AN ABSTRACT OF THE THESIS OF

Preecha Sakarindr (Name) for the M. S. (Degree) in Statistics (Major)

Date thesis is presented December 3, 1964

Title MULTINOMIAL ESTIMATION FROM CENSORED SAMPLES

Abstract approved _____ (Major professor)

Consider the estimation of the category proportions in a multinomial population from a sample which is "censored" in the sense that under an appropriate, unknown permutation of the sample categories, the population proportions are all known. We are considering the estimation of an ordered set of sample proportions, known except for their order. The estimation problem reduces to one of matching a set of known sample proportions with a set of known population proportions. The method of maximum likelihood yields the matching that common sense or one's intuition would suggest; highest sample proportion associated with highest population proportion, second highest sample proportion with second highest population proportion, and so forth. The work of this thesis is to examine, by a complete enumeration of cases for some simple problems, how good the method of maximum likelihood is. We study the effectiveness of maximum likelihood matching under variation of the three factors (1) "roughness" of the set of population

proportions, (2) number of categories of the multinomial population, and (3) size of the sample. The effectiveness of maximum likelihood matching is measured by the ratio "proportion of the time maximum likelihood matching correct" divided by "proportion of the time random matching correct". The empirical study confirms the conclusions suggested by intuition that: (1) the greater the "roughness" of the set of population proportions, the more effective is the method of maximum likelihood; (2) the greater the number of categories the more effective is the method of maximum likelihood; and (3) the greater the sample size the more effective is the method of maximum likelihood.

MULTINOMIAL ESTIMATION FROM CENSORED SAMPLES

by

PREECHA SAKARINDR

A THESIS

submitted to

OREGON STATE UNIVERSITY

in partial fulfillment of
the requirements for the
degree of

MASTER OF SCIENCE

June 1965

APPROVED:

Redacted for Privacy

Assistant Professor of Statistics

In Charge of Major

Redacted for Privacy

Chairman of Department of Statistics

Redacted for Privacy

Dean of Graduate School

Date thesis is presented December 3, 1964

Typed by Carol Baker

ACKNOWLEDGMENT

The author wishes to express his sincere appreciation to Dr. Edwin Joseph Hughes, Assistant Professor of Statistics, Oregon State University, for his invaluable assistance and guidance throughout the successful completion of this thesis.

TABLE OF CONTENTS

LIST OF TABLES

MULTINOMIAL ESTIMATION FROM CENSORED SAMPLES

CHAPTER 1

INTRODUCTION

This thesis is one approach to the problem of how to estimate the category proportions in a multinomial population from a sample which is "censored" as described in Chapter 2. An interesting example of this thesis problem is realized in the problem of solving a simple substitution cipher such as the secret message which is written in English. The problem is to match off the message (sample) proportions of letter A, B, C, ... , Y, Z with the English (population) proportions, which is explained in detail in Chapter 3. Fisher's well-known method of "maximum likelihood" has been used to attack this problem of estimating population proportions by matching sample proportions with population proportions. Another objective of this thesis is to study the effectiveness of maximum likelihood matching as compared with the method of random matching. Both maximum likelihood and random methods of matching and also the effectiveness of maximum likelihood are detailed in Chapter 4. Chapter 5 contains the table of results from the complete enumeration of cases for simple problems as given at the end of the previous chapter. Also included in Chapter 5 is the detailed discussion about the probabilities

and effectiveness of maximum likelihood matching. Finally the conclusions are given. The last chapter contains proposals for further work relating to this problem. In the Appendix some examples of the empirical study for simple problems are given.

CHAPTER 2

THE PROBLEM

The problem specified in the thesis title in simple language involves the estimation of the category proportions in a multinomial population from a sample which is "censored" in the sense that only a set of proportions is observed, and there is an unknown correspondence between sample proportions and population proportions. The set of population proportions is known, so that the problem of estimating category proportions reduces to a problem of establishing the correspondence between sample proportions and population proportions.

For example, in sampling from a k-nomial population, a random sample of size n is drawn and the sample frequencies $X_{i_1}, X_{i_2}, \ldots, X_{i_k}$ are observed where

k-population proportions $P_1, P_2, P_3, \ldots, P_k$ are known

and k-sample proportions $\bar{X}_{i_1}, \bar{X}_{i_2}, \bar{X}_{i_3}, \ldots, \bar{X}_{i_k}$ are known

as

$$X_{i_j} = \sum_{t=1}^{n} x_t \; ; \; x_t = \begin{cases} 1 \text{ if } t^{th} \text{ observation in } i_j \text{ category} \\ 0 \quad \text{Otherwise; } i_j \neq j \quad j = 1, 2, \ldots k. \end{cases}$$

$$\bar{X}_{i_j} = \frac{1}{n} X_{i_j} = \frac{1}{n} \sum_{t=1}^{n} x_t \text{ is the } i_j^{th} \text{ sample proportion.}$$

The problem of estimating category proportions reduces to a problem of establishing the correspondence between the set of population proportions $P_1, P_2, P_3, \ldots, P_k$ and the set of sample proportions $\bar{X}_{i_1}, \bar{X}_{i_2}, \bar{X}_{i_3}, \ldots, \bar{X}_{i_k}$, that is, establishing which permutation of 1, 2, , ... , k is the set $i_1, i_2, \ldots, i_k$.

One could contrast this problem with the usual problem of estimation of the population proportions from the known sample proportions in which the sample is not "censored".

The random sample of size n is drawn from the k-nomial distribution and sample frequencies $X_1, X_2, \ldots, X_k$ are observed where k-population proportions $P_1, P_2, P_3, \ldots, P_k$ are unknown and k-sample proportions $\bar{X}_1, \bar{X}_2, \bar{X}_3, \ldots, \bar{X}_k$ are known as

$$X_j = \sum_{i=1}^{n} x_i \quad ; \quad x_i = \begin{cases} 1 \text{ in } j^{th} \text{ category; } j=1, 2, \ldots, k \\ 0 \quad \text{Otherwise} \end{cases}$$

$$\bar{X}_j = \frac{1}{n} X_j = \frac{1}{n} \sum_{i=1}^{n} x_i \text{ is the } j^{th} \text{ sample proportion.}$$

The problem is to estimate the set of population proportions $P_1, P_2, \ldots, P_k$ from the set of sample proportions $\bar{X}_1, \bar{X}_2, \ldots, \bar{X}_k$ or estimate P_i; i = 1, 2, ... , k .

The solution is that the population proportions will be estimated by the corresponding sample proportions, or

$P_i = \bar{X}_i, \quad i = 1, 2, \ldots, k.$

An interesting example of the problem of this thesis is the solution of a simple substitution cipher. The sample consists of a set of letter frequencies which must be matched with known letter frequencies of the language. A correct matching of sample frequencies to population frequencies identifies the simple substitution cipher and renders the cipher readable.

The method of maximum likelihood is used to provide a procedure for establishing the correspondence between sample proportions and population proportions. It makes the highest sample proportion correspondto the highest population proportion, the second highest sample proportion correspond to the second highest population proportion, and so forth. Such a procedure has considerable intuitive appeal. Its merits are studied by comparison with a procedure of random matching of sample proportions to population proportions, and the influence of the factors (1) variation among population proportions, (2) number of categories, and (3) sample size is examined. A complete enumeration of cases is made for certain simple problems to obtain exact data on the merits of the maximum likelihood procedure and the influence of the several factors.

CHAPTER 3

EXAMPLE OF SIMPLE SUBSTITUTION CIPHER

The earliest appearances of cipher were among the well-educated Greeks and in the late Roman Republic [7] where the two great classes of cipher seem to have been invented, respectively; (1) Transposition Cipher, in which the letters of the original message are thrown into some meaningless order and (2) Substitution Cipher, in which each letter of the original message is replaced by some other letter, symbol or figure [7, 10]. Substitution cipher is divided into two classes, simple and multiple substitution ciphers. Simple substitution cipher has each letter of the original message represented by one and always the same letter, symbol, or figure. For example, if the original message is "TODAY IS A GOOD DAY", and the letters of the original (plain text) message are represented as in the table below,

PLAIN	A	B	C	D	E	F	X	Y	Z
CIPHER	E	F	G	H	I	J	B	C	D

then the resulting (cipher text) message will read XSHEC MW E KSSH HEC.

Multiple substitution cipher involves multiple use of simple substitution cipher as, for example, the use of one simple

substitution cipher for the first ten letters of the message, a second simple substitution cipher for the second ten letters, and so forth. For example, the first ten letters are represented as in the table

PLAIN	A	B	C	D	E	F	X	Y	Z
CIPHER	D	E	F	G	H	I	A	B	C

The second ten letters are represented as in the table

PLAIN	A	B	C	D	E	F	X	Y	Z
CIPHER	E	F	G	H	I	J	B	C	D

and so forth.

Let us illustrate the decipherment of a simple substitution cipher in English below. The sample is a message of unknown content and consists of a set of letter frequencies which must be matched with known letter frequencies of the language. A correct matching of sample frequencies to population frequencies identifies the simple substitution cipher and renders the cipher readable as in the following message.

1	2	3	4	5	6
SZPQP	ERJKQ	PCRKJ	VZXPU	PJSZP	GKRSC

7	8	9	10	11	12	13
GCSPT	QIQXL	SKNQC	LZPQR	ZKTFM	ZPRES	CSPFK

14	15	16	17	18	19	
JNKUP	QCREG	LFPRT	HRSES	TSEKJ	IELZP	Q

(The groups are numbered for convenience in referring to them.)

The first step in decipherment is to count the frequency with which each letter appears and draw up a frequency table as follows

P = 13	T = 4
S = 10	F = 3
Q = 8	G = 3
K = 8	I = 2
R = 8	U = 2
Z = 7	X = 2
C = 6	N = 2
E = 6	M = 1
J = 5	H = 1
L = 4	V = 1

where P and S are the first and second highest frequencies followed by Q, K, R, Z, etc.

We now turn to the table of letter frequencies [7, p. 252, Table I]. Here we have found that letter E is the most frequent letter in English with T next. We use the maximum likelihood matching so that the highest sample proportion corresponds to the highest population proportion, the second highest sample proportion to the second highest population proportion, and so forth. Therefore, it seems likely that P is E and S is T. For convenience, we are

going to rewrite the message with the provisional values, P = e and S = t, in replacement.

1	2	3	4	5	6
tZeQe	ERJKQ	eCRKJ	VZXeU	eJtZe	GKRTC

7	8	9	10	11	12	13
GCteT	QIQXL	tKNQC	LZeQR	ZKTFM	ZeREt	CteFK

14	15	16	17	18	19	
JNKUe	QCREG	LFeRT	HRtEt	TtEKJ	IELZe	Q .

Consider the next letters in order of frequency are A, O, N, R, I and S. In the message under consideration this corresponds very well with the high frequencies of the letters K, Q, R, Z, C and E; but both in the message and in the frequency table these six letters are closely grouped. It is very difficult to tell which was which without any more information. We take a short cut by consulting the table of Bigrams and Trigrams [7, p. 260-264, Tables VIII, XII, respectively]. These show that TH, THE are the most frequent of two and three letter combinations in the language. Reference to groups 1 and 5 the combination T-blank-E occurs twice, and in both the blank is represented by the same letter of the cipher Z, then Z should represent H as the maximum likelihood matching.

If Z is H, then Q is probably R or S; for with the insertion of the H group 1 reads THE-blank-E, which is a strong possibility for THERE or THESE. However in groups 10 and 19-20

the combination ZPQ occurs and the ZP has been solved as HE, reference to Trigram table shows that HER is one of the most common, while HES is relatively rare. Then Q is more likely to be R. And in group 17-18 occurs the combination SESTSE which has been partially solved to read T-blank-T-blank-T-blank, which with the repeated E's, constitutes a pattern word. Therefore, we look at the table of common or high frequency pattern words [7, p. 263, Table XI] and will discover that this pattern usually occurrs as TITUTI or TETATE. Since we have found P = E in this cipher, then the pattern must represent the first of these two combinations as E = I and T = U.

The group of letters that show high frequencies in the message which now remains unsolved is R, C, K and J. Of the high-frequency letters for which no values in the message have been found there remain A, O, N and S. Two of these drop into place with the acceptance of the TITUTI combination, which can hardly end in anything but ON, yielding K = O and J = N .

If this is correct, groups 1-2 now read THERE I-NOR or THERE I NO R, which makes it apparent that R = S.

Finally this leaves only one letter in the high-frequency group (A) and in the message (C), that is C = A.

Once more filling in, we have

1	2	3	4	5
there/	is/ no/ r	eason/	VhXeU	en/ the/

6	7	8	9	10	11	12
Gost/ a	Gateu	r/ IrXL	toNra	Lhers	houFM/	hesit

13	14	15	16	17	18	19	
ate/ Fo	nNoUe	r/ a/ siG	LFe/ su	Hstit	ution/	IiLhe/	r

Obviously nothing will do at the end of group 11 but the letters L and D to complete the word "should" which gives the correspondences F = L and M = D.

Also replacing G of the message in group 6 with M yields a satisfactory result, and U in groups 4 and 14 work out nicely as V. LON-blank in group 13-14 now becomes clear as LONG, and H = B is required in group 17. Thus the remainder can be filled in: V = W; X = Y; L = P; I = C.

Finally the message is solved using the correspondences

PLAIN	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
CIPHER	C	H	I	M	P	A	N	Z	E	E		F	G	J	K	L		Q	R	S	T	U	V		X	

.

Now we can write down the original message as

1	2	3	4	5
There/	is/ no/ r	eason/	why/ ev	en/ the

6	7	8	9	10	11	12
most/ a	mateu	r/ cryp	togra	pher/ s	hould/	hesit

13	14	15	16	17	18	19	
ate/lo	ng/ ove	r/ a/ sim	ple/ su	bstit	ution/	ciphe	r/.

CHAPTER 4

THE METHOD OF MAXIMUM LIKELIHOOD

Principle of Maximum Likelihood

One of the most widely used techniques for obtaining a desirable estimate and one which is closely connected with other desirable criteria is the method of maximum likelihood. In order to define the maximum likelihood estimates, let us first define and interpret the likelihood function. The likelihood function of n-random variables $x_1, x_2, \ldots, x_n$ is the joint density of the n-random variables, $\phi(x_1, x_2, \ldots, x_n; \Theta)$, which is considered to be a function of parameter vector Θ. Particularly, if $x_1, x_2, \ldots, x_n$ is a random sample from the density $f(x, \Theta)$. Then the likelihood function is

$$\phi(x_1, x_2, \ldots, x_n; \Theta) = f(x_1, \Theta) \cdot f(x_2, \Theta), \ldots, f(x_n, \Theta)$$

$$= \prod_{i=1}^{n} f(x_i, \Theta) . \qquad (1)$$

The maximum likelihood estimate $\hat{\Theta} = \hat{\Theta}(x_1, x_2, \ldots, x_n)$ of Θ is that value Θ among all values in the paramater space $\mathbf{\Theta}$ which maximizes the likelihood function (if such a value exists). That is

$$\phi(x_1, x_2, \ldots, x_n; \hat{\Theta}) \geq \phi(x_1, x_2, \ldots, x_n; \Theta) \text{ for all } \Theta \in \mathbf{\Theta}.$$

If certain regularity conditions are satisfied, then the maximum likelihood estimate $(\hat{\theta})$ is usually obtained by setting the derivation with respect to the unknown parameter equal to zero and solving for the unknown as

$$0 = \frac{\partial}{\partial\hat{\theta}} \phi(x_1, x_2, \ldots, x_n;\theta) = \frac{\partial}{\partial\hat{\theta}} \prod_{i=1}^{n} f(x_1, \theta) .$$

Equivalently, $\hat{\theta}$ is found

$$0 = \frac{\partial}{\partial\hat{\theta}} \ell n\, \phi(x_1, x_2, \ldots, x_n;\theta) = \sum_{i=1}^{n} \frac{\partial}{\partial\hat{\theta}} \ell n\, f(x_i;\theta) .$$

Let us illustrate the method of maximum likelihood to estimate the parameter of a binomial distribution.

Example 1. Suppose that a random sample of size n is drawn from the point binomial distribution.

$$f(x, p) = p^x q^{1-x} \quad ; \quad x = 0, 1; \quad 0 \leq p \leq 1$$

The sample values are $x_1, x_2, \ldots, x_n$, a sequence of 0's and 1's.

Thus the likelihood function is

$$L(p) = \prod_{i=1}^{n} p^{x_i} q^{1-x_i} = p^{\Sigma x_i} q^{n-\Sigma x_i}$$

or

$$L^* = \log L(p) = \Sigma x_i \log p + (n - \Sigma x_i) \log q .$$

$$\frac{dL^*}{d\hat{p}} = \frac{\Sigma x_i}{\hat{p}} - \frac{(n-\Sigma x_i)}{\hat{q}} = 0$$

$$\frac{\Sigma x_i}{\hat{p}} = \frac{(n-\Sigma x_i)}{1-\hat{p}} \qquad \because \hat{q} = 1-\hat{p}$$

$$\Sigma x_i - \hat{p}\Sigma x_i = \hat{p}n - \hat{p}\Sigma x_i$$

$$\hat{p} = \frac{\Sigma x_i}{n} = \overline{X}$$

That is, $\overline{X}$ is the maximum likelihood estimate of p.

Next we illustrate the method of maximum likelihood to estimate the two parameters of a normal distibution.

Example 2. A random sample of size n from normal distribution has the likelihood function,

$$L(\mu, \sigma^2) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{1}{2}\left(\frac{x_i-\mu}{\sigma}\right)^2}$$

$$= \left(\frac{1}{2\pi\sigma^2}\right)^{n/2} e^{-\frac{1}{2}\frac{\Sigma(x_i-\mu)^2}{\sigma^2}}$$

or $$L^*(\mu, \sigma^2) = \log L(\mu, \sigma^2)$$

$$= \frac{-n}{2}\log\sigma^2 - \frac{n}{2}\log 2\pi - \frac{1}{2\sigma^2}\Sigma(x_i-\mu)^2 .$$

Differentiating with respect to $\hat{\mu}$ and $\hat{\sigma}^2$, we obtain

$$\frac{\partial L^*}{\partial \hat{\mu}} = \frac{1}{\hat{\sigma}^2} \Sigma(x_i - \hat{\mu}) = 0$$

and

$$\frac{\partial L^*}{\partial \hat{\sigma}^2} = \frac{-n}{2\hat{\sigma}^2} + \frac{1}{2\hat{\sigma}^4} \Sigma(x_i - \hat{\mu})^2 = 0$$

By solving the simultaneous equations for $\hat{\mu}$ and $\hat{\sigma}^2$ we find the estimators

$$\hat{\mu} = \frac{1}{n} \Sigma x_i$$

$$= \bar{X}$$

and

$$\hat{\sigma}^2 = \frac{1}{n} \Sigma (x_i - \bar{X})^2 \quad .$$

Consider the method of maximum likelihood as applied to the problem of this thesis.

Let $X_{i_1}, X_{i_2}, \ldots, X_{i_k}$ be the k sample frequencies, $\bar{X}_{i_1}, \bar{X}_{i_2}, \ldots, \bar{X}_{i_k}$ the k sample proportions, and $P_1, P_2, \ldots, P_k$ the k population proportions from a k-nomial population. The problem is to determine the correspondence between the set of sample proportions and the set of population proportions. Let $P_1 \geq P_2 \geq \ldots P_k$. We shall see that if $\bar{X}_1, \bar{X}_2, \ldots, \bar{X}_k$ is the reordering of $\bar{X}_{i_1}, \bar{X}_{i_2}, \ldots, \bar{X}_{i_k}$ such that $\bar{X}_1 \geq \bar{X}_2 \geq \ldots \bar{X}_k$, then the method of maximum likelihood gives the correspondence $\bar{X}_i$

corresponds to P_i for $i = 1, 2, \ldots, k$. Thus the method of maximum likelihood yields the matching that common sense or one's intuition would suggest: highest sample proportion associated with highest population proportion, ... etc.

Let L_1 = likelihood function for order #1

$$= \phi(X_{i_1}, X_{i_2}, \ldots, X_{i_k}, P)$$

$$= \frac{n!}{X_{i_1}! X_{i_2}! \ldots X_{i_k}!} P_1^{X_{i_1}} P_2^{X_{i_2}} \ldots P_k^{X_{i_k}}$$

L_2 = likelihood function for order # 2

$$= \phi(X_{i_1}, X_{i_3}, \ldots, X_{i_k}; P)$$

$$= \frac{n!}{X_{i_1}! X_{i_3}! \ldots X_{i_k}!} P_1^{X_{i_1}} P_2^{X_{i_3}} \ldots P_k^{X_{i_k}}$$

$$\vdots$$

$L_{k!-1}$ = likelihood function for order #k! -1

$$= \phi(X_{i_k}, X_{i_2}, \ldots, X_{i_{k-1}}; P)$$

$$= \frac{n!}{X_{i_k}! X_{i_2}! \ldots X_{i_{k-1}}!} P_1^{X_{i_k}} P_2^{X_{i_2}} \ldots P^{X_{i_{k-1}}}$$

$L_{k!}$ = likelihood function for order #k!

$$= \phi(X_{i_k}, X_{i_1}, \ldots, X_{i_{k-1}}; P)$$

$$= \frac{n!}{X_{i_k}! X_{i_1}! \ldots X_{i_{k-1}}!} P_1^{X_{i_k}} P_2^{X_{i_1}} \ldots P_k^{X_{i_{k-1}}}.$$

That is the biggest function of L, say L_j will give the maximum likelihood matching of k-random sample frequencies $X_{i_1}, X_{i_2}, \ldots, X_{i_k}$ with k-population proportions $P_1, P_2, \ldots, P_k$.

We shall see that L_j has highest sample frequency X_1 (highest sample proportion $\bar{X}_1 = \frac{X_1}{n}$) associated with highest population proportion P_1, second highest sample frequency X_2 (second highest sample proportion $\bar{X}_2 = \frac{X_2}{n}$) associated with second highest population proportion P_2, and so forth. This statement can be proved as given below. Suppose

$$P_1 \geqslant P_2 \geqslant P_3 \geqslant \ldots \geqslant P_k \geqslant 0$$

and

$$X_1 \geqslant X_2 \geqslant X_3 \geqslant \ldots \geqslant X_k \qquad \text{(known)}$$

where

$$X_i \in (0, 1, 2, \ldots, n), \quad i = 1, 2, \ldots, k$$

$$\sum_{i=1}^{k} X_i = n > 0$$

and $X_{i_1}, X_{i_2}, \ldots, X_{i_k}$ represents $X_1, X_2, \ldots, X_k$ in <u>some order</u> of all possible k! orders.

Then we have to prove that

$$P_1^{X_1} \cdot P_2^{X_2} \cdots P_k^{X_k} \geqslant P_1^{X_{i_1}} \cdot P_2^{X_{i_2}} \cdots P_k^{X_{i_k}} .$$

<u>Proof</u> Clearly $X_1 \geqslant X_{i_1}$

$$X_1 + X_2 \geqslant X_{i_1} + X_{i_2}$$

and in general,

$$X_1 + X_2 + \ldots\ X_j \geq X_{i_1} + X_{i_2} + \ldots\ X_{i_j} \quad \text{for} \begin{cases} j = 1, 2, \ldots, k \\ i_j \neq j \end{cases}$$

because the sum of the j largest X's are at least as large as or larger than the sum of one arbitrary set of j X's.

Let
$$L = P_1^{X_1} \cdot P_2^{X_2} \ldots P_k^{X_k}$$

and
$$L' = P_1^{X_{i_1}} \cdot P_2^{X_{i_2}} \ldots P_k^{X_{i_k}}$$

or

$$\text{Log } L = X_1 \log P_1 + X_2 \log P_2 + \ldots\ X_k \log P_k$$

$$\text{Log } L' = X_{i_1} \log P_1 + X_{i_2} \log P_2 + \ldots\ X_{i_k} \log P_k$$

rewritten in form of additive terms of logarithms for convenience in comparing.

$$\text{Log}L = \underbrace{(\log P_1 + \log P_1 + \ldots \log P_1)}_{X_1 \text{ terms}} + \underbrace{(\log P_2 + \log P_2 + \ldots \log P_2)}_{X_2 \text{ terms}} + \ldots \underbrace{(\log P_k + \log P_k + \ldots \log P_k)}_{X_k \text{ terms}}$$

$$\text{Log}L' = \underbrace{(\log P_1 + \log P_1 + \ldots \log P_1)}_{X_{i_1} \text{ terms}} + \underbrace{(\log P_2 + \log P_2 + \ldots \log P_2)}_{X_{i_2} \text{ terms}} + \ldots \underbrace{(\log P_k + \log P_k + \ldots \log P_k)}_{X_{i_k} \text{ terms}}.$$

Let us consider the first X_1 terms of Log$L$, which are all $\log P_1$. Since $X_1 \geq X_{i_1}$, then all $\log P_1$ of Log L' are included within the first X_1 terms. That is, the first X_1 terms of Log L' must involve $\log P_1$, $\log P_2$, ... ; all $\leq \log P_1$. Hence the first X_1 terms of Log$L' \leq$ the first X_1 terms of Log L. The next X_2 terms of the Log L

are all log P_2. The next X_2 terms of Log L' involve log P_2, logP_3, ...; all $\leqslant$ log P_2, because no more log P_1 left in Log L'. Therefore the next X_2 terms of Log L' $\leqslant$ next X_2 terms of Log L. The next X_3 terms of Log L are all log P_3.

Since $X_{i_1} + X_{i_2} \geqslant X_i + X_2$, then all log P_1 and log P_2 of Log L' will be within the first $X_1 + X_2$ terms. So that the next X_3 terms of Log L' must involve only log P_3, log P_4, ...; all $\leqslant$ log P_3. Hence the next X_3 terms of Log L' $\leqslant$ the next X_3 terms of Log L and so forth.

The last X_k term of Log L' $\leqslant$ the last X_k term of Log L. Hence, finally the sum of all n terms of Log L' $\leqslant$ the sum of all n terms of Log L.

Therefore Log L' $\leqslant$ Log L or Log L $\geqslant$ log L'

Whence $L \geqslant L'$

Or $$P_1^{X_1} \cdot P_2^{X_2} \cdot \ldots P_k^{X_k} \geqslant P_1^{X_{i_1}} \cdot P_2^{X_{i_2}} \cdot \ldots P_k^{X_{i_k}} .$$

Then the statement above has been proved.

For an example in Trinomial distribution, suppose $X_{i_1}, X_{i_2}, X_{i_3}$ be a sample frequency from a sample of size n = 10 with population proportions $P_1 = .5$; $P_2 = .3$; $P_3 = .2$ and $X_{i_1} = 7$; $X_{i_2} = 2$; $X_{i_3} = 1$. Therefore, the maximum likelihood function will be

$$L = \frac{10!}{7!\,2!\,1!} (.5)^7 (.3)^2 (.2)^1$$

where i_1, i_2, i_3 is in the order of 1, 2, 3 respectively.

Note: In the above example, suppose $X_{i_1} = 8$; $X_{i_2} = 1$ and $X_{i_3} = 1$.

When, because of "ties" such as $X_{i_2} = 1$ and $X_{i_3} = 1$, the likelihood function does not have a unique solution, but rather a set of k solutions (here k = 2!), then the maximum likelihood estimate is obtained by a random selection from among the k solutions. If one of the k-solutions involves the correct matching of sample proportions to population proportions, then the chance of the maximum likelihood estimate being correct is $\frac{1}{k}$ times the chance of the occurrence of a solution of the likelihood equation, which is

$$L = \frac{1}{k} \cdot \frac{n!}{X_{i_1}! X_{i_2}! X_{i_3}!} P_1^{X_{i_1}} P_2^{X_{i_2}} P_3^{X_{i_3}}$$

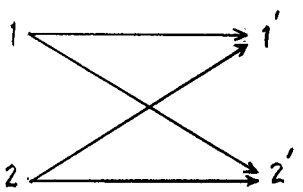
This fact is taken into account in the probability computations in the Appendix.

The Method of Random Matching

The method of random matching consists in selecting at random one of the k! reorderings of the sample proportions $\bar{X}_{i_1}, \bar{X}_{i_2}, \ldots, \bar{X}_{i_k}$ to correspond to the population proportions $P_1, P_2, \ldots, P_k$, where $P_1 \geq P_2 \geq \ldots \geq P_k$. The reordering $\bar{X}_1, \bar{X}_2, \ldots, \bar{X}_k$ of $\bar{X}_{i_1}, \bar{X}_{i_2}, \ldots, \bar{X}_{i_k}$ has $\bar{X}_i$ corresponding to

P_i, $i = 1, 2, \ldots, k$. We can call this method that the "Guessing Method". It gives us the matching without any information. The probability of matching depends on the number of categories or all possible ways of getting matching. That is, if the number of categories is large, the probability of correct random matching must be small and if there is a small number of categories, the probability of correct random matching will be large.

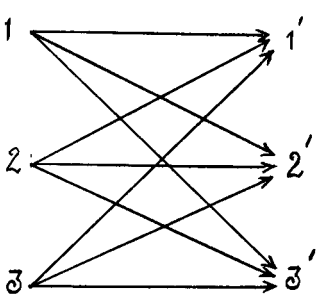
For instance, in the binomial distribution there are two categories.



All possible ways of getting matching = $2! = 2$ ways.

Therefore, the probability of random matching = $\frac{1}{2} = 0.5000$.

In the trinomial distribution there are three categories.



All possible ways of getting matching = $3! = 6$ ways.

The probability of random matching = $\frac{1}{6} = 0.16667$.

In case of k-nomial distribution which consists of

k categories, and all possible ways of matching = k! ways, the probability of random matching = $\frac{1}{k!}$.

How Good is the Method of Maximum Likelihood as Compared to the Method of Random Matching

Besides using the method of maximum likelihood in matching off the sample proportions with the population proportions, the work of this thesis is also to examine how good the method of maximum likelihood matching is by a complete enumeration of cases for some simple problems as given below. The problems considered are:

(I) binomial distribution (2 categories) with population proportions

Case 1. identical; $\frac{5}{10}$, $\frac{5}{10}$

Case 2. moderately different; $\frac{7}{10}$, $\frac{3}{10}$

Case 3. very much different; $\frac{9}{10}$, $\frac{1}{10}$

each case for sample sizes 2, 3, 4, 5 times the number of categories which are 2(2) = 4, 3(2) = 6, 4(2) = 8, 5(2) = 10 respectively.

(II) trinomial distribution (3 categories) with population proportions

Case 1. identical; $\frac{3.33}{10}$, $\frac{3.33}{10}$, $\frac{3.33}{10}$

Case 2. moderately different; $\frac{5}{10}$, $\frac{3}{10}$, $\frac{2}{10}$

Case 3. very different; $\frac{7}{10}$, $\frac{2}{10}$, $\frac{1}{10}$

each case for sample sizes 2, 3, 4, 5 times the number of

categories which are $2(3) = 6$, $3(3) = 9$, $4(3) = 12$, $5(3) = 15$, respectively.

And in order to have a measure of "how good the method of maximum likelihood is" we shall use the ratio

$$\frac{\text{Proportion of time maximum likelihood correct matching;}}{\text{Proportion of time random correct matching}}$$

or the effectiveness of the method of maximum likelihood as compared to random matching. This will help us interpret the results as we vary the three factors number of categories (binomial, trinomial, etc.), sample sizes (2, 3, 4 and 5 times the number of categories), and population proportions (identical, moderately different, very much different).

CHAPTER 5

RESULTS AND CONCLUSIONS

This chapter gives the results and conclusions in summary. Table I gives the values of the ratio

$$\frac{\text{Proportion of the time maximum likelihood correct matching}}{\text{Proportion of the time random correct matching}}$$

for all the problems considered, including variations in the three factors: number of categories, sample size, population proportions.

Table I

Number of Categories	Sample Sizes				Population Proportions
	2·(#Categories)	3·(#Categories)	4·(#Categories)	5·(#Categories)	
Binomial	1.000	1.000	1.000	1.000	identical
Trinomial	1.000	1.000	1.000	1.000	
Binomial	1.568	1.674	1.748	1.802	moderately
Trinomial	2.321	2.674	2.889	3.052	different
Binomial	1.944	1.991	1.995	1.998	very much
Trinomial	3.091	3.982	4.134	4.480	different

Table II gives the values of the probabilities of giving correct matching by the methods of maximum likelihood and random matching, also the ratio of

Table II

Population Proportions	Sample Sizes	BINOMIAL DISTRIBUTION(Two Categories) Pr. (maximum likelihood correct matching)	Pr. (random correct matching)	Pr. (max. likelihood correct matching) / Pr. (random correct matching)	TRINOMIAL DISTRIBUTION(Three Categories) Pr. (maximum likelihood correct matching)	Pr. (random correct matching)	Pr. (max. likelihood correct matching) / Pr. (random correct matching)
Case I (Identical)	2·(#categories)	0.50000	0.50000	1	0.16667	0.16667	1
	3·(#categories)	0.50000	0.50000	1	0.16667	0.16667	1
	4·(#categories)	0.50000	0.50000	1	0.16667	0.16667	1
	5·(#categories)	0.50000	0.50000	1	0.16667	0.16667	1
Case II (Mod. different)	2·(#categories)	0.78410	0.50000	1.56800	0.386938	0.16667	2.321581
	3·(#categories)	0.836920	0.50000	1.67384	0.443996	0.16667	2.663975
	4·(#categories)	0.873964	0.50000	1.74793	0.481426	0.16667	2.888551
	5·(#categories)	0.9011913	0.50000	1.80238	0.508643	0.16667	3.051805
Case III (Very different)	2·(#categories)	0.972000	0.50000	1.94400	0.515211	0.16667	3.091261
	3·(#categories)	0.991440	0.50000	1.99144	0.663616	0.16667	3.981692
	4·(#categories)	0.997272	0.50000	1.99454	0.689037	0.16667	4.134221
	5·(#categories)	0.999109	0.50000	1.99822	0.746678	0.16667	4.480065

$$\frac{\text{Proportion of the time maximum likelihood correct matching}}{\text{Proportion of the time random correct matching}}$$

or the effectiveness of maximum likelihood as compared to random matching, for all the problems considered, including variations in the three factors: number of proportions, sample sizes, and number of categories.

Discussion

In order to make it easier for the reader to understand the conclusions in the next section, we are going to give a detailed discussion on the probabilities and effectiveness of maximum likelihood matching as compared to the random matching from the table of results (Table II, p. 25).

Case 1. The population proportions are identical both in binomial and trinomial distributions. We have discovered that the probabilities of giving correct matching by the method of maximum likelihood in binomial and trinomial distribution are constant regardless of the sample sizes used 0. 5000 and 0. 16667 respectively. In this case random matching gives the same probability as that of maximum likelihood. The effectiveness of maximum likelihood matching as compared to random matching equal to 1 in both distributions.

Case 2. The population proportions are moderately different. We have found that the probabilities of giving correct matching by the method of maximum likelihood will increase slightly in both distributions when compared with the first case at the same sample sizes. Also these probabilities will increase as sample size increases and will decrease as the number of categories is increased. Although the probability of correct matching by maximum likelihood decreases as the number of categories is increased, the probability of correct matching by the random method decreases also, and in a greater proportion; so that the effectiveness ratio actually increases when the number of categories is increased. Therefore, in this case the effectiveness of maximum likelihood matching as compared to random matching will significantly increase when the sample size and number of categories increase.

Case 3. In this case the population proportions are very different. We have found that the probabilities of the maximum likelihood matching are higher than those probabilities of the first and second case at the same sample sizes. These probabilities will increase when the sample size increases and such probabilities will decrease as the number of categories increases. This case the effectiveness of maximum likelihood matching as compared to random matching is also higher than those of the first two cases at the

same sample size. And it will significantly increase as the sample size and number of categories increase.

Since we know that the greatest probability of correct matching by maximum likelihood is 1, then we can assert the maximum effectiveness of maximum likelihood matching as compared to random matching as given below.

(1) Binomial Distribution,

The maximum effectiveness of maximum likelihood correct matching

$$= \frac{\text{The greatest probability of maximum likelihood correct matching}}{\text{The probability of random correct matching}} = \frac{1}{1/2} = 2$$

(2) Trinomial Distribution

The maximum effectiveness of maximum likelihood correct matching

$$= \frac{\text{The greatest probability of maximum likelihood correct matching}}{\text{The probability of random correct matching}} = \frac{1}{1/6} = 6$$

From the table of results (Table II), we have found that the highest effectiveness of maximum likelihood as compared to random matching are 1.998 and 4.480 in binomial and trinomial distributions respectively at the sample size 5 times the number of categories. This effectiveness is still less than the maximum effectiveness as shown above. The reasons are (1) the sample size is not big enough (2) the population proportions are not much different from each other.

Furthermore, if we reconsider the effectiveness of maximum likelihood correct matching in each distribution, we will realize that the effectiveness of maximum likelihood will indicate the "roughness" in the population proportions such as in the trinomial distribution at sample size five times the number of categories. The effectiveness of maximum likelihood matching as compared to random matching is 1 when the population proportions are $P_1:P_2:P_3 = \frac{3.33}{10} : \frac{3.33}{10} : \frac{3.33}{10}$; and it is 3.0518 when population proportions are $P_1:P_2:P_3 = 5:3:2$; and finally be 4.4801 at the population proportions are $P_1:P_2:P_3 = 7:2:1$.

Conclusions

From the above discussion we conclude the following.

(1). Other things (sample size and number of categories) being equal, as the population proportions become "more different" ("roughness" increases) the effectiveness of maximum likelihood matching as compared to random matching increases.

(2). Other things (sample size and population proportions) being equal, as the number of categories increases the effectiveness of maximum likelihood matching as compared to random matching increases.

(3) Other things (number of categories and population proportions) being equal, as the sample size increases the effectiveness of maximum likelihood matching as compared to random matching increases.

(4) When the population proportions are all equal, the effectiveness of maximum likelihood matching is the same as that of random matching.

CHAPTER 6

PROPOSAL FOR FURTHER STUDY

In this chapter we suggest methods of information theory to replace the method of complete enumeration of cases in studying multinomial population where the number of categories is large. Reference could be made to Shannon's paper "Communication Theory of Secret Systems"[8, 9] for earlier work done using information theory. Further enumeration work might be carried out by using a computer. One might classify languages in accordance with the "roughness" of their population proportions, characterizing a language's effectiveness for maximum likelihood matching, and hence ease of solution of simple substitution cipher. One might determine the sample size required to obtain a certain degree of matching of sample proportions with population proportions. One might develop a sequential method of matching sample proportions with population proportions.

BIBLIOGRAPHY

1. Abramson, N. Information theory and coding. New York, McGraw-Hill. 1963. 201 p.

2. B. A. Mathematical Tables. Vol IX. Table of powers. Cambridge University Press, 1958. 398 p.

3. Fraser, D. A. Statistics an introduction. New York, Wiley, 1958. 398 p.

4. Kullback, S. Information theory and statistics. New York, Wiley, 1959. 395 p.

5. Mood, A. M. and F. A. Graybill. Introduction to the theory of statistics. New York, McGraw-Hill, 1963. 443 p.

6. Pierce, J.R. Symbols, signals and noise. New York, Brothers, 1961. 305 p.

7. Pratt, F. Secret and urgent. The story of codes and ciphers. New York, Bobbs-Merrill, 1939. 282 p.

8. Shannon, C. E. A mathematical theory of communication. Bell System Technical Journal 27: 379-423. 1948.

9. Shannon, C. E. Communication theory of secret systems. Bell System Technical Journal 28: 656-715. 1949.

10. Smith, L. D. Cryptography. The science of secret writing. New York, W W Norton, 1943. 164 p.

APPENDIX

APPENDIX

AN EMPIRICAL STUDY

The detailed computations given here are just for samples of size two times the number of categories when both number of categories and population proportions vary, for the results (effectiveness of maximum likelihood as compared to random matching: 1.000, 1.000, 1.568, 2.321, 1.944 and 3.097) shown in the left hand column of Table I, p. 24. The complete computation of cases has been worked out in the same pattern as the given example and is not shown here because the work involves a lengthy mathematical operation. However the exact data of results can be found in Table II, p.25.

Consider the binomial distribution (2 categories); at sample size = two times the number of categories = 2 x 2 = 4.

The possible values in the two sample categories are:

X_1	X_2
4	0
3	1
2	2

The probabilities and the effectiveness of maximum likelihood matching at each possibility are:

Case 1. identical; $P_1 : P_2 = \frac{5}{10} : \frac{5}{10}$

$$\text{Pr.}(4, 0) = \frac{4!}{4!\,0!}\left(\frac{5}{10}\right)^4\left(\frac{5}{10}\right)^0 = 0.0625$$

$$\text{Pr.}(3, 1) = \frac{4!}{3!\,1!}\left(\frac{5}{10}\right)^3\left(\frac{5}{10}\right)^1 = 0.2500$$

$$\text{Pr.}(2, 2) = \frac{1}{2}\cdot\frac{4!}{2!\,2!}\left(\frac{5}{10}\right)^2\left(\frac{5}{10}\right)^2 = 0.1875$$

The probability of maximum likelihood correct matching = ΣPr.

= 0.5000 .

The effectiveness of maximum likelihood matching is

$$\frac{\text{The probability of maximum likelihood correct matching}}{\text{The probability of random correct matching}} = \frac{0.5000}{0.5000} = 1$$

Case 2. moderately different; $P_1 : P_2 = \frac{7}{10} : \frac{3}{10}$

$$\text{Pr.}(4, 0) = \frac{4!}{4!\,0!}\left(\frac{7}{10}\right)^4\left(\frac{3}{10}\right)^0 = 0.2401$$

$$\text{Pr.}(3, 1) = \frac{4!}{3!\,1!}\left(\frac{7}{10}\right)^3\left(\frac{3}{10}\right)^1 = 0.4116$$

$$\text{Pr.}(2, 2) = \frac{1}{2}\cdot\frac{4!}{2!\,2!}\left(\frac{7}{10}\right)^2\left(\frac{3}{10}\right)^2 = 0.1323$$

The probability of maximum likelihood correct matching = ΣPr.

= 0.7840

The effectiveness of maximum likelihood matching is

$$\frac{\text{The probability of maximum likelihood correct matching}}{\text{The probability of random correct matching}} = \frac{0.7840}{0.5000} = 1.568$$

Case 3. very different; $P_1 : P_2 = \frac{9}{10} : \frac{1}{10}$

$$\text{Pr.}(4, 0) = \frac{4!}{4!0!}\left(\frac{9}{10}\right)^4\left(\frac{1}{10}\right)^0 = 0.6561$$

$$\text{Pr.}(3, 1) = \frac{4!}{3!1!}\left(\frac{9}{10}\right)^3\left(\frac{1}{10}\right)^1 = 0.2916$$

$$\text{Pr.}(2, 2) = \frac{1}{2}\cdot\frac{4!}{2!2!}\left(\frac{9}{10}\right)^2\left(\frac{1}{10}\right)^2 = 0.0243$$

The probability of maximum likelihood correct matching is = ΣPr.

= 0.9720

The effectiveness of maximum likelihood matching is

$$\frac{\text{The probability of maximum likelihood correct matching}}{\text{The probability of random correct matching}} = \frac{0.9720}{0.5000} = 1.9440.$$

Consider the trinomial distribution (3 categories), at the sample size two times of the number of categories = 2(3) = 6. The possible values in the three sample categories are:

X_1	X_2	X_3
6	0	0
5	1	0
4	2	0
4	1	1
3	3	0
3	2	1
2	2	2

The probabilities and the effectiveness of maximum likelihood matching at each possibility are:

Case 1. identical; $P_1 : P_2 : P_3 = \frac{3.33}{10} : \frac{3.33}{10} : \frac{3.33}{10}$

$$\text{Pr.}(6,0,0) = \frac{1}{2}\cdot\frac{6!}{6!0!0!}\left(\frac{3.33}{10}\right)^6\left(\frac{3.33}{10}\right)^0\left(\frac{3.33}{10}\right)^0 = 0.000686$$

$$\text{Pr.}(5,1,0) = \frac{6!}{5!1!0!}\left(\frac{3.33}{10}\right)^5\left(\frac{3.33}{10}\right)^1\left(\frac{3.33}{10}\right)^0 = 0.008230$$

$$\text{Pr.}(4,2,0) = \frac{6!}{4!2!0!}\left(\frac{3.33}{10}\right)^4\left(\frac{3.33}{10}\right)^2\left(\frac{3.33}{10}\right)^0 = 0.020576$$

$$\text{Pr.}(4,1,1) = \frac{1}{2}\cdot\frac{6!}{4!1!1!}\left(\frac{3.33}{10}\right)^4\left(\frac{3.33}{10}\right)^1\left(\frac{3.33}{10}\right)^1 = 0.020576$$

$$\text{Pr.}(3,3,0) = \frac{1}{2}\cdot\frac{6!}{3!3!0!}\left(\frac{3.33}{10}\right)^3\left(\frac{3.33}{10}\right)^3\left(\frac{3.33}{10}\right)^0 = 0.013717$$

$$\text{Pr.}(3,2,1) = \frac{6!}{3!2!1!}\left(\frac{3.33}{10}\right)^3\left(\frac{3.33}{10}\right)^2\left(\frac{3.33}{10}\right)^1 = 0.082302$$

$$\text{Pr.}(2,2,2) = \frac{1}{6}\cdot\frac{6!}{2!2!2!}\left(\frac{3.33}{10}\right)^2\left(\frac{3.33}{10}\right)^2\left(\frac{3.33}{10}\right)^2 = 0.020576$$

The probability of maximum likelihood correct matching $= \Sigma \text{Pr.}$

$= 0.1666$

The effectiveness of maximum likelihood matching is

$$\frac{\text{The probability of maximum likelihood correct matching}}{\text{The probability of random correct matching}} = \frac{0.1666}{0.1666} = 1.$$

Case 2. moderately different; $P_1 : P_2 : P_3 = \frac{5}{10} : \frac{3}{10} : \frac{2}{10}$

$$\text{Pr.}(6,0,0) = \frac{1}{2}\cdot\frac{6!}{6!0!0!}\left(\frac{5}{10}\right)^6\left(\frac{3}{10}\right)^0\left(\frac{2}{10}\right)^0 = 0.007812$$

$$\text{Pr}(5,1,0) = \frac{6!}{5!1!0!}\left(\frac{5}{10}\right)^5\left(\frac{3}{10}\right)^1\left(\frac{2}{10}\right)^0 = 0.056250$$

$$\text{Pr.}(4,2,0) = \frac{6!}{4!2!0!}\left(\frac{5}{10}\right)^4\left(\frac{3}{10}\right)^2\left(\frac{2}{10}\right)^0 = 0.084375$$

$$\text{Pr.}(4,1,1) = \frac{1}{2}\cdot\frac{6!}{4!1!1!}\left(\frac{5}{10}\right)^4\left(\frac{3}{10}\right)^1\left(\frac{2}{10}\right)^1 = 0.05625$$

$$\text{Pr.}(3,3,0) = \frac{1}{2}\cdot\frac{6!}{3!3!0!}\left(\frac{5}{10}\right)^3\left(\frac{3}{10}\right)^3\left(\frac{2}{10}\right)^0 = 0.03375$$

$$\Pr.(3,2,1) = \frac{6!}{3!2!1!}\left(\frac{5}{10}\right)^3\left(\frac{3}{10}\right)^2\left(\frac{2}{10}\right)^1 = 0.135000$$

$$\Pr.(2,2,2) = \frac{1}{6}\cdot\frac{6!}{2!2!2!}\left(\frac{5}{10}\right)^2\left(\frac{3}{10}\right)^2\left(\frac{2}{10}\right)^2 = 0.013500$$

The probability of maximum likelihood correct matching = ΣPr.

= 0.386938

The effectiveness of maximum likelihood matching is

$$\frac{\text{The probability of maximum likelihood correct matching}}{\text{The probability of random correct matching}} = \frac{0.386938}{0.166666} = 2.321581$$

Case 3. very different; $P_1 : P_2 : P_3 = \frac{7}{10} : \frac{2}{10} : \frac{1}{10}$

$$\Pr.(6,0,0) = \frac{1}{2}\cdot\frac{6!}{6!0!0!}\left(\frac{7}{10}\right)^6\left(\frac{2}{10}\right)^0\left(\frac{1}{10}\right)^0 = 0.058825$$

$$\Pr.(5,1,0) = \frac{6!}{5!1!0!}\left(\frac{7}{10}\right)^5\left(\frac{2}{10}\right)^1\left(\frac{1}{10}\right)^0 = 0.201684$$

$$\Pr.(4,2,0) = \frac{6!}{4!2!0!}\left(\frac{7}{10}\right)^4\left(\frac{2}{10}\right)^2\left(\frac{1}{10}\right)^0 = 0.14406$$

$$\Pr.(4,1,1) = \frac{1}{2}\cdot\frac{6!}{4!1!1!}\left(\frac{7}{10}\right)^4\left(\frac{2}{10}\right)^1\left(\frac{1}{10}\right)^1 = 0.07203$$

$$\Pr.(3,3,0) = \frac{1}{2}\cdot\frac{6!}{3!3!0!}\left(\frac{7}{10}\right)^3\left(\frac{2}{10}\right)^3\left(\frac{1}{10}\right)^0 = 0.02744$$

$$\Pr.(3,2,1) = \frac{6!}{3!2!1!}\left(\frac{7}{10}\right)^3\left(\frac{2}{10}\right)^2\left(\frac{1}{10}\right)^1 = 0.008232$$

$$\Pr.(2,2,2) = \frac{1}{6}\cdot\frac{6!}{2!2!2!}\left(\frac{7}{10}\right)^2\left(\frac{2}{10}\right)^2\left(\frac{1}{10}\right)^2 = 0.00294$$

The probability of maximum likelihood correct matching = ΣPr.

= 0.515211

The effectiveness of maximum likelihood matching is

$$\frac{\text{The probability of maximum likelihood correct matching}}{\text{The probability of random correct matching}} = \frac{0.515211}{0.166666} = 3.091241$$