THE TECHNOLOGICAL INSTITUTE

NORTHWESTERN UNIVERSITY
EVANSTON, ILLINOIS

SYSTEMS RESEARCH MEMORANDUM No. 171

The Technological Institute The College of Arts and Sciences

Northwestern University

GENERATION AND APPROXIMATION OF REACH AND DISTRIBUTION OF FREQUENCIES

by

A. Charnes, W. W. Cooper*, J. K. DeVoe**,
D. B. Learner***, and W. Reinecke***

February 14, 1967

* Carnegie Institute of Technology
** Cargill, Wilson and Acree, Inc.
*** Batten, Barton, Durstine and Osborn, Inc.

Part of the research underlying this report was undertaken for the Office of Naval Research, Contract NONR-1228(10), Project NR 047-021, and for the U. S. Army Research Office - Durham, Contract No. DA-31-124-ARO-D-322 at Northwestern University. Reproduction of this paper in whole or in part is permitted for any purpose of the United States Government.

SYSTEMS RESEARCH GROUP

A. Charnes, Director

1. Introduction:

A previous paper [1] described a goal programming model for media scheduling which explicitly considered the cumulative audiences within and between media. For this purpose appropriately structured formulas for handling audience duplication had to be designed in a manner that was compatible with the basic goal programming model while faithfully representing audience characteristics.

For convenient reference we herewith define and distinguish the following terms

Gross Rating Points = Gross Audience
Reach = Net Audience
Average Frequency = Gross Rating Points/Reach.

It is common practice to proceed toward a synthesis of media schedules by reference to these kinds of measures. The development of mathematical models and related computer aids, etc., make it feasible and desirable to consider refinements and extensions of these practices. These kinds of possibilities were, in fact, exploited when proceeding to the goal programming model for media selection that we have elsewhere characterized as LP II. See [11]. Thus, in particular, we have replaced the concept of average frequency--a single number--with the entire distribution of frequencies in each of the audiences which are to be targeted.

1/ [11].

The availability of such distributions now make it possible to consider the cumulative audience of a single medium as well as the duplication of audiences among many different media, cumulatively and over a variety of time periods.[1/] Such possibilities were not always available in the past, and when available could not, in fact, be fully exploited because of the additional complexities that were thereby introduced into an already complicated problem. Indeed, it was deemed wise to avoid the additional developments that would have been necessary even to incorporate these features into the model that we refer to as LPI.[2/]

2. Background:

As part of these developments for LPII it became necessary to consider how ways might be devised for handling such reach-frequency considerations as part of an operational model. There are, of course, exact formulations which are available for employment on certain assumptions. At best, however, these formulations are unwieldy for use on problems of the size contemplated here. They also suffer from other defects. For instance, they are based on assumptions as to data availability that are generally difficult to fulfill in practice.

References for some of these exact formula possibilities are given in the bibliography along with further references to some of the empirical

---

1/ See [12].

2/ See [4] and [12] and [7]. See also [10] and [11].

approaches that were also explored. This includes work by Agostini as well as some of the work that other persons have undertaken either to extend or to replace the original Agostini formulations. None of these proved to be sufficiently well suited for the use requirements of LPII, however, and so these explorations were also unsuccessful.

The availability of adequate electronic computation facilities suggested the possible use of tabulated data which could be handled by suitable "table look up" devices. Such tabulations as those prepared by A. C. Nielsen, Politz, Simmons and others, however, covered only relatively small portions of the total population that was of interest for LPII. Hence this alternative possibility also had to be abandoned.

Some of the simulation approaches to media scheduling have also been considered the possibilities for improved media scheduling by reference to ways in which the distribution of frequencies might be developed for explicit consideration. Cases in point are the models devised by Simulmatics as well as the proposed Compass Model of J. Diebold and Associates and the CAM Model of the London Press Exchange. The first two provide the wanted counts only after a schedule has been selected. The CAM Model includes a further difficulty in that it relies on readership data that are not generally available in the U.S., as well as assessing a schedule only after it has been selected. LPII, however, is designed to consider the conditional interaction between the media selected and the way they are scheduled with reference to all of the possibilities admitted by the constraints.

3. Direction of Further Developments:

The above summary discussion is intended to indicate some of the directions that were also taken in this research on practicable ways of obtaining entire distributions of frequencies for use in planning media schedules. Having thus examined all of the above possibilities, it next became necessary to consider what might be done in developing entirely new approaches. This was done in two phases. First an exact and complete mathematical formulation of the nonlinear audience accumulation was developed which could be transformed into a system compatible with linear programming techniques--but without loss of the nonlinear accumulation properties. Second suitable approximating devices were developed and tested empirically for obtaining the wanted distributions of frequencies.

4. Analytic Development:

To make this all more precise, we reproduce the following notations and definitions from [12]. Let

(1)
$$d_{kij}(t) = \text{gross } k^{\underline{th}} \text{ audience segment obtained by the } j^{\underline{th}} \text{ cumulative purchase of medium } i \text{ in period } t.$$
$$x_{ij}(t) = \text{amount of the } j^{\underline{th}} \text{ cumulative purchase of medium } i \text{ in period } t$$

and $x_{ij}(t) = 0$ or 1. We approximate this generally by requiring

(2) $$\sum_j x_{ij}(t) \leq 1, \quad x_{ij}(t) \geq 0.$$

Thus, the gross $k^{\underline{th}}$ audience segment obtained by media purchases in period t may be written

(3) $$\sum_i \sum_j d_{kij}(t)\, x_{ij}(t) .$$

Next we consider the net audience or "reach". Let

$r_{kij}(t)$ = $k^{\underline{th}}$ net audience segment obtained by the $j^{\underline{th}}$ cumulative purchase of medium i in period t

(4) and

$R_k(t)$ = net $k^{\underline{th}}$ audience segment obtained by media purchases in period t.

We now assume that the "non-reach" is given by the product of the individual non-reaches--e.g.,

(5) $$1 - R_k(t) = \prod_{i,j} (1 - r_{ki_j}(t))^{x_{ij}(t)}$$

where "$\prod$" means "product"--of the indicated terms--and the $x_{ij}(t)$ conform to (2) for each t. Making logarithmic transformations of (5) we obtain

(6) $$\ln (1 - R_k(t)) = \sum_i \sum_j x_{ij}(t) \ln (1 - r_{kij}(t)) ,$$

an expression which is linear in the decision variables $x_{ij}(t)$.

In general the data for the $r_{kij}(t)$ are not immediately available.

They are generated, however, by means of the following formula

$$(7) \qquad r_{kij} = r_{kil} + a_{ki} \ln j$$

where the r_{kil} are obtained from data supplied by syndicated services along with r_{kij} for some particular $j > 1$. These data are used to estimate the a_{ki} so that the above formula may then be used to secure the other r_{kij}.

5. Approximation for Distribution of Frequencies:

T approximate the distribution of frequencies recourse was had to various types of statistical distributions (such as the Poisson, etc.). These proved to be unsatisfactory either in the fits they yielded or in the fact that they were unwieldy to implement. Previous work in connection with the DEMON[1/] model, however, suggested that the log normal distribution might be expected to provide satisfactory descriptions of this additional aspect of consumer behavior.

As with the normal distribution, the log-normal distribution is also completely specified by 2 parameters. Thus the approximation problem became one of specifying these parameters from the available syndicated survey sources. It was found simplest to do this in terms of a logarithmic scale whereby the "studentized variate" of the associated normal distribution would be specified. Although the studentized variate involves an estimate of the standard deviation, it was found possible to develop linear formulae

---

1/ See [10] and [11].

which realistically represented the necessary means and standard deviations. Thus, the following approximations were used,

$$\mu_k(t) = A + B \sum_i \sum_j P_{ki}(t)\, x_{ij}(t) + C \sum_i \sum_j x_{ij}(t)$$

$$\sigma_k(t) = D + E \sum_i \sum_j x_{ij}(t) \tag{8}$$

$$\text{with} \quad P_{ki}(t) = \frac{d_{kil}(t)}{U_k}$$

and U_k = total number in universe for the $k^{\underline{th}}$ audience segment

where μ_k and σ_k are, respectively, the mean and standard deviation of the associated normal distribution and the A, B, C, D and E are constants which have been determined empirically.

The correspondence with the discrete distribution is made via

$$\frac{\ln (s - 1.5) - \mu_k(t)}{\sigma_k(t)} = Z_{(1 - H_{ks}(t))} \tag{9.1}$$

where Z is the studentized normal variate and

(9.2) $H_{ks}(t)$ = proportion of the net $k^{\underline{th}}$ audience segment which is reached s or more times in period t.

Employing (8) and effecting algebraic rearrangements, then

$$Z = \frac{\ln (s - 1.5) - A - B \sum_{i,j} P_{ki}(t)\, x_{ij}(t) - C \sum_{i,j} x_{ij}(t)}{D + E \sum_{i,j} x_{ij}(t)} \tag{10}$$

where Z is the fractile associated with $1 - H_{ks}(t)$ for N(0, 1).

Validation of the constants A, B, C, D and E--as employed in (8)--was obtained by applying X^2 tests to the predicted distributions of frequencies as obtained from (10). In all cases the resulting X^2 values were highly satisfactory. In every case the theoretical and empirical distributions were in agreement at levels well beyond 99%.

6. Conclusion:

One convenient way to summarize the above developments is to note again that they all proceeded by reference to data availabilities and requirements for use in large scale problems of media planning. They were also designed for use in a linear programming model of goal programming variety. Thus the stability-sensitivity properties of large-scale programming models needs to be allowed for in assessing the validity of the results obtained by reference to their effects on the optimal outcomes in these applications.

Quite apart from the usages of these reach and distribution of frequency estimators in linear programming models for media planning, these formulations have evident value for other aspects of media planning. Nor is there any reason to restrict their potential applications to the media field alone. They should, in fact, be regarded as important estimating devices for any problem involving statistical phenomena of saturation type--i.e., cases involving concave nonlinear behavior which are otherwise not representable within convex optimizations. It was partly for these reasons that the statistical tests were undertaken with the resulting X^2 measures that were referred to at the end of the preceding section.

Bibliography

[1] Aaker, D., "Probabilistic Approach to Industrial Media Selection" (Stanford University: Graduate School of Business, May, 1966).

[2] Agostini, J. M., "Analysis of Magazine Accumulative Audience," Journal of Advertising Research 2, No. 4, Dec., 1962, pp. 18-23.

[3] _____, "How to Estimate Unduplicated Audiences," Journal of Advertising Research 1, No. 3, March, 1961, pp. 11-14.

[4] American Association of Advertising Agencies, Papers from 1961 Regional Convention, Group IV, Mathematical Programming for Better Media Selection

[4.1] Wilson, C. L., "Introduction," pp. 1-3
[4.2] Charnes, A. and W. W. Cooper, "Budgeting and Planning Media Schedules," pp. 5-24.
[4.3] Learner, D. B. "The Translation from Theory to Practice," pp. 25-32.
[4.4] Godfrey, M., "Computer Processing and Computer Results," pp. 33-42.
[4.5] Maneloveg, H.D. "How The Practical Media Man Handles the Output," pp. 43-50.
[4.6] Lucas, D. B., "A Summary Perspective--Some Implications, Questions and Comments," pp. 51-58.

[5] Batten, Barton, Durstine & Osborn, Inc., Glossary of Media Terms (New York: BBDO, Inc., 1966).

[6] Bower, John, "The Audiences of U.S. and Canadian Magazines: Seven Tests of Agostini Formula," Journal of Advertising Research, Vol. III, No. 1, March, 1963, pp. 13-20.

[7] Buzzell, R. D., "Mathematical Models and Marketing Management," Ch. V (Cambridge: Harvard University, Division of Research, Graduate School of Business Administration, 1964).

[8] Caffyn, J. M. and M. Sheovsky, "Net Audiences of British Newspapers - A Comparison of the Agostini and Gainsbury Methods," Journal of Advertising Research, Vol. III, No. 1, March, 1963, pp. 21-25.

[9] Charnes, A. and W. W. Cooper, Management Models and Industrial Applications of Linear Programming (New York: John Wiley & Sons, Inc., 1961).

[10] _____ _____, J. K. DeVoe, and D. B. Learner, "DEMON: Decision Mapping Via Optimum Go-No Networks--A Model for Marketing New Products," Management Science 12, No. 11, July, 1966, pp. 865-887.

[11] _____ _____ _____ _____, "DEMON, Mark II: An Extremal Equation Approach to New Product Marketing," Management Science, Series A, submitted.

[12] _____ _____ _____ _____, and W. Reinecke, "LPII: A Goal Programming Model for Media Planning," Management Science, Series B, submitted.

[13] Claycamp, H. J. and C. W. McClelland, "Estimating Research and the Magic of K," Palo Alto: Stanford University, Graduate School of Business.

[14] Engleman, Fred L., "An Empirical Formula for Audience Accumulation," Journal of Advertising Research, Vol. V, No. 2, June, 1965, pp. 21-28.

[15] Hofmans, Pierre, "Measuring the Cumulative Net Coverage of any Combination of Media," Journal of Marketing Research, August, 1966, pp. 269-78.

[16] Kaatz, R.B., "Improving Agostini's Formula for Net Audience," Journal of Advertising Research, Vol. III, No. 3, September 1963, pp. 43-44.

[17] Kuhn, Walther, "Net Audiences of German Magazines: A New Formula," Journal of Advertising Research, Vol. III, No. 1, March, 1963, pp. 30-33.

[18] Marc, Maral, "Net Audiences of French Business Papers - Agostini's Formula Applied to Special Markets," Journal of Advertising Research, Vol. III, No. 1, March, 1963, pp. 26-29.

[19] Metheringham, Richard A., "Measuring the Net Cumulative Coverage of a Print Campaign," Journal of Advertising Research, Vol. IV, No. 4, December, 1964, pp. 23-28.

[20] Mevik, Oddvar Bie and Niels Vinding, "Two Dimensions of Media Selection: Coverage and Frequency," Journal of Advertising Research, Vol. VI, No. 1, March, 1966, pp. 29-34.

[21] National Industrial Conference Board, Studies in Business Policy Policy No. 121, Evaluating Media (New York: NICB, 1967).

[22] Politz, Alfred, Research, Inc., A Study of Four Media: Their Accumulative and Repeat Audiences, (Time, Inc., 1953).

[23] _____, Magazine Combinations, Vol. Six, Life Study of Consumer Expenditures (Time, Inc., 1959).

[24] Schyberger, B. W., "The Accumulative and Repeat Audiences of Swedish Weekly Magazines," Journal of Advertising Research, Vol. III, No. ., December, 1963, pp. 25-33.

[25] Simmons, W. R. and Associates Research, Inc. Mass Markets and the Media Reaching Them (New York, 1964).

[26] _____, Selective Markets and the Media Reaching Them (New York, 1964).

[27] Stock, J. Stevens, "Characteristics of the Reading Audiences of Newsweek, Time and U.S. News & World Report" (Technical Appendix) Newsweek, (New York: 1959).

[28] Valavanis, Stephan, Econometrics (McGraw-Hill, Inc., New York, 1959).

Unclassified
Security Classification

DOCUMENT CONTROL DATA - R&D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author)
Northwestern University

2a. REPORT SECURITY CLASSIFICATION
Unclassified

2b. GROUP
Not Applicable

3. REPORT TITLE
GENERATION AND APPROXIMATION OF REACH AND DISTRIBUTION OF FREQUENCIES

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)
Technical Report

5. AUTHOR(S) (Last name, first name, initial)
Charnes, Abraham, Cooper, William W., DeVoe, James K.*, Learner, D. B.**, and Reinecke, Walter**

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
February, 1967	8	28

8a. CONTRACT OR GRANT NO.	9a. ORIGINATOR'S REPORT NUMBER(S) ***
NONR 1228(10)	Systems Research Memorandum No. 171
b. PROJECT NO. NR 047-021	
c.	9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report) *** Management Sciences Research Report Number 97
d.	

10. AVAILABILITY/LIMITATION NOTICES
Distribution of this document is unlimited.

11. SUPPLEMENTARY NOTES ***	12. SPONSORING MILITARY ACTIVITY
Also under Contract NONR 760(24) Project NR 047-048 Carnegie Institute of Technology	Logistics and Mathematical Statistics Branch Office of Naval Research Washington, D. C. 20360

13. ABSTRACT

Media planning has heretofore been restricted, by and large, only to the consideration of average frequency. A recently developed mathematical model--see LPII (12)--explicitly considers both cumulative duplicating audiences over a variety of time periods and the simultaneous selection and scheduling of media. This model thus required the development of methods for handling such nonlinear concave aspects of audience characteristics within a linear programming model and for generating distribution of frequency data within the computational processes of the model. This is done by means of logarithmic transformations and, for the discrete distributions of frequencies, by log-normal approximating devices. Each of these two techniques is important in its own right as a type of method which may be employed in a wide variety of modelling or computational situations. Thus, because neither of these techniques is restricted to the area of media planning per se it seemed worthwhile to present them on their own terms--as in the present paper--instead of treating them only as a detail in developing and interpreting the LPII model (12).

*Cargill, Wilson and Acree, Inc.
**Batten, Barton, Durstine and Osborn, Inc.

***Joint with Carnegie Institute of Technology, Management Sciences Research Group; see items 9a, 9b and 11.

DD FORM 1 JAN 64 1473

Unclassified
Security Classification

Unclassified

Security Classification

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Linear Programming Log Linear Constraints Log normal approximations Nonduplication Advertising Media Planning						

INSTRUCTIONS

1. ORIGINATING ACTIVITY: Enter the name and address of the contractor, subcontractor, grantee, Department of Defense activity or other organization *(corporate author)* issuing the report.

2a. REPORT SECURITY CLASSIFICATION: Enter the overall security classification of the report. Indicate whether "Restricted Data" is included. Marking is to be in accordance with appropriate security regulations.

2b. GROUP: Automatic downgrading is specified in DoD Directive 5200.10 and Armed Forces Industrial Manual. Enter the group number. Also, when applicable, show that optional markings have been used for Group 3 and Group 4 as authorized.

3. REPORT TITLE: Enter the complete report title in all capital letters. Titles in all cases should be unclassified. If a meaningful title cannot be selected without classification, show title classification in all capitals in parenthesis immediately following the title.

4. DESCRIPTIVE NOTES: If appropriate, enter the type of report, e.g., interim, progress, summary, annual, or final. Give the inclusive dates when a specific reporting period is covered.

5. AUTHOR(S): Enter the name(s) of author(s) as shown on or in the report. Enter last name, first name, middle initial. If military, show rank and branch of service. The name of the principal author is an absolute minimum requirement.

6. REPORT DATE: Enter the date of the report as day, month, year, or month, year. If more than one date appears on the report, use date of publication.

7a. TOTAL NUMBER OF PAGES: The total page count should follow normal pagination procedures, i.e., enter the number of pages containing information.

7b. NUMBER OF REFERENCES: Enter the total number of references cited in the report.

8a. CONTRACT OR GRANT NUMBER: If appropriate, enter the applicable number of the contract or grant under which the report was written.

8b, 8c, & 8d. PROJECT NUMBER: Enter the appropriate military department identification, such as project number, subproject number, system numbers, task number, etc.

9a. ORIGINATOR'S REPORT NUMBER(S): Enter the official report number by which the document will be identified and controlled by the originating activity. This number must be unique to this report.

9b. OTHER REPORT NUMBER(S): If the report has been assigned any other report numbers *(either by the originator or by the sponsor)*, also enter this number(s).

10. AVAILABILITY/LIMITATION NOTICES: Enter any limitations on further dissemination of the report, other than those imposed by security classification, using standard statements such as:

(1) "Qualified requesters may obtain copies of this report from DDC."

(2) "Foreign announcement and dissemination of this report by DDC is not authorized."

(3) "U. S. Government agencies may obtain copies of this report directly from DDC. Other qualified DDC users shall request through _____."

(4) "U. S. military agencies may obtain copies of this report directly from DDC. Other qualified users shall request through _____."

(5) "All distribution of this report is controlled. Qualified DDC users shall request through _____."

If the report has been furnished to the Office of Technical Services, Department of Commerce, for sale to the public, indicate this fact and enter the price, if known.

11. SUPPLEMENTARY NOTES: Use for additional explanatory notes.

12. SPONSORING MILITARY ACTIVITY: Enter the name of the departmental project office or laboratory sponsoring *(paying for)* the research and development. Include address.

13. ABSTRACT: Enter an abstract giving a brief and factual summary of the document indicative of the report, even though it may also appear elsewhere in the body of the technical report. If additional space is required, a continuation sheet shall be attached.

It is highly desirable that the abstract of classified reports be unclassified. Each paragraph of the abstract shall end with an indication of the military security classification of the information in the paragraph, represented as *(TS), (S), (C), or (U)*.

There is no limitation on the length of the abstract. However, the suggested length is from 150 to 225 words.

14. KEY WORDS: Key words are technically meaningful terms or short phrases that characterize a report and may be used as index entries for cataloging the report. Key words must be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location, may be used as key words but will be followed by an indication of technical context. The assignment of links, roles, and weights is optional.

DD FORM 1 JAN 64 1473 (BACK)

Unclassified

Security Classification