UNITED ARAB REPUBLIC

THE INSTITUTE OF NATIONAL PLANNING

Memo. No. 1164

ON

Optimum Sample Size For Estimating Dura Crop Yield Rate In Sudan

By

Dr. A. H. Abdel-Razek

and

Mr. Abdel-Azim Hassan

CAIRO
SALAH SALEM St.-NASR CITY

On
Optimum Sample For Estimating
Dura Crop Yield Rate
In Sudan

By
Dr. A.H. Abdel-Razek
Institute of National Planning Cairo, A.R.E.

and

Mr. Abdel-Azim Hassan
Dept. of statistics Khartoum, Sudan

Introduction :

The Sudan is predominately an agricultrual country. Agriculture is the main sector of its economy and its principal source of foriegn exchange the Government is building up most of its social and economical development programme and basing its ecternal trade agreements upon agriculture.

The sampling technique has proved since long, to be one of the main pillar for most of the modern sciences. Due of its major application fields is the agricultural one. Estimation of the crop yield by the so-colled crop cutting experiments is only one of its many uses.

The problem of sample size required to estimate a certain statistic is one of major problems which faces statisticians on planning a certain survey. These are many factors which should be considered on this aspect. An important one is the cost on the land and precision on the other.

It will be the main objective of this paper to compute an optimum sample size to estimate the dura (Sorghum) crop yield rate in Sudan both by the variance function and cost function, on the basis

of a crop-cutting survey run in Sudan in 22 councils important for the crop.

The dura is the stable food of the country and is cultivated almost all over the country on different grounds. Thus in some parts it is cultivated manually and in small pieces of land while most of the crop is cultivated mechanically and in very large pieces of land. Though this looks as if it complicates the problem of designing a sampling survey, but on the basis of long experience it was found that the same design almost fits most of the parts without loss of precision. Nevertheless additional studies have to be made on some of the open areas and the nomads area of the crop.

It is thus natural to find out that this crop statistics has found greater attention. This survey , under consideration, was the result of long sustained efforts iniated in 1959 by a small survey on dura in the mechanized crop production schemes of gedarif district, Kassala province. By 1967/68 the department of statistics has completed 18 such surveys.

The Survey :

This survey covered 22 councils important for dura crop distributed over the 6 northern provinces. The design adopted was multistage stratified sampling. Each council constituted a stratum and wherever possible it was divided into few substrata on the pattern of the crop roughly. Within a stratum, a two-stage sampling was used for the area with the shiekhship as the primary sampling unit and agricultural holding within the shiekhship selected as the secondary unit. For crop-cutting the sampling was extended to two more stages. Fields growing dura within each selected holding constitute the third stage sampling unit and a 6 x 7 m plot the fourth stage.

The selection of the units at each stage was done at random with

equal probability and without repleement. The selection of the shiekhship was done at the headquarters of the Department of stratistics. The work of listing of holders and conduct of crop cutting was entrusted to enumerators under close supervision of the supervisors and the inspectors. The enumerators and supercisory staff was adequately trained in the conduct of field work before they proceed to the field work.

After reaching the selected shiekhship holders (persons operating holdings) were selected at random. For the purpose of selecting a field, first a continuious serial numbering over all the fields which grew dura. The last serail was noticed and hence 3 fields were selected by tables of random numbers.

The random location of the plot was done on the date of harvesting already fixed in consultation with the holders by random numbers table as usual. It should be made sure that all of the plot lies within the selected field.

The cobs are harvested and spread over a piece of cloth for a while before threshing the crop is then threshed according to the local practice winnowed and cleaned. The clean produce is weighed and then left for a period to dry and then reweighidg.

Equations used :

Computations depended on the following equations :

$$\frac{\frac{S_b^2}{n} + \frac{S_w^2}{mn}}{(\bar{Y}_{mn})^2} = \frac{t^2}{(100)^2} \quad \ldots\ldots\ldots\ldots \quad (1)$$

$$m_o = \sqrt{\frac{c_1}{c_2} \times \frac{s^2_w}{s^2_b}} \qquad (2)$$

Where $\frac{s^2_b}{n} + \frac{s^2_w}{mn}$ gives the variance of the mean $(\bar{Y}_{mn})$

on taking the finite multipliers as unity since the sampling fraction is small in each stratum (less than 0.05).

s^2_b is the variance between Shiekhships

s^2_w " " " " fields

n is the number of Shiekhships

m " " " " plots

t is the percent standerd error.

c_1 is the cost per shiekhship

c_2 " " " " plot

Equation (1) gives the distributions of m and n for different "t" while aquation (2) gives the optimum number of plots at minimum cost.

Computations

In multistage design variance is best estimated by (ANOVA) technique by pooling the ANOVA over individual strata.

Table (1) Form of the (ANOVA) table

Source of variation	d. of f.	M S	E M S
Between Strata	L-1		
Between Shiekhships within strata	$\sum_{h=1}^{L}$	(n_h-1) s_b^2	$\bar{S}_w^2 + m S_b^2$
Between fields within Shiekhships within Strata	$\sum_{h=1}^{L}$	$n_h(m-1)$ s_w^2	$\bar{S}_w^2$

Thus the estimates of S_w^2 and S_w^2 which can be obtained from the pooled (ANOVA) are:

$$E\ (\bar{S}_w^2) = S_w^2 \qquad \text{.....} \qquad (3)$$

$$E\ (S_h^2) = \frac{s_b^2 - s_w^2}{m} \qquad \text{.....} \qquad (4)$$

Results :

The pooled (ANOVA) is here shown for kassala province with distribution of number of plots and shiekhships for different percent standerd error. The other provinces will be summerized in a seperate table.

In Kassala province 3 councils were chosen for the survey, each was divided into two strata, 72 Shiekships were selected and all the 216 planned experiments were conducted. The average yield Y_{mn} was found to be 2.8811 kgms/plt

Table (2) Pooled (ANOVA) for Kassala Province

Source of variation	d. of. f	Sum Square	M.S.S.
Between Strata	5	283.0902	56.6180
Between Sh. within strata	66	764.4655	12.0373
Between fields within Sh. within strata	144	670.3065	4.7539
Total	215	1747.8622	

Thus $E(\bar{S}_w^2)$ = 4.7539

$$E(S_b^2) = \frac{12.0373 - 4.7539}{3} = 2.4278$$

With the help of table (2) and its results we can construct (3) below :

Table (3) Distribution of shiekhships for different levels of statistic and different values of m for Kassala Province

t \ m		1	2	3	4	5
2	n	2103	1417	1188	1074	1006
3		935	630	528	477	447
4		526	354	297	269	251
5		336	227	190	172	161
6		234	157	132	119	112

It is noticed that the number of shiekhships decreases as the

number of plots m increases. It is also observed that the rate of decrease is noticable as m increases from 1-3 , but it diminishes on further increase. Indicating that 3 plots is about an optimum size. Thus if we suffice ourselves by 5% standerd error it is found that it is necessary to select 190 shiekhships to estimate the dura yield rate in Kassala province.

On application equation (2) one is and faced by the non availability of the elements of the equation i.e, c_1 and c_2. On the basic previous similar situations, $\frac{c_1}{c_2}$ is estimated as 6. This leads to the result that at minimum cost 4 plots are necessary to be surveyed showing showing that 172 shiekhships are needed.

The similar results for the other previnces are shown below :

Table (4)

Province	Without Cost		With cost in account	
	No of Sh.	No of Plots	No sh.	No of plots
Blue Nile	457	3	457	3
Kordofan	149	3	87	7
Darfur	292	3	226	5
Northern and Khartoum	273	3	245	4
	1361		1187	

Thus it is necessary to surcey 1361 , shiekhships and 4083 plots and on taking cost in account 1187 shiekhships and 4778 plots.

Summary and Conclusion :

A general review of the results thus obtained reveals the following :

1- the number of shiekhships required for different degrees of precision showed the same pattern . It decreases on the increase of the number of plots m. The rate of decrease is noticable as m increases from 1-3 . On further increase in m the rate of decrease diminshes and not appreciable, indicating that there is no need to sample more than 3 plots per shiekhship .

2- On fixing the level of standard error at 5% which is generally accepted in such type of sureys and conducting 3 experiments per shiekhship the number of the latter varies from 149 in kardofan to 457 in Blue Nile provinces.

3- A noticable fact is that the sample size computed in this research is larger than that which was conducted in the actual survey . This may be explained by the large percentage error which resulted from the survey being as high as 16% in Northern and khartoum provinces combined. This high error on the other hand might be due to :

a) The very small size of the sample actually run
b) The crop in a great number of the plots has totally failed giving zero yield.

4- On considering thz cost to estimate the optimum sample size at minimum cost one is faced with the non availability of the rough data on the cost elements. They are estimated by consideration of similar previous situation the optimum number of plots per shiekhship varied from 3 in Blue Nile province to 7

in Kordofan province. The corresponding number of shiek-ships was 457 in Blue Nile and 87 in kordofan province.

5- The relatively large number of plots required in kordofan is due to the largeness of the ratio S_w^2 / S_b^2 which was 9.2502.

A factor which should be considered carefully on planning future surveys is the non sampling error which may mainly be due to the lack of adequate staff and their centralization in the headquarters only.

The most suitable size of the experimental plot needs further study. It is hoped that uniformity trails should be made on different size and shapes to choose the best of them.

Together with this it is necessary to run a survey with the main objective of computing the cost function elements which were only estimated in this paper.

The computations given here depended mainly on the 6 Northern provinces. The irrigated area of the Gezira was excluded since all the yield is recorded. These are some parts of the open areas which need certain tackling and investigations have to be made on them.

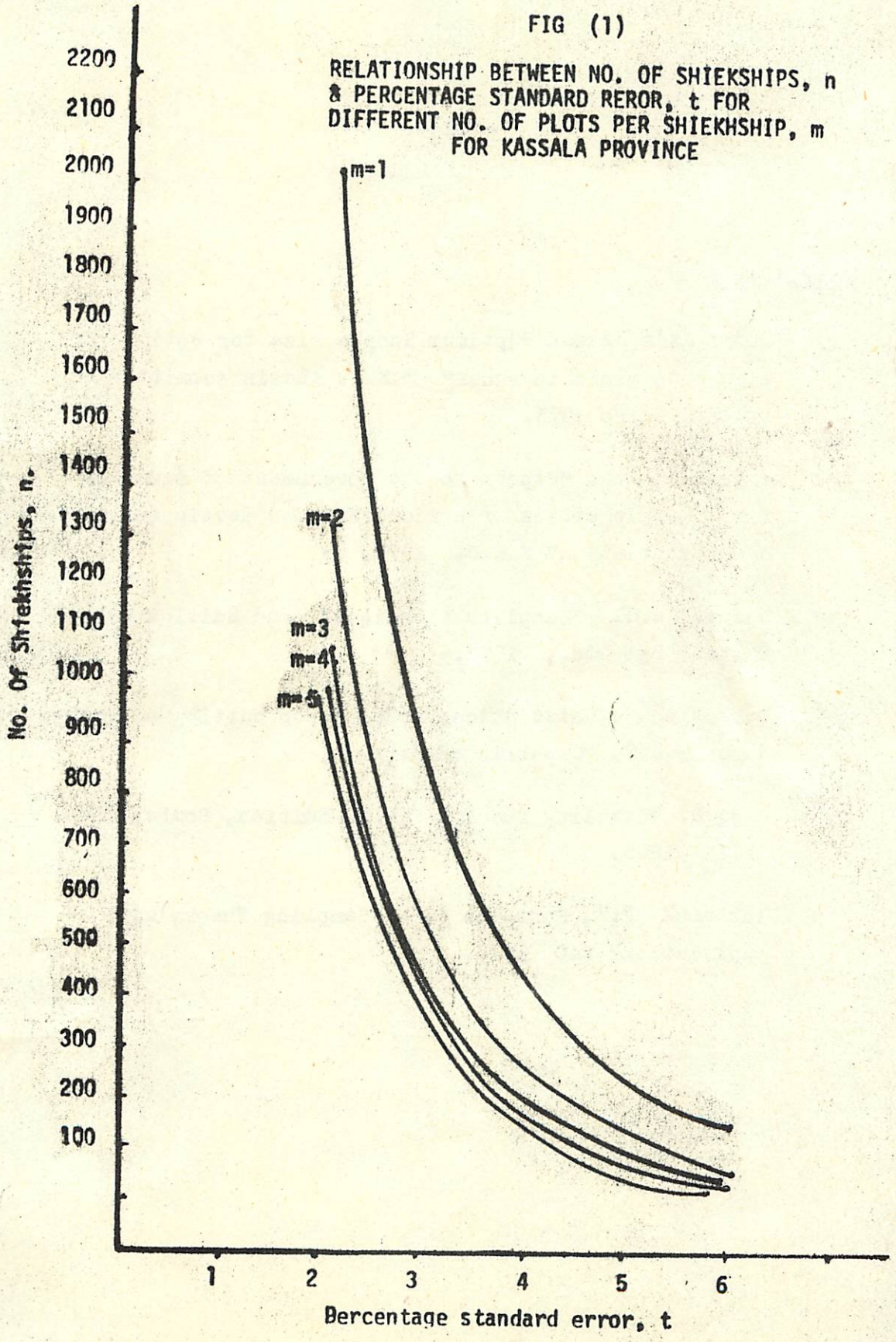
If is also worth mentioning that the sub-sampling unit was the same as that guessed for the survey.

Fig I shows the variation of n with per cent standard error "t" for given number of plots m to be harvested for each shiekhship.

Finally we hope that this work might be of some benifits to whom so even interested.

FIG (1)

RELATIONSHIP BETWEEN NO. OF SHIEKSHIPS, n & PERCENTAGE STANDARD REROR, t FOR DIFFERENT NO. OF PLOTS PER SHIEKHSHIP, m FOR KASSALA PROVINCE



Refrences :

(1) Abdel Azim Hassan "Optimum Sample size for estimating dura crop yield in sudan" M.S.C. thesis submitted to I.S.S.R Cairo 1975.

(2)

(2) A yachit, G.R. "Report to the government of Sudan on first sample consus of agriculture and developments of crops statistics" F.A.O. 1970.

(3) Cochran W.G. "Sampling technique" 2nd Edition Wiley Eastern Pyt Ltd., 1972.

(4) Department of statidtics, sudan "crop cutting survey on dura 1968/69 "Unpublished data.

(5) Raj, D. "Sampling Theory" T.M.H. Edition, Bombay New - Delhi 1968.

(6) Sukhatme P.V, Sukhatme B.V. "Sampling Theory with Application" FAO 1970.