UNCLASSIFIED

AD NUMBER
AD421157
LIMITATION CHANGES
TO: **Approved for public release; distribution is unlimited.**
FROM: **Distribution authorized to U.S. Gov't. agencies and their contractors; Administrative/Operational Use; 07 JUN 1963. Other requests shall be referred to Army Medical Research Lab., Fort Knox, KY.**
AUTHORITY
AMRL ltr 26 Feb 1970

THIS PAGE IS UNCLASSIFIED

UNCLASSIFIED

AD 421157

DEFENSE DOCUMENTATION CENTER
FOR
SCIENTIFIC AND TECHNICAL INFORMATION
CAMERON STATION, ALEXANDRIA, VIRGINIA

UNCLASSIFIED

Best
Available
Copy

NOTICE: When government or other drawings, specifications or other data are used for any purpose other than in connection with a definitely related government procurement operation, the U. S. Government thereby incurs no responsibility, nor any obligation whatsoever; and the fact that the Government may have formulated, furnished, or in any way supplied the said drawings, specifications, or other data is not to be regarded by implication or otherwise as in any manner licensing the holder or any other person or corporation, or conveying any rights or permission to manufacture, use or sell any patented invention that may in any way be related thereto.

CATALOGED BY DDC
AS AD No. 421157

US ARMY MEDICAL RESEARCH LABORATORY

FORT KNOX, KENTUCKY

REPORT NO. 575

A GENERAL METHOD OF OBTAINING EXACT SAMPLING PROBABILITIES OF THE SHANNON-WIENER MEASURE OF INFORMATION, $\hat{H}$

James N. Cronholm, M. S.

DDC
RECEIVED
OCT 25 1963
TISIA B

UNITED STATES ARMY
MEDICAL RESEARCH AND DEVELOPMENT COMMAND

7 June 1963

FOR ERRATA

AD 421157

THE FOLLOWING PAGES ARE CHANGES

TO BASIC DOCUMENT

AD 421159

US ARMY MEDICAL RESEARCH LABORATORY
Fort Knox, Kentucky

ERRATUM

Report No. 575

A GENERAL METHOD OF OBTAINING EXACT SAMPLING PROBABILITIES OF THE SHANNON-WIENER MEASURE OF INFORMATION, $\hat{H}$

James N. Cronholm, M. S.

Page 1, 2nd paragraph, line 1, should read: Inspection of equation ([illegible]) will show that $\hat{H}$ is a one-to-one func-

Report Submitted 25 March 1963

Author

James N. Cronholm, M.S.

Complex Processes Branch
Psychology Division

Qualified requestors may obtain copies of this report from ASTIA.

Foreign announcement and dissemination of this report by ASTIA is limited.

AD Accession No.

US Army Medical Research Lab, Ft. Knox, Ky.
A GENERAL METHOD OF OBTAINING EXACT SAMPLING PROBABILITIES OF THE SHANNON-WIENER MEASURE OF INFORMATION, Ĥ. - J. N. Cronholm
Report No. 575, 7 Jun 63, 9 pp & i - 1 table
DA Project No. 3A012001B801, Unclassified Report

UNCLASSIFIED
1. Statistics
2. Non-parametric
3. Information

A two-variable generating function is described which yields the sampling probabilities of the Shannon-Wiener information measure. Expansion and collection of terms in like powers of the first variable imposes the restriction that the sum of the k category frequencies equal n; collection of terms in like powers of the second variable then produces terms whose coefficients are the required probabilities. The method may be used with either equal or unequal category probabilities for any finite n and k, and thus represents a general solution to the small sample problem. Tables of sampling probabilities are presented.

AD Accession No.

US Army Medical Research Lab, Ft. Knox, Ky.
A GENERAL METHOD OF OBTAINING EXACT SAMPLING PROBABILITIES OF THE SHANNON-WIENER MEASURE OF INFORMATION, Ĥ. - J. N. Cronholm
Report No. 575, 7 Jun 63, 9 pp & i - 1 table
DA Project No. 3A012001B801, Unclassified Report

UNCLASSIFIED
1. Statistics
2. Non-parametric
3. Information

A two-variable generating function is described which yields the sampling probabilities of the Shannon-Wiener information measure. Expansion and collection of terms in like powers of the first variable imposes the restriction that the sum of the k category frequencies equal n; collection of terms in like powers of the second variable then produces terms whose coefficients are the required probabilities. The method may be used with either equal or unequal category probabilities for any finite n and k, and thus represents a general solution to the small sample problem. Tables of sampling probabilities are presented.

AD Accession No.

US Army Medical Research Lab, Ft. Knox, Ky.
A GENERAL METHOD OF OBTAINING EXACT SAMPLING PROBABILITIES OF THE SHANNON-WIENER MEASURE OF INFORMATION, Ĥ. - J. N. Cronholm
Report No. 575, 7 Jun 63, 9 pp & i - 1 table
DA Project No. 3A012001B801, Unclassified Report

UNCLASSIFIED
1. Statistics
2. Non-parametric
3. Information

A two-variable generating function is described which yields the sampling probabilities of the Shannon-Wiener information measure. Expansion and collection of terms in like powers of the first variable imposes the restriction that the sum of the k category frequencies equal n; collection of terms in like powers of the second variable then produces terms whose coefficients are the required probabilities. The method may be used with either equal or unequal category probabilities for any finite n and k, and thus represents a general solution to the small sample problem. Tables of sampling probabilities are presented.

AD Accession No.

US Army Medical Research Lab, Ft. Knox, Ky.
A GENERAL METHOD OF OBTAINING EXACT SAMPLING PROBABILITIES OF THE SHANNON-WIENER MEASURE OF INFORMATION, Ĥ. - J. N. Cronholm
Report No. 575, 7 Jun 63, 9 pp & i - 1 table
DA Project No. 3A012001B801, Unclassified Report

UNCLASSIFIED
1. Statistics
2. Non-parametric
3. Information

A two-variable generating function is described which yields the sampling probabilities of the Shannon-Wiener information measure. Expansion and collection of terms in like powers of the first variable imposes the restriction that the sum of the k category frequencies equal n; collection of terms in like powers of the second variable then produces terms whose coefficients are the required probabilities. The method may be used with either equal or unequal category probabilities for any finite n and k, and thus represents a general solution to the small sample problem. Tables of sampling probabilities are presented.

REPORT NO. 575

A GENERAL METHOD OF OBTAINING EXACT SAMPLING PROBABILITIES OF THE SHANNON-WIENER MEASURE OF INFORMATION, $\hat{H}$

James N. Cronholm, M.S.

Psychology Division
US ARMY MEDICAL RESEARCH LABORATORY
Fort Knox, Kentucky

7 June 1963

Basic Research in Psychological and Social Sciences
DA Project No. 3A012001B801

Report No. 575
DA Project No. 3A012001B801

ABSTRACT

A GENERAL METHOD OF OBTAINING EXACT SAMPLING PROBABILITIES OF THE SHANNON-WIENER MEASURE OF INFORMATION, $\hat{H}$

OBJECT

To describe a method of obtaining the exact sampling probabilities of the maximum likelihood estimate of the Shannon-Wiener information measure, and to present tables of these probabilities.

RESULTS

A two-variable generating function was described which yielded the sampling probabilities of the Shannon-Wiener information measure. Expansion and collection of terms in like powers of the first variable imposed the restriction that the sum of the k category frequencies equal n; collection of terms in like powers of the second variable then produced terms whose coefficients were the required probabilities. Tables of these probabilities were presented for the equiprobable case.

CONCLUSIONS

The method may be used with either equal or unequal category probabilities for any finite n and k, and thus represents a general solution to the small sample problem for this widely used statistic.

APPROVED: George S. Harker
GEORGE S. HARKER, Ph. D.
Director, Psychology Division

APPROVED: Sven A. Bach
SVEN A. BACH
Colonel, MC
Director

A GENERAL METHOD OF OBTAINING EXACT SAMPLING PROBABILITIES OF THE SHANNON-WIENER MEASURE OF INFORMATION, $\hat{H}$

The outcomes of many experiments are characterized by the distribution of n independent observations among k mutually exclusive and exhaustive categories. In situations of this sort, a number of statistics may be computed as functions of the category frequencies, n_i. A popular, and sometimes useful statistic is the maximum likelihood estimate of the Shannon-Wiener measure of information [2, 3]:

$$\hat{H} = -\sum_{i=1}^{k} \frac{n_i}{n} \log_2 \frac{n_i}{n}$$

$$= \log_2 n - \frac{1}{n} \sum_{i=1}^{k} n_i \log_2 n_i \quad , \tag{1}$$

where

$$\sum_{i=1}^{k} n_i = n \quad . \tag{2}$$

While approximate tests of the significance of $\hat{H}$ are available for large n [1], knowledge of the exact sampling probabilities is essential for accuracy when n is small, and of considerable theoretical interest for any n. The purpose of this paper is to present a general method for finding exact sampling probabilities of $\hat{H}$ using a two-variable generating function.

Inspection of equation (1) will show that $\hat{H}$ is a single valued function of the sum

$$S_r = \sum_{i=1}^{k} n_i \log_2 n_i \quad . \tag{3}$$

Accordingly, the probability of obtaining the rth value of $\hat{H}$ is equal to the probability of obtaining the rth value of the sum S, i.e., $f(\hat{H}_r) = f(S_r)$. The two-variable generating function of equation (4) produces the probabilities of these sums:

$$F(t, u) = \prod_{i=1}^{k} \sum_{n_i=0}^{\infty} \frac{1}{n_i!} p_i^{n_i} u^{n_i} t^{n_i \log_2 n_i} \quad . \tag{4}$$

The k p_is in this function are probabilities subject to the restriction that

$$\sum_{i=1}^{k} p_i = 1 \ , \quad p_i > 0 \ (i = 1, 2, \cdots, k) \ ,$$

and may be regarded as the parameters of a multinomial distribution with k categories from which the n independent observations are drawn.

Each of the k sums in $F(t, u)$ may be identified with a category, while the individual terms in each sum represent possible outcomes in the specified category. The general term in the sum representing the ith category,

$$\frac{1}{n_i!} p_i^{n_i} u^{n_i} t^{n_i \log_2 n_i} ,$$

indicates that this category contains n_i observations which could have occurred in any one of $n_i!$ ways with a probability of $p_i^{n_i}$, and that the corresponding ith term in the sum, S_r, will be $n_i \log_2 n_i$.

If the multiplication in (4) is carried out and terms in like powers of u are collected, $F(t, u)$ may be expressed as

$$F(t,u) = \sum_{n=0}^{\infty} g_n(t) \frac{u^n}{n!} .$$

This operation invokes the restriction (2) that the sum of the category frequencies equal the number of observations, n. Multiplication of the coefficients of u^n by $n!$ accounts for the possible orderings of all n observations.

Next, if terms in like powers of t are collected in $g_n(t)$, the resulting terms will have the form

$$T = \left\{ \sum^{[n_i]_r} \sum^{q} \frac{n!}{\prod_{i=1}^{k} n_i!} \prod_{i=1}^{k} p_i^{n_i} \right\} t^{S_r} , \tag{5}$$

in which the first sum is over the

$$q = \frac{k!}{\prod_{j=0}^{n} k_j!} , \tag{6}$$

with

$$\sum_{j=0}^{n} k_j = k ,$$

possible orderings of a distinct set of n_is among the k categories. In q (6), the k_js are the number of n_is equal j. The second sum in (5) is over all distinct sets of n_is, $[n_i]_r$, which yield the rth value of S.

This expression (5) can be interpreted readily. The multinomial probability within the summation signs gives the probability of obtaining

a particular set of category frequencies in the indicated categories. Summing over the distinct permutations of the k categories gives the probability of obtaining a particular set of n_is irrespective of category. Finally, summing these probabilities over all distinct sets of n_is which yield the sum S_r gives the probability of obtaining S_r. Since the observed set of n_is is included in these sets, the coefficient of t^{S_r} in (5) gives the required probability.

$g_n(t)$ is a polynomial in t in which the exponents of t are the possible values of S_r, and the coefficients of t are the corresponding probabilities. Thus, $g_n(t)$ may be written

$$g_n(t) = \sum_{r=1} f(S_r) t^{S_r} ,$$

$$= \sum_{r=1} f(\hat{H}_r) t^{S_r} ,$$

and the sampling probability of the rth value of $\hat{H}$ is given by (5), with t set equal to 1.

Thus,

$$f(\hat{H}_r) = \sum^{(n_i)_r} \sum^{q} \frac{n!}{\prod_{i=1}^{k} n_i!} \prod_{i=1}^{k} p_i^{n_i} . \tag{7}$$

A special case of some importance arises when the category probabilities are known or assumed to be equal, i.e., $p_i = 1/k$ $(i = 1, 2, \ldots, k)$. In this case, (7) reduces to

$$f(\hat{H}_r) = \sum^{(n_i)_r} \frac{k!}{\prod_{j=0}^{n} k_j!} \frac{n!}{\prod_{i=1}^{k} n_i!} \frac{1}{k^n} .$$

In summary, the sampling probabilities of $\hat{H}$, given n, k, and the population probabilities, p_i, may be determined by first expanding the two-variable generating function, F(t, u) (4), collecting terms in $u^n/n!$ to find $g_n(t)$, and finally, collecting terms in this function in like powers of t to find $f(\hat{H})$.

For computational purposes, it should be noted that if the infinite series in F(t, u) are truncated at $n_i = c$, the functions $g_n(t)$ will be correct for all values of n up to and including c. Terms in higher powers of u than c should be discarded.

As an example of how the two-variable generating function is used consider the equiprobable case in which $k = 4$, $p_i = 1/k = 1/4$, and

$n = 4$. Substituting these values in $F(t, u)$ (4), and truncating at $n_i = 4 = c$ gives

$$F(t,u) = \left(1 + \tfrac{1}{4}u + \tfrac{1}{2!}\tfrac{1}{4^2}u^2 t^{2\log_2 2} + \tfrac{1}{3!}\tfrac{1}{4^3}u^3 t^{3\log_2 3} + \tfrac{1}{4!}\tfrac{1}{4^4}u^4 t^{4\log_2 4}\right)^4 .$$

Expanding and collecting terms in $u^n/n!$ up to $u^4/4!$ yields

$$F(t,u) = 1 + u + \left(\tfrac{12}{16} + \tfrac{4}{16}t^2\right)\tfrac{u^2}{2!} + \left(\tfrac{24}{64} + \tfrac{36}{64}t^2 + \tfrac{4}{64}t^{3\log_2 3}\right)\tfrac{u^3}{3!} + \left(\tfrac{24}{256} + \tfrac{144}{256}t^2 + \tfrac{36}{256}t^4 + \tfrac{48}{256}t^{3\log_2 3} + \tfrac{4}{256}t^8\right)\tfrac{u^4}{4!} ,$$

in which the polynomial coefficients of $u^n/n!$ are the functions $g_n(t)$:

$$g_1(t) = t^0$$

$$g_2(t) = \tfrac{3}{4}t^0 + \tfrac{1}{4}t^2$$

$$g_3(t) = \tfrac{3}{8}t^0 + \tfrac{9}{16}t^2 + \tfrac{1}{16}t^{3\log_2 3}$$

$$g_4(t) = \tfrac{3}{32}t^0 + \tfrac{9}{16}t^2 + \tfrac{9}{64}t^4 + \tfrac{3}{16}t^{3\log_2 3} + \tfrac{1}{64}t^8 .$$

The exponents of t in these expressions are the possible values of S_r for $n = 1, 2, 3, 4$, and $k = 4$. The value of $\hat{H}$ determined by each S_r may be found by substituting S_r (3) in the equation for $\hat{H}$ (1). For $g_4(t)$ the S_rs are: $S_1 = .0000$, $S_2 = 2.0000$, $S_3 = 4.0000$, $S_4 = 3\log_2 3 = 4.7549$, and $S_5 = 8.0000$. The corresponding values of $\hat{H}$ are: $\hat{H}_1 = 2.0000$, $\hat{H}_2 = 1.5000$, $\hat{H}_3 = 1.0000$, $\hat{H}_4 = .8113$, and $\hat{H}_5 = .0000$. The sampling probabilities associated with these values of $\hat{H}$ are given by the coefficients of t^{S_r} in $g_4(t)$. Thus, $f(2.0000) = 3/32$, $f(1.5000) = 9/16$, $f(1.0000) = 9/64$, $f(.8113) = 3/16$, and $f(.0000) = 1/64$. The functions $g_1(t)$, $g_2(t)$, and $g_3(t)$ may be used in the same way to identify the sampling probabilities of $\hat{H}$ when $n = 1, 2$, and 3.

Table 1 is the result of similar calculations for the special case of equiprobable categories, $2 \leq k \leq 11$, and $2 \leq n \leq 12$. The first page of Table 1 and the entries for $9 \leq n = k \leq 11$ were computed by hand, set aside, and recomputed. As a further check, the expected values of these sampling distributions were compared with the values obtained by Rogers and Green [2]. At this point the probabilities for the entire table became available through computer-expansion of the generating function, and the hand calculations were used to check these results.*

*The author gratefully acknowledges the contribution of A. T. Chen of the University of Louisville Speed Scientific School Computer Laboratory who developed the program used to expand the generating function.

The exact sampling probability of an obtained $\hat{H}$ may be found by entering Table 1 at n, k, and the value of the observed $\hat{H}$. Tests of the significance of obtained values of $\hat{H}$ under the hypothesis of equal category probabilities may be devised by summing the tabled sampling probabilities over appropriately chosen rejection regions.

While the use of Table 1 is restricted to the case of equal category probabilities, it should be emphasized that the generating function F(t, u) (4) may be employed in the general case of unequal probabilities for any finite n and k. The proposed method therefore represents a general solution to the small sample problem.

REFERENCES

[1] Miller, G. A. and W. G. Madow. On the maximum likelihood estimate of the Shannon-Wiener measure of information. Air Force Cambridge Research Center Technical Report 54-57, 1954.

[2] Shannon, C. E. and W. Weaver. The Mathematical Theory of Communication. Urbana: University of Illinois Press, 1949.

[3] Wiener, N. Cybernetics. New York: Wiley, 1948.

Reference to Rogers and Green on page 4

Rogers, M. S. and B. F. Green. The moments of sample information when the alternatives are equally likely. In H. Quastler (Ed.), Information theory in psychology. Glencoe, Ill.: The Free Press, pp 101-107, 1955.

TABLE 1

Exact sampling probabilities of $\hat{H}$ for $2 \le k \le 11$, $2 \le n \le 12$, and $p_i = 1/k$.

N	$\hat{H}$	k = 2	3	4	5	6	7	8	9	10	11
2	1.000	.5000	.6667	.7500	.8000	.8333	.8571	.8750	.8889	.9000	.9091
	.000	.5000	.3333	.2500	.2000	.1667	.1428	.1250	.1111	.1000	.0909
3	1.585		.2222	.3750	.4800	.5556	.6122	.6562	.6914	.7200	.7438
	.918	.7500	.6667	.5625	.4800	.4167	.3673	.3281	.2963	.2700	.2479
	.000	.2500	.1111	.0625	.0400	.0278	.0204	.0156	.0123	.0100	.0083
4	2.000			.0937	.1920	.2778	.3498	.4102	.4609	.5040	.5409
	1.500		.4444	.5625	.5760	.5556	.5248	.4922	.4609	.4320	.4057
	1.000	.3750	.2222	.1406	.0960	.0694	.0525	.0410	.0329	.0270	.0225
	.811	.5000	.2963	.1875	.1280	.0926	.0700	.0547	.0439	.0360	.0301
	.000	.1250	.0370	.0156	.0080	.0046	.0029	.0020	.0014	.0010	.0008
5	2.322				.0384	.0926	.1499	.2051	.2561	.3024	.3442
	1.922			.2344	.3840	.4630	.4998	.5127	.5121	.5040	.4918
	1.522		.3704	.3516	.2880	.2315	.1874	.1538	.1280	.1080	.0922
	1.371		.2469	.2344	.1920	.1543	.1249	.1025	.0854	.0720	.0615
	.971	.6250	.2469	.1172	.0640	.0386	.0250	.0171	.0122	.0090	.0068
	.722	.3125	.1234	.0586	.0320	.0193	.0125	.0085	.0061	.0045	.0034
	.000	.0625	.0123	.0039	.0016	.0008	.0004	.0002	.0002	.0001	.0001
6	2.585					.0154	.0428	.0769	.1136	.1512	.1878
	2.252				.1152	.2315	.3213	.3845	.4266	.4536	.4694
	1.918			.2637	.3456	.3472	.3213	.2884	.2561	.2268	.2012
	1.792			.1172	.1536	.1543	.1428	.1282	.1136	.1008	.0894
	1.585		.1234	.0879	.0576	.0386	.0268	.0192	.0142	.0108	.0084
	1.459		.4938	.3516	.2304	.1543	.1071	.0769	.0569	.0432	.0335
	1.252		.1234	.0879	.0576	.0386	.0268	.0192	.0142	.0108	.0084
	1.000	.3125	.0823	.0293	.0128	.0064	.0036	.0021	.0014	.0009	.0006
	.918	.4687	.1234	.0439	.0192	.0096	.0054	.0032	.0020	.0013	.0009
	.650	.1875	.0494	.0176	.0076	.0038	.0021	.0013	.0008	.0005	.0004
	.000	.0312	.0041	.0010	.0003	.0001	.0001	.0000	.0000	.0000	.0000
7	2.807						.0061	.0192	.0379	.0604	.0853
	2.522					.0540	.1285	.2019	.2655	.3175	.3585
	2.236				.1612	.2701	.3213	.3364	.3319	.3175	.2987
	2.128				.0537	.0900	.1071	.1122	.1106	.1058	.0996
	1.950			.1538	.1612	.1350	.1071	.0841	.0664	.0529	.0427
	1.842			.3076	.3225	.2701	.2142	.1682	.1328	.1058	.0853
	1.664			.0513	.0537	.0450	.0357	.0280	.0221	.0176	.0142
	1.557		.2881	.1538	.0806	.0450	.0268	.0168	.0111	.0075	.0053
	1.449		.1920	.1025	.0537	.0300	.0178	.0112	.0074	.0050	.0036
	1.379		.2881	.1538	.0806	.0450	.0268	.0168	.0111	.0075	.0053
	1.149		.0576	.0308	.0161	.0090	.0054	.0034	.0022	.0015	.0011
	.985	.5469	.0960	.0256	.0089	.0038	.0018	.0009	.0005	.0003	.0002
	.863	.3281	.0576	.0154	.0054	.0022	.0011	.0006	.0003	.0002	.0001
	.592	.1094	.0192	.0051	.0018	.0008	.0004	.0002	.0001	.0001	.0000
	.000	.0156	.0014	.0002	.0001	.0000	.0000	.0000	.0000	.0000	.0000
8	3.000							.0024	.0084	.0181	.0310
	2.750						.0245	.0673	.1180	.1693	.2172
	2.500					.0900	.1836	.2523	.2950	.3175	.3259
	2.406					.0240	.0490	.0673	.0787	.0847	.0886
	2.250				.1290	.1800	.1836	.1682	.1475	.1270	.1076
	2.156				.1720	.2401	.2448	.2243	.1967	.1693	.1448
	2.000			.0384	.0537	.0525	.0459	.0386	.0320	.0264	.0220
	1.906			.3076	.2580	.1800	.1224	.0841	.0590	.0423	.0310
	1.811			.1025	.0860	.0600	.0408	.0280	.0197	.0141	.0103
	1.750			.1538	.1290	.0900	.0612	.0421	.0295	.0212	.0155
	1.561		.2561	.1025	.0430	.0200	.0102	.0056	.0033	.0020	.0013
	1.549			.0205	.0172	.0120	.0082	.0056	.0039	.0028	.0021
	1.500		.1920	.0769	.0322	.0150	.0076	.0042	.0024	.0015	.0010
	1.406		.2561	.1025	.0430	.0200	.0102	.0056	.0033	.0020	.0013
	1.299		.1536	.0615	.0258	.0120	.0061	.0034	.0020	.0012	.0008
	1.061		.0256	.0102	.0043	.0020	.0010	.0006	.0003	.0002	.0001
	1.000	.2734	.0320	.0064	.0018	.0006	.0002	.0001	.0001	.0000	.0000
	.954	.4375	.0512	.0102	.0029	.0010	.0004	.0002	.0001	.0001	.0000
	.811	.2187	.0256	.0051	.0014	.0005	.0002	.0001	.0000	.0000	.0000
	.544	.0625	.0073	.0015	.0004	.0001	.0001	.0000	.0000	.0000	.0000
	.000	.0078	.0004	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000

TABLE 1 (Continued)

Exact sampling probabilities of $\hat{H}$ for $2 \le k \le 11$, $2 \le n \le 12$, and $p_i = 1/k$.

N	$\hat{H}$	K = 2	3	4	5	6	7	8	9	10	11
9	3.170								.0009	.0036	.0085
	2.948							.0108	.0337	.0653	.1016
	2.725						.0472	.1136	.1770	.2286	.2666
	2.642						.0105	.0252	.0393	.0508	.0592
	2.503					.0900	.1574	.1892	.1967	.1905	.1778
	2.419					.0900	.1574	.1892	.1967	.1905	.1778
	2.281				.0581	.0765	.0748	.0662	.0566	.0476	.0400
	2.197				.2322	.2701	.2361	.1892	.1475	.1143	.0889
	2.113				.0516	.0600	.0524	.0421	.0328	.0254	.0198
	2.059				.0774	.0900	.0787	.0631	.0492	.0381	.0296
	1.975			.1154	.0774	.0450	.0262	.0158	.0098	.0064	.0042
	1.891			.2307	.1548	.0900	.0524	.0315	.0197	.0127	.0085
	1.880				.0077	.0090	.0079	.0063	.0049	.0038	.0030
	1.837			.1730	.1161	.0675	.0393	.0236	.0148	.0095	.0063
	1.753			.1154	.0774	.0450	.0262	.0158	.0098	.0064	.0042
	1.658			.0692	.0464	.0270	.0157	.0095	.0059	.0038	.0025
	1.585		.0854	.0256	.0086	.0033	.0014	.0007	.0004	.0002	.0001
	1.530		.3841	.1154	.0387	.0150	.0066	.0032	.0016	.0009	.0005
	1.447			.0077	.0052	.0030	.0017	.0011	.0006	.0004	.0003
	1.435		.1152	.0346	.0116	.0045	.0020	.0009	.0005	.0003	.0002
	1.392		.0960	.0288	.0097	.0038	.0016	.0008	.0004	.0002	.0001
	1.352		.1536	.0461	.0155	.0060	.0026	.0013	.0006	.0004	.0002
	1.224		.0768	.0231	.0077	.0030	.0013	.0006	.0003	.0002	.0001
	.991	.4922	.0384	.0058	.0013	.0004	.0001	.0001	.0000	.0000	.0000
	.986		.0110	.0033	.0011	.0004	.0002	.0001	.0000	.0000	.0000
	.918	.3281	.0256	.0038	.0009	.0002	.0001	.0000	.0000	.0000	.0000
	.764	.1406	.0110	.0016	.0004	.0001	.0000	.0000	.0000	.0000	.0000
	.503	.0352	.0027	.0004	.0001	.0000	.0000	.0000	.0000	.0000	.0000
	.000	.0039	.0002	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
10	3.322									.0004	.0015
	3.122								.0047	.0163	.0346
	2.922							.0236	.0656	.1143	.1616
	2.846							.0045	.0125	.0218	.0308
	2.722						.0562	.1183	.1639	.1905	.2020
	2.646						.0450	.0946	.1311	.1524	.1616
	2.522					.0563	.0881	.0966	.0929	.0841	.0741
	2.446					.1500	.2248	.2366	.2186	.1905	.1616
	2.371					.0250	.0375	.0394	.0364	.0318	.0269
	2.322				.0116	.0488	.0646	.0651	.0587	.0505	.0424
	2.246				.1548	.1500	.1124	.0788	.0546	.0381	.0269
	2.171				.1548	.1500	.1124	.0788	.0546	.0381	.0269
	2.161					.0030	.0045	.0047	.0044	.0038	.0032
	2.122				.1161	.1125	.0843	.0591	.0410	.0286	.0202
	2.046				.0516	.0500	.0375	.0263	.0182	.0127	.0090
	1.971			.1442	.0774	.0375	.0187	.0098	.0055	.0032	.0019
	1.961				.0310	.0300	.0225	.0158	.0109	.0076	.0054
	1.922			.0721	.0387	.0188	.0094	.0049	.0027	.0016	.0010
	1.895			.0641	.0344	.0167	.0083	.0044	.0024	.0014	.0008
	1.846			.2884	.1548	.0750	.0375	.0197	.0109	.0064	.0038
	1.771				.0026	.0025	.0019	.0013	.0009	.0006	.0004
	1.761			.0865	.0464	.0225	.0112	.0059	.0033	.0019	.0012
	1.722			.0360	.0194	.0094	.0047	.0025	.0014	.0008	.0005
	1.685			.0577	.0310	.0150	.0075	.0039	.0022	.0013	.0008
	1.571		.2134	.0769	.0284	.0117	.0053	.0026	.0014	.0008	.0005
	1.522		.1600	.0360	.0097	.0031	.0012	.0005	.0002	.0001	.0001
	1.485		.2561	.0577	.0155	.0050	.0019	.0008	.0004	.0002	.0001
	1.371		.0640	.0144	.0039	.0012	.0005	.0002	.0001	.0000	.0000
	1.361		.1280	.0288	.0077	.0025	.0009	.0004	.0002	.0001	.0000
	1.357			.0027	.0015	.0007	.0004	.0002	.0001	.0001	.0000
	1.295		.0854	.0192	.0052	.0017	.0006	.0003	.0001	.0001	.0000
	1.157		.0366	.0082	.0022	.0007	.0003	.0001	.0001	.0000	.0000
	1.000	.2461	.0128	.0014	.0002	.0001	.0000	.0000	.0000	.0000	.0000
	.971	.4102	.0213	.0024	.0004	.0001	.0000	.0000	.0000	.0000	.0000
	.922		.0046	.0010	.0003	.0001	.0000	.0000	.0000	.0000	.0000
	.881	.2344	.0122	.0014	.0002	.0001	.0000	.0000	.0000	.0000	.0000
	.722	.0879	.0046	.0005	.0001	.0000	.0000	.0000	.0000	.0000	.0000
	.469	.0195	.0010	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	.000	.0020	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000

TABLE 1 (Continued)

Exact sampling probabilities of $\hat{H}$ for $2 \le k \le 11$, $2 \le n \le 12$, and $p_i = 1/k$.

N	$\hat{H}$	K = 2	3	4	5	6	7	8	9	10	11
11	3.459										.0001
	3.278									.0020	.0077
	3.096								.0114	.0359	.0692
	3.027								.0019	.0060	.0115
	2.914							.0325	.0801	.1257	.1616
	2.845							.0217	.0534	.0838	.1077
	2.732						.0442	.0829	.1040	.1108	.1067
	2.664						.0883	.1626	.2003	.2096	.2020
	2.595						.0118	.0217	.0267	.0279	.0269
	2.550					.0206	.0442	.0569	.0601	.0576	.0525
	2.482					.1375	.1766	.1626	.1336	.1048	.0808
	2.413					.0917	.1178	.1084	.0890	.0698	.0539
	2.404						.0012	.0022	.0027	.0028	.0027
	2.368					.0688	.0883	.0813	.0668	.0524	.0404
	2.300				.0426	.0573	.0515	.0407	.0306	.0227	.0168
	2.231				.1703	.1375	.0883	.0542	.0334	.0210	.0135
	2.222					.0138	.0177	.0163	.0134	.0105	.0081
	2.187				.0852	.0688	.0442	.0271	.0167	.0105	.0067
	2.163				.0378	.0306	.0196	.0120	.0074	.0046	.0030
	2.118				.1703	.1375	.0883	.0542	.0334	.0210	.0135
	2.049					.0009	.0012	.0011	.0009	.0007	.0005
	2.040				.0511	.0412	.0265	.0163	.0100	.0063	.0040
	2.005				.0142	.0115	.0074	.0045	.0028	.0017	.0011
	1.981			.0881	.0378	.0153	.0065	.0030	.0015	.0008	.0004
	1.972				.0227	.0183	.0118	.0072	.0044	.0028	.0018
	1.936			.1983	.0852	.0344	.0147	.0068	.0033	.0017	.0010
	1.868			.1322	.0681	.0321	.0157	.0081	.0044	.0026	.0015
	1.858			.0396	.0170	.0069	.0029	.0014	.0007	.0003	.0002
	1.823			.0991	.0426	.0172	.0074	.0034	.0017	.0009	.0005
	1.790			.1586	.0681	.0275	.0118	.0054	.0027	.0014	.0008
	1.686			.0396	.0170	.0069	.0029	.0014	.0007	.0003	.0002
	1.677			.0396	.0170	.0069	.0029	.0014	.0007	.0003	.0002
	1.673				.0008	.0006	.0004	.0002	.0002	.0001	.0001
	1.617			.0264	.0114	.0046	.0020	.0009	.0004	.0002	.0001
	1.573		.1956	.0330	.0071	.0019	.0006	.0002	.0001	.0000	.0000
	1.539		.1565	.0264	.0057	.0015	.0005	.0002	.0001	.0000	.0000
	1.495		.2347	.0396	.0085	.0023	.0007	.0003	.0001	.0000	.0000
	1.491			.0113	.0049	.0020	.0008	.0004	.0002	.0001	.0001
	1.435		.1565	.0264	.0057	.0015	.0005	.0002	.0001	.0000	.0000
	1.348		.0469	.0079	.0017	.0004	.0001	.0001	.0000	.0000	.0000
	1.322		.0782	.0132	.0028	.0008	.0002	.0001	.0000	.0000	.0000
	1.309		.0335	.0057	.0012	.0003	.0001	.0000	.0000	.0000	.0000
	1.278			.0009	.0004	.0002	.0001	.0000	.0000	.0000	.0000
	1.241		.0447	.0076	.0016	.0004	.0001	.0001	.0000	.0000	.0000
	1.096		.0168	.0028	.0006	.0002	.0001	.0000	.0000	.0000	.0000
	.994	.4512	.0156	.0013	.0002	.0000	.0000	.0000	.0000	.0000	.0000
	.946	.3223	.0112	.0009	.0001	.0000	.0000	.0000	.0000	.0000	.0000
	.866		.0019	.0003	.0001	.0000	.0000	.0000	.0000	.0000	.0000
	.845	.1611	.0056	.0005	.0001	.0000	.0000	.0000	.0000	.0000	.0000
	.684	.0537	.0019	.0002	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	.440	.0107	.0004	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	.000	.0010	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
12	3.418										.0008
	3.252									.0054	.0189
	3.189									.0008	.0028
	3.085								.0178	.0503	.0881
	3.022								.0102	.0287	.0504
	2.918							.0305	.0674	.0961	.1153
	2.855							.0488	.1068	.1509	.1763
	2.792							.0054	.0119	.0168	.0196
	2.752						.0227	.0447	.0579	.0629	.0624
	2.689						.1009	.1626	.1781	.1676	.1469
	2.626						.0505	.0813	.0890	.0838	.0734
	2.617							.0005	.0010	.0014	.0017
	2.585					.0034	.0416	.0640	.0690	.0644	.0562
	2.522					.0688	.0858	.0772	.0623	.0482	.0367
	2.459					.1375	.1514	.1220	.0890	.0629	.0441
	2.451						.0061	.0098	.0107	.0101	.0088
	2.418					.0688	.0757	.0610	.0445	.0314	.0220

TABLE 1 (Continued)

Exact sampling probabilities of $\hat{H}$ for $2 \le k \le 11$, $2 \le n \le 12$, and $p_i = 1/k$.

N	$\hat{H}$	K = 2	3	4	5	6	7	8	9	10	11
12	2.396					.0204	.0224	.0181	.0132	.0093	.0065
	2.355					.0917	.1009	.0813	.0594	.0419	.0294
	2.292				.0681	.0458	.0256	.0141	.0080	.0048	.0029
	2.284					.0275	.0303	.0244	.0178	.0126	.0088
	2.252				.0255	.0229	.0158	.0102	.0065	.0042	.0028
	2.230				.0908	.0611	.0336	.0181	.0099	.0056	.0033
	2.221					.0092	.0101	.0081	.0059	.0042	.0029
	2.189				.2044	.1375	.0757	.0407	.0223	.0126	.0073
	2.126				.0681	.0504	.0303	.0176	.0104	.0063	.0039
	2.117				.0409	.0275	.0151	.0081	.0044	.0025	.0015
	2.085				.0511	.0344	.0189	.0102	.0056	.0031	.0018
	2.054				.0817	.0550	.0303	.0163	.0089	.0050	.0029
	2.000			.0220	.0076	.0025	.0009	.0004	.0002	.0001	.0000
	1.959			.1983	.0886	.0367	.0160	.0074	.0037	.0020	.0011
	1.951				.0136	.0092	.0050	.0027	.0015	.0008	.0005
	1.947					.0003	.0003	.0002	.0002	.0001	.0001
	1.918			.0744	.0255	.0086	.0032	.0013	.0006	.0003	.0001
	1.896				.0091	.0061	.0034	.0018	.0010	.0006	.0003
	1.888			.1190	.0409	.0138	.0050	.0020	.0009	.0004	.0002
	1.855			.0991	.0341	.0115	.0042	.0017	.0007	.0003	.0002
	1.825			.0793	.0272	.0092	.0034	.0014	.0006	.0003	.0001
	1.792			.0198	.0068	.0023	.0008	.0003	.0001	.0001	.0000
	1.784			.1190	.0409	.0138	.0050	.0020	.0009	.0004	.0002
	1.781				.0039	.0026	.0014	.0008	.0004	.0002	.0001
	1.730			.0793	.0272	.0092	.0034	.0014	.0006	.0003	.0001
	1.650			.0119	.0041	.0014	.0005	.0002	.0001	.0000	.0000
	1.626			.0198	.0068	.0023	.0008	.0003	.0001	.0001	.0000
	1.614			.0170	.0058	.0020	.0007	.0003	.0001	.0001	.0000
	1.585		.0652	.0083	.0017	.0005	.0002	.0001	.0000	.0000	.0000
	1.554		.3130	.0396	.0068	.0015	.0004	.0001	.0000	.0000	.0000
	1.551			.0113	.0039	.0013	.0005	.0002	.0001	.0000	.0000
	1.500		.1043	.0132	.0023	.0005	.0001	.0000	.0000	.0000	.0000
	1.483		.0939	.0119	.0020	.0004	.0001	.0000	.0000	.0000	.0000
	1.459		.1565	.0198	.0034	.0008	.0002	.0001	.0000	.0000	.0000
	1.418			.0042	.0014	.0005	.0002	.0001	.0000	.0000	.0000
	1.384		.0894	.0113	.0019	.0004	.0001	.0000	.0000	.0000	.0000
	1.325		.0626	.0079	.0014	.0003	.0001	.0000	.0000	.0000	.0000
	1.281		.0447	.0057	.0010	.0002	.0001	.0000	.0000	.0000	.0000
	1.252		.0168	.0021	.0004	.0001	.0000	.0000	.0000	.0000	.0000
	1.208			.0003	.0001	.0000	.0000	.0000	.0000	.0000	.0000
	1.189		.0224	.0028	.0005	.0001	.0000	.0000	.0000	.0000	.0000
	1.041		.0074	.0009	.0002	.0000	.0000	.0000	.0000	.0000	.0000
	1.000	.2256	.0052	.0003	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	.980	.3867	.0089	.0006	.0001	.0000	.0000	.0000	.0000	.0000	.0000
	.918	.2417	.0056	.0004	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	.817		.0007	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	.811	.1074	.0025	.0002	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	.650	.0322	.0007	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	.414	.0058	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	.000	.0005	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000

AD Accession No.

US Army Medical Research Lab, Ft. Knox, Ky.
A GENERAL METHOD OF OBTAINING EXACT SAMPLING PROBABILITIES OF THE SHANNON-WIENER MEASURE OF INFORMATION, H. - J. N. Cronholm
Report No. 575, 7 Jun 63, 9 pp & i - 1 table
DA Project No. 3A012001B801, Unclassified Report

UNCLASSIFIED
1. Statistics
2. Non-parametric
3. Information

A two-variable generating function is described which yields the sampling probabilities of the Shannon-Wiener information measure. Expansion and collection of terms in like powers of the first variable imposes the restriction that the sum of the k category frequencies equal n; collection of terms in like powers of the second variable then produces terms whose coefficients are the required probabilities. The method may be used with either equal or unequal category probabilities for any finite n and k, and thus represents a general solution to the small sample problem. Tables of sampling probabilities are presented.

AD Accession No.

US Army Medical Research Lab, Ft. Knox, Ky.
A GENERAL METHOD OF OBTAINING EXACT SAMPLING PROBABILITIES OF THE SHANNON-WIENER MEASURE OF INFORMATION, H. - J. N. Cronholm
Report No. 575, 7 Jun 63, 9 pp & i - 1 table
DA Project No. 3A012001B801, Unclassified Report

UNCLASSIFIED
1. Statistics
2. Non-parametric
3. Information

A two-variable generating function is described which yields the sampling probabilities of the Shannon-Wiener information measure. Expansion and collection of terms in like powers of the first variable imposes the restriction that the sum of the k category frequencies equal n; collection of terms in like powers of the second variable then produces terms whose coefficients are the required probabilities. The method may be used with either equal or unequal category probabilities for any finite n and k, and thus represents a general solution to the small sample problem. Tables of sampling probabilities are presented.

AD Accession No.

US Army Medical Research Lab, Ft. Knox, Ky.
A GENERAL METHOD OF OBTAINING EXACT SAMPLING PROBABILITIES OF THE SHANNON-WIENER MEASURE OF INFORMATION, H. - J. N. Cronholm
Report No. 575, 7 Jun 63, 9 pp & i - 1 table
DA Project No. 3A012001B801, Unclassified Report

UNCLASSIFIED
1. Statistics
2. Non-parametric
3. Information

A two-variable generating function is described which yields the sampling probabilities of the Shannon-Wiener information measure. Expansion and collection of terms in like powers of the first variable imposes the restriction that the sum of the k category frequencies equal n; collection of terms in like powers of the second variable then produces terms whose coefficients are the required probabilities. The method may be used with either equal or unequal category probabilities for any finite n and k, and thus represents a general solution to the small sample problem. Tables of sampling probabilities are presented.

AD Accession No.

US Army Medical Research Lab, Ft. Knox, Ky.
A GENERAL METHOD OF OBTAINING EXACT SAMPLING PROBABILITIES OF THE SHANNON-WIENER MEASURE OF INFORMATION, H. - J. N. Cronholm
Report No. 575, 7 Jun 63, 9 pp & i - 1 table
DA Project No. 3A012001B801, Unclassified Report

UNCLASSIFIED
1. Statistics
2. Non-parametric
3. Information

A two-variable generating function is described which yields the sampling probabilities of the Shannon-Wiener information measure. Expansion and collection of terms in like powers of the first variable imposes the restriction that the sum of the k category frequencies equal n; collection of terms in like powers of the second variable then produces terms whose coefficients are the required probabilities. The method may be used with either equal or unequal category probabilities for any finite n and k, and thus represents a general solution to the small sample problem. Tables of sampling probabilities are presented.