



CONFIDENCE LIMITS FOR THE CV DATA OF THE FIELD EXPERIMENTS ON WHEAT CROP: PEARSON TYPE IV DISTRIBUTION

¹Raykundaliya, D. P., ¹Patel, S.R. & ²Patel, N.P.

¹PG Department of Statistics, Sardar Patel University, Vallabh Vidyanagar – 388 120

²Department of Agricultural Statistics, B. A. College of Agriculture, Anand Agricultural University, Anand.

Corresponding Author*: dharmeshrky_04@yahoo.com

ABSTRACT

The coefficient of variation (CV) of the field experiments on wheat crop showed positive skewed distribution. The Pearson Type-IV distribution to the data is best fitted. The 95% confidence limits showed that for reliability of the wheat crop experiments the CV should be more than 3.46 (*i.e.* 4) and less than 17.89 (*i.e.* 18) per cent difference of the estimator.

KEY WORDS: Coefficient of Variation, Mean, Variance, Confidence limits, Pearson Type-IV Distribution, Wheat Crop etc.

INTRODUCTION

In agriculture large number of experiments are conducted to test the effects of the treatments on growth and yield parameters of field crops and also their effects on soil productivity. The results of such experiments are influenced by several factors [3] requiring some index to judge the reliability of the results. Also one has to use the criterion of homogeneity for forming the blocks and for this purpose CV is used as a tool. If the CV is between certain limits one can say that block has homogeneity in the yield character of the crop under study. To obtain such limits, that is to obtain confidence limits for the necessary parameter is important. Here the study of such confidence interval is done for the distribution which is fitted to the given data through CV. Thus for the reliability of the results of the experiments the analysis of CV is necessary. Deciding on which distribution is to use for modeling CV data of an experiment is a common problem in Agricultural Science. Traditionally, modeling has been based on the mean-variance analysis assuming a symmetric distribution, but in literature it is well documented that returns in share is not fully captured by the first two moments of the distribution due to distribution of return having long tail and high peak or a kind of positive skewed distribution[4]. In the area of finance, especially for the risk management, some heavy tailed and skewed distributions are strongly recommended. Distribution of CV is also having high peak and long tail. Patel et al [7] discuss the method of fitting lognormal distribution to CV data of mustered crop experiment. Given a need to allow for fat tail, skewness and all unimodal distribution it has become interesting to study frameworks that are flexible enough to accommodate distributions with broad range of properties. The Pearson [11] family is a particular family of distributions with easily

expressible density functions, which can be used to model a wide scale of distributions with various skewness and kurtosis. The Pearson Type IV distribution was employed recently to characterize the shape of pull distribution, which is often asymmetric with non Gaussian tail in Physics [5].

The paper is organized as follows: Section II discusses about the source of the data collection as well as type of the distribution fitted to CV data. In section III we discuss statistical analysis of data. Section IV is devoted for the method of fitting of Pearson Type IV distribution. Section V is concluded for the analysis.

II. MATERIAL AND METHODS

The CV values for yield character (secondary data) of 1037 experiments conducted on wheat crop during the years 1992-93 to 2002-03 at different research stations of Gujarat Agricultural University (GAU) were collected from the reports of research stations [1] for this study. With the help of SPSS 15 [10] and Ms Excel software, the data were analyzed to know the nature of the distribution. The moments of these data are used and method of moments is worked out, which indicates that the Pearson Type-IV distribution fits well. The detailed calculation of fitting of distribution is given in Section IV. Further 95% confidence limits for the mean of Pearson Type-IV distribution were obtained.

III. Statistical Analysis

The CV values of 1037 field experiments on wheat crop were analyzed with the help of SPSS 15 software [10]. A histogram with superimposing normal curve (Fig. 1) is obtained. The shape of curve showed positive skewed distribution with high picks and long tail on the right hand side, which leads to Pearson family of distributions [7]. Also Normal Probability Plot of CV data is depicted in Fig. 2.

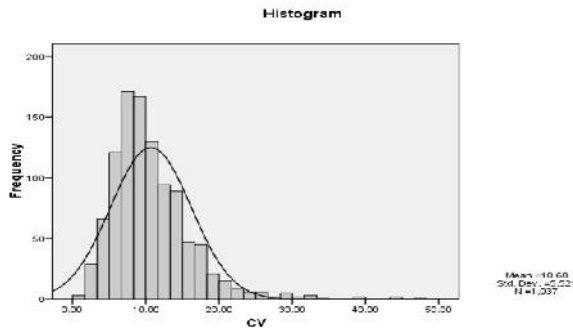


FIGURE 1: Histogram of CV data of Wheat Crop

A nonparametric test – one sample Kolmogorov-Sminrov goodness of fit test is used to assess the fit of the Normal distribution to the data.

TABLE 1: One Sample Kolmogorov-Sminrov Test for CV (n=1037)

Mean	10.68
Variance	30.48
Kolmogooov-Sminrov Z	3.05**

The Kolmogorov-Sminrov Test ($Z=3.05^{**}$) indicates that the assumed distribution (Normal distribution) does not fit well and need to search a distribution having high peak and long tail on the right side which is one of the members of the Pearson family of distributions. The Pearson Type IV distribution is tried to fit the data of wheat experiments. The computation for fitting of Pearson Type IV distribution is discussed in the Section IV. For the wheat data mean and variance of CV are given in Table 1.

IV. Computation for fitting of Pearson Type IV Distribution

Pearson's Type-IV distribution is obtained by fitting $b_2x^2 + b_1x + b_0$ as $b_2\{(x-b)^2 + a^2\}$ by taking imaginary roots $b+ia$ and $b-ia$ in Pearsonian differential equation $\frac{df}{dx} = \frac{(x-\alpha)f}{b_0 + b_1x + b_2x^2}$ given in [11] which after integrating out differential equation to get the probability density function of Pearsonian's Type-IV as

$$f(x) = \frac{1}{aF(r,\nu)} \left(1 + \frac{x^2}{a^2}\right)^{-m} \exp\left(-\nu \tan^{-1} \frac{x}{a}\right)$$

$$-\infty \leq x \leq \infty,$$

Where, a is scale parameter whereas the parameters m and ν are shape parameters which jointly determine the degree of skewness and kurtosis of the distribution. Also $F(r,\nu)$ has been used from Tables for Statisticians and Biometricians [8]. The kurtosis of the distribution is decreasing in the parameter m and the distribution is negatively (positively) skewed when $\nu > 0 (< 0)$. Note that when $\nu = 0$ the distribution reduces to Student's t distribution with $2m-1$ degrees of freedom. More generally, the Pearson Type IV distribution is as an asymmetric version of t -distribution. As the distribution is characterized by the parameters involved in the

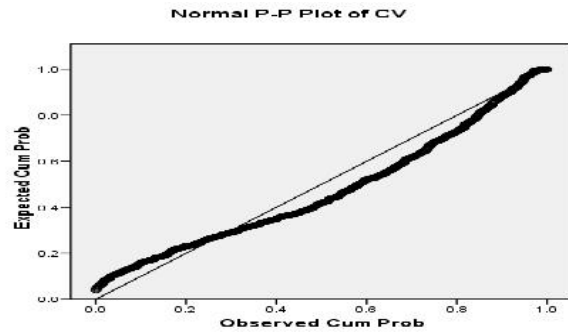


FIGURE 2: Normal Probability Plot of CV data of Wheat Crop

distribution and these are uniquely determined by first four moments of the distribution. When Karl Pearson developed his family of distributions, the maximum likelihood method was unknown. Pearson used method of moments, which is not really adequate in many cases, but may be used to provide starting value to a maximum likelihood fitter. The maximum likelihood estimates of the parameters can be obtained by minimizing the negative likelihood of Pearson Type IV distribution [6]. Due to complexity of solving intrinsic equations the method of moments is used to estimate parameters of the distribution [11]. After obtaining the raw moments about the origin and using recurrence relation for the given data for $n=1037$ observations one can obtain:

$$\hat{\mu}'_1 = 10.6796 \quad \hat{\mu}'_2 = 30.4813 \quad (\text{obtained through method of moments})$$

$$\hat{\beta}'_1 = 1.84031 \quad \hat{\beta}'_2 = 6.5489$$

$$b_1 = \frac{-\sqrt{\hat{\mu}'_2 \hat{\beta}'_1} (\hat{\beta}'_2 + 3)}{D'} = -2.8151,$$

$$\text{where } D' = 2 \left[5\hat{\beta}'_2 - 6\hat{\beta}'_1 - 9 \right] = 25.4053$$

$$b_2 = (2\hat{\beta}'_2 - 3\hat{\beta}'_1 - 6) / D' = -0.2793$$

$$b_0 = -\hat{\mu}'_2 (4\hat{\beta}'_2 - 3\hat{\beta}'_1) / D' = -24.8055$$

$$\text{Kappa } (k) = \frac{b_1^2}{4b_0b_2} = 0.2859$$

Since b_0 and b_2 are of same sign and $\text{Kappa} < 1$, indicates that Pearson's Type-IV distribution fits well to the CV data of the wheat experiment.

$$m = \frac{-1}{2b_2} = 1.7897$$

$$r = 2m - 2 = 1.5793$$

$$c = \hat{\mu}'_1 r b_2 = -4.7122$$

$$a = \frac{-\hat{\mu}'_1 r}{\nu} = -89.9794$$

$$\nu = \sqrt{\frac{\hat{\beta}'_1 (r-2)^2}{16(r-1)}} = 0.18745$$

$$C = \frac{1}{aF(r,\nu)} = -0.0447$$

$$\Phi = \tan^{-1}(\nu) = 0.1853 \quad (\text{it is in radian})$$

Karl Pearson vide's Table LIV, tables for Statisticians and Biometricians [8].

$$\Phi \text{ (In degree)} = \frac{0.185298 * 180^0}{3.14} = 10.62218$$

According to table

$\frac{\Phi}{\phi}$	\rightarrow	1	2
$\phi \downarrow$			
10		0.3042	0.2048
11		0.3049	0.2067
10.62218		0.3046	0.2059

The final $F(r, \nu)$ value for $\phi = 10.6222$ is 0.2484

$$\hat{\mu}_2'$$

= Based on recurrence relation of Karl Pearson's Type-IV is

$$\frac{a^2}{r(r-1)} + [r + \nu^2] = 14172.46$$

$$\hat{\mu}_2 = \hat{\mu}_2' - (\hat{\mu}_1')^2 = 14058.4$$

$$S.D. = 118.5681$$

By Slutsky's theorem, $Pr\{| - | < \} \rightarrow 1$ and hence the asymptotic distribution of $\hat{\mu}_2'$ is normal with mean and variance $\frac{1}{n}$. Therefore, asymptotic 95% confidence interval for mean of CV is given by:

$$\text{Lower bound} = \bar{x} - Z_{0.025} \left(\frac{S.D.}{\sqrt{n}} \right)$$

$$= 3.4630 \text{ and}$$

$$\text{Upper bound} = \bar{x} + Z_{0.025} \left(\frac{S.D.}{\sqrt{n}} \right)$$

$$= 17.8962$$

V. CONCLUSION

The statistical analysis of 1037 experiments conducted during 1992-2003 on different research stations showed that Pearson Type-IV distribution fits well to the CV data. Confidence limits indicate that the results of experiments

should be accepted if CV value lies between 3.46% to 17.89%. The rounded values for both the limits are 4% and 18%. The results thus lead to the conclusion that the results of field experiments on wheat crop be accepted (with high reliability) if CV of such experiments lie between 4 to 18 per cent.

REFERENCES

- [1] Anonymous (1992-93 to 2002-03). Reports of Main Wheat Research Station, GAU, Vijapur (Gujarat).
- [2] Bhattacharya, M., Mishra, N. and Kodase, B. (2008). MaxVar for Non-normal and Heteroskedastic Returns. European Journal of Operation Research 191, 386-397.
- [3] Federer, W.T. (1955). Experimental design. The Mac Millan Co., New York. Second edition. pp. 40.
- [4] Francis, J.(1975). Skewness and Investors' Decisions, Journal of Financial and Quantitative Analysis, 10, 163-172.
- [5] Heinrich, J. (2004). A Guide to the Pearson Type IV distribution, University of Pennsylvania.
- [6] Nagahara, Y. [1999]. The PDF and CF of Pearson Type IV distributions and the ML estimation of the parameters. Statistics & Probability Letters, 43, 251-264.
- [7] Patel, N.P.; Raykundaliya, D.P.; Patel, S.R. Darji, V.B. and Dixit, S.K. (2010). Confidence Limits for the CV Data of Field Experiments on Mustard Crop, Agric. Sci. Digest., 30 (3): 230-231.
- [8] Pearson, K. (1914). Tables for Statistician and Biometrician, Cambridge University Press.
- [9] Pearson, K. (1895). Contributions to the Mathematical Theory of Evolution.II. Skew Variation in Homogeneous Material. Philosophical Transactions of the Royal Society of London. A. 186. 343-414.
- [10] SPSS 15, www.spss.com
- [11] Stuart, A. and Oed, J.K. (1994). Kendall's Advanced Theory of Statistics, Volume 1: Distribution Theory, A Hodder Arnold Publication, Ed.6.