In presenting the dissertation as a partial fulfillment of the requirements for an advanced degree from the Georgia Institute of Technology, I agree that the Library of the Institution shall make it available for inspection and circulation in accordance with its regulations governing materials of this type. I agree that permission to copy from, or to publish from, this dissertation may be granted by the professor under whose direction it was written, or, in his absence, by the Dean of the Graduate Division when such copying or publication is solely for scholarly purposes and does not involve potential financial gain. It is understood that any copying from, or publication of, this dissertation which involves potential financial gain will not be allowed without written permission.

SAMPLE DESIGNS AND ANALYSES FOR SERVICE TESTS OF MILITARY WEAPONS (FINISHED ASSEMBLIES)

A THESIS

Presented to

the Faculty of the Division of Graduate Studies

Georgia Institute of Technology

In Partial Fulfillment

of the Requirements for the Degree

Master of Industrial Engineering

by

Victor B. Kovac

June 1955

SAMPLE DESIGNS AND ANALYSES FOR SERVICE TESTS

OF MILITARY WEAPONS (FINISHED ASSEMBLIES)

Approved:

Chairman

Date Approved by Chairman: JUN 4 1955

PREFACE

This study considers certain qualities of finished products and the service tests required to evaluate them. It is concerned with military weapons, especially those associated with direct, aimed fire. The statistical analyses and graphical methods of solution used are within the scope of readers familiar with only elementary statistics.

Since the items under consideration are military weapons, their sensitive nature precludes identifying them or reporting on their performance characteristics in physical units of measurement. Numerical values have been coded so as to have only academic interest. Military clearance for publishing this paper was granted by Office, Chief of Army Field Forces, Fort Monroe, Virginia, on June 4, 1954.

This opportunity is taken to express my appreciation to the officers and enlisted men of Board No. 3, Continental Army Command, Fort Benning, Georgia, for their cooperation in gathering material, and to the faculty of the School of Industrial Engineering, Georgia Institute of Technology, for their stimulus and advice in this undertaking.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ILLUSTRATIONS

ABSTRACT

SAMPLE DESIGNS AND ANALYSES FOR SERVICE TESTS OF MILITARY WEAPONS (FINISHED ASSEMBLIES)

VICTOR B. KOVAC

An analysis is made of the factors which contribute to the quality of military weapons. Accuracy and reliability are investigated to determine simple and effective methods of proof testing and analyzing results.

Ballistic errors appear to be bivariate distributions on a vertical target plane, normally distributed about the center of impact along vertical and horizontal axes. Aiming errors and human errors appear to follow similar distributions about the aiming point. Aiming errors for expert gunners are negligible compared with human and incidental errors and therefore, the two terms should not be treated as synonymous.

The human and incidental errors in two types of weapons investigated appear to be significantly larger than ballistic errors, hence they tend to determine system performance at field targets more than do ballistic errors. It is therefore not sound policy to base decisions on ballistic accuracy of weapons alone. A more appropriate

test of gunner-weapon system accuracy is to determine the characteristic hit probability on appropriate targets disposed at several distances. For greatest consistency, only proficient gunners should be used. Average gunners appear to distribute their hits more widely in the horizontal direction than in the vertical direction.

A nomograph for converting probable errors into hit probability for a wide range of target sizes is submitted. It may be used to estimate total system errors from hit probability, hence total field firing errors and ballistic errors may be compared directly. The difference between total system variance and ballistic variance yields a resultant which is considered as the human error variance. Further investigation is necessary to evaluate the effects which make the human error the dominant factor in field firing.

Malfunctions under normal firing conditions appear to follow the Poisson distribution. Severe trials may be used to induce higher rates of occurrence, but the resulting counts of malfunctions must be tested for consistency. The CHI-Square statistic may be used for comparing the frequencies of failures occurring to two weapons, but a graphic comparison made on Binomial Probability Paper seems to be more informative.

Two statistical methods found to be effective in evaluating dispersion are analysis of variance and regression of errors on range (distance).

CHAPTER I

INTRODUCTION

Definition of the Problem

Complexity of modern weapons.--During the last century, the instruments of war have become increasingly complex. Consider the transition from the smooth bore firearms of the Civil War, through the machine gun of World War I, to the present semiautomatic rifle with its myriad supporting arms. For both the individual soldier and his unit, the complement of arms has increased in number, range, accuracy and automaticity.

This study is concerned with the Army direct fire weapons, i.e., those that can be used to engage the enemy with direct, aimed fire. The present complement of these weapons ranges from the small arms such as carbines, rifles and machine guns to the larger caliber antitank rockets and recoilless guns. Those of the future depend, to a great extent, on the manner in which we evaluate the capabilities and limitations of existing material and on our alertness in utilizing technical developments.

Until World War II, the Armed Forces tended to adopt a new weapon as a single entity, either to improve upon an existing one, or to fill a gap. It is noteworthy that the U. S. water cooled machine gun, which helped produce a stalemate in World War I and which was used effectively even in World War II, was developed by one man, -- Browning. Vannevar Bush, in his "Modern Arms and Free Men," states:

> Guns were made bigger, and muzzle velocities and precision increased somewhat. But the whole gamut of new ordnance devices -- rockets, recoilless guns, guided missiles, proximity fuzes, bazookas, frangible bullets -- waited for the pressure of war, appearing then largely outside the organized system of ordnance development, and sometimes in spite of it.

During the last war, in addition to developing better weapons; scientific research improved the effectiveness of existing implements through operational analysis, provided means for accelerated training and selection of crews for critical material, and influenced the design of new weapons to accommodate the capabilities of the using troops.

Essentially, these scientific techniques consisted of simply analyzing existing material or procedures and comparing them with alternatives. This analysis undertook the measurement of crew and weapon capabilities and limitations and translated the findings into simpler, more readily understandable mathematical form. These techniques, applying the laboratory procedures of the physical scientists to the instruments of war, are essentially similar to the scientific management tools used by industry to simplify executive decision. It must be noted, however, that such effort was rarely attempted by a single researcher, but rather by teams of specialists from diverse fields selected for their probable contribution toward resolution of the problem. Furthermore, there would have been little success had it not been for industry's ability to mass produce material to acceptable tolerances under stringent standards through quality control. This study proposes to adapt some of the wartime lessons, in a simple form, to the proof testing of finished products, specifically to direct fire weapons.

Purpose of the study.--Since the value of a single product depends not only on its ability to excel in one characteristic, but to perform satisfactorily over a wide range of characteristics, it is obvious that all pertinent factors must be taken into account. Thus, while it may be risky to accept an outstanding item which contains a serious deficiency, it may not be better to accept an item that is barely acceptable, merely because it contains no serious defects. Judgement must be exercised constantly. Furthermore, a finished product may be so complex, that only specially trained operators may obtain satisfactory results and, to expect identical performance from less trained personnel, particularly under difficult conditions, is to be unrealistic.

Thus the purpose of this study is to select valid proof tests for direct fire weapons under simulated service conditions, in order to measure performance with regard to reliability and the overall accuracy of the man-gunsight-weapon system.

Events leading up to the study.--During the course of research for a simple method of measuring and computing accuracy of the weapons system, consisting of the man-gunsight-weapon combination, it became apparent that there was little public or classified information available. True, there were a good number of publications, both books and periodicals, devoted to arms, sporting rifles and ordnance, but they were quite inconsistent in their treatment of accuracy. It soon developed, that accuracy alone was not a suitable basis for judging weapon effectiveness, but, like the speed of an automobile, it was one of the important characteristics. For a proper perspective, it is better to set up standards that are essential or desirable, listing all the factors appropriate to a specific weapon family. Then the testing, evaluation and analysis take on a more reliable aspect. The problem is somewhat analogous to the testing of the quality of a final assembly or product, and is therefore subject to techniques used by industry in quality control. It goes beyond the analogy of the man-machine combination, in that psychological and environmental factors are involved. There appears to be a need for an appraisal of the problem for the benefit of Service test agencies and possibly for similar civilian agencies. At any rate, the investigation of human errors in gunnery alone appears to warrant further exploration.

Literature Survey

Background of proof tests.--Early European proof tests of cannon and hand weapons consisted of firing shots with reinforced loads (1, pp 2-12). These were acceptance tests based on structural sturdiness and perhaps for safety. At any rate, the privilege of stamping a proof mark was both an honor and a safeguard against inferior products, a trade mark. As firearms improved, government ordnance establishments undertook to provide stimulus and direction to armament makers and to safeguard their forces against faulty material. Since the turn of the Twentieth Centure, Lissak and Cranz have published mathematical treatises on internal, external and terminal ballistics (2) (3). Their efforts were directed toward improving the knowledge of ordnance and toward the replacement of earlier empirical rules. Cranz utilized statistical sampling methods to reduce the costs of proof tests. The probable error technique was also

---

(1) Numbered footnotes refer to Bibliography.

adopted and is still in use by our armed forces. In 1930, the Coast Artillery Journal published "Gunnery and Fire Control for Antiaircraft Artillery" which differentiates between accidental error and systematic error, the latter being one that could be corrected in subsequent firing.

Since the start of World War II, U. S. Ordnance published the "Ordnance Proof Manual, OPM 7-27," which illustrates the use of statistical methods in evaluating external ballistics. Simon's "Manual" is an excellent guide for Ordnance officers as well as engineers (4). It contains practical applications of statistical tools and examples of the type of proof tests conducted by Ordnance.

Two factors should be apparent. First, free discussion of almost any facet of Ordnance, including proof tests, is hampered by the restrictions imposed by the need for security. This tendency applies especially to new developments, and is true for our own as well as for foreign governments. Second, most of the tests conducted by Ordnance are in the nature of strictly engineering tests and any human error is carefully screened from them. In the treatment that follows, it will be noted that a greater freedom of expression prevails. It appears that scientists prefer a more liberal interpretation of security measures.

<u>Wartime scientific developments</u>.--The impact of science on modern weapons is revealed by a considerable number of

published works, principally concerning the operations of the Office of Scientific Research and Development (OSRD), and the National Defense Research Committee (NDRC) (5). Gray calls it the physicists' war, and from the viewpoint of this study, they contributed to the accuracy, velocity and terminal effect of modern weapons. Bray expounds on the contributions of the Applied Psychology Panel of NDRC in this manner:

> One of the greatest intellectual achievements of the war was the application of scientific method to making the material weapon fit the man, and to testing, training, and selecting the man to utilize the particular weapon.

There is no doubt that the human error is a significant factor in accuracy and in the very handling of weapons, hence performance tests must take it into serious account. From the published works concerning operations research, it is evident that certain aspects apply to the front line combatants and their individual or crew served weapons (6). These aspects include analyses directed toward obtaining the maximum effectiveness from weapons, toward improving hit probability, and toward increasing the effectiveness of hits.

While this paper is concerned with the customers' or users' viewpoint, it may be well to keep in mind

---

(5) Refer to Bibliography: Vannevar Bush, J. P. Baxter, L. R. Thiesmeyer, G. W. Gray, C. W. Bray.

contributions by industry (7). They are: improved reliability in functioning, particularly of automatic weapons; greater consistency in performance of weapons; trends in development, especially in weight reductions through the use of new alloys; and industrial engineering techniques of time, motion, and dollar economies, and in aids to executive decision. One other engineering contribution which is of vital interest to the Army is the determination of the service life of weapons and the means for minimizing deterioration.

Present method of proof testing.--In addition to the evaluation of weapons conducted at the U. S. Army Proving Grounds, which is an Ordnance establishment, the Continental Army Command (CONARC) supports Test Boards, which evaluate military weapons in the hands of troops. These Boards utilize combat experienced officers to plan, conduct, and evaluate tests. They are supported by technical service advisors. These Boards are equipped with libraries of technical literature, including files of the reports made by National Defense Research Council. The facilities available include: field ranges, maintenance shops and, of course, trained military personnel.

---

(7) Refer to Bibliography: G. I. Butterbaugh, E. B. Kurtz, L. E. Simon.

<u>Lack of civilian counterpart</u>.--Letters of inquiry, addressed to several testing agencies, drew responses which were not very fruitful. A letter from the American Society for Testing Materials, dated August 10, 1953 states in part:

> This (performance testing of finished assemblies) is a field in which our Society has attempted no standization. Each assembly, from past experience, seems to require its own simulated service test. It is very difficult to break down the performance of a product into all the essential properties which influence the end result.

An editorial expresses this modest claim, ". . . test methods for consumers' goods are at an elementary stage of development aside from the work done by Consumers' Research, . ." (8 p 2).

CHAPTER II

ACCURACY

Hit Probability

<u>Weapon system under field conditions</u>.--The capability of a man-weapon system to obtain hits on a field target is an important measure of its effectiveness. Frequently, military standards prescribe minimum per cent hits on a given target at a stated distance. Unfortunately, at least two major blocks exist which hinder direct measurement of per cent hits: the existence of external factors which confuse the response and thus veil the true effects, and difficulties inherent in evaluating the true effects and expressing the conditions under which they apply.

Although military standards specify target distance, service tests require the gunner to fire at several ranges unknown to him -- hence each gunner's ability to determine range and to sense corrections, is introduced. The terrain, target visibility and aiming point clarity may be obscure or deceptive. Frequently, the sample size is inadequate. Other factors which tend to defy description and therefore duplication are: wide range of proficiency among gunners, unequal state of learning with the test weapon between gunners, their physical and mental state at time of firing and a host of intangibles which change during the course of a test and between tests.

Morse and Kimball (6, p 130) are strong proponents of the concept wherein system performance is tested with weapons under normal maintenance, serviced by average crews. However, they are careful to evaluate the nature and distribution of sampling errors involved. In the ensuing example, a direct measure of hits and misses is described, along with the analysis used. Subsequent examples illustrate other means of measuring accuracy, and when feasible, how these measures can be transformed into hit probability.

Example 1. Weapon A.--Four well trained gunners were used to fire two models of Weapon A at standard targets disposed at six distances. Four targets were grouped around each selected range so that the exact distance was unknown to the gunners. Targets were designated at random and each gunner ranged in and fired one shot. He was then assigned a new target at a different range. The total number of shots fired and number of targets hit at each range were recorded.

Table 1. Expert Gunners Firing at Field Targets, Weapon A

Range	Number Shots Fired	Number of Hits	Per Cent Hits
R1	21	19	90.5
R2	24	16	66.7
R3	41	18	43.8
R4	28	7	25.0
R5	14	2	14.3
R6	10	0	0

The progression of per cent hits in Table 1 is consistent with increasing range. Frequently, however, these percentages fluctuate without any apparent trend. In order to establish limits for the true percentage of hits, 90% confidence limits are introduced.

Table 2. Per Cent Hits; 90% Confidence Limits; Trend, Weapon A

Range	% Hits	90% Confidence Limits		Trend *
R1	90.5	76	98	90.5
R2	66.7	51	80	65
R3	43.8	32	56	43
R4	25.0	14	38.5	25.5
R5	14.3	4	34	11
R6	0	0	21	2

In Table 2, the 90% confidence limits were obtained graphically from Table 9b of Dixon and Massey (9, pp 193,321). For computing confidence limits when Np exceeds 5, the same

---

* Values of probability read from trend line in Fig. 1.

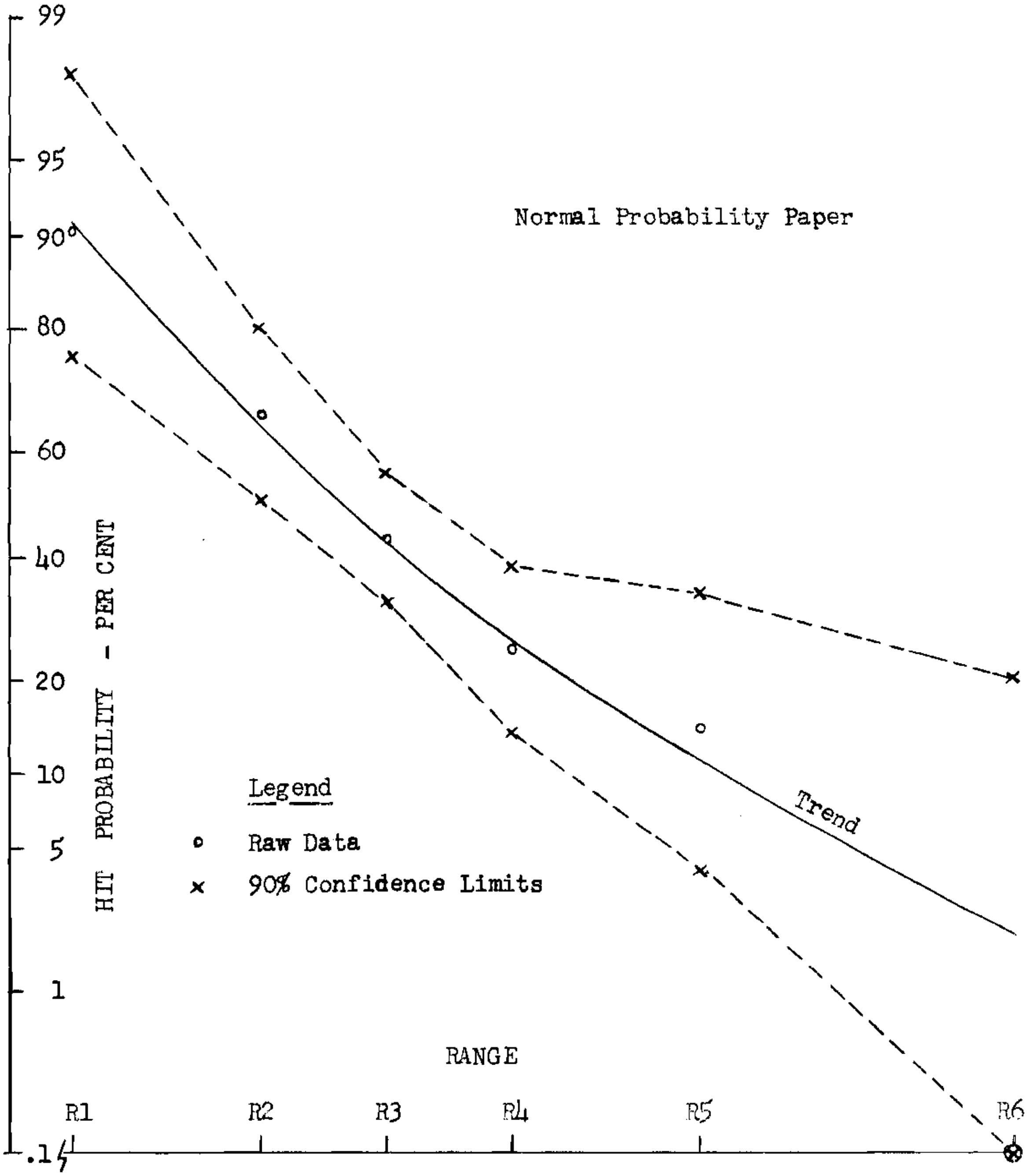

Fig. 1. Field Firing Probability. Weapon A.

authors offer the following normal approximation of the binomial distribution:

$$p \pm z \sqrt{\frac{p\,(1-p)}{N}} \tag{1}$$

where z equals 1.645 for the 90% confidence limit. Therefore, for row R3 in Table 2, we have:

$$0.438 \pm 1.645 \sqrt{\frac{0.438\,(0.562)}{41}}$$

$$= 0.438 \pm 1.645 \quad (0.0773)$$

$$= 0.438 + 0.127 = 0.565 \quad \text{or} \quad 56.5\%, \text{ and}$$

$$= 0.438 - 0.127 = 0.311 \quad \text{or} \quad 31.1\%$$

These values are consistent with those obtained graphically. A plot of the values in Table 2 on normal probability paper permits the drawing of a smooth trend of true hit probability (Fig. 1).

Sample size.--In the example just cited, Weapon A was compared with a military standard. When two weapons are to be compared, it may be desirable to determine whether a difference between their hit proportions attains a significance level of 10%. Hence, some estimate of sample size is needed. We will consider the test weapon superior to the control item. Therefore, the one tail test of the statistic t, may be used.

$$t = \frac{p_1 - p_2}{\sigma_{p_1 - p_2}} \qquad (2)$$

Duncan simplifies this relation by an inverse sine transformation (10, pp 382-3):

$$\phi = 2 \text{ arc sin } \sqrt{p} \text{ degrees and} \qquad (3)$$

$$\sigma_p = \frac{57.3}{\sqrt{N}} \text{ degrees}$$

which transforms Equation (2) into

$$t = \frac{2(\phi_1 - \phi_2)}{57.3\sqrt{\frac{1}{N_1} + \frac{1}{N_2}}} \qquad (4)$$

Letting $N_1 = N_2$, and solving for N_1,

$$N_1 = \frac{1}{2} \left| \frac{57.3\ t}{\phi_1 - \phi_2} \right|^2 \qquad (5)$$

For N of about 120, and a one tail test at 10% level, $t = 1.29$.

$$N_1 = \frac{1}{2} \left| \frac{73.917}{\phi_1 - \phi_2} \right|^2 \qquad (6)$$

For $p_1 - p_2$ the sample size depends on the level of p_1. This is illustrated by the sample values in Table 3 increasing in steps until p_2 equals 0.5, or 45 degrees. The lower half of the table, not shown, is symmetrical with the upper half.

Table 3. Sample Size N for Which $p_1 - p_2 = .1$ is Just Significant at the 10% Level. ($N_1 = N_2$)

p_1	p_2	ϕ_1	ϕ_2	$\phi_1 - \phi_2$	$\frac{73.917}{\phi_1 - \phi_2}$	N	N_1
.90	.80	71.56	63.44	8.12	9.103	83	42
.80	.70	63.44	56.79	6.65	11.115	124	62
.70	.60	56.79	50.77	6.02	12.278	150	75
.60	.50	50.77	45.00	5.77	12.810	164	82

Binomial probability paper provides a much quicker solution of N and is far more flexible in its application (11, pp 191-2). For example, assuming p_1 of .20 and p_2 of .10, Example 1 in Fig. 2 shows their corresponding splits, (80-20) and (90-10). For 90% level of confidence, there is a line drawn at a distance 1.28 σ, or 6.5 millimeters from each split and parallel to it. These parallels intersect at a point (71,12). This corresponds to the desired minimum sample, i.e., (71 + 12 - 1) = 82. This value agrees reasonably well with that of N in the first row of Table 3.

Weapon Dispersion

Measures of dispersion.--The Army uses several measures of dispersion; mean radius for small arms and the probable error for larger caliber weapons. One objection to probable error is that ± 4 PE is designated as a 100% zone. Actually, 1 PE = .6745 σ, hence ± 4 PE = ± 2.7 σ, which includes only

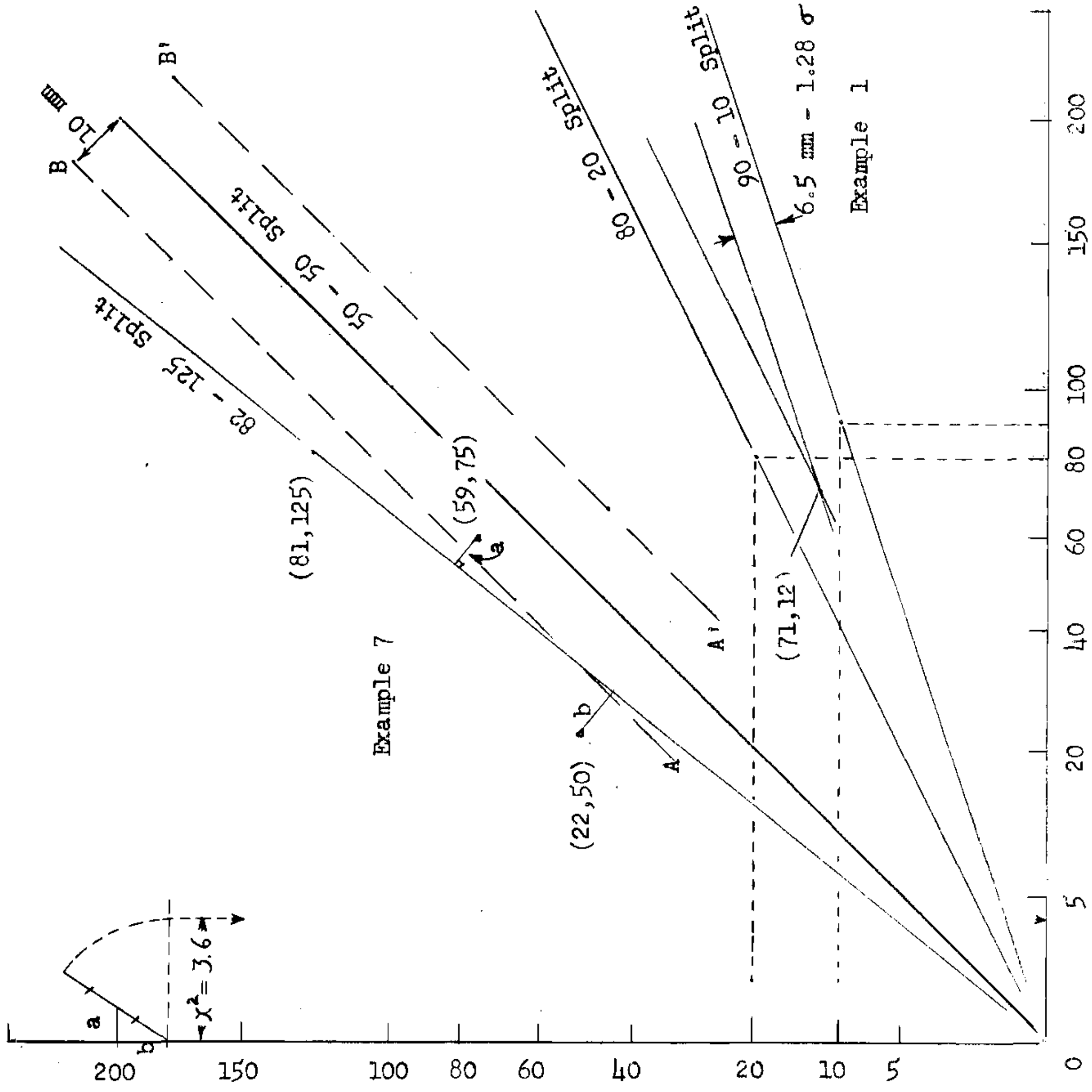

Fig. 2. Binomial Probability Paper.

99.3% of a normal population. Obviously, setting confidence limits or zones of dispersion near 100% is likely to cause confusion.

Mean radius is considered an adequate yardstick of dispersion because it appears appropriate for circular shot groups. Actually, it is troublesome to compute from rectangular coordinates; it has a distinct one-tail distribution with the origin at the center of impact; and the mean radius cannot be resolved into horizontal and vertical components. One final weakness of the mean radius is the prevalent failure to consider the distribution of shot group centers of impact.

Although this study uses the statistic, standard deviation, frequently results are expressed in probable errors. For ease in computing, the former is used exclusively. Since ballistic and other errors have horizontal and vertical components which are generally normally distributed, combining variances or resolving errors lend themselves to the use of standard deviation. Also, the various confidence levels, normal probability functions and other statistics are tabled in terms of standard deviation.

<u>Ballistic error</u>.--If we place a gun on a substantial mount, balance out any non-concentric masses of metal, and fire along a wind-proof fence, the resulting shot group on a

vertical panel yields a measure of ammunition dispersion. This method was originated by F. W. Mann, who made extensive tests with rifles at the turn of the century (12). He found that the deviation of a bullet begins at the muzzle as a result of a number of causes. Several of these causes are listed in the Ordnance Proof Manual (13).

If the gun is placed on its mount, the resulting dispersion includes the effect of the mount on the gun during firing. Now, if a gunner aims and fires each shot, the human error is introduced. From experience, the aiming errors of experts tend to be small, comparable to units of error (UOE) for optical instruments (14). However, when the gun alignment is disturbed at the moment of firing, then dispersion increases markedly.

<u>Human errors</u>.--An appreciably different error is introduced into the system when the gunner attempts to fire at a field target at an unknown range. Here range determination, sight adjustment and other factors take effect. For a single gunner, this might be a characteristic bias, but for a class of well-trained gunners, the bias takes on some distribution. Based on experience, the horizontal and vertical components of human error tend to follow the normal distribution about the point of aim.

Example 2. Dispersion of Two Types of Ammunition, Weapon B.--There are two parts to this example, intended to illustrate two methods of analysis. Due to differences in the physical characteristics of ammunition types K and L, it appears likely that some difference in dispersion may exist. Hence the variances may be subjected to an F ratio test. An analysis of variance may be made using the total shot deviations from the mean points of impact.

Four expert gunners fired a ten-round shot group at a vertical panel, using type K and type L ammunition in turn. Two identical guns were used so that both types were fired at one time. The coordinates of each point of impact were measured.

Table 4. Vertical Distribution of Shots, Ammunition K and L.

	Ammunition K				Ammunition L			
Shot	Gunner: A	B	C	D	A	B	C	D
1	1	55	28	12	29	0	43	65
2	33	20	61	0	0	10	40	57
3	22	24	32	45	7	40	70	48
4	0	0	14	8	31	51	40	4
5	38	29	1	54	53	49	21	11
6	33	50	26	26	52	60	54	60
7	17	31	40	55	41	39	60	74
8	27	21	29	24	42	3	0	0
9	88	8	12	25	27	65	88	63
10	37	25	0	21	50	31	32	31
Total	296	263	243	270	332	348	448	413
SS	5476	2516	3082	3202	3016	4908	5604	6864
s^2	608	279	342	356	335	545	622	762

<u>Analysis I</u>. Pooled Variances. Fisher's F Ratio.

Variances for each weapon type may be pooled, providing they are homogeneous.* Dividing the larger variance by the smaller yields a computed F ratio which can be compared with tables for the F statistic.

$$S_K^2 = SS/DF \tag{7}$$

$$= 14{,}276 \;/\; 36 \quad = \quad 396.5$$

$$S_L^2 = 20{,}392 \;/\; 36 \quad = \quad 566$$

$$F = S_L^2 \;/\; S_K^2 \tag{8}$$

$$= 566 \;/\; 396.5 \quad = \quad 1.42$$

$$F_{05} \text{ for DF} = 4 \times 9, \; (36,36) \quad = \quad 1.75$$

The computed F does not exceed the table value, hence there is not sufficient evidence to suspect a difference between the two types of ammunition at the 5% level of significance.

<u>Analysis II</u>. Analysis of Variance (15, pp 51, 86).

Measurement data as in Table 4 may be analyzed by still another method. This procedure resolves the total variance into components due to one or more variables and a residual error. The residual is used to test the significance of main effects.

---

* These variances have been subjected to the Bartlett Test and found homogeneous.

Table 5. Sums of Vertical Measurements

Type of Ammunition	Gunner: A	B	C	D	Total
K	296	263	243	270	1,072
L	332	348	448	413	1,541
Total	628	611	691	683	2,613

Each number of Table 5 corresponds to the sum of measurements in Table 4. For the calculations which follow, the rule for the divisor is: the number of individuals in a set such as a row, column, or grand total, times the number of measurements which it represents. Here ten rounds were measured for each gunner.

(1) Square each shot group sum (individual), and total:

$$1/10\ (296^2 + 263^2 + \text{- -} + 448^2 + 413^2) = 89{,}133$$

(2) Square both row totals and divide their sum by the number of individuals per row:

$$1/40\ (1072^2 + 1541^2) = 88{,}096$$

(3) Square column sums and divide their sum by the number per column:

$$1/20\ (628^2 + 611^2 + 691^2 + 683^2) = 85{,}584$$

(4) Correction factor, "C": Square the grand total and divide by the total number of original individuals:

$$C = 2613^2/80 = 85{,}347$$

(5) Square each individual in Table 4 and total:

$(1^2 + 55^2 + - - + 32^2 + 31^2) = $ 123,801

Assemble these results as indicated in Table 6.

Table 6. Analysis of Variance

Source of Variance	Sum of Squares		DF	Mean
Between Rows (Ammunition)	(2) -C	2,749	1	2,749
Between Columns (Gunners)	(3) -C	237	3	79
Interaction (A X G)		800	3	266
Subtotal	(1) -C	3,786	7	
Individual Error (Residual)		34,668	72	480
Total	(5) -C	38,454	79	

All the sums of squares are thus accounted for except those for interaction and individual errors. These are obtained by subtraction. Note that the sum of squares for individual error is equal to the sum of the deviations for ammunition K and L in Equation (7):

$$14{,}276 + 20{,}392 = 34{,}668$$

The degree of freedom (DF), in Table 6, is set by one less than the number of levels in each variable. Thus for four gunners, the DF is (4-1), or 3. For interaction, the DF is the product of the DF's for ammunition and gunners.

The mean square for individual error is used to test interation and the two main effects. By inspection, none but the effects due to ammunition appear significant.

$$F = 2749/480 = 5.727$$

$$F_{05} \text{ for DF } (1, 72), = 3.98$$

Hence, there is evidence of a significant difference between ammunition types K and L at the 5% level of significance.

Example 3. Regression, Dispersion at Several Ranges.--
This is a continuation of Example 2, except that: two gunners were used instead of four, and firing was conducted at six target distances. Dispersion, in terms of vertical and horizontal probable errors, was computed for each shot group and combined.

Table 7. Probable Errors of Ammunition K and L at Six Ranges

Range	Ammunition K Horizontal	Ammunition K Vertical	Ammunition L Horizontal	Ammunition L Vertical
5	6.7	6.5	9.8	10.8
8	12.3	10.0	14.7	16.3
10	15.3	13.5	17.0	13.9
10		15.3		15.9
12	22.2	15.8	21.7	23.6
15	29.4	27.5	28.6	31.2
20	35.6	46.5	34.2	40.0

The replication at Range 10 was combined into one estimate of dispersion for horizontal components, hence the blank spaces. A plot of the respective components against range does not permit a good estimate of the trend. In order to obtain the best estimate of each trend, it appeared advisable to obtain a fit by using least squares.

Snedecor gives a detailed explanation of linear and curvilinear regression and its implications (16, pp 103, 374).

Table 8 is arranged so as to provide the values needed to solve the constants a and b in the linear equation:

$$X = a + bx \tag{9}$$

Table 8. Linear Regression of Horizontal Probable Errors

Range Code R	PE X	Deviations from Mean r	x	Ammunition K r^2	rx	Ammunition L PE X	x	rx
0	.3	-10	-17.1	100	171.0	0	-18.0	180
5	6.7	- 5	-10.7	25	53.5	9.8	- 8.2	41
8	12.3	- 2	- 5.1	4	10.2	14.7	- 3.3	6.6
10	15.3	0	- 2.1	0	0	17.0	- 1.0	0
12	22.2	2	4.8	4	9.6	21.7	- 3.7	7.4
15	29.4	5	12.0	25	60.0	28.6	10.6	53
20	35.6	10	18.2	100	182.0	34.2	16.2	162
70	121.8	0	0	258	486.3	126.0	0	450
10	17.4	Mean				18.0		

The linear regression formula is,

$$X_K = \overline{X} + \frac{\sum (rx)}{\sum r^2} (R - \overline{R}) \tag{10}$$

$$X_K = 17.4 + 486/258 \; (R - 10)$$

$$= -1.4 + 1.88 \; R$$

$$X_L = 18 + 450/258 \; (R - 10)$$

$$= 0.5 + 1.75 \; R$$

These regression lines are plotted in Figure 3.

Analysis III.--An additional component of range (R^2), may be added to linear regression to establish curvilinear regression. This method is applied to the vertical probable errors.

Table 9. Curvilinear Regression. Vertical PE's of Ammunition K

R	R^2	Y	Deviation r_1	r_2	y	Products r_1r_2	r_1y	r_2y	Squares r_1^2	r_2^2
0	0	0	-10	-132	-17	1320	170	2244	100	17,424
5	25	6.5	-5	-107	-10	535	50	1070	25	11,450
8	64	10.0	-2	-68	-7	136	14	476	4	4,613
10	100	13.5	0	-32	-3	0	0	96	0	1,024
10	100	15.3	0	-32	-1	0	0	32	0	1,024
12	144	15.8	2	12	-1	24	-2	-12	4	144
15	225	27.5	5	93	10	465	50	930	25	8,650
20	400	46.5	10	266	29	2660	290	7714	100	70,755
80	1058	135.1	0	0	0	5140	570	12550	258	115,084
10	132	17	Mean			A	B	C	D	E

The curvilinear regression equation is,

$$Y = a + b\,(R-\bar{R}) + c\,(R^2 - \overline{R^2}) \qquad (11)$$

where:

$$a = \bar{Y} = 17$$

$$\bar{R} = 10 \qquad \text{and } \overline{R^2} = 132$$

$$b = \frac{B\,E - A\,C}{D\,E - A\,A} = \frac{511 \times 115{,}084 - 5140 \times 12{,}550}{258 \times 115{,}084 - 5140 \times 5140}$$

$$= \frac{11{,}000}{35{,}064} = 0.315$$

$$c = \frac{C\,D - A\,B}{D\,E - A\,A} = \frac{12550 \times 258 - 5140 \times 570}{35{,}064}$$

$$= \frac{3{,}332}{35{,}064} = 0.095$$

Substituting, we have:

$$Y_K = 17 + 0.315\ (R - 10) + 0.095\ (R^2 - 132)$$

$$= 0.15 + 0.315\ R + 0.095\ R^2$$

Following the same procedure with ammunition L, we obtain:

$$Y_L = 1.5 + 1.48\ R + 0.0233\ R^2$$

The advantages of regression lines become apparent when the curves are drawn as in Figure 3. Regression lines permit the drawing of the best fitting curves to data and are superior to approximating lines by eye. There is a continuous estimate of the dependent variable over the entire range of the independent variable. The behavior of weapon dispersion is presented graphically and analytically. A regression line draws its power from the combined weight of several samples. Comparison between two weapon system dispersions may be presented graphically or analytically. In Figure 3, there is little or no difference in the horizontal components of ammunition K and L dispersions. But in vertical dispersion, ammunition K is obviously superior to ammunition L from zero range to R5.

Analysis of weapon system dispersion as a function of range provides the maximum of information and is economical from the point of view of ammunition expenditure. Samples may be obtained at convenient levels of the independent variable, or be spaced at equal intervals. Since

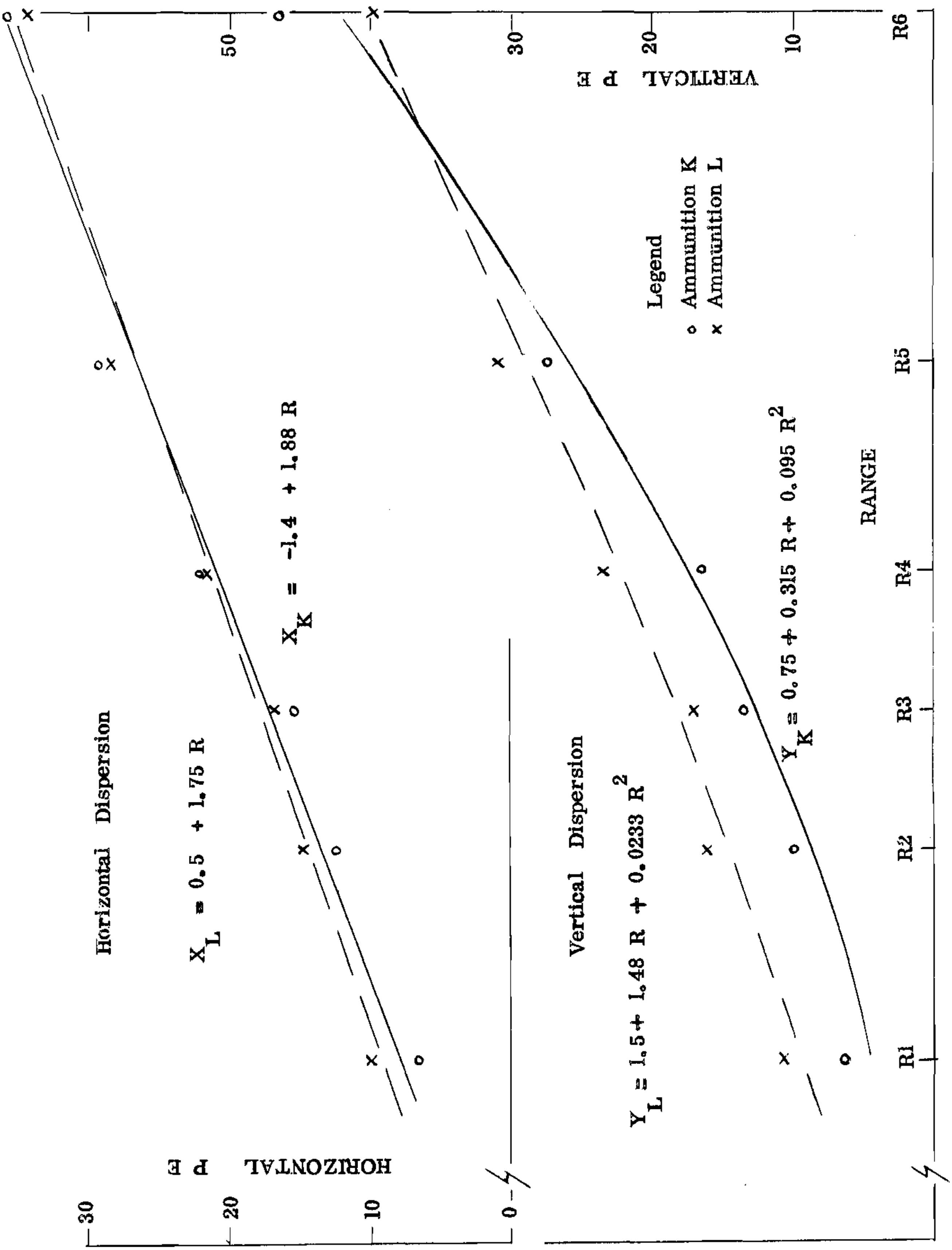

Fig. 3. Regression of Probable Errors. Weapon B.

regression lines smooth out fluctuations in sample yields, final presentation of results should use values established by regression lines.

Relation Between Dispersion and Hit Probability

Conversion of dispersion into hit probability.--Points of impact on a vertical target are generally recorded as coordinate points with respect to an arbitrary origin or to an aiming point. Horizontal ballistic dispersion is the deviation of these points from the center of impact ($\bar{X}$, $\bar{Y}$), and the vertical ballistic dispersion is the vertical deviation from the center of impact. Generally, each component population is normally distributed about the mean and each component is independent of the other. The horizontal and vertical components of dispersion are not necessarily equal. Horizontal standard deviation and probable error are expressed by:

$$\sigma_X = \sqrt{\frac{\sum(X - \bar{X})^2}{n - 1}} \tag{12}$$

$$PE_X = 0.6745\ \sigma \tag{13}$$

where n is the number of shots in a shot group.

When several shot groups have homogeneous variances, a pooled estimate of the variance is given below:

$$\sigma = \sqrt{\frac{\sum_{1}^{k} n_i s_i^2}{\sum_{1}^{k} n_i}} \tag{14}$$

where k is the number of shot groups.

When the outline of a rectangular target w units wide and h units high is placed over a shot group so that the target center is over the center of impact, we obtain an estimate of the proportion of hits falling inside the target, or hit probability. An estimate of the same hit probability may be obtained by determining the area under the normal curve defined by the horizontal dispersion error and target width, and a corresponding area for the vertical direction, respectively. The statistic Z, used to enter a table of areas under the normal curve, is equal to X/σ, (17, p 114). Here X is equal to half the width of the target. Thus, for the horizontal component of dispersion:

$$Z_X = \pm \frac{1/2 \ w}{\sigma} \tag{15}$$

Similarly, a value of Z_Y for the vertical component of dispersion and target height may be obtained. Each Z thus determined establishes the respective component probabilities, p_X and p_Y, of falling within these limits. Since the component distributions are independent,

$$p = p_X \ p_Y \tag{16}$$

where p is the probability of hits on a given rectangular target.

The case just described, applies to the condition when the mean point of impact and center of target coincide. A second case applies to aiming errors. Aiming errors are comparable to ballistic errors, but their distribution is about the aiming point. Hence, in the total distribution of shots fired by a number of gunners, it is obvious that individual shots are distributed about mean points of impact, and the MPI's are distributed about the aiming point. In this study the aiming point is always at the geometric center of the target.

Nomograph, per cent hits - target size - probable error.-- For all ballistic cases in which the distribution of component errors may be assumed normal, there are three related variables: per cent hits, target size, and dispersion probable error, and it is possible to solve for any one unknown variable when the other two are known. A nomograph can be readily produced to relate these three variables (18).

Let us choose a variable, such as hit probability, and select a nominal value, say $p = 0.90$. Assume that the component probabilities, p_X and p_Y, are equal, and Equation (16) becomes:

$$p = p_X^2 \quad \text{and} \tag{17}$$

$$p_X = \sqrt{p}$$

$$p_X = \sqrt{0.90} = 0.948$$

From the normal distribution area table, we extract the $Z = \pm 1.94$ corresponding to an area of 0.948. Then if $X = 50$, $\sigma = X/Z = 50/1.94$, or 25.8 units. Thus for 90% of a shot group hitting a square target 100 units on a side, the component of dispersion in terms of standard deviation, σ, is 25.8 units. In terms of probable error, the estimate is 0.6745 (25.8), or 17.4 units. Refer to the dotted line in Figure 4.

The nomograph shown in Figure 4 was constructed on three-cycle, semi-logarithmic paper. The scales were aligned so that a PE of 50 units, a target size of 100 units, and the nominal probability of 0.25 are on the same horizontal line. Note the 1:2 ratio of PE's and target sizes along straight lines radiating from the nominal probability of 0.25. The left scale was made identical with the logarithmic scale. These known points on the left scale, connected by straight lines with the nominal probability of 0.25, made it a simple matter to identify the points along the target scale. Now, with two scales identified, it was easy to align the computed PE's and target sizes with nominal probabilities to fix the last scale.

Note that a distinction must be made between nominal probability and component (X or Y) probability. The former expresses the proportion of hits falling between two sets of target limits, whereas the latter yields the proportion of hits between a set of limits in one direction only. Now, if, beside the scale of nominal p in Figure 4 there were also a scale of square roots of p, then the nomograph could be used for relating component PE's, target widths and component probabilities. An example of this method will help clarify such use.

Supposing we had a known ballistic error with $PE_X = 17.4$ units and $PE_Y = 30$ units. We wish to determine the hit probability on a target 100 units wide and 60 units high. Aligning the X components first, the p scale reads 0.90. Extracting the square root, this becomes 0.95. In a similar manner, the Y component probability is $\sqrt{0.25}$, or 0.50. Note that the target height just equals $\pm 1\ PE_Y$. Thus by definition, 50% of the vertical dispersion falls within these limits. In accordance with Equation (16), the resulting hit probability p, is 0.95 (0.50), or 0.475.

We can now take the dispersions of Weapon A and convert them into hit probabilities on a target of given dimensions. Method (2) in Figure 4 was used in determining the hit probability of Weapon A because the two component dispersions were not equal. The results are plotted on the graphs in Figures 5 and 6*. Note the divergence between ballistic capability of the system and performance of expert gunners in the field.

* For reasons of security, neither target size nor units of dispersion may be disclosed. Hence the data is not tabulated. For our purposes, a target of 100 units is a suitable approximation.

RESTRICTIONS

(1). For $PE_X = PE_Y$ and Target Height = Width:

Align any two known variables with a straight line to obtain the third (unknown) variable.

(2). For $PE_X \neq PE_Y$ or Target Height $\neq$ Width:

Extract square root of probability (expressed as a decimal). PE_X (or PE_Y) is for one component, and target limits w (or h) are in one dimension. To find total p:

$$p = p_X \, p_Y$$

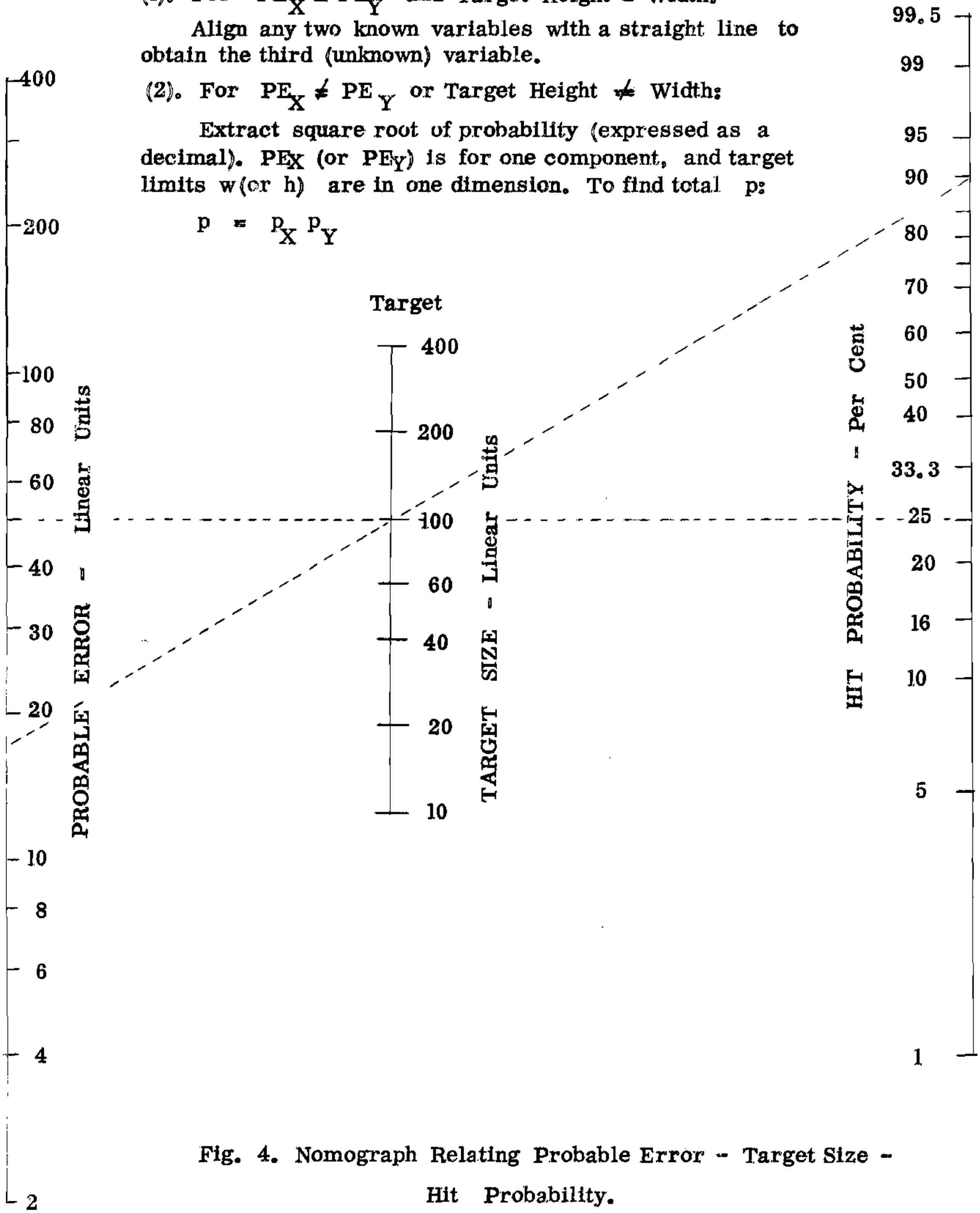

Fig. 4. Nomograph Relating Probable Error - Target Size - Hit Probability.

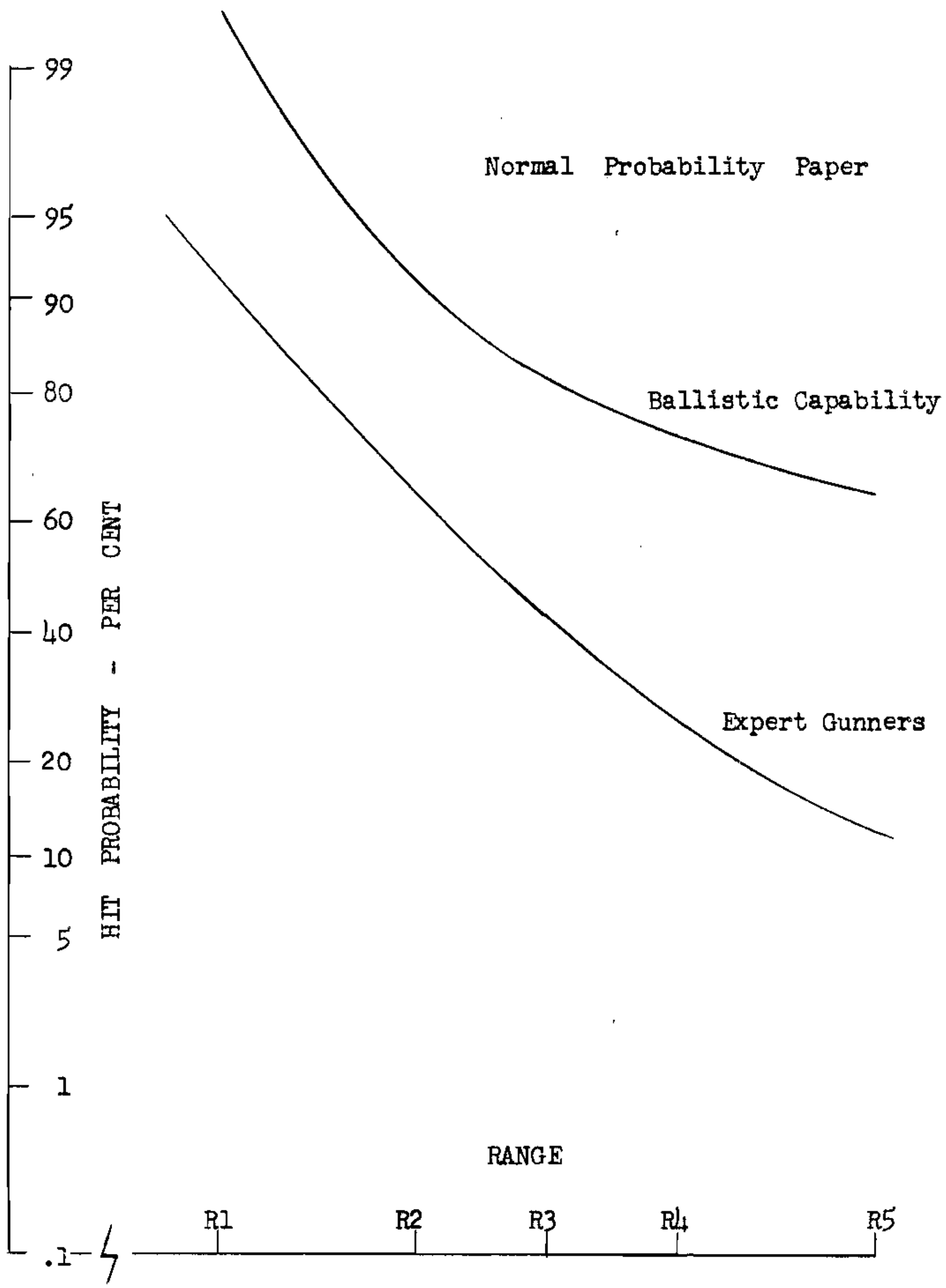

Fig. 5. Ballistic and Field Firing Probability. Weapon A.

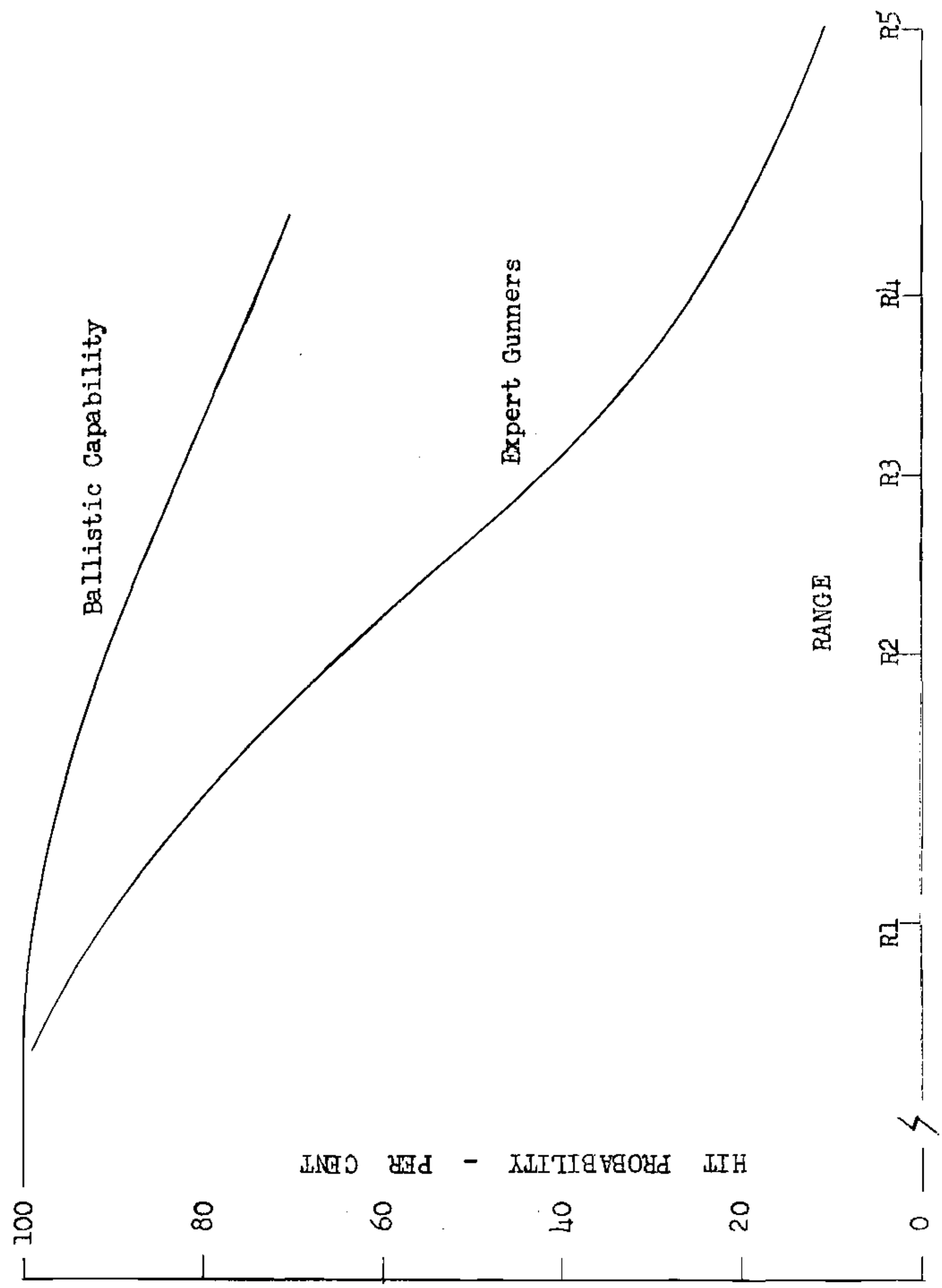

Fig. 6. Field Firing Probability for Experts, Weapon A.

CHAPTER III

THE HUMAN ERROR IN GUNNERY

Nature of distribution of errors.--At known distances, where the firers have ranged in on the target, the distribution of individual shots is known to be normal about the point of aim. Since ballistic dispersion is normally distributed about the center of impact, the distribution of mean points of impact about the aiming point is therefore also normal. When targets are disposed at unknown distances, the increased incidence of overs and shorts reflects the effect of range estimation errors. The standard deviation of visual range estimation errors is around 20% of the range, and the mean may have a small bias, depending on the nature of the terrain and visibility. Lateral errors are proportional to target distance. Except for gross errors, it may be assumed that the total distribution of human errors forms an oval pattern centered on the target, with the range axis longer than the lateral axis. Both components are normally distributed, or nearly so, about the point of aim.

With the assumption of normality and independence, error variances are additive, and we can thereby combine or resolve component errors. Let S_h be the human error, and the subscripts "t" and "b" represent total and ballistic sets. Then we have:

$$S_t^2 = S_b^2 + S_h^2 \quad \text{or,} \quad S_h^2 = S_t^2 - S_b^2 \qquad (18)$$

Example 4. Experts Firing Weapon A.--With the aid of the nomograph in Figure 4, the field firing probability in Figure 6 was converted into total probable errors. Based on observation of strikes, the horizontal and vertical components were assumed equal. The values of total error at four ranges are given in Table 10. Each of two expert gunners fired three ten-round shot groups with two weapons and the resulting ballistic probable errors are included in Table 10. Estimates of the human error may be obtained by using Equation (18), or graphically, from the relationship between two sides of a right triangle and the hypotenuse. A graphical solution is given in Figure 7, where OF = BH is the hypotenuse and OB and OH are the two perpendicular arms. The human errors were thus scaled off at convenient intervals and plotted as shown by the dotted line in Figure 7.

Table 10. Total, Ballistic and Human Probable Errors

Range	Component	PE_t	PE_b	PE_h	$\frac{PE_h}{PE_b}$	$\left\lvert\frac{PE_h}{PE_b}\right\rvert^2$
R1	X	15.0	8.7	12.2	1.40	1.96
	Y		10.2	11.2	1.10	1.21
R2	X	23.6	13.4	19.6	1.46	2.13
	Y		15.2	18.4	1.21	1.46
R3	X	31.3	16.5	26.7	1.62	2.62
	Y		17.9	25.8	1.44	2.07
R4	X	42.5	19.5	37.8	1.94	3.76
	Y		20.1	37.5	1.87	3.39

Table 10 lists the magnitudes of the total, human, and ballistic errors and provides an estimate of the weapon system performance in the hands of expert gunners. In an ideal system, the human and ballistic errors would be equal. Here the human-ballistic variance ratio (last column in Table 10), reaches a significant difference at about 2.1. The last column in Table 10 shows that the variance ratio for the X component reaches significance before the Y component, at range R2 and R3, respectively. Thus an effort to reduce the lateral error by the gunners at these two ranges should yield an appreciable gain in hit probability. A glance at Figure 7 reveals that total and human errors increase exponentially with range. In fact, it is evident that the human error governs system performance, at least for ranges R3 and over.

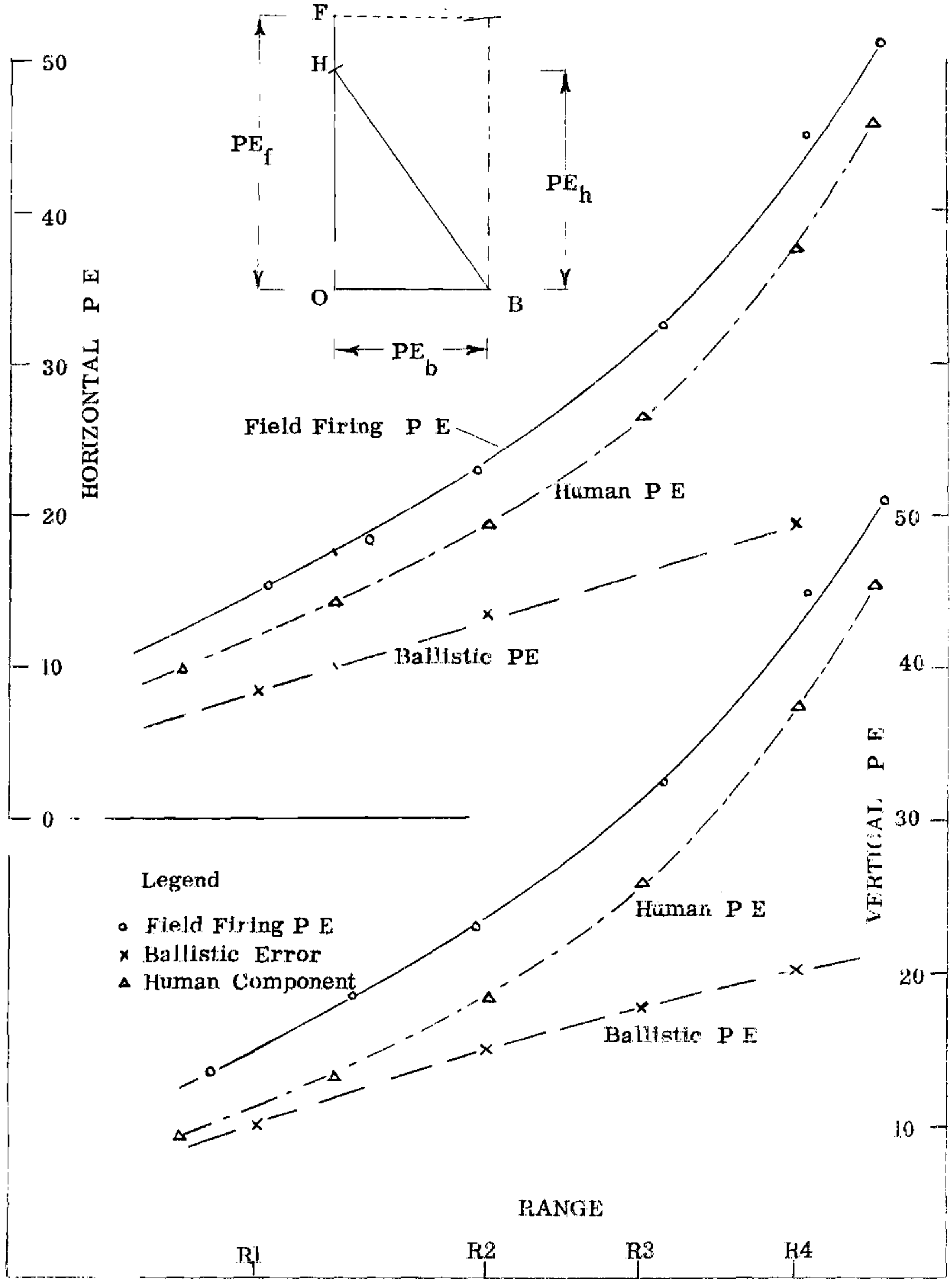

Fig. 7. Total Errors Resolved Into Human and Ballistic Errors. Weapon A.

Analysis I.--From the foregoing, it appears that ballistic dispersion is merely one factor to be considered in evaluating weapon performance. For Weapon A, there is a fair balance between human and ballistic errors only up to range R2. Beyond that range, performance response depends upon the human error. Therefore, any comparison of two or more models of Weapon A family, should be based on total system performance or on the magnitude of the human errors involved. It is obvious that ballistic dispersion should be kept small, but more emphasis should be on reducing the magnitude of the human errors.

Analysis II.--The term "human error" as used so far, has been at times mistaken for the term, "aiming error." Actually, the latter is merely a component of the former. A more restrictive definition identifies it as the accuracy with which a gunner can realign the sight axis of his gunsight with the line to the aiming point. The aiming error is comparable to the unit of error (UOE), used in determining the resolving power of an optical instrument.

In order to establish the magnitude of aiming error, three experts made three trials with a fixed sight comparable to the one on Weapon A. In each trial, a movable target was adjusted to intersect the sight axis, and the gunner then made four attempts to have the target realigned with the original point. The total number of measurements was 36. The computed aiming errors are listed in Table 11.

Table 11. Aiming Errors with Weapon A Type Gunsight

Range	Comp	PE_h	PE_a	$\frac{PE_h}{PE_a}$	$\left\lvert\frac{PE_h}{PE_a}\right\rvert^2$
R1	X	12.2	0.95	12.8	164
	Y	11.2	0.92	12.2	149
R2	X	19.6	1.30	15.0	225
	Y	18.5	1.30	14.2	202
R3	X	26.7	1.54	17.4	303
	Y	25.8	1.56	16.8	282

The most notable fact about the aiming error in Table 11, is that the magnitude is very small. Relative to the human error, it is almost negligible and was therefore not plotted on a graph.

Example 5. Errors of Experts and Marksmen with Weapon C.--
The test described below was performed by a military agency, and is included here because it is one of very few experiments in which deviations of all shots from a field target were measured. The purpose of this experiment was to evaluate expert and average marksmen with Weapon C. The data, suitably coded, is given in Table 12.

Table 12. Field Firing Errors of Experts and Marksmen

Range	Experts P	PE_X	PE_Y	Marksmen P	PE_X	PE_Y
R1	88	3.0	4.3	80	5.5	6.4
R2	60	10.0	7.8	47	13.2	9.5
R3	44	7.4	8.8	27	12.8	11.0
R4	37	12.8	9.8	16	17.2	14.2

In order to evaluate the ballistic error, two experts fired four ten-round shot groups, using Weapon C fixed in an appropriate bench rest to minimize extraneous vibrations. The weapons were first ranged in on targets at known distances B, C and E. Probable errors were computed from the measurements taken at each target, and listed in Table 13.

Table 13. Ballistic Error of Weapon C Fired from Bench Rest

Component	Target Distance B	C	E
PE_X	1.9	2.8	4.35
PE_Y	1.8	3.28	6.4

The total and ballistic errors are plotted in Figure 8.

Analysis I.--The data in Table 12 indicates, by an irregular progression of the horizontal probable errors, that an external disturbing factor was present. Although the cause was not determined, this phenomenon occurs fairly often and may be attributed to deceptive appearance of the target due to unequal target contrast. These irregularities were arbitrarily smoothed out by straight or curved

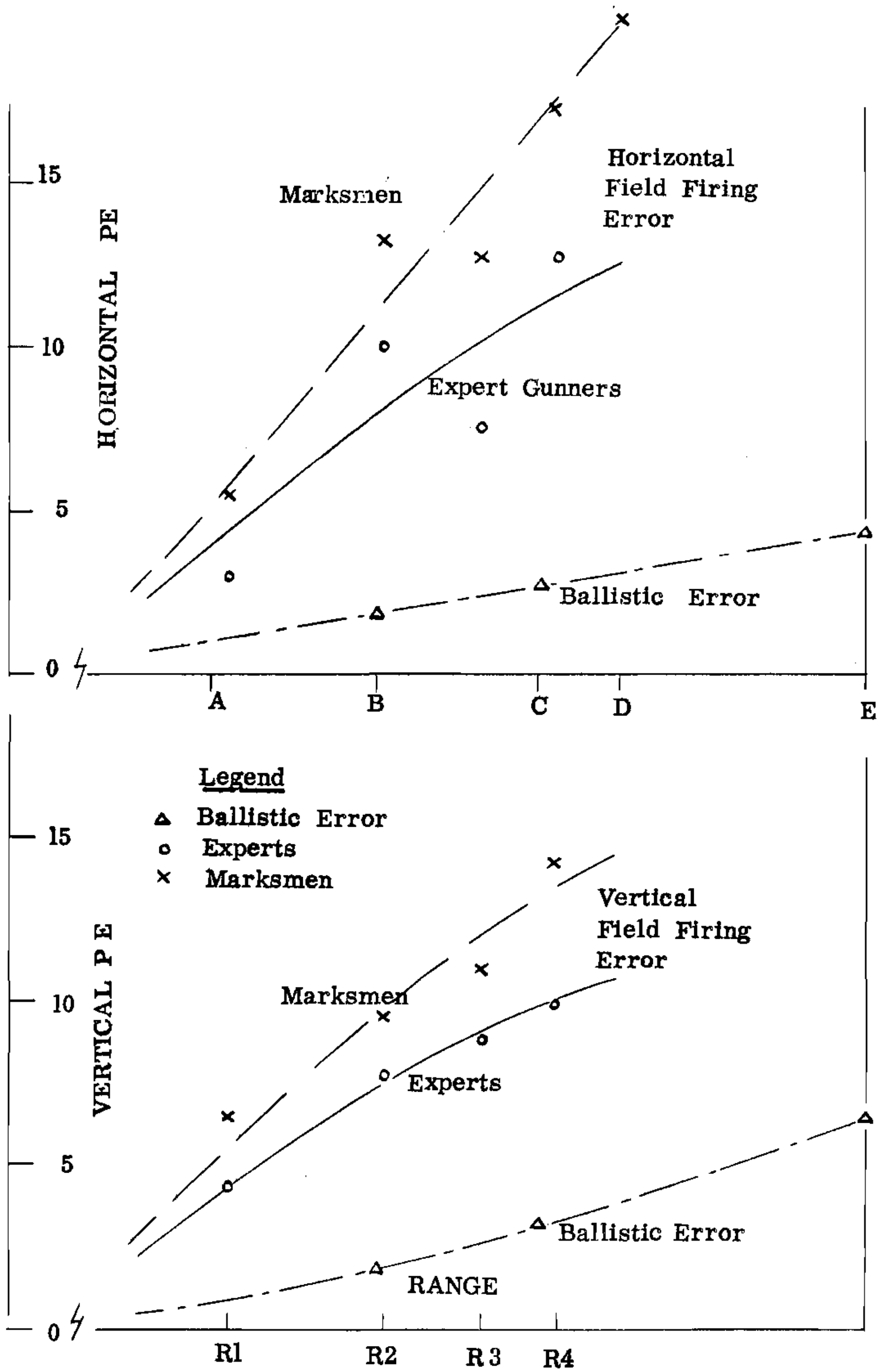

Fig. 8. Field Firing Errors and Ballistic Errors. Weapon C.

lines drawn in Figure 8. Separation of the human error was done graphically at convenient points, and the human error alone plotted in Figure 9. In the table below, the ballistic error provides a common measuring stick for the human errors of experts and marksmen, which were scaled from Figure 9.

Table 14. Comparison of Human Errors, Experts and Marksmen

Range	Comp	PE_h	PE_b	$\frac{PE_h}{PE_b}$	$\left\|\frac{PE_h}{PE_b}\right\|^2$
EXPERTS					
A	X	4.0	1.0	4.0	16.0
	Y	4.0	0.8	5.0	25.0
B	X	7.8	1.9	4.1	16.8
	Y	7.1	1.8	3.95	15.6
C	X	10.8	2.6	4.15	17.2
	Y	9.4	3.1	3.05	9.3
D	X	12.0	2.85	4.2	17.6
	Y	10.0	3.85	2.6	6.7
MARKSMEN					
A	X	5.0	1.0	5.0	25.0
	Y	5.0	0.8	6.2	38.4
B	X	10.9	1.9	5.7	32.5
	Y	9.5	1.8	5.3	28.1
C	X	16.7	2.6	6.4	40.9
	Y	12.7	3.1	4.1	16.8
D	X	19.5	2.85	6.8	46.2
	Y	14.0	3.85	3.6	13.0

In the last column of Table 14, all the human-ballistic variance ratios are obviously significant. Their values are much higher than those in Table 10. The horizontal ratios for marksmen are the largest. In Figure 9,

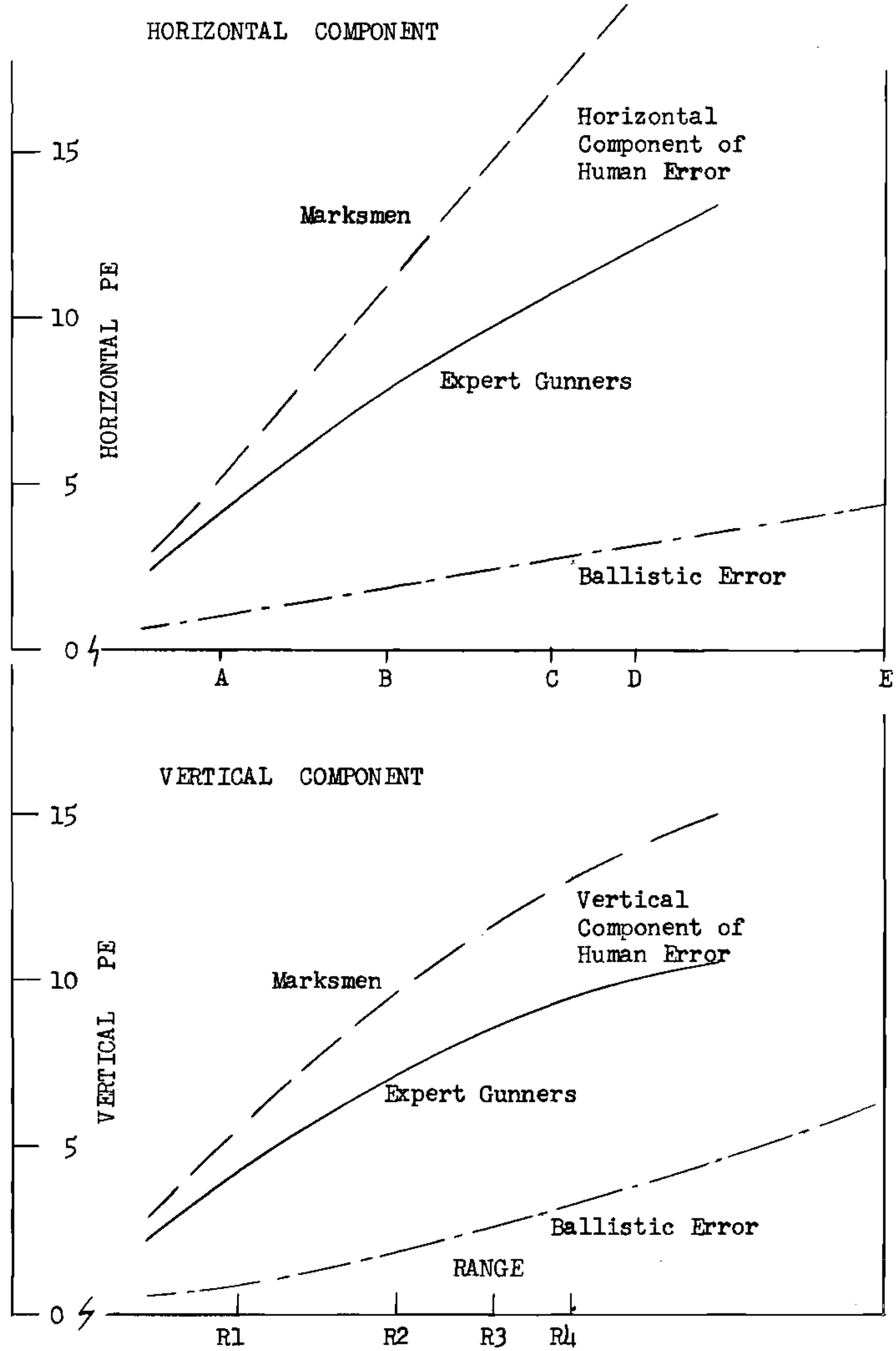

Fig. 9. Human Errors Versus Ballistic Frrors. Weapon C.

the curve of horizontal human errors has the steepest slope and contains the highest values. Thus it appears that marksmen suffer the greatest loss in hit probability due to excessive lateral dispersion. For targets which are tall and narrow, this factor weakens marksmen's performance even more than the figures or the curve indicate. On the other hand, the vertical component of human errors for experts indicates a better trend (Fig. 9).

Analysis II.--The question may arise whether any difference between weapon types C and A may be assigned to differences in their sights. In order to determine the aiming error with Weapon C gunsight, a test was performed similar to the one made with Weapon A sight (Table 11). Three experts performed the experiment with the resulting aiming errors listed in Table 15 and plotted in Figure 10.

Table 15. Aiming Errors with Weapon C Gunsight

Range	Comp	PE_h	PE_a	$\frac{PE_h}{PE_a}$	$\left\lvert\frac{PE_h}{PE_a}\right\rvert^2$
A	X	4.0	1.0	4.0	16
	Y	4.0	.3	13.3	177
B	X	7.8	1.1	7.1	50.4
	Y	7.1	.52	13.6	185
C	X	10.3	1.2	8.6	74
	Y	9.4	.7	13.6	185

In Table 15, the human error-aiming variance ratios for experts are significant at all ranges. In Figure 10,

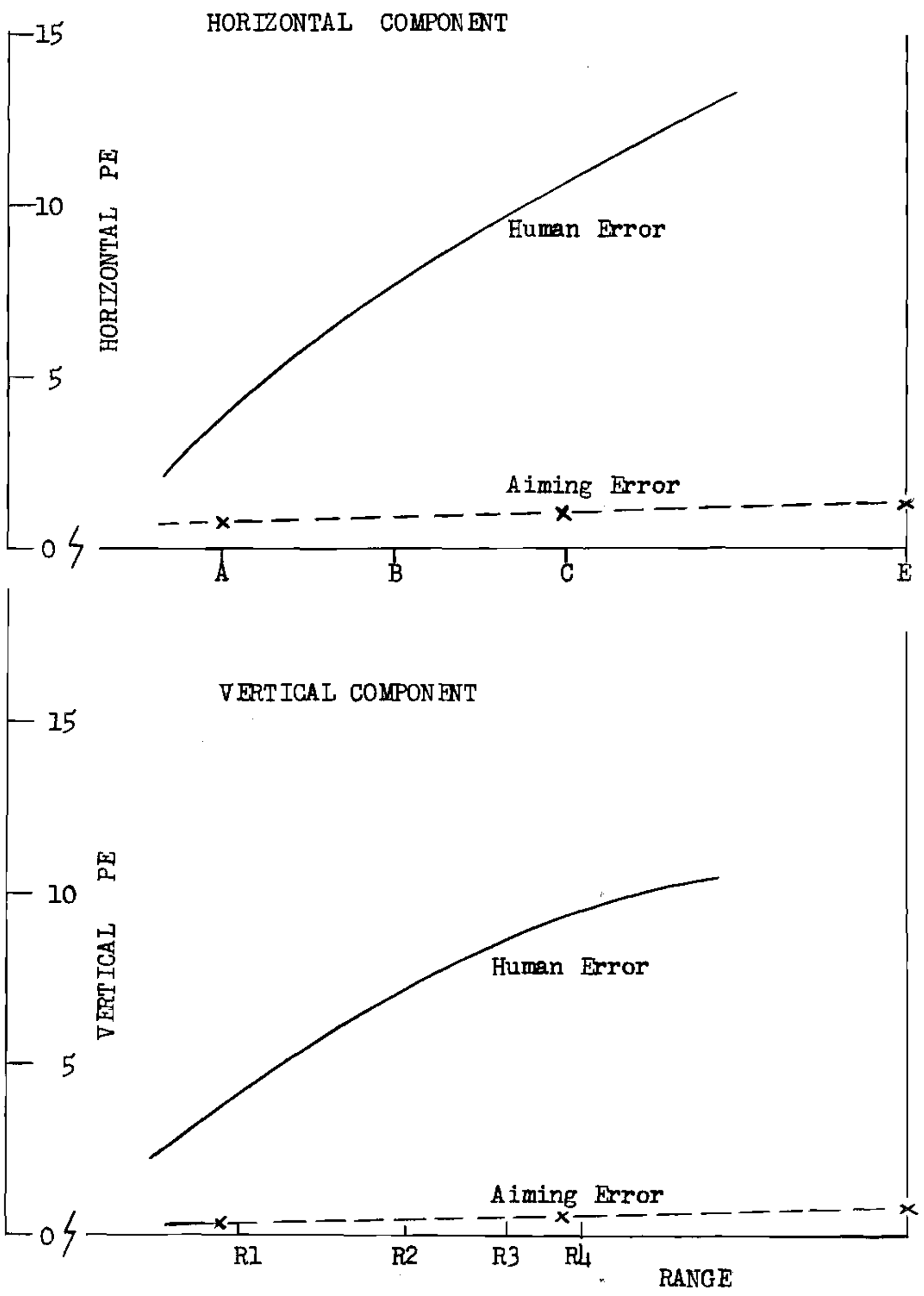

Fig. 10. Aiming Versus Total Human Errors for Experts. Weapon C.

the aiming error is obviously negligible. It is therefore evident that human errors are not aiming errors, but that other significant effects are present during field firing.

Among expert shots at known distance ranges (where the firer can range in prior to a match), there is a saying, "What you can see, you can hit." Evidently, the small errors associated with alignment of gunsights, tend to support this saying.

For a final comparison of experts and marksmen firing Weapon C at field targets, the total errors in Figure 8 were converted into probability using the nomograph (Fig. 4), and the results plotted in Figure 11. A plot of actual per cent hits indicates a reasonably close agreement, particularly for experts. This agreement suggests that the nomograph is usable, and reasonably accurate.

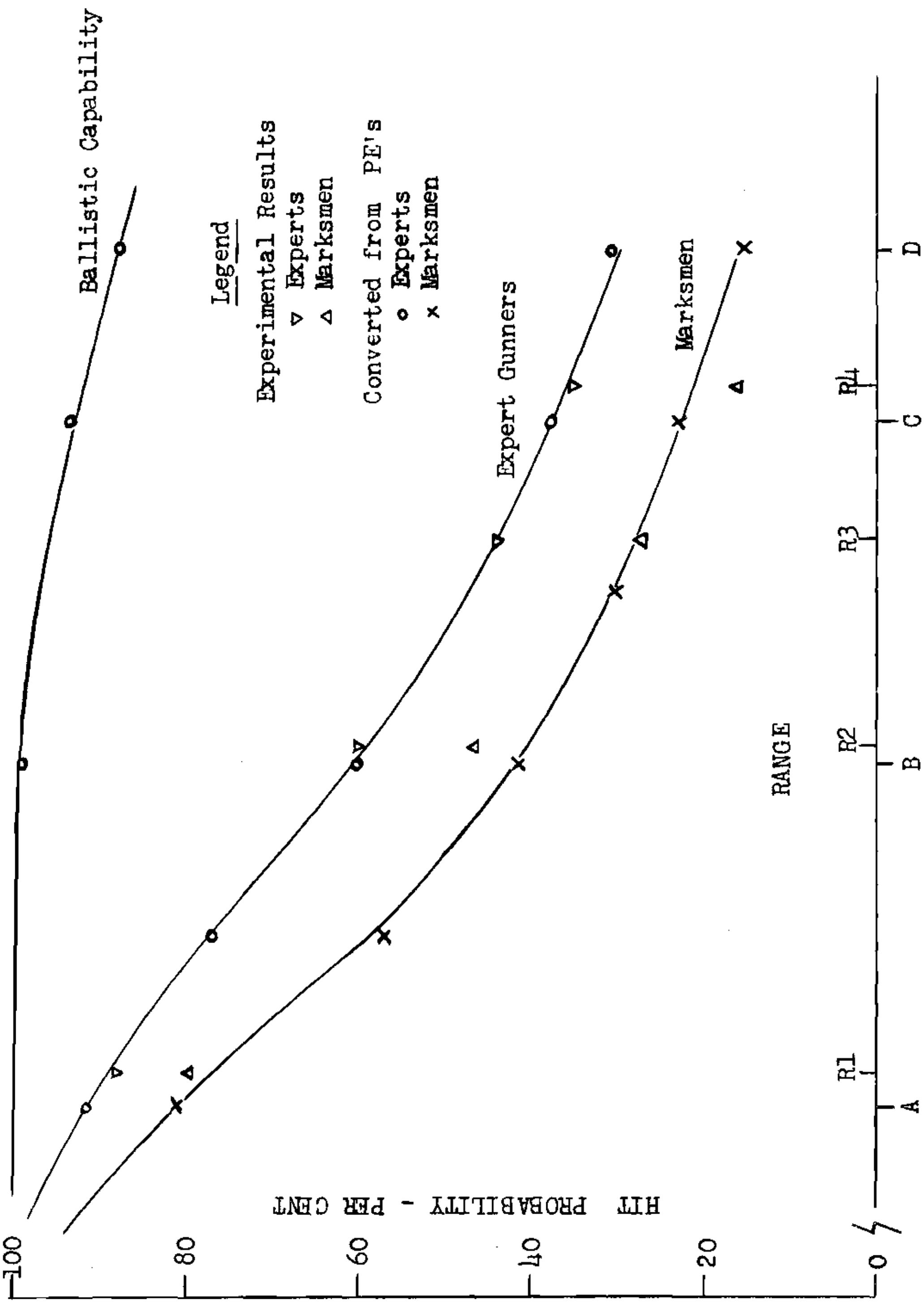

Fig. 11. Hit Probability for Experts and Marksmen With Weapon C. Actual and Converted from PE's.

CHAPTER IV

RELIABILITY

Counts of malfunctions.--In service tests of weapons, the term, "reliability" is associated with the occurrence of malfunctions, that is, failure to fire or the breakage of a component in the firing mechanism. Failure of ammunition is not charged against weapon performance. The likelihood of malfunction occurring in modern weapons is very small, nevertheless it exists, and appears to occur at random. This independence and small probability of occurrence suggest a Poisson distribution. During routine firing tests, the occurrence of malfunctions per hundred rounds fired is very often zero, occasionally one or two, and rarely three or more. Breakages are even less frequent. But any failure is considered serious, hence every attempt to fire during the various service trials is carefully accounted for. This scrutiny is particularly essential when two types of weapons are subjected to rough usage. It has been found feasible to expose weapons to unusually severe test conditions and thereby induce a higher incidence of failures. The big problem is to control these artificial conditions so that both types are equally exposed.

Example 6. Failures During Normal and Severe Tests.--Six weapons of type D and type E were subjected to four trials under conditions designed to induce failure to function. The number of attempts to fire was kept equal for each weapon type and the malfunctions were counted for each test.

Table 16. Number of Malfunctions for Tests Varying in Severity

Weapon Type	Trial a	b	c	d	Total
D	22	30	32	41	125
E	13	20	20	28	81

The frequency of malfunctions is analogous to defectives occurring to two types of machines operating at four different speeds. Two questions are posed by the data. (1) Do the weapon types retain their consistency under different stresses? (2) Do their frequencies differ significantly?

We now introduce the hypothesis that the expected frequency E of one weapon type is equal to the other. The deviations of observed frequencies from the expected are used to calculate the statistic, CHI-Square. The calculated value is then compared with a critical value at a given significance level. The contribution to CHI-Square by each trial is:

$$\chi^2 = \frac{(O_D - E)^2}{E} + \frac{(O_E - E)^2}{E} \qquad (19)$$

Where E is the expected frequency based on the hypothesis

O is the observed frequency for weapons D and E.

Since there are four trials, there will be four contributions to total CHI-Square, and therefore four degrees of freedom. The number of degrees of freedom is necessary to select the appropriate critical CHI-Square from the table (17).

The hypothesis to be tested is that the frequencies of both weapon types are equal. The expected frequency for each trial in Table 16 is the mean of the two observed values. It is shown in the table below, along with the contributions to CHI-Square by the two cells of each trial.

Table 17. Malfunctions of Weapon Types D and E Under Severe Trials; Expected Frequencies; and Values of CHI-Square

Trial	Weapon Type D	Weapon Type E	Total MF	Expected MF	Difference	CHI-Square	DF
a	22	13	35	17.5	4.5	2.314	1
b	30	20	50	25.0	5.0	2.000	1
c	32	20	52	26.0	6.0	2.769	1
d	41	28	69	34.5	6.5	2.449	1
Sum of four CHI-Squares						9.543	4
Totals	125	81	206	103	22.0	9.398	1
Difference						0.134	3

The CHI-Square table permits the following comparisons:

(1) CHI-Squares for individual trials do not reach 3.841, the critical value for one degree of freedom at the 5% level.

(2) Their sum exceeds 9.488 (DF=4), at the 5% level.

(3) For column totals, the value exceeds 6.635, which is the CHI-Square for one degree of freedom at the 1% level.

(4) The difference, based on three degrees of freedom, is a small value and is not significant.

The two questions mentioned earlier can now be answered. Consistency of the data is evidenced by the small value of CHI-Square difference (4), which indicates no interaction between severity of trial and the type of weapon under test. This conclusion is reinforced by the small variation between individual trial CHI-Squares (1).

The primary basis for rejecting the hypothesis of equality rests on the total CHI-Square (2), and its small probability, -- less than 5% chance of occurrence. This conclusion is reinforced by (3), which can stand on its own merit providing there is evidence of consistency.*

Example 7. Two Weapon Types and Two Categories of Failures. During the trials listed in Example 6, each malfunction was investigated in order to determine whether it was correctible by the gunner, or whether more extensive repairs or

---

* This analysis of the problem provides a mathematical equivalent for the graphical solutions which follow (16, p 189).

replacement of a part was required. For convenience, the failures are divided into major and minor categories.

Table 18. Two-Way Contingency Table. Two Categories of Failures Occurring to Two Weapon Types.

Weapon Type	Category of Failure Major	Minor	Total
D	a = 50	b = 75	125
E	c = 22	d = 59	81
Total	72	134	206

Table 18 is also known as a 2 X 2 Contingency Table. The problem that a two-way contingency table seeks to solve is whether one classification is independent of the other. Here, we are interested in evidence whether a difference exists between the true ratio of major to minor defects for the two weapon types. If a difference exists, then each category, major and minor, may be used to compare the two weapons.

A 2 X 2 table requires no "a priori" hypothesis to determine expected frequency, since it is entirely prescribed by the row and column totals. There is but one degree of freedom. Expected cell frequencies for Table 18 are:

$$E_a = 72 \times 125/206 = 43.69$$

$$E_b = 134 \times 125/206 = 81.31$$

$$E_c = 72 \times 81/206 = 28.31$$

$$E_d = 134 \times 81/206 = 52.69$$

Applying these and observed values to Equation (19):

$$\chi^2 = \frac{(50-43.69)^2}{43.69} + \frac{6.31^2}{81.31} + \frac{6.31^2}{28.31} + \frac{6.31^2}{52.69}$$

$$= 0.911 + 0.490 + 1.407 + 0.756$$

$$= 3.564$$

This value is less than the critical value of CHI-Square for one degree of freedom at the 5% level, which is 3.841. Hence the two categories may be considered independent. There is not sufficient evidence that a difference between the true ratios exists. The investigation may be carried one step further, by using the procedure of Example 6 and the hypothesis that the expected occurrence of malfunctions for both weapon types is equal.

Table 19. Malfunctions of Weapon Types D and E by Category; Values of CHI-Square

Category	Weapon D	Type E	CHI-Square	DF	Significance Level (%)
Major	50	22	8.68	1	0.3
Minor	75	59	1.68	1	19.0
Sum of two CHI-Squares			10.36	2	0.6
Totals	125	81	8.92	1	0.27
Interaction			1.44	1	26.0

Table 19 yields this evidence: there appears no significant interaction; the hypothesis of equality is disproved by the large sum of CHI-Squares and CHI-Square for totals. The disparity for major malfunctions is highly significant.

Use of binomial probability paper.--Binomial probability paper (Fig. 12), lends itself to the statistical comparison of paired counts and provides a visual appraisal as well. The grid lines on this graph paper are based on the square root scale. A pair of numbers, read as coordinates, represents a point on the graph. Equal counts fall along a line called the "50-50 split." A split is a line from the origin representing a theoretical proportion which corresponds to a ratio of the expected frequencies. For example, the 80-20 split is equal to the ratio, 4:1.

A set of paired counts may give evidence against a fixed theoretical proportion in two ways. (1) Variability of observed values may be tested in order to infer a lack of homogeneity if the samples came from different populations.* (2) Lack of agreement between the observed proportion and the theoretical one, such as equality (1:1 ratio), may be tested at a given significance level, usually 5%. If a set of paired counts is homogeneous, a CHI-Square test of their sums with respect to the 50-50 split is sufficient.**

---

* To apply the CHI-Square test for homogeneity, plot the paired counts and the split through their sum. Combine the perpendiculars as shown under "Crab Addition" in Fig. 12, and read off the CHI-Square value at the marginal scale.

** Refer to text and footnote on page 55.

Graphical analysis is facilitated by the fact that the deviation of a coordinate point from a split provides a measure of CHI-Square with one degree of freedom, and that ten millimeters on the graph is equal to the 5% level of significance. Thus two parallel lines, one above and one below the 50-50 split, serve to distinguish those coordinate points which are outside the 5% level. Refer to lines AB and A'B' in Figure 12. If an observed split is homogeneous, divergence of its end point from the 50-50 split is related to either parallel by inspection.

<u>Example 8. Failures of Two Weapon Types During Six Trials</u>.-- Table 20 lists the counts of malfunctions of Weapon C and Weapon D obtained during six trials. It is desired to determine whether both weapons types were equally exposed, i.e., whether the data is homogeneous, and if so, whether there is any difference in the frequency of occurrence.

Table 20. Malfunctions of Weapon Types C and D; Six Trials

Weapon Type	Trial a	b	c	d	e	f	Total
C	6	8	16	22	80	66	198
D	5	5	14	32	91	135	282

The computed CHI-Square for the homogeneity test of Table 20 is 12.856. For five degrees of freedom, this value

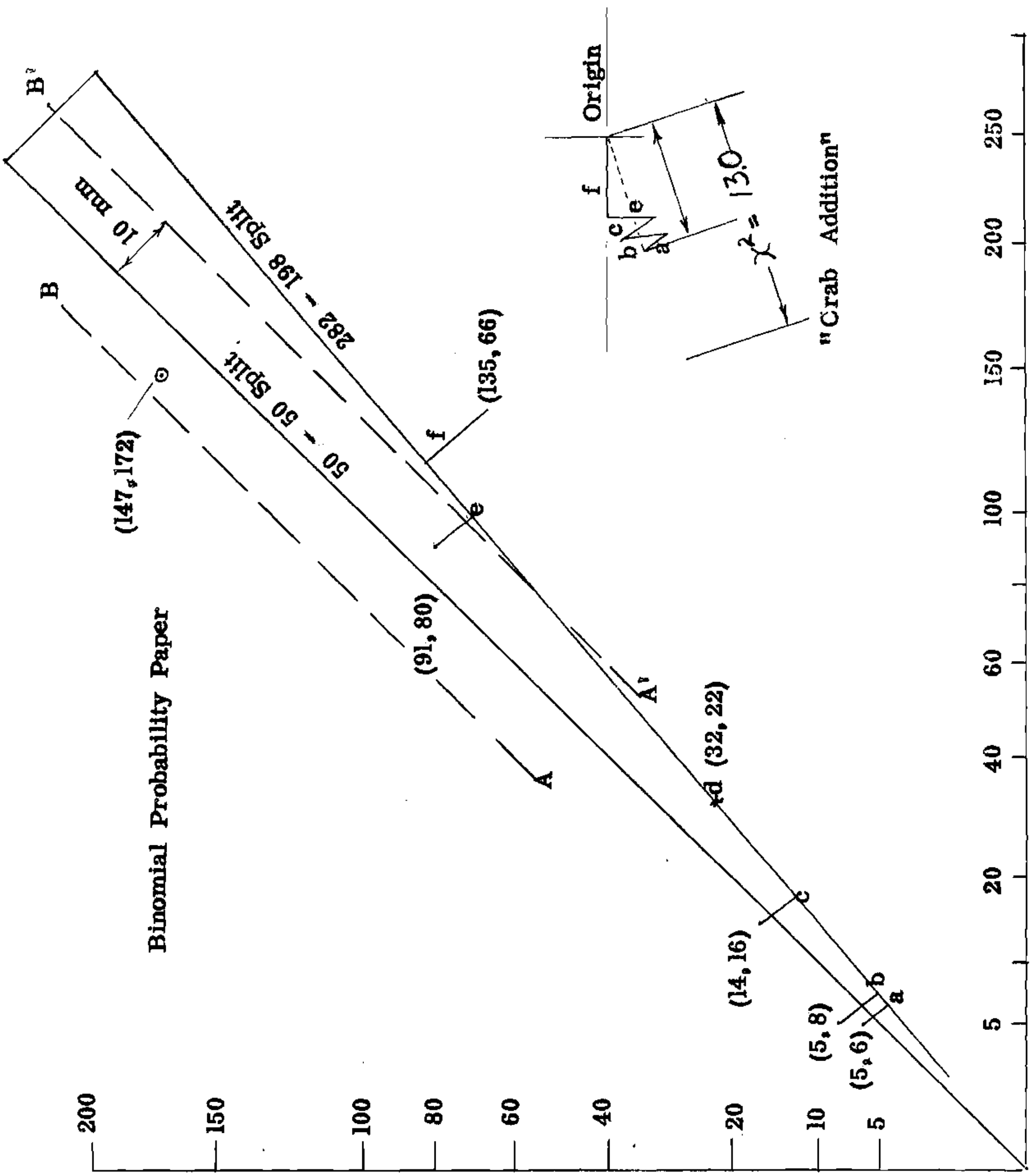

Fig. 12. Use of Binomial Probability Praper To Test Reliability.

is significant at the 2.5% level.* The lengthy calculations determine only that the data is not homogeneous. In Figure 12, the graphical solution consists of plotting coordinate points and the 282-198 split. The six deviations (a through f, Fig. 12), combined graphically by "crab addition," produce a CHI-Square which measures 13.0.** This is very close to the calculated value. In addition, the coordinate point (135,66), is more than 10 mm from the 282-198 split, thus indicating that the count in trial "f" is not homogeneous with the others, i.e., its contribution to total CHI-Square is significant at the 5% level. If the distribution were homogeneous, the set of paired counts differs from equality because the end point of the split is outside the 5% level, line A'B'. If the count for trial "f" is removed from the total, then the end point of a new split would be at the coordinate (147,172). Since the coordinate points do not deviate significantly from the 147-172 split (not drawn), the data would appear homogeneous. Also, the coordinate (147,172) lies within the 5% limit from equality, line AB.

A graphical representation of a 2 X 2 continency

---

* Normally, there is one DF for each trial. Here one DF is lost when the expected frequency (or split) is determined not by an "a priori" hypothesis, but by the count totals.

** In "crab addition," the first two deviations (e and f), are perpendicular. The third deviation (c) is drawn perpendicular to the hypotenuse, and so on. Note that (d) is zero.

table, such as Table 18, is given in Example 7 in Figure 2. Homogeneity with respect to the 81-125 split corresponds to the hypothesis of "independence." The two deviations, "a" and "b," (Fig. 2), combined at the upper left corner of the graph, yield a CHI-Square of 3.6. This agrees closely with the computed value, 3.56. A second comparison, with respect to the 50-50 split, reveals that minor malfunctions do not differ significantly from equality whereas the two frequencies for major malfunctions are highly significant. Here the deviations of points (59,75) and (22,50) are measured with respect to the 50-50 split. The end point is outside the line AB, hence the combined frequencies are different from equality at the 5% level.

Sample size for reliability tests.--It is not possible to estimate sample size needed to determine significantly different frequencies of malfunctions. However, it is possible to infer when one frequency approaches significance with respect to a known control weapon's rate. For a 5% level, lines AB and A'B' in Figure 12 furnish a locus of coordinates which fulfill this requirement. The equation of this line for values of X from 40 to 100 is:

$$Y = 1.23\,X + 10.4 \qquad (20)$$

Where X and Y are the counts of malfunctions of a control weapon and test weapon, respectively.

For a lower limit,

$$Y = \frac{X - 10.4}{1.23} \tag{21}$$

Equation (20) or (21) can be readily applied at the test site. For example, if the counts of control and test weapon malfunctions are 100 and 72, respectively, by Equation (21):

$$Y = \frac{100 - 10.4}{1.23} = \frac{89.6}{1.23}$$

$$= 72.9 \text{ or } 73$$

Since the actual count is less than computed Y, the test item's frequency is significantly smaller at the 5% level. It must be remembered, however, that a valid comparison can be made only after the paired counts are tested for homogenuity with other paired counts obtained.

CHAPTER V

CONCLUSIONS

General.--Most weapon characteristics subjected to investigation during service tests lend themselves to analysis by statistical methods. Statistics serve to condense data into usable terms, set confidence limits, test hypotheses, estimate probabilities and assist in the presentation of data. Comparisons between two or more items can be made according to the frequency of occurrence, or based on measurements of some characteristic of performance.

Measurement data.--Measurements can be summarized in terms of the central tendency (mean) and the dispersion (variance or probable error). When populations can be characterized by the normal probability, Binomial, or Poisson distributions, then the mathematical model provided by these distributions enhances our information concerning these entities.

Ballistic dispersion is a bivariate distribution, in which the horizontal and vertical components are normally distributed about the mean point of impact. Aiming errors for expert gunners using optical or metallic sights are very small, and their horizontal and vertical components are normally distributed about the aiming point. For distinct aiming points, they are comparable to or slightly greater than errors obtained with a surveyor's transit. When expert gunners register on a target, subsequent firing results in distributions which are somewhat larger than the ballistic error, but are known to be small.

When gunners engage targets at unknown ranges without previously registering, then the resulting dispersions tend to be very pronounced. With Weapon A, expert gunners achieved a normal distribution, but their human error was comparable to the ballistic error only up to mid-range (half of weapon's usable range). Beyond that, the human error rises exponentially and governs the man-weapon system performance. Human and incidental errors need considerable investigation. One valid conclusion obtained is that aiming errors are insignificant compared to human errors, and the two should be carefully distinguished.

Weapon C permitted a comparison between hit probability converted from measured total probable errors and actual per cent hits and thereby indicates that, at least for expert gunners, total dispersions at field targets tend to be normally distributed in horizontal and vertical directions about the point of aim. As a by-product, use of the nomograph for converting probable errors into hit probability appears acceptable.

Count data.--Malfunctions under normal firing conditions appear to follow the Poisson distribution. Severe trials may be used to induce higher rates of occurrence providing the weapons are equally exposed and the resulting counts are tested for consistency. The statistic, CHI-Square, may be used for comparing the malfunction rates of two weapons, but Binomial Probability Paper permits a more vivid, graphic analysis.

Hit probability.--A graph of hit probability versus range illustrates total system performance more forcibly than any other measure of accuracy. Its use should be encouraged. But care must be used to select only proficient gunners and to provide reasonable safeguards against spurious effects.

APPENDIX 1

GLOSSARY OF SYMBOLS AND TERMS

X or Y	a horizontal or vertical measurement.
$\bar{X}$	average of a group.
n	number of individual measurements in a group.
k	number of groups.
SS	sum of squares. $SS = \sum (X-\bar{X})^2$
DF	degrees of freedom. Normally one less than the number of measurements, e.g., (n-1), or (k-1).
s^2	variance. A mathematical ratio, SS / DF.
σ	standard deviation. The square root of variance. σ is a unit of measure of population dispersion such that 99.75% of the individuals of a normally distributed population fall within limits of the true mean $\pm 3\sigma$.
t	student's t ratio. Used to compare the differences between two means or a sample mean and a standard.
F	Fisher's F ratio. Used to compare sample variances.
$F_{.05}$	A value in a statistical F table located by the degrees of freedom of the variances under investigation.
PE	probable error. Equal to $0.6745\,\sigma$ for a normal distribution. It is that error which is just as likely to be exceeded as not.
p	probability, expressed from 0.0 to 1.0.
P	Percentage. (0 to 100%).

APPENDIX 2

DESCRIPTION OF WEAPONS

A	This is a large caliber weapon for an Infantry piece. It is crew served. The muzzle velocity is over Mach 1.
B	This weapon is similar to A, but has a different mount.
C	This is a small caliber weapon with a projectile velocity over Mach 2.
D	This weapon is comparable to C, but has an appreciably different silhouette.
E	This weapon is of the same family of weapons as C, but has different physical characteristics.
F	This is a small caliber weapon of a different family than Weapon C, and differs from C in physical and performance characteristics.

BIBLIOGRAPHY

References Cited in Text

1 Engelhardt, A. Baron, "The Story of European Proof Marks," _The Gun Digest_ 7th ed, 1953.

2 Lissak, O. M., _Ordnance and Gunnery_, New York, John Wiley & Sons, Inc., 1907

3 Cranz, C., and K. Becker, _Handbook of Ballistics_ London, HMS Stationery Office, 1921 (trans from German)

4 Simon, Leslie E., _Engineers' Manual of Statistical Methods_, New York, John Wiley & Sons, Inc., 1941

5 Baxter, James P. 3d, _Scientists Against Time_, Boston, Little, Brown, & Co., 1946

Bray, C. W., _Psychology and Military Proficiency_, Princeton, N.J., Princeton University Press, 1948

Bush, Vannevar, _Modern Arms and Free Men_, New York, Simon & Schuster, 1949

Gray, George W., _Science At War_, New York, Harper & Brothers, 1943

Thiesmeyer, L. R., _Combat Scientists_, Boston, Little, Brown & Co., 1947

6 Morse, P. M., and G. E. Kimball, _Methods of Operations Research_, 1st ed, Revised, Joint, The Technology Press of MIT and John Wiley & Sons, Inc., Second Printing, 1951

7 Butterbaugh, G. I., _Bibliography of Statistical Quality Control_, (revised ed, 1951), Seattle, University of Washington Press, 1946

Kurtz, Edwin B., _The Science of Valuation and Depreciation_, New York, The Ronald Press Co., Inc., 1937

Simon, Leslie E. Maj. Gen., "Ordnance Reliability Through Quality Control," _Ordnance_, XXXVIII, 199: 56-9, Jul-Aug 1953

8 Editorial, "Off the Editor's Chest," Consumers' Research Bulletin, 30: 2, Aug 1952

9 Dixon, W. J. and F. J. Massey, Jr., Introduction to Statistical Analysis, New York, McGraw-Hill Book Co., Inc., 1951

10 Duncan, A. J., Quality Control and Industrial Statistics, Chicago, Ill., Richard D. Irwin, Inc., 1952

11 Mosteller, F., & J. W. Tukey, "The Uses and Usefulness of Binomial Probability Paper," Journal of the American Statistical Association. 44: 174-212, June 1949

12 Mann, F. W., The Bullet's Flight, The Ballistics of Small Arms, Huntington, W. Va., Standard Publishing Co., 1909 Reprint 1942

13 Ordnance Proof Manual, Ord - M608 - PM, Vol III, Small Arms Division, Industrial Service, Washington, D. C.

14 Jacobs, D. H., Fundamentals of Optical Engineering, New York, McGraw-Hill Book Co., Inc., 1943

15 Brownlee, K. A., Industrial Experimentation, 4th ed., London, HMS Stationery Office, 1949, reprinted 1952

16 Snedecor, George W., Statistical Methods, Ames, Iowa, The Iowa State College Press, 1946

17 Arkin, H., & R. R. Colton, Tables for Statisticians, New York, Barnes & Noble, Inc., 1950

18 Heacock, F. A., Graphical Methods for Solving Problems, Princeton, N. J., Princeton University Press, 1952

References Not Cited

1 The American Society for Testing Materials, Manual on Presentation of Data, Philadelphia, Pa., 1940, reprint 1950

2 Captain Straight Shooter, "Inaccurate Rifles Make Inaccurate Riflemen," Combat Forces Journal, IV, 10:45 - May 1954

3 Cooper, B. M., "Graphical Analysis of Automatic Weapons Fire," Coast Artillery Journal, 87: 36-8 Mar-Apr 1944

4 Dudley, J. W., Jr., *Examination of Industrial Measurements*, New York, McGraw-Hill Book Company, Inc., 1946

5 Freedman, Paul, *The Principles of Scientific Research*, Washington, Public Affairs Press, 1950

6 Grant, E. L., *Statistical Quality Control*, 2d ed., New York, McGraw-Hill Book Co., Inc., 1952

7 Guilford, J. P., *Psychometric Methods*, 1st ed., New York, McGraw-Hill Book Co., 1936

8 Marshall, S. L. A., *Men Against Fire*, New York, Wm. Morrow & Co., 1947

9 McDonald, John, "The War of Wits," *Fortune*, 43:99-103, March 1951

10 Mead, L. C., & J. W. Wulfeck, "Human Engineering: The Study of the Human Factor in Machine Design," *Science Monthly*, 75: 372-9, Dec 1952

11 Morse, J. W., "Geometric Presentation of Correlation," *Journal of the American Statistical Association*, 32: 364-5, June 1937

12 Shewhart, W. A., & W. E. Deming, *Statistical Method from the Viewpoint of Quality Control*. Washington, The Graduate School, The Department of Agriculture, 1939

13 Solow, H., "Operations Research," *Fortune*, 43: 105-7, Apr 1951

14 Stevens, S. S., *Handbook of Experimental Psychology*, New York, John Wiley & Sons, Inc., 1951

15 Swan, A. W., "Human Relation in Industrial Operational Research," *Engineer*, 194, No. 5035; 117-20, July 25, 1952

16 Swift, E. F., Lieut. Col., "Division Objective," *Combat Forces Journal*, IV, 3: 18-34, Oct 1953

17 U.S. Army Field Forces Research Div., "Trajectories of Flat Trajectory, High Velocity Weapons," *Research Study No. 10.*, Fort Monroe, Va.