MAUSAM, 49, 4 (October 1998), 493-498

551.577.3:519.24

# Four statistical models for wet-spell analysis

S.D. GORE and PARVIZ NASIRI

*Deptt. of Statistics, University of Pune, Pune - 411 007, India*

*(Received 15 November 1996, Modified 8 December 1997 )*

**सार** — आर्द्र काल खण्ड विश्लेषण वर्षा के विश्लेषण का महत्वपूर्ण भाग है। आर्द्र काल खण्ड की वेला का वितरण वर्षा के कालिक वितरण के लिए उपयोगी जानकारी प्रदान करता है। विभिन्न प्रायिकता वितरणों के माध्यम से इस वितरण का परम्परागत मॉडल तैयार किया गया है। इस शोध-पत्र में हमने ऐसे चार मॉडलों, नामतः कोचरन मॉडल, ट्रंकेटेड प्वासों वितरण, ट्रंकेटेड नाकारात्मक द्विपद वितरण और लॉगरिथम श्रृंखला वितरण का तुलनात्मक अध्ययन किया है। भारत के पांच वर्षामापी केन्द्रों के अनुप्रयोग की मदद से इन तुलनात्मक अध्ययनों को और निखारा गया है।

**ABSTRACT.** Wet-spell analysis is an important part of rainfall analysis. The distribution of the length of wet-spells provides useful information on the temporal distribution of rainfall. This distribution has traditionally been modelled through different probability distributions. Here we compare four such models, namely, Cochran's model, truncated Poisson distribution, truncated negative binomial distribution, and logarithmic series distribution. These comparisons are accomplished with help of application to five rainguage stations in India.

**Key words** - Bernoulli trials, Cochran's formula, Logarithmic series distribution, Negative binomial distribution, Poisson distribution, Run length, Truncated distribution, Wet spell.

## 1. Introduction

Analysis of wet spells is an important aspect of rainfall analysis. Length of a wet spell provides information on persistence of rainfall. A statistical analysis of wet spells involves searching for the best statistical model for the observed wet spell patterns. While this model may differ from one place to another, it can help identify places which are similar or different with reference to the temporal distribution of rainfall or, equivalently, wet spells.

In this paper, we discuss four statistical models for wet spell analysis. These four models are: (1) Cochran's model, (2) zero- truncated Poisson distribution, (3) zero-truncated negative binomial distribution and (4) logarithmic series distribution. We compare these four models by fitting them to some rainfall data obtained during the monsoon season in India. The model that fits best to a particular data set may be considered to provide the best description of wet spell patterns at the concerned place.

A wet spell is a run of wet days preceded by a dry day and followed by a dry day. An observation period can then be considered to be an alternating sequence of dry and wet spells. In this paper, our interest is in wet spells only and, as such, the information on dry spells is ignored. A day is defined to be (meteorologically) dry if the amount of precipitation does not exceed 2.5 mm. If the amount of precipitation during a day exceeds 2.5 mm, then that day is defined to be a wet day (Chowdhury 1981).

## 2. Cochran's formula

Cochran (1938) derived a formula for the probability of a success run of length '$r$' in a sequence of '$m$' independent Bernoulli trials as follows. Consider a sequence of '$m$' trials. There are two distinct ways in which '$r$' consecutive days are wet. First, the wet spell either starts on the first day or ends on the last day. In the former case, the first '$r$' days are wet and the $(r + 1)$th day is dry. In the latter case, the $(m - r)$th day is dry and the last '$r$' days are wet. In the former case, no information is available on the day preceding the first day, while in the latter case, no information is available on the day following the '$m$'-th day. Hence the probability of this event is $p^r q + qp^r = 2p^r q$. Second, the wet spell starts and ends on some intermediate days. The starting day of such a wet spell then must be one of the days among 2,3, ...., $(m - r)$. Thus, such a wet spell can occur in $(m - r - 1)$ distinct ways. In each of these possible occurrences, the wet spell is preceded by a dry day and is followed by a dry day. The probability of this event, therefore, is $(m - r - 1)\, p^r q^2$. Adding the two terms derived above, we obtain the probability of observing a wet spell of length '$r$' as

$$f_{rm} = 2p^r q + (m - r - 1)p^r q^2 \qquad (1)$$

If we actually count wet spells of different lengths over a certain period of '$m$' days, then we would also obtain the frequencies of actual occurrences of wet spells of length 1,

length 2, and so on. This will give us a table of the following form.

| $r$ | 1 | 2 | 3 | 4 | ... |
|---|---|---|---|---|---|
| $O_r$ | $O_1$ | $O_2$ | $O_3$ | $O_4$ | ... |
| $f_{rm}$ | $f_{1m}$ | $f_{2m}$ | $f_{3m}$ | $f_{4m}$ | ... |

where, $O_r$ is the observed frequency of wet spells of length '$r$'.

We can now compare the observed and expected frequencies by estimating '$p$' and substituting the estimated value in the above formula of $f_{rm}$ for $r$ = 1,2, .... . The goodness of fit can be tested using the standard Chi-square test. In the following subsection, we discuss the method of maximum likelihood estimation of the parameter $p$ of Cochran's model.

2.1. *Estimation: Method of maximum likelihood*

Let $P(X = r) = f_r$ for $r$ =1,2,..., $k$, where

$$f_r = 2p^r q + (m - r - 1)p^r q^2 \qquad (2)$$

The likelihood function is then given by

$$L = \prod_{r=1}^{k} f_r^{O_r} \quad \text{That is,}$$

$$L = \prod_{r=1}^{k} [2p^r q + (m - r - 1)p^r q^2]^{O_r} \qquad (3)$$

The likelihood equation is obtained by differentiating log- likelihood with respect to the parameter $p$. Thus we have

$$\frac{d \ln L}{dp} = \sum_{r=1}^{k} \frac{2rp^{r-1}q - 2p^r + r(m-r-1)p^{r-1}q^2 - 2(m-r-1)p^r q}{2p^r q + (m-r-1)p^r q^2} O_r \qquad (4)$$

If we write the likelihood equation in the form of $f(p) = 0$, then we have

$$f(p) = \sum_{r=1}^{k} \frac{2(rq-p) + (m-r-1)\, q\, (rq-2q)}{pq[2+(m-r-1)q]} O_r \qquad (5)$$

In order to use the Newton-Raphson method for numerically solving the likelihood equation, we also obtain

$$f'(p) = \sum_{r=1}^{k} O_r \frac{(-r-1-rnq-nq+np)\,(pq + npq^2/2)}{(pq + npq^2/2)^2}$$

$$- \frac{(rq - p + r\,nq^2/2 - npq)\,(q - p + nq^2/2 - npq)}{(pq + npq^2/2)^2} \qquad (6)$$

where $n = m - r - 1$. The maximum likelihood estimate $\hat{p}$ is then obtained by an iterative procedure.

2.2. *Fitting Cochran's model*

To illustrate an application of Cochran's formula, we consider rainfall data at the Osmanabad rainguage station in India. In this data set, we have 522 rainy days over a period of 122 days (months of June, July, August, and September) observed over 10 years. This means that the observed proportion of rainy days in the sample is 522/1220 = 0.42278688. The maximum likelihood estimate of $p$ is obtained to be $\hat{p}$ = 0.4682775. The following table gives the observed and expected frequencies of wet spells of different lengths, along with the chi-square test statistics for testing the goodness of fit.

| $r$ | 1 | 2 | 3 | 4 | 5 | — |
|---|---|---|---|---|---|---|
| $O_r$ | 192 | 58 | 23 | 10 | 18 | — |
| $f_{rm}$ | 163.86 | 76.11 | 35.35 | 16.42 | 9.26 | $\chi^2_3$= 24.21 |

## 3. Zero-truncated Poisson distribution

A discrete random variable $X$ is said to have a Poisson distribution with parameter $\lambda$ if its probability mass function is given by

$$f(x) = P[X = x] = \frac{e^{-\lambda}\lambda^x}{x!}, \quad x = 0,1,2,.... \qquad (7)$$

The zero-truncated Poisson distribution is obtained by discarding the value 0 from the set of possible values of the Poisson random variable. We note from Eqn. (7) that $f(0) = P[X = 0] = e^{-\lambda}$, and hence the probability mass function of the zero-truncated Poisson distribution becomes

$$f_T(n) = \frac{e^{-\lambda}\lambda^x}{x!(1-e^{-\lambda})}, \quad x = 1, 2, 3, .... \qquad (8)$$

## 4. Zero-truncated negative binomial distribution

The negative binomial distribution is obtained by inverse sampling of independent Bernoulli trials. The probability mass function of this distribution is given by

$$f(x;p,k) = P[X=x] = \frac{\Gamma(x+k)}{x!\Gamma(k)} p^k (1-p)^x, \qquad (9)$$

$$x = 0, 1, 2, \ldots$$

We note form Eqn. (9) that $f(0) = p^k$, and hence the probability mass function of the zero-truncated distribution becomes

$$f_T(x; p, k) = \frac{\Gamma(x+k)p^k(1-p)^x}{x!\,\Gamma(k)\,(1-p^k)}, \; x = 1, 2, \ldots . \qquad (10)$$

4.1. *Estimation: Brass estimators*

As estimating equation, Brass (1958) employed $\mu = \bar{x}$, $\mu_2 = s^2$, and $f_T(1) = n_1/n$. Estimators of the parameters $p$ and $k$ follow from these equations. We have

$$p^k = \frac{f_T(1)}{\mu p} \quad \text{and} \quad (1-p^k) = \frac{\mu p - f_T(1)}{\mu p} \tag{11}$$

These can be rewritten as

$$p = \frac{\mu}{\mu_2}[1 - f_T(1)] \quad \text{and} \quad k = \frac{p\mu - f_T(1)}{1-p} \tag{12}$$

On substituting $\mu = \bar{x}$, $\mu_2 = s^2$, and $f_T(1) = n_1/n$ into these equations, the resulting estimators become

$$p^* = \frac{\bar{x}}{s^2}\left[1 - \frac{n_1}{n}\right] \quad \text{and} \quad k^* = \frac{p^*\bar{x} - (n_1/n)}{1-p^*} \tag{13}$$

Brass demonstrated that these estimators are consistent, although they are not unbiased. However, when $n$ is large the effect of bias is only slight. The efficiency for most combinations of parameter values is above 90%.

4.2. *Illustrative examples*

(1) To illustrate the method of estimation of parameters in the truncated negative binomial distribution, we consider a sample of chromosome breakage that was originally given by Sampford (1955). Sample data are as follows:

| Number breaks $(x)$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Frequency $(nx)$ | 11 | 6 | 4 | 5 | 0 | 1 | 0 | 2 | 1 | 0 | 1 | 0 | 1 |

In summary $n = 32$, $n_1 = 11$, $n\bar{x} = 110$, $\bar{x} = 3.4375$ and $s^2 = 9.9315$.

Brass estimates $p^* = 0.2345$ and $k^* = 0.6040$.

Maximum likelihood estimates $\hat{p} = 0.2113$, $\hat{k} = 0.493$.

(2) To illustrate estimation in the truncated negative binomial distribution, we consider a sample of rainfall data obtained at Osmanabad. Sample data are as follows:

| $i$ | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| $fi$ | 192 | 58 | 23 | 28 |

In summary $n = 301$, $n_1 = 192$, $n\bar{x} = 522$, $\bar{x} = 1.73422$ and $s^2 = 1.7500$.

Brass estimates $p^* = 0.35887$ and $k^* = -0.02$. Note that the parameter $k$ is positive, and hence we write its estimated value as 0.0001, even though the estimated value is out of range.

Maximum liklihood estimates $\hat{p} = 0.3370581$, $\hat{k} = -0.0760$ replaced for the same reason by 0.0001.

| $O$ | 192 | 58 | 23 | 28 | Sample value | Table value |
|---|---|---|---|---|---|---|
| $E_3$ | 188.22 | 60.39 | 25.83 | 26.56 | $\chi^2 = 0.56$ | $\chi^2_1 = 3.84$ |
| $E_4$ | 183.48 | 60.83 | 26.88 | 29.81 | $\chi^2 = 1.19$ | $\chi^2_1 = 3.84$ |

*Note:* $E_3$ denotes the expected frequency obtained by using the Brass estimates, while $E_4$ denotes the expected frequency obtained by using the maximum likelihood estimates of the parameteres.

## 5. Logarithmic series distribution

The random variable $X$ has a logarithmic series distributions if its probability mass function is given by (*see*, for instance, Patil 1962, Patil and Wani 1963).

$$f(x) = \alpha \frac{\theta^x}{x}, \quad x = 1, 2, \ldots, ; 0<\theta<1 \tag{14}$$

where $\alpha = \dfrac{-1}{\ln(1-\theta)}$

The individual probabilities are terms in the series expansion of -1 ln (1-θ).

For a logarithmic series distribution with parameter θ, we have

$$E(X) = \frac{\alpha\theta}{(1-\theta)} \tag{15}$$

$$V(X) = \frac{\alpha\theta}{(1-\theta)^2}(1 - \alpha\theta) \tag{16}$$

Patil (1962) has tabulated values of $E(X)$ for values of θ in the interval (0, 1).

Table 1 gives the values of $\bar{x}$ for different values of θ as tabulated by Patil (1962).

5.1. *Estimation: Method of maximum likelihood*

Let $x_1, x_2, \ldots x_n$ be a random sample from the logarithmic series distribution with parameter θ. The likelihood function is then given by

$$L(\theta) = [\alpha]^n \frac{\theta(\Sigma_{i=1}^n x_i)}{\Pi_{i=1}^n x_i} \tag{17}$$

Differentiating the log-likelihood and equating the derivative to zero, we have the following likelihood equation

$$\bar{x} = \frac{\alpha\theta}{(1-\theta)} \tag{18}$$

Note that this equation coincides with that derived from the method of moments.

5.2. *Application to publication data*

The following data gives the distribution of 1534 biologists according to the number of research paper to

**TABLE 1**

**Means of logarithmic series distribution for θ = .01 (.01).99**

| θ | 00 | 01 | 02 | 03 | 04 | 05 | 06 | 07 | 08 | 09 |
|---|---|---|---|---|---|---|---|---|---|---|
| 00 | - | 1.0050 | 1.0102 | 1.0154 | 1.0207 | 1.0261 | 1.0316 | 1.0372 | 1.0429 | 1.0487 |
| 10 | 1.0546 | 1.0606 | 1.0667 | 1.0730 | 1.0704 | 1.0858 | 1.0925 | 1.0992 | 1.1061 | 1.1132 |
| 20 | 1.1204 | 1.1277 | 1.1352 | 1.1429 | 1.1507 | 1.1587 | 1.1669 | 1.1752 | 1.1838 | 1.1926 |
| 30 | 1.2016 | 1.2108 | 1.2202 | 1.2299 | 1.2398 | 1.2500 | 1.2604 | 1.2711 | 1.2821 | 1.2934 |
| 40 | 1.3051 | 1.3170 | 1.3294 | 1.3421 | 1.3551 | 1.3687 | 1.3825 | 1.3968 | 1.4116 | 1.4269 |
| 50 | 1.4427 | 1.4591 | 1.4760 | 1.4935 | 1.5117 | 1.5306 | 1.5503 | 1.5706 | 1.5919 | 1.6140 |
| 60 | 1.6370 | 1.6611 | 1.6862 | 1.7126 | 1.7401 | 1.7690 | 1.7994 | 1.8313 | 1.8650 | 1.9005 |
| 70 | 1.9380 | 1.9778 | 2.0200 | 2.0649 | 2.1128 | 2.1640 | 2.2189 | 2.2779 | 2.3416 | 2.3800 |
| 80 | 2.4853 | 2.5670 | 2.6566 | 2.7553 | 2.8648 | 2.9870 | 3.1244 | 3.2802 | 3.4587 | 3.6656 |
| 90 | 3.9087 | 4.1991 | 4.6716 | 4.9960 | 5.5686 | 6.3424 | 7.4560 | 9.2208 | 12.5255 | 21.4976 |

their credit. In the review of Applied Entonology, Vol. 24, 1936.

| No. of paper per author ($x_i$) | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| No. of authors $f(x_i)$ | 1062 | 263 | 120 | 50 | 22 | 7 | 6 | 2 | 0 | 1 | 1 |

Mean of sample is $\bar{x} = 1.55085$, and so $\hat{\theta} = 0.5602$.

After obtaining $\hat{\theta}$ we can compute the theoretical frequencies and compare them with the observed frequencies. The result of a Chi-square test of goodness of fit applied to this data is given below.

| $i$ | $P_i$ | $E_i$ | $O_i$ |
|---|---|---|---|
| 1 | 0.68198 | 1046.5 | 1062 |
| 2 | 0.19102 | 293.03 | 263 |
| 3 | 0.07130 | 109.44 | 120 |
| 4 | 0.02997 | 45.98 | 50 |
| 5 | 0.01343 | 20.61 | 22 |
| 6 | 0.01230 | 18.44 | 17 |

Chi-square calculated = 4.88

Chi-square tabulated = 11.10

*Note*: In the above table, $P_i$ denotes the probability of the value $i$, $E_i$, denotes the expected frequency under the fitted logarithmic series distribution, and $O_i$ denotes the observed frequency of $i$ in the data.

Since the calculated value of the Chi-square test statistic is insignificant, we may accept *Ho* , that is, we may conclude that the logarithmic series distribution may adequately describe the observed data.

5.3. *Application to chromosome breakage data*

We can find several nature example of the logrithmic series data in the literature. We consider a sample of chromosome breakage that was originally given by Sampford (1955). Sample data are as follows:

| Number breaks ($x$) | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Frequency ($n_x$) | 11 | 6 | 4 | 5 | 0 | 1 | 0 | 2 | 1 | 0 | 1 | 0 | 1 |

$$\bar{X} = \frac{\Sigma_{n=1}^{13} x n_x}{\Sigma_{n=1}^{13} n_x} = \frac{110}{32} = 3.4375. \qquad (19)$$

$$\hat{\theta} = 0.875. \qquad (20)$$

## 6. Application to rainfall data

Now we illustrate application of the logarithmic series distribution to rainfall data analysis. For this purpose, we have collected daily rainfall data at five rainguage stations in India. The selection of the these rainguage station has not been on the basis of the climate conditions or typical rainfall patterns, but purely on the basis of availability of rainfall data. For the purpose of our discussion, we shall denote these five rainguage stations by the simple notation described below. These rainguage stations are Osmanabad, Buldhana, Wardha, Gondia and Bhir. Note that the data at different stations was not available for the same period. The following table shows the number of years corresponding to the five stations.

| Station | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| No. of year | 10 | 21 | 24 | 40 | 29 |

For each rainguage station, we define a random variable $X$, taking only two possible values, namely 0 and 1, as follows :

If the amount of rainfall on a particular day dose not exceed 2.5 mm, then the random variable $X$ takes the value 0. If the amount of rainfall on a particular day exceeds 2.5 mm then the random variable $X$ takes the value of 1. Note that on any day, the random variable $X$ takes only one of the two possible value 0 and 1, since the two events described above are mutually exclusive. Defining the random variable $X$ as above, the daily rainfall data was coded to produce values of the random variable $X$ for every day of the year. A wet spell is then defined as a period of sussessive rainy days. The number of rainy days in a single stretch is then defined as the length of the wet spell. Denoting the length of a wet spell by $x_i$, $i = 1,2, \ldots$ and the corresponding frequencies by $f_i$, $i = 1,2, \ldots$, we obtain the following data for Osmanabad.

| $x_i$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|
| $f_i$ | 192 | 58 | 23 | 10 | 11 | 3 | 1 | 2 | 1 |

The sample mean is $\bar{x} = 1.73422$ and so

$\hat{\theta} = 0.6379046$

After obtaining $\hat{\theta}$ we can compute the theoretical or expected frequencies and compare them with observed ones.The result of a goodness of fit test is as follows.

| $i$ | $P_i$ | $E_i$ | $O_i$ |
|---|---|---|---|
| 1 | 0.62796 | 189.01390 | 192 |
| 2 | 0.20028 | 60.28642 | 58 |
| 3 | 0.08518 | 25.63799 | 23 |
| 4 | 0.04075 | 12.26594 | 10 |
| 5 | 0.04583 | 13.79589 | 18 |

Chi-square calculated = 2.11

Chi-square tabulated = 9.49

Note that Cochran's model was applied to this data set in Section 2.2 and gave a Chi-square value of 24.21. In Section 4.2, the zero-truncated negative binomial distribution is fitted to the same data, with Brass estimates producing a Chi-square value of 0.56 and the maximum likelihood estimates giving a Chi-square value of 1.19. In comparison, the logarithmic series distribution gives a Chi-square value of 2.11. It then appears that the zero-truncated negative binomial distribution fits better than the logarithmic series distribution.

In this connection, it may be noted that the Chi-square value in case of the zero-truncated negative binomial distribution has 2 degrees of freedom, while that in case of the logarithmic series distribution is based on 3 degrees of freedom. This is so because the logarithmic series distribution has only one parameter, while the zero-truncated negative binomial distribution has two parameters, and that affects the degrees of freedom of the Chi-square test statistic.

We now summarise the results of fitting all the four distributions described above to rainfall data at the other four rainguage stations. In each case, $O$ denotes the observed frequency, $E_1$ denotes the expected frequency according to the Cochran's model, $E_2$ denotes the expected frequency according to the zero-truncated Poisson distribution, $E_3$ denotes the expected frequency according to the zero-truncated negative binomial distribution (Brass estimator), $E_4$ denotes the expected frequency according to the zero-truncated negative binomial distribution (maximum likelihood estimator), and $E_5$ denotes the expected frequency according to the logarithmic series distribution.

6.1. *Buldhana*

| | 1 | 2 | 3 | 4 | 5 | Sample value | Table value |
|---|---|---|---|---|---|---|---|
| $O$ | 289 | 114 | 62 | 25 | 29 | | |
| $E_1$ | 263.40 | 129.06 | 63.23 | 30.40 | 32.91 | $\chi^2 = 5.69$ | $\chi^2_3 = 7.81$ |
| $E_2$ | 225.47 | 167.59 | 83.04 | 30.86 | 12.04 | $\chi^2 = 65.37$ | $\chi^2_3 = 7.81$ |
| $E_3$ | 291.47 | 112.86 | 53.23 | 27.33 | 34.11 | $\chi^2 = 2.44$ | $\chi^2_3 = 5.99$ |
| $E_4$ | 290.55 | 114.49 | 53.83 | 27.35 | 32.78 | $\chi^2 = 1.89$ | $\chi^2_3 = 5.99$ |
| $E_5$ | 303.55 | 105.56 | 48.95 | 25.53 | 35.41 | $\chi^2 = 6.02$ | $\chi^2_3 = 7.81$ |

It may noted here that most of the distributions of fits satisfactorily, while their relative performances are more or less similar to what was observed for Osmanabad. Only, the maximum likelihood estimates have given a better fit than Brass estimates for the zero-truncated negative binomial distribution.

6.2. *Wardha*

| | 1 | 2 | 3 | 4 | 5 | Sample value | Table value |
|---|---|---|---|---|---|---|---|
| $O$ | 387 | 136 | 65 | 31 | 46 | | |
| $E_1$ | 341.91 | 166.69 | 81,26 | 39.30 | 35.84 | $\chi^2 = 19.48$ | $\chi^2_3 = 7.81$ |
| $E_2$ | 293.52 | 214.40 | 104.40 | 38.13 | 14.55 | $\chi^2 = 142.62$ | $\chi^2_3 = 7.81$ |
| $E_3$ | 327.78 | 156.09 | 81.05 | 43.83 | 56.25 | $\chi^2 = 22.09$ | $\chi^2_3 = 5.99$ |
| $E_4$ | 378.76 | 137.89 | 63.82 | 32.69 | 51.84 | $\chi^2 = 0.97$ | $\chi^2_3 = 5.99$ |
| $E_5$ | 391.51 | 135.18 | 62.23 | 32.23 | 43.85 | $\chi^2 = 0.33$ | $\chi^2_3 = 7.81$ |

The results for Wardha are similar to those obtained for Osmanabad and Buldhana. Note that the zero-truncated Poisson distribution gives a very bad fit compared to the other three distributions.

6.3. *Gondia*

The zero truncated negative binomial and the logarithmic series distributions have provided with good fits, while the other two models have not provided with good fits, although the relative performances of the four models are very similar to the earlier results.

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | Sample value | Table value |
|---|---|---|---|---|---|---|---|---|---|
| $O$ | 546 | 244 | 132 | 69 | 50 | 27 | 49 | | |
| $E_1$ | 514.67 | 273.47 | 146.26 | 77.97 | 57.75 | 22.15 | 24.73 | $\chi^2 = 33.43$ | $\chi^2_5 = 11.07$ |
| $E_2$ | 358.40 | 315.67 | 229.97 | 112.81 | 44.26 | 14.47 | 5.42 | $\chi^2 = 551.92$ | $\chi^2_5 = 11.07$ |
| $E_3$ | 543.12 | 274.15 | 131.74 | 75.37 | 44.88 | 27.43 | 20.32 | $\chi^2 = 44.94$ | $\chi^2_4 = 9.49$ |
| $E_4$ | 544.78 | 245.36 | 130.69 | 74.98 | 44.88 | 27.60 | 48.71 | $\chi^2 = 1.10$ | $\chi^2_4 = 9.49$ |
| $E_5$ | 584.37 | 225.23 | 115.75 | 66.92 | 41.27 | 26.51 | 56.95 | $\chi^2 = 9.39$ | $\chi^2_5 = 11.07$ |

6.4. *Bhir*

| | 1 | 2 | 3 | 4 | 5 | Sample value | Table value |
|---|---|---|---|---|---|---|---|
| $O$ | 418 | 144 | 64 | 24 | 38 | | |
| $E_1$ | 366.35 | 171.36 | 80.15 | 37.49 | 32.65 | $\chi^2 = 20.63$ | $\chi^2_3 = 7.81$ |
| $E_2$ | 343.75 | 216.16 | 90.61 | 28.49 | 8.99 | $\chi^2 = 142.26$ | $\chi^2_3 = 7.81$ |
| $E_3$ | 414.94 | 150.03 | 64.34 | 29.75 | 28.94 | $\chi^2 = 04.21$ | $\chi^2_3 = 5.99$ |
| $E_4$ | 415.63 | 148.07 | 63.88 | 29.97 | 30.45 | $\chi^2 = 3.17$ | $\chi^2_3 = 5.99$ |
| $E_5$ | 427.98 | 138.24 | 59.54 | 28.85 | 33.39 | $\chi^2 = 2.56$ | $\chi^2_3 = 7.81$ |

At Bhir, the logarithmic series distribution gives the best fit among the four models being compared here. This shows that the logarithmic series distribution can provide a good description of wet spell distribution at some places. It should now be interesting to find out specific features of rainfall patterns that are better described by the logarithmic series distribution. This is a point that the authors are now working on. Some apparent characterization will definitely help scientist in general, and meteorologists in particular, identify areas where wet spells follow the logarithmic series distribution.

*Acknowledgements*

The second author wishes to thank the Ministry of Culture and High Education, Islamic Republic of Iran, for the scholarship which enabled him to continue his studies with the Department of Statistics, University of Pune. He also wishes to thank the Department of Statistics, University of Pune, for giving him an opportunity to pursue research in the area of his interest.

Both the authors wish to thank the Department of Statistics, University of Pune, for making the necessary facilities available for carrying out the research reported in this paper.

**References**

Brass, W., 1958, "Simplified methods of fitting the truncated negative binomial distribution," *Biometrika*, **45**, 59-68.

Chowdhury, A., 1981, "On the occurrence of wet and dry spells in Bihar," *Mausam*, **32**, 285-290.

Cochran, W.G., 1938, "On the probability of rain," *Quart. J. Roy. Meteor. Soc.*, **64**, 631-634.

Patil, G.P., 1962, "Some methods of estimation for the logarithmic series distribution," *Biometrics*, **18**, 68-75.

Patil, G.P. and Wani, J.K., 1963, "Maximum likelihood estimation for the complete and truncated logarithmic series distribution," *Proceedings of the International Symposium on Discrete Distributions, Montreal*, 398-409. Also, *Snakhya*, Series A, **27**, 271-280.

Sampford, M.R., 1955, "The truncated negative binomial distribution," *Biometrika*, **42**, 58-69.