
16	0.333333	0.4	-0.53333	-0.6
17	-0.2	0.1	-0.36667	-0.3
18	-0.53333	-0.5	-0.1	-0.2
19	-0.3	-0.3	-0.6	-0.6
20	0.266667	0.2	0.7	0.7
21	0.266667	0.3	-0.43333	-0.3
22	0	0.1	-0.53333	-0.6
23	-0.6	-0.6	0.033333	0.1
24	0.333333	0.5	0.466667	0.6
25	-0.06667	-0.1	-0.16667	-0.3
26	-0.33333	-0.3	-0.03333	0.1
27	0.266667	0.3	0.333333	0.3
28	-0.53333	-0.6	-0.06667	-0.1
29	-0.66667	-0.6	-0.13333	-0.1
30	0.466667	0.6	-0.2	-0.2

31	0.066667	0.1	0.5	0.4
	MEAN(X)	**EYE(X)**	**MEAN(Y)**	**EYE(Y)**
STDEV	0.400122	0.408269	0.39379	0.395629

As may be seen from Table 12.2, there appears little significant difference between the estimates of group center provided by the human eye and the mean calculated from the coordinates of each round's impact point.

FIGURE 12.2. HORIZONTAL CENTER OF GROUP, CALCULATED MEAN AND ESTIMATED BY EYE.



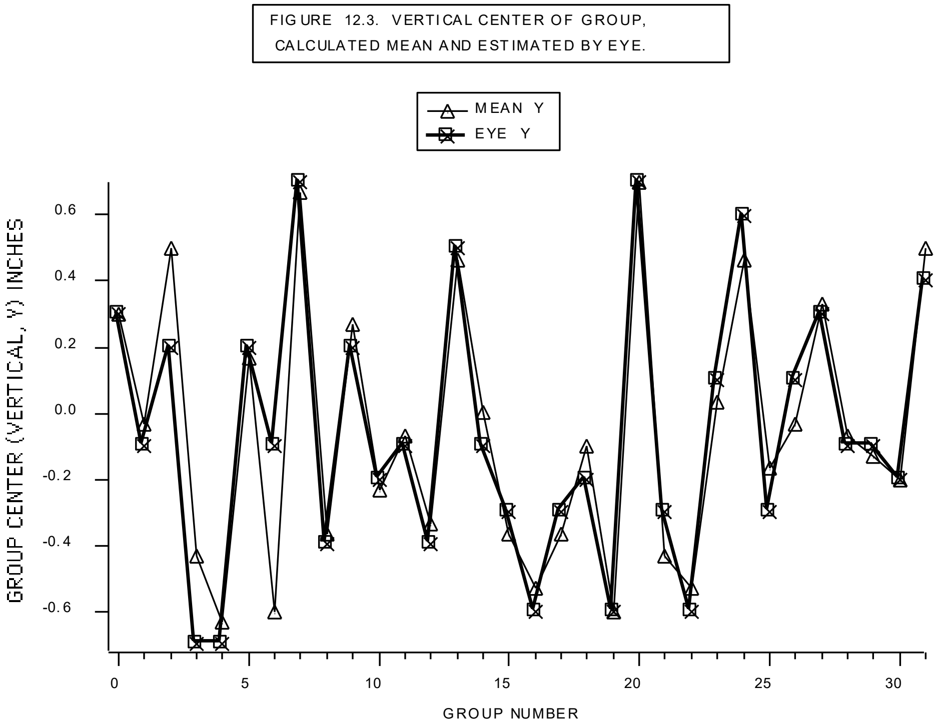


FIGURE 12.3. VERTICAL CENTER OF GROUP,
CALCULATED MEAN AND ESTIMATED BY EYE.
MEAN Y
EYE Y
GROUP CENTER (VERTICAL, Y) INCHES
0.6
0.4
0.2
0.0
-0.2
-0.4
-0.6
0
5
10
15
20
25
30
GROUP NUMBER

This result should bolster the confidence of those charged with adjusting the sights of direct-fire weapons in the field. It appears that a simple visual estimate of group center of impact is as reliable as more complicated numerical calculations, at least for small numbers of rounds (small sample sizes).

A reference at the end of this chapter discusses a different psychovisual study.

NOTES AND REFERENCES FOR CHAPTER 12

Cleveland, W. S., Harris, C. S., and McGill, R., "Judgments of Circle Sizes on Statistical Maps," Journal of the American Statistical Association, Vol. 77, No. 379, Applications Section, September 1982.

A note on quantization noise. In arbitrarily plotting the estimated group center coordinates on a one-tenth-inch grid, a quantization error of $0.1/(12)^{1/2}$ is introduced. This is $0.1/3.464 = 0.02887$ of root-mean-square quantization noise affecting each estimated coordinate value. This appears too small to have any significant effect. See pp. 327-329 of: Schwartz, M., "Information Transmission, Modulation, and Noise," McGraw-Hill, 1959.

GLOSSARY

Abscissa: In plane Cartesian coordinates, the x-axis.

Accuracy: Refers to the amount of error to be expected at the target. An accurate device or measurement system has a small error.

Adjustment (of fire): The process of correcting the aiming of the gun or battery to cause the rounds to strike the target.

Ammunition: The projectile, including its propellant, warhead or bursting charge, cartridge case, primer and fuze. Term also includes aircraft bombs, rockets and missiles of all types, as well as grenades and mines.

Artillery: Originally meant an implement of war, especially a military engine for throwing missiles. Use now generally limited to guns, howitzers, mortars and rockets in land warfare.

Ball: A type of bullet found in small arms ammunition. In early gun designs, the bullet was spherically shaped, hence the name 'ball'. In current designs, the bullet is elongated with a sharp or rounded point to reduce air resistance. Today, 'ball ammunition' means ordinary rifle, pistol or machine gun ammunition for use by infantry. The term is used to distinguish that ammunition from other types of ammunition such as armor piercing, incendiary, or tracer.

Ballistics: The study of the motion of projectiles of all types. Term comes from a Greek word meaning "to throw." The subject is further subdivided into interior, exterior, launch, and terminal ballistics.

Battery: A group of guns controlled and directed to bring the fire of all upon a given target or area. Formerly, in naval usage, all guns of a given caliber mounted upon a particular ship. Thus: the main battery of the battleship Alabama consists of nine sixteen-inch -45 caliber rifles.

Bias (in standard deviation): The method commonly used to calculate the standard deviation of a set of numbers is biased. This bias may be corrected if it is known that the sample is properly drawn from a population having a normal probability density distribution.

Bin: A subdivision of the range of a numerical sample of data. Generally all bins are the same width. See histogram.

Bivariate normal probability density function: A two dimensional probability density function, usually represented as a surface above the x-y plane. In general, the standard deviations of x and y are unequal. Also, the means of x and y are not equal.

Bivariate probability density function: A probability function in which the probability depends upon the values taken on by two distinct random variates.

Bullet: Projectile for small arm ammunition. Derived from French 'boule' for ball. Word is often used colloquially by non-professionals to mean cartridge.

Caliber: Diameter of gun bore. Thus: the U.S. Navy 5-inch gun has a bore diameter slightly smaller than the Soviet's 130 millimeter gun (127 millimeters versus 130 millimeters).Also used as a measure of length of gun tube. A 5-inch/54-caliber gun has a barrel length of 54 multiplied by 5 or 270 inches.

Cannon: A gun so large as to require a mount. Word perhaps derived from the Italian word for tube. Today, used in the term 'automatic cannon' such as the large-bore machine guns in aircraft and on ship and ground air defense systems.

Cannon-cocker: Artilleryman, in U.S. Marine Corps slang.

Cartridge: The unit of small arm ammunition, comprised of a tubular case which encloses a charge of propellant and is sealed on one end with a with a primer and at the other end with a bullet.

Cell: Name formerly used to mean what is today called 'bin'.See histogram.

Central tendency: The characteristic of a random variate that it clusters near a central point or value. A uniformly distributed variate has no central tendency. Normally-distributed variates tend to cluster toward the mean. Common measures of central tendency are the mean and median.

CEP: Circular error probable, which see.

Circular Error Probable: Usually referred to as CEP, although sometimes written as CPE. The radius of the circle on the x-y plane of the circular normal probability plot which encloses one-half of the variates. The same definition may be applied to a bivariate distribution in which the standard deviations are unequal, although the calculation may be more difficult.

Circular Normal Probability Density: A particular case of the bivariate normal probability density function in which the two standard deviations are equal.

Class interval: Name formerly used to mean what is today called bin or cell. Subdivision of range of sample of numerical data. See histogram.

Coriolis: Apparent deflection of a round as seen by an observer on the earth. The effect is caused by the rotation of the earth.

Confidence (or confidence level): A numerical estimate of the probability that a variate might lie within given limits, or above or below a given limit.

Cordite: A smokeless powder formerly used by the British as a propellant charge in small arm and some cannon ammunition. It is seldom encountered today. The word is used by journalists to refer to the odor found in the vicinity of a high explosive detonation. Whatever the journalist may smell, it is not likely to be cordite.

CPE: Circular probable error: Has same meaning as CEP, which see.

Cumulative Probability Function: A mathematical function or graph which describes the probability of occurrence of a variate of a particular magnitude or smaller. It is the definite integral of the probability density function.

Dispersion: The characteristic of a random variate to scatter across its range.

Distribution (of Probability): Generally, a mathematical relationship or function relating a random variate and its probability of occurrence. Sometimes used to mean "cumulative distribution function "

Dynamics: That branch of Physics which studies the motion of bodies and systems under the action of forces.

Electrothermal-chemical: Propellant system for guns under development.

Error: The numerical difference between a measured quantity and the "true" or correct value.

Exterior Ballistics: The study of the motion of projectiles after leaving the gun or launcher. Originally applied only to projectiles which are not self-powered.

Extreme spread: Term used in measuring the dispersion of rounds. It is the largest distance between any pair of rounds when impacts are plotted upon a plane. It is convenient for use in evaluating dispersion in small arms firing, as the extreme spread may be measured directly upon the target.

Extreme-value statistics: The study of the probability distribution of largest values of a variate. Early applications are to hydrology and other natural phenomena.

Factorial function: Defined for integers only:

$n! = n*(n-1)*(n-2)*...(1)$.

Fall of Shot: The pattern made by a group of gun projectiles.

Gamma function: One of the so-called "special functions" of mathematics. It is a generalized factorial function.

Gaussian: Name often applied to a population distributed according to the normal probability density function.

Gun: Any projectile-throwing device which uses chemical propellant or air or gas (or other, pending) to thrust a projectile through a tube and thus on toward a target.

Histogram: A graph of a tally of the number of data values which occur within given limits of the variates. Usually, the data are sorted into bins of equal width.

Howitzer: A field artillery piece with barrel relatively shorter than that of a gun. Its lower muzzle velocity gives the projectile a plunging trajectory upon the target, useful for engaging targets located on the reverse side of hills or behind other obstructions.

Human Factors: An engineering discipline in which equipment design is modified to give greatest efficiency of use by a human operator.

Interior Ballistics: (In Britain, Internal Ballistics) The study of the motion of projectiles inside the gun, including physical, chemical, and thermodynamic energy transfers.

JDAM: Joint direct attack munition. A glide bomb consisting of a MARK 84 2000 pound bomb or a BLU-109 2000 pound penetrating bomb. Both have Global Position System and Inertial Navigation System guidance.

JSOW: Joint stand-off weapon: A glide bomb in three variants. Two are cluster bombs. The third has a high explosive warhead. A later version may have a penetrating warhead. All have combined Global Position System and Inertial Navigation System guidance.

Launch Ballistics: The study of the motion of projectiles in the transition from gun to free trajectory. Today, more often concerns the transition of a rocket or missile from cell, silo, launcher rail or pylon to powered flight.

Laying (of a gun or battery): To point or aim so as to bring fire upon the target.

Log-normal: (Log-normal probability density or distribution) A probability function in which the logarithms of the random variate are distributed according to the normal probability density.

Mark I: (Mark One). In naval ordnance, the initial version of a particular equipment accepted for regular service.

Mean: The mathematical average of a function or set of numbers. For a sample of data the mean is the sum of the numbers divided by the quantity of numbers.

Mechanics: That branch of Physics which studies the action of forces on rigid bodies and motion. Includes the study of Statics, or bodies at rest.

Median: The center value. One-half of the probability density function has variates less than the median, and one-half of the variates are greater than the median. The centermost numerical value or data point of a sample.

Mil: A unit of measurement of angles. It is particularly useful for small angles. Today the mil is generally taken to be a milliradian, or 1/1000 of a radian. Thus there are $6{,}283^{+}$ mils in a circle. For small angles, the milliradian is approximately equal to the tangent of the angle According to General Hatcher, the Infantry adopted a slightly different definition of the mil. The Infantry definition is such that a circle contains 6,280 mils. The Infantry mil is no longer used. The Artillery adopted a yet different mil. The Artillery mil is defined so that a circle contains 6400 mils.

Missile: Any object thrown. Today usually means "Guided Missile."

Mortar: A type of gun having a relatively short barrel, used exclusively for high-angle fire. Formerly used as artillery pieces in land warfare and in coastal defense forts. Now used primarily as infantry weapons. Some were adapted for use in small craft (river patrol boats) in Viet Nam. Common calibers in U.S. forces are 60 and 81 millimeter, and 4.2 inch. Many foreign armed forces are equipped with 120 millimeter mortars.

Multivariate (Probability function): A probability function in which the probability is determined by the values of two or more random variates.

Munition: Any equipment for war, but especially weapons and ammunition

Nonparametric: A probability function or statistic which does not use parameters. Also termed "distribution free."

Normal or Gaussian Random (Process): A process in which the probability of occurrence of a given variate is described by the normal probability density function.

Normal Probability Density Function or Distribution: A particular probability density function or distribution found to describe many natural processes. Has the widest applicability of any probability density function.

Ordinate: In plane Cartesian coordinates, the y-axis.

Ordnance: Any equipment provided for military, naval, or other armed force. Today, the term is usually restricted to weapons and their ammunition.

Parameter: A variable which determines the position or size or other characteristic of a mathematical function. For example, the mean and standard deviation are parameters of the normal probability density function.

Parametric Statistics: Statistics based upon probability density functions which contain parameters.

Parent Population: Population from which a sample is drawn. The term "Universe" is sometimes used in place of population.

Pattern: The impact points of a group of rounds.

PGM: Precision guided munition. A missile, bomb, or other projectile with accuracy (maximum error) of 3 meters.

Population: The total of all individual items under study. In theory, a population may be finite or infinite. Some writers prefer to use the term "Universe" in place of population.

Precision: Generally used to mean the same as accuracy. In a measurement device or system, may refer to the degree of fineness of a measurement. Thus for a given dimension, a measurement made to within 0.001 inch is more precise than a measurement made to within 0.1 inch.

Probability Density Function or Distribution: A mathematical function or graph which describes the probability of occurrence of a variate. The integral of the probability density function gives the cumulative distribution function.

Propellant: Chemical material used to generate high pressure gas and thereby propel a projectile from a tube.

Psuedo-random: "False Random." A process in which a sequence of numbers is generated by a device or algorithm. There is nearly no correlation between adjacent numbers or between numbers not widely separated in sequence. However, the sequence of numbers will eventually repeat. Hence the number sequence is not truly random. Nevertheless, a short sequence of numbers is apparently random, and may be considered random with small chance of bias.

Psychovisual: The process or operation of the eye and brain in seeing and forming images.

Quality Control: System to measure and adjust a manufacturing process to maintain a product within specified limits.

Quantization noise: Uncertainty introduced into a measurement by the process of representing a signal or variable by finite steps or levels of amplitude, phase, frequency, or other characteristic.

Random: A characteristic of a process in which the end, result, or outcome is never definitely known or predictable in advance.

Range (1): Distance from launch point of munition to target. May have components such as slant range, horizontal range, height, or range to intercept point.

Range (2): For a sample of numerical data, the maximum value minus the minimum value.

Round: A projectile of any kind. The term originally referred to spherical shot used in early gun designs.

R/CEP: Ratio of radius to circular error probable. A normalizing ratio used to standardize calculations.

Salvo: A group of rounds fired (or launched) within a short time interval at a single target.

Sample: A finite number of measurements or observations drawn from some population, universe, or process.

Sample size: The total number of data values.

Sight Adjustment: Change made to a weapon's sights or sighting system to bring the round impact point on to the target.

Sighting-in: The process of adjusting the sights of the gun or weapon so that the projectile strikes the point of aim at a given range.

Small arms: Arms which are hand-held and carried by one person. Today, the term is usually applied only to firearms such as rifles, shotguns, submachine guns, pistols and revolvers. In infantry usage, may include grenade launchers and perhaps machine guns.

Spotting: The process of observing the fall of shot or impact of rounds and adjusting the fire so as to bring the maximum effective fire upon the target.

Standard circular normal probability density function: A circular normal probability density function in which the standard deviations are equal to 1.0 and the means are zero.

Standard deviation: A measure of scatter or dispersion of a sample of data or measurements. Also a measure of the amount of spread of a probability density function. The standard deviation is the square root of the variance.

Stick (of bombs): A group of bombs arranged to be released from an aircraft to fall in a row across a target. The term is descriptive of the elongated pattern made by the line of impacts created when several bombs are released in sequence as the aircraft flies in a nominal straight and level path.

Sturges' rule: A method for estimating the optimal number of bins to use when plotting a histogram.

Symmetry: If a probability density distribution is an 'even' function, then p(x) is equal to p(-x), and the distribution is symmetrical about its mean. This symmetry may be expected to appear in a histogram of a random sample drawn from a population having a symmetrical probability density. Similarly, a non-symmetrical probability density function may be expected to generate a non-symmetrical histogram. Thus the symmetry (or non-symmetry) of a histogram of a random sample from an unknown population is an important clue as to the form of the probability distribution of the population under study.

Terminal Ballistics: The study of the effects of a particular design of munition upon a target.

Theory of Errors: Early term applied to the study of errors and uncertainty of measurements, especially the application of the theory of probability to that study.

Trajectory: The path of a round through space from launch to impact.

Tube: Word sometimes used to refer to the barrel of a gun, howitzer, or mortar. Thus howitzers may be referred to as 'tube' artillery to distinguish them from rocket batteries. The distinction is less apt today, as many current rockets are launched from tubes.

Universe: A term used to mean the same as population. The totality of all variates that may be observed or that may occur. The total of all outcomes of a trial.

Variance: For a sample, the variance is the mean of squares of differences taken between each variate and the mean of all the variates. For a probability density function, the variance expresses the amount of spread the probability density function has about its own mean. The square root of the variance is the standard deviation.

Variate: A term used to mean a random variable.

Warhead: That portion of a munition which contains the payload. The payload may be explosive, incendiary, smoke, gas, toxin, or biological agent. Some munitions such as bullets or mortar, howitzer, gun and cannon projectiles do not have separate warheads.

BIBLIOGRAPHY

Abramowitz, M., and Stegun, I. A., "Handbook of Mathematical Functions," National Bureau of Standards, 1964.

Backman, M. E., "Terminal Ballistics," NWC, China Lake, CA, 1976.

"Ballisticians in War and Peace. A History of the United States Army Ballistics Research Laboratories 1914-1956," Vol. 1, United States Army Ballistics Research Laboratories, Aberdeen Proving Ground, MD, November, 1975.

Burington, R. S. and May, D. C., Jr. , "Handbook of Probability and Statistics with Tables," Handbook Publishers, Inc., Sandusky, OH, 1958.

Campbell, L. M., "Some Applications of Extreme-Value and Non-Parametric Statistics to Naval Engineering," Naval Engineers Journal, Vol. 89, No. 6, Washington D. C., December, 1977.

Campbell, L. M., "Applications of Hybrid Statistics and the Median," Naval Engineers Journal, Washington, D.C., January, 1983 and pp. 153-156, July 1983

Canan, J. W., "Smart and Smarter," Sea Power, Vol. 38, No. 4, Navy League of the United States, Arlington, VA, April 1995.

Cleveland, W. S., Harris, C. S. , and McGill, R. , "Judgments of Circle Sizes on Statistical Maps," Journal of the American Statistical Association, Vol. 77, No. 379, Applications Section, September 1982.

Crow, E. L., Davis, F. A. , Maxfield, M. W. , "Statistics Manual," Dover, NY, 1960.

Gnedenko, B. V. and Kinchin, A. Ya., "An Elementary Introduction to the Theory of Probability," Dover, NY, 1962.

Grubbs, F. E., "Statistical Measures of Accuracy for Riflemen and Missile Engineers," 1964.

Gumbel, E. J., and Carlson, P. G. , "Extreme Values in Aeronautics," Journal of the Aeronautical Sciences, Vol. 21, No. 6, pp. 389-398, June, 1954.

Gumbel, E. J., “Statistics of Extremes,” Columbia University Press, NY, 1958.

Hatcher, J. S., “Hatcher’s Notebook,” Stackpole, Harrisburg, PA, 1947.

Herrman, E. E., “Exterior Ballistics,” U.S Naval Institute, Annapolis, MD, 1935.

Hojo, T., “Distribution of the Median, Quartiles and Interquartile Distance in Samples from a Normal Population,” Biometrika, Vol. 23, pp. 315-363, 1931.

Jackson, Dunham, “Note on the Median of a Set of Numbers,” Bulletin of the American Mathematical Society, Vol. 27, pp. 160-164, 1921.

Jacobsen, G. L., “Factors in Accuracy,” The NRA Handloader’s Guide,

p. 145, National Rifle Association of America, Washington, D. C., 1969.

Juran, J. M., “Quality Control Handbook,” 3rd Ed., McGraw, 1974.

King, J. R., “Probability Charts for Decision Making,” Industrial Press, Inc., NY, 1971.

King, J. R., “Frugal Sampling Schemes” , Technical and Engineering Aids for Management, Tamworth, NH, 1980.

Lipschutz, S.,”Probability,” Schaum’s Outlines, McGraw, 1965.

“NRA Handloader’s Guide,” National Rifle Association of America, Washington, D.C., 1969.

Schwartz, M., “Information Transmission, Modulation, and Noise,” McGraw, NY,1959.

Spiegel, M. R., “Theory and Problems of Statistics,” Schaum, NY, 1961.

Urick, R. J., “Principles of Underwater Sound,” 3rd ed., McGraw, 1983.

USNA, Dept. of Ordnance and Gunnery, “Naval Ordnance and Gunnery,” Vol. 2-Fire Control, NavPers 10798-A, USGPO, Wash., DC, 1958.

Weaver, F. L., Naval Engineers Journal, pp. 153-156, July 1983.

Whittaker, E., and Robinson, G., “The Calculus of Observations,” Dover, NY,1967.

Wilks, S. S., “Determination of Sample Sizes for Setting Tolerance Limits,” Annals Math. Stat., Vol. 12, pp. 91-96, (1941).

Wilks, S. S., “Statistical Prediction with Special Reference to the Problem of Tolerance Limits,” Annals Math. Stat., Vol. 13, pp. 400-409 (1942).

INDEX

(numbers refer to chapters)
(also see glossary)

INDEX OF NAMES

(numbers refer to chapters)

(see also notes and references at end of each chapter and at end of book)

SYMBOLS, ABBREVIATIONS, AND ACRONYMS

(also see glossary)

a: alpha, reciprocal of Gumbel slope
CEP: circular error probable
COMB(n,r): combinations of n items, taken r at a time
CPE: circular probable error; meaning same as CEP
e: 2.71828..., base of natural logarithms
E.S. : extreme spread
ETC: electrothermal-chemical
EXP( ): the number e raised to the power within the parentheses ( )
GPS: global positioning system
G: gamma function
INS: inertial navigation system
_: infinity
JDAM: a glide bomb
JSOW: a glide bomb
k: optimal quantity of bins
LG_{10}: common logarithm; logarithm to base 10
LN( ): natural logarithm of quantity enclosed within parentheses
LRIP: low rate initial production
m: number of trials in binomial probability distribution
mil: measure of angles
μ: mu, mean
n: sample size
N: sample size
NBS: National Bureau of Standards (now National Institute of Science and Technology)
N!: N factorial. = $N_*(N-1)_*(N-2)_*..._*2_*1$

%: percent; parts per hundred

PGM : precision guided munition

p: pi, 3.14159...

f: phi, mathematical function, probability density

F: capital phi, cumulative distribution function of probability

P: probability, usually refers to cumulative probability

p: probability; usually a density function

p(R): probability density at radius R

p(x): probability density function of the random variate x

p(x,y): probability density of point x,y

q: probability that an event will not occur; used in the binomial distribution

R: radius

RAN: RANNUM, which see

RANNUM: computer software routine for generating Gaussian (normally) distributed random variates (Gaussian random number generator)

R_c: fraction of given population which is covered by range of sample

R/CEP: ratio of radius to circular probable error

s_N: Factor used by Gumbel to compute standard deviation of sample of size N

s: sigma, standard deviation

_: sum of

_: mathematical integral

u: location parameter for extreme-value probability distribution

x: mathematical variable or random variate

X: measured variable or variate

X_j: jth value of variate X

y: mathematical variable or random variate

Y: measured variable or variate

Y_N: factor used by Gumbel to compute location parameter u

Y_j: jth value of variate Y

x: variable of integration

LMC Biography

Lewis Michael Campbell was born on February 28, 1935 in Mobile, Alabama. His mother Effie was employed as a business office clerk. His father Douglas was a steam boilermaker. His brother “Jim” was six years older. Lewis attended public schools, graduating from Murphy High School in 1952. A few weeks later he enlisted in the U.S. Marine Corps. He was trained as a radar repairman and served at five posts in the U.S. and in Japan. He was released to reserve status in 1955. Returning to Mobile, he worked at television receiver repair and three years as a civilian employee of the U.S. Air Force.

In 1959 he moved to Tuscaloosa, AL and began attending the University of Alabama. He graduated in 1963 with a B.S. in Electrical Engineering. In June of 1963 he took a position at the U.S. Naval Ordnance Laboratory in White Oak, MD. He married Judith Hartley in 1968. They had a daughter, Michelle, and a son, Steven. He retired from the Lab in White Oak in 1992.

He had several hobbies, one of which was reading, and his favorite subjects were philosophy and history. He had enjoyed shooting small arms over the years and made applications of mathematical statistics to ballistic dispersion.

LMC passed away in Bridgewater, VA in 2019.