Journal of Statistical
Theory & Methods
1998, Vol. 9, No. 1, pp.57 ~ 62

Determination of Sample Sizes for Plant Characteristics of Food Crops [1]

Suk Hwan Chang [2]

Abstract

The minimum number of samples for the measurement of plant characteristics of major crops were calculated from the data obtained from the field experiments on rice, barley, wheat, soybean and sweet potato conducted by Kyungpuk Rural Development Administration.

Key Words and Phrases: Sample size, Food crops, Field experiment

1. Introduction

Sample sizes for the plant characteristics of food crops in the field experiments are frequently required for the measurement of the characteristics which are closely related with yield in order to estimate the yield at the earliest growth stage of the crops. If, for an example, the yield components of rice such as the number of panicle per unit area, the number of filled grains per panicle and the 1000 grains weight could be estimated as early as possible with a great precision, the final yield of rice can be estimated immediately with higher precision, which is important information for food strategy and planning of economic development.

Abraham(1966) studied the sample sizes of crops required for the given precision using the method of sampling in the field experiment(Kempthorne, 1962), and Oh et al. (1969) also estimated the number of hills from the data collected from the farmer's field to estimate the yield components of rice, using the same technique. Chang and Ha (1981) studied the sample sizes for the estimation of yield components of wheat, and Chang (1997) recently reported the results of study on the sample size for plant characteristics of rice crop. The number of samples being taken for various

[1]This study was performated by the Bisa Research Grant of Keimyung University in 1997.

[2]Professor, Department of Statistics, Keimyung University, #1000, Shindang Dong, Dalso Gu, Taegu, 704-701, Korea.

studies are, in fact, different among the research workers even in an institution, which result in inefficient utilization of resources.

The present study, thus, aims at the determination of the minimim sample sizes for plant characteristics of major crops to meet the given precision at different growth stages.

2. Data used and statistical methods 2.1 Data used

The data used for the present study are from the experiments for ascertaining the productivity of high-yielding rice varieties conducted in Taegu and Andong areas in 1996, and from the experiments on the regional adaptability of the superior lines of barley and wheat conducted in successive three years from 1993 to 1995 and of soybean conducted in 1994 through 1996. The experiments were carried out in the randomized block design with three or four replications by Kyongpuk Rural Development Administration(PRDA).

2.2 Statistical methods

For the randomized block design with t treatments in r blocks, the statistical model, assuming that n samples were taken from each treatment, would be

$$y_{ijk} = \mu + \gamma_i + \tau_j + \eta_i j + \varepsilon_i jk, \qquad \begin{matrix} i = 1, 2, \cdots, r \\ j = 1, 2, \cdots, t \\ k = 1, 2, \cdots, n \end{matrix} \tag{1}$$

where y_{ijk} is the kth observed value of the characteristic under study of the jth treatment in the ith block, μ the overall mean, γ_i the ith block effect, τ_j the jth treatment effect, η_{ij} the experimental error $\sim N(0, \sigma^2)$ and ε_{ijk} the sampling error $\sim N(0, \sigma_s^2)$.

The analysis of variance for the equation (1) would then be as the Table 1. Since the variation of the plant characteristics depend largely on the geographical locations where agricultural climates are different each other and on the weather conditions that may vary from year to year. Therefore, similar experiments, in which some of the treatments may be replaced by some newly designed treatments or newly developed lines of crops, are carried out in several geographical locations and/or for several years in order to seek for desirable treatment(s) which are also tolerable against the year variations. In these cases, the experimental results may be combined over the experimental years or locations. Suppose the experiments of the model (1) were carried out for m years. Then the combined analysis of variance may be as the Table 2.

Table 1. Analysis of variance table for the randomized block design with n samples from each plot.

S.V.	DF	SS	MS	E(MS)
Blocks	r-1	SSBL	MSBL	
Treatments	t-1	SSTR	MSTR	
Error	(r-1)(t-1)	SSE	MSE	$\sigma_s^2 + n\sigma^2$
Sampling Error	rt(n-1)	SSSE	MSSE	σ_s^2
Total	rtn-1	SSTO	-	

Table 2. Analysis of variance table for the combined data of the model (1) for m years.

S.V.	DF	SS	MS	E(MS)
Year (Y)	m-1	SSYR	MSYR	$\sigma_\epsilon^2 + nt\sigma_{\alpha\beta}^2 + nrt\sigma_\alpha^2$
Reps within Year	m(r-1)	SSR/Y	MSR/Y	
Treatments (T)	t-1	SSTR	MSTR	$\sigma_\epsilon^2 + n\sigma_e^2 + nr\sigma_{\alpha\gamma}^2 + mnr\sigma_\gamma^2$
Year × Treatment	(m-1)(t-1)	SS(Y × T)	MS(Y × T)	$\sigma_\epsilon^2 + n\sigma_e^2 + nr\sigma_{\alpha\gamma}^2$
Error	m(r-1)(t-1)	SSE	MSE	$\sigma_\epsilon^2 + n\sigma_e^2$
Sampling Error	mrt(n-1)	SSSE	MSSE	σ_ϵ^2

In Table 2 it can be shown that $\sigma_\varepsilon^2 = \sum_{i=1}^m \sigma_s^2$, and $\sigma_e^2 = \sum_{i=1}^m (\sigma_s^2 + n\sigma_i^2)$, where σ_s^2 , and σ_i^2 are respectively, σ_s^2 and σ^2 in the Table 1 for the ith year.

The standard error of a treatment mean is given by

$$s(\bar{y}_{.j..}) = \sqrt{\frac{MSE}{mrn}} \tag{2}$$

$$= \sqrt{\frac{\sigma_\varepsilon^2}{mrn} + \frac{\sigma_e^2}{mr}} \tag{3}$$

and the percentage standard error(P) would be

$$P = \frac{s(\bar{y}_{.j..})}{\bar{y}_{....}} \times 100 \tag{4}$$

From (4) the sample size may easily be calculated for a particular year (m=1) as

$$n = \frac{\hat{\sigma}_\varepsilon}{\delta - \hat{\sigma}_e^2} \tag{5}$$

where $\delta = \frac{rP^2\bar{y}_{...}^2}{100^2}$ and $\hat{\sigma}_e^2$ is the pure experimental error estimated as usual in the Table 2.

3. Results and Discussion

Based on the results of the experiments mentioned in section 2.1, the sample sizes have been computed by the formula (5), and are given in the Table 3. The minimum number of samples to be taken for given precision and replications were computed for the plant characteristics at various growth stages of the crops under study as seen in the Table 3. It is considered that although only one sample is to be taken for plant height and culm length of rice when p=4% with three replications, 23 samples and 79 samples, which are even much more than those practiced at present, are required at 30DAT and 50DAT, respectively, for the count of tiller number, and 20 samples are to be taken for panicle numbers at the harvesting stage. For the number of glumous flowers 16 samples should be taken to meet p=6% when r=3, and 10 samples for p=6% when r=4. In barley the necessary samples to be taken is one for culm length and 2 samples for the number of ears, 7 samples for ear length for which the sampling unit is an area of $0.18m^2$, and 16 samples for the grain numbers per ear when r=3. These samples are reduced considerably when the number of replications increases to 4. For the culm length of measurement of ear length with 3 replications. For the number of grain per ear, however, at least 6 samples are necessary when r=3, and four samples when r=4. The number of samples for the ear number/$0.18m^2$ were not able to compute since $\hat{\sigma}_e^2$ was not available.
In soybean two samples are required for the measurement of stem length to meet p=4number of pods per plant 50 samples are needed for p=6with r=4, and 10 and 8 samples are, respectively, required to estimate the number of seeds per pod with p=6The number of samples to the taken depends, as seen in (5), on δ, $\hat{\sigma}_e^2$ and $\hat{\sigma}_\varepsilon^2$, where δ depends largely on the value of $\bar{y}_{\cdots\cdots}$, the mean value of the plant characteristic under consideration. It is often experienced that even though the sampling variance $\hat{\sigma}_\varepsilon^2$ is considerably high, if the value of $(\delta - \hat{\sigma}_e^2)$ is relatively large as compared to $\hat{\sigma}_\varepsilon^2$ the number of samples are small, whereas if δ is smaller or just greater than $\hat{\sigma}_e^2$, the sample size n can not be computed or appears to be tremendously large. Thus, reasonable sample size for the plant characteristics of sweet potato could not be computed because of this phenonenon.

4. Conclusions

In conclusion two samples should be taken for the measurement of plant height as well as for the sampling error and 10 samples for other characteristics of rice to meet p=6% when three replications. For barley and wheat two samples as in rice are enough for the culm lemgth, and 5 - 6 samples for other characteristics to meet p=6% when r=3 and four samples when r=4. In soybean two samples are recommended for the measurement of stem length, 50 and 20 samples for the number of pods/plant are required to meet p=6%, respectively, when r=3 and r=4. Ten samples should

be taken for count of seeds/pod. It is seen in the Table 4 that the total number of sample for each plant characteristics when r=4 are, in general, smaller than those when r=3. Thus, higher efficiency is expected from the small amount of samples from four replications than larger samples taken from three replications. However, it depends on the cost of additional replication and the cost of the sampling.

Acknowledgment

The author would like to express his wormest thanks to Mr. Kim, Chil Yong, Head and Senior Researcher, and Mr. Park, Sang Gu, Research worker of Crop Divison, Kyungpuk Rural Development Administration for their providing data. Sincere thanks are also extended to Mr. Ryu, Keun Jae, a Ph. D. condidate, Department of Statistics, Keimyung University, for his self-sacrificing help in computing data and typing this manuscripts.

Table 3. Number of samples to be taken for plant characteristics of the major food crops for given precision.

Crops	growth stage	plant characteristics	$r=2$			$r=3$			$r=4$		
			$p=2$	$p=s$	$p=6$	$p=2$	$p=s$	$p=6$	$p=2$	$p=s$	$p=6$
Rice	30DAT	Plant Height	-	2	1	13	1	1	6	1	1
		No. of Tiller	-	106	11	-	23	6	-	13	4
	50DAT	Plant Height	3	1	11	2	1	1	2	1	1
		No. of Tiller	-	-	11	-	79	5	-	15	3
	Heading	Plant Height	3	1	1	2	1	1	2	1	1
		No. of Tiller	-	-	73	-	-	8	-	-	4
	Harvesting	Culm Length	-	1	1	12	1	1	4	1	1
		No. of Panicles	26	4	2	13	3	1	9	2	1
		Panicle Length	-	-	7	-	19	4	-	9	3
		No. of Glumous Flowers	-	-	61	-	-	16	-	158	10
Barely	Harvesting	Culm Length	-	2	1	79	1	1	8	1	1
		No. of Ears 1)	19	3	2	10	2	1	7	2	1
		Ear Length	-	13	4	-	7	3	-	4	2
		No. of Grains per Ear	-	39	9	-	16	5	-	10	4
Wheat	Harvesting	Culm Length	41	2	1	8	1	1	5	1	1
		Ear Length	-	7	3	88	4	2	27	3	2
		No. of Grains per Ear	-	-	15	-	-	6	-	24	4
Soybean	Harvesting	Stem Length	-	4	1	-	2	1	19	2	1
		No. of Nods per per mani Sten	40	5	2	18	4	2	12	3	1
		No. of Pods per Plant	-	-	-	-	-	50	-	-	23
		No. of Seeds per Pods	-	51	16	-	28	10	376	19	8

1) For the number of ears the sampling unit is an area of 0.18 m^2

References

1. Abraham, T. P.(1966). Investigation on Field Experimental Techniques with Rice Crop II. Sampling in Field Experiments for Estimation of Plant Characters and Incidence of Pests and Diseases. *Indian J. of Agricultural Science*, Vol. 36, No. 4, 180 - 192.

2. Chang, S. H.(1997). Study on the Sample Size for Plant Characteristics of Rice Crop. *The Proceedings of Mathematical Science, Institute of Mathematical Science, Keimyung University*, Vol. 17,

3. Chang, S. H. and Y. W. Ha.(1981). Determination of Size and Number of Sampling Units for Spike Count in Wheat. *The Korean Journal of Crop Science*, Vol. 26, No. 1, 293 - 297.

4. Kempthorne, O.(1962). *The Design and Analysis of Experiments.* John Wiley and Sons. Inc. 212 - 216.

5. Oh. W. K., Chang. S. H. and H. C. Lee.(1969). A Study on the Number of Sample Units for Yield Components(I). *J. of Korean Society of Soil and Fertilizers*, Vol. 2, No. 1, 75 - 78.