NASA CONTRACTOR REPORT

NASA CR-61335

ESTIMATION IN TRUNCATED POISSON DISTRIBUTIONS WITH CONCOMITANT EXPOSURE INTERVALS AND TRUNCATION POINTS

By A. Clifford Cohen

University of Georgia
Athens, Georgia

August 1970

Final Report

Prepared for

NASA-GEORGE C. MARSHALL SPACE FLIGHT CENTER
Marshall Space Flight Center, Alabama 35812

1. Report No. NASA CR-61335

2. Government Accession No.

3. Recipient's Catalog No.

4. Title and Subtitle

ESTIMATION IN TRUNCATED POISSON DISTRIBUTIONS WITH CONCOMITANT EXPOSURE INTERVALS AND TRUNCATION POINTS

5. Report Date

August 1970

6. Performing Organization Code

7. Author(s)

A. Clifford Cohen

8. Performing Organization Report No.

9. Performing Organization Name and Address

University of Georgia
Athens, Georgia

10. Work Unit No.

11. Contract or Grant No.

NAS8-11175

12. Sponsoring Agency Name and Address

Aero-Astrodynamics Laboratory
George C. Marshall Space Flight Center
Marshall Space Flight Center, Alabama 35812

13. Type of Report and Period Covered

CR - Final Report

14. Sponsoring Agency Code

15. Supplementary Notes Contract Monitor: Lee W. Falls, Aerospace Environment Division, Aero-Astrodynamics Laboratory.

16. Abstract

Let $\{x_i\}$, $i = 1, \ldots, n$, denote random observations of a truncated Poisson variable with parameter t_i, where the t_i are known positive numbers, and where truncation points c_i are specified such that $x_i \geq c_i$. Using x_i plus concomitant information provided by t_i and c_i, the likelihood estimating equation for λ becomes

$$\sum_1^n x_i = \lambda \left\{ \sum_1^n t_i + \sum_1^n [f(c_i - 1)/\bar{F}(c_i)] \right\}$$

where

$$f(c_i - 1) = e^{-\lambda t_i} (\lambda t_i)^{c_i - 1} / (c_i - 1)! \quad \text{and} \quad \bar{F}(c_i) = \sum_{x=c_i}^{\infty} e^{-\lambda t_i} (\lambda t_i)^x / x!$$

Solutions (estimates) are readily found using trial and error procedures with linear interpolation, but any standard iterative procedure would be applicable. The asymptotic variance of $\hat{\lambda}$ is obtained from the second partial of the likelihood function. An illustrative example is included.

17. Key Words (Suggested by Author(s))

Truncated Poisson Distribution
Statistical Analysis
Concomitant Variables

18. Distribution Statement

PUBLIC RELEASE:

E. D. Geissler
Director, Aero-Astrodynamics Laboratory

19. Security Classif. (of this report) UNCLASSIFIED

20. Security Classif. (of this page) UNCLASSIFIED

21. No. of Pages 12

22. Price*

*For sale by the Clearinghouse for Federal Scientific and Technical Information
Springfield, Virginia 22151

ESTIMATION IN TRUNCATED POISSON DISTRIBUTIONS WITH CONCOMITANT EXPOSURE INTERVALS AND TRUNCATION POINTS

by

A. Clifford Cohen

The University of Georgia

1. INTRODUCTION

Estimation in the truncated Poisson distribution has previously been considered by numerous writers including Bliss [1], David and Johnson [4], Moore [6], Plackett [7], Subrahmanian [8], the writer [2,3], and others. Our concern here is with estimation of the Poisson parameter when the $i^{\underline{th}}$ observation of the discrete random variable X extends over a specified exposure interval of magnitude t_i, subject to the restriction that $x_i \geq c_i$, where c_i is a specified truncation point. More specifically, our attention is directed toward maximum likelihood estimation based on a random sample consisting of n independent observations $\{x_i, t_i, c_i\}$, where x_i is the number of occurrences of the event of interest observed during the $i^{\underline{th}}$ interval of exposure, while t_i and c_i are corresponding concomitant values. Of course, x_i and c_i must be non-negative integers, but t_i is permitted to assume any positive value. Accident data in which reports are rendered for time intervals of varying size with truncation as stipulated fall into the category of samples under consideration here. Inspection data in which defects per unit are reported for units of differing sizes might also result in samples of the type considered here.

With concomitant values t_i and c_i given, the applicable Poisson probability function may be written as

$$\Pr\{X=x_i \mid t_i, c_i\} = \begin{cases} \dfrac{e^{-\lambda t_i} (\lambda t_i)^{x_i}}{x_i! \, \overline{F}(c_i)} , & x_i = c_i, c_i+1, \ldots, \\ 0 , & \text{elsewhere}, \end{cases} \tag{1}$$

where

$$\overline{F}(c_i) = \sum_{x_i=c_i}^{\infty} \frac{e^{-\lambda t_i} (\lambda t_i)^{x_i}}{x_i!} = 1 - F(c_i - 1) \ . \tag{2}$$

2. MAXIMUM LIKELIHOOD ESTIMATION OF λ

The likelihood function for a sample consisting of the n observed values $\{x_i, t_i, c_i\}$ for which the probability function (1) is applicable, follows as

$$L = \prod_{i=1}^{n} \frac{e^{-\lambda t_i} (\lambda t_i)^{x_i}}{x_i! \, \overline{F}(c_i)} \ . \tag{3}$$

On taking logarithms of both sides of (3) and differentiating with respect to λ, we obtain

$$\frac{\partial \ln L}{\partial \lambda} = -\sum_1^n t_i + \frac{1}{\lambda}\sum_1^n x_i - \sum_1^n \frac{1}{\bar{F}(c_i)} \frac{\partial \bar{F}(c_i)}{\partial \lambda} . \tag{4}$$

It follows from (2) that

$$\frac{\partial \bar{F}(c_i)}{\partial \lambda} = \frac{e^{-\lambda t_i} (\lambda t_i)^{c_i - 1}}{(c_i - 1)!} = f(c_i - 1) . \tag{5}$$

When (5) is substituted into (4) and the resulting expression is equated to zero and simplified, we obtain the estimating equation

$$\sum_1^n x_i = \lambda\left[\sum_1^n t_i + \sum_1^n f(c_i - 1)/\bar{F}(c_i)\right] , \tag{6}$$

in which $f(c_i-1)$ and $\bar{F}(c_i)$ are defined by (5) and (2) respectively.

With the aid of an ordinary table of individual and cumulative values of the Poisson function such as those of Molina [5], equation (6) can be solved with relative ease for the required estimate $\hat{\lambda}$, using trial and error techniques. Once two values λ_1 and λ_2 have been found in a sufficiently narrow interval such that $G(\lambda_1) < \sum_1^n x_i < G(\lambda_2)$, where $G(\lambda)$ has been written for the right side of (6), the required estimate $\hat{\lambda}$ can be obtained by interpolating linearly between λ_1 and λ_2. Of course, standard iterative procedures might also be employed in the calculation of $\hat{\lambda}$.

3. SOME SPECIAL CASES

By appropriately restricting the values of t_i and c_i in (6), we obtain results that apply in various cases of special interest. Certain of these specialized results are, of course, quite well known.

No Truncation — Unequal Intervals

In this case, $c_i = 0$ for all i, and equation (6) leads to the estimator

$$\hat{\lambda} = \sum_1^n x_i / \sum_1^n t_i \ . \qquad (7)$$

No Truncation — Equal Intervals

If in addition to the restriction, $c_i = 0$ for all i, we impose the further restriction that $t_i = t$ for all i, the applicable estimator reduces to

$$\hat{\lambda} = \bar{x}/t \ , \qquad (8)$$

where $\bar{x} = \sum_1^n x_i/n$. When t = 1, as in the usual case of estimation in the Poisson distribution, then (7) leads to the familiar estimator $\hat{\lambda} = \bar{x}$.

Only Zero Class Truncated In Making Each Observation — Unequal Intervals

In this case $c_i = 1$ for each i, and the estimating equation (6) reduces to

$$\sum_1^n x_i = \lambda \sum_1^n [t_i/(1-e^{-\lambda t_i})] \ , \qquad (9)$$

a result which was obtained independently by one of the writer's graduate assistants, Mr. S. S. Srivastava.

<u>Only Zero Class Truncated In Making Each Observation — Equal Intervals</u>

When we further specialize the preceeding result by requiring that $t_i = t$ for all i, then (9) becomes

$$\bar{x} = \lambda t/[1-e^{-\lambda t}] \ . \tag{10}$$

If we specialize still further by requiring that t = 1, then (10) reduces to

$$\bar{x} = \lambda/[1-e^{-\lambda}] \ , \tag{11}$$

as previously given in [3], where a table of the function on the right side of (11) is provided in order to facilitate solution for the required estimate $\hat{\lambda}$.

4. VARIANCE OF ESTIMATES

The asymptotic variance of $\hat{\lambda}$ can be expressed as

$$V(\hat{\lambda}) = -\left[E\left(\frac{\partial^2 \ln L}{\partial \lambda^2}\right)\right]^{-1} \doteq -\left[\frac{\partial^2 \ln L}{\partial \lambda^2}\right]^{-1}_{\lambda=\hat{\lambda}} \tag{12}$$

for each of the cases considered here. The second derivative in the most general case under consideration follows from (4) as

$$\frac{\partial^2 \ln L}{\partial \lambda^2} = -\frac{\sum_1^n x_i}{\lambda^2} - \sum_1^n \left[\frac{f(c_i-2) - f(c_i-1)}{\bar{F}(c_i)}\right] + \sum_1^n \left[\frac{f(c_i-1)}{\bar{F}(c_i)}\right]^2 , \tag{13}$$

a result which was previously given in [2]. With the aid of tables such as those of Molina [5], this derivative can easily be evaluated once (6) has been solved for $\hat{\lambda}$. The variance then follows from (12).

5. AN ILLUSTRATIVE EXAMPLE

In order to illustrate the practical application of estimators derived here, we consider inspection results on the number of surface defects per unit found in products of differing surface areas but which were manufactured on the same production line. In some of the inspections, no truncation takes place. In others, the zeros are eliminated, and in still others, both zeros and ones are eliminated. Based on the total sample which results from combining the several sub-samples involved, we proceed to estimate the Poisson parameter λ which characterizes the underlying common production process, and to determine the asymptotic variance of this estimate. The total sample available for this purpose consists of inspection results on 422 separate manufactured units. These data are presented in Table 1, with the total sample broken down into eight sub-samples having values of (t,c) as follows: (1,0),(1.5,1),(2,1),(2.5,1),(3,1),(1,2),(2,2), and (3,2). The value t = 1 indicates a surface area of one unit, while t = 2 indicates a surface area of two units, etc. For the total number of defects found in the 422 units inspected, we have $\sum_1^{422} x_i = 1597$. For the sum of the t's, we have $\sum_1^{422} t_i = 100(1)$ $+ 94(1.5) + 50(2) + 32(2.5) + 28(3) + 30(1) + 34(2) + 54(3) = 765$. These values are substituted into (6) which must then be solved for

TABLE 1

NUMBER OF SURFACE DEFECTS FOUND IN

FINAL INSPECTION OF 422 MANUFACTURED UNITS

No. Defects	Frequencies n_x								Row Totals
	c=0	c=1				c=2			
x	t=1	t=1.5	t=2	t=2.5	t=3	t=1	t=2	t=3	
0	13								13
1	28	16	5	1	0				50
2	27	21	6	3	1	14	6	3	81
3	17	22	9	5	3	9	5	3	73
4	10	16	11	6	5	5	7	7	67
5	3	11	9	5	4	1	6	7	46
6	2	6	4	5	5	1	5	9	37
7		1	3	4	4		3	8	23
8		1	2	2	3		1	6	15
9			1	1	1		1	5	9
10					2			3	5
11								2	2
12								1	1
Column Totals	100	94	50	32	28	30	34	54	422
$\Sigma x n_x$	200	294	203	154	162	86	153	345	1597

the required estimate $\hat{\lambda}$. As a first approximation, we try $\lambda_1 = 2.0000$, which comes from the sub-sample of size 100 with t = 1 and c = 0. In general, satisfactory first approximations might be obtained from any one or a combination of the available sub-samples. In some circumstances, a judicious guess might even suffice. We subsequently arrive at the final estimate $\hat{\lambda} = 2.027$, by linear interpolation as follows

λ		$G(\lambda)$
2.100		1652.91
2.027	$\sum_{1}^{422} x_i =$	1597.00
2.000		1576.75

where $G(\lambda)$ has been written for the right side of the estimating equation (6). For the variance of the above estimate, we employ (12) and (13) to calculate $V(\hat{\lambda}) = 0.0027$.

REFERENCES

[1] Bliss, C. I., (1948). Estimation of the mean and its error from incomplete Poisson distributions, Connecticut Agricultural Experiment Station Bulletin 513, 12 p.

[2] Cohen, A. C., Jr. (1954). Estimation of the Poisson parameter from truncated samples and from censored samples, J. Am. Stat. Assn. 49, 158-68.

[3] Cohen, A. C., Jr. (1960). Estimating the parameter in a conditional Poisson distribution, Biometrics 16, 203-11.

[4] David, F. N., and Johnson, N. L., (1952). The truncated Poisson, Biometrics 8, 275-85.

[5] Molina, E. C., (1942). Poisson's Exponential Binomial Limit, D. Van Nostrand Co., Inc.

[6] Moore, P. G., (1952). The estimation of the Poisson parameter from a truncated distribution, Biometrika, 39, 247-51.

[7] Plackett, R. L., (1953). The truncated Poisson distribution, Biometrics 9, 485-88.

[8] Rider, Paul R., (1953). Truncated Poisson distributions, J. Am. Stat. Assn. 48, 826-30.

[9] Subrahmanian, K., (1965). A note on estimation in the truncated Poisson, Biometrika 52, 279-82.

DISTRIBUTION

DEP-T
AD-S-DIR
PD-DIR(2)
PM-DIR
PM-MO-DIR (2)
CC-P
AST-U
S-DIR
A&TS-MS-IP
A&TS-MS-IL (8)
A&TS-MS-H
A&TS-MS-T, Mr. Wiggins (6)
A&TS-PAT

S&E-DIR
S&E-R-DIR
S&E-CSE-DIR (3)
S&E-ASTR-DIR (2)
S&E-COMP-DIR (2)
S&E-QUAL-DIR (2)
S&E-ASTN-DIR
S&E-AERO-DIR
S&E-AERO-T
S&E-AERO-F (2)
S&E-AERO-D (2)
S&E-AERO-A
S&E-AERO-AU
S&E-AERO-G
S&E-AERO-Y (3)
S&E-AERO-YT (50)
S&E-AERO-YE
S&E-AERO-P
S&E-AERO-YA (4)
S&E-AERO-YS (2)

Tech. & Sci. Info. Facility (25)
P. O. Box 33
College Park, Md.
Attn: NASA Rep. (S-AK/RKT)

NASA Headquarters
OART - Director
Dr. H. H. Kurzweg (3)
OMSF - Director
OSSA - Director
Dr. M. Tepper
Mr. Wm. Spreen
Dr. Summers
OTDA - Director

NASA - Langley Research Center
Attn: Director
Technical Library

NASA - Kennedy Space Center
Attn: Dr. A. H. Knothe
Dr. R. Bruns
Tech. Library

NASA - Manned Spacecraft Center
Attn: Director
Tech. Library

Dr. Churchill Eisenhart
Senior Res. Fellow
Adm. Bldg. A-1000
National Bureau of Standards
Washington, D. C. 20234

Univ. of Ga.
Dept. of Statistics
Athens, Ga. 30601
Attn: Dr. C. F. Kossack (2)
Dr. A. C. Cohen (50)
Mr. M. C. Carter
Dr. W. O. Williford
Dr. B. J. Williams
Dr. R. E. Bargmann

Dr. H. L. Crutcher (2)
NWRC - Federal Bldg.
Asheville, N. C. 28801

U. S. Naval Weapons Laboratory
Dahlgren, VA 22448

Ofc. of Naval Research
Washington, DC 20360

Sandia Corporation
Albuquerque, NM 87115

Naval Weather Service Command Hq.
Washington Navy Yard
Washington, DC20390

Meteorology Div. Deseret Test Center
Dugway, UT 84022

Defense Atomic Support Agency
Joint Task Force 8 (J-3)
Albuquerque, NM 87115

ESSA Air Resources Lab - Las Vegas
P. O. Box 14985
Las Vegas, NV 89114

N. Sissenwine (2)
CRE,Air Force Cambridge Res. Lab.
L. G. Hanscom Field
Bedford, Mass. 01730

Dr. Edward Epstein
Dept. of Meteorology & Oceanography
Univ. of Mich.
Ann Arbor, Mich. 48105

Dr. R. Q. Bradley
Head, Dept. of Statistics
Fla. State Univ.
Tallahassee, Fla. 32306

AFSC (SCWTS)
Andrews AFB
Washington, DC 20331

Pa. State Univ.
Dept. of Meteorology
State College, Pa. 16801
Attn: Dr. Hans Panofsky
Dr. C. Kreitzburg

RSIC (4)
Redstone Arsenal, Ala. 35812

Lockheed Co.
Research Park
Huntsville, Ala.

Mrs. N. E. Rich
Bldg. 5400 - Code AMS MI-RDF
Redstone Arsenal, Ala. 35809

Dr. Wm. Mendenhall
Head, Dept. of Statistics
Univ. of Fla.
Gainesville, Fla.

Dr. Max Woodbury
Div. of Biomathematics
P. O. Box 3200
Duke Medical Center
Durham, N. C. 27706

Environmental Sci. Service Adm.
Rockville, Md. 20852
Attn: Mr. K. Nagler, -W12
Dr. Sidney Tweles, -Wle
Mr. E. M. Vernon, -W11
Mr. H. Saylor, -W33
Mr. W. H. Klein, -W42
Tech. Library

Dr. Geo. E. Nicholson, Jr.
Chairman, Dept. of Statistics
Univ. of N. C.
Chapel Hill, NC 27515

Dr. F. Mosteller
Dept. of Statistics
Harvard Univ.
Rm. 311, 2 Divinity Ave.
Cambridge, Mass. 02138

Meteorological & Geoastrophysical
Abstracts
P. O. Box 1736
Washington 13, DC

Mr. R. Martin
General Dynamics/Astronautics
P.O. Box 1128
San Diego, Calif. 92112

Dr. P. R. Merry
Director, OA Standby Unit
Univ. of Denver
Denver, Colorado 80202

Dr. L. M. Lecam
Dept. of Statistics
Univ.of Calif.
Berkeley, Calif. 94720

Dr. O. Hartley
Dir., Inst. of Statistics
Texas A&M Univ.
College Sta., Texas 77843

Dr. O. M. Essenwanger
AOMC Res. Lab.
ORDMX-RRA
Redstone Arsenal, Ala. 35812

Dr. Arnold Court
17168 Septo St.
Northridge, Calif.

Dr. A. L. Finkner
Dept. of Experi. Statistics
N. C. State College
Durham, North Carolina 27702

Dr. T. A. Gleeson
Dept. of Meteorology
Fla. State Univ.
Tallahassee, Fla. 32306

Dr. B. G. Greenberg
Dept. of Biostatistics
School of Public Health
Univ. of N. C.
Chapel Hill, North Carolina 27515

Dr. F. E. Grubbs
Assoc. Tech. Dir.
Ballistic Res. Lab.
Aberdeen Proving Ground, Md. 21005

Dr. S. B. Littaver
Industrial & Management Engr.
Columbia Univ.
320 Seeley W. Mudd Bldg.
N. Y. 27, N. Y. 13625

Dr. Boyd Harshbarger
Head, Dept. of Statistics
Va. Polytechnic Inst.
Blacksburg, Va. 24061