Normalising Transformations for the Yield of Arecanut

S. Bhagavan*

Abstract

The yield data of arecanut was subjected to normality test. The study indicated that the distribution is non-normal and exhibits a pattern similar to poisson population. Square root transformations were attempted and found that $\sqrt{X+3/8}$ transformation for both number of nuts and wet weight of nuts would normalise the variability or distribution.

Introduction

Most of the statistical analyses are carried out with the assumption that the data follows normal distribution. But in reality all the data which we come across may not follow the normal and multivariate normal distribution. Therefore, the study on nature of distribution of the data which we deal more often and the methods of normalising transformations became necessary. Various transformations, taking into consideration the parametrical relationship between mean and variance have been suggested (Rao, 1952; Snedecor and Cochran, 1967) with illustrative examples. The present study aims at determining whether the yield data of arecanut is distributed normally or not and finding suitable normalising transformation in case of non-normality.

Materials and Methods

Ten year yield data (from 1963-'64 to 1973-'74) on number of nuts and wet weight of nuts for 508 areca palms grown under uniform conditions at CPCRI Regional Station, Vittal were considered.

Frequency distribution, co-efficient of skewness (β_1) and kurtosis (β_2) were obtained and subjected to tests of skewness and kurtosis (Chakravarthy, et al., 1967) to study whether the data agrees or not with normal distribution.

Several normalising transformations such as $\sqrt{X}$, $\sqrt{X+1/2}$, $\sqrt{X+2/3}$, $\sqrt{X+3/8}$ and $\sqrt{X+3/16}$ were attempted and the transformed data were subjected to tests of skewness and kurtosis.

Results and Discussion

Co-efficient of skewness and kurtosis were worked out separately for both the yield data viz., number of nuts and wet weight of nuts for all the ten years and are presented in Table 1. These were subjected to tests of skewness and kurtosis and found that the data doesn't follow normal distribution.

* Central Plantation Crops Research Institute, Regional Station, Vittal 574 243, Karnataka

The frequency distribution of ten years yield data do not conform to a normal distribution. The frequencies being larger in the lower and smaller in the higher classes, have characteristics of poisson distribution. When mean and variance are related as in a poisson distribution or when the variance is proportional to the mean, i. e. $\sigma x^2 = K\bar{x}$ then the suitable normalising transformation is given by $\sqrt{\bar{x}}$. The yield data under study, both in case of number of nuts and wet weight of nuts were found to exhibit the pattern of variance being proportional to the mean and hence square root transformations such as $\sqrt{X}$, $\sqrt{X+1/2}$, $\sqrt{X+2/3}$, $\sqrt{X+3/8}$ and $\sqrt{X+3/16}$ were attempted.

The values of β_1 and β_2 for the three transformations (nearly suitable ones) $\sqrt{X}$, $\sqrt{X+2/3}$ and $\sqrt{X+3/8}$ for both the yield characters are given in Tables 2 and 3.

Table 1. Co-efficient of skewness and kurtosis for data on number and weight of nuts

Year	Number of nuts		Weight of nuts	
	β_1	β_2	β_1	β_2
1964 - 65	0.73*	3.19	0.51*	2.87
1965 - 66	1.05*	4.06*	0.50*	3.15
1966 - 67	0.84*	3.91*	0.70*	3.50*
1967 - 68	0.70*	3.82*	0.27*	2.98
1968 - 69	6.29*	11.82*	5.63*	10.44*
1969 - 70	0.57*	3.22	0.30*	2.72
1970 - 71	0.59*	3.52*	0.29*	2.94
1971 - 72	0.54*	3.47*	0.23*	2.79
1972 - 73	0.71*	3.78*	0.60*	4.03*
1973 - 74	1.09*	4.21*	0.69*	3.56*

* Significantly different from either $\beta_1=0$ or $\beta_2=3$

Table 2. Transformations, Co-efficient of skewness and kurtosis for data on number of nuts

Year	Transformations					
	$\sqrt{X}$		$\sqrt{X+2/3}$		$\sqrt{X+3/8}$	
	β_1	β_2	β_1	β_2	β_1	β_2
1964 - 65	0	2.39*	0	2.34*	0	2.35*
1965 - 66	0.02	2.94	0.01	2.82	0.01	2.85
1966 - 67	0.01	2.76	0	2.67	0	2.69
1967 - 68	0.15	2.67	0.10	2.58	0.11	2.60
1968 - 69	0.45*	2.80	0.67*	3.13	0.61*	3.05
1969 - 70	0.09	2.98	0.06	2.86	0.06	2.88
1970 - 71	0.18*	3.20	0.13	3.08	0.14	3.10
1971 - 72	0.22*	2.70	0.17*	2.62	0.15	2.64
1972 - 73	0.21*	3.32	0.14	3.16	0.15	3.19
1973 - 74	0.07	2.30*	0.04	2.29	0.05	2.29*

* Significantly different from either $\beta_1=0$ or $\beta_2=3$

Normalising transformation for number of nuts

The results in Table 2 indicate that $\sqrt{x+3/8}$ is the most suitable transformation. Seven out of 10 populations have shown agreement with the normal distribution. Two populations showed disagreement only with respect to kurtosis whereas the remaining population (1968 - 1969) showed skewness even after transformation.

Normalising transformation for weight of nuts

The results in Table 3 reveal that $\sqrt{x}$ transformation is found to exhibit satisfactory agreement with normality. Eventhough six out of ten populations showed disagreement with normality, the variability from normality is not very much. Since $\sqrt{x+3/8}$ transformation has already been found as the most suitable one in the case of number of nuts, the same transformation may be preferred in the case of weight of nuts too.

Finally, it may now be concluded that $\sqrt{x+3/8}$ transformation approximates the given data on both number of nuts and weight of nuts to a normally distributed one.

Acknowledgement

The author expresses his deep sense of gratitude to Mr. K. Shama Bhat, Scientist S-3, CPCRI Regional Station, Vittal for providing facilities and encouragement.

Table 3. Transformations, Co-efficient of skewness and kurtosis for data on weight of nuts

Years	Transformations					
	$\sqrt{x}$		$\sqrt{x+2/3}$		$\sqrt{x+3/8}$	
	β_1	β_2	β_1	β_2	β_1	β_2
1964 - 65	0.02	2.45*	0	2.20*	0	2.22*
1965 - 66	0.12	2.88	0	2.46*	0.01	2.51*
1966 - 67	0.01	2.65	0.01	2.42*	0	2.44*
1967 - 68	0.30*	2.62	0.06	2.27*	0.10	2.32*
1968 - 69	0.40*	2.68	1.53*	4.34*	1.24*	3.91*
1969 - 70	0.08	3.08	0.04	2.56	0.07	2.64
1970 - 71	0.34*	3.38	0.09	2.79	0.13	2.88
1971 - 72	0.36*	2.68	0.11	2.32*	0.15	2.38*
1972 - 73	0.12	3.43	0.04	2.86	0.07	2.93
1973 - 74	0.15	2.26*	0	2.21*	0.02	2.20*

* Significantly different from either $\beta_1=0$ or $\beta_2=3$

References

CHAKRAVARTHI, I.M., LAHA, R. G. and ROY, J. 1967. *Hand Book of Methods of applied Statistics*. John Wiley & Sons, New York PP 460.

RAO, C. R. 1952. *Advanced Statistical Methods in Biometric Research*. John Wiley & Sons, New York PP 390.

SNEDECOR, G. W. and COCHRAN, W. G. 1967. *Statistical methods*. Oxford & IBH Publishing Co., Calcutta PP 594.