APPLICATION OF POISSON PROCESS TO DROUGHT PREDICTION——THE CASE STUDY OF YUCHENG CITY

Yujian Yang[1*], Yingqiang Song[2]

[1]School of Civil and Architectural Engineering of Shandong University of Technology, Zibo city, China -yyjtshkh@126.com
[2] School of Civil and Architectural Engineering Shandong University of Technology, Zibo city, China - yingqiang_song@163.com

KEY WORDS: Drought, Poisson process, Prior, Posterior, Prediction, R language

ABSTRACT:

Open-source R language can implement quantitative research using the flexibility, adaptability and simplicity of bayesian inference models. Counts of drought which are regarded as the realizations of drought "event" in a Poisson process follows an Gamma distribution (prior distribution) in the case study of Yucheng city with the support of R language. That is, the annual drought counts in Yucheng City during the 10 years from 1974 to 1983 is regarded as the 10 realizations of the Poisson process, and the drought counts in Yucheng city during the 10 years is described by the Poisson distribution. The Gamma (1,1) of theta is used as the prior distribution, and its posterior distribution is calculated to be Gamma (60,11). The sample value of theta is obtained by random sampling from the posterior distribution of theta, and a Poisson sample is randomly generated for each value as the predicted value. The posterior mean value of theta sampling results, the mean and quantiles of the predicted value, and the probability estimation of theta and the predicted value are calculated to obtain the highest density region of theta. Studies have shown that Poisson model of drought prediction in Yucheng city has more flexibility, adaptability and simplicity and can provide posterior mean, highest-posterior density of the data, and the drought prediction of Yucheng city is a feasible method on basis of the stochastic Poisson process, thereby providing better reference value for drought reduction.

1. INTRODUCTION

The arid and semi-arid area which will continue to expand globally accounts for about 40% of the global land areas, and 78 % of the expansion of the arid and semi-arid area will occur in developing countries. Drought is not only one of the important agricultural disasters, but also is key issues for developing countries. Meteorological drought is an important type of drought, it is the phenomenon that the evaporation of a certain region is more than the precipitation or the precipitation is abnormally small in a relatively long period. The precipitation is often used as drought indicators, drought is one of the main climatic factors restricting agricultural production. Currently, drought and water shortage are a very important environmental problem facing the world (Jianping Huang, Haipeng Yu, Xiaodan Guan, Guoyin Wang, Ruixia Guo, 2016). Drought is a natural hazard, which is a result of a prolonged shortage of precipitation, high temperature and change in the weather pattern. Drought harms society, the economy and the natural environment, but it is difficult to identify and characterize (Asad Ellahi, Ibrahim, Almanjahie, Tajammal Hussain, Muhammad Zaffar Hashmi, Shahla Faisal, Jaz Hussain, 2020).

Droughts have been occurring with increased frequency and bringing with them considerable losses. From 2000 to 2018, China' s agriculture suffered an average annual area of 18450.22 thousand hectares, water shortage reached 2015.46 million people resulting from drought. Under the background of global climate change, the situation of drought in China is becoming more and more serious, and corresponding to the monitoring, assessment and response of drought are more important. North China is one of the main grain production areas in China. In recent years, droughts have occurred frequently in this region, which has brought great losses to agriculture. It is necessary to establish a certain degree of accuracy of agricultural drought prediction model and provide accurate agricultural drought warning and prediction to all levels of government departments (Pei-Yu Wu, Gene Jiing-Yun You, Ming HsiuChan, 2020) or (Qianchuan Mi, Xining Gao, Li Yue, 2022). Drought itself has randomness, random theory is a reasonable method to study drought disaster. On basis of drought's nature, the best approach to monitoring and assessing droughts is in terms of stochastic theories. As a consequence, this study applied Poisson distribution and Poisson process in the drought detection, and examine/interpret drought-related phenomena (Pei-Yu Wu, Gene Jiing-Yun You, Ming HsiuChan, 2020).

A growing body of literature studies probabilistic programming from a bayesian inference of perspectives. Discusses of the state-of-the-art advances in machine learning and artificial intelligence field, namely, bayesian optimization probabilistic programming, data compression and automatic model discovery (Ghahramani, 2015) or (Tenenbaum, Kemp, Griths, Goodman, 2011). Bayesian inference is an alternative approach assessing the uncertainty and interactions which are described by the posterior density based on Markov Chain Monte Carlo (Laura Poggio, Alessandro Gimona, Luigi Spezia, Mark J.Brewer, 2016) or (Andrew Gelman, Jennifer Hill, 2006) or (Carpenter, Gelman, Hoffman, Lee, Goodrich, Betancourt, Brubaker, J. Guo, P. Li, Ridell, 2017) or (Hoffman, Gelman, 2014).

The last few decades, bayesian inference has gained more and more attention from people in almost all the sciences, in particularly statistics domain. Many of these models were unsolvable just a few years ago, but now in a relative easy way. Currently, probabilistic programming based on R platform allows us to flexibly build modern bayesian inference which work independently of the model to fit interested model to study data, it is possible due to theoretical and computational

[1*]**Corresponding author**

advancements. Apparently, the necessity for flexible and transparent models contributed to the trend and the combination of flexible model specification and automatic inference of bayesian inference. Two typical characteristics was summarized in bayesian inference model, including Poisson distribution model. (1) Probability distributions, whether prior distribution or posterior distribution performed to describe unknown quantities. (2) The values of the parameters conditioned on the data from prior, posterior and likelihood function is updated using bayes' theorem. Generally, we can describe the process of constructing bayesian modeling in 3 steps: (1) Some data and assumptions on how this data could have been generated, a model of random variables is designed by combining and transforming way.(2)Using Bayes' theorem to condition on the data, a posterior distribution was obtained based on prior distribution and likelihood function to reduce the uncertainty for possible parameter values.(3) Checking and diagnosing how the model makes sense according to different criteria (Osvaldo, Martin, Ravin Kumar, Junpeng Lao, 2022). In a word, bayesian inference is a particular form of statistical inference based on prior distribution, likelihood function and posterior distribution using the combination of flexible model specification and automatic inference.

Previous study applied the rectangular pulses Poisson process model to the quantification and analysis of drought. A theoretical methodology of drought severity–duration–frequency analysis was proposed based on the model structure of the rectangular pulses Poisson process (ChulsangYoo, Daeha Kim, Tae-Woong Kim, Kyu-NamHwang, 2008). Hidden Markov models, especially those with a Poisson density governing the latent state-dependent emission probabilities, have enjoyed substantial and undeniable success in modeling natural hazards. Especially, for discretized empirical recurrence rates ratio (Bhaduri, 2020). A series of Poisson distributions are fit to sets of natural disasters, accidents and anthropogenic catastrophes using the Poisson density function(Zachary, 2018) or (Ho, Bhaduri, 2015) or (Guo, 2010). Poisson distribution is the distribution of any kind of "rare" event, or the Poisson distribution is a limiting case of a binomial distribution. More formally, it is a probability distribution on basis of bayesian inference such that: (1) the probability of observing an event; (2) the probability that an event occurred in a given interval is independent from the probability that an event occurs in any other disjoint interval; (3) the events are never simultaneous. This setup is called a "Poisson process"(McElreath, 2017) or (Martin, 2018) .

In a word, application of Poisson process based on Poisson distribution to drought is not only feasibility, but also new challenges for understanding of drought predictability and developing drought forecasting methods. Actually, counts of the droughts which are regarded as the implementation of drought event in a Poisson process follows an Gamma distribution (prior distribution) in this study. That is to say, 10 counts implementation in a Poisson process are Gamma-distributed, this case study based on Poisson probability programming in Yucheng city is very rare from the literatures, it is necessary to develop the related droughts to improve prediction model flexibility and adaptability, and the findings of this study can also be an important references for the policymakers and government to take decisions.

As summarized above, few studies have found drought prediction using Poisson process in Yucheng city, located in the core area of the Yellow-Huai River Plain. Thus. the paper is organized as follows: Section 2 generalizes materials and methods, Section 3 reviews and develops the drought prediction using Poisson distribution and Poisson process, and attempts to model drought prediction based on R language probabilistic programming, Section 4 draws conclusion and discussions to clarify the further improvements to the model.

2. MATERIALS AND METHODS

2.1 Study Area Situation

Yucheng city which is located in the northwest of Shandong Province is in the core area of Yellow-Huai River Plain. As a typical Fluvo-Aquic soil area and country-level demonstration of rural revitalization, the area which has sufficient light is in the warm temperate zone in the half moist monsoon climate region. The precipitation is less, and annual precipitation is unevenly distributed, mainly concentrated in July and August. The average annual precipitation is 616 mm, and the distribution between years and months is extremely uneven. The highest value of annual precipitation is 1144.4 mm, and the lowest value of annual precipitation is only 239 mm. The annual cumulative average evaporation was 2268.2 mm, the monthly maximum evaporation was 391 mm in June, and the monthly minimum evaporation was 57.4 mm in January. This uneven distribution of precipitation and evaporation can easily lead to drought. In this area, micro- topography is more complex, and the terrain slope is gently tilt from southwest to northeast and the soil is sandy clay loam and the wheat-maize rotation was conducted yearly in the crop fields in the area.

2.2 Poisson Distribution and Poisson Process

Discrete nonnegative numbers: {0, 1, 2, 3, ...} are usually modelled in the examples, we called the numbers type of variable as count data, one scenario where Poisson distribution is useful for count data (Martin, 2018). The posterior distribution embodies both the prior distribution and likelihood, the prior distribution is combined with the likelihood to obtain the posterior distribution of the parameter of interest. Conjugate priors which have the nice property of resulting posteriors distribution of the same distributional family can avoid intractability, and they are the convenient choice. For Poisson model, if the observed value is y= $(y_1, y_2, \ldots y_n)$, according to a set of discrete count data y, we wish to estimate their mean for the data using Poisson distribution then the likelihood (Xizhi Wu, 2020) is:

$$p(y|\lambda) \propto \prod_{i-1}^{n} \lambda^{y_i} e^{-\lambda} \qquad (1)$$

If the likelihood is Poisson and the prior Gamma distribution, then the posterior distribution will be a Gamma distribution (Osvaldo Martin, Ravin Kumar, Junpeng Lao, 2022). In the case study, conjugate prior distribution is:

$$Gamma(\alpha,\beta): \ p(\lambda) \propto e^{\beta\lambda} \lambda^{\alpha \ \ 1} \qquad (2)$$

Posterior distribution is:

$$p(\lambda|y) \sim Gamma(\alpha + ny, \beta + n) \qquad (3)$$

Poisson model was referred in the following formula (4) and formula (5)

$$p(y|\theta) = \prod_{i=1}^{n} Poisson(y_i|\theta) \qquad (4)$$

$$p(\theta) = Gamma(\theta|\alpha, \beta) \qquad (5)$$

https://doi.org/10.5194/isprs-archives-XLVIII-3-W1-2022-73-2022 | © Author(s) 2022. CC BY 4.0 License.

A Poisson process is a model for a series of discrete events (drought counts) where the average time between events is known, but the exact timing of events(drought each year) is random, Poisson process is a random mechanism that generates the number of events which landing in disjoint intervals are independent random variables with a Poisson distribution. The Poisson process is the model we use for describing randomly occurring events and, by itself, isn't that useful. We need the Poisson distribution to do interesting things like find the probability of a given number of events (probability of drought counts) during the past 10 years. We can use the Poisson distribution to find the probability of observing a number of events over an interval generated by a Poisson process.

Counts of drought in the case of Yucheng city during the past 10 years (1974–1983) was acquired with the following method and steps. The monthly precipitation in each year was calculated according to the observation value of daily precipitation in Yucheng City from 1974 to 1983, because drought and flood were relative to the difference value of atmospheric precipitation required by crops during the growth stage process. If precipitation is used as a characterization index of drought and flood disasters, it is generally believed that droughts (or flood disasters) will be occurred only when precipitation is less than (or more than) a certain threshold (or critical value). According to the sum of monthly precipitation during the past 10 years (1974–1983), the total precipitation in each month was compared, and the threshold of corresponding monthly precipitation was compared and formed to evaluate the droughts risk. The precipitation was lower than this threshold, and we believe that drought will be occurred. In this way, counts of droughts in Yucheng City from 1974 to 1983 is 4，7，4，7，7，4，8，5，6，7，respectively, as shown in **Figure** 1.

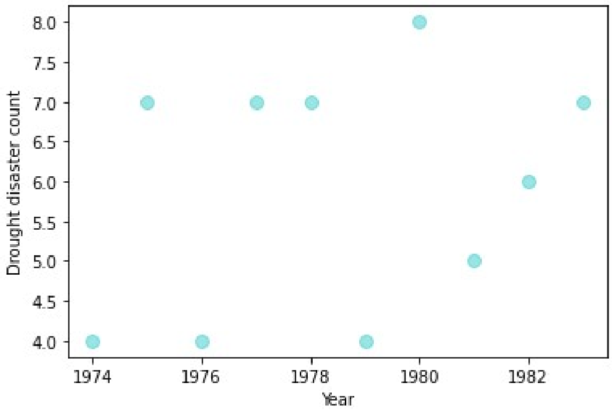

Figure 1. Counts of drought in Yucheng city from 1974 to 1983.

Poisson distribution and Poisson process were used to simulate and achieve the prediction of drought in Yucheng city. Occurrences of droughts in the time series is thought to follow a Poisson process. The occurrence counts of drought in Yucheng City from 1974 to 1983 was predicted, and the Poisson distribution was used to describe the occurrence of drought in Yucheng City in the past 10 years as the realization of Poisson process. We can use Poisson distribution to monitor the numbers of drought in the case region, and we are interested in model flexibility, adaptivity and the drought occurrence counts from Poisson process. Therefore we will build a model to simulate and predict the counts of drought to explore model flexibility using Poisson distribution and Poisson process.

3. RESULTS AND ANALYSIS

The Poisson distribution is widely used to model count data. In the example, we are interested in estimating the posterior distribution of expected counts of drought. Poisson distributions which count of drought is gamma distributed was developed to compare the consistency between prediction and observed data. Poisson distribution is used to describe the probability of drought counts occurring on a fixed time interval, that is, one year. Thus, the Poisson distribution assumes that the drought counts occur independently. This discrete distribution is parametrized using only one value corresponding to the mean and the variance of the distribution (Martin, 2018) or (Xizhi Wu, 2020). Since we are now considering $Gamma(1,1)$ prior distributions of theta, according to the Gamma distribution is conjugate to the Poisson distribution, Hence, the corresponding posterior distribution is $Gamma(1+\sum y_i, 1+10)$, that is $Gamma(60,11)$, so prior and posterior distribution density of theta were created based on R language, as illustrated in **Figure 2**.

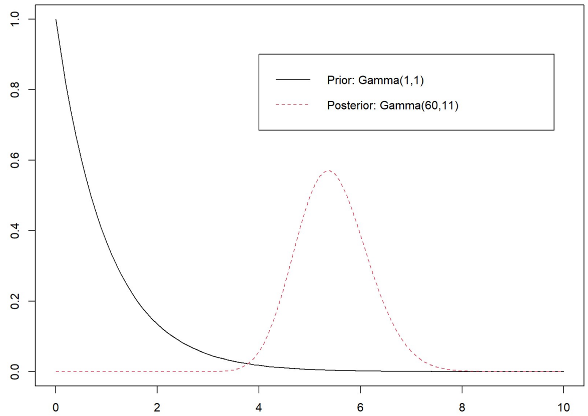

Figure2. $Gamma(1,1)$ prior and $Gamma(60,11)$ posterior of theta. Here, black solid line denotes the prior distribution, red dotted line denotes the posterior distribution.

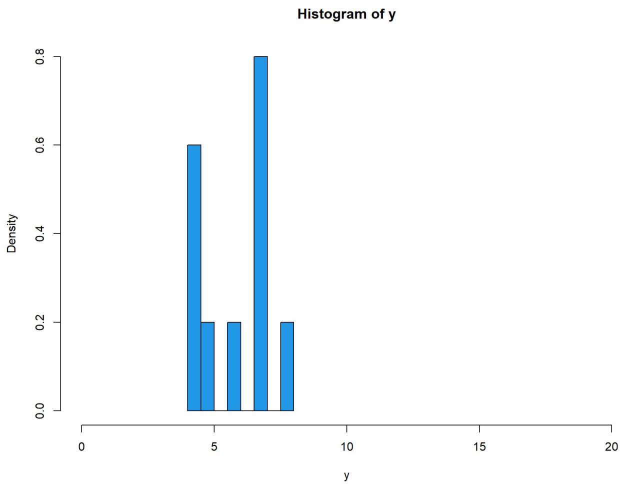

Figure 3. The observed value of theta from the samples.

Random sampling was carried out to get a lot of theta sample values from the posterior distribution of theta, and then a Poisson sample is randomly generated for each value as the predicted value. The observed value of theta, the sample value of theta and the histogram of the predicted value are generated as shown in **Figure 3, 4 and 5**. The samples obtained by metropolis are plotted by us as a histogram (blue color) and the true distribution as the continuous (red color) line, as shown in **Figure 4.**

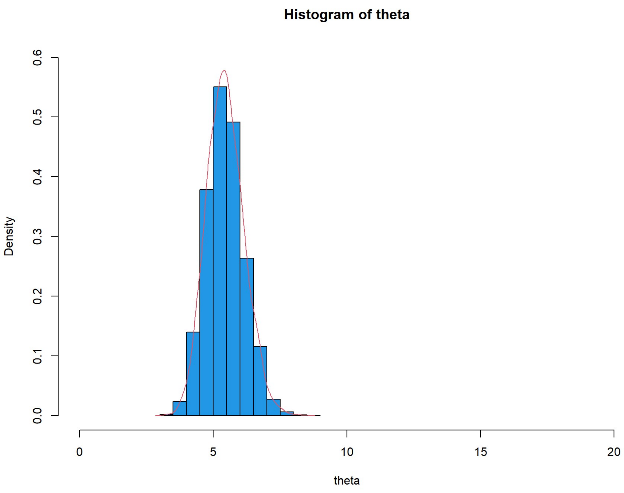

Figure 4. The sample value of theta. Red line: posterior samples of the parameter; blue histogram: posterior predictive distribution.

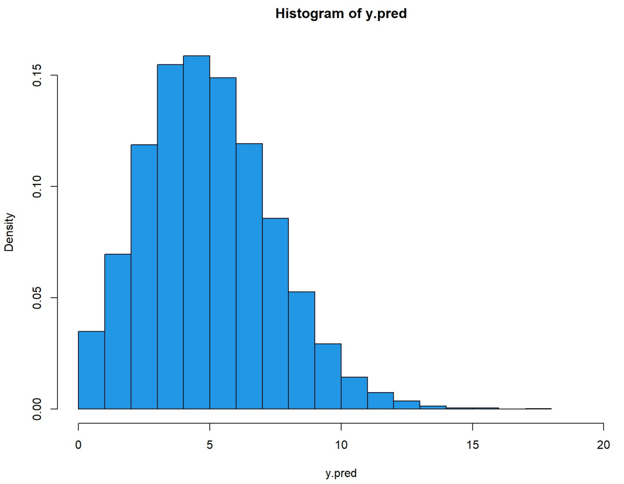

Figure 5. The predicted value of theta from the samples.

Convergence of the Metropolis-Hastings algorithm refers to whether the algorithm has reached its target distribution. If this model has reached is convergence, then the generated sample comes from the correct target distribution. Hence, monitoring the convergence of the Metropolis-Hastings algorithm is essential for producing results from the posterior distribution of drought counts prediction. There are many ways to monitor convergence, such as trace plots, autocorrelation plots, rank plots, the effective sample size (ESS), potential scale reduction factor, Monte Carlo standard error (MCSE). Trace plots which are effective diagnostic samples convergence for metropolis are probably the most popular plots in bayesian literature. To check whether the model makes sense, we diagnose numerical inference according to trace plot to evaluate MCMC methods convergence, a trace plot is made by drawing the sampled values at each iteration step. We will sample using a Metropolis step method, which implements Metropolis-Hastings designed to handle discrete values (Salvatier, Wiecki, Fonnesbeck, 2016) or (Clark, 2007). To explore the results of our inference of Poisson prediction of drought, we are going to generate a trace plot of the posterior distribution. In the trace plot below we can see the plausible values from the posterior of drought prediction. The traceplot itself should look like white noise, or trace plots do not present irregularities, meaning any recognizable pattern should not be seen of theta sample trace plots of drought prediction. The convergence of the chain can be checked visually using trace plots obtained, as shown in **Figure 6**. No patterns or irregularities are observed from the trace plot, therefore convergence can correctly be assumed. Similarly, the results of Poisson sample prediction of drought counts, as illustrated in **Figure 7,** no patterns or irregularities are observed from the plot. **Figure 6**. and **Figure 7** have examples of traces with good mixing, a chain run to 10000 iterations show the paratmeter theta to have 95% interval [4.8, 7.2] correspond to the results.

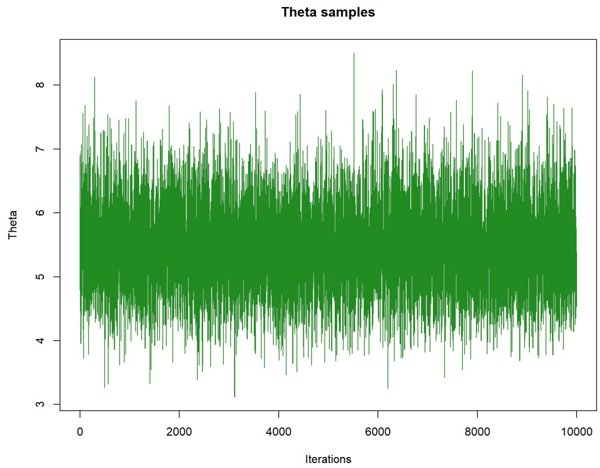

Figure 6. 10000 iterations trace plot of theta.

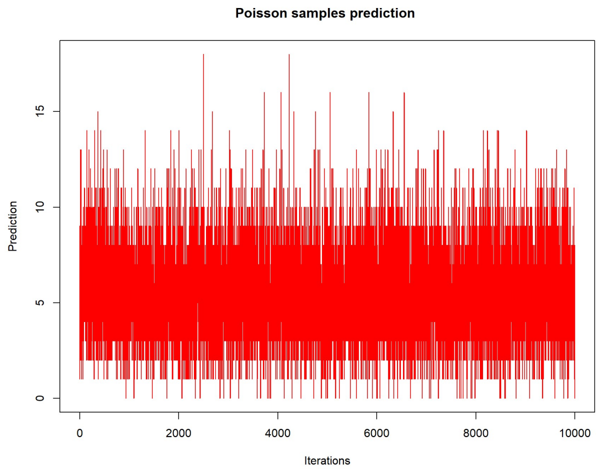

Figure 7. 10000 iterations trace plot of Poisson samples prediction of drought.

A posterior distribution is the result of a bayesian analysis, and includes all the information about the parameters given a dataset and a model. Additionally, the posterior distribution also reflects all that we know about a balance between the prior and the likelihood. Apparently, the posterior is proportional to the likelihood times the prior according to bayes' theorem (Geyer, 2011) or (Lunn, Christopher, Nicky, Andrew, Spiegelhalter, 2013). The posterior summarized and reflected the consequences of a model and data. For each parameter of bayesian model, the mean (or mode or median) is included in the posterior results. Theta mean is equal to 5.446, y.pred(predicted value) mean is equal to 5.447, theta standard deviation (sd) is 0.733, yet y.pred. standard deviation is 2.565 from Poisson model.

There is a very intuitive interpretation about Highest-Posterior Density (HPD), frequentist confidence intervals are different from bayesian credible intervals, the spread of a posterior distribution is summarized by a HPD interval. An HPD is the shortest interval containing a given portion of the probability density. One of the most commonly-used is the 95% HPD, often accompanied by the 50% HPD. The 95% HPD which is default value from R platform, 10000 iterations trace plot of Poisson samples prediction of drought is [4.8-7.2] in the case study, this means that the parameter in question is between 4.8 and 7.2 with a probability of 95% according to our data and model. (Osvaldo, Martin, Ravin Kumar, Junpeng Lao, 2022) or (Martin, 2018). As illustrated in **Figure 8**. We can see the mean and the HPD, on top of a curve representing the entire posterior distribution. The 95% HPD for Poisson process of drought disasters in Yucheng city is [0.76-0.92](green bar), we think that the parameter theta is between 0.76 and 0.92 with a probability of 95% according to our data and model. The 99% HPD for Poisson process of drought disasters in Yucheng city is [0.73-0.94] (red bar), this means that the parameter theta is between 0.73 and 0.94 with a probability of 99% according to our model and data. Identically, the 50% HPD for Poisson process of drought disasters in Yucheng city is [0.83-0.88] (blue bar), we think that think the parameter theta is between 0.83 and 0.88 with a probability of 50% according to our data and model.

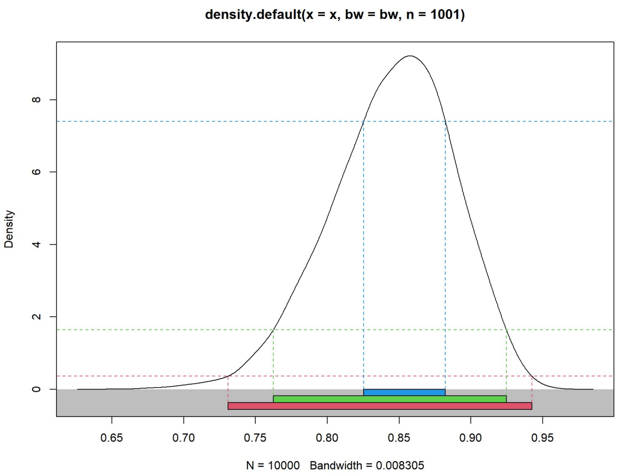

Figure 8. 50%，95%，99% highest-posterior density of the parameter theta.

What we know about the value of the parameter theta was reflected in the prior distribution before seeing the data. The likelihood is an expression of the plausibility of the data given the parameters. The posterior which is a probability distribution for the parameters in our model and not a single value is the combination of priors and likelihood model which is the result of the bayesian analysis and reflects all that we know about our data and model. We can also think of the posterior as the updated prior in the light of data combined with likelihood. As a prior in the case study, we will use a $Gamma(1,1)$ distribution, which is a common considering prior distributions of theta, our model has more flexibility and can better accommodate the observed mean and HPD of the data, Thus the corresponding posterior distribution is $Gamma(60,11)$, furthermore, random sampling was carried out to get a lot of theta sample values from the posterior distribution of theta, and then a Poisson sample is randomly generated for each value as the predicted value. posterior samples of the parameter and posterior predictive distribution of the total number of successes were performed in the study of drought prediction. we are summarizing the posterior and HPD interval, The 95% HPD for 10000 iterations trace plot of Poisson samples prediction of drought is [4.8-7.2], this means that the parameter in question is between 4.8 and 7.2 with a probability of 95% according to our data and model.

4. CONCLUSIONS AND DISCUSSION

There is a vital need for research that links meteorological drought with random theory. This study addresses that need, drought prediction was implemented by code and then performed by bayesian inference in a fairly automated fashion with the support of open-source R platform. Poisson distribution, as an important part of bayesian probabilistic programming, is the distribution of any kind of “rare” event (here, specifically, “event” is the drought disaster counts in Yucheng city from 1974 to 1983), counts of drought are thought to follow a Poisson distribution, are regarded as the 10 realizations of Poisson process. Application of Poisson process to droughts in Yucheng city from 1974 to 1983 is feasibility to enrich drought predictability theory and drought forecasting methods. Drought itself has randomness, random theory is a reasonable method to develop drought prediction. The counts of droughts prediction was developed to explore the droughts model flexibility based on Poisson process. Poisson distribution is the distribution of any kind of the number of drought in Yucheng city from 1974 to 1983. The Gamma (1,1) of theta is used as the prior distribution, and its posterior distribution is calculated to be Gamma (60,11). The sample value of theta is obtained by random sampling from the posterior distribution of theta, and a Poisson sample is randomly generated for each value as the predicted value. The posterior mean value of theta sampling results, the mean and various quantiles of the predicted value, and the probability estimation of theta and the predicted value greater than the current sample mean are calculated to obtain the Highest-Posterior Density (HPD) region of theta. Studies have shown that our model has more flexibility and can provide posterior mean, Highest-Posterior Density region of estimation parameters, and the drought prediction of Yucheng City is a feasible method on basis of the stochastic Poisson process, thereby providing better reference value for drought disaster reduction. In a word, the proposed methodology which is based on Poisson distribution and Poisson process shows great promise as a predictive and analytical tool for drought prediction.

Reading and reviewing the current machine learning literature, it is important for bayesian statistics which is emerging as a powerful framework to express and understand next-generation deep neural networks. Bayesian inference based on prior and likelihood is about conditioning models to the available data and obtaining posterior distributions, there is the rise of probabilistic programming being a burst of innovation in fitting methods for bayesian models that represent notable improvement over existing MCMC methods (Osvaldo, Martin, Ravin Kumar, Junpeng Lao, 2022) or (Lunn, Thomas, Best, Spiegelhalter, 2003). As we have already mentioned, the Poisson distribution and the Poisson-gamma model are widely used for modelling discrete count data. Actually, problem frequently encountered is that the probability of zero cannot be estimated by the Poisson model, so zero inflated Poisson model (ZIP) which can deal with zero and non-zero have been developed in practice. In the example, we get a zero because drought is not occurred in each year, when counting drought disasters, one option is to get a zero in different years, other option is events from Poisson distribution, thus it is necessary to consider of two options. The ZIP model, it is enough to assume that we have a mixture of two processes (one model by a Poisson distribution with probability, the other giving extra zeros with probability, that is, the model can be fitted using the zero). Maybe the ZIP model can better accommodate the zero and non-zero with probability about the prediction of drought counts.

REFERENCES

Andrew Gelman and Jennifer Hill. 2006: *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge University Press, pp:1-648.

Asad Ellahi, Ibrahim M. Almanjahie, Tajammal Hussain, Muhammad Zaffar Hashmi, Shahla Faisal, jaz Hussain. 2020. Analysis of agricultural and hydrological drought periods by using non-homogeneous Poisson models: Linear intensity function. *J. Atmos. Solar-Terr. Phys.*, 198. doi.org/10.1016/j.jastp.2020.105190.

Bhaduri, M. 2020. On modifications to the Poisson-triggered hidden Markov paradigm through partitioned empirical recurrence rates ratios and its applications to natural hazards monitoring. *Sci Rep*. 10:15889. doi.org/10.1038/s41598-020-72803-z.

Carpenter B., Gelman A., Hoffman M., Lee D., Goodrich B., Betancourt M., Brubaker M. A., J. Guo, P. Li, and Ridell A. 2017. Stan: A probabilistic programming language. *J. STAT. SOFTW.*, 76(1),1-32. doi.org/10.18637/jss.v076.i01.

ChulsangYoo, Daeha Kim, Tae-Woong Kim, Kyu-NamHwang. 2008.Quantification of drought using a rectangular pulses Poisson process model. *J. Hydrol*. 355(1–4),34-48. doi.org/10.1016/j.jhydrol.2008.02.025.

Clark J S. 2007:*Models for Ecological Data An Introduction, Statistical Computation for Environmental Sciences in R:Lab Manual for Models foe Ecological Data*. Princeton University Press.

David Lunn, Christopher Jackson, Nicky Best, Andrew Thomas, David Spiegelhalter., 2013: *The BUGS BOOK-A Practical Introduction to Bayesian Analysis.* CRC Press.

Geyer, C. J., 2011: *Introduction to MCMC. In Handbook of Markov Chain Monte Carlo*. Edited by S. P. Brooks, A. E. Gelman, G. L. Jones, and X. L. Meng. Chapman & Hall/CRC, Boca Raton, FL, pp. 3-48.

Ghahramani, Z. 2015. Probabilistic machine learning and artificial intelligence. *Nature*, 521,452–459.

Guo, H. 2010. Understanding global natural disasters and the role of earth observation. *Int. J. Digit. Earth*. 3, 221–230.

Ho, C.-H. & Bhaduri, M. 2015. On a novel approach to forecast sparse rare events: Applications to Parkfield earthquake prediction. *Nat. Hazards*. 78, 669–679.

Hoffman M. D. and Gelman A. 2014.The No-U-turn sampler: Adaptively setting path lengths in Hamiltonian monte carlo. *J Mach Learn Res*. 15(1),1593–1623.

Jianping Huang, Haipeng Yu, Xiaodan Guan, Guoyin Wang & Ruixia Guo. 2016. Accelerated dryland expansion under climate change, *Nat. Clim. Chang.*, 6,166–171.

Laura Poggio, Alessandro Gimona, Luigi Spezia, Mark J.Brewer. 2016. Bayesian spatial modelling of soil properties and their uncertainty: The example of soil organic matter in Scotland using R-INLA. *Geoderma*. 277,69-82. doi.org/10.1016/j.geoderma.2016.04.026.

Lunn, D.J., Thomas, A., Best, N., Spiegelhalter, D., 2003. WinBUGS — a Bayesian modelling framework: concepts, structure, and extensibility. *Stat Comput.*, 10, 325–337.

Martin O. 2018: *Bayesian Analysis with Python: Introduction to Statistical Modeling and Probabilistic Programming Using PyMC3 and ArviZ*, 2nd Edition. Packt Publishing.

McElreath R. 2017: Rethinking: Statistical Rethinking Course and Book Package. URL https://github. com/rmcelreath/rethinking.

Osvaldo A.Martin, Ravin Kumar, Junpeng Lao. 2022：*Bayesian Modeling and Computation in Python*. CRC Press.

Pei-Yu Wu, Gene Jiing-Yun You, Ming HsiuChan. 2020. Drought analysis framework based on copula and Poisson process with non stationarity. *J. Hydrol*. 9,588. doi.org/10.1016/j.jhydrol.2020.125022.

Qianchuan Mi, Xining Gao, Li Yue. 2022. Application of deep learning method to drought prediction. *J Appl Meteor Sci*, 33(1),104-114. doi.org/10.11898/1001-7313.20220109.

Ravin Kumar, Colin Carroll, Ari Hartikainen, and Osvaldo Martin. 2019.Arviz a unified library for exploratory analysis of bayesian models in python. *JOSS*. 4(33),1143.

Salvatier J, Wiecki TV, Fonnesbeck C. 2016. Probabilistic programming in Python using PyMC3. *PeerJ Computer Science* 2: e55 https://doi.org/10.7717/peerj-cs.55.

Tenenbaum, J. B., Kemp, C., Griths, T. L. & Goodman, N. D. 2011. How to grow a mind: Statistics, structure, and abstraction. *Science*. 331,1279-1285.

Xizhi Wu. 2020: *Bayesian data analysis implemented by R and Python.* Beijing: People's University of China Press.

Zachary, D.S. 2018. The Inverse Poisson Functional for forecasting response time to environmental events and global climate change. *Sci Rep*. 8,11342. doi.org/10.1038/s41598-018-29680-4.