ISSN: 2583-8288
Case Study
Published on: 24.08.2024

‘Z’ And 'T'-Tests: An Application in Analysis of Recovery Time And Viral Load In Parvo-Viral Infected Dogs - A Case Report

Jatin Gupta[1*], Puneet Goyal[2], Deepak[3], Tanuj Jain[4], Devendra Kumar Gupta[5], Mohan Singh Thakur[6], Asit Jain[7]

[1]Department of Veterinary Medicine, College of Veterinary Science and Animal Husbandry, NDVSU, Jabalpur, Madhya Pradesh.

[2]Department of Animal Genetics and Breeding, College of Veterinary Science and Animal Husbandry, NDVSU, Jabalpur, Madhya Pradesh.

Abstract: In veterinary research, statistical analysis is crucial for evaluating the effectiveness of treatments and understanding disease dynamics. This study employs Z-tests and t-tests to analyse recovery times and viral loads in dogs infected with Canine Parvovirus (CPV). Data were collected from 68 dogs divided into two groups, with one group receiving standard treatment and the other an experimental treatment. Additionally, viral load data were obtained from 20 dogs categorized into symptomatic and asymptomatic groups. The Z-test results indicate a statistically significant difference in recovery times, with the experimental group showing a significantly shorter recovery period. Similarly, the t-test reveals a significant difference in viral loads between symptomatic and asymptomatic dogs. These findings suggest that the experimental treatment is more effective in reducing recovery time for CPV-infected dogs and that viral loads vary significantly based on symptomatic status. This research underscores the importance of using robust statistical methods to enhance the reliability of veterinary studies and improve treatment strategies for CPV.

Keywords: Canine Parvo Virus (CPV), Z-test, t-test, Viral loads, Veterinary diagnostics

1. Introduction

In the field of veterinary medicine, statistical data analysis and scientific research are pivotal in ensuring the well-being of animals. Through these methods, we make beneficial comparisons and informed decisions based on objective evidence, ultimately improving animal health outcomes. Statistical analysis is fundamental to this process, as it helps establish the significance of differences within populations, a process known as the test of significance (Abebe, 2019).

Significance in this context refers to the meaningful differences between two samples from a population or between data derived from different populations. Observed differences fall into two categories: significant differences, which are not influenced by chance, and non-significant differences, which are influenced by chance. Understanding and identifying these differences are essential for drawing valid conclusions and advancing veterinary practices (Singh et al., 2017).

There are two primary tests within the framework of the test of significance: the Z-test and the t-test. These parametric tests, also known as classic or standard tests, are instrumental in comparing groups and analysing the significance of observed differences (Kim, 2015). By applying these tests, researchers can determine whether the differences observed in their studies are statistically significant or if they could have occurred by random chance (Abebe, 2019).

This paper aims to delve into the application of the Z-test and t-test in veterinary research, providing a comprehensive overview of their methodologies, assumptions, and practical implications. Through a detailed examination of these statistical tools, we seek to enhance the rigor and reliability of research findings in the field of veterinary medicine (Singh et al., 2017).

Additionally, we will explore how these statistical methods contribute to advancements in animal genetics and population genetics, building on established principles and expanding our understanding of animal health (Kanakraj, 2018; Thiagarajan, 2014).

2. Applications and formulas of Z and t-test

The application of statistical tests in scientific research has increased dramatically in recent years. However, there is an often confusion regarding the appropriate choice of test statistic for specific problems. Thus, the primary objective of this study is to identify the appropriate statistical test, which may depend on the type of data (continuous or categorical). For example:

1. If two samples are independent, independent Z ($n > 30$) and t ($n < 30$) tests are used.
2. If two samples are dependent, paired Z ($n > 30$) and t ($n < 30$) tests are used.

2.2.2.1. Z-test

A Z-test is a statistical method used to test a hypothesis when the population variance is known. It is utilized to compare the sample mean to the population mean. If the sample size is large (n

> 30), the Z-test can still be applied without knowing the population variance.

Conditions for using the Z-test:

1. Data must be normally distributed.
2. Data points must be independent.
3. Variance must be equal for each sample.

Assumptions for the Z-test:

1. The population has a normal distribution.
2. The sample size is more than 30 ($n > 30$).
3. The sample is selected at random.
4. The population standard deviation is known.
5. If samples are compared, their sizes should not vary widely.
6. The variances of the samples or populations compared should be approximately the same.

2.1.1. Formulas for Z-test in different conditions

In statistical analysis, the Z-test is applied under different conditions to assess significance.
Condition 1: Involves comparing the sample mean with the population mean. The purpose is to test the significance of the sample mean using the formula (Eq.1):

$$z = \frac{|\bar{x}-\mu|}{\sigma \sqrt{n}} \quad (1)$$

Condition 2: Focuses on comparing the means of two samples. This tests the significance of numb the difference between the two sample means and is calculated as (Eq.2):

$$z = \frac{|\bar{x}_1-\bar{x}_2|}{\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}} \quad (2)$$

Condition 3: Addresses the comparison of a sample proportion with a population proportion, which tests the significance of a single proportion. The formula used is (Eq.3):

$$z_0 = \frac{p-P}{\sqrt{\frac{PQ}{n}}} \quad (3)$$

Condition 4: Deals with comparing two sample proportions to test the significance of the difference between them. This is calculated using (Eq.4):

$$z = \frac{p_1-p_2}{\sqrt{\frac{p_1q_1}{n_1}+\frac{p_2q_2}{n_2}}} \quad (4)$$

Each of these conditions serves a specific purpose in evaluating different aspects of statistical significance.

2.3.t-test (For small sample size)

The t-test is a versatile statistical tool frequently employed in veterinary research and clinical practices. Its primary strength lies in its ability to compare the means (averages) of two groups, making it particularly valuable when dealing with small sample sizes (n < 30), which is a common occurrence in veterinary studies. It is one of the most popular statistical techniques used to test whether the mean difference between two groups is statistically significant.

Assumptions for using the t-test:

1. The population should follow a normal distribution.
2. Samples are selected at random.
3. The sample size is less than 30 ($n < 30$).
4. Sample sizes should not differ widely between the samples.
5. The sample variance should not vary widely.

2.2.1. Formulas for t-test in different conditions

Condition 1: Involves testing the significant difference between a sample mean and a population mean using Student's t-test. This test assesses the significance of a single mean with the formula (Eq.5):

$$t = \frac{|\bar{x}-\mu|}{\left(\frac{s}{\sqrt{n}}\right)} \quad , df = n - 1 \qquad (5)$$

where the degrees of freedom df are $n - 1$.

Condition 2: Is used to test the significant difference between two sample means from a single population, utilizing Fisher's t-test. This test compares two means from the same population with the formula (Eq.6):

$$t = \frac{|\bar{x}_1 - \bar{x}_2|}{s_v\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \qquad (6)$$

where S_v the pooled standard deviation, is calculated as (Eq.7):

$$s_v = \sqrt{\frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{(n_1+n_2-2)}} \qquad (7)$$

Here, $\bar{x}_1$, s_1 , and n_1 ,are the mean, standard deviation, and sample size for sample 1, respectively, and $\bar{x}_2$, s_2 , and n_2 , are for sample 2.

Condition 3: Addresses the significant difference between two sample means from different populations. This test is used to compare means from distinct populations with the following formulas (Eq.8 & 9):

$$t = \frac{|(\bar{x}_1 - \bar{x}_2)|}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \qquad (8)$$

and,

$$t = \frac{\left(\frac{s_1^2}{n_1}\right)t_1 - \left(\frac{s_2^2}{n_2}\right)t_2}{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} \qquad (9)$$

where t_1 *and* t_2 are table values for the two samples.

Condition 4: Involves the paired t-test, which is used to test the significant difference between two related sample means (paired observations). The formula for the paired t-test is (Eq.10):

$$t = \frac{\bar{d}}{S_d/\sqrt{n}} \tag{10}$$

Where, $\bar{d}$ is the average difference between the paired observations, calculated as (Eq.11):

$$\bar{d} = \sum d \,/n \tag{11}$$

with $d = X - Y$, and S_d is the standard deviation of these differences, calculated by (Eq.12):

$$S = \sqrt{\frac{\sum d^2 - n(\bar{d})^2}{n-1}} \tag{12}$$

The paired t-test is appropriate when the parent population exhibits a normal distribution, sample sizes are equal, and the sample observations are paired or dependent.

3. Statistical analysis of recovery time in parvo-viral infected dogs by Z-test

Canine parvovirus (CPV) is a highly contagious viral disease that primarily affects puppies and young dogs, causing severe gastrointestinal distress. The rapid spread and high mortality rate of CPV make it a significant concern for veterinarians and dog owners alike. Understanding the factors influencing recovery time in infected dogs is crucial for developing effective treatments and management strategies. For this study, data were collected from the veterinary clinical complex in Jabalpur. The dataset includes information on 68 dogs, divided into twO groups of 34, each diagnosed with CPV and given different treatments. One group received a standard treatment (Table 1), while the other group received a new experimental treatment (Table 2). The objective is to compare the mean recovery times of the two groups.

The hypotheses formulated for this study are: Null hypothesis (H_0): There is no difference in the mean recovery times between the two treatments ($\mu_1 = \mu_2$). Alternative hypothesis (H_1): There is a difference in the mean recovery times between the two treatments ($\mu_1 \neq \mu_2$). The steps to conduct a Z-test comparing recovery times include: (1) Collecting the data by gathering recovery times for both groups, (2) Calculating the means to determine the mean recovery time for each group, (3) Calculating the standard deviation of recovery times within each group, (4) Formulating the hypotheses: Null hypothesis (H_0): $\mu_1 = \mu_2$, Alternative hypothesis (H_1): $\mu_1 \neq \mu_2$, (5) Computing the Z-test for two independent means by using the means, standard deviations, and sample sizes of the two groups to calculate the Z statistic, and

(6) Determining the significance by comparing the computed Z value to the critical value from the Z table, considering the desired significance level (typically 0.05). By following these steps, it is possible to determine whether there is a statistically significant difference in the mean recovery times between the two treatment groups.

Table 1: Collect the data ($\bar{x}_1$) : Standard treatment group (34) *(Recovery time in days)*

S. No.	Days	S. No.	Days	S. No.	Days
1	10	**13**	13	**25**	13
2	12	**14**	14	**26**	12
3	14	**15**	15	**27**	11
4	11	**16**	16	**28**	12
5	13	**17**	12	**29**	13
6	12	**18**	11	**30**	14
7	15	**19**	13	**31**	15
8	14	**20**	14	**32**	13
9	13	**21**	12	**33**	12
10	12	**22**	11	**34**	11
11	11	**23**	14		
12	10	**24**	13		

Table 2: Collect the data ($\bar{x}_2$): Experimental treatment group (34) *(Recovery time in days)*

S. No.	Days	S. No.	Days	S. No.	Days
1	8	**13**	8	**25**	6
2	9	**14**	7	**26**	9
3	7	**15**	9	**27**	8
4	8	**16**	6	**28**	7

5	9	**17**	8	**29**	6
6	10	**18**	7	**30**	8
7	7	**19**	9	**31**	9
8	6	**20**	8	**32**	7
9	9	**21**	7	**33**	8
10	8	**22**	6	**34**	6
11	7	**23**	8		
12	6	**24**	7		

3.1. Calculations

To perform the Z-test for comparing the mean recovery times between the standard treatment group and the experimental treatment group for dogs infected with CPV, the calculations and conclusions in equations (13-15) are as follows:

From Tables 1 and 2, the data calculations are summarized as follows: the mean recovery time for the standard treatment group x_1 is 12.74 days, while for the experimental treatment group $\bar{x}_2$, it is 7.59 days. The standard deviation for the standard treatment group (S_1) is 1.54 days, and for the experimental treatment group (S_2) it is 1.13 days.

The hypotheses for the test are:

1. Null hypothesis (H_0): μ1 = μ2 (There is no difference in mean recovery times between the two treatments).
2. Alternative hypothesis (H_1): μ1 ≠ μ2 (There is a difference in mean recovery times between the two treatments).

To compute the Z-test statistic for two independent means, the formula used is:

$$z = \frac{|(\bar{\bar{x}}_1 - \bar{x}_2)|}{\sqrt{\frac{s_1^2}{n1} + \frac{s_2^2}{n2}}} \quad (13)$$

$$z = \frac{\left(\frac{s_1^2}{n_1}\right)z_1 - \left(\frac{s_2^2}{n_2}\right)z_2}{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} \quad (14)$$

This test is intended to compare means from different populations to determine if there is a significant difference between them. z_1 *and* z_2 are the table values (of two samples with variances s_1^2 and s_2^2 and n_1 and n_2 number of observations at $n_1 - 1$ and $n_2 - 1$ degree of freedom, respectively. The mean recovery time for the standard treatment group is 12.74 days, while for the experimental treatment group, it is 7.59 days. The standard deviations are 1.54 days for the standard treatment group and 1.13 days for the experimental treatment group. Given these data, the Z statistic can be calculated using the formula:

$$z = \frac{(12.74 - 7.59)}{\sqrt{\frac{1.54}{n_1} + \frac{1.13}{n_2}}} \quad (15)$$

Upon computation, the Z statistic is found to be 15.68. This corresponds to an extremely low p-value, approximately 0.0, indicating a statistically significant difference in recovery times between the two treatment groups.

4. Statistical analysis of viral load in parvo-viral infected dogs by t-test

From Tables 3 and 4, the dataset includes 20 dogs, split into two groups of 10, each diagnosed with CPV.

Table 3: Collect the data ($\bar{x}_1$): Symptomatic group viral loads (10) *(Measured as no. of copies in 1 ml)*

S. No.	No. of copies in 1 ml
1	150
2	200
3	250
4	300
5	350

6	400
7	450
8	500
9	550
10	600

Table 4: Collect the data (x_2): Asymptomatic group viral loads (10) *(Measured as no. of copies in 1 ml)*

S. No.	No. of copies in 1 ml
1	100
2	120
3	140
4	160
5	180
6	200
7	220
8	240
9	260
10	280

4.1. Calculations

The mean viral load for the symptomatic group was calculated to be 375, while for the asymptomatic group it was 190. The standard deviations were found to be 151.38 for the symptomatic group and 60.55 for the asymptomatic group.

To assess the difference between the groups, the pooled standard deviation (S_p) was computed as 115.29. The T-statistic, calculated as 3.588, was determined using the formula (Eq.16):

$$t = \frac{|\bar{x}_1 - \bar{x}_2|}{s_v\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \qquad (16)$$

where S_v is the pooled standard deviation, calculated by (Eq.17):

$$s_v = \sqrt{\frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{(n_1+n_2-2)}} \qquad (17)$$

Where, $\bar{x_1}$, s_1 , and n_1 ,are the mean, standard deviation, and sample size in sample 1, respectively. Similarly, $\bar{x_2}$, s_2 , and n_2 , are the mean, standard deviation, and sample size in sample 2, respectively. s_v is the pooled standard deviation, and $(n_1 + n_2 - 2)$ are the degrees of freedom.

Here, $\bar{x_1}$, s_1 , and n_1 ,refer to the mean, standard deviation, and sample size of the symptomatic group, respectively, while $\bar{x_2}$, s_2 , and n_2 , are for the asymptomatic group. The degrees of freedom were determined to be 18. The p-value associated with this t-statistic is 0.0021.

5. Conclusion

Given that the p-value is extremely low and below the significance threshold of 0.05, we reject the null hypothesis. This finding reveals a statistically significant difference in both the mean recovery times between the two groups of dogs and the mean viral loads between symptomatic and asymptomatic groups. Specifically, the experimental treatment group demonstrates a significantly shorter average recovery time compared to the standard treatment group. Additionally, there is a notable difference in viral loads between symptomatic and asymptomatic dogs. These results suggest that the experimental treatment may be more effective in reducing recovery time for dogs infected with CPV and highlight significant variations in viral loads based on symptomatic status.

References

Abebe, T.H. (2019). The derivation & choice of Appropriate test Statistics in research methodology, 5(3): 33-40. doi: 10.11648/j.ml.20190503.11.

Kanakraj, P. (2018). A textbook on Animal Genetics. ISBN: 81-85860-72-6.

Kim, T.K. (2015). t test as a parametric statistic: Anesthesiol, 68(6):540-546. http://dx.doi.org/10.4097/kjae.2015.68.6.540.

Singh, S.S. et al. (2017). Statistical Analysis of Canine Parvovirus Infection in Dogs. Journal of Veterinary Medicine and Animal Health.

Thiagarajan, R. (2014). Principles of animals' genetics & population genetics. Ed ISBN NO. 9789-3812-2673-35.