Significant aspects about the pitching and batting of the Vegueros team

Aspectos significativos sobre el pitcheo y el bateo del equipo de Vegueros

Aspectos significativos de arremesso e batimento para a equipe Vegueros

Luis Enrique Cardoso Rojas[1]* , *Marcia Virginia Robalino Robalino[2]*

María Daniela Rovalino Robalino[3]

[1]University of Pinar del Río "Hermanos Saíz Montes de Oca". Cuba. Pinar del Río, Cuba.

[2]Totoras Educational Unit. Ecuador.

[3]Alberto Guerra" EU. Ecuador.

*Corresponding author: luis.cardoso@upr.edu.cu

Baseball in Cuba is the national sport, it is part of the identity and pride of the Cuban; a large part of the citizens has acquired culture in this sport, especially if it is about the trajectory of the Pinar del Río baseball team. Its history has been based on the careers of great athletes such as Luis Giraldo Casanova, Omar Linares and Pedro Luis Lazo, among others. This team, which has sometimes changed its name from Pinar del Río to Vegueros as a representation of the identity of the land of great tobacco vegas, owes its genius to the players it has had and its successful directors and coaches.

In honor of that merit, the authors of this work have decided to dedicate the research to the glorious pitching and hitting of Vegueros, not only from the historical point of view; but also, from the conceptualization and analysis of the data that they collect, in detail, what happens in a baseball game and that are considered as a measure of how good or bad a player or a team is in general.

The development in the statistical annotation of competitive performance and its evaluation based on indicators that collect the actions during the game, occurs from the very beginning of this sport, these data provide interesting information to be able to objectively comply with the study and analysis of the game. previous preparation period, to make a correct future planning and as a quantitative basis in the decision-making process for technical groups.

In baseball, there is a large number of coefficients to measure the performance of athletes and teams, with the statistical data that is collected, based on indicators in each of the games that take place during a series. In the case of offense, there is the on-base percentage (OBP) which offers an index of the number of times the batter reaches base by the number of times he tries to do so and uses more indicators related to the action of reaching base without fail that is ultimately the main objective of each batter, the total coefficient of bases covered (TB) which is responsible for recording the number of bases that batters cover with their hit connections (H) and the Slugging (SLUG), also known as power average, which aims to provide an estimate of the strength at bat of each batter.

On the other hand, there are coefficients to measure the performance of the pitchers where the average with which they hit a pitcher (AVE) is found with a calculation very similar to that of the batters , but here the lower it is the better its performance; the average of earned runs (AER) which objective is to determine how many earned runs allowed by a pitcher in every nine innings of performance, for which the total earned runs allowed (ER) are divided by the total innings pitched (INN) and multiply by nine; the won-lost average and the WHIP with its acronym in English (Walk, Hit, Innings Pitched) where the bases for balls and hits connected are added and divided by the innings pitched.

Other mathematical techniques have been used in order to measure performance, such as the Pythagorean theorem, which in baseball is applied to the sum of victories and losses that make up the universe of games played, to establish a logic correlation between wins and losses and the positive difference between runs scored; The purpose of this formula is to predict what the theoretical percentage of games won would be based on the runs scored and received.

Also, the José Antonio Salamanca (JAS) coefficient is used, which is the most used and widespread system in Cuba to carry out a statistical evaluation of the performance in the competition of baseball players in honor of its inventor; his analysis is based on the three fundamental areas of the game (offense, defense and pitchers).

The offense is subdivided into touch and strength hitters; defense, receivers, infielders and outfielders and finally, pitchers are analyzed as starters and relievers. This system has as a deficiency the approach it gives to the offense, since the importance for modern baseball

implies winning in the integrity of its players with the objective of successfully facing the different game situations that occur in a game of baseball is obviated. baseball (Pérez, 2007).

On the other hand, the presence of sabermetrics constitutes a powerful form of statistical analysis in baseball and focuses on how individual and collective player statistics affect a team's win-loss record. In this sense, the calculation of indicators is carried out, through established formulas, which yield percentages of each player and through a statistical analysis of historical records, the performance in the game, the subseries, the tournament and even the lifetime trajectory (Soto and González, 2015).

However, the analyzes that are made based on sabermetrics should be used based on the direction of the team in general and that, according to Paz and Martínez (2018), has some advantages, such as knowing how to place the players in the alignment, in correspondence with his performance; the selection of the throw according to the base runners, that is, knowing how to hit behind the runner; the selection of the offensive tactical plays, depending on the batters and game situation; greater communication between batter-coach-runner; as well as, to know in greater depth the offensive weaknesses of the team, to make a more objective selection of the pinch hitters according to the moment of the game and to increase the preparation of the technical group for each game.

Also, in Cuba, since the LIV National Baseball Series, the Balls in Play Analyzer and Strike Zone system are used. The first shows the graph of the distribution of throws for each hitter against right-handed pitchers and against left-handed pitchers, which is very useful for positioning players on defense. While the second allows manipulating information throw by throw. The latter constitutes the most innovative since it has provided the most compelling and up-to-date scientific results by offering objective and scientific answers to many questions that tended to be answered empirically (Ríos, Sánchez and Rojas, 2019).

On the other hand, García and Cordero (2019) determined indicators for the control and evaluation of the technical-tactical actions in baseball pitchers, in game situations that allow knowing the influence they exert on the performance of the pitcher to obtain good results and responds to the conditions demanded by contemporary baseball, this leads to improving the competitive results of the pitcher and the team in which he is inserted and allows coaches systematic control during training, which enables timely correction, from the planning of the training.

Although there are many coefficients used for the analysis of the statistical data collected in each baseball game and in each national series published in Cuba, it has been possible to verify that there is little evidence of the use of statistics, mainly hypothesis tests, to analyze a baseball team in a certain period of time; therefore, the objective of this article was to carry out a statistical analysis of the Vegueros team of Pinar de Río, based on the use of statistical tests. The fulfillment of the objective focused mainly on the descriptive analysis of the data from the batting and pitching departments in the national series in which the team participated up to the LX series. In addition, the results in terms of games won and lost and the positions reached in each series were analyzed.

Regarding this work, an analysis of the data was carried out, through three cuts in time. The first cut was made from the VII series, in which it appeared for the first time in national series, to the XXIV series; the second, from the XXV series, where a kind of final series or Playoff began to develop, up to the XL (the 20th century ends) and the third, from the XLI series to the LX.

For this, previous works, information and data disclosed by print, audiovisual or electronic media, both national and international, were reviewed. The originality of the study was reflected in the approach, criteria, conceptualizations, reflections, analysis, conclusions and recommendations, the analysis of documents was used for the interpretation of the statistics of the national series, corresponding to the team under study and the general statistics in the batting and pitching departments, where the average was calculated in each one and compared with the historical average of all the teams that participated in the series, in each stage, from the performance of the student's t parametric statistical test. of mean differences, in each of the indicated stages and in the departments analyzed.

As a result of the review of the data disclosed by print, audiovisual and electronic media, a database was prepared in an Excel book, where in each of the sheets, the general statistics by team were located and the corresponding calculations were made. Below are the values of the arithmetic mean (average) and standard deviation in the batting and pitching departments and the collective mean for each of the stages analyzed. (Table 1)

Table 1. - Measures calculated for the Vegueros team and collective in the stages analyzed in the batting department

Stages	Batting Department			
	Average	Standard deviation	Collective average for the stage	Deviation Collective standard for the stage
First stage (from the VII national series to the XXIV)	249	33.8	241	fifteen
Second stage (from the XXV national series to the XL)	300	17.1	278	16
Third stage (from the XLI national series to the LX)	286	26.5	285	10.9

The mean and standard deviation values were also calculated in the pitching department in the analyzed stages, these values are shown in the following table (Table 2).

Table 2. - Measures calculated for the Vegueros team, and collective by stages in the pitching department

Stages	Picheo Department			
	Average	Standard deviation	Collective average for the stage	Collective standard deviation for stage
First stage (from the VII national series to the XXIV)	2.60	0.62	2.96	0.42
Second stage (from the XXV national series to the XL)	3.24	0.66	4.27	0.81
Third stage (from the XLI national series to the LX)	3.71	0.88	4.50	0.53

As can be seen in table 1, in the case of batting, only in the first stage it was below the collective historical average; which demonstrated a superior result for the second and third stages. However, in the case of pitching, in all stages the average was below the historical average; this established superior results for this department. However, it could not be affirmed that the Vegueros team was superior in these stages, since the corresponding statistical tests would need to be carried out to find out if the results, whether lower or higher, differ significantly from the historical average in each of the departments. analyzed.

Below are the steps followed to carry out the student`s t test, in the first stage, in the batting department:

1. Problem Statement. The interest was to verify if the batting average of the Vegueros team was higher than the collective historical average in the stage analyzed.

2. Starting hypothesis.

   - Ho: The batting average of the Vegueros team was lower than the collective average in the stage analyzed.
   - H1: The batting average of the Vegueros team was higher than the collective average in the stage analyzed

3. Selection of the statistical assessment test. Since the problem posed was to determine if there are differences between the means in two stages, recorded on an ordinal scale with a small sample, it was decided to use the student's t test.
4. Set the level of significance. A significance level $\alpha = 0.01$ was assumed.
5. Definition of the rejection region. The statistical test of student`s t evaluation was compared with the probability values, which for a significance level $t_p \succ 2,58 = 0.01$ with 17 degrees of freedom, the following decision rule was established:

Reject Ho otherwise accept Ho.

6. Calculation of the test statistician: the calculation of Zp, from the data record to compare with the critical value in the normal probability distribution table, was carried out according to the formula (Equation 1):

$$t_p = \frac{\bar{X}_2 - \bar{X}_1}{\sigma \sqrt{\left(\frac{1}{N_1} + \frac{1}{N_2}\right)}}$$

(1)

Data:

$\bar{X}_1$ Collective batting average in the stage analyzed.

$\bar{X}_2$ Batting average of the Vegueros team in the stage analyzed.

S_1Collective standard deviation of batting in the analyzed stage.

S_2 Standard deviation of batting for the Vegueros team in the stage analyzed.

N_1 Collective sample size (analyzed stage)

N_2 Sample size (Vegueros team analyzed stage).

σ Population standard deviation.

To calculate the value of t, from the data it was first necessary to calculate the population standard deviation that was calculated from the expression (Equation 2), (Equation 3) and (Equation 4):

$$\sigma = \sqrt{\frac{N_1(S_1)^2 + N_2(S_2)^2}{N_1 + N_2 - 2}} \quad (2)$$

Solving we have that:

$$\sigma = \sqrt{\frac{18(33,8)^2 + 18(15)^2}{18 + 18 - 2}} = 26 \quad (3)$$

Now, the general expression could be used to calculate the value of t.

$$t_p = \frac{249 - 241}{26\sqrt{\frac{1}{18} + \frac{1}{18}}} = = \frac{8}{8,6} = 0.93 \quad (4)$$

The t value of the sample was lower than the t value of the theoretical distribution, therefore, the null hypothesis was accepted and the alternative hypothesis was rejected, in this case, it was established that in the analyzed stage there were no differences in the batting average of the Vegueros team, nor in the collective average.

For the rest of the stages, the steps of the tests were not proposed, since it would occupy a large space in this article, so a summary table of each test with the results obtained is presented below.

Table 3. - Results of the hypothesis tests for the batting department

Stages	Decision rule	Sample t value	Results
First stage (from the VII national series to the XIV)	Reject Null Hypothesis (Ho) if Zp> 2.58 otherwise accept Ho.	t= 0.93	Ho is accepted, there are no significant differences.
Second stage (from the XXV national series to the XL)	Reject Null Hypothesis (Ho) if Zp > 2.62 otherwise accept Ho.	t=3.6	Ho, the batting average of the Vegueros team, is rejected is superior at this stage
Third stage (from the XLI national series to the LX)	Reject Null Hypothesis (Ho) if Zp> 2.53 otherwise accept Ho.	t=0.15	Ho is accepted, there are no significant differences.

As can be seen in table 3, only in the second stage the batting average of the Vegueros team was higher than the collective average for that stage and although the averages of this team were higher, it could not be affirmed that there are significant differences in comparison with the collective average in each stage analyzed.

In the case of the pitching department, it was proceeded in the same way as in the batting department, with the particularity that in this case the lower the value, the more effective it was, so the test was directed to demonstrate which pitch was more effective. Therefore, the following hypotheses were made:

- **Ho:** The Vegueros team average was higher than the collective average, therefore less efficient.

- **H1:** The Vegueros team average was lower than the collective average, therefore more efficient.

The results achieved in the hypothesis tests in each of the stages analyzed are shown below (Table 4).

Table 4. - The results achieved in the hypothesis tests in each of the stages

Stages	Decision rule	Sample t value	Results
First stage (from the VII national series to the XIV)	Reject Null Hypothesis (Ho) if Zp < 2.58 otherwise accept Ho.	t= 2	Ho is rejected, the average of the Vegueros team is more efficient.
Second stage (from the XXV national series to the XL)	Reject Null Hypothesis (Ho) if Zp < 2.62 otherwise accept Ho.	t= 0.55	Ho is rejected, the pitching average of the Vegueros team is more efficient.
Third stage (from the XLI national series to the LX)	Reject Null Hypothesis (Ho) if Zp < 2.53 otherwise accept Ho.	t=0.59	Ho is accepted, there are no significant differences.

In the case of pitching, in the three stages analyzed the value of t calculated was lower than the theoretical one, therefore, it could be affirmed that the pitching average of the Vegueros team was more efficient than the collective one, which demonstrated the quality of this department throughout the national series.

To start any debate about the results of this team, it is necessary to know that the Vegueros team appears, for the first time as a team, in the VII national series, accompanied by another team with the name of Pinar del Río. Vegueros was the representative team of this Province where its best players played; suffice it to say that in the XIV national series, the team that bore the name of Pinar del Río was dissolved and was replaced by Forestales as a second team that accompanied the Vegueros team. These two teams remained until the national series number XXXI and it is from this series that it definitively took the name of Pinar del Río.

Also, reference is made to the fact that the Vegueros team was one of the two teams that represented the Province, between 1967 and 1992; between the VII and XIII series (1967-1974), together with the Pinar del Río team and between the XIV and XXXI series (1974-1992), together with the Forestales team. As of the 1992-1993 season, the Province began to be represented by a single team, with the name of Pinar del Río, that is why when the champion teams in national series since (1962-2021) are summarized on this site, they name Vegueros Baseball Club and Vegueros, from Pinar del Río.

It should be noted that there are many articles and comments that remain on social networks in favor of the teams that represented Pinar del Río in national series, of course as a counterpart to the national media that have almost always tried to obscure the results of this Province in Baseball. In this sense, (García 2019) refers to the substantial fanaticism and frivolous treatment that people from Pinar del Río receive in the sports media, especially baseball, and alleges that all the titles achieved in this sport are not taken into account.

On the other hand, (García 2021) also refers to the Vegueros/Pinar del Río team, as the representative team of a Province with a little more than 191 thousand inhabitants and one of the most economically backward and that the media do not count the titles that It has reached in national series counting the selective ones; It also refers to the number of titles obtained (17), counting the selective ones, the Caribbean series and the five number one positions.

Likewise, (Manzano 2015) states that when the Playoffs were inaugurated in 1985, the Vegueros team was present there with the second place, of the pair of subtitles that he conquered and during the 18 years that he survived in the National Series, that is to say that they won six pennants and a League title and yet, this splendid harvest could not be included in the biographical catalog of the cast with the name of the Province, due to the incongruous structural system of amateur Cuban baseball. The above evidences, even more, the treatment that has been given to this team in the media and in the National Baseball Directorate.

In this sense, (Martínez de Osaba 1921) refers to the Vegueros team as an irreplaceable name, due to the productive capacity of the area and its imprint on Cuban baseball and affirms that in its honor, all the broadcasters, federations and people in general, they should call Vegueros de Pinar del Río the first team that, as is logical, represents this Province since in those times its name arises from the fame that the tobacco produced by the owners of Las Vegas had achieved and that, in addition, these they were the main supporters of this team in the Vueltabajo region.

It can be affirmed that, from the X national series, the Vegueros team becomes the strongest team in the Province and obtains better results than the team that bears the name of Pinar del Río and even becomes the team which for the first time makes a national series crown is obtained. In the 14th national series, the Pinar del Río team dissolved and the Vegueros team continued, accompanied by a second team named Forestales.

The Vegueros team begins to show results in the XVII series where it obtains 1st place and also obtains the crown in the XX, XXI, XXIV, XXVI and XXVII series, which shows that the best players in the Province were concentrated in this team. Also, in the XXXII series, a single team with the name of Pinar del Río appears for the first time, obtains its first crown with that name in the XXXVI series and wins the XXXVII series, the L or Gold Series and the LII. These results demonstrate the quality of the Vegueros team as representative of Pinar del Río.

By way of conclusions, it was possible to demonstrate that the Vegueros team exhibits excellent results, endorsed both by the results of the a priori statistical data, and by the interpretation of the hypothesis tests carried out with the data from the batting and pitching departments in the three stages analyzed from the sixty-national series.

The results of the bibliographical review and the analysis of the prepared database offer sufficient support to indicate that the Vegueros baseball team, since the beginning of the national series, is the representative cast of the province of Pinar del Río and the proofs of Hypotheses made with the data from the batting and pitching departments demonstrate the high quality of the Vegueros team in national series.

REFERENCES

García, R. y Cordero, V. O. 2020. Instrumento para el control y evaluación del accionar técnico-táctico de los lanzadores de béisbol. Revista PODIUM, enero-abril 2020; 15(1): pp. 72-83. http://podium.upr.edu.cu/index.php/podium/article/view/869

García, A. (2019). Pinar del Río en series nacionales, la historia escondida. https://swingcompleto.com/pinar-del-rio-en-series-nacionales-la-historia-escondida/

Manzano, R. 2015. Vegueros de Pinar del Río. Sitio web Desde mi palco de fanático. Memorias y Reflexiones sobre el beisbol cubano. https://desdemipalcodefanatico.wordpress.com/equipos/vegueros-de-pinar-del-rio/

Martínez de Osaba, J. A. 2021. ¿Quién gana más en la pelota cubana? https://www.beisbolencuba.com/foros/serie-nacional-de-beisbol/quien-gana-pelota-cubana.html

Paz, C. I. y Martínez A. 2018. La sabermetría en función de la ofensiva del equipo Holguín en la LVI serie nacional de béisbol. Revista Deporvida. 15(38), pp.107-118. https://deporvida.uho.edu.cu/index.php/deporvida/article/view/707

Pérez, I. (2007). El Béisbol en Cuba. ISSN: 0718-4921. http://www.portaldeportivo.cl/larevista/.

Ríos, A., Sánchez, B. y Rojas, E. 2019. La Inteligencia Deportiva en el béisbol cubano Lecturas: Educación Física y Deportes, 23, (250). https://creativecommons.org/licenses/by-nc-nd/4.0/deed.es

Soto, C. y González, M. (2015) Sabermetría y nuevas tendencias en el análisis estadístico del juego de béisbol. *Retos* (28), pp. 122-127. https://www.redalyc.org/pdf/3457/345741428023.pdf

Conflict of interests:

The authors declare not to have any interest conflicts.

Authors' contribution:

The authors have participated in the writing of the work and analysis of the documents