

Prediction System of Police Officer Admissions Number using Analysis of Variance Method

TTA Putri¹, HD Hutahaean², RD Sari³, R Rahmadani⁴
PTIK-FT, Universitas Negeri Medan, Indonesia^{1,2,3,4}

tansatrisna@unimed.ac.id

Abstract: Police officer admission attracts people to be able to contribute for their country. There could be many youths who want to participate on it. Based on that number, we believe that a prediction system about the admission of police officer is important to notice a decent candidate. This research generated a prediction system to find out how many police officers are received at each year. The object of this research is the North Sumatra Police Chief. This study used the ANOVA method, that is a two-way ANOVA procedure which is classified using statistical product and service solutions (SPSS) version 16 to determine the admission of prospective police officers from year to year. Supporting factors in the recruitment of prospective police officers is the ability of intelligence candidates to be seen from year to year.

Keyword: police officer admissions, prediction system

1. Introduction

Becoming a police officer is a very noble job because it serves as a servant of the state, maintaining national security, regulating the rules and regulations. Becoming a police officer is not easy because besides having to have a healthy and healthy body. The police must also be smart and agile, because they must be prepared in all conditions even as a threatening emergency.

Prediction is a process of estimating systematically about something that is most likely to occur in the future based on past and present information that is owned, so that errors (the difference between something that happens with the estimated results) can be minimized. Prediction does not have to provide a definite answer to events that will occur, but rather try to find answers as close as possible to happen. In predicting a problem, especially the number of admissions by police officers is quite difficult where the amount of revenue will certainly be difficult to determine.

The increasing interest of the people of Indonesia in wanting to become a police officer, it is very difficult to determine the number of police personnel received each year, because the number of police personnel every year has decreased and has also increased. In determining the number of reception of members of the Police there are several factors that are very influential in the reception of members of the Police for each year, namely, the criminal factor the higher the number of crimes that occur in North Sumatra, the number of reception of members of the Police that will be needed is also higher but if the number of criminals in North Sumatra is low, the number of admissions required by police is not large. As for other factors that affect the

recruitment of the police officer, such as the retirement factor of the police and regional expansion factors.

Therefore, the writer wants to predict the number of police officers received to find out how many people are needed to become a police officer that will be received every year. To predict the amount of acceptance, the writer uses a computer and software as the most efficient tool to examine the prediction system.

Analysis of Variance (ANOVA) is a method for breaking down total diversity of data into components that measure various sources of diversity. ANOVA is used if there are more than two variables. In Indonesian literature this method is known by various other names, such as analysis of variance, variance, and Analysis of Variance. Anova is one of the statistical analyzes that is often used to check the difference in mean values of two or more data groups. In general there are two applications of ANOVA, namely hypothesis testing and estimation [2].

In this ANOVA analysis a Two Way ANOVA procedure is often used which is called the design of a factor that is one of the Two Way ANOVA analysis tools. Anova is a method of testing hypotheses and predictions whose results can be interpreted. By using ANOVA, whether there is a significant difference between the predicted value and observational data.

Prediction

Prediction is a process of estimating systematically about something that is most likely to occur in the future based on past and present information that is owned, so that the error (the difference between something that happens with the estimated results) can be minimized. Predictions do not have to provide answers with certainty of events that will occur, but rather try to find answers as close as possible to happen. Many parties use the test for predictive purposes in order to screen and select candidates who have the potential to succeed in education and other jobs. [1]

Prediction Technique

Based on the technique used to predict the prediction can be divided into two parts, namely qualitative prediction and quantitative prediction.

Qualitative Prediction

Qualitative predictions are based on qualitative data from the past. Qualitative methods are used if past data from the variable to be predicted is absent, insufficient or less reliable. The results of predictions made are very dependent on the individuals who compose them. This is important because the prediction results are determined based on judgmental thoughts or opinions, knowledge and experiences of the constituents. Therefore this qualitative method is also called judgmental, subjective, and intuitive.

Quantitative Prediction

Quantitative predictions are based on quantitative data from the past. The results of predictions made depend on the method used in the prediction. With different methods different predictions will be obtained. The thing to note from the use of these methods is whether or not the method used and is determined by the deviation between the predicted results with the reality that occurs. A good method is a method that gives the values of differences or possible deviations. Quantitative prediction can only be used if there are three conditions as follows:

1. There is information about other circumstances.
2. The information can be quantified in the form of data.
3. It can be assumed that past patterns will be sustainable in the future.

ANOVA Method

Analysis Of Variance (ANOVA) is a method for breaking down total diversity of data into components that measure various sources of diversity. ANOVA is used if there are more than two variables. In Indonesian literature this method is known by various other names, such as analysis of variance, variance, and Analysis Of Variance. Anova is one of the statistical analyzes that is often used to check the difference in mean values of two or more data groups. In general there are two applications of ANOVA, namely hypothesis testing and estimation.[2][5]

Two Ways ANOVA

The concept of variance analysis is based on the concept of the F distribution and can usually be applied to various types of cases as well as in the analysis of the relationships between the various observed variables. In statistical calculations, analysis of variance is greatly influenced by the assumptions used such as the normality of the distribution, the homogeneity of variance and freedom from error.[3]

The normality assumption of distribution provides an explanation of the data characteristics of each group. The assumption of homogeneity of variance explains that the variance in each group is considered the same. While the free assumptions explain that each variance to the average in each group is mutually independent. If the t test is used for testing two samples, then the F test or ANOVA Hoyt is used for testing more than two samples.[4]

The assumptions used in the Anova Hoyt test are:

1. The populations to be tested are normally distributed.
2. The variance of these populations is the same.
3. Samples are not related to each other.

The steps for completing Two Way ANOVA are:

1. Hypothesis
 - a. H_0 : All treatments (columns, rows, interactions) that have the same average value.
 - b. H_1 : There are treatments (columns, rows, interactions) that have averages that are not of equal value (different).
2. Determine the significant level of α
3. Two Way Anova Table

Table 1 Two Way ANOVA

Source of Uniformity (SK)	Sum of Squares (JK)	Free Degrees (db)	Middle Squared (KT)	f count	f Table
Row Average	JKB	db numer1= r - 1	S ² B = KTB= $\frac{JKB}{r - 1}$	f count = $\frac{KTB}{KTG}$	db numer1= db denum = f tabel =
Column Average	JKK	db numer2= k - 1	S ² K = KTK= $\frac{JKK}{r - 1}$	f count = $\frac{KTK}{KTG}$	db numer2= db denum = f tabel =
Error	JKG	db denum= (r - 1).(k - 1)	S ² G = KTG= $\frac{JKG}{(r - 1)(k - 1)}$		
Total	JKT	r.k - 1			

source : Thomas Yunigunarto,Uji Anova,2009

4. Settlement of JKT, JKK, JKB, JKG

a) Sum of total squares

$$JKT = \sum_{i=1}^r \sum_{j=1}^k x_{ij}^2 - \frac{T_{**}^2}{rk}$$

b) Number of Column Squares

$$JKK = \sum_{i=1}^k \frac{T_{**}^2}{r} - \frac{T_{**}^2}{rk}$$

c) Number of Row Squares

$$JKB = \sum_{i=1}^r \frac{T_{i*}^2}{k} - \frac{T_{**}^2}{rk}$$

d) Number of Error Square

$$JKG = JKT - JKB - JKK$$

Where :

- k: The number of columns
- r: Many lines
- x_{ij}: Data in the i-th row, j-th column
- Q * h: Total (number) of the jth column
- r: Number of lines / blocks
- Ti *: Total (number) i-th row
- Q **: Total (number) of all observations

5. Determine the testing criteria by comparing F_{count} with F_{table}

- a. H₀ is accepted when F_{count} ≤ F_{table}
- b. H₁ is accepted when F_{count} > F_{table}

6. Decision

The decision is the final result of a study obtained from H₀ and H₁.

2. Analysis, Results and Discussion

Becoming a police officer is a very noble job because it serves as a servant of the state, maintaining national security, regulating the rules and regulations. Becoming a police officer is not easy because besides having to have a healthy and healthy body. The police must also be smart and agile, because they must be prepared in all conditions even as a threatening emergency. A prospective police officer, before being accepted must undergo a series of tests. In official terms, the entire selection process to become a member of the Police from the local level to the central level is free of charge.

Prediction is a process of estimating systematically about something that is most likely to occur in the future based on past and present information that is owned, so that the error (the difference between something that occurs with the estimated results) can be minimized. Prediction does not have to provide a definite answer to events that will occur, but rather try to find answers as close as possible to happen.

The number of police officers received from year to year could not be ascertained how many. Sometimes it has increased and sometimes it has decreased. This makes someone who wants to register, pessimistic about the opportunity to be accepted as a member of the police force so that predictions need to be made to find out the amount needed by the Regional Police in the coming year.

With the help of SPSS software version 16, it is hoped that it can help the public, especially for those who want to try the selection of Police acceptance, so that they can predict the number of police officers based on data from year to year.

In general, the total number of police force admissions is influenced by various factors that shape it. Some of these factors are factors that affect the number of recruits from the police force, increasing or decreasing. These factors include:

- a) Factors of Crime
- b) The factor of criminality can affect the number of recruits from the police force. The lack of security personnel from the police to combat the increasing and rampant crime, becomes a factor that requires the Regional Police to add new police officers so that new police officers can help other police tasks to deal with criminal actions that are often happens in Indonesia, especially in the area of North Sumatra.
- c) Factors of Police Retirement
- d) The total number of recruitment of police officers so far has also aimed to replace personnel who will retire. Therefore the regional police need a new police officer to replace the retired police personnel.
- e) Regional expansion

Because of the vast territory of an area so it is difficult to develop an area, in order to achieve a civilized and organized society that requires a new security, such as the addition of police personnel so that the area is safe and secure. That is why the Regional Police open receipts every year.

Completion Analysis with the Analysis of Variance (ANOVA) Method

Analysis Of Variance (ANOVA) is a method for breaking down total diversity of data into components that measure various sources of diversity. ANOVA is used if there are more than two variables. In Indonesian literature this method is known by various other names, such as analysis of variance, variance, and Analysis Of Variance. ANOVA is a development of the Behrens-Fisher problem, so the F-test is also used in making decision. Analysis Variance was

first introduced by Sir Ronald Fisher, the father of modern statistics. In practice, analysis of variance can be a hypothesis test (more often used) or estimation (estimation, especially in the field of applied genetics). In general, analysis of variance tests two variances (or variances) based on the null hypothesis that the two variances are the same. The first variance is the variance between samples and the second variance is the variance within each sample (within samples). With this kind of idea, ANOVA with two examples will give the same results with the t-test for two mean (mean).

ANOVA is relatively easy to modify and can be developed for a variety of more complex forms of experiments. If you take the step of testing the average difference one by one (with t test) will take time, a lot of effort. In addition, there is a great risk of being wrong. For that, it has been found a way of analysis that contains smaller errors and can save time and energy, namely by ANOVA (Analysis Of Variances) basically the sample patterns can be grouped into:

- a) All samples, both those in the first group to those in other groups, come from the same population. For this condition the null hypothesis is limited to no treatment effect.
- b) The samples in Group One come from different populations from the sample populations in other groups. For this condition the null hypothesis can be read: there is no treatment effect between groups.

Table 2. Statistics on the number of police officers every year has increased and decreased.

Academic Year	Gender		Number of Admission
	Male	Female	
2011/2012	64	6	70
2012/2013	241	0	241
2013/2014	879	16	895
2014/2015	517	278	795
Total	1701	300	2001

Implementation of the Anova Method

- 1) Anova = Test difference > 2 mean
- 2) It is not recommended to use alpha t inflation test
- 3) 1- (1-a)
- 4) Recommended test f

Settlement of received amount:

- 1) Hypothesis
 - a. H_0 = four academic years have the same level of difficulty.
 - b. H_1 = At least one is not the same.
- 2) $\alpha = 0.05$
- 3) The Two Way Anova table
- 4) Settlement of JKT, JKB, JKK and JKG:

a) Sum of total squares

$$\begin{aligned}
 \text{JKT} &= \sum_{i=1}^r \sum_{j=1}^k x_{ij}^2 - \frac{T^{**2}}{rk} \\
 &= (64^2 + 241^2 + 879^2 + 517^2 + 6^2 + 0^2 + 16^2 + 278^2) - \frac{2001^2}{8} \\
 &= 1179683 - 500500.125 = 679182.875
 \end{aligned}$$

b) Number of Column Squares

$$\begin{aligned}
 JKB &= \sum_{i=1}^r \frac{T_{i*}^2}{r} - \frac{T_{**}^2}{rk} \\
 &= \frac{70^2 + 241^2 + 895^2 + 795^2}{4} - \frac{2001^2}{8} = \frac{1496031}{4} - \frac{4004001}{8} \\
 &= 374007.75 - 500500.13 = -126492.38
 \end{aligned}$$

c) Number of Row Squares

$$\begin{aligned}
 JKK &= \sum_{i=i}^k \frac{T_{*j}^2}{k} - \frac{T_{**}^2}{rk} \\
 &= \frac{1701^2 + 300^2}{2} - \frac{2001^2}{8} = \frac{2983401}{2} - \frac{4004001}{8} \\
 &= 1491700.5 - 500500.125 = 991200.375
 \end{aligned}$$

d) Number of Error Square

$$\begin{aligned}
 JKG &= JKT - JKB - JKK \\
 &= 679182.875 - (-126493.13) - 991200.375 \\
 &= -185524.37
 \end{aligned}$$

Source of Uniformity (SK)	Sum of Squares (JK)	Free Degrees (db)	Middle Squared (KT)	f count	f table
Row Average	JKB = -126493.13	db numer1= r-1 = 4-1 = 3	s ² B= KTB = $\frac{JKB}{r-1}$ = -42164.377	f count = $\frac{KTB}{KTG}$ = 0.6818141	α = 0.05 db numer1=3 db denum= 3 f tabel =
Column Average	JKK= 991200.375	db numer2 =k-1 =2-1 =1	s ² K=KTK = $\frac{JKK}{k-1}$ = 991200.375	f count = $\frac{KTK}{KTG}$ = -16.028091	α = 0.05 db numer2=1 db denum = 3 f tabel =
Error	JKG = -185524.37	db denumer =(r-1)(k-1) =3.1=3	s ² G=KTG = $\frac{JKG}{(r-1)(k-1)}$ =-61841.457		
Total	JKT = 679182.875	r.k-1 = 4*2-1 =7			

Conclusion of Analysis

- a. Reject H_0' , and conclude that the amount of revenue each year experiences unequal difficulties.
- b. Receive H_0'' , and conclude that the four school years have an unequal number of members of the Police force.

To make it easier to analyze the process, flowcharts are used on the running system. The flowchart used can be seen in Figure 1 below:

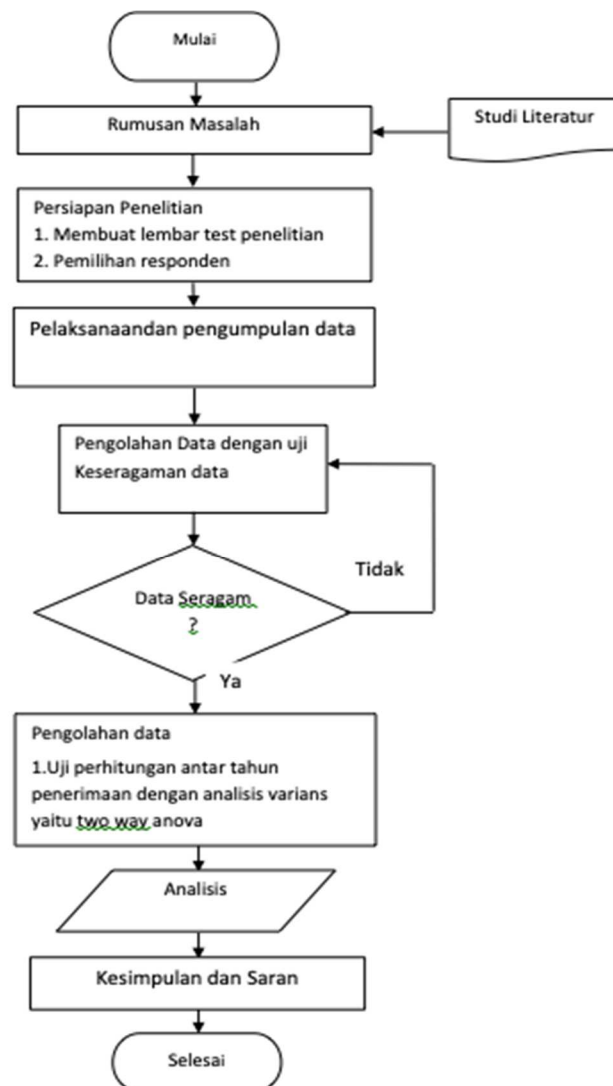


Fig 1 Process Analisis

3. Conclusion

Based on the previous description, the author can provide the following conclusions:

1. Crime is the most dominant factor in determining the number of admissions of police officers in addition to the police retirement period
2. The implementation of the ANOVA method in predicting the number of police personnel every year tends to increase.
3. Two way ANOVA procedure can help to compare the amount of revenue from year to year.

4. References

- [1] Kerlinger, Fred, N. *Behavioral Research Principle (3rd ed)*. (Landung R. Simatupang. Translated). Yogyakarta: Gajah Mada University Press. (Original book published on 1986). (1996)
- [2] Breitsohl, H. Beyond ANOVA: An introduction to structural equation models for experimental designs. *Organizational Research Methods*, 22(3), 649-677. (2019)
- [3] Klasson, K. T. Two-way ANOVA for Unbalanced Data: The Spreadsheet Way. *USDA-ARS Research Notes*. (2019)
- [4] Hoyt, C. L., Burnette, J. L., Auster-Gussman, L., Blodorn, A., & Major, B. *The obesity stigma asymmetry model: The indirect and divergent effects of blame and changeability beliefs on antifat prejudice*. *Stigma and Health*, 2(1), 53. (2017)
- [5] Herzog, M. H., Francis, G., & Clarke, A. ANOVA. In *Understanding Statistics and Experimental Design* (pp. 67-82). Springer, Cham. (2019)