Reprinted From

A Modern Course on
STATISTICAL DISTRIBUTIONS IN SCIENTIFIC WORK

Edited by
G. P. PATIL, S. KOZ, AND J. K. ORD

Published by
D. Reidel Publishing Company
38 Papeterspad, P. O. Box 17, Dordrecht, Holland
306 Dartmouth Street, Boston, Massachusetts 02116
STATISTICAL DISTRIBUTIONS IN SCIENTIFIC WORK

Based on the Nato Advanced Study Institute
A Modern Course on Statistical Distributions in Scientific Work
and The International Conference on Characterizations
of Statistical Distributions With Applications

Held at
The University of Calgary, Calgary, Alberta, Canada
July 29-August 10, 1974

Sponsored by
International Statistical Institute
The Pennsylvania State University
The University of Calgary
Indian Statistical Institute

With the Support of
North Atlantic Treaty Organisation
National Research Council of Canada
United States Army Research Office

DIRECTOR
G. P. Patil

SCIENTIFIC DIRECTORS
G. P. Patil, S. Kotz, J. K. Ord

JOINT DIRECTORS
E. C. Enns (Local Chairman), J. K. Wani, P. C. Consul

ADVISORS
T. Cacoullos C. B. Kemp I. Olkin
J. B. Douglas T. Kotlarski C. R. Rao
A. Hald E. Lukacs L. R. Shenton
W. L. Harkness L. J. Martin D. A. Sprott
N. L. Johnson W. Molenar M. Teicher
J. E. Mosimann

A CHARACTERISTIC PROPERTY OF CERTAIN GENERALIZED POWER SERIES DISTRIBUTIONS

G. P. Patil and V. Seshadri

The Pennsylvania State University, University Park, Pa., U.S.A. and McGill University, Montreal, Quebec, Canada.

SUMMARY. There are some probability distributions which remain invariant in their form under suitable transformations. In this paper, we show that the logarithmic series distribution and the geometric distribution enjoy the property of invariance of form as a characterizing property under a special type of transformation which we introduce below as a modulo sequence. Further, we provide a necessary and sufficient condition for the generalised power series distribution to have this characteristic property.

KEY WORDS. Modulo sequence, logarithmic series distribution, geometric distribution.

1. NOTATION AND TERMINOLOGY.

Definition 1. Let \( I^+ \) be the set of positive integers. A sequence of random variables \( \{X_k\} \) is said to be a modulo sequence of random variables generated by \( X_1 \) if, for every \( k \in I^+ \), the random variable (rv) \( X_k \) has, as its distribution, the conditional probability distribution of \( X_i/k \) given that \( X_i/k \in I^+ \), that is, given that \( X_i/k \) has assumed a positive integral value, that is, given that \( X_i \equiv 0 \mod (k) \). Further, the rv \( X_k \) is said to be the modulo (k) rv generated by the rv \( X_1 \).

Definition 2. Let \( T \) be a subset of the set \( I \) of non-negative integers. Define \( f(\beta) = \sum a(x) \beta^x \) where the summation extends over \( T \) and \( a(x) > 0 \), \( \beta > 0 \) with \( 0 \in \beta \), the parameter space, such that \( f(\beta) \) is finite and differentiable. Then a rv \( X \)
with probability function (p.f.)
\[ \text{Prob} \{ X = x \} = p(x; \theta) = a(x) \theta^x / f(\theta) \quad x \in T \]
is said to have the generalized power series distribution (GPSD) with range \( T \) and the series function (s.f.) \( f(\theta) \).

**Definition 3.** The GPSD with range \( T = \mathbb{N}^+ \) and the s.f. \( f(\theta) = -\log(1-\theta) = 1/\theta(\theta) \) is said to be the logarithmic series distribution (LSD) with parameter \( \theta \), \( 0 < \theta < 1 \).

**Definition 4.** The GPSD with range \( T = \mathbb{N}^+ \) and the s.f. \( f(\theta) = \theta/(1-\theta) = 1/\theta(\theta) \) is said to be the geometric distribution (GD) with parameter \( \theta \), \( 0 < \theta < 1 \).

2. Results.

**Theorem 1.** Let \( \{ X_k \} \) be a modulo sequence of r.v.'s generated by \( X_1 \). If \( X_1 \) has the LSD with parameter \( \theta \), then its modulo (k) r.v. \( X_k \) has the LSD with parameter \( \theta^k \) for every \( k \in \mathbb{N}^+ \).

**Proof.** From definition 1, we have
\[ \text{Prob} \left\{ X_k = x \right\} = \text{Prob} \left\{ X_1 \equiv x \pmod{k} \right\}, \]
which because of definition 3
\[ a(x) \theta^x / x \sum_{x=1}^{\infty} a(x) \theta^x = a(\theta^k) \theta^{kx} / kx. \]
Hence the theorem.

**Theorem 2.** Let \( \{ X_k \} \) be a modulo sequence of r.v.'s generated by \( X_1 \). If \( X_1 \) has the GD with parameter \( \theta \), then its modulo (k) r.v. \( X_k \) has the GD with parameter \( \theta^k \) for every \( k \in \mathbb{N}^+ \).

**Proof.** Follows as for Theorem 1.

**Theorem 3.** Let \( \{ X_k \} \) be a modulo sequence of r.v.'s generated by \( X_1 \) such that \( X_1 \) has a GPSD with parameter \( \theta \) and that the modulo (k) r.v. \( X_k \) has the LSD with parameter \( \theta^k \) for \( k > 1 \). Then \( X_1 \) has the unity–truncated LSD with parameter \( \theta \).

**Proof.** Let the s.f. of the GPSD of \( X_1 \) be
\[ g(\theta) = \sum y a(y) \theta^y \]
where \( y \in T \) with \( T \) as yet unspecified.

Now, by hypothesis, \( \text{Prob} \{ X_k = x \} \)
A PROPERTY OF CERTAIN GENERALIZED POWER SERIES DISTRIBUTIONS

\[ a(k^x) \frac{(\theta^k)^x}{x} \quad \text{for } k > 1 \]  \hspace{1cm} (2)

which by definition 1 of the module sequence

\[ = \text{Prob} \left[ X_k^1 = a/ k \in \mathbb{N}^+ \right] \]

\[ = \frac{a(kx)}{\sum_{x=1}^{\infty} a(kx) \theta^x} \quad \text{for } k > 1 \]  \hspace{1cm} (3)

Comparing (2) and (3), we note that, for a fixed \( \theta \),

\( xa(kx) \) is constant for every \( k > 1 \)  \hspace{1cm} (4)

Putting \( k = 2 \) and \( x = 1 \) and \( x = y \) we get from (4),

\( ya(2y) = a(2) \quad \text{for } y \geq 1 \)  \hspace{1cm} (5)

Further, putting \( k = y \) and \( x = 1 \) in (4) gives

\( a(y) = xa(yx) \)

which by putting \( x = w \) in turn gives

\( a(y) = 2a(2y) \quad \text{for } y > 1 \)  \hspace{1cm} (6)

Finally, we get from (5) and (6) that

\( a(y) = \frac{2a(2)}{y} \quad \text{for } y > 1 \)  \hspace{1cm} (7)

From (1) and (7) it follows that \( X_k \) has the unity-truncated LSD with parameter \( \theta \), since a constant multiple of the \( \text{sf} \) of a GPSD does not affect its probability function.

Theorem 4. Let \( (X_k) \) be a modulo sequence of \( rv \)'s generated by \( X_k \) such that \( X_k \) has a GPSD with \( \text{sf} g(\theta) = \sum a(y) \theta^y \) with \( a(1) = 2a(2) \); and further that \( X_k \) has the LSD with parameter \( \theta_k \) for \( k > 1 \). Then \( X_1 \) has the LSD with parameter \( \theta \).

Proof. Following the proof of Theorem 3, we have from (7)

\( a(y) = \frac{2a(2)}{y} \quad \text{for } y > 1. \) Also by hypothesis \( a(1) = 2a(2) \), therefore

\( a(y) = \frac{2a(2)}{y} \quad \text{for } y > 1. \) Hence the Theorem.

Theorem 5. Let \( (X_k) \) be a modulo sequence of \( rv \)'s generated by \( X_k \) such that \( X_k \) has a GPSD with parameter \( \theta \) and that the modulo \( (k) \) \( RV \) \( X_k \) has the GD with parameter \( \theta_k \) for \( k > 1 \). Then \( X_1 \) has the unity-truncated GD with parameter \( \theta \).

Proof. Similar to the proof of Theorem 3.
Theorem 6. Let \( \{X_k\} \) be a modulo sequence of \( rv \)'s generated by \( X_1 \) such that \( X_1 \) has a GPDF with \( g(y) = \frac{a(1)}{a(y)} \) with \( a(1) \) a(2); and further that \( X_k \) has the GD with parameter \( \theta^k \) for \( k > 1 \). Then \( X_1 \) has the GD with parameter \( \theta \).

Proof. Similar to the proof of Theorem 4.

Lastly, let \( f_k(\theta^k) = \sum_{x=1}^{\infty} a(x) (\theta^k)^x \) be a SF in powers of \( \theta^k \) for each \( k \in I^+ \). Let \( \{X_k\} \) be a modulo sequence of \( rv \)'s generated by \( X_1 \). We have then the following.

Theorem 7. A necessary and sufficient condition for the modulo \( \{X_k\} \) to be a GPDF with the SF \( f_k(\theta^k) \) in powers of \( \theta^k \) is that the coefficient \( a(kx) \) decomposes into two suitable factors which separate \( k \) and \( x \).

Proof. To prove the necessity, we have for all \( x \in I^+ \) and

\[
\sum_{x=1}^{\infty} a(kx) \theta^{kx} = \sum_{x=1}^{\infty} a(x) (\theta^k)^x, \quad \text{therefore}
\]

\[
\sum_{x=1}^{\infty} a(kx) \theta^{kx} = f_k(\theta^k), \quad \text{with} \quad b(1) = 1
\]

(8)

\[
a(kx) = b(k) a(x),
\]

(9)

where \( b(k) = \sum_{x=1}^{\infty} a(kx) \theta^{kx} / f_k(\theta^k) \) with \( b(1) = 1 \).

To prove the sufficiency, (8) clearly follows from (9).

Corollary 1. The LSD is characterized by \( b(k) = 1/k \) where \( b(k) \) is defined by (9) for \( k \in I^+ \).

Corollary 2. The GD is characterized by \( b(k) = 1 \) where \( b(k) \) is defined by (9) for \( k \in I^+ \).

REFERENCES

A Modern Course on

STATISTICAL DISTRIBUTIONS IN SCIENTIFIC WORK

Contents

VOLUME 1: MODELS AND STRUCTURES

Section

1. INAUGURAL ADDRESS — G. P. Patil
2. POWER SERIES AND RELATED FAMILIES — S. W. Joshi, T. Cacoullos,
   A. W. Kemp and C. D. Kemp, P. C. Consul and L. R. Shenton,
   J. Garland and R. Tripathi, G. P. Patil and V. Seshadri
3. RECENT TRENDS IN UNIVARIATE MODELS — S. J. Press, J.
   Behboodian, K. V. Mardia, P. McNolty, J. R. Huynen, and
   E. Hansen, T. P. Hettmansperger and M. A. Kanman, A. L. Rukhin
4. MOMENTS-RELATED PROBLEMS — M. S. Ramanujan, C. G. Heyde,
   W. L. Harkness
5. LIMIT DISTRIBUTIONS AND PROCESSES — B. Harris and A. P. Somm,
6. MULTIVARIATE CONCEPTS AND MODELS — S. Kotz, E. Jogdeo,
   J. J. J. Roux, M. Siotani
7. CERTAIN MULTIVARIATE DISTRIBUTIONS — A. Dussauchoy and
   E. Berland, P. C. Durling, R. P. Gupta, C. G. Khatri,
   J. Tiago de Oliveira
8. SAMPLING DISTRIBUTIONS AND TRANSFORMATIONS — M. Shaked,
   K. G. Bowman and W. E. Duenasberry, V. B. Waiker, P. Frishman,
   J. Ogawa

VOLUME 2: MODEL BUILDING AND MODEL SELECTION

1. MODELLING AND SIMULATION — J. K. Ord and G. P. Patil,
   G. P. Patil and M. Boswell, E. J. Dudewicz, G. P. Patil,
   M. Boswell, and D. Friday, J. S. Ramberg
2. MODEL IDENTIFICATION AND DISCRIMINATION — R. Srinivaasan and
   C. E. Antle, M. George, V. Seshadri, and M. Yalovsky,
   S. K. Katti and K. McDonald, J. Garland and R. C. Nachys,
   M. L. Tiku, J. J. Cott
3. MODELS IN THE SOCIAL SCIENCES AND MANAGEMENT — J. K. Ord,
   J. K. Ord, E. W. Resek, C. Chatfield
4. MODELS IN THE PHYSICAL AND BIOMEDICAL SCIENCES - J. E.
Hosimann, J. E. Hosimann, M. E. Wise, S. Talwalkar, D. M.
Schults, S. S. Shapiro, C. G. Laurent, E. Elvers
5. MODELS IN THE ENVIRONMENTAL SCIENCES - M. F. Dacey, D. V.
Gokhale, G. Ramachandran, M. G. Warren
6. A MODERN COURSE ON STATISTICAL DISTRIBUTIONS - Participants

VOLUME 3: CHARACTERIZATIONS AND APPLICATIONS

1. LINNEN MEMORIAL INAUGURAL LECTURE - C. R. Rao
2. MATHEMATICAL TOOLS FOR CHARACTERIZATION PROBLEMS - E. Lukacs,
J. Aczél, R. J. Rossberg
3. CHARACTERIZATIONS USING ORDER STATISTICS - J. Galambos, J.
Govindarajulu, M. Ahsanullah, J. S. Huang and J. S. Hwang,
Z. Govindarajulu, J. S. Huang, and A. E. M. E. Saleh
4. CHARACTERIZATIONS BY OTHER STATISTICAL PROPERTIES - A. L.
Rubin, F. S. Gordon and S. F. Gordon, A. W. Kagam and L. B.
Klebanov, L. E. Nondessen, C. G. Khatri, B. Gyires
5. CHARACTERIZATIONS ON SPACES AND PROGRESSES - 1. S. Kotlarski,
R. L. S. Prakasa Rao, K. Ahmad, E. Seneta
6. CHARACTERIZATION PROBLEMS FOR DAMAGED OBSERVATIONS - G. P.
Patil and M. V. Rangaparkhi, R. C. Srivastava and J. Singh,
P. C. Consul, J. K. Ord
7. CHARACTERIZATIONS USING ENTROPY MEASURES AND RELATED PROBLEMS
- D. V. Gokhale, R. Shimizu, A. W. Kemp, J. Aczél
8. CHARACTERIZATIONS FOR DISCRETE DISTRIBUTIONS AND FAMILIES -
A. V. Godambe and G. P. Patil, A. M. Nevill and C. D. Kemp,
K. G. Janardan
9. CHARACTERIZATIONS FOR CONTINUOUS DISTRIBUTIONS AND FAMILIES -
K. V. Mardia, M. S. Bingham and K. V. Mardia, A. F. Basu and
R. W. Block, J. K. Wani and G. P. Patil

FOR YOUR LIBRARY COPY

These Proceedings of a Modern Course on Statistical Distributions in Scientific Work provide carefully edited materials in three comprehensive volumes on related problems of contemporary interest and concern, such as, Models and Structures, Model Building and Model Selection, and Characterizations and Applications. Every quantitative scientist and statistician should find them very useful and worthwhile for both research and instruction involving any aspects of distributions. Order your copy today.