STATISTICAL MEASURES OF ACCURACY FOR RIFLEMEN AND MISSILE ENGINEERS

by

Frank E. Grubbs, Ph.D.

PREFACE

There has long existed the need for a systematic treatment and analysis of the measures of dispersion, or as sometimes mistakenly called the "accuracy" of patterns of shots, which result from firing rifles at vertical targets, or guns or missiles for ground impact. The problems encountered in such evaluations are essentially statistical in nature and should therefore be so treated. Moreover, due to the rather wide applicability of an appropriate analytical treatment of the measures of dispersion to a large class of firing problems, it is required that the whole subject matter of weapon delivery accuracy should be approached in a fairly elementary manner. Therefore, in this book we have attempted to accomplish these aims in order that many interested in the firing of rifles, guns, missiles or arrows may be able to make the comparisons required for the many and various measures of dispersion and also convert one to the other.

The basic statistical theory for most of the measures of dispersion we will discuss has been developed by various investigators over many years. However, for some of the other measures it was found necessary to develop either new theory or suitable approximations in order to provide this more or less complete treatment discussed herein. For those readers who are interested and so inclined, therefore, we have relegated to an Appendix the related statistical theory with the hope that it will serve as reference material for any future research.

I am indebted to Prof. E. S. Pearson for his permission to publish the means and standard deviations of the univariate range meter (Table 2) and also the mean deviation (Table 3) which are available in the Biometrika Tables for Statisticians (Reference 14). Also, Prof. Pearson, on the behalf of the Biometrika Trustees, and Prof. H. E. Daniels kindly gave their permission to publish the means and standard deviations of the radius of the covering circle (our Table 7), which appeared in Daniels' Biometrika paper (Reference 1).

The permission of the National Academy of Sciences to use some of the moment constants (Table 10) of the bivariate range based on studies of the late Prof. Samuel S. Wilks of Princeton University in connection with the tracking data analysis study (Reference 11) is also appreciated. I am also grateful to Prof. John W. Tukey for granting the permission on behalf of Princeton University to include values from the original Table 10, improved herein in accuracy by Ref. [16].

Ideas for this book came from a learned friend, the late Mr. Philip G. Rust of the Winnstead Plantation, Thomasville, Geogia, who had already investigated the use of the extreme spread over many years and really was the one who stimulated so much of the work behind this book. To Major General Leslie Simon's memory I express my genuine appreciation for his

many helpful suggestions and persistent encouragement, which was so often required in the preparation of the manuscript.

Although the first printing of this booklet in 1964 has seemed to serve its original purposes quite well, there also appeared some need to update some parts of the material and expand it in a few places. We have now included the Figure of Merit (FOM), which has been rather widely used by various small arms experts, and indeed seems very efficient on practical grounds, and to the tables we have added the upper 5% or 95% probability levels, or estimates of it, for the various measures of precision. Finally, due to many requests for confidence bounds on the CEP or Circular Probable Error, we have added in the space on page 50 the 1-α upper confidence bound for the true, unknown or universe CEP. One might note also rearranged examples. Even though exact distributions in probability for all of the bivariate measures of dispersion have been hard to come by, it is hoped that the various approximations used are of sufficient accuracy to deal with the practical aspects of target firings. It is mostly with this hope in mind that we have decided to update the booklet in the newer version for continuing usage.

Frank E. Grubbs

4109 Webster Road
Havre De Grace, Maryland 21078-1621

Second Edition, March 1991

Third Printing, December 1991

<u>Note</u>: All analyses in this book may be carried out in terms of distances on the target, or angular measurements such as mils, degrees or "minutes of angle" (MOA). One mil is equal to 3.375 MOA's.

For any number of rounds or sample sizes beyond our tabulated values, suggest dividing the whole sample into random groups, perhaps in the order fired with group sizes of about ten or so. Then use the average determination for estimation, as this often promotes efficiency. For the range, for example, the best group size is about eight.

TABLE OF CONTENTS

* A confidence bound for the true CEP is given on page 50.

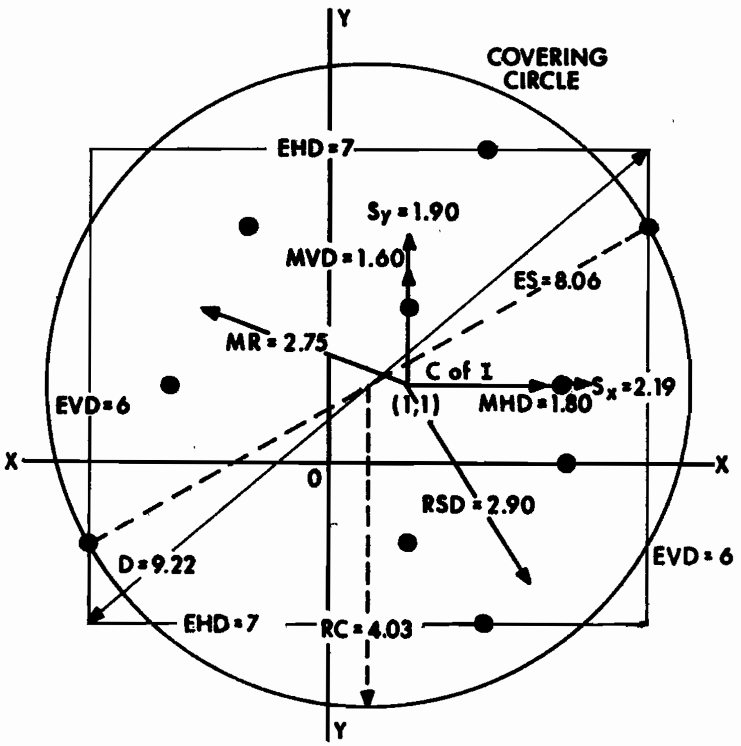
1. INTRODUCTION

In the firing of rifles at vertical targets, or guns or missiles for ground impact, there results a two-dimensional pattern of shots or impact points which exhibit an amount of scatter depending on the round-to-round aiming error and the ordinary ballistic dispersion. The two-dimensional pattern of shots gives rise to various measures of dispersion which are used by riflemen and ballisticians to summarize the "accuracy" of the pattern of shots. The measures of dispersion usually employed consist of the (sample) standard deviation in each direction, the extreme horizontal dispersion, the extreme vertical dispersion, figure of merit the mean horizontal deviation, the mean vertical deviation, the mean radius, the extreme spread, the radial standard deviation, the covering circle, and the "diagonal" of the pattern. All of these measures of dispersion require statistical analysis, since they are really random variables from one firing group to another. In the following, therefore, we will first define each of the above named measures of dispersion and then make a study of their statistical properties, which is necessary to relate one measure to the others and to judge their relative efficiency as estimators of the parameters of accuracy and precision of fire.

A point of considerable importance we record here is that the mean or expected values of all of the various measures of dispersion depend on the sample size or number of rounds, some measures depending very markedly on the sample size. For this reason, it becomes necessary to provide a common basis for comparing one measure of dispersion with another, and this is done by means of the population standard deviation σ (Greek letter "sigma"),which we will define and discuss in great detail. It is with the aid of the large-sample standard deviation, σ, or estimates of it, that we are able to make the proper comparisons of the measures of dispersion or "accuracy". Moreover, in applied work it is generally better to quote the estimated population standard deviation σ rather than the measures of dispersion which depend on sample size, since in this way we may correct for bias, avoid confusion and promote standardization.

2. THE PATTERN OF SHOTS

In Figure 1, we have depicted the impact points or coordinates of ten bullets fired on a vertical target. The horizontal and vertical (or x and y) coordinates of the ten impacts, in the order the bullets were fired from a rifle



S_x	= STANDARD DEVIATION OF X	S_y	= STANDARD DEVIATION OF Y
EHD	= EXTREME HORIZONTAL DISPERSION	EVD	= EXTREME VERTICAL DISPERSION
MHD	= MEAN HORIZONTAL DEVIATION	MVD	= MEAN VERTICAL DEVIATION
RSD	= RADIAL STANDARD DEVIATION	FOM	= FIGURE OF MERIT
MR	= MEAN RADIUS		= (EHD + EVD) / 2 = 6.5
ES	= EXTREME SPREAD		
RC	= RADIUS OF COVERING CIRCLE		
D	= DIAGONAL		

TEN BULLET IMPACTS WITH MEASURES OF DISPERSION

Figure 1.

are: (1,2), (-3,-1), (1,-1), (2,4), (3,0), (4,3), (-1,3), (2,-2), (-2,1), and (3,1).* The mean of the x's or horizontal locations of impacts is $\bar{x} = 1$, and the mean of the y's or vertical positions is $\bar{y} = 1$. Thus, the center of impact (or C of I) of the ten rounds is located at (1,1). Note that the C of I is at the radial distance of $\sqrt{1 + 1} = \sqrt{2} = 1.414$ from the origin, 0, or aimpoint. Later, we will determine whether based on such scatter of impacts, this deviation of the C of I from the aim point is of any significance, or otherwise merely an accidental variation.

If another group of ten rounds were fired at such a target, then the pattern of impact points would be somewhat different and the new C of I of the rounds would be located at a point different from the previous C of I, i. e. (1,1). Thus, the pattern of shots on the target would vary from group to group in a random manner, and moreover the C of I and all of the measures of dispersion we will discuss in the sequel are random variables due also to the inherent dispersion and small sample size. It is our purpose to study just how much variation is to be expected and how the measures of scatter such as the standard deviation, the mean radius, the extreme spread, the mean deviation, the radial standard deviation, etc., relate to each other and compare in efficiency of estimation of the underlying "sigma".

In what follows, we will deal with distances on the target, although of course the distances on the target could be converted to angles, in mils for example, if we know the range to the target. Indeed, for small-arms fire the dispersion in mils is nearly constant as a function of the range to the target, and angular data would therefore represent a more general treatment. Since the conversion from distances on the target to angular data is straightforward and our illustrations are served better by use of bullet impact positions on the target, we will approach the subject by dealing with target impact positions, therefore.

Thus, we are now in good position to discuss the various measures of "accuracy" or dispersion for patterns of shots. We will start with locations and deviations of the projections of impact points on the horizontal and vertical axes, since the (univariate) measures of scatter for such point projections are of basic importance and lead rather naturally to standards of comparison for all of the measures of

*Statistically speaking, these points may be considered to be a random sample of size 10 from a bivariate normal population, with perhaps equal dispersion in both directions. A lot of ammunition could be called a "population", or even a "system" involving a rifle, the ammunition lot and a rifleman reaiming between rounds. Each may add some significant scatter to the impacts, especially the rifleman!

"accuracy" (actually precision*) which are to be discussed. Although the standard errors or standard deviations we will discuss first are somewhat complex computationally, they are nevertheless very efficient and of a fundamental character for our study here.

3. VARIANCES AND STANDARD DEVIATIONS

If the sum of the squares of the deviations of the x coordinates (i. e. projections of points on a horizontal axis) from their mean is divided by n = 10, the number of points, we obtain the sample "variance" of the horizontal or the x coordinates. Thus, the sample variance s_x^2 is given by

$$s_x^2 = \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2 \qquad (1)$$

$$= \frac{1}{10} [(1-1)^2 + (-3-1)^2 + (1-1)^2 + \dots +(3-1)^2] = 4.8.$$

This value may best be calculated generally on a computing machine without the accumulation of rounding errors by means of the formula

$$s_x^2 = \frac{1}{n^2} A_{xx} = \frac{1}{n^2}[n \Sigma x^2 - (\Sigma x)^2] . \qquad (2)$$

Note that the sum and sum of squares may be computed simultaneously on many pocket calculators. Thus, we see that

$$s_x^2 = \frac{1}{100}\{10[1^2+(-3)^2+1^2+2^2+\dots+3^2] - [1-3+1+2+\dots+3]^2\}. \qquad (3)$$

$$= \frac{1}{100}[10(58) - (10)^2] = 480/100 = 4.8.$$

Note: s_x may be computed directly on many pocket calculators.

---

* The term "precision" refers to the dispersion of the bullets about their own mean or C of I, whereas the term "accuracy" includes not only the round-to-round precision but in addition the bias or closeness of the true mean or C of I to the point aimed at on the target. (The term "precision" defined here moreover should not be confused with another term we define later and widely known as the precision of an unbiased estimate.) These concepts will no doubt become much clearer to the reader as we proceed to cover them fully.

The square root of the variance is known as the standard deviation. Thus,

$$s_x = [(1/n)\Sigma(x_i - \bar{x})^2]^{1/2} = \sqrt{4.8} = 2.19. \qquad (4)$$

In a like manner, we find the variance and standard deviation of the y's or projections of the (ten) impact points on the vertical axis. Thus, we have

$$s_y^2 = \frac{1}{n^2} A_{yy} = \frac{1}{n^2} [n\Sigma y_i^2 - (\Sigma y_i)^2] \qquad (5)$$

$$= \frac{1}{100}\{10[2^2 + (-1)^2 + (-1)^2 + \dots + 1^2] - [2-1-1+\dots+1]^2\}$$

$$= 360/100 = 3.6$$

and $s_y = \sqrt{3.6} = 1.90.$

For this sample of 10 rounds, we get $s_x = 2.19$ and $s_y = 1.90$. For another firing of ten such rounds, we might get, for example, $s_x = 3.03$ and $s_y = 2.10$, so that these standard errors s_x and s_y vary randomly from one group to another.

Had there been a very large number of shots or impacts on the target, then the standard deviations, s_x and s_y, would approach their large sample or true values, which we will call σ_x and σ_y, respectively.* Then these true standard deviations, σ_x and σ_y, for many, many thousands of shots are called or are known as the population standard deviations of the randomly varying x and y projections, as compared to the small sample or group values, s_x and s_y. The population values represent the true round-to-round standard deviations in the horizontal and vertical directions for huge lots of ammunition, for example. For samples of ten rounds, or for other small sample sizes, note, as already mentioned, that the sample standard deviations, s_x and s_y, vary from sample to sample of rounds fired at the target. Thus, for small samples or small numbers of shots, s_x varies in a random manner from firing to firing about the true population standard deviation σ_x, and the amount of variation depends on the sample size or number of rounds, n. The quantity σ_x, as seen from its definition, is a measure of the round-to-round variation of individual impact points in the horizon-

* *The reader must be eternally aware! The most important and key parameter σ is hardly ever attained, so that it is "always hidden" !!*

tal direction. In an analogous manner to the description we have given for the variation of an individual x we could express the variation of the standard deviation, s_x, from sample to sample by means of the standard deviation itself of a large number of such values for a fixed sample size. On the average, s_x, for fixed small sample sizes n does not exactly equal the population standard deviation σ_x. That is to say, for small samples s_x is a biased estimate of the population σ_x. The bias is due to the fact that the computation of each s_x involves deviations about the sample mean, not the population mean, so that on the average s_x is somewhat less than σ_x. The amount of bias depends on the sample size n, but approaches zero for large samples of size n.

It is known from statistical theory that the sample variance computed as

$$S_x^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2 = A_{xx}/n(n-1), \quad (S_x^2 = \frac{n}{n-1} s_x^2), \qquad (6)$$

which is based on n-1 "degrees of freedom" (d.f.) - (one degree of freedom being used in the calculation of the sample mean $\bar{x}$ as an estimate of the population mean) - is on the average equal to the population variance σ_x^2, whereas the sample variance computed from $s_x^2 = \Sigma(x_i - \bar{x})^2/n$, and based on the entire sample size n is on the average equal to the quantity $(n-1)\,\sigma_x^2/n$, and is thus biased by $-\sigma_x^2/n$. (It might seem curious to the reader, but it is true that S_x, based on the square root of formula (6), is not an unbiased estimate of σ_x for small sample sizes!)* One of our main interests in this book is that of examining and expressing the measures of dispersion in terms of the population standard deviation σ_x. For a Normal or Gaussian distribution of shots on the target, the mean value of s_x and the standard error of s_x may be calculated as a multiple of the population (or large sample) value of σ_x as indicated in the APPENDIX ON RELATED STATISTICAL THEORY. The mean values and standard deviations of the sample standard deviation depend on the sample size n and are given in Table 1. The 95% probability levels of s are also given. In Table 1 we drop the subscript x from s and σ since the theory covers generally the relationship between the sample and population standard deviations. Also, for most rifle firings we can assume that $\sigma_x = \sigma_y = \sigma$, say.*

*For nearly unbiased estimates of σ, multiply the estimate S_x by $(n-.75)/(n-1)$, or s_x by $(n-.25)/(n-1)$.

Note in Table 1 that the mean values of s (or s_x) do approach the true σ (or σ_x) and hence become unbiased for very large sample sizes. Also, the standard deviations of s decrease in value, becoming more precise for the very largest sample sizes, as would be expected. σ is an "unknown" value.

For rifle firing, as already mentioned, the population standard errors, σ_x and σ_y, for the horizontal and vertical directions, are about equal, with the result that we may use a single value, σ, to represent either of σ_x or σ_y.

The population standard deviation, σ, then is the true or population standard error of an <u>individual</u> shot or bullet in either the horizontal or the vertical direction. The standard deviation of the <u>mean</u> or average of n shots for either the x's or the y's is $\sigma/\sqrt{n}$ and the standard error of either of the sample standard deviations, s_x or s_y, for n rounds is approximately equal to $\sigma/\sqrt{2n}$. Exact values of the standard deviations of s_x and s_y (i.e. s, generally) are given in the fourth column of Table 1. Thus, for the ten shots indicated on the target, the standard deviation of either component of the C of I ($\bar{x}$=1 or $\bar{y}$=1) is $\sigma/\sqrt{10}$ and the standard deviation of s_x or s_y is about $\sigma/\sqrt{20}$.

<u>Example 1</u>. Using the data for the firing represented in Figure 1, find the estimates of σ_x, σ_y, and σ.

Answer: Since n = 10, using Table 1, we find

Estimate of $\sigma_x = s_x/.9227 = 2.19/.9227 = 2.37$
$= 2.19 \times 1.084$ also.

Estimate of $\sigma_y = s_y/.9227 = 2.06 = 1.90 \times 1.084$ also.

To find the estimate of σ, assuming σ_x and σ_y are about equal, we could take the average of 2.37 and 2.06, or better still, since the variances are additive, use

$$[.5(s_x^2 + s_y^2)]^{1/2}/.9227 = \sqrt{4.2}/.9227 = 2.22.$$

<u>Example 2.</u> Find the standard error of the components of the C of I. Find also the standard deviation of either of the

* The case $\sigma_x \neq \sigma_y$ is discussed briefly in the APPENDIX.

TABLE 1

TABLE OF MEANS OR EXPECTED VALUES AND STANDARD DEVIATIONS OF

$s = \sqrt{\Sigma(x_i-\bar{x})^2/n}$ FOR A NORMAL POPULATION

Sample Size	Mean Value of s	Reciprocal of Mean Value Coefficient	Standard Deviation of s	95% Prob Level of s
n	$E(s/\sigma)$ *	$1/E(s/\sigma)$	$SD(s/\sigma)$ **	$s_{.95}/\sigma$
2	.5642	1.772	.4263	1.39
3	.7236	1.382	.3782	1.41
4	.7979	1.253	.3367	1.40
5	.8407	1.189	.3052	1.38
6	.8686	1.151	.2808	1.36
7	.8882	1.126	.2612	1.34
8	.9027	1.108	.2452	1.33
9	.9139	1.094	.2318	1.31
10	.9227	1.084	.2203	1.30
11	.9300	1.075	.2104	1.29
12	.9359	1.068	.2017	1.28
13	.9410	1.063	.1940	1.27
14	.9453	1.058	.1871	1.26
15	.9490	1.054	.1809	1.26
16	.9523	1.050	.1753	1.25
17	.9551	1.047	.1701	1.24
18	.9576	1.044	.1654	1.24
19	.9599	1.042	.1611	1.23
20	.9619	1.040	.1570	1.23

Note: * $E(s/\sigma)$ = Expected value of. ** SD = Stand. Dev.

The mean values and the standard deviations in this table and the following tables are in terms of a population standard deviation of unity. Hence, all tabular entries in the second and forth columns are to be multiplied by the population standard deviation σ or an estimate of it. For the third column, if we let the expected value of s be $E(s) = c\sigma$, say, then the reciprocal of the mean value coefficient is $1/c$, and the entry in the third column when multiplied by s results in an unbiased estimate of σ. [$c = E(s/\sigma)$]

Standard deviations of sample statistics or variables given in this and the following tables are calculated about their own expected values, and hence not about unbiased estimates of the parameter unless so indicated.

The fifth column gives the 95% probability level or the upper 5% point of the distribution of s for a test of significance, if so desired. Check s over stated σ with table.

sample standard deviations, s_x and s_y.

Answer: Now $\bar{x} = \bar{y} = 1$. The standard deviation of the components $\bar{x}$ and $\bar{y}$ of the C of I are then given by $\sigma_{\bar{x}} = \sigma_{\bar{y}} = \sigma/\sqrt{n} = \sigma/\sqrt{10} = 2.22/\sqrt{10} = .70$. (The standard deviation of an average is equal to the standard deviation of an individual observation divided by the square root of the sample size.) The standard deviation of either of s_x or s_y is approximately given by the quantity $\sigma_s = \sigma/\sqrt{2n} = 2.22/\sqrt{20} = .50$. The exact value from our Table 1 is .2203(2.22) = .49 in this particular case.

We begin to see then that the population or large-sample standard deviation, σ, for an individual shot is the key parameter for the study of dispersion and accuracy. As a matter of fact, σ, the standard error of an individual, is the real basis or standard of comparison for all the measures of precision and "accuracy". Indeed, as we will see, the average values of each of the various measures of "accuracy" turn out to be multiples of σ. Similarly, standard deviations of the measures of dispersion or "accuracy" turn out to be fractions of σ also.

With the definition of and an appreciation for the importance of the population standard deviation σ, and also a recognition of its usefulness, we are, therefore, now ready to proceed with the analysis of the other measures of dispersion and "accuracy". First, however, we will discuss the two dimensional measure of dispersion called the Circular Probable Error (CEP or CPE), and the one directional measure of dispersion or precision called the Probable Error (PE).

4. THE CIRCULAR PROBABLE ERROR (CEP or CPE) AND THE PROBABLE ERROR (PE)

A measure of dispersion (precision) which is widely used for firings at targets is the Circular Probable Error, which is designated by CEP or CPE. The CEP is defined as the radius of the circle about the (true) center of impact (C of I) of the rounds, or sometimes about the point of aim, which, includes one-half or 50% of the shots fired upon the target. This circle has a radius of 1.1774σ = the CEP, and is therefore the 50 percent probability circle. Thus, the standard deviation σ is also quite basic to the determination of the CEP, and moreover an efficient estimator of the quantity we designate as σ will likewise give an efficient estimator of the CEP. **(The Spherical Probable Error or SEP = 1.5382σ)**

If all of the shots are projected on the x-axis (or the y-axis), the interval about both sides of the mean which includes 50% of the shots is called the Probable Error or simply the PE. That is, the interval from the true but unknown mean minus the PE to the average plus one PE contains 50% of the shots in the x (or in the y) direction, considering a very large number of shots. The Probable Error is actually $.6745\sigma$, i. e. PE = $.6745\sigma$. The PE is a one-directional or a univariate measure of dispersion, whereas the CEP is a two-directional or bivariate measure of precision.*

For the two-dimensional case and also for unequal standard deviations, or $\sigma_x \neq \sigma_y$, in the x and y directions, then Grubbs [6] gives a good approximate formula for the CEP that is of sufficient accuracy for most practical cases. (See the APPENDIX.) A confidence bound for the CEP appears of page 50.

5. THE EXTREME HORIZONTAL DISPERSION (EHD) AND THE EXTREME VERTICAL DISPERSION (EVD) [The Univariate Range (R)]

These are very simple measures of dispersion and by far the easiest to compute. If we project the impact points onto the horizontal (x) and the vertical (y) axes, then we see that the EHD is simply the difference between the greatest and the least values of the x points, or farthest right minus farthest left projected points, and the EVD is given by the highest minus the least values of y projected points. In this connection, we see from Figure 1, the farthest x point to the right is x = 4, i. e. the point (4, 3) and the leftmost point or value of x is x = -3, i. e. the value of x for the point (-3, -1). Therefore, the EHD = 4 - (-3) = 7. For the EVD, we see in a like manner, the maximum variation, the range or maximum dispersion for y occurs for the two y points (2, 4) and (2, -2). Thus, the EVD = 4 - (-2) = 6. By simply arranging the x's (or the y's) in increasing order, then one sees that the formula for the range or the maximum dispersion (or the EVD or EHD) is

$$R = x_n - x_1, \quad \text{where } x_1 \leq x_2 \leq x_3 \leq \cdots \leq x_n. \qquad (7)$$

The EHD and the EVD are thus terms of the rifleman which are widely known statistically as the "maximum dispersion", the "maxumum variation" or the "range" of the observations. Note in particular that the EHD, EVD, range, etc., are univariate or one-directional measures of the scatter of shots.

The probability distribution of the range, and hence the EHD or EVD, has been the subject of very extensive study by Dederick [2], Hartley [7], Pearson [13], and Tippett [17]. The sample range (EHD) for small samples such as for example n = 5 or n = 10 is a random variable, as was true for the s_x sample statistic discussed above in Section 4. We point out that the amount of variation in the EHD from any one sample to another depends, of course, on the sample size and clearly also on the value of the population sigma, σ. In Table 2 we give the mean or expected values and the standard deviations of the range, i. e. the EHD or EVD, as a multiple of σ our unknown population standard deviation. The 95% or that is the upper 5% probability levels of the range also appear in Table 2. An example would be instructive for the range.

** Why don't riflemen use the CEP? Or quote the unbiased estimate of sigma? Learn that σ measures round-to-round variation-imprecision-but accuracy includes also offset of aim point! Study pages 43-45.*

Example 3. By use of the computed values of s_x and s_y, predict the EHD and the EVD and compare with observed values of the quantities.

Answer: The estimate of σ_x is $s_x/.9227 = 2.19/.9227 = 2.37$. Then for a sample of size 10, we find from Table 2 that the expected value of the EHD would be $3.078\, s_x = (3.078)(2.37) = 7.3$, as compared to the actual observed value of 7 for the EHD, showing acceptable prediction.

In a like manner, $3.078 s_y/.9227 = (3.078)(1.90)/.9227 = 6.3$, as compared with the observed value of 6 for the EVD.

Actually, instead of using s_x and s_y individually, it is seen that we may just as well, under the assumption and demonstration of a circular distribution, or $\sigma_x = \sigma_y = \sigma$, adopt the estimate of $\sigma = 2.22$ of Example 1 and multiply it by the factor 3.078 for 10 rounds, giving 6.8 for the estimate of either of EHD or EVD, as compared to the observed values of 7 and 6, respectively. Such differences are expected and as a matter of fact attributable to random variations or fluctuations for the small sample size of 10 used here.

Example 4. Estimate the population σ by using a rather popular measure of precision or scatter of the shot known widely as the "Figure of Merit" (FOM), which is simply the average of the EHD and the EVD, i. e.

$$\text{FOM} = (\text{EHD} + \text{EVD})/2 = (R_x + R_y)/2 \;. * \qquad (7a)$$

Find also the standard error of this estimate.

Answer: Now the FOM or the average of the EHD and EVD is (7 + 6)/2 = 6.5. From Table 2 for the given sample size of 10 rounds, the mean value of the range is 3.078 times σ. Therefore, the estimate of σ is 6.5/3.078 (or .3249 x 6.5) = 2.11. The estimated standard error for this unbiased estimate is given by $k_n\sigma/d_n\sqrt{2} = (.7971) \times (2.11)/[(3.078)(1.414)] = .39$. The $\sqrt{2}$ used here in the denominator comes from the fact that we are now dealing with the average of two ranges, i. e. the standard deviation of the average of two individual observations

* See footnote, page 20

TABLE 2

TABLE OF MEAN VALUES AND STANDARD DEVIATIONS OF THE RANGE R = EHD or EVD

Sample Size	Mean Value of the Range	Reciprocal of mean value coefficient	Standard Deviation	95% Prob Level of the Range
n	$E(R/\sigma) = d_n$	$1/d_n$	$SD(R/\sigma) = k_n$	$R_{.95}/\sigma$
2	1.128*	.8862	.8525	2.77
3	1.693	.5908	.8884	3.31
4	2.059	.4857	.8798	3.63
5	2.326	.4299	.8641	3.86
6	2.534	.3946	.8480	4.03
7	2.704	.3698	.8332	4.17
8	2.847	.3512	.8198	4.29
9	2.970	.3367	.8078	4.39
10	3.078	.3249	.7971	4.47
11	3.173	.3152	.7873	4.55
12	3.258	.3069	.7785	4.62
13	3.336	.2998	.7704	4.68
14	3.407	.2935	.7630	4.74
15	3.472	.2880	.7562	4.80
16	3.532	.2831	.7499	4.85
17	3.588	.2787	.7441	4.89
18	3.640	.2747	.7386	4.93
19	3.689	.2711	.7335	4.97
20	3.735*	.2677	.7287	5.01

* Note that the mean or expected values of the range show that it is very sensitive to sample size, the value for n = 20 being over three times that for n = 2.

Tabular entries in the second, fourth and fifth columns are to be multiplied by the population standard deviation or an estimate of it. The values here are all reproduced from the Biometrika Tables [14] with permission of Prof. E. S. Pearson.

---

is $\sigma/\sqrt{2}$, and the standard deviation of the average of two ranges is $k_n\sigma/\sqrt{2}$. The computed figure of .39 indicates that this estimate of sigma is subject to this particular standard error, whereas the standard deviation of an individual observation is much larger, being 2.22.

6. THE MEAN HORIZONTAL DEVIATION (MHD) AND THE MEAN VERTICAL DEVIATION (MVD) (The Mean Deviation)

These measures of dispersion are also known as the mean deviation from the mean and the mean absolute deviation. The mean horizontal deviation (MHD) is defined as the average of the unsigned or absolute (positive) deviations from the sample mean of the x components. That is to say the MHD is

$$MHD = (1/10)[|1-1| + |-3-1|^* + |1-1| + \ldots + |3-1|] = 1.8.$$

In a like manner, the mean vertical deviation (MVD) is the average of the unsigned deviations of the y's measured from their own sample average. Thus, for the y's we have that

$$MVD = (1/10)(1 + 2 + 2 + \ldots + 0) = 16/10 = 1.6.$$

The algebraic formula for the mean deviation (MD) is

$$MD = \sum_{i=1}^{n} |x_i - \bar{x}|/n . \qquad (8)$$

The sample mean horizontal and mean vertical deviations are rather easy to calculate as compared to the sample standard deviation, and moreover, the mean deviation is nearly as efficient as the standard deviation of the sample in estimating the population standard deviation, σ. The mean deviation has been investigated by Godwin [4]. In Table 3, we give the expected or mean values of the sample mean deviation and also the standard errors of the MD. The 95% probability levels of the mean deviation are also included.

Example 5. Estimate σ from the sample MHD and MVD. How precise is the estimate?

Answer: Since the dispersions in the x and y directions are about equal, we may as well use the average of the MHD and MVD to gain precision, i. e. the estimator of σ is taken as $\sigma = (1/2)(MHD + MVD)/(.7569) = 1.7/(.7569) = 2.25$. Note how this compares with the estimate 2.22, a quantity which was estimated by using the more efficient estimators, s_x and s_y to determine σ. The precision, or standard error, of the unbiased estimate of σ based then on the average of the MHD and MVD is $(.1894/.7569\sqrt{2})\sigma$ or $(.1894)(2.25)/(.7569)(1.414) = 0.40$, or hence about the same precision as we obtained for the FOM, or Figure of Merit for the sample of size 10.

Example 6. What is the relation between the MHD and the EHD on the average? Hence, predict the size of the EHD from the MHD for say, 15 rounds, thereby showing a further use of the tables.

---

* |-3-1| means positive value of, which is 4, etc.

TABLE 3

TABLE OF EXPECTED VALUES AND STANDARD DEVIATIONS OF THE MEAN HORIZONTAL DEVIATION (MHD) AND MEAN VERTICAL DEVIATION (MVD)

[The Mean Deviation (MD)]

Sample Size n	Mean Value of MD $E(MD/\sigma)$	Reciprocal of Mean Value Coefficient $1/E(MD/\sigma)$	Standard Deviation of MD $\sigma(MD/\sigma)$	95% Prob Level of MD/σ *
2	.5642	1.772	.4263	1.39
3	.6515	1.535	.3419	1.28
4	.6910	1.447	.2970	1.22
5	.7137	1.401	.2663	1.19
6	.7284	1.373	.2436	1.16
7	.7387	1.354	.2258	1.14
8	.7464	1.340	.2115	1.12
9	.7523	1.329	.1996	1.10
10	.7569	1.321	.1894	1.09
11	.7608	1.314	.1807	1.07
12	.7639	1.309	.1731	1.06
13	.7666	1.304	.1664	1.05
14	.7689	1.301	.1604	1.04
15	.7708	1.297	.1550	1.04
16	.7725	1.294	.1501	1.03
17	.7741	1.292	.1457	1.02
18	.7754	1.290	.1416	1.02
19	.7766	1.288	.1378	1.01
20	.7777	1.286	.1344	1.01

(Tabular entries in the second, fourth and fifth columns are to be multiplied by the population sigma. All tabular values given above are reproduced from the Biometrika Tables [14] with the permission of Prof. E. S. Pearson)

* To illustrate a use of the 95% probability level values of a sample statistic, suppose that we take as the estimate of sigma the value of 2.22 we obtained by using the average of the variances in the two directions as in Example 1. Then we now suppose that in a further shooting of, say, 12 rounds we obtained an MVD of 2.00. Could we accept the value of 2.22 as the population sigma of the new target firing?

The answer is yes, since for 12 rounds the 95% probability level estimated for the new firing is 1.06 x 2.22 = 2.35, a value larger than expected of the random MVD = 2.00.

Answer: From Tables 2 and 3 for 10 rounds, the ratio of expected values of EHD and MHD is 3.078/0.7569 = 4.07. Hence, using the computed value of MHD = 1.8, we multiply this by 4.07 and get 7.32, as compared to an EHD of 7 which was observed for the original data of Figure 1. For n = 15 rounds, we would predict: EHD = 3.472(sigma) = 3.472(1.8/.7569) = 8.3, a larger value, of course.

<u>Example 7.</u> For a sample of size 15, what is the relationship between the average values of the sample mean deviation and the sample standard deviation?

Answer: From Tables 1 and 3, we see for n = 15 that the ratio of the mean value of the sample mean deviation to that of the standard deviation is .7708/.9490 = .812, or on the average the mean deviation (MD) is 18.8% smaller than s, the sample standard deviation.

The measures of dispersion discussed so far, except the CEP, are for either the x or the y direction separately, and hence they are one-directional or univariate measures of the dispersion or scatter of shots, as we have previously indicated. We now turn to measures of dispersion of the impacts which take into account both the x and y directions simultaneously. The measures of dispersion or precision which involve both the horizontal and vertical directions in a single estimate of the population σ are known as bivariate or two-directional sample measures or statistics. As will be seen in the sequel, the bivariate measures are more precise than the one-directional or univariate values (sample statistics) in estimating the unknown σ, no doubt as would logically be expected since the sample size in effect may be considered to be "doubled", especially due to equal amounts of scatter in the two directions. Thus, we see rather easily that the bivariate estimates are far more "efficient".

For the two-directional measures, we will discuss first the Radial Standard Deviation (RSD), which involves the sum of the sample variances in the x and the y directions, this turning out to be the most efficient estimator of sigma.

7. THE RADIAL STANDARD DEVIATION (RSD)

The Radial Standard Deviation or RSD is defined for our purposes here as the square root of the total sum of squares of the deviations in each of the x and y directions from the respective sample means, divided by n, the number of impacts or points. We see therefore that the RSD is really given by the square root of the sum of the sample variances in x and y, i. e. s_x^2 and s_y^2.

Thus, the formula for the radial standard deviation is

$$RSD = \{(1/n)[\Sigma(x_i-\bar{x})^2 + \Sigma(y_i-\bar{y})^2]\}^{1/2} = (s_x^2 + s_y^2)^{1/2}. \quad (9)$$

Since $s_x^2 = 4.8$ and $s_y^2 = 3.6$ for the 10 shots on the Figure 1 target, then it can be seen that the RSD is calculated as

$$RSD = \sqrt{4.8 + 3.6} = \sqrt{8.4} = 2.90.$$

It should be noted that since we take the square root of the sum of s_x^2 and s_y^2, then the RSD should on the average be expected to be about $\sqrt{2}$ times the standard deviation of the points of impact in either the horizontal or vertical direction. As previously indicated, we see now that the radial standard deviation takes into account all of the information on dispersion (observations) in both directions.

The radial standard deviation has been studied by Grubbs [5], and the first two moments, or the mean and standard deviation, are given in Table 4, along with the 95% probability levels of the RSD distribution.

We should keep in mind that the RSD is the most efficient of the estimators of the population sigma we could use concerning the analyses of target dispersion, but it is also somewhat more involved. (This makes little difference with the modern-day pocket calculator).

As a point of some practical interest, we record at this time that the probability distribution of the radial variance, i. e. the square of the RSD, is theoretically well established, and it is therefore possible to compare the dispersion patterns of two targets or two riflemen in a rather simple manner. The reader is referred to the Appendix concerning this.

<u>Example 8.</u> For the 10 rounds fired at the target, find the estimate of sigma and then use it to predict the size of the observed extreme horizontal dispersion (EHD) or the extreme vertical dispersion (EVD).

Answer: From Table 4 for n = 10 rounds, the estimate of σ is given by RSD/1.323 = RSD(.7559) = 2.90/1.323 = 2.19 as compared to the value 2.22 previously found in Example 1. (Slight differences may be expected for the two different approaches.) The prediction of the EHD or a EVD is then 3.078 x 2.19 = 6.7 versus the 7 for EHD that observed or the 6 for the EVD. The number 3.078 may be found in Table 2 for n = 10.

<u>Example 9.</u> Since the sample radial standard deviation takes into account both the dispersions in the horizontal and the vertical directions, is it not more efficient than either of the univariate sample standard deviations, s_x or s_y, assuming, of course, that $\sigma_x = \sigma_y = \sigma$?

Answer: We define efficiency here as the ratio of the variance of the best unbiased estimator (the RSD) to the variance of any other unbiased estimator. Since the variance is the square of the standard deviation, then we see using Table 4 for a sample size of 10, or any other sample size in fact, that the variance of a unbiased estimate of σ based on the RSD may be found by simply taking the square of the ratio of the standard deviation of the sample statistic to its mean value. This means that the definition of precision of any unbiased estimate is actually the coefficient of variation (times the population σ). As a general definition of "precision", then, we will take it as being the ratio of the standard error of any sample statistic or estimator to its mean value, this ratio being finally multiplied by the population σ. (Others often define **precision** differently, but we prefer here to use only first powers for simplicity and not either squares or reciprocal of squares, as we will soon see in the sequel.) Thus, for the 10 rounds fired upon the target and from Table 1, for example, the estimator $s_x/.9227$ gives an unbiased estimate of σ, and the ratio .2203/.9227 = .239, when multiplied by the population σ, will be referred to as the "precision", since it is the standard error of the unbiased estimator using the sample standard deviation in either the x or y direction alone. In a like manner, the precision of the RSD may be found from Table 4 for 10 rounds as being .2219/1.323 = .168 (times σ). Therefore, the .239 versus the .168 indicates that the RSD is considerably more precise than a one-directional estimator of σ such as the sample standard deviation s_x (or s_y).

Finally, the <u>efficiency</u> of s_x (or s_y) as compared to the RSD is the square of the ratio .168/.239 = .70, this turning out to be .49, or about 50%, say. That is, the univariate s_x (or s_y) is only 50% as efficient in estimating the population sigma as is the bivariate measure RSD. (More will be discussed on this subject later.)

<u>Example 10.</u> Estimate the CEP by using the "efficient" RSD.

Answer: The estimate of σ based on the RSD is 2.19 from Example 8. Therefore, the estimate of the CEP = 1.1774σ = (1.1774)(2.19) = 2.58. A circle of this radius about the center of impact (C of I) will include approximately 50% of the shots. See p. 50 for a confidence bound concerning the CEP.

TABLE 4

TABLE OF MEAN VALUES AND STANDARD DEVIATIONS OF THE RADIAL STANDARD DEVIATION (RSD)

Sample Size n	Mean Value of RSD $E(RSD/\sigma)$	Reciprocal of Mean Value Coefficient $1/E(RSD/\sigma)$	Standard Deviation of RSD $\sigma(RSD/\sigma)$	95% Prob Level $RSD_{.95}/\sigma$
2	.8862	1.128	.4633	1.73
3	1.085	.9217	.3940	1.78
4	1.175	.8511	.3455	1.77
5	1.226	.8157	.3108	1.76
6	1.259	.7943	.2848	1.75
7	1.282	.7800	.2643	1.73
8	1.300	.7692	.2478	1.72
9	1.313	.7616	.2338	1.71
10	1.323	.7559	.2219	1.70
11	1.332	.7508	.2118	1.69
12	1.339	.7468	.2028	1.68
13	1.345	.7435	.1949	1.67
14	1.350	.7407	.1881	1.67
15	1.354	.7386	.1817	1.66
16	1.358	.7364	.1760	1.65
17	1.361	.7348	.1708	1.65
18	1.364	.7331	.1660	1.64
19	1.367	.7315	.1617	1.64
20	1.369	.7305	.1576	1.63

(Tabular entries in the second, fourth and fifth columns are to be multiplied by the population σ, or an estimate of it.)

8. THE MEAN RADIUS (MR) *

To compute the Mean Radius (MR), we find merely the average of the radial distances between the observed center of impact (C of I) of the rounds and all of the impact points on the target. Since the C of I for the 10 rounds is located at (1, 1) and the points of impact in the order of firing are (1, 2), (-3, -1), (1, -1), (2, 4), (3, 0), (4, 3), (-1, 3), (2, -2), (-2, 1) and (3, 1), it is easily seen that the radial distances could be measured directly and swiftly on the target, and are in fact 1, $\sqrt{20} = 4.47$, 2, $\sqrt{10} = 3.16$, $\sqrt{5} = 2.24$, $\sqrt{13} = 3.61$, $\sqrt{8}, = 2.83$, $\sqrt{10} = 3.16$, 3 and 2. For example, the distance between the C of I (1, 1) and the particular point (2, -2) is $\sqrt{(1-2)^2 + [1-(-2)]^2} = \sqrt{10} = 3.16$.

* See formula (27), p. 35

The average of these ten radial distances is 2.75, which is, of course, the mean radius, MR. We remark that each of the radii are easily measured with a ruler on the target. Moreover, as we shall see, the mean radius is very efficient for estimating the population sigma.

The mean values and the standard deviations of the mean radius are derived in the Appendix, and the computed values for samples of size 2 through 20 are given in Table 5. They depend, of course, on the number of rounds, n, and the population σ. For very large sample sizes, it can be shown that the mean values of the MR will approach 1.253σ. For $n = 15$, the expected value of the MR is already 1.211σ, i. e. is off by only about 3% from the large sample value.

Example 11. The mean radius (MR) of the 10 shots which were fired at the target is 2.75. The estimate of σ based on the mean radius is therefore, using Table 5 for $n = 10$, given by a $\hat{\sigma} = (.841)(2.75) = 2.31$, and this compares with the value of 2.19 obtained by using the RSD.

Example 12. What is the relative efficiency of the estimate of σ based on the MR in Example 11?

Answer: From Table 5 for the MR, the standard error of the unbiased estimate for $n = 10$ is $.2063/1.189 = 0.174$, whereas from Table 4 for the RSD, the most efficient estimator, the equivalent standard error is $.2219/1.323$ or 0.168. Hence, the efficiency of the MR is $(.168/.174)^2 = 0.94$ or 94%, which is very good indeed.

9. THE EXTREME SPREAD (ES) OR THE BIVARIATE RANGE

The Extreme Spread, known as ES, or the bivariate range, is defined as the maximum of the distances between all possible pairs of points or shots on the target. Note that in the figure the pairs of impact points giving rise to the extreme spread are the two shots located at the coordinates of (-3, -1) and (4, 3). The numerical value of the ES is hence the quantity

$$ES = \sqrt{[4-(-3)]^2 + [3-(-1)]^2} = \sqrt{7^2 + 4^2} = \sqrt{65} = 8.06.$$

The extreme spread, it should be noted, is very easy to measure with a ruler on the target and it is also rather easy to compute. Indeed, the ES provides a very rapid measure of an estimate of dispersion for the two-dimensional scatter diagram, when divided by the appropriate constant for the sample size used. The extreme spread, like the other measures of dispersion, is also a random variable, i. e. it varies in a haphazard manner from one group of shots to another. The

TABLE 5

TABLE OF MEAN VALUES, STANDARD DEVIATIONS AND THE 95% PROBABILITY LEVELS OF THE MEAN RADIUS (MR) *

Sample Size n	Mean Value of MR $E(MR/\sigma)$	Reciprocal of Mean Value $1/E(MR/\sigma)$	Standard Deviation $SD(MR/\sigma)$	95% Prob Level of MR $MR_{.95}/\sigma$
2	.8862	1.128	.4632	1.73
3	1.023	.9775	.3738	1.68
4	1.085	.9217	.3243	1.65
5	1.121	.8921	.2906	1.62
6	1.144	.8740	.2656	1.60
7	1.160	.8621	.2461	1.58
8	1.172	.8532	.2304	1.56
9	1.182	.8460	.2174	1.55
10	1.189	.8410	.2063	1.54
11	1.195	.8368	.1968	1.53
12	1.200	.8333	.1885	1.52
13	1.204	.8306	.1811	1.51
14	1.208	.8278	.1746	1.50
15	1.211	.8258	.1686	1.50
16	1.214	.8237	.1633	1.49
17	1.216	.8224	.1585	1.48
18	1.218	.8210	.1541	1.48
19	1.220	.8197	.1500	1.47
20	1.222	.8183	.1462	1.47

(Tabular entries in the second, fourth and fifth columns are to be multiplied by the population σ, or an estimate of it. These values were calculated from the theory covered in the Appendix for the estimation of the mean values and standard deviations of the MR as a function of the sample size n. The 95% probability levels for the MR were determined from a two moment fit of a Chi variate due to Patnaik in Reference [12]. This two moment fitting procedure is explained in the Appendix, Section H, for the approximate theory concerning which we have used for the Diagonal - See also Section 11 below.)

Note: The mean radius MR is a measure (the average) of radial distances. There is also the so-called "radial error", a measure based on the square root of sums of squares in the x and y directions, and is described, for example, by Weil in Reference [18].

* We understand the British and some of its Commonwealth riflemen define our MR as their "Figure of Merit". We like the more descriptive "Mean Radius".

ES should not be confused with either the diameter (=2RC) of the "covering circle" in Section 10 or the "diagonal" (D) of the pattern of shots discussed in Section 11 in the sequel.

In Table 6, we give the mean values, their reciprocals, the standard deviations and the 95% probability level values of the extreme spread. The mean values and the standard errors were originally computed by the late Prof. Samuel Stanley Wilks of Princeton University in connection with a study of the Panel on Tracking Data Analysis, Reference [11]. For the original calculations, a Monte Carlo sampling procedure was used to determine the means and standard deviations for the ES. These particular calculations were programmed on an IBM 7090 computer using the Fortran language by Prof. Wilks and Mr. Paul Raynault of Princeton University. In 1975, an improved Monte Carlo experiment was conducted by Taylor and Grubbs [16] using much larger sample sizes for establishing both the moments and the percentage points of the distribution. Note that the mean value of the ES is very sensitive to the sample size.

Example 13. By using the extreme spread, ES, find the estimate of the normal population sigma which is unbiased.

Answer: The unbiased estimate of σ = (.262)(8.06) = 2.1

Example 14. What is the standard error of the estimate that we have calculated in Example 13?

Answer: In Example 9 we learned how to compute the sigma of an unbiased estimate. Using this procedure, it is easy to find from Table 6 using a sample size of 10 that

$$\text{Standard error of } ES/3.813 = (.745/3.813)(2.1) = 0.41$$

Example 15. What is the efficiency of the extreme spread for 10 rounds?

Answer: As computed previously in Example 9 for the sample standard deviation, we have that the efficiency of the extreme spread for n = 10 rounds is given by

$$(.2219/1.323)^2 \div (.745/3.813)^2 = .74 \text{ or } 74\%,$$

where we have used Tables 6 and 4.

10. THE RADIUS OF THE COVERING CIRCLE (RC or RCC)

The Covering Circle is defined as the smallest circle of all such circles which contains on it or inside it each and every point of impact. The radius of the covering circle or

TABLE 6

TABLE OF MEAN VALUES, STANDARD DEVIATIONS AND THE 95% PROBABILITY LEVELS OF THE EXTREME SPREAD (ES)

Sample Size	Mean Value of ES	Reciprocal of Mean Value	Standard Deviation	95% Prob Level of ES
n	$E(ES/\sigma)$	$1/E(ES/\sigma)$	$SD(ES/\sigma)$	$ES_{.95}/\sigma$
2	1.772	.564	.932	3.462
3	2.406	.416	.887	3.984
4	2.787	.359	.856	4.285
5	3.066	.326	.828	4.519
6	3.277	.305	.806	4.670
7	3.443	.291	.783	4.805
8	3.582	.279	.771	4.937
9	3.710	.270	.754	5.029
10	3.813	.262	.745	5.118
11	3.888	.257	.735	5.174
12	3.964	.252	.725	5.229
13	4.039	.248	.714	5.285
14	4.115	.243	.704	5.340
15	4.190	.239	.694	5.396
16	4.242	.236	.689	5.443
17	4.295	.233	.684	5.490
18	4.347	.230	.678	5.536
19	4.399	.227	.673	5.583
20	4.452	.225	.668	5.630
25	4.639	.216	.650	5.790

(Tabular entries in the second, fourth and fifth columns are to be multiplied by the population σ, or an estimate of it. Since these are Monte Carlo values, the third decimal places may be in error.)

---

its diameter provides a fairly rapid measure of the dispersion of the shots on the target. The diameter of the covering Circle is not generally the same as the Extreme Spread or the bivariate range, it should be noted. Daniels [1] has made a study of the radius of the Covering Circle, and so we give in our Table 7 the key moment constants which are taken from Daniels' Biometrika paper [1]. Letting RCC (or RC) denote the radius of the Covering Circle, then expected values and also standard deviations of the RCC are given in Table 7 along with the 95% probability levels of the distribution.

The RCC is also a random variable, as it varies from one group of shots to another in a random manner, depending very much on the size of the population σ, of course.

TABLE 7

TABLE OF MEAN VALUES, STANDARD DEVIATIONS AND THE 95% PROBABILITY LEVELS OF THE RADIUS OF THE COVERING CIRCLE (RC)

Sample Size	Mean Value of RC	Reciprocal of Mean Value	Standard Deviation	95% Prob Level of RC
n	$E(RC/\sigma)$	$1/E(RC/\sigma)$	$SD(RC/\sigma)$	$RC_{.95}/\sigma$
2	.8862	1.128	.4632	1.731
3	1.211	.8258	.4461	2.000
4	1.409	.7097	.4274	2.157
5	1.548	.6460	.4123	2.268
6	1.655	.6042	.4001	2.352
7	1.742	.5741	.3901	2.421
8	1.814	.5513	.3816	2.478
9	1.876	.5330	.3743	2.527
10	1.929	.5184	.3680	2.570
11	1.977	.5058	.3625	2.608
12	2.020	.4950	.3575	2.642
13	2.058	.4859	.3528	2.673
14	2.093	.4778	.3484	2.702
15	2.125	.4706	.3440	2.728
16	2.153	.4645	.3411	2.749
17	2.180	.4587	.3384	2.770
18	2.206	.4533	.3357	2.791
19	2.231	.4482	.3333	2.812
20	2.255	.4435	.3309	2.833
30	2.427	.4120	.3126	2.975
40	2.543	.3932	.3010	3.071
50	2.629	.3804	.2924	3.143
100	2.881	.3471	.2704	3.358

(Tabular entries in the second, fourth and fifth columns are to be multiplied by the population σ or an estimate of that σ. The values of Table 7 were obtained from Daniels' Biometrika paper [1] with permission, and entries for n = 16 - 19 were obtained by interpolation. For interested readers, the diameter of the Covering Circle, DC or DCC = 2 RCC, so that the mean values of the DCC would be double that we give for RC in the second column, the standard deviations of the DCC would be two times the figures in column four, and hence also the 95% probability levels double those listed in the fifth column above.)

The Covering Circle is particularly useful for rifle firings at relatively short ranges or matches when the individual impacts are not discernible on the target, but rather the result of the firing is a big hole with all of the shots having gone through it. It should be mentioned in this connection that the diagonal (see below) also could be used and is apparently more efficient in estimating σ.

Example 16. Use the radius of the Covering Circle to determine an unbiased estimate of σ.

Answer: For the group of 10 shots on Figure 1, it turns out that the diameter of the covering circle is equal to the extreme spread, so that the radius of the covering circle is thus 8.06/2 = 4.03. (For all firings it cannot be expected that the extreme spread will equal the diameter of the covering circle.) Anyway, using Table 7 for a sample size of n = 10, we get

$$\sigma = .5184(4.03) = 2.09,$$

as compared to the value 2.19 estimated with the RSD.

Example 17. Is the radius of the covering circle much more efficient in estimating σ than the extreme spread for as few as n = 10 rounds?

Answer: The "precision" of the extreme spread turns out to be .745/3.813 = .195, and that likewise of the radius of the covering circle is .3680/1.929 = .191. Therefore, the efficiency of the extreme spread relative to that of the covering circle radius is

$$(.191/.195)^2 = .96 \text{ or } 96\%.$$

Thus, the extreme spread is almost as efficient for n = 10.

11. THE DIAGONAL (D)

The rectangle, with sides parallel to the x and y axes, which is determined by the extreme horizontal dispersion or variation (EHD) and the extreme vertical dispersion or range (EVD), and includes all the shots or impacts on the boundary or within such a rectangle, is also used to estimate the dispersion of the shots on the target. In fact, we may use the diagonal D of this rectangle as a measure of bivariate scatter just as we do the RSD, the ES, etc. The diagonal, D, is simply the square root of the sum of squares of the range values, EHD and EVD, in the two directions. Thus, we have

$$D = \sqrt{(EHD)^2 + (EVD)^2} \quad (10)$$

and from the figure, we see that numerically

TABLE 8. TABLE OF MEAN VALUES, STANDARD DEVIATIONS AND THE APPROXIMATE 95% PROBABILITY LEVELS OF THE DIAGONAL (D)

Sample Size	Mean Value of D	Reciprocal of Mean Value	Standard Deviation	95% Prob Level of D
n	$E(D/\sigma)$	$1/[E(D/\sigma)]$	$SD(D/\sigma)$	$D_{.95}(D/\sigma)$
2	1.772	.5643	.9294	3.46
3	2.540	.3937	.9254	4.16
4	3.035	.3294	.9021	4.60
5	3.397	.2945	.8795	4.91
6	3.680	.2717	.8616	5.15
7	3.911	.2557	.8483	5.35
8	4.107	.2435	.8306	5.52
9	4.276	.2339	.8143	5.66
10	4.423	.2261	.8061	5.79
11	4.555	.2196	.7912	5.90
12	4.672	.2140	.7871	6.00
13	4.779	.2092	.7784	6.09
14	4.877	.2050	.7690	6.17
15	4.967	.2013	.7614	6.25
16	5.050	.1980	.7564	6.32
17	5.128	.1950	.7461	6.39
18	5.200	.1923	.7424	6.45
19	5.268	.1898	.7356	6.51
20	5.332	.1875	.7297	6.56

(Tabular entries in the second, fourth and fifth columns are to be multiplied by the population σ or an estimate of that σ. The values given in Table 8 were determined by using the Chi approximation of Patnaik [12], which is described in the Appendix. Since the diagonal D consists of components under the radical involving the ranges in both the x and y directions, then the problems of approximating the true distribution of the diagonal are very similar to those of the range. The 95% probability levels were also obtained by using this same approximation of Patnaik [12]).

The use of the diagonal may be especially desirable, for example, when the impact points are not clearly discernible.

Note:

There may be some patterns of shots on the target, which result in the extreme spread ES, the diameter of the covering circle 2RC and the diagonal D all being equal.

$$D = \sqrt{(7)^2 + (6)^2} = \sqrt{85} = 9.22 .$$

Table 8 gives the expected or mean values and the standard deviations of the diagonal, D, for n = 2 to 20 shots or rounds, along with approximate 95% probability levels, all of which were obtained or calculated from theory we have in our Appendix.

Example 18. Estimate the CEP by using the Diagonal D.

Answer: The CEP = 1.1774 times the estimate of σ, or
1.1774(9.22/4.423) = 2.45
as compared to the value of 2.58 in Example 10 for which we used the most efficient estimator, the RSD.

Example 19. Which is the more efficient estimate of σ for a sample of 10 rounds, the diagonal D, the extreme spread ES, or the radius of the covering circle RC?

Answer: The precision of the diagonal D is calculated as .8061/4.423 = .182, which is less than the value of 0.191 for the RC from Example 17 and the value of 0.195 for the ES. Therefore, the diagonal D is a more efficient estimator of σ than either the radius of the covering circle or the extreme spread. Make frequent use of the diagonal!

12. THE RELATIVE PRECISION OF THE VARIOUS UNBIASED ESTIMATORS OF POPULATION STANDARD DEVIATION

As a basis for comparison, we summarize in tabular detail in our Table 9 the relative precisions of the various single direction or univariate measures of dispersion, and the bivariate or two-directional measures, we have studied in our account herein. Recalling from Example 9 that the most useful or appropriate method of comparison of the different estimates involves first dividing each estimate by its mean or expected value, then our definition of precision consists of simply determining the standard deviation of an unbiased estimate of the unknown population sigma. Hence, precision as defined herein is actually the coefficient of variation multiplied by the population σ. In this connection, we prefer to deal with the standard deviation - the first power or the linear measure or value - rather than the second power or the variance, or the reciprocal of the variance, etc., as a measure of imprecision. Thus, in order to determine the imprecision for each of the measures of dispersion or pattern closeness, and for each different sample size, one takes the standard deviation for each sample size and divides that by the corresponding mean value tabulated. The following Table 9 gives the final values of imprecision and they may be used as a basis for comparison or also to calculate the efficiency (by simply squaring the tabulated value) of the measures

TABLE 9. THE RELATIVE PRECISIONS OF THE VARIOUS MEASURES OF PATTERN DISPERSION ON THE TARGET *

Sample Size n	s_x or s_y	EHD, EVD or R	MHD or MVD	RSD	MR	ES	RC	D	FOM
2	.755	.755	.755	.523	.523	.523	.523	.524	.534
3	.523	.525	.525	.363	.365	.369	.368	.364	.371
4	.422	.427	.430	.294	.298	.307	.303	.297	.302
5	.363	.372	.373	.254	.259	.270	.266	.259	.263
6	.323	.335	.334	.226	.232	.246	.242	.234	.237
7	.294	.308	.306	.206	.212	.227	.224	.217	.218
8	.272	.288	.283	.191	.197	.215	.210	.202	.204
9	.254	.272	.265	.178	.184	.203	.200	.190	.192
10	.239	.260	.250	.168	.174	.195	.191	.182	.184
11	.226	.248	.238	.159	.165	.189	.183	.174	.175
12	.216	.239	.227	.151	.157	.183	.177	.168	.169
13	.207	.231	.217	.145	.150	.177	.171	.163	.163
14	.198	.224	.209	.139	.145	.171	.166	.158	.158
15	.191	.218	.201	.134	.139	.166	.162	.153	.154
16	.184	.212	.194	.130	.134	.162	.158	.150	.150
17	.178	.207	.188	.125	.130	.159	.155	.145	.146
18	.173	.203	.183	.122	.127	.156	.152	.143	.144
19	.168	.199	.177	.118	.123	.153	.149	.140	.141
20	.163	.195	.173	.115	.120	.150	.147	.137	.138

of dispersion of interest and dividing such results into the corresponding square of the RSD values. Thus, for n = 8, we get the efficiency of the FOM = $(.191/.204)^2$ = 88%.

From an examination of Table 9, we notice that the bivariate measures of dispersion (5th - 10th columns) are considerably more precise than the univariate or one - directional estimators (2nd - 4th columns) as might well be expected, of course. For the univariate measures, the standard deviation of sample values is the most precise or efficient estimator of σ, while for the bivariate measures of scatter the radial standard deviation is uniformly best. It is observed in addition that the mean radius stands next in precision to that of the RSD, and the extreme spread, which is the simplest to determine from the pattern of shots, is the most inefficient of all the various bivariate estimators studied. The diagonal is somewhat better than the radius of the covering circle in precision and efficiency. The FOM or figure of merit is surprisingly precise for the sample sizes considered.

Many comparisons may be made with the imprecision values listed in Table 9. As an example, the RSD for a sample size

* The precision of an average measure based on m subgroups of n each is found from Table 9 by dividing the appropriate entry by $\sqrt{m}$.

of 8 is just as precise as the RC for a sample of 10 rounds, and the mean radius for 11 rounds is about as precise as the extreme spread for 15 rounds (.165 vs .166).

As previously indicated, one may calculate the relative efficiency of any estimate by taking the appropriate Table 9 entry, dividing it into the corresponding tabular value for the RSD, and squaring the result. Also, however, the relative efficiency of one measure of pattern tightness in contrast to that of another one may be easily found. Such may be determined by dividing the smaller of any two corresponding values by the larger one and squaring the result. Thus, the efficiency of the extreme spread relative to that of the mean radius for a sample of 16 rounds is $(.134/.162)^2 = 69\%$.

If the values in the second column of Table 9 are divided by $\sqrt{2}$, then the result is the standard deviation of a slightly biased estimate of σ based on the average of s_x and s_y. In this connection, it will be found that the unbiased estimate based on the RSD is slightly more precise as in column 5. Also, the RSD is somewhat more precise than the estimate of σ computed from the square root of the average of s_x^2 and s_y^2 as in Example 1 because of the divisor.

Moranda [10] and Scheuer [15] have studied the efficiency of various estimators of the CEP for the case where one may assume that the true C of I is known, i. e. that the x_i and y_i are deviations which are normally and independently distributed with zero mean and unknown variance σ^2 (or standard deviation σ). In our treatment here we assume that both the mean and the variance of the bivariate population are unknown as this generally is the case in practice. (Concerning estimation, some readers will be interested in Example 22 at the end of the following Section 13.)

13. SIGNIFICANCE OF THE OBSERVED CENTER OF IMPACT

It is our purpose here to determine whether the observed center of impact (C of I) differs significantly from the aim point or whether the sight on the rifle is set properly. Referring to Figure 1, suppose that the aim point was the origin, i. e. the point (0, 0). We know that the observed C of I for the ten rounds fired turns out to be the point $\bar{x} = 1$, $\bar{y} = 1$ or (1, 1). In order to judge whether there exists any statistical evidence to indicate that the true C of I or aim point is actually at the origin (0, 0) or not, we first make the assumption that the x and y directions are really independent or uncorrelated and the standard deviations in the x and y directions are equal, i. e. $\sigma_x = \sigma_y = \sigma$. In this case then we use the statistical test based on

$$F = \frac{n(\bar{x} - \alpha)^2 + n(\bar{y} - \beta)^2}{S_x^2 + S_y^2}, \qquad (11)$$

with

$$S_x^2 = \frac{\Sigma(x_i - \bar{x})^2}{n - 1}, \quad S_y^2 = \frac{\Sigma(y_i - \bar{y})^2}{n - 1}, \quad S_{xy} = \frac{\Sigma(x_i - \bar{x})(y_i - \bar{y})}{n - 1}$$

$$= \frac{A_{xx}}{n(n - 1)}, \quad = \frac{A_{yy}}{n(n - 1)}, \quad = \frac{A_{xy}}{n(n - 1)}^{*},$$

which follows the Snedecor F-distribution with 2 degrees of freedom for the numerator and 2n - 2 degrees of freedom for the denominator. α and β are the hypothesized x and y values for the location of the true C of I. If the observed values of F are significant, then we would reject the null hypothesis that the C of I is located at the point (α, β). Now for an example, as this is a study of accuracy of fire.

<u>Example 20</u>. Using the data of the figure, then determine if it can be said that the true unknown C of I is at (0, 0), or that is at the origin.

Answer: We have the pertinent numerical data,

$$n = 10, \bar{x} = 1, \bar{y} = 1, S_x^2 = 5.33, S_y^2 = 4, \text{ and } S_{xy} = 0.889.$$

In order to demonstrate the proper use of (11), we should show first that the standard deviations in the x and y directions are equal. This is done by calculating

$$F = S_x^2/S_y^2 = 5.33/4 = 1.33,$$

which is not significant for 9 and 9 degrees of freedom.

To test the hypothesis that $\alpha = \beta = 0$, or i. e. that the true unknown C of I is at the origin, we calculate

$$F = \frac{10(1 - 0)^2 + 10(1 - 0)^2}{5.33 + 4.00} = 2.14,$$

which for 2 and 18 degrees of freedom is not significant, even at the 90% probability level of F. Hence, we are in the position of concluding that true aim point could well be considered to be at the origin (0, 0).

The reader will note that we did not use the calculated value $S_{xy} = .889$ in this particular example. This is due to the fact that the sigmas in the two directions were found to

* See Formula (12) on next page.

be equal and there doesn't seem to be any correlation at all between the x and y values, as judged from an examination of the Figure. However, the much more general case of unequal sigmas and an inclined pattern of shots also in the two directions would require instead of (11) the formula or sample statistic

$$T^2 = \frac{s_y^2[\sqrt{n}(\bar{x}-\alpha)]^2 - 2s_{xy}[\sqrt{n}(\bar{x}-\alpha)][\sqrt{n}(\bar{y}-\beta)] + s_x^2[\sqrt{n}(\bar{y}-\beta)]^2}{s_x^2 s_y^2 - s_{xy}^2}, \quad (12)$$

which is known as Hotelling's T^2. The quantity

$$F = (n - 2)T^2/2(n - 1) \quad (12a)$$

follows the Snedecor F distribution with 2 and n - 2 degrees of freedom.

For our Example 20, we find that F = 8(3.68)/2(9) = 1.64 and this value of F for 2 and 8 d. f. is not significant, so that the true C of I may still be taken as the origin (0, 0).

Finally, two other examples on the text:

Example 21. The radial standard deviation (RSD) for the ten shots on the Figure turned out to be 2.90. Therefore, it is seen that the radial variance is $(2.90)^2 = 8.41$ and this estimate has 18 d. f. Suppose now that another rifleman fired 10 rounds at the target, obtaining an RSD of 4.50. Are the two population standard deviations for the riflemen equal?

Answer: The radial variance for the second rifleman will be found to be 20.3 and also has 18 d. f. Moreover, the ratio of variances is 2.41, and under the null hypothesis of equal variances such a quantity would be distributed as "F". But $F_{.95}(18, 18) = 2.22$, and hence we would have to conclude that the first rifleman exhibits the smaller sigma (tighter pattern). (A similar test applies also to the squares of the two Diagonals - see the Appendix.)

Example 22. In estimating σ by using the RSD and by formula (9), would it ever pay not to use the $\bar{x}$ and $\bar{y}$ in the formula (9) as estimates of the population means, but rather ignore them by substituting zero values in their place?

Answer: Yes. Moranda [10] shows that for a small sample size of about 8 or less, then the RSD based on (9) is not as precise as the quantity

$$\Sigma(x_i^2 + y_i^2)/n$$

in estimating the CEP = 1.1774σ, since two degrees of freedom are lost if $\bar{x}$ and $\bar{y}$ are used in (9).

14. APPENDIX ON RELATED STATISTICAL THEORY

A. The Sample Standard Deviation (s or s_x or s_y).

It is well known that the Chi-square probability distribution is given by

$$f(\chi^2)\, d\chi^2 = \frac{1}{\Gamma(\frac{\nu}{2})\, 2^{\frac{\nu}{2}}} (\chi^2)^{\frac{\nu}{2}-1} e^{-\frac{\chi^2}{2}} d\chi^2 \quad (13)$$

for ν degrees of freedom (d.f.). For $s^2 = \frac{1}{n}\Sigma(x_i-\bar{x})^2$, then $\chi^2 = ns^2/\sigma^2$ is distributed as Chi-square, and we have that the kth moment of the sample standard deviation, s, for a normal population will be given by

$$E(s^k) = \frac{(n/\sigma^2)^{\frac{n-1}{2}}}{\Gamma(\frac{n-1}{2})\, 2^{\frac{n-1}{2}}} \int_0^\infty s^{n+k-2} e^{-ns^2/2\sigma^2} ds = \left(\frac{2\sigma^2}{n}\right)^{\frac{k}{2}} \frac{\Gamma(\frac{n+k-1}{2})}{\Gamma(\frac{n-1}{2})}. \quad (14)$$

For k = 1, we get the mean value of the standard error, s, or

$$E(s) = \left(\frac{2}{n}\right)^{1/2} \frac{\Gamma(\frac{n}{2})}{\Gamma(\frac{n-1}{2})} \sigma = c_n \sigma, \text{ say.} \quad (15)$$

[The very complex c_n is about equal to (n-1)/(n-.25)]

For k = 2, $$E(s^2) = \left(\frac{2}{n}\right) \frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n-1}{2})} \sigma^2 = \frac{n-1}{n} \sigma^2, \quad (16)$$

so that the variance of s is

$$\text{Var } s = \sigma_s^2 = [\frac{n-1}{n} - c_n^2]\sigma^2, \quad (17a)$$

and the standard deviation of s is

$$\sigma_s = \sqrt{\frac{n-1}{n} - c_n^2}\ \sigma. \quad * \qquad \textbf{(17b)}$$

Values of $c_n = (\frac{2}{n})^{1/2}\Gamma(\frac{n}{2})/\Gamma(\frac{n-1}{2})$ and the quantity $\sqrt{\frac{n-1}{n} - c_n^2}$ for n = 2(1)20 are given in Table 1. They are based on the Biometrika Tables for Statisticians [14].

B. The Sample Range (w, or R_x, R_y, or EHD, or EVD)

The probability distribution of the sample range, i. e. the largest minus the least sample values (which is also the extreme horizontal or extreme vertical dispersion) has been investigated by Tippett [17], Dederick [2] and Hartley [7]. Letting $w = x_n - x_1$ = largest minus smallest sample values, then the probability that w is less than W for normal samples is

$$P_n(W) = n \int_{-\infty}^{\infty} f(x)\ [\int_x^{x+W} f(t)dt]^{n-1}dx, \qquad (18)$$

where $f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$.

The probability integral of the range, w, has been tabulated by Pearson [13] for n = 2(1)20 and Dederick [2] for small values of n.

Tippett [17] in 1925 gave values of the mean and standard deviation of the sample range which we label as

$$E(w) = d_n\sigma \qquad (19)$$

and

$$\sigma_w = k_n\sigma, \qquad (20)$$

where d_n and k_n depend, of course, on the sample size, n. The values of d_n and k_n given in Table 2 were taken from the Biometrika Tables [14].

* The quantity $c_n \approx 1 - 3/4(n-1) + 17/32(n-1)^2$, or $1/c_n \approx (n - .25)/(n - 1)$. Also, $E(\chi) \approx \sqrt{n}(n-1)/(n-.25)$ and $\Gamma(n/2)/\Gamma[(n-1)/2] \approx \sqrt{n}(n-1)/\sqrt{2}(n-.25)$.

Patnaik [12] has made a study of the distribution of the range by approximating it as $w/\sigma \approx c\chi/\sqrt{\nu}$, where c is a scale factor depending on n and χ is a Chi variate with ν, an equivalent (fractional) number of degrees of freedom for Chi. Values of c and ν are given in our Table 11, page 42.

C. The Sample Mean Deviation (MD, MHD or MVD)

The probability distribution of the sample mean deviation for any general sample size, n, was worked out apparently first by Godwin [4] in 1945, although R. A. Fisher in 1920 derived the mean value and the standard deviation of the sample mean deviation for samples of size n from a normal population. Godwin's distribution is a multiple integral which was tabulated by numerical quadrature.

Using MD to denote the sample mean deviation, which is also the mean horizontal deviation (MHD) or mean vertical (MVD), then Fisher [3] showed that

$$E(MD) = \sqrt{2(n-1)/n\pi}\ \sigma \rightarrow .7979\ \sigma, \qquad (21)$$

$$\sigma_{MD} = \sqrt{\frac{2(n-1)}{n^2\pi}\left\{\frac{\pi}{2} + \sqrt{n(n-2)} - n + \sin^{-1}\frac{1}{n-1}\right\}}\ \sigma \qquad (22)$$

for random normal samples.

Values of the mean value (21) and the standard error (22) of the sample mean deviation are available in the Biometrika Tables [14]. These moment constants are given in our Table 3.

D. The Radial Standard Deviation (RSD)

The radial standard deviation (RSD) given by

$$RSD = R = \sqrt{\frac{1}{n}[\Sigma(x_i - \bar{x})^2 + \Sigma(y_i - \bar{y})^2]} \qquad (23)$$

has been investigated by Grubbs [5]. When $\sigma_x^2 = \sigma_y^2 = \sigma^2$, we note that $n \cdot R^2/\sigma^2$ for normal samples is distributed as a χ^2 with 2(n-1) degrees of freedom. In this case, the elementary probability distribution of R is from (13)

$$d\,F(R) = \frac{2(n/2\sigma^2)^{n-1}}{\Gamma(n-1)} e^{-(nR^2/2\sigma^2)} R^{2n-3} dR, \quad 0 < R < \infty . \tag{24}$$

The first moment and the variance of the RSD are

$$E(R) = \frac{\Gamma(n-\frac{1}{2})}{\Gamma(n-1)} \frac{2}{\sqrt{n}} \sigma \simeq \sqrt{\frac{2(n-1)}{n}} \{1 - \frac{1}{8(n-1)} + \frac{1}{128(n-1)^2}$$

$$+ \frac{5}{1024(n-1)^3}\} \sigma , \tag{25}$$

and $\text{Var}\ R = \frac{2}{n}\{n-1 - [\frac{\Gamma(n-\frac{1}{2})}{\Gamma(n-1)}]^2\}\sigma^2.$ (26)

For $n = 2(1)15$, reference [5] gives the mean, the standard deviation and the .5%, 5%, 95% and 99.5% probability levels of R. Mean values and standard errors of the radial standard deviation, as a multiple of σ (and not $\sqrt{2}\,\sigma$ as in [5]), are given in Table 4.

When $\sigma_x \neq \sigma_y$, the quantity $R^2/(\sigma_x^2 + \sigma_y^2)$ has mean equal to $(n-1)/n$ and variance equal to $2(n-1)(1+\lambda^4)/n^2(1+\lambda^2)^2$, where $\lambda = \sigma_x/\sigma_y$. We may thus take the quantity $\chi_a^2 = n(1+\lambda^2)^2R^2/(1+\lambda^4)(\sigma_x^2+\sigma_y^2)$ as being distributed approximately as Chi-square [6] with $(n-1)(1+\lambda^2)^2/(1+\lambda^4)$ degrees of freedom. Therefore, probability statements about R or R^2 may be made using χ_a^2 as being distributed like Chi-square.

Since the radial variance follows a Chi-square distribution, then we may easily compare the dispersion patterns of two riflemen or that on two targets by using the "F" test.

E. The Mean Radius (MR)

The mean radius (MR) is defined as the mean of the radial distances from the observed center of impact (C of I) to the individual points of impact on the target. The observed C of I is $(\bar{x}, \bar{y})$, and an individual impact point is (x_i, y_i), thus the (bivariate) sample mean radius is

$$MR = \frac{1}{n} \sum_{i=1}^{n} \sqrt{(x_i - \bar{x})^2 + (y_i - \bar{y})^2} . \tag{27}$$

We observe that $n(x_i-\bar{x})^2/(n-1)\sigma^2$ and also $n(y_i-\bar{y})^2/(n-1)\sigma^2$ are each distributed as Chi-square with one degree of freedom. Hence, the quantity

$\chi_i = \frac{\sqrt{n}}{\sqrt{n-1}\,\sigma} \sqrt{(x_i-\bar{x})^2 + (y_i-\bar{y})^2}$ is a Chi-variable with 2 degrees of freedom, and we thus have

$$MR = \sqrt{\frac{n-1}{n}} \cdot \frac{\sigma}{n} \sum_{i=1}^{n} \chi_i\,(2). \tag{28}$$

Therefore, the mean value of the MR is

$$E(MR) = \sqrt{\frac{n-1}{n}} \cdot \frac{\sigma}{n} \sum_{i=1}^{n} E(\chi_i) = \sqrt{\frac{n-1}{n}} \cdot \sqrt{2}\, \frac{\Gamma(\frac{3}{2})}{\Gamma(1)} \sigma \to 1.253\,\sigma . \tag{29}$$

Values of the multiplier of σ are given in Table 5 for $n = 2(1)20$.

In order to find the variance and hence the standard deviation of the mean radius, MR, we use the relation

$$\text{Var}\ MR = E(MR)^2 - [E(MR)]^2 . \tag{30}$$

Now

$$E(MR)^2 = E\{\frac{1}{n} \cdot \sum_{i=1}^{n} \sqrt{(x_i-\bar{x})^2 + (y_i-\bar{y})^2}\}^2 \tag{31}$$

$$= \frac{1}{n^2} E\{\sum_{i=1}^{n} [(x_i-\bar{x})^2 + (y_i-\bar{y})^2]$$

$$+ 2\sum_{i<j}\sqrt{[(x_i-\bar{x})^2+(y_i-\bar{y})^2][(x_j-\bar{x})^2+(y_j-\bar{y})^2]}\}$$

$$= \frac{1}{n^2} \{n \cdot \frac{2(n-1)}{n} \sigma^2 + 2\binom{n}{2}\left(\frac{n-1}{n}\right)\sigma^2 \cdot E\sqrt{(u_i^2+v_i^2)(u_j^2+v_j^2)}\},$$

where $u_i = (x_i-\bar{x})/\sqrt{(n-1)/n}\,\sigma$ and $v_i = (y_i-\bar{y})/\sqrt{(n-1)/n}\,\sigma$,

so that u_i and v_i are normally distributed with zero mean and unit variance. Note, however, that u_i and u_j (and v_i and v_j) are correlated, and that $E(u_i\, u_j) = \rho = -\frac{1}{n-1}$.

In order to find $E\sqrt{(u_i^2 + v_i^2)(u_j^2 + v_j^2)}$, we evaluate the quadruple integral

$$\int\int_{-\infty}^{\infty}\int\int \frac{\sqrt{(u_i^2+v_i^2)(u_j^2+v_j^2)}}{4\pi^2(1-\rho^2)} e^{-\frac{1}{2(1-\rho^2)}\{u_i^2-2\rho u_i u_j+u_j^2+v_i^2-2\rho v_i v_j+v_j^2\}}$$

$$\cdot\, du_i du_j dv_i dv_j \quad . \qquad (32)$$

Let $u_i = r\cos\theta$, $v_i = r\sin\theta$, $u_j = s\cos\phi$ and $v_j = s\sin\phi$, then the integral becomes

$$\frac{1}{4\pi^2(1-\rho^2)} \int_0^{\infty}\int\int_0^{2\pi}\int r^2 s^2\, e^{-\frac{1}{2(1-\rho^2)}\{r^2-2\rho rs\cos(\theta-\phi)+s^2\}}$$

$$\cdot\, dr\, ds\, d\theta\, d\phi$$

$$= \sum_{\nu=0}^{\infty} \frac{\rho^{\nu}}{4\pi^2(1-\rho^2)^{\nu+1}\,\nu!}\{\int_0^{\infty} r^{\nu+2} e^{-\frac{r^2}{2(1-\rho^2)}} dr\}^2 \{\int_0^{2\pi}\int \cos^{\nu}(\theta-\phi)$$

$$\cdot\, d\theta\, d\phi\} \, .$$

Carrying out the indicated integration, we obtain finally

$$E\sqrt{(u_i^2 + v_i^2)(u_j^2 + v_j^2)} =$$

$$\sum_{\nu=0}^{\infty} \frac{(1-\rho^2)^2\rho^{\nu}2^{\nu+1}[\Gamma(\frac{\nu+3}{2})]^2}{\nu!} \begin{cases} 1 & \text{if } \nu = 0 \\ 0 & \text{if } \nu \text{ is odd} \\ \left(\frac{\nu-1}{\nu}\right)\left(\frac{\nu-3}{\nu-2}\right)\cdots\left(\frac{3}{4}\right)\frac{1}{2} & \text{if } \nu \text{ is even} \end{cases}$$

$$(33)$$

This enables us to find that the variance of the MR is

$$\text{Var MR} = \frac{2(n-1)}{n^2}\{1 + \frac{n-1}{2} \cdot E\sqrt{(u_i^2 + v_i^2)(u_j^2 + v_j^2)} - \frac{n\pi}{4}\}\sigma^2 \, . \qquad (34)$$

Values of the standard deviation of the mean radius were determined with this formula and are given in Table 5.

F. The Extreme Spread or Bivariate Range

The Extreme Spread or Bivariate Range is the maximum of the distances between the $\binom{n}{2}$ pairs of points or impacts on the target plane. Denoting a general point by (x_i, y_i), $i = 1,2,3, \ldots, n$, then the extreme spread is the maximum of the distances

$$\sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \; ; \; i \neq j; \; i, j = 1,2,3,4, \ldots, n.$$

For n = 2, we note that the extreme spread is given by

$ES_{n=2} = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} = \sqrt{2}\,\chi\sigma$, a random Chi variate, where Chi has two degrees of freedom. Now in this case, $E(ES) = 1.77245\ \sigma$ and $\sigma_{ES} = .9265\ \sigma$.

For any general sample size, n, the exact distribution of the extreme spread, ES, has apparently not been worked out analytically.

Mean values, standard deviations, and the third and fourth moments of the extreme spread were computed originally under the direction of the late Samuel S. Wilks of Princeton University as a result of **discussions between** the **author** and Prof. Wilks concerning the Tracking Data Analysis study of Reference [11]. In connection with this study, the first four moments of the trivariate range, the trivariate midrange, the bivariate range, and the bivariate midrange were computed by a Monte Carlo sampling procedure on an IBM 7090 Computer. The Fortran Language was used by Prof. Wilks and Mr. Paul Raynault of Princeton University, the latter person assisting in the computations. More recently, Taylor and Grubbs [16] performed very extensive Monte Carlo computations to get moments and percentage points of the extreme spread more accurately. Their moment constants are given in Table 10. The values listed in the table should not, of course, be treated as exact ones, since due to sampling error the third decimal places may be slightly in error for sample sizes greater than n=2. The skewness and kurtosis moment constants, α_3 and α_4, are given by

$$\alpha_3 = \sqrt{\beta_1} = \mu_3/\mu_2^{3/2} \quad \text{and} \quad \alpha_4 = \beta_2 = \mu_4/\mu_2^2 .$$

In connection with approximations to the probability distribution of the extreme spread, Taylor and Grubbs [16] have found that the approximate Chi-square technique [6] gives a good fit for the Monte Carlo percentage points obtained.

TABLE 10

MOMENT CONSTANTS FOR THE BIVARIATE RANGE OR EXTREME SPREAD

(FROM REFERENCE [16])

Sample Size n	Mean μ_1'	Standard Deviation σ	α_3	α_4
2	1.772	0.932	0.632	3.294
3	2.406	0.887	0.451	3.143
4	2.787	0.856	0.393	3.163
5	3.066	0.828	0.390	3.171
6	3.277	0.806	0.374	3.194
7	3.443	0.783	0.373	3.177
8	3.582	0.771	0.392	3.231
9	3.710	0.754	0.382	3.215
10	3.813	0.745	0.388	3.288
15	4.190	0.694	0.395	3.255
20	4.452	0.668	0.400	3.240
25	4.639	0.650	0.439	3.307
28	4.734	0.642	0.426	3.357
30	4.788	0.635	0.463	3.441
31	4.822	0.631	0.434	3.321
34	4.891	0.623	0.422	3.318

G. The Radius of the Covering Circle (RC)

The Covering Circle is defined as the smallest circle on the target containing on it or within it all of the sample points or bullet impacts. The distribution of the radius, RC = r, of the covering circle, and the distribution of its center ρ has been studied by Daniels [1]. Daniels shows that the probability density function of the radius r for normal samples is

$$dF_n(r) = n(n-1)e^{-r^2/\sigma^2}(1-e^{-r^2/2\sigma^2})^{n-2} r \, dr/\sigma^2 , \qquad (35)$$

and the cumulative distribution function of r is

$$F_n(r) = n(1-e^{-r^2/2\sigma^2})^{n-1} - (n-1)(1-e^{-r^2/2\sigma^2})^{n}. \qquad (36)$$

Daniels points out that the cumulative distribution of r may be related to and computed from Karl Pearson's Incomplete Beta Function as

$$I_z(2, n-1) \quad \text{with } z = e^{-r^2/2\sigma^2} .$$

Also,

$$r = \{2 \ln [(n-1)/2F + 1]\}^{1/2}\sigma, \qquad (37)$$

where F has Fisher's variance ratio distribution with 4 and 2(n-1) degrees of freedom.

Values of the mean and standard deviation of r/σ or RC/σ in our Table 7 were obtained from Table 1 of Daniel's paper [1].

H. The Diagonal (D)

The diagonal, D, is simply the square root of the sum of squares of the extreme horizontal dispersion, EHD, and the extreme vertical dispersion, EVD. Thus, we have

$$D = \sqrt{(EHD)^2 + (EVD)^2} = \sqrt{R_x^2 + R_y^2} , \qquad (38)$$

where R_x and R_y are the independent ranges or maximum dispersions for normal samples in the x and y directions, respectively. Following Patnaik [12], we will approximate the range by a Chi variable, i. e.

$$\frac{R_x}{\sigma} \simeq \frac{c}{\sqrt{\nu}} \chi_x \text{ and } \frac{R_y}{\sigma} \simeq \frac{c}{\sqrt{\nu}} \chi_y, \qquad (39)$$

where c is a scale factor depending on the number of points, n, and ν is an equivalent fractional number of degrees of freedom for Chi.

Let us define the mean and variance of the univariate range (R) by

$$E(R) = d_n\sigma , \qquad \text{Var } R = k_n^2 \sigma^2 . \qquad (40)$$

Then, of course

$$E(R^2) = (d_n^2 + k_n^2)\sigma^2 . \qquad (41)$$

Further, due to the approximation, $R/\sigma \simeq c\,\chi_\nu/\sqrt{\nu}$, then

$$E(R/\sigma) = \frac{c}{\sqrt{\nu}} \cdot E(\chi_\nu) = \frac{c}{\sqrt{\nu}} \sqrt{2}\, \Gamma(\tfrac{\nu+1}{2})/\Gamma(\tfrac{\nu}{2}). \qquad (42)$$

and

$$E(R^2/\sigma^2) = \frac{c^2}{\nu} \cdot E(\chi_\nu^2) = \frac{c^2}{\nu} \cdot \nu = c^2 . \qquad (43)$$

Therefore, for a formula for the scale factor c, we have

$$c^2 = d_n^2 + k_n^2. \qquad (43a)$$

Using the Chi approximation for both R_x and R_y, we find the mean value of the Diagonal D:

$$E(D) = E\sqrt{R_x^2 + R_y^2} \simeq E\sqrt{c^2\chi_x^2/\nu + c^2\chi_y^2/\nu}\;\sigma$$

$$= \frac{c}{\sqrt{\nu}}\sigma \cdot E\sqrt{\chi_\nu^2 + \chi_\nu^2} = \frac{c}{\sqrt{\nu}}\sigma \cdot E(\chi_{2\nu}) \qquad (44)$$

$$= c\sqrt{2}\,\Gamma(\nu+\tfrac12)\sigma/\sqrt{\nu}\,\Gamma(\nu)$$

$$\approx \sqrt{2}\, c\sigma\{1 - 1/8\nu + 1/128\nu^2 + 5/1024\nu^3\},$$

from which we may find the mean values of the diagonal D. Note that c and ν both depend on the number n of points or impacts; c is found from (43a) and ν by formula (5) of Patnaik's paper [12]. Our d_n = Patnaik's M.

The variance of D is Var D = $E(D^2) - [E(D)]^2$, or

$$\text{Var } D = \frac{2c^2}{\nu}\{\nu - [\Gamma(\nu+\tfrac{1}{2})/\Gamma(\nu)]^2\}\ \sigma^2$$

$$= \{2\ c^2 - \frac{2c^2}{\nu}\ [\frac{\Gamma(\nu+\tfrac{1}{2})}{\Gamma(\nu)}]^2\}\sigma^2 \ . \qquad (45)$$

Mean values of the diagonal using (44) and standard deviations of the diagonal from formula (45) are given in our Table 8. Values of n, c and ν are given in Table 11.

TABLE 11

Table of Values of n, c and ν for the Diagonal
(Values of d_n and k_n are given in Table 2)

n	c	ν
2	1.41421	1.0000
3	1.91154	1.9846
4	2.23887	2.9291
5	2.48125	3.8267
6	2.67253	4.6772
7	2.82980	5.4841
8	2.96288	6.2512
9	3.07793	6.9818
10	3.17905	7.6799
11	3.26910	8.3482
12	3.35016	8.9893
13	3.42379	9.6055
14	3.49117	10.1991
15	3.55323	10.7717
16	3.61072	11.3251
17	3.66422	11.8606
18	3.71424	12.3795
19	3.76118	12.8829
20	3.80537	13.3719

Since $\nu D^2/c^2\sigma^2$ is approximately distributed as Chi-square with 2ν d.f., then the "F" test may be used for comparing the dispersion patterns on two targets. For the same number of shots on each target we merely find the ratio of the square of the two diagonals and look up F for 2ν and 2ν d.f. For unequal numbers of shots on the two targets we compute

$$F \approx \frac{D_1^2}{c_1^2} \Big/ \frac{D_2^2}{c_2^2} \quad \text{with } 2\nu_1 \text{ and } 2\nu_2 \text{ d.f.}$$

I. Probability of Hitting a Circular Target

So far we have discussed the various measures of precision and their statistical characteristics. To make our treatment more complete, we should also discuss probability of hitting, at least in an elementary way. Assuming that x and y are independently and normally distributed with means (or aim point) α and β, and round-to-round variances σ_x^2 and σ_y^2, respectively, then the density function of x and y is given by

$$f(x, y) = \frac{1}{2\pi\sigma_x\sigma_y} \exp\left\{-\frac{(x-\alpha)^2}{2\sigma_x^2} - \frac{(y-\beta)^2}{2\sigma_y^2}\right\} . \qquad (46)$$

The chance that a round falls within a distance R of the C of I at (α,β), is then

$$Pr = \int_R\int f(x, y)\, dx\, dy \qquad (47)$$

which is to be taken over the region R defined by

$$(x-\alpha)^2 + (y-\beta)^2 \leq R^2 \qquad (48)$$

Case I: $\sigma_x^2 = \sigma_y^2 = \sigma^2$ (Circular Case)

For this case, we make the polar transformation

$$x - \alpha = r\cos\theta \qquad y - \beta = r\sin\theta \qquad (49)$$

and we find easily

$$Pr\{(x-\alpha)^2 + (y-\beta)^2 \le R^2\} = Pr(r^2 \le R^2) = Pr(r \le R)$$

$$= 1 - e^{-R^2/2\sigma^2} = 1 - (1/2)^{(R/CEP)^2} \quad . \quad * \tag{50}$$

If we set this probability equal to .5, then we find the radius of the circle which gives by definition the Circular Probable Error (CEP), i. e. solving

$P = .5 = 1 - e^{-R^2/2\sigma^2}$ for $R = R(.50)$, we get

$$R(.50) = CEP = 1.1774\sigma \quad . \tag{51}$$

Otherwise, the probability of hitting within any circle of radius R about the true C of I = (α, β) is given by Formula (50).

Case II: $\sigma_x \neq \sigma_y$ (Non-Circular Case)

Let us assume in all generality that we not only have $\sigma_x \neq \sigma_y$, but also that we wish to find the chance of hitting within a circle of radius R about a point (a, b) which is offset from the C of I or point of aim (α, β). For this general case, Grubbs [6] gives a method which approximates the true probability with sufficient accuracy for most practical cases. The technique requires that we set

$$\sigma_o^2 = \sigma_x^2 + \sigma_y^2 \ , \tag{52}$$

$$m = 1 + \frac{1}{\sigma_o^2}[(\alpha-a)^2 + (\beta-b)^2] \ , \tag{53}$$

$$v = 2\{\frac{\sigma_x^4 + \sigma_y^4}{\sigma_o^4} + 2[\frac{\sigma_x^2(\alpha-a)^2 + \sigma_y^2(\beta-b)^2}{\sigma_o^4}]\} \ , \tag{54}$$

and compute the quantity

---

* As an illustration, and using the data of Example 1, it is easily verified that a rifleman firing at a 12" bull's eye located at a range of 100 yards will have a probability of hitting equal to

$$1 - e^{-(6/2.22)^2/2} = 0.97.$$

$$t_R = \{\sqrt[3]{R^2/\sigma_o^2 m} - (1 - v/9m^2)\}/\sqrt{v/9m^2} \quad , \tag{55}$$

which is to be referred to a table of the standardized Normal (cumulative) distribution or integral to find the desired probability.

In connection with formulae (52)-(55), it is outlined in Reference [6] for

$$E(x) = \alpha, \ E(y) = \beta, \ E(x-\alpha)^2 = \sigma_x^2, \ E(y-\beta)^2 = \sigma_y^2 \tag{56}$$

$$\sigma_o^2 \ \psi^2 = (x-a)^2 + (y-b)^2 \tag{57}$$

and

$$E(\psi^2) = m \qquad \text{as in (53),} \tag{58}$$

$$E(\psi^2 - m)^2 = \text{Var } \psi^2 = v \text{ as in (54),} \tag{59}$$

then the quadratic form

$$2m\psi^2/v \approx \chi^2(2m^2/v) \tag{60}$$

is approximately distributed as χ^2 (Chi-square) with $2m^2/v$ degrees of freedom. Then applying the Wilson-Hilferty transformation converting χ^2 to approximately normality, we have

$$Pr\{(x-a)^2 + (y-b)^2 \le R^2\}$$

$$\approx Pr\{t = [\sqrt[3]{\chi^2/n_1} - (1-2/9n_1)]/\sqrt{2/9n_1} \le t_R\} \quad , \tag{61}$$

where $n_1 = 2m^2/v$ degrees of freedom (d.f.) .

With the above theory, it is easy to find the approximate CEP when $\sigma_x \neq \sigma_y$. We have only to equate (55) to zero and solve for R. We get

$$CEP = R_{.50} \approx \sigma_o\sqrt{m}(1-v/9m^2)^{3/2} . \tag{62}$$

When $\sigma_x = \sigma_y = \sigma_o/\sqrt{2}$ [i.e. the σ_o of Formula (52)], then $CEP = R_{.50} \approx (32/27)\sigma_x = 1.185\sigma_x$ as compared to the well-known relation $CEP = \sqrt{2 \ln 2}\ \sigma_x = 1.1774\sigma_x$.

(The Spherical Probable Error is SPE = 1.538σ.)

REFERENCES

[1] Daniels, H. E., "The Covering Circle of a Sample from a Circular Normal Distribution", Biometrika, Vol 39(1952) Parts 1 & 2, 137-43.

[2] Dederick, L. S., "Relation Between the Probable Error of a Single Shot and the Bracket of a Salvo or Series", Unpublished Aberdeen Proving Ground Report, 1926. *

[3] Fisher, R. A., "A Mathematical Examination of the Methods of Determining the Accuracy of an Observation by the Mean Error and by the Mean Square Error", Month. Not. Royal Astro. Soc., Vol. 80 (1920), 758-70.

[4] Godwin, H. J., "On the Distribution of the Estimate of Mean Deviation Obtained from Samples from a Normal Population", Biometrika, Vol. 33 (1945), 254-65.

[5] Grubbs, Frank E., "On the Distribution of the Radial Standard Deviation", Annals of Mathematical Statistics, Vol. 15 (1944), 75-81.

[6] Grubbs, Frank E., "Approximate Circular and Non-Circular Offset Probabilities of Hitting", Operations Research, Vol. 12, No. 1 (1964), 51-62.

[7] Hartley, H. O., "The Range in Random Samples",Biometrika, Vol. 32 (1941-42), 334-48.

[8] Hotelling, Harold, "The Generalization of Student's Ratio", The Annals of Mathematical Statistics, Vol. 2, (1931), 360-78.

[9] Kamat, A. R., "Some More Estimates of Circular Probable Error", Journal of the American Statistical Association, Vol. 57 (1962), 191-5.

[10] Moranda, P. B., "Comparison of Estimates of Circular Probable Error", Journal of the American Statistical Association, Vol. 54 (1959), 794-800.

[11] National Academy of Sciences, Advisory Committee to the Air Force Systems Command, "Final Report of the Panel on Tracking Data Analysis", National Academy of Sciences, Washington, D. C. (June 1964).

[12] Patnaik, P.B., "The Use of Mean Range as an Estimator of Variance in Statistical Tests", Biometrika, Vol. 37, Parts 1 and 2 (1950), 78-87.

[13] Pearson, E. S., "The Probability Integral of the Range in Samples of n Observations from a Normal Population", Biometrika, Vol. 32 (1941-42), 301-10.

[14] Pearson, E. S. and Hartley, H. O., Biometrika Tables for Statisticians, Vol. 1 (1954), Cambridge University Press.

[15] Scheuer, E. M., "Moments of the Radial Error", Journal of the American Statistical Association, Vol.57, (1962) 187-90.

[16] Taylor, M. S. and Grubbs, Frank E., "Approximate Probability Distributions for the Extreme Spread", Naval Research Logistics Quarterly, Vol. 22 (Dec. 1975), 713-9.

[17] Tippett, L. H. C., "On the Extreme Individuals and the Range of Samples Taken from a Normal Population", Biometrika, Vol. 17 (1925), 364-87.

[18] Weil, Herschel, "The Distribution of Radial Error", The Annals of Mathematical Statistics, Vol. 25 (1954), 168-70.

* An account of Dederick's original work on the sample range is also given in (General) Leslie E. Simon's book, AN ENGINEERS' MANUAL OF STATISTICAL METHODS, page 204, John Wiley & Sons, New York, N. Y., 1941.

INDEX

*For an upper confidence bound, see page 50

Note on Upper Confidence Bound for the true, unknown CEP:

As is sometimes desired, one may calculate an upper confidence on the true, unknown CEP for the general case given in Formula (62) by noting that the true CEP is really a scaling factor times the σ_o. Hence, the observed

$$s^2 = s_x^2 + s_y^2 \qquad (63)$$

has a mean of $\sigma_x^2 + \sigma_y^2$ and variance of $2(\sigma_x^4 + \sigma_y^4)/(n-1)$, so that the number of degrees of freedom for Chi-square is

$$\nu = 2m^2/v = (n-1)(1 + 2\sigma_x^2\,\sigma_y^2/\{\sigma_x^4 + \sigma_y^4\}), \qquad (64)$$

and the upper 100(1-α)% confidence bound on the true CEP is

$$(\text{Observed CEP})\sqrt{\nu/\chi_\alpha^2(\nu)}\ , \qquad (65)$$

where χ_α^2 is the lower α probability level of the well-known probability distribution of Chi-square. (The reader should note for equal variances in x and y, the degrees of freedom are properly 2(n-1).)

The above gives an approximate confidence bound for the true, unknown CEP. The CEP of (62) depends very much on the m of (53) and its square root, for the 9m^2 dominates v very greatly as the offset increases. For large offset, the parentheses to the 3/2 power in (62) approaches unity, so that the CEP becomes $\sigma_o\sqrt{m}$, and the rate of convergence is fast. This more or less indicates the importance of the mean value m. Of course, one might perhaps work out a more exact bound for the CEP, so that we leave this as an exercies for an energetic graduate student!

In a footnote on page 10 we raised the question, "Why do not riflemen use the CEP?" In fact, there still exists much confusion on the subject of rifle accuracy since the various measures of scatter depend so much on the number of shots in a group of rounds, and the underlying sigma unfortunately is a hidden parameter. Thus, better standardization is called for. As a suggestion, we put forward the idea that for circular patterns of shots - for which the sigmas may be considered to be equal in both directions - standardization may be effected by always quoting the unbiased CEP,

$$\text{CEP} = 1.1774\,\hat{\sigma}\ , \qquad (66)$$

where the quantity $\hat{\sigma}$ is always the unbiased estimate of the underlying population parameter σ based on hopefully one of the more efficient measures of dispersion herein, which accounts for scatter in both directions. At least, one could make it a point to simply use the unbiased estimate of sigma which has the effect of eliminating sample size confusion.

We note that it is fairly easy to remember that for very large samples (#rounds) the standard deviations approach the true population sigma without bias, that the mean deviations approach .7979 sigma and the mean radius approaches 1.253 σ. And obviously, the radial standard deviation approaches $\sqrt{2}\,\sigma$. However, the other measures depend very markedly on the sample size, and one must quote the number of rounds to let the reader correct for bias, or otherwise give the unbiased estimate of sigma. Always think in terms of the hidden sigma!

Note on Wild Shots:

What about "flyers" or wild rounds, which occur perhaps a bit too often? Such rounds will inflate the amount of scatter in the pattern and the analyst will sometimes want to be able to detect them, but more important will desire to find the physical cause and correct it, if at all possible. Fortunately, there are statistical tests for detecting whether some of the observations in a sample, or therefore in a given group of shots, can be considered to be aberrant, and indeed suspect. This area of statistics may be found in many textbooks and concerns procedures for the detection of outliers. Alternatively, the reader might see a United States Army Engineering Design Handbook, dated December 1983, which has the title, "Selected Topics in Experimental Statistics, With Army Applications". It is available from the National Technical Information Service, Department of Commerce, located at Springfield, Virginia 22161, and identified as DARCOM Pamphlet 706-103.

Accuracy:

As we have observed throughout our coverage of bullet impacts the term accuracy is not very easy to fully understand or define and there is a lot of confusion on the subject too. Nevertheless, we should perhaps attempt to try and add a bit more illumination finally. The underlying sigma is a measure of the "internal" scatter or imprecison of the impacts about their C of I. But, the C of I may be offset from the aimpoint, resulting in a bias causing inaccuracy therefore. Hence, a complete description of accuracy must account also for such bias or offset, especially since it is not very easy to place the C of I of shots on the aim point or "zero in". Nevertheless, scatter or imprecision of shots and the offset or bias can both be taken care of, for example, by the use of the general CEP of formula (62), involving scatter and offset. In fact, and generally speaking, a small CEP in (62) means high probability of hitting, and that is a very important measure of system accuracy itself! The rifleman should therefore study this whole matter very carefully. Indeed, an attempt should always be made to keep these components firmly in mind, including their relative sizes, when one thinks of the problem of accuracy. Detection of the size or amount of bias depends markedly on the underlying sigma and the number of rounds fired or sample size. Thus, for the problem of accuracy, one needs to estimate the sizes of the components of offset along with the round-to-round sigmas, then substitute these values in formulas (52) - (54), and compute the CEP of (62). This is indeed a very good measure of accuracy even though the sigmas of x and y are unequal and the problem of offset also exists. Hence, we would advise that riflemen might well put some thought into this suggestion! A better understanding of overall accuracy would then probably result.

Finally, do not forget that all of the measures described herein can be calculated when the inpacts of the individual shots are not discernible. In the case of a big hole or a glob of shots, one may have to use the extreme spread or another measure like the diagonal or covering circle, and in such cases the diagonal may be a little better. However, for a small number of shots, such as five per group, the differences in efficiency are not very great.

NOTES & COMMENTS