Journal MVZ Córdoba
2022; may-august. 27(2):e2572
https://doi.org/10.21897/rmvz.2572

Original

Determination of minimum number of animals in comparing treatment means by power analysis

Ozgur Koskan[1*]; Hayati Koknaroglu[1]; Yasin Altay[2].

[1]Universidad de Ciencias Aplicadas de Isparta, Facultad de Agricultura, Departamento de Ciencia Animal, Isparta, Turquía
[2]Universidad de Eskisehir Osmangazi, Facultad de Agricultura, Departamento de Ciencia Animal, Eskisehir, Turquía
*Correspondencia: ozgurkoskan@isparta.edu.tr

Received: November 2021; Accepted: February 2022; Published: May 2022.

ABSTRACT

Objective. The purpose of this study was to determine the minimum number of animals (minimum sample size) in treatment comparisons with different effect sizes (0.25-2.0), the number of treatments (2-7), and the power of the test (80-95%). In addition, linear, quadratic, and cubic regressions equations that estimate the minimum sample size that should be used in treatment comparisons were developed. **Materials and methods.** Within the scope of this research, average daily gain (GDP) of feedlot cattle experiments conducted at Iowa State University totaling 1283 steers were used. The power of the test was calculated after random samples were taken from the GDP data and the differences between the treatments in terms of standard deviation were established. This process was iterated 1000 times via a macro written in the Minitab package program in the number of treatments and power levels to be compared. **Results.** It was found that the cubic regression equations gave more reliable results than others. As a result, after determining the number of treatments, the power of the test, and the effect size, a sufficient number of experimental units can be easily determined by using the estimation equations created without power analysis. **Conclusions.** In this way, excess money expenditure and financial loss in scientific studies can be prevented and the opportunity to find financing more easily can be provided.

Keywords: Effect size; minimum number of animals; sample size; power analysis; simulation (*Source; CAB*).

RESUMEN

Objetivo. El propósito de este estudio fue determinar el número mínimo de animales (tamaño mínimo de la muestra) en comparaciones de tratamientos con diferentes tamaños de efecto (0.25-2.0), el número de tratamientos (2-7) y la potencia de la prueba (80- 95%). Además, se desarrollaron ecuaciones de regresión lineal, cuadrática y cúbica que estiman el tamaño mínimo de muestra que debe usarse en las comparaciones de tratamientos. **Materiales y métodos.** Dentro del alcance de esta investigación, se utilizó la ganancia media diaria (GMD) de los experimentos con ganado

How to cite (Vancouver).
Koskan O, Koknaroglu H, Altay Y. Determination of minimum number of animals in comparing treatment means by power analysis. Rev MVZ Cordoba. 2022; 27(2):e2572. https://doi.org/10.21897/rmvz.2572

ISSNL: 0122-0268

de engorde a corral realizados en la Universidad Estatal de Iowa con un total de 1283 novillos. La potencia de la prueba se calculó después de que se tomaron muestras aleatorias de los datos de GMD y se establecieron las diferencias entre los tratamientos en términos de desviación estándar. Este proceso se repitió 1000 veces mediante una macro escrita en el programa del paquete de Minitab en la cantidad de tratamientos y niveles de potencia a comparar. **Resultados.** Se encontró que las ecuaciones de regresión cúbica dieron resultados más fiables que las demás. Como resultado, después de determinar el número de tratamientos, la potencia de la prueba y el tamaño del efecto, se puede determinar fácilmente un número suficiente de unidades experimentales utilizando las ecuaciones de estimación creadas sin análisis de potencia. **Conclusiones.** De esta manera, se pueden prevenir los gastos excesivos de dinero y las pérdidas financieras en estudios científicos y se puede brindar la oportunidad de encontrar financiamiento más fácilmente.

Palabras clave: Tamaño del efecto; Número mínimo de animales; Tamaño de la muestra; Análisis de potencia; Simulación (*Fuente: CAB*).

INTRODUCTION

Power, in statistics, is the ratio to reject the H_0 hypothesis when H1 is correct and the probability to find a difference when a real difference exists in the population. Type II error is the probability of not finding a difference even though there is a difference. Thus power is defined as type II error (1). When there are significant differences among treatments, power is the probability of this difference to be real (2). The power of a study is determined by three factors: the sample size, the alpha level, and the effect size (3).

Power analysis is the determination of the sample size based on statistical parameters in a planned study. The sample size becomes larger when effect size and type I error level are kept low and the power of the test high (4). For a suitable sample size, a lower effect size increases the accuracy of the parameter estimation. Small effect size for a suitable sample size increases the accuracy of the parameter estimation (5). The success of the research will increase if the sample size is sufficient, that is, the appropriate number. The purpose of the research, statistical distribution patterns, measurement methods used in the research, research model, and statistical analysis methods are very important in determining the sample size. To determine the sample size, the population parameters should be known, and type I error and type II error probabilities of the effect size should be determined. When all other factors are kept constant, the power of the test decreases as the probability of type I error (α) determined at the beginning decreases (6). Experimental power analysis is the determination of the power of the decisions obtained in the direction of a concluded research. In hypothesis controls, when the null hypothesis is tested against the alternative hypothesis, and the decision is made, there are two types of errors. During the controls, when the correct null hypothesis (H_0) is rejected, this is a type I error (7) and the probability of making type I error is denoted by α. Type II error is the error made by accepting the control hypothesis when the alternative hypothesis (H_1) is correct (8). and the probability of making type II error is denoted by β. The accuracy of the decisions made in hypothesis controls depends on α and β probabilities. In the hypothesis control, the power of the test theoretically varies between 0 and 1. In general, it is desired to have the power of the test 80% and above, however if it is below 50%, this does not allow a reliable comment on the results of the study (8,9,10). If maximizing the probability of reaching the correct result in a hypothesis control is wanted, the probability of type I error should preserve what was originally agreed upon and the probability of the power of the test should be high (11). Smaller β probability increases the power of the hypothesis control. This depends on the distance from the sample value that the hypothesis to be controlled for a given α and sample size.

It is necessary to determine how effective the levels of the examined factor are in explaining the variable under consideration. Determining the power of the test or the number of individuals in the sample to reach a certain power is to estimate the effect size on the variable. Starting the power analysis with the determining effect size removes most of the obstacles that may arise. Normal hypothesis tests in a study can detect differences between two treatments, however, the results do not give the researcher clear information about this difference. As the effect size is also a measure of the standardized

difference it can give this information. The effect size is important in comparing the results of any two studies conducted on the characteristics considered.

In recent years, power analysis has been widely used in hypothesis testing protocols, especially of studies involving biological material (12,13, 14,15,16,17,18,19,20,21,22,23,24). In many areas, power analysis is seen as a universal step that must be done before trials are conducted. In addition, the power of the test can be easily calculated with the help of sample size, effect size, and variance (standard deviation) (25). Before starting the research, determining the sample size with the help of power analysis is the most effective method in determining the dynamics of the study such as budget, labor, time, etc. In this way, by determining the sample size of the research, it will be easier to find finance, and time and money will not be wasted. However, using more than the optimum sample size will not be financially correct, in terms of time and labor. For these reasons, it has become essential to determine the sample size in scientific studies using animal material. For these reasons, it has become essential to determine the sample size in scientific studies using animal material.

Since animal purchase price and feeding are high in animal science experiments, it is important to determine the minimum sample size used in experiments. Thus, the purpose of this study was to determine the minimum sample size in treatment comparisons with standard deviations of different mean differences with power analysis.

MATERIALS AND METHODS

Average daily gain (ADG) of feedlot cattle experiments conducted at Iowa State University totaling 1283 steers was used to determine the difference of means in terms of standard deviation. In the study, the minimum sample size was determined by conducting power analysis. For this purpose combinations of two power rate (95 and 80%), number of treatments (2,3,4,5,6,7), difference of means in terms of standard deviation (0.25, 0.5, 0.75, 1.0, 1.5, 2.0, 2.5, 3.0, 5.0) were used.

The average daily gain (ADG) of the steers in the study was 1.328±0.00549 kg, and the percentage of the standard deviation to the mean (Coefficient of Variation) was found as 14.83%. Normal distribution compliance control was performed with the Anderson Darling test, and it was found that ADG was normally distributed (p>0.05). Skewness and Kurtosis values in the descriptive statistics in Table 1 support the Anderson Darling test result of the ADG.

Table 1. Descriptive statistics of Average Daily Gain (ADG)

Variable	N	Mean	SEMean	StDev	CoefVar	Min	Median	Max	Skewness	Kurtosis
ADG	1283	1.328	0.005	0.197	14.830	0.545	1.335	2.102	-0.040	0.170

The effect size, in other words, the difference between the means in terms of standard deviation is obtained by using Equation 1.

$$\Delta = \frac{(\mu_A - \mu_B)}{S_p} \quad [1]$$

Calculation of the power test for comparing two treatments (t-test);

When the probability of type I error is denoted as α in the hypothesis control, the hypothesis is accepted or rejected as a result of the comparison of the table value of the $(n_A\text{-}1) + (n_B\text{-}1)$ degree of freedom t distribution with the value of the test statistic. If the H_0 hypothesis is rejected (invalid), the non-central parameter of the relevant distribution is calculated by Equation 2.

$$\delta = \frac{|\mu_A - \mu_B|}{S_p} \sqrt{\frac{n}{2}} \quad [2]$$

The power of the test is calculated with Equations 3, 4, and 5 by using, $(n_A\text{-}1) + (n_B\text{-}1)$ degree of freedom t distribution calculated with the help of Equation 2 (26).

Power of test for two-sided hypothesis = $(1-\beta)=P(t < -t_{\alpha/2})+ P(t > t_{\alpha/2})$ [3]

Power of test for one-sided (right) hypothesis = $(1-\beta)= P(t > t_{\alpha})$ [4]

Power of test for one-sided (left) hypothesis = $(1-\beta)= P(t < -t_{\alpha})$ [5]

Calculation of the power of the test in 3 or more treatments comparisons (F test);

When the probability of type I error is denoted as α in the hypothesis control, the hypothesis is accepted or rejected as a result of the comparison of the table value of the (GASD -1) and (HSD-1) degree of freedom F distribution with the value of the test statistic. If the H_0 hypothesis is rejected (invalid), the non-central parameter of the relevant distribution is calculated by Equation 6.

$$\phi = \sqrt{\frac{n\sum_{i=1}^{k}(\mu_i-\bar{\mu})}{k\sigma^2}} \quad [6]$$

After determining ϕ calculated with the help of Equation 6, the power of the test is calculated by looking at Hartley or power tables (dfb -1) and (dfw -1) degrees of freedom are calculated by looking at Hartley or power tables.

After determining the ϕ calculated with the help of Equation 6, (dfb -1) and (dfw -1) degrees of freedom connected Hartley or power tables, the power of the test is calculated (27,28).

In the study, ADG of 1283 steers was considered as a population. The samples were created by random sampling with a replacement method by taking into account the parameters of the population.

The experimental approach to the power of the test in the simulation study is that two populations having a mean of μx and μy variances of and with normal distribution and assumed to be μx = μy + 1Δ. In the next stage, the desired amount of sample size was taken from the populations and the statistical significance of the difference between the two sample averages was checked. When this process is repeated for the number of attempts (for example 10000), the power of the test is calculated in terms of the proportional (empirical) rejection probability of the H_0 hypothesis.

The minimum sample size was estimated by creating linear, quadratic, and cubic regression equations between the actual numbers of treatments with different power and effect sizes and the estimated number of treatments calculated by the random resampling method from the population. For this purpose, this process was iterated 1000 times with the help of a macro written in the Minitab package program. In addition, linear, quadratic, and cubic equations with power between 80% and 95% were created separately in terms of each standard deviation between the averages. Linear, quadratic, and cubic regression models are given in Equation 7, Equation 8, and Equation 9, respectively.

$$\hat{Y}_i=\beta_0+\beta_1 X_1+\varepsilon_i \quad [7]$$

$$\hat{Y}_i=\beta_0+\beta_1 X_1-\beta_2 X_1^2+\varepsilon_i \quad [8]$$

$$\hat{Y}_i=\beta_0+\beta_1 X_1-\beta_2 X_1^2+\beta_3 X_1^3+\varepsilon_i \quad [9]$$

These equations represent : observation i of dependent variable y (number of minimum animals), , , and : regression parameters, : observation i of the independent variable (number of treatments to compare) and : random error.

RESULTS

Linear, quadratic, and cubic regression equations created to determine the minimum sample size are designed depending on different effect size and power. Linear, quadratic, and cubic regression equations and equations with different effect size and power are given in Table 2, Table 3, and Table 4, respectively. In addition, linear, quadratic, and cubic equations with different effect sizes (0.25-2.00) and power between 80% and 95% were created. When the determination coefficients of the regression equations were examined, it was found that the most accurate estimation equations were cubic, quadratic, and linear estimation equations, respectively. In the linear regression equations created to determine the minimum sample size, it is found that determination coefficients approached 1 as the power of the test increased and the effect size decreased (Table 2). The determination coefficient of the equations with an effect size of 1.50 and the power of the test is 80 - 85%, which means that there is almost a functional relationship between them.

Table 2. Linear regression equation and determination coefficients for differences in terms of each standard deviation generated between means.

Power (1-β)	Δ	Regression Equations	DC
	0.25	195.60 + 35.97 Treatment	97.7
	0.50	50.26 + 8.94 Treatment	97.1
	0.75	22.74 + 4.06 Treatment	97.3
80%	1.00	13.38 + 2.29 Treatment	98.0
	1.25	9.53 + 1.40 Treatment	98.5
	1.50	7.00 + 1.00 Treatment	100
	2.00	4.94 + 0.46 Treatment	91.4
	0.25	226.30 + 39.26 Treatment	97.7
	0.50	57.73 + 9.80 Treatment	97.2
	0.75	25.69 + 4.51 Treatment	98.0
85%	1.00	15.02 + 2.51 Treatment	97.6
	1.25	10.72 + 1.54 Treatment	96.1
	1.50	8.00 + 1.00 Treatment	100
	2.00	5.13 + 0.60 Treatment	92.2
	0.25	269.20 + 43.14 Treatment	97.70
	0.50	68.95 + 10.71 Treatment	97.67
	0.75	31.88 + 4.66 Treatment	97.11
90%	1.00	19.02 + 2.51 Treatment	97.61
	1.25	12.32 + 1.74 Treatment	96.94
	1.50	9.19 + 1.14 Treatment	97.96
	2.00	6.13 + 0.60Treatment	92.20
	0.25	339.50 + 49.11 Treatment	97.54
	0.50	85.75 + 12.31 Treatment	97.32
	0.75	38.98 + 5.49 Treatment	98.01
95%	1.00	22.19 + 3.14 Treatment	96.39
	1.25	14.67 + 2.00 Treatment	98.13
	1.50	10.83 + 1.37 Treatment	96.81
	2.00	6.70 + 0.77 Treatment	96.13

DC: Determination Coefficient (R^2)

The determination coefficients of linear estimation equations with an effect size of 2.00 and power of 80%, 85%, and 90% had the lowest values and were 91.40, 92.20, and 92.20%, respectively. In the quadratic estimation equations, the determination coefficients of all the equations except the estimation equation (91.40%), which had a power of 80% and an effect size of 2.00, had values of 99.20% and above (Table 3). As in simple linear estimation equations, the effect size is 1.50 and the power of the test is 80 - 85%, the coefficient of expression is approximately 1, and there is a similarity in the quadratic estimation equations.

It was found that the use of cubic estimation equations, as well as simple linear and quadratic estimation equations, would be more effective (Table 4). The determination coefficients of cubic estimation equations with different effect sizes and power of the test generally were approximately 1. By using the treatment numbers, estimation equations with an effect size of 2.00 and the power of the test 80%, explained 93.70% of the estimated minimum sample size. This cubic estimation equation had the lowest determination coefficient. In other estimation equations, using the number of treatments in determining the minimum sample size in terms of ADG will lead to more accurate results.

Linear, quadratic, and cubic estimation equations for the minimum sample size which is calculated by considering the number of treatments, different power, and effect size are given in Table 5, Table 6, and Table 7. It was found that the number of treatments determined by the linear estimation equations decreased as the effect sizes increased, and the number of animals increased as the power of the test increased. In the comparison of two treatments, required animal numbers were 18, 20, 25, and 29 when the power of the test was between 80% - 95% and effect size (Δ) was 1 (Table 5). In the comparison of three treatments, required minimum animal numbers were 7, 7, 8, and 10 when the power of the test was between 80% - 95% and effect size (Δ) was 2. In the comparison of seven treatments, the required minimum animal numbers were 448, 502, 572, and 684 when the power of the test was between 80% - 95%, and effect size (Δ) was 0.25.

While the number of treatments was 2 in the quadratic estimation equations, fewer estimates occurred compared to the minimum sample size found by using linear estimation equations with the same conditions as the minimum sample size that should be found with different power and effect size (Table 6). There is no similar relationship between the cubic estimation equations and the minimum animal numbers found using the quadratic estimation equations. In the comparison of two treatments, the required minimum animal numbers were 68, 74, 87, and 106 when the power of the test was between 80% - 95%, and effect size (Δ) was 0.5. In the comparison of five treatments, required minimum animal numbers were 12, 13, 16, and 19 when the power of the test was between 80% - 95%, and effect size (Δ) was 1.5. In the comparison of seven treatments, required minimum animal numbers were 9, 9, 11, and 12 when the power of the test was between 80% - 95% and effect size (Δ) was 2.

Table 3. Quadratic regression equation and determination coefficients for differences in terms of each standard deviation generated between means.

Power (1-β)	Δ	Regression Equations	Determination Coefficient (R^2)
	0.25	132.50 + 68.76 Treatment - 3.64 Treatment^2	99.9
	0.50	32.61 + 18.10 Treatment - 1.02 Treatment^2	99.8
	0.75	15.31 + 7.91 Treatment - 0.43 Treatment^2	99.6
80%	1.00	10.29 + 3.89 Treatment - 0.18 Treatment^2	99.2
	1.25	7.99 + 2.20 Treatment - 0.09 Treatment^2	99.3
	1.50	7.00 + 1.00 Treatment - 0.00 Treatment^2	100
	2.00	4.94 + 0.46 Treatment - 0.00- Treatment^2	91.4
	0.25	157.60 + 74.94 Treatment - 3.96 Treatment^2	99.8
	0.50	38.54 + 19.76 Treatment - 1.11 Treatment^2	99.9
	0.75	18.26 + 8.37 Treatment - 0.43 Treatment^2	99.9
85%	1.00	10.69 + 4.76 Treatment - 0.25 Treatment^2	99.7
	1.25	7.63 + 3.15 Treatment - 0.18 Treatment^2	98.9
	1.50	8.00 + 1.00 Treatment - 0.00 Treatment^2	100
	2.00	3.57 + 1.40 Treatment - 0.09 Treatment^2	96.6
	0.25	193.40 + 82.52 Treatment - 4.38 Treatment^2	99.80
	0.50	50.07 + 20.52 Treatment - 1.09 Treatment^2	99.80
	0.75	22.90 + 9.32 Treatment - 0.52Treatment^2	99.70
90%	1.00	14.69 + 4.76 Treatment - 0.25 Treatment^2	99.70
	1.25	9.23 + 3.35 Treatment - 0.18 Treatment^2	99.10
	1.50	7.64 + 1.95 Treatment - 0.09 Treatment^2	99.20
	2.00	4.57 + 1.40 Treatment - 0.09 Treatment^2	96.60
	0.25	250.30 + 95.40 Treatment - 5.14 Treatment^2	99.80
	0.50	62.23 + 24.53 Treatment - 1.36 Treatment^2	99.80
	0.75	30.31 + 9.99 Treatment - 0.50 Treatment^2	99.70
95%	1.00	15.07 + 6.84 Treatment - 0.41 Treatment^2	99.90
	1.25	11.57 + 3.61 Treatment - 0.18 Treatment^2	99.80
	1.50	8.04 + 2.82 Treatment - 0.16 Treatment^2	99.60
	2.00	5.46 + 1.41 Treatment - 0.07 Treatment^2	97.90

Table 4. Cubic regression equation and determination coefficients for differences in terms of each standard deviation generated between means

Power (1-β)	Δ	Regression Equations	Determination Coefficient (R^2)
	0.25	85.62 + 106.90 Treatment - 12.89 Treatment^2 + 0.69 Treatment^3	100
	0.50	18.05 + 29.97 Treatment - 3.893 Treatment^2 + 0.21 Treatment^3	100
	0.75	7.71 + 14.10 Treatment - 1.93 Treatment^2 + 0.11 Treatment^3	99.9
80%	1.00	3.95 + 9.05 Treatment - 1.43 Treatment^2 + 0.09259 Treatment^3	99.8
	1.25	8.62 + 1.69 Treatment + 0.04 Treatment^2 - 0.01 Treatment^3	99.3
	1.50	1.48 + 4.53 Treatment - 0.71 Treatment^2 + 0.05 Treatment^3	99.8
	2.00	7.48 - 1.61 Treatment + 0.50 Treatment^2 - 0.04 Treatment^3	93.7
	0.25	104.40 + 118.30 Treatment - 14.46 Treatment^2 + 0.78 Treatment^3	100
	0.50	27.14 + 29.05 Treatment - 3.36 Treatment^2 + 0.17 Treatment^3	100
	0.75	13.19 + 12.50 Treatment - 1.43 Treatment^2 + 0.07 Treatment^3	100
85%	1.00	5.62 + 8.89 Treatment - 1.25 Treatment^2 + 0.07 Treatment^3	100
	1.25	5.10 + 5.21 Treatment - 0.68 Treatment^2 + 0.04 Treatment^3	99.1
	1.50	8.00 + 1.00 Treatment + 0.00 Treatment^2 - 0.00 Treatment^3	100
	2.00	2.95 + 1.92 Treatment - 0.21 Treatment^2 + 0.01 Treatment^3	96.6
	0.25	133.20 + 131.50 Treatment - 16.25 Treatment^2 + 0.88 Treatment^3	100
	0.50	34.24 + 33.41 Treatment - 4.21 Treatment^2 + 0.23 Treatment^3	100
	0.75	14.67 + 16.02 Treatment - 2.14 Treatment^2 + 0.12 Treatment^3	99.90
90%	1.00	9.62 + 8.89 Treatment - 1.25 Treatment^2 + 0.07 Treatment^3	100
	1.25	5.43 + 6.44 Treatment - 0.93 Treatment^2 + 0.06 Treatment^3	99.50
	1.50	4.48 + 4.53 Treatment - 0.71 Treatment^2 + 0.05 Treatment^3	99.80
	2.00	3.95 + 1.92 Treatment - 0.21 Treatment^2 + 0.01 Treatment^3	96.60
	0.25	178.10 + 154.20 Treatment - 19.39 Treatment^2 + 1.06 Treatment^3	100
	0.50	45.76 + 37.94 Treatment - 4.611 Treatment^2 + 0.24 Treatment^3	100
	0.75	22.71 + 16.17 Treatment - 2.00 Treatment^2 + 0.11 Treatment^3	99.90
95%	1.00	11.90 + 9.42 Treatment - 1.04 Treatment^2 + 0.05 Treatment^3	100
	1.25	11.57 + 3.61 Treatment - 0.18 Treatment^2 - 0.00 Treatment^3	99.80
	1.50	6.14 + 4.37 Treatment - 0.54 Treatment^2 + 0.03 Treatment^3	99.80
	2.00	4.19 + 2.45 Treatment - 0.32 Treatment^2 + 0.02 Treatment^3	98.10

Table 5. Number of animals per treatment determined by linear regression having different effect sizes and power.

Test statistics	Number of treatments	1-β	Δ=0.25	Δ=0.5	Δ=0.75	Δ=1	Δ=1.25	Δ=1.5	Δ=2
t	2	0.80	268	69	31	18	13	9	6
		0.85	305	78	35	20	14	10	7
		0.90	356	91	42	25	16	12	8
		0.95	438	111	50	29	19	14	9
F	3	0.80	304	77	35	21	14	10	7
		0.85	344	88	40	23	16	11	7
		0.90	399	102	46	27	18	13	8
		0.95	487	123	56	32	21	15	10
F	4	0.80	340	86	39	23	16	11	7
		0.85	384	97	44	25	17	12	8
		0.90	442	112	51	30	20	14	9
		0.95	536	135	61	35	23	17	10
F	5	0.80	376	95	43	25	17	12	8
		0.85	423	107	49	28	19	13	9
		0.90	485	123	56	32	22	15	10
		0.95	586	148	67	38	25	18	11
F	6	0.80	412	104	47	27	18	13	8
		0.85	462	117	53	31	20	14	9
		0.90	529	134	60	35	23	17	10
		0.95	635	160	72	42	27	20	12
F	7	0.80	448	113	52	30	20	14	9
		0.85	502	127	58	33	22	15	10
		0.90	572	144	65	37	25	18	11
		0.95	684	172	78	45	29	21	13

Table 6. Number of animals per treatment determined by quadratic regression having different effect sizes and power.

Test statistics	Number of treatments	1-β	Δ=0.25	Δ=0.5	Δ=0.75	Δ=1	Δ=1.25	Δ=1.5	Δ=2
t	2	0.80	256	68	30	18	12	9	6
		0.85	292	74	34	20	14	10	6
		0.90	341	87	40	24	16	12	8
		0.95	421	106	49	28	19	14	8
F	3	0.80	306	78	36	21	14	10	7
		0.85	347	88	40	23	16	11	7
		0.90	402	102	47	27	18	13	8
		0.95	491	124	56	32	21	16	10
F	4	0.80	350	89	41	23	16	11	7
		0.85	394	100	45	26	18	12	8
		0.90	454	115	52	30	20	14	9
		0.95	550	139	63	36	24	17	10
F	5	0.80	386	98	45	26	17	12	8
		0.85	434	110	50	29	19	13	9
		0.90	497	126	57	33	22	16	10
		0.95	599	151	68	39	26	19	11
F	6	0.80	414	105	48	28	18	13	8
		0.85	465	118	53	31	20	14	9
		0.90	532	134	61	35	23	17	10
		0.95	638	161	73	42	27	20	12
F	7	0.80	436	110	50	29	19	14	9
		0.85	488	123	56	32	21	15	9
		0.90	557	141	63	36	24	17	11
		0.95	667	168	76	43	29	20	12

It was found that when the number of treatment was 2 (t-test), the minimum sample size obtained with cubic estimation equations was less than that determined by linear and quadratic prediction equations. However, in the F test, the cubic models differ from other models. When the analysis was conducted, it was found that the minimum sample size in each treatment was 418 when power was 95%, the difference of means in terms of standard deviation was 0.25, and the number of the treatment was 2. When the analysis was conducted, it was found that the minimum sample size in each treatment was 13 when power was 90%, the difference of means in terms of standard deviation was 1.5, and the number of the treatment was 3. When the analysis was conducted it was found that the minimum sample size in each treatment was 10 when power was 95%, the difference of means in terms of standard deviation was 2, and the number of the treatment was 2. When the analysis was conducted it was found that the minimum sample size in each treatment was 560 when power was 90%, the difference of means in terms of standard deviation was 0.25, and the number of the treatment was 7. When the analysis was conducted it was found that the minimum sample size in each treatment was 30 when power was 90%, the difference of means in terms of standard deviation was 1.0, and the number of the treatment was 4. When the analysis was conducted it was found that the minimum sample size in each treatment was 8 when power was 80%, the difference of means in terms of standard deviation was 2, and the number of the treatment was 7 (Table 7).

Table 7. Number of animals per treatment determined by cubic regression having different effect sizes and power.

Test statistics	Number of treatments	1-β	Δ=0.25	Δ=0.5	Δ=0.75	Δ=1	Δ=1.25	Δ=1.5	Δ=2
t	2	0.80	254	65	29	17	12	8	6
		0.85	290	74	33	19	14	10	6
		0.90	339	87	40	23	16	12	8
		0.95	418	106	48	27	19	13	8
F	3	0.80	309	79	36	21	14	10	7
		0.85	351	89	40	23	16	11	7
		0.90	406	103	47	28	18	13	9
		0.95	495	125	57	33	21	16	10
F	4	0.80	351	90	41	24	16	12	7
		0.85	396	101	45	26	18	12	8
		0.90	456	116	53	30	20	15	9
		0.95	553	140	63	36	24	17	11
F	5	0.80	384	98	44	25	17	12	8
		0.85	432	110	50	28	19	13	9
		0.90	495	125	57	33	22	16	10
		0.95	597	151	68	39	26	19	11
F	6	0.80	411	104	47	27	18	13	8
		0.85	462	117	53	30	20	14	9
		0.90	528	133	60	34	23	16	10
		0.95	634	160	72	42	27	20	12
F	7	0.80	438	111	50	29	19	14	8
		0.85	491	124	56	32	21	15	9
		0.90	560	142	64	37	25	18	11
		0.95	670	169	77	43	29	20	12

DISCUSSION

In general, considering all conditions, in all estimation equations, the sample size increased as the number of treatments increased, the increase in the power of the test increased the sample size, while the increase in the effect size caused a decrease in the number of samples. As the effect size increased, the sample size was not affected much, regardless of the power of the test. Results showed that the power of the test depends on the effect size of the population and the sample size taken from these populations. These results are similar to the results reported by other researchers (29,30).

Our power of test results did not agree with some researchers' results when they calculated the power of the test if the population variance was homogeneous for two or more treatment comparisons with different effect sizes (30,31,32). The reason for this contradiction is thought to be due to the fact that in the aforementioned studies used populations with different variances while forming treatments, used different amounts of simulation or sampling methods.

Increasing the effect size will lead to the opportunity to work with a smaller sample size due to the increase in the difference between the means to be compared (33). Previous studies support our study. In terms of effect size, in general, if the Cohen's d value is less than 0.20, it has a large effect size, if it is between 0.20 and 0.50, it has a medium effect size, and if d >0.80, it has a weak effect size (27). This means that as the number of comparisons with large effect sizes and the power of the test increase, the optimum sample size increases. When this case was evaluated in terms of medium and low effect size, it was found that there were similar results with a large effect size. In general, before the power analysis is conducted, the researchers determine the power of the test as 80% (34). In a test with 80% power, it was found that as the number of treatments increases, the sample size increases, and the sample size decreases with the increase in effect size, and the results of our study are in agreement with previous studies.

The shape of the distribution is thought to be effective in determining the sample size. When the distributions were non-symmetric, (34) planned an experiment with the number of observations of 4, 6, 8, and 10, and simulated by iterating 10000 times each time. The power of the test was calculated for the differences between the treatment averages when the standard deviation ranged from 0.0 to 2.5. Within the framework of the results obtained, there was no similarity between the results, as the current study was not simulated from a normal distribution.

Results showed that when other variables were similar, the minimum sample size increased as the number of treatments increased, the minimum sample size decreased as the difference of means in terms of standard deviation increased, and the minimum sample size decreased as the power of the test increased. Similar results to this study were found by (8).

In conclusion, researchers can do a hypothesis check by determining the number of experimental units with power between 80% and 95%, using relevant equations and tables without performing power analysis before starting a study. Thus, researchers will be able to reveal important differences as a result of hypothesis control. In this way, when the optimum sample size is determined for any variable, time and money will be spent wisely.

In the current simulation study, although the power of the test, type I and type II errors, and sample sizes were determined, the fact that in previous studies, no prediction equation was developed to determine the minimum sample size with different effect sizes and power increases the originality of this study.

The study can be applied not only for determining the minimum sample size but also for traits derived from any material considered. With the use of linear, quadratic, and cubic regressions equations found, the sample sizes can be easily determined by the researchers. For this reason, this study will provide the opportunity to easily determine the sample sizes in the desired power and effect size without the need for any software. Results showed that the difference between the results of the current study and the statistical program or software results used by the researchers for the sample size was not much and this proves the importance and strong part of the current study.

Conflict of Interest

The authors have no conflict of interest

Acknowledgements

This article is presented as short full text at the 10th International Congress of Animal Science. Dr. Yasin ALTAY was included in the study due to his contributions in the process of full text.

REFERENCES

1. Akobeng AK. Understanding type I and type II errors, statistical power and sample size. Acta Paediatrica. 2016; 105(6):605-609. https://doi.org/10.1111/apa.13384

2. Greenland S, Senn SJ, Rothman KJ, Carlin JB, Poole, C, Goodman SN, Altman DG. Statistical tests, P values, confidence intervals, and power: a guide to misinterpretations. Eur J Epidemiol. 2016; 31(4):337-350. https://doi.org/10.1007/s10654-016-0149-3

3. Cohen J. Statistical power analysis: Current Directions in Psychological Science 1992; 1(3):98-101. https://doi.org/10.1111/1467-8721.ep10768783

4. Wilcox RR. ANOVA: A paradigm for low power and misleading measures of effect size?. Rev Educ Res. 1995; 65(1):51-77. https://doi.org/10.3102/00346543065001051

5. Kelley K, Rausch, JR. Sample size planning for the standardized mean difference: Accuracy in parameter estimation via narrow confidence intervals. Psychol Methods. 2006; 11(4):363–385. https://doi.org/10.1037/1082-989X.11.4.363

6. Houle TT, Penzien DB, Houle CK. Statistical power and sample size estimation for headache research: An overview and power calculation tools. Headache. 2005; 45(5):414-418. https://doi.org/10.1111/j.1526-4610.2005.05092.x

7. Mascha EJ, Vetter TR. Significance, errors, power, and sample size: the blocking and tackling of statistics. Anesth Analg. 2018; 126(2):691-698. https://doi.org/10.1213/ANE.0000000000002741

8. Aslan E, Koşkan Ö, Altay Y. Determination of the Sample Size on Different Independent K Group Comparisons by Power Analysis. Turk Tarim Arast Derg. 2021; 8(1):34-41. https://doi.org/10.19159/tutad.792694

9. Szucs D, Ioannidis JP. Sample size evolution in neuroimaging research: An evaluation of highly-cited studies (1990–2012) and of latest practices (2017–2018) in high-impact journals. NeuroImage. 2020; 221:117-164. https://doi.org/10.1016/j.neuroimage.2020.117164

10. Murphy KR. Myors B. Statistical power analysis: A simple and general model for traditional and modern hypothesis tests. Second Edition (2rd ed.). Mahwah, New Jersey, USA: Lawrence Erlbaum Associates Publishers; 2004.

11. Lane SP, Hennes EP. Power struggles: Estimating sample size for multilevel relationships research. J Soc Pers Relat. 2018; 35(1):7-31. https://doi.org/10.1177/0265407517710342

12. Peterman RM. Statistical power analysis can improve fisheries research and management. Can J Fish Aquat Sci. 1990; 47(1):2–15. https://doi.org/10.1139/f90-001

13. Fairweather PG. Statistical power and design requirements for environmental monitoring. Mar Freshw Res. 1991; 42(5):555–567. https://doi.org/10.1071/MF9910555

14. Muller KE, Benignus VA. Increasing scientific power with statistical power. Neurotoxicol Teratol. 1992; 14(3):211–219. https://doi.org/10.1016/0892-0362(92)90019-7

15. Taylor BL, Gerrodette T. The uses of statistical power in conservation biology: the vaquita and northern spotted owl. Conserv Biol Ser. 1993; 7(3):489–500. https://doi.org/10.1046/j.1523-1739.1993.07030489.x

16. Searcy-Bernal R. Statistical power and aquacultural research. Aquaculture. 1994; 127(4):371–388. https://doi.org/10.1016/0044-8486(94)90239-9

17. Thomas L, Juanes F. The importance of statistical power analysis: an example from animal behaviour. Anim. Behav. 1996; 52(4):856–859. https://doi.org/10.1006/anbe.1996.0232

18. Thomas L, Retrospective power analysis. *Conservation Biology.* Conservation Biology. 1997; 11(1):276-280. https://www.jstor.org/stable/2387304

19. Pinu FR, Beale DJ, Paten AM, Kouremenos K, Swarup S, Schirra HJ, Wishart D. Systems biology and multi-omics integration: viewpoints from the metabolomics research community. Metabolites. 2019; 9(4):76. https://doi.org/10.3390/metabo9040076

20. Hoenig JM, Heisey DM. The abuse of power: the pervasive fallacy of power calculations for data analysis. Am Stat. 2001; 55(1):19-24. https://doi.org/10.1198/000313001300339897

21. Faul F, Erdfelder E, Lang AG, Buchner AG. Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behav Res Methods. 2007; 39(2):175-191. https://doi.org/10.3758/BF03193146

22. La Rosa PS, Brooks JP, Deych E, Boone EL, Edwards DJ, Wang Q, Sodergren E, Weinstock G, Shannon WD. Hypothesis testing and power calculations for taxonomic-based human microbiome data. PloS One. 2012; 7(12):e52078. https://doi.org/10.1371/journal.pone.0052078

23. Fanelli D. Do pressures to publish increase scientists' bias? An empirical support from US States Data. PloS One. 2010; 5(4):e10271. https://doi.org/10.1371/journal.pone.0010271

24. Abt G, Boreham C, Davison G, Jackson R, Nevill A, Wallacea E, Williamsa M. Power, precision, and sample size estimation in sport and exercise science research. J Sports Sci. 2020; 38(17):1993-1935. https://10.1080/02640414.2020.1776002

25. Lewis KP. Statistical power, sample sizes, and the software to calculate them easily. BioScience. 2006; 56(7):607-612. https://doi.org/10.1641/0006-3568(2006)56[607:SPSSAT]2.0.CO;2

26. Zar JH. Biostatistical Analysis: Pearson New International Edition. New Jersey, NJ, USA: Pearson Higher Edition; 2013.

27. Hartley HO. The maximum F-ratio as a short-cut test for heterogeneity of variance. Biometrika. 1950; 37(3/4):308-312. https://doi.org/10.2307/2332383

28. Pearson ES, Hartley HO. Charts of the power function for analysis of variance tests, derived from the non-central F-distribution. Biometrika. 1951; 38(1/2):112-130. https://doi.org/10.2307/2332321

29. Başpınar E. Gürbüz F, The power of the test in comparing samples of different sample sizes taken from binary combinations of Normal, Beta, Gamma (Chi-square) and Weibull distributions. J Agric Sci. 2000; 6(1):116-127. https://doi.org/10.1501/Tarimbil_0000000940

30. Lenth RV. Statistical power calculations. J Anim Sci. 2007; 85(13):24-29. https://doi.org/10.2527/jas.2006-449

31. Başpınar E, Type I error and power of the test obtained by applying Student's t, Welch and sorted t-tests on two samples of different sample sizes taken from normal populations with different variance ratios. J Anim Sci. 2001; 7(1):151-157. https://doi.org/10.1501/Tarimbil_0000000271

32. Koşkan Ö, Gürbüz F. Resampling approach and power of t-test and comparison of type I error. J Anim Prod. 2008; 49(1):29-37. https://dergipark.org.tr/en/pub/hayuretim/issue/7617/99817

33. Kalaycioğlu O, Akhanli SE. The importance and main principles of power analysis in health research: Application examples on medical case studies. Turk J Public Health. 2020;. 18(1):103-112. https://doi.org/10.20518/tjph.602400

34. Mewhort, DJ, A comparison of the randomization test with the F test when error is skewed. Behav Res Methods. 2005; 37(3):426-435. https://doi.org/10.3758/BF03192711