Determination of Microinsurance Premiums in Cattle Farms Based on Protein Content Index and Milk Production Using the Black-Scholes Method

Fabiyan Rahadi[1*], Dwi Susanti[2], Nurfadhlina Binti Abdul Halim[3]

[1]Mathematics Undergraduate Program, Faculty of Mathematics and Natural Scinces, Universitas Padjadjaran, Indonesia
[2]Department of Mathematics, Faculty of Mathematics and Natural Scinces, Universitas Padjadjaran, Indonesia
[3]Faculty of Science and Technology, Universiti Sains Islam Malaysia, Bandar Baru Nilai 71800, Malaysia

*Corresponding author email: rahadifabiyan@gmail.com

Abstract

Micro insurance is an insurance product that is intended for low-income people with features and administration that are simple, easy to obtain, economical prices and immediate settlement of compensation. Cattle business insurance is an effort to protect breeders in event of a risk of death and loss of cattle, this insurance product is a protection against worries the burden of a large premium. This study aims to calculate the premium price using the Black-Scholes method. In this study, a correlation analysis was conducted to determine the influence the protein content on cow's milk production. Then, modelling the protein content using ARIMA-GARCH and determining the premium using the Black-Scholes method. The results showed that the protein content of milk in cow's milk production has a positive correlation. Protein content follows the ARIMA(1,0,0)-GARCH(1) model. Based on the results of the analysis it can be concluded that the bigger protein content in milk, the higher premium that needs to paid.

Keywords: Micro insurance, insurance premiums, protein, ARIMA-GARCH, Black-Scholes.

1. Introduction

Microinsurance is an insurance product intended for people with low incomes. It has simple features and admin, is easy to get at an economical price and is immediately completed by providing compensation (Eling et al., 2014). The Ministry of Agriculture of the Republic of Indonesia, which also oversees the livestock subsector, 2016 implemented the AUTS (Cattle Livestock Business Insurance) program as a form of the government's partisanship in efforts to protect breeders and the risk of death and/or loss of cattle. The risk of losing cows has implications for financial losses. One solution to minimize the financial risks that occur is insurance coverage for cattle. Cows are an important asset for breeders because the cows they produce are expected to have high nutritional and protein values, so they have high selling power for the cows. Cattle microinsurance is a facility for farmers to transfer the risk of loss of cattle by paying a relatively small premium.

In the livestock sector, to overcome the risk of loss of cattle, this insurance is required to be provided as a welfare guarantee for farmers, where when the cattle they keep die due to poisoning and the risk of loss to the breeder, the farmer can benefit from this insurance. To be able to participate in this insurance, breeders need to prepare well, especially in paying insurance premiums.

This research will determine the price of cattle microinsurance premiums indexed by the protein content of cow's milk and production of cow's milk using the Black-Scholes method, which is suitable for the lower middle-class society which is expected to be input for the management of cattle microinsurance so that it can be implemented directly in the environment society itself.

2. Materials and Methods

2.1. Materials

The data used in this study is secondary data obtained from BBPTU-HPT in Baturaden, Central Java. This data is monthly data on protein levels in cow's milk and the production of cow's milk from January 2012 to December 2019.

2.2. Methods

2.2.1. Correlation Analysis

Correlation analysis aims to determine the linear relationship between two or more variables. Correlation is often used in research because it relates to regression. According to Cohen et al. (2014), the correlation coefficient between the independent variable (X) and the dependent variable (Y) can be expressed in the following equation:

$$r_{XY} = \frac{\sum(X_i - \hat{X})(Y_i - \hat{Y})}{\sqrt{\sum(X_i - \hat{X})^2 \sum(Y_i - \hat{Y})^2}}$$

(1)

2.2.2. ARMA

The AR(p) and MA(q) models can be combined into a model known as the Autoregressive Moving Average (ARMA), so it has the assumption that data influence the current period data in the previous period (Choi, 2012). The ARMA model with order p and q is written ARMA (p,q) or ARIMA (p,0,q), which has the following formulation:

$$Z_t = \mu + \phi_1 Z_{t-1} + \cdots + \phi_p Z_{t-p} + \varepsilon - \theta_1 \varepsilon_{t-1} - \cdots - \theta_p \varepsilon_{t-q}$$

(2)

with

$\phi_1, \phi_2, \dots, \phi_p$	: parameters that describe AR
$\theta_1, \theta_2, \dots, \theta_q$	: parameters that describe MA
Z_t	: dependent variable
μ	: constant
$Z_{t-1}, Z_{t-2}, \dots, Z_{t-p}$	: Independent variable
ε_{t-q}	: the remainder at time t

2.2.3. ARIMA

Box and Jenkins introduced the ARIMA model. This model has a p-order Autoregressive (AR) process, a q-order Moving Average (MA) process, or a combination of both. D-order differentiation is performed if the time series data is non-stationary, even though the AR and MA aspects of the ARIMA model require stationary data. The general form of the ARIMA model (p,d,q) is as follows (Hamjah, 2014):

$$\phi_p(B)(1 - B)^d X_t = \theta_q \varepsilon_t$$

(3)

with,

p	: degree of autoregressive (AR)
d	: difference degree
q	: moving average degrees (MA)
t	: time
X_t	: random change t
ϕ_p	: parameters that describe AR
θ_q	: parameters that describe MA
ε_t	: random residual at the t time which is assumed to be normal stochastically independent
B	: backshift operator

The stages in ARIMA modeling are as follows:

1. Stationary Test, a stationary test, was carried out on data on the protein content of cow's milk for January 2012-December 2019 by looking at the data plots, the ADF test results, and the ACF and PACF plots.

2. Differencing process, differencing the protein content of cow's milk for January 2012-December 2019 so that the data is stationary.

3. Model-identification is based on ACF and PACF plots. After obtaining several alternative models, parameter estimation is carried out to obtain the best model.

2.2.4. GARCH

The ARCH model was developed into the Generalized Autoregressive Conditional Heteroscedasticity (GARCH) model by Bollerslev (1986). The GARCH model is used to avoid larger ARCH orders. The GARCH model is defined by the following equation:

$$\sigma_t^2 = a_0 + a_1\varepsilon_{t-1}{}^2 + \cdots + a_p\varepsilon_{t-p}{}^2 + b_1\sigma_{t-1}{}^2 + \cdots + b_q\sigma_{t-q}{}^2 + \varepsilon_t \tag{4}$$
$$a_0 > 0, a_1, \dots, a_p, b_1, \dots, b_q \geq 0$$

where $a_1, \dots, a_p, b_1, \dots, b_q$ are control parameters.
The stages in ARIMA modeling are as follows:
1. Test the effect of ARCH on the residual data from the ARIMA modeler using the ARCH-LM test. Tests were conducted to determine the presence of ARCH effects.
2. Model-identification is done based on residual data plots. After obtaining several alternative models, parameter estimation is carried out to obtain the best model.
3. Diagnostic tests are performed to ensure that the model does not contain ARCH effects, or white noise, and is normally distributed
4. Forecasting is done to obtain data for the next month.
From the data obtained in stage 4, the standard deviation value is obtained which is then used to calculate insurance premiums.

2.2.5. Determination of Insurance Coverage

The amount of coverage for micro-insurance for cattle is calculated based on the costs incurred by a farmer in one year. These costs include feed, cages, labor costs, health costs, seed costs, waste treatment costs, medicine costs, and operational costs.

2.2.6. Black-Scholes

The method developed in 1973 by Fisher Black and Myron Scholes is used to determine the value of options in a stock price contract. The Black-Scholes model states that by adjusting the proportions in the data, investors can create a riskless hedging portfolio, where market risk is eliminated (Anwar & Andallah, 2018). The Black-Scholes formula is a solution for pricing put options in Europe, the formula can be written as follows:

$$P = Ke^{-rT}N(-d_2) - S_0N(-d_1) \tag{5}$$

with

$$d_1 = \frac{\ln\left(\frac{S_0}{K}\right) + (r + \frac{\sigma_T{}^2}{2})T}{\sigma_T\sqrt{T}} \tag{6}$$

$$d_2 = \frac{\ln\left(\frac{S_0}{K}\right) + (r - \frac{\sigma_T{}^2}{2})T}{\sigma_T\sqrt{T}} = d_1 - \sigma_T\sqrt{T} \tag{7}$$

Information:
P	: European put option price
S_0	: initial share price
K	: option strike price
r	: annual risk-free interest rate
$\sigma_T\sqrt{T}$	: standard deviation of the stock price
T	: time of year
$N(-d_1)$	: normal cumulative density function of d_1
$N(-d_2)$	: normal cumulative density function of d_2

2.2.7. Determination of Insurance Premiums

For the determination of insurance premiums based on this index, the equation can be used:

$$P = Ke^{-rT}N(-d_2)$$ (8)

with

$$d_2 = \frac{\ln\left(\frac{R_0}{R_T}\right) + (r - \frac{\sigma^2}{2})T}{\sigma\sqrt{T}}$$ (9)

when

K	: Price of insurance coverage
r	: Annual risk free interest rate
t	: Time (in years)
$N(-d_2)$	: Standard normal cumulative density function $(-d_2)$
R_0	: High value of initial protein content
R_T	: The benchmark value is high protein content
σ	: The standard deviation of the protein content or the volatility of the protein content

3. Results and Discussion

3.1. Correlation Analysis of Milk Protein Content and Cow's Milk Production

The quarterly correlation value between cow's milk's protein content and cow's milk's yield from January 2012 to December 2019 was calculated with the help of Microsoft Excel 2013, which refers to equation (1) and is written in Table 1.

Table 1: Quarterly correlation analysis between cow's milk yield and protein content

	Correlation Analysis		
Milk	Proteins		
	Quarterly 1	Quarterly 2	Quarterly 3
Quarterly 1	0.065844326	-0.279370916	-0.34036672
Quarterly 2	0.09817706	-0.148983891	-0.268161436
Quarterly 3	0.268233916	0.488455297	0.412373577

Based on the calculation results in Table 1, the strongest correlation value is found in the third quarter, which is 0.412373577. This means that there is a relationship between the yield of cow's milk and the protein content of cow's milk. The more protein in milk, the higher the cow's milk production. The protein content of cow's milk in the third quarter is then used as an index of protein content in calculating the value of the cattle microinsurance premium.

3.2. ARIMA Modeling of GARCH Protein Content

Modeling cow's milk protein content using the ARIMA model was carried out with the help of Eviews 9 software. The data used is monthly data on cow's milk protein content from January 2012 to December 2019. The first step is to test whether the data on cow's milk protein content is stationary. To find out the stationarity of cow's milk protein content, data can be done by looking at the data plots in Figure 1.
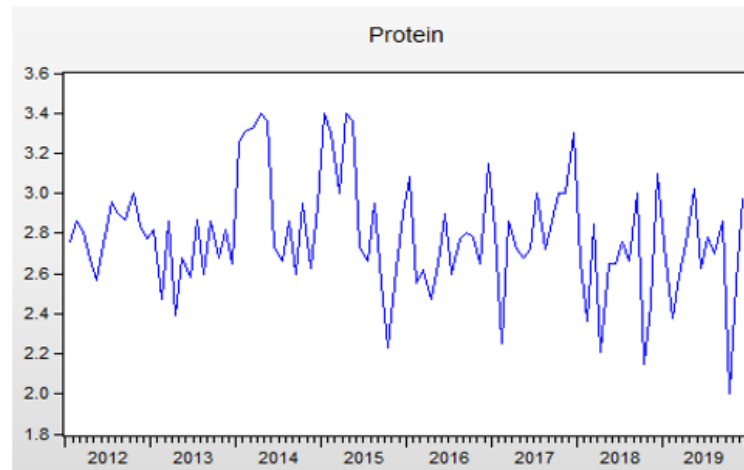
Figure 1: Graph of Cow's Milk Protein Content Data Plot

Figure 1 shows that the data is quite stationary because the time series plot of cow's milk protein content is not stable every month. Furthermore, to further ensure the stationarity of the data, a stationarity test was carried out using the unit root test or the Augmented Dickey-Fuller (ADF) test with the help of Eviews 9. The test results can be seen in Table 2.

Table 2: ADF test for cow's milk protein content

		t-statistic	Ptob.*
Augmented Dickey-Fuller test statistic		-6.903668	0.0000
Test Critical Values:	1%	-3.500669	
	5%	-2.892200	
	10%	-2.583192	

Based on the test results in Table 2, it obtained $|t_{count}|$ of 6.903668, which is greater than $|t_{table}|$ at a significance level of 5%, namely 2.892200. In addition, the probability value on the ADF test is smaller than the significance level $\alpha = 0.05$. This shows that the data on cow's milk protein content is stationary.

After obtaining stationary data, the model can be identified by looking at the ACF and PACF plots, data on the protein content of cow's milk processed using the Eviews software in Figure 2. Based on the figure, the ACF plot is interrupted at lag 1, while the PACF plot is interrupted at lag 1. So that the ARIMA model allows ARIMA(1,0,0), ARIMA(0,0,1), and ARIMA(1,0,1).
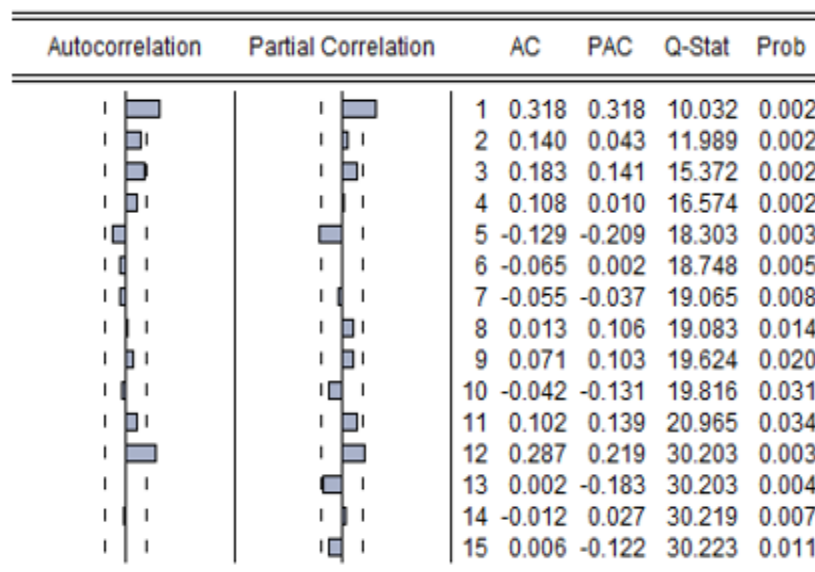

Figure 2: ACF and PACF plots of cow's milk protein content data

After identifying the model and obtaining several alternative models, each model's parameters are estimated. In parameter estimation, a significant model is selected, namely a model that has a p-value less than a significance level of 0.05. Parameter estimation results are given in Table 3.

Table 3: ARIMA Model Parameter Estimation Results

No	Model	AR(p) 1	MA(q) 1	Sig
1	ARIMA(1,0,0)	0.316591		Yes
2	ARIMA(0,0,1)		0.293546	No
3	ARIMA(1,0,1)	0.562198	-0.281259	No

Based on Table 3, a significant model is obtained, namely the ARIMA(1,0,0) model so it can be said that this model is a fairly good model.
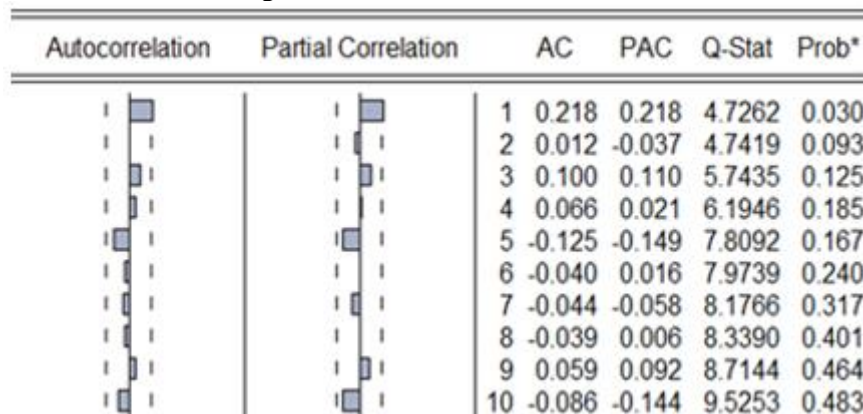
3.3. GARCH Modeling of Protein Content

After obtaining the best ARIMA, the ARCH-LM test was carried out to determine whether the data variance was heteroscedastic. The ARCH-LM test was carried out with the help of Eviews software to obtain the results in Table 4.

Table 4: ARCH-LM Test ARIMA Cow's Milk Protein Content (1,0,0)

Heteroskedasticity Test ARCH			
F-statistic	7.141910	Prob.F(1.93)	0.0089
Obs*R-squared	6.775199	Prob. Chi-Square(1)	0.0092

Based on the table, the Prob value is obtained. Chi-Square is smaller than the 0.05 significance level. So, there is an ARCH effect in the ARIMA (1,0,0) cow's milk protein content model. After the model contains the ARCH effect, the next step is identifying the GARCH model by looking at the ACF, and PACF plots of the best ARIMA modeling squared residual correlation. The plot can be seen in Fig. Based on the plot in Figure, the GARCH(p,q) model is determined with $p = 0.1$ and $q = 0.1$. The estimation results obtained a significant model, namely a model that has a p-value of less than 0.05. The results of the significance test can be seen in Table.

Figure 3: Correlogram of Residual Squared Model of Cow's Milk Protein Content ARIMA(1,0,0)

Table 5: GARCH Parameter Estimation Results

No	Model	Constant	ARCH(1)	GARCH(1)	AIC	Sig
1	ARCH(1)	0.002950		0.970431	0.161589	Yes
2	GARCH(1)	0.076753	0.14705		0.175109	Yes

Based on Table 5, there are two significant models. The smallest AIC value will be selected as the best model. The GARCH(1) model has the smallest AIC value, so it can be said that the model is quite good. Furthermore, diagnostic tests were carried out on this model, including the Heteroscedastic test, White Noise test, and Normality test. The results of the heteroscedastic test can be seen in Table 6.

Table 6: ARCH-LM Test Model of Cow's Milk Protein Content ARIMA(1,0,0)-GARCH(1)

Heteroskedasticity Test ARCH			
F-statistic	0.226140	Prob.F(1.93)	0.6355
Obs*R-squared	0.230443	Prob. Chi-Square(1)	0.6312

Based on the table, the Prob value is obtained. The Chi-Square is 0.6335 which is greater than the significance level of 0.05. So it can be concluded that there is no ARCH effect in the model. Next is the White Noise test.

To determine whether a model meets the white noise assumption, you can look at the ACF and PACF Correlgram plots of residuals provided that no lag exceeds the Bartlet line. The ACF and PACF Correlgram plots of residuals can be seen in Fig. The figure shows that almost all residual lags do not come out of the Bartlet line, and several lags still come out of the Bartlet line, namely lag 7 and lag 12 in the ACF and PACF plots. The residual can still meet the white noise assumption because only a little lag comes out of the bartle line. So it can be concluded that the ARIMA(1,0,0)-GARCH(1) residuals fulfill the white noise assumption, meaning there is no correlation between them.
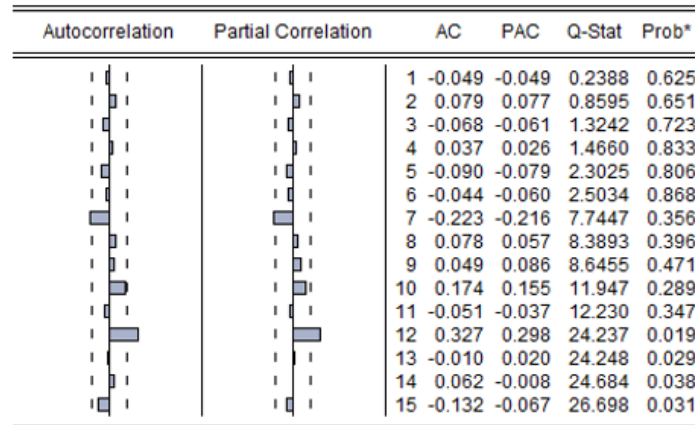
Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob*
		1	-0.049	-0.049	0.2388	0.625
		2	0.079	0.077	0.8595	0.651
		3	-0.068	-0.061	1.3242	0.723
		4	0.037	0.026	1.4660	0.833
		5	-0.090	-0.079	2.3025	0.806
		6	-0.044	-0.060	2.5034	0.868
		7	-0.223	-0.216	7.7447	0.356
		8	0.078	0.057	8.3893	0.396
		9	0.049	0.086	8.6455	0.471
		10	0.174	0.155	11.947	0.289
		11	-0.051	-0.037	12.230	0.347
		12	0.327	0.298	24.237	0.019
		13	-0.010	0.020	24.248	0.029
		14	0.062	-0.008	24.684	0.038
		15	-0.132	-0.067	26.698	0.031

Figure 4: Correlogram of residual ARIMA(1,0,0)-GARCH(1) Model

After the ARIMA(1,0,0)-GARCH(1) residuals fulfill the white noise assumption, the residuals are tested to determine whether the residuals are normally distributed. The figure shows that the residuals are normally distributed because the graph follows the bell curve, and the value of the Jarque-Berra statistic is very large.
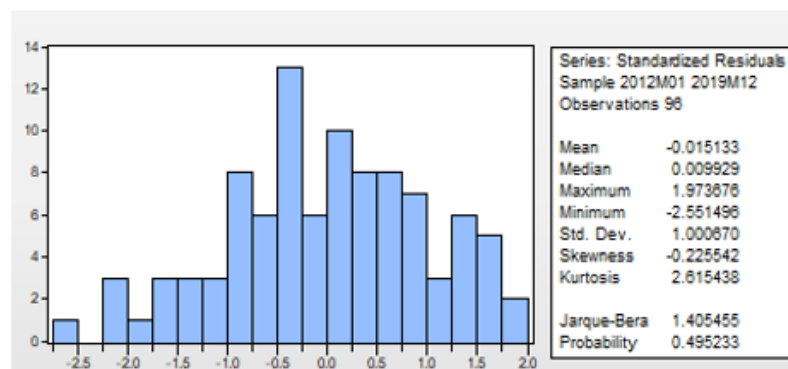

Series: Standardized Residuals	
Sample 2012M01 2019M12	
Observations 96	
Mean	-0.015133
Median	0.009929
Maximum	1.973676
Minimum	-2.551496
Std. Dev.	1.000670
Skewness	-0.225542
Kurtosis	2.615438
Jarque-Bera	1.405455
Probability	0.495233

Figure 5: Residual Histogram ARIMA(1,0,0)-GARCH(1) Model

After the model obtained is stated to be good enough, the model can be used to predict data for future periods. The prediction of the ARIMA(1,0,0)-GARCH(1) model for the next month (97th data) is:

$\hat{Z}_{97} = 2.792338 + 0.351095 Z_{96}$
$\hat{Z}_{97} = 3.8385862$

with the variance equation:

$\hat{\sigma}_{97}^2 = 0.002950 + 0.970431 \alpha_{96}^2$
$\hat{\sigma}_{97}^2 = 0.07088017$
$\hat{\sigma}_{97}^2 = \sqrt{0.07088017} = 0.2662332999 \approx 0.266$

3.4. Determination of the amount of insurance coverage

Based on these costs, the amount insured for a farmer in one year is IDR 3,000,000.

3.5. Determination of Insurance Premiums

Determination of the premium price for cattle microinsurance is calculated for the benchmark value of milk protein content $(R_T) = 5$ percentile, $(R_T) = 10$ percentile, $(R_T) = 15$ percentile, $(R_T) = 20$ percentile, $(R_T) = 25$ percentile, $(R_T) = 30$ percentile, and $(R_T) = 35$ percentile of the selected milk protein content data, namely quarterly data 3. Calculation of the benchmark value of milk protein content was carried out using Microsoft excel 2013 software, the following values were obtained:

$(R_T) = 5$ percentile= 2.19 %, $(R_T) = 10$ percentile= 2,46 %, $(R_T) = 15$ percentile= 2.55 %, $(R_T) = 20$ percentile= 2.57 %, $(R_T) = 25$ percentile= 2.63 %, $(R_T) = 30$ percentile= 2.65 %, $(R_T) = 35$ percentile= 2.68 %.

Before determining the price of cattle microinsurance premiums using equation (7), it is necessary to first calculate the cumulative distribution function d_2 with equation (8), namely:

$$d_2 = \frac{\ln\left(\frac{R_0}{R_T}\right) + \left(r - \frac{\sigma^2}{2}\right)t}{\sigma\sqrt{t}}$$

The initial cow's milk protein content value (R_0) used is the latest cow's milk protein content value, namely $R_0 =$. In contrast, R_T is the benchmark value of the selected cow's milk protein content. The assumption is that the risk-free interest rate is $r = 6.5\%$ and the time period (t) used is 1 year. The standard deviation of the protein content of cow's milk is $\sigma =$ which is obtained from the results of calculating the variance. The calculation of the cumulative distribution value d_2 when $(R_T) = 5$ percentile $= 2.19$ % is as follows:

$$
\begin{aligned}
d_2 &= \frac{\ln\left(\frac{R_0}{R_T}\right) + \left(r - \frac{\sigma^2}{2}\right)t}{\sigma\sqrt{t}} \\
&= \frac{\ln\left(\frac{2.98}{2.19}\right) + \left(0.065 - \frac{0.266}{2}\right)1}{0.266\sqrt{1}} \\
&= \frac{\ln(1.36) + \left(0.065 - \frac{0.266}{2}\right)}{0.266} \\
&= \frac{0.3080217567 - 0.090798081}{0.266} \\
&= \frac{0.2400217567}{0.266} = 0.9023374312 \approx 0.9
\end{aligned}
$$

$N(-d_2) = N(-0.9) = 0.1841$

If the cumulative distribution value d_2 has been obtained, then the cattle microinsurance premium can be calculated using equation (8). The calculation of the premium that must be paid when the benchmark value of cow's milk protein content is 2.19% is:

$$
\begin{aligned}
Premi &= Ke^{-rT}N(-d_2) \\
&= (\text{IDR } 3{,}000{,}000)(e^{-0.065(1)})(0.1841) \\
&= (\text{IDR } 3{,}000{,}000)(0.9370674634)(0.1841) \\
&= \text{IDR } 517{,}542.36
\end{aligned}
$$

So, the premium must be paid when the benchmark value of cow's milk protein content is 2.19% IDR 517,542.36 in one year. The premium price that must be paid with the benchmark value of other cow's milk protein content can be seen in Table 7.

Table 7: Premiums to be paid in 1 year

Percentile	Protein Content (%)	Coverage	Premium
5	2.19	IDR 3,000,000	IDR 517,542.36
10	2.46	IDR 3,000,000	IDR 983,882.324
15	2.55	IDR 3,000,000	IDR 1,160,745.46
20	2.57	IDR 3,000,000	IDR 1,171,709.15
25	2.63	IDR 3,000,000	IDR 1,282,470.53
30	2.65	IDR 3,000,000	IDR 1,293,715.33
35	2.68	IDR 3,000,000	IDR 1,349,658.26

The table shows the amount of cattle microinsurance premiums that need to be paid by farmers in one year. It can be seen that the higher the protein content of cow's milk, the higher the premium payment will be. It can be seen that when the protein content is 2.19%, the premium payment is IDR 517,542.36. Likewise, when the protein content is

2.68%, the premium payment is IDR 1,349,658.26. This happened because the protein content of cow's milk was chosen as the benchmark value, namely the protein content of cow's milk in the third quarter, which increased cow's milk production.

4. Conclussion

The value of the cattle microinsurance premium can be determined based on milk protein content data for 2009-2019 which follows the ARIMA(1,0,0)-GARCH(1) time series model. Then calculated using the Black-Scholes method. The calculation results show that the higher the protein content contained in milk, the greater the premium payment.

References

Anwar, M. N., & Andallah, L. S. (2018). A study on numerical solution of Black-Scholes model. *Journal of Mathematical Finance*, *8*(2), 372-381.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics*, *31*(3), 307-327.

Choi, B. (2012). *ARMA model identification*. Springer Science & Business Media.

Cohen, P., West, S. G., & Aiken, L. S. (2014). *Applied multiple regression/correlation analysis for the behavioral sciences*. Psychology press.

Eling, M., Pradhan, S., & Schmit, J. T. (2014). The determinants of microinsurance demand. *The Geneva Papers on Risk and Insurance-Issues and Practice*, *39*(2), 224-263.

Hamjah, M. A. (2014). Forecasting major fruit crops productions in Bangladesh using Box-Jenkins ARIMA model. *J. Econ. Sustain. Dev*, *5*(7).