

Home

Contents

jameslindlibrary.org

## Records

Title Page(s)   Key Passage(s)   Context

[Download key passages/title pages as a PDF](#)

**Mainland D (1948)**. Statistical methods in medical research. Canadian Journal of Research E, 26:1-166.

**Title pages**

AUTHOR'S COPY. *With compliments  
D. Mainland*

**C.W.J.ARMSTRONG**

NATIONAL RESEARCH COUNCIL  
of CANADA

**STATISTICAL METHODS IN MEDICAL RESEARCH**  
I. QUALITATIVE STATISTICS (ENUMERATION DATA)

By DONALD MAINLAND



*Reprinted from the*  
**CANADIAN JOURNAL OF RESEARCH**  
*E, 26 : 1-166. 1948.*

Key passages

## STATISTICAL METHODS IN MEDICAL RESEARCH

### I. QUALITATIVE STATISTICS (ENUMERATION DATA)<sup>1</sup>

BY DONALD MAINLAND<sup>2</sup>

#### Abstract

This article is designed to help investigators in applying to qualitatively classified clinical and laboratory data the appropriate statistical treatment—tests of significance in binomial and multinomial distributions, estimation of confidence limits, analysis of contingency tables, and estimation of sample sizes required for further investigation. Section A is a brief introduction (definitions and principles). Section B comprises 40 examples classified so that the investigator can choose data and problems comparable to his own. Questions that arise in the examples, regarding experimental design (especially random sampling) and the interpretation of the tests, are discussed in Section C (Notes).

Because the standard deviation of the binomial and the chi square contingency test are often used without appreciation of the risk entailed, tables, which can be used also in nonmedical investigation, are presented: binomial confidence limits (with graphs) and exact probabilities for small-sample fourfold contingency tables. For samples not covered by the tables, precautions and rules regarding the use of chi square have been derived from more than five hundred comparisons between chi square and the exact method. To help in the exact computation of probabilities where that is necessary, four-decimal logarithms of factorials of numbers up to 1000 are given.

#### Contents

	PAGE
PREFACE.....	3
A—INTRODUCTION.....	4
1. Two types of data—enumeration data; mensuration data.....	4
2. Some principles and definitions—samples and population; random samples; frequency; probability, chances, and odds; types of problem.....	5
3. Argument from sample to population—levels of significance; confidence limits..	7
4. Comparison of two or more samples—arrangement of data; methods of analysis; levels of significance; one-sided comparisons.....	9
5. Random sampling—importance; simplicity; equality of numbers; bias and sample size.....	11
B—EXAMPLES.....	12
Summary of examples.....	12
1. Argument from sample to population.....	13
(1) Two-class Samples: Examples 1-12.....	13
(2) Samples with more than two classes: Example 13.....	26
2. Comparison of samples.....	29
(1) Two samples, each in two classes: Examples 14-26.....	29
(2) More than two samples; more than two classes: Examples 27-29.....	45

<sup>1</sup> Manuscript received August 28, 1947.

Contribution from the Anatomy Department, Dalhousie University, Halifax, N.S.  
Prepared for the Division of Medical Research of the National Research Council, with the aid of grants from the Council.

<sup>2</sup> Professor of Anatomy.

Contents—*Concluded*

	PAGE
<i>B—EXAMPLES—Concluded</i>	
3. Combination of information from two or more samples: Examples 30 and 31...	51
4. Confidence limits for differences between samples: Example 32.....	56
5. Sizes of samples required:	
Examples 33-36.....	57
6. Measurements treated as qualitative statistics:	
Examples 37-40.....	61
<i>C—NOTES</i> .....	65
1. Random sampling variation.....	65
2. The binomial expansion.....	67
3. Argument from sample to population.....	70
4. Significance and nonsignificance.....	71
5. Reasonableness of significance conventions.....	72
6. Confidence limits.....	73
7. The normal frequency curve.....	74
8. The standard deviation, $\sqrt{Npq}$ .....	74
9. Chi square used for testing a sample against a population value.....	77
10. Limitations of the chi square test.....	78
11. Chi square used with two-class samples.....	79
12. The general procedure in contingency tests.....	81
13. The exact method for contingency tests.....	82
14. Chi square in contingency tests.....	86
15. The use of the standard deviation, $\sqrt{Npq}$ , in comparison of samples.....	89
16. Confidence limits of sample differences.....	90
17. Required sample size—no population difference.....	91
18. Calculation of $P$ from the binomial expansion.....	92
19. The use of estimates as true population values.....	94
20. Tests of accuracy of chi square in fourfold contingency tables.....	95
21. Preparation and accuracy of the tables.....	96
22. Factorials.....	97
23. Random sampling techniques.....	97
24. Recommendations regarding mathematical tables and other sources of information.....	100
ACKNOWLEDGMENTS.....	101
REFERENCES.....	101
INDEX OF SUBJECTS.....	103
TABLES.....	104
IA. Confidence limits for twofold classification of enumeration data—number of $A$ 's in sample = 0.....	104
IB. Confidence limits for twofold classification of enumeration data—number of $A$ 's in sample: 1-20.....	105
II. Confidence limits for twofold classification of enumeration data—number of $A$ 's in sample: 20 and over.....	125
III. Correction terms for estimation of confidence limits—number of $A$ 's in sample greater than 20; percentage of $A$ 's: 10 or less.....	135
IV. Probabilities for fourfold contingency tables—equal samples up to $N = 20$ .....	136
V. Significant differences in fourfold contingency tables—unequal samples up to $N_1 = 20, N_2 = 19$ .....	142
VI. Significant differences in fourfold contingency tables—equal samples; $N = 20$ and over.....	157
VII. Chi square probabilities.....	158
VIII. Four-place logarithms of factorials of numbers up to 1000.....	159
GRAPHS 1-6 (Figs. 6-11). Confidence limits of Table II	

*A well-designed experiment contains in itself its estimate of error, and a procedure such as the following could be suggested:*

(a). If it is desired to select the apparently more susceptible men, do so by a preliminary test.

(b). Allocate in advance to the selected group, *strictly at random*, treatment and nontreatment. This takes care of all such factors as variable degrees of susceptibility, activities before the test and time of day when the test is made. Even loss of men owing to unforeseen circumstances (unless associated with their susceptibility to motion sickness) does not introduce a bias, although it removes the advantage of equal numbers.

(c). Compare the two samples as in the Examples indicated under that heading in the Summary of Examples (Subsection 2).

(d). If there is a significantly greater immunity in the drug-treated men, estimate for them the confidence limits as shown in the present example. The answers will show estimates of immunity *after* the drug, not necessarily wholly *due to* the drug, even if the nontreated men, in the particular experiment, all developed motion sickness.

[Home](#)

[Contents](#)