Determination of risk-based for rice farmers using the nonparametric Bayesian method

I Maulidi*, I Syahrini¹, Mahmudi¹, V Apriliani²

¹Department of Mathematics, Universitas Syiah Kuala, Jalan Syech Abdurrauf, Syiah Kuala-Banda Aceh 23111, Indonesia
²Department of Mathematics Education, UIN A-Raniry, Jalan Syech Abdurrauf, Syiah Kuala-Banda Aceh 23111, Indonesia

*Email: ikhsanmaulidi@unsyiah.ac.id

Abstract. The Paddy Farmers Business Insurance or Asuransi Usaha Tani Padi (AUTP) is one of the insurance programs organized by the Indonesian government to improve the welfare of rice farmers in Indonesia. In this study, we formulated a premium estimation model of risk-based AUTP insurance. The method used was the Bayes method with a nonparametric approach. The data used was the AUTP claim data in Aceh province. The determination of the premium has been done by providing a settlement algorithm based on past data for two periods. From the analysis conducted, it was found that the Bayesian premium obtained is still not in accordance with what is currently applied by the government, especially in the area of study.

1. Introduction

Agriculture has an important role in the economic development of a country. As an agrarian country, Indonesia extremely depends on this sector to increase the community economy. The farmers in Indonesia currently face the risk of crop failure due to many factors, such as flood, drought, climate change, extreme weather, and attack of plant-disturbing organisms. These risks cause crop failure highly. In 2015, The Ministry of Agriculture of the Republic of Indonesia issued a rice farmer’s business insurance program known as Asuransi Usaha Tani Padi (AUTP). This program is expected to protect farmers from loss caused by crop failure and supporting the compensation funds for the next investment capital [1].

The premium in an insurance product is an important thing to be evaluated fairly and proportionally with the estimation of claim frequency events. There is much research about determination of premium in agricultural insurance that has been done, including those conducted by [2] and [3]. The Bayesian method is also one of the predicting methods in statistical theory that could be used to estimate the premium insurance. For example, [4] has studied how to determine Bayesian premium parametrically. Moreover, [5] have also developed a method known as Kriging Bayesian Method spatially and nonparametrically to estimate the premium. The others research about the Bayesian analysis method can be seen in [6–8].

The study of AUTP’s premium has also been carried out in [9] using a parametric approach by assuming the data of crop failure follows a normal distribution. Another approach that can be
proposed in the determination of AUTP’s premium is a nonparametric approach using Bayesian analysis. A nonparametric approach is a simple method that could be applied because there is no need for the distribution of data and applicable for small and limited data [10].

Nowadays, the AUTP premium applied by the government is 3% of the benefit provided, which is IDR 6,000,000 for each hectare of claimed land or IDR 180,000. In addition, the government also provides a premium subsidy of 80%, so farmers only need to provide a premium of Rp 36,000 for each hectare of the land insured. One of the objectives of this research is to observe whether the premiums currently applied are in accordance with the claim data so that they can provide suggestions so that the AUTP premium will be better in the future.

2. Materials and methods

The data used in this research is secondary data from Aceh Food Crops Agency in the form of AUTP participant data in 4 districts in Aceh (Aceh Besar, Aceh Utara, Bireuen, and Pidie) on the 2015 fiscal year. This data consists of two period data namely the first planting season period (October 2015-March 2016) and the second planting season period (April 2016-September 2016). Since the analysis that we have done is limited to a discrete random variable, then the random variable of the land area claim is assumed discrete and the number is rounded up.

The process to obtain Bayesian premium non-parametrically is
a. Present the data in a claim frequency table of each farmer group due to claimed land area for every district. By using this table, the table of the probability of claims can be obtained.
b. Calculate the average of claimed land areas for each period (period 1 and period 2).
c. Calculate the probability of a claim event in the area of study.
d. Calculate the probability of a claim event on two periods.
e. Determine the probability of claim in the area of study on two periods season
f. Determine the land area of claim in the area of study in the next period.
g. Estimate the Bayesian Premium.

3. Results and discussion

3.1. Table of Conditional Probability of Claims

Suppose that $X$ is the land area of claims and $\Theta$ is a random variable for regions ($\theta_1$ represents Aceh Besar, $\theta_2$ represents Aceh Utara, $\theta_3$ represents Bireuen, and $\theta_4$ represents Pidie). From the data of claim, a frequency table and probability table for random claim variable ($X$) and the regional risk variable ($\Theta$) is presented in Table 1 below.

| Region ($\Theta$) | Period | Land area claim (hectares) | Total |
|-------------------|--------|-----------------------------|-------|
|                   |        | $x_1 = 0$ | $x_2 = 1$ | $x_3 = 2$ | $x_4 = 3$ | $x_5 = 4$ |       |
| $\theta_1$        | 1      | 52         | 18        | 5          | 4          | 0        | 79    |
|                   | 2      | 68         | 4          | 2          | 2          | 3        |       |
| $\theta_2$        | 1      | 31         | 11         | 8          | 0          | 6        | 56    |
|                   | 2      | 51         | 2          | 2          | 1          | 0        |       |
| $\theta_3$        | 1      | 16         | 5          | 1          | 3          | 1        | 26    |
|                   | 2      | 23         | 0          | 0          | 0          | 3        |       |
| $\theta_4$        | 1      | 36         | 4          | 7          | 4          | 2        | 53    |
|                   | 2      | 46         | 2          | 3          | 0          | 2        |       |
Based on Table 1, it is possible to calculate the probability of claim or crop failure occurring on an area for each study area using equation

\[ P(X = x_j | \Theta = \theta_k) = \frac{n(x_j|\theta_k) + n(x_j|\theta_k)}{2n(\theta_k)}, \quad j = 1,2,3,4,5. \]

\[ k = 1,2,3,4. \]  \hspace{1cm} (1)

where,

\( n(x_j|\theta_k) \) : The frequency of farmers that make a claim for land area \( x_j \) in period 1 in the region \( \theta_k \).

\( n(x_j|\theta_k) \) : The frequency of farmers that make a claim for land area \( x_j \) in period 2 in the region \( \theta_k \).

\( n(\theta_k) \) : The frequency of farmers that make a claim in region \( \theta_k \).

Using equation (1), The probability of land area claims is obtained as follows.

| Region (\( \theta \)) | Land area claim (hectares) | Total |
|----------------------|-----------------------------|-------|
| \( \theta_1 \)       | 0.75949 0.13924 0.04430 0.03798 0.01899 | 1     |
| \( \theta_2 \)       | 0.73214 0.11607 0.08929 0.00893 0.05357 | 1     |
| \( \theta_3 \)       | 0.75000 0.09615 0.01923 0.05769 0.07693 | 1     |
| \( \theta_4 \)       | 0.77358 0.05660 0.09434 0.03774 0.03774 | 1     |

3.2. The Average Claim Area for Each Period

The average area of land that is experiencing crop failure for each period can be calculated using this equation

\[ \bar{x}_i = \frac{\Sigma_{j=1}^{5} \Sigma_{k=1}^{4} x_j \ n(x_j|\theta_k)}{\Sigma_{k=1}^{4} n(\theta_k)}, \quad i = 1,2. \]  \hspace{1cm} (2)

Because the data used in this study only have two planting seasons, there are two values for the average area of land claimed.

\[ \bar{x}_1 = \frac{\Sigma_{j=1}^{5} \Sigma_{k=1}^{4} x_j \ n(x_j|\theta_k)}{\Sigma_{k=1}^{4} n(\theta_k)} = 149 \]

\[ = 214. \]

\[ \bar{x}_1 = 0.69626 \text{ hectares}. \]

By the same way, we obtained \( \bar{x}_2 = 0.29439 \) hectares. In our analysis, \( \bar{x}_1 \) and \( \bar{x}_2 \) should be an integer number, by using the round function, let \( y_1 = [\bar{x}_1] = 1 \) and \( y_2 = [\bar{x}_2] = 0 \). The analysis uses \( y_1 \) and \( y_2 \) to compute the Bayesian premium.

3.3. The Probability of Claim in the area of study

Determining the probability of a claim occurring in the study area, the assumption is that the claims that occur are proportional to the area of land registered and each region has the same characteristics and treatment. The probability of a claim (crop failure) in the study area can be calculated using equation

\[ P(\Theta = \theta_k) = \frac{m(\theta_k)}{\Sigma_{i=1}^{4} m(\theta_i)}, \quad k = 1,2,3,4. \]  \hspace{1cm} (3)

where \( m(\theta_k) \) is The total of land area of all farmers in region \( \theta_k \).
So that it is obtained that:

$$P(\Theta = \theta_1) = \frac{m(\theta_1)}{\sum_{i=0}^{m(\theta_1)}} = \frac{702}{2064} = 0.34030.$$  

Using the same idea, it can be obtained $P(\theta_2) = 0.46020$, $P(\theta_3) = 0.08096$, and $P(\theta_4) = 0.11854$. From the results of the above calculation, the greatest probability of crop failure occurs in the area $\theta_2$ which is Aceh Utara. This is because the most registered land area is in the Aceh Utara region.

3.4. The Probability of Claim on Two Periods Season

This probability is a non-conditional probability of claim events of $y_1$ in period 1 and $y_2$ in period 2 and it is assumed equals for all districts. The Probability of Claim on Two Periods Season can be computed as follows

$$P(y_1, y_2) = \sum_{k=1}^{4} P[(y_1, y_2)|\theta_k]P(\theta = \theta_k) = \sum_{k=1}^{4} P(y_1|\theta_k)P(y_2|\theta_k) P(\theta = \theta_k).$$  

$$P(1,0) = P(1|\theta_1)P(0|\theta_1)P(\theta_1) + P(1|\theta_2)P(0|\theta_2)P(\theta_2) + P(1|\theta_3)P(0|\theta_3)P(\theta_3) + P(1|\theta_4)P(0|\theta_4)P(\theta_4) = 0.08612.$$  

3.5. The Probability of Claim in Area of Study on Two Periods Season

The probability of a claim (crop failure) in the study area ($\Theta = \theta_k$) if it is known that the area of crop failure that occurred in the two previous planting season periods can be determined using equation

$$P[\theta = \theta_k|(x_1, x_2)] = \frac{P[\theta_k(x_1, x_2)]}{P(x_1, x_2)}.$$  

Calculation of the probability of a claim (crop failure) $\theta_1$ if it is known that the chance of crop failure occurred in the previous 2 planting season periods with the average area of crop failure for period 1 is $y_1 = 1$ and period 2 is $y_2 = 0$ as follows:

$$P[\theta_1|(y_1 = 1, y_2 = 0)] = \frac{P[\theta_1 \cap (y_1 = 1, y_2 = 0)]}{P(y_1 = 1, y_2 = 0)}.$$  

With the same way (using eq. (5)), it can be resulted the probability of a claim (crop failure) in the study area as follows.

| $\theta$ | $P[\theta = \theta_k|(y_1 = 1, y_2 = 0)]$ |
|---------|-------------------------------------|
| $\theta_1$ | 0.41786 |
| $\theta_2$ | 0.45408 |
| $\theta_3$ | 0.06779 |
| $\theta_4$ | 0.06027 |

3.6. The Land Area of Claim in Area of Study in the Next Period

The land area of a claim or the amount of crop failure that is expected to occur in the next period for each study area ($\Theta = \theta_k$) does not depend on the average value of the claimed area. The results of the calculation of the claim area in the study area in the next period can be calculated using equation

$$E[X|\theta_1] = \sum_{j=1}^{5} x_j P(x_j|\theta_1)$$
\[= x_1 P(x_1|\theta_1) + x_2 P(x_2|\theta_1) + x_3 P(x_3|\theta_1) + x_4 P(x_4|\theta_1) + x_5 P(x_5|\theta_1)\]
\[= 0 + 0,13924 + 0,0886 + 0,11394 + 0,07596\]
\[= 0,41774 \text{ hectares.}\]

Then using the same equation, the conditional expectation value for each region (risk) is presented in Table 4.

| \(\theta\) | \(E[X|\theta]\) (hectare) |
|----------|--------------------------|
| \(\theta_1\) | 0,41774                  |
| \(\theta_2\) | 0,53572                  |
| \(\theta_3\) | 0,61540                  |
| \(\theta_4\) | 0,50946                  |

From Table 4, it is shown that the smallest area of crop failure that will occur in the next period in Aceh Besar area is 0.41774 hectares and the largest area of crop failure that will occur next period in Bireuen area is 0.61540 hectares.

3.7. Bayesian Premium Estimation
Bayes premium is the amount of claims or crop failures that are expected to occur in the next whole period of the study area (\(\theta\)). Based on the estimated land area in the next period for case 1 can be calculated using equation as follows:

\[E[X|y_1 = 1, y_2 = 0] = \sum_{k=1}^{4} P[\theta_k|(y_1, y_2)] E(X|\theta_k) = P[\theta_1|(y_1, y_2)] E(X|\theta_1)\]

Furthermore, the Bayes premium (hectares) that has been obtained is multiplied by the insurance costs stipulated by the insurance for one planting season of Rp 6,000,000 / ha, so that the Bayes premium is obtained (Rupiah) with amount 2,941,434. The results obtained are still far from what implemented by Indonesia government now that is by applying the premium of the 3% of the benefits insurance which is Rp 180,000. In addition, the Indonesia government also provides a subsidy of 80% of this premium. Therefore, from the data claim on 2015 in Aceh province, the applied premium is still not in accordance and needs to analysis in further study.

4. Conclusions
The Bayesian premium of Asuransi Usaha Tani Padi (AUTP) program has been studied in this research by using a nonparametrically approach. From the data, it is obtained that the applied premium is still not in accordance with the real claim events. In other words, the government gives protection funds greater than the collective premium. The further study could be done by lengthen the time interval of data and analyze it parametrically.

References
[1] Yanuarti R, Aji J M M and Rondhi M 2019 Risk aversion level influence on farmer’s decision to participate in crop insurance: A review Agric. Econ. (Czech Republic) 65 481–9
[2] Barry K G 1994 Premium Rate Determination in the Federal Crop Insurance Program: What Do Averages Have to Say About Risk? J. Agric. Resour. Econ. 19 382–95
[3] J T and Zadeh S V 2007 Determination of Premium Rate in Agricultural Insurance Plans Iran. J. Agric. Econ. 1 83–100
[4] Ozaki V A, Ghosh S K, Goodwin B K and Shirola R 2008 Spatio-temporal modeling of agricultural yield data with an application to pricing crop insurance contracts Am. J. Agric. Econ. 90 951–61
[5] Park E, Wade Brorsen B and Harri A 2019 Using Bayesian kriging for spatial smoothing in crop insurance rating Am. J. Agric. Econ. 101 330–51
[6] Verrall R J 2004 A Bayesian Generalized Linear Model for the Bornhuetter-Ferguson Method of Claims Reserving North Am. Actuar. J. 8 67–89
[7] Migon H S and Moura F A S 2005 Hierarchical Bayesian collective risk model: An application to health insurance Insur. Math. Econ. 36 119–35
[8] Ker A P, Tolhurst T N and Liu Y 2016 Bayesian Estimation of Possibly Similar Yield Densities: Implications for Rating Crop Insurance Contracts Am. J. Agric. Econ. 98 360–82
[9] Mutaqin A K, Kudus A and Karyana Y 2015 Metode Parametrik Untuk Menghitung Premi Asuransi Usahatani Padi (AUTP) di Indonesia Prosiding SNaPP, pp 15–23
[10] Tsybakov A . 2008 Introduction to nonparametric estimation (Springer Science & Business Media)

Acknowledgment
The authors thank to LPPM Syiah Kuala University for financial support of this research through Hibah Penelitian Lektor PNPB 2020.