Kurdish Studies
August 2024
Volume: 12, No: 5, pp 5125 -5134
ISSN: 2051-4883 (Print) | ISSN 2051-4891 (Online)
www.KurdishStudies.net

DOI: 10.53555/ks.v12i5.3304

Statistical Modeling Of The Performance Of All-Rounders In Female Test Cricket

Qamruz Zaman[1*], Seema Zubair[2], Mansoor Ahmad[3], Sundas Hussain[4], Sofia Gul[5], Syed Habib Shah, Abdur Rehaman Sabir[1], Neelam[1]

[1,6,7]Department of Statistics, University of Peshawar, Pakistan
[2]Department of Mathematics, Statistics and Computer Science, University of Agriculture Peshawar, Pakistan.
[3]Shaheed Benazir Bhutto Women University Peshawar, Pakistan
[4] College of Home Economics, University of Peshawar, Pakistan
[5] Institute of Numerical Sciences, Kohat University of Science and Technology, Pakistan

Abstract: This study aimed to identify factors affecting the performance of women all-rounders in test cricket and examine the impact of bowling, batting, and wicket-keeping all-rounders on match outcomes. Logistic regression was employed to compare the performance of all-rounders with non-all-rounders, focusing on binary outcomes. The Chi-square test, using R^2 for goodness of fit, helped in model selection. All-rounders were categorized into wicket-keeper, bowling, and batting roles, with multinomial logistic regression used to assess performance differences. Significant factors included home-ground advantage, batting average, number of wickets, and first-innings performance. Batting and bowling all-rounders outperformed wicket-keepers.

Keywords: Multinomial logistic regression, Chi-square test, Batting and Bowling of females, test match.

1. Introduction: Sports and Statistics are related to each other Statistics can be defined as the orderly assembling and inspection of numeric data to evaluate or find connections between events by applying different techniques and tools, thereby explaining, foreseeing, and influencing their occurrence. Statistics has applications in almost every field of research ranging from social sciences to numerical, physical, and biological sciences (Huck, S. W et. al. 1974). Literature has shown an immediate link between sports and statistics. In sports, different statistical methodologies are used to measure the performance of players, and teams are also ranked based on these methodologies (Albert and & Koning 2007). In cricket, statistics play a key role in determining different aspects of a player's performance due to the availability of the player's related data (Saikia, H; et. al. 2016).

Cricket is the second most prominent game in the world. Cricket is a game that is played between two teams of eleven players each, on a rectangular pitch in a prepared circular area. It is believed to have been played in the 16th century for the first time in England. By the end of the 18th century, it becomes the national game of England (Khairullah, M. 2011). International cricket matches were being held around the mid-19th century and presently more than 100 countries of the world play this game. Nowadays, cricket is played in three different formats, i.e. One Day International (ODI), T-20, and test cricket. Test cricket is the conventional and lengthy form of this game. In this format, the game spans five days and each team have to play two innings one after another. To make cricket more exciting and energetic, the second version of cricket i.e., ODI was launched around the mid-19th century. In One-day cricket, each team has to play 50 over innings before being bolded out. In 2003, a new format of cricket was introduced in England, T20, the most thrilling and passionate form of cricket (Albert, J; Glickman, et al. 2017). Numerous games are being played all around the globe at national and international levels in both T-20 and one-day formats. It is widely known that male and female players in crickets vary concerning physiological and physical capabilities like strength, stamina, and power (Munro, C. E; Christie, 2018). It is a fact that very little has been written about the history of women's cricket. The phenomenon of women's cricket remains largely untouched. The first women's cricket match took place in 1745 and during the 19th century expanded its wings (Nicholson 2015).

1.2 History of Cricket. Cricket, believed to have been first played in 16th century England, gained the status of England's national game by the late 18th century. The expansion of the British Empire spread cricket globally, leading to international matches by the mid-19th century. Today, over 100 countries play cricket, making it the second most popular sport worldwide after soccer. Cricket is a bat-and-ball game played between two teams of eleven players on a rectangular pitch within a circular field.
1.3 Formats of Cricket. At the international level, three formats of cricket are being played these days: Test cricket, One Day International cricket and T-20 Cricket

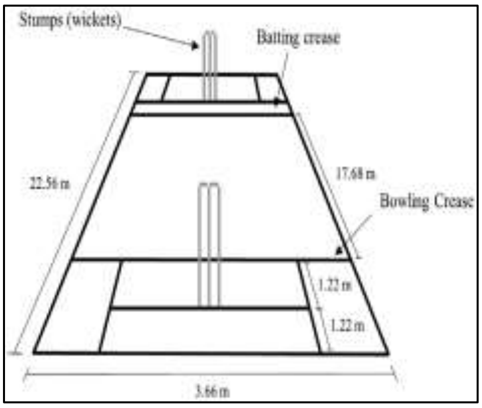
1.3.1 Test Cricket. Test cricket is the conventional and lengthy form of cricket that spans five days in which each team has to play two innings one after another. Frequently the results of the test match in draws. Sometimes teams bat intends to achieve a draw when they worry, they are on the edge of losing the game, and batting intentionally assists teams to save from dismissal and elongates the period of innings.

1.3.2 One Day International. To make cricket more exciting and energetic, a second version of cricket was launched around the mid-19th century. In One-day cricket, each team has to play 50 overs innings before being bold out. One-day cricket is primarily

a form of confined overs and each side has only one inning to bat. There is less chance of draws in limited-overs cricket. One-day cricket is more competitive with batsmen playing shorts with full energy with a high risk of losing wickets.
1.3.3 T-20 Cricket. T20, the most thrilling and passionate form of cricket was played for the first time in 2003 in domestic tournament in England.

1.4 Women Cricket. Cricket, traditionally seen as a gentleman's game, has seen a rise in women's participation, with numerous national and international T-20 and One Day matches being played. However, many of these women's games are not televised, and their statistics are poorly recorded, creating a research gap. While male and female cricketers differ in physiological and physical capabilities such as strength, stamina, and power, substantial research on male players could help evaluate female players' performance. Women's cricket dates back to the first recorded match in 1745. The sport began to grow significantly in the 19th century, with the establishment of the White Heather Club by eight ladies in Yorkshire in 1887, followed by the formation of several women's clubs. In 1890, the English Cricket and Athletic Association formed two teams of lady professionals known as 'The Original English Lady Cricketers'. Women's international cricket tours began in 1934-35 when a team of female cricketers from England visited Australia, playing three matches in a test series, marking the start of international women's cricket.

1.4.1 Performance of Female all-rounder in cricket. The research addresses a significant gap in cricket literature by focusing on female all-rounders in test cricket. It highlights their crucial role in batting, bowling, and wicket-keeping, which are essential for team balance across all game formats. Despite their importance, cricket coaching has traditionally emphasized batting and bowling talent over fielding skills. The study employs statistical methods like logistic regression and multinomial regression to analyze the performance dynamics of female all-rounders, aiming to provide comprehensive insights into their capabilities and contributions within cricket.
1.5 Pitch information. There is a 22-yard-long pitch at the center of the field where most of the actions takes place. There will be wickets on both ends of the pitch which has three wooden stumps and two cross pieces called the bails.



2. Literature Review
By using the Naïve Bayes classification, (Saikia and & Bhattacharjee 2010) put all-rounders in four categories including performer, underperformer, batting, and bowling all-rounder. The research received 35 all-rounders who participated in the first three seasons of the Indian Premier League (IPL), the research established a classification model that was used to suggest six new all-rounder groups.
The wicketkeeper seems to be an extremely important supporter of even the fielding hand. This is the only player who could have been actively engaged in every inning delivery. There seems to be a lack of wicket-keeping literature, with much of this work Concentrating on bowling and batting quickly. This review mainly deals with the identification and conditioning of performance requirements associated with wicket-keeping (MacDonald 2013).

(Wormgoor 2010) applied Pearson's correlation method and estimated that wide front leg expansion during ball release, shoulder adjustment in a cross direction moving away from the batsman hitting on the front foot, wider ankle height as well as increased shoulder extension throughout delivery stride led to an especially in increasing ball release speed.
(Shah, S et al. 2017) examined the performance of cricket players while using factor analysis on data from 85 batsmen as well as 85 bowlers shortlisted from IPL (2016) and 95 batsmen as well as 95 bowlers from those in the ICC World Cup (2015). The study also showed that bowling performance had been influenced by batting abilities respectively.
Using one-way ANOVA, (Malhotra, A; & Krishna, 2017), analyzed increasing speed as well as exposure for down-outs order batters to pace. The study found that fast bowlers played superior in terms of strike rate and average but there have been no significant differences observed in "CBR" (combined bowling rate) as well as economy rate

(Wickramasinghe 2014) , has examined batsmen's success during test cricket. Utilizing three-step hierarchical linear models it is observed that somehow a player with a stronger ranking team substantially improved the player's performance.
(Allsopp and & Clarke 2004) developed various variables which often influence the outcome of test matches. Using the multinomial regression technique, it has been discovered that factors such as batting first and strong bowling, first-inning

score lead, the batting lineup, and the advantage of playing at the home ground have been found to have significant factors that influence the result of the game. The ability of player footwork and timing play a crucial role in evaluating the performance of a batsman.

(Singh, A; & Sharma, R. 2017), have been using the product moment correlation technique used by Pearson to evaluate the ball sense ability. The results show that brain-perceived visual information is an important component of ball prediction. Expert batsmen utilize visual information for an effective bat-ball interception with appropriate timing.
(Shah 2017), had already discovered a new philosophy for assessing a batsman and bowler's performance. A batsman's and sometimes a bowler's performance measurement was calculated by taking into account the quality of bowling or batting of batsman on the pitch respectively. The individual performance of batsmen as well as bowlers against either the quality bowler or batsman was aggregated for measuring the performance index of batsmen and bowlers.
Muller & Abernethy (2006) using liquid crystal spectacles investigated batsmen's capabilities confronting bowlers. The study suggested that pre-bounce ball flight was effective for batters to make efficient bat-ball contact (Malhotra, A; & Krishna, 2017).

(Sumner, J; & Mobley, M. 1981), studied every test cricket match from 1877 to 1980 and used two sample tests to measure the proportion of legs before wicket dismissals for both (home as well as away) teams for each testing country. Their analysis showed that more often in Australia, India, and Pakistan away batsmen had been out leg before wicket. We proved that dominants in Pakistan and India. Kimber & Hansford (1993), claimed that sometimes an unrealistic parametric argument depends on conventional batting averages. They suggested a non-parametric methodology for measuring batting performance based on runs. The method has already been used at different cricket levels to a very large sample of the best (Kimber and C. 1993).

The leg-before-wicket (LBW) decisions have historically been a source of controversy in test cricket matches, especially those played in Australia. Crows and Middeldrop (1996) conducted a study analyzing LBW dismissal rates in test series matches between Australia and visiting teams across various Australian grounds from 1977 to 1994. They utilized a generalized linear model to examine innings involving the first six batsmen. The study suggested that there was perceived bias in umpiring decisions favoring Australia during these test series, highlighting ongoing debates and scrutiny surrounding LBW judgments in cricket (Crowe 1996).

In the 1990s, De Silva and Swartz (1997) statistically studied 427 international One-Day (ODI) cricket matches. First of all, they reached two general conclusions from their finding that there is no major advantage in winning the coin toss. The second home-ground effect increases the probability of winning log odds by approximately 0.5 (De Silva, B. M., et al. 2001) .
Different authors have also used logistic regression including Koning (2000) in the Netherlands, employing an ordered probit model regarding soccer tests (Koning, R. H. 2000). Willoughby (2002) used conditional logistic regression to evaluate the result (win, draw, or loss) of Canadian Football League (CFL) Football games (Willoughby 2002).

2.1 Aims and Objectives

1. To determine the factors responsible for the performance of the women all-rounders in test cricket
2. To study the effect of bowling, batting, and wicket-keeping all-rounders on match result
3. To study the overall performance of women in test cricket

3. Research Methodology

This study aims to identify the key factors affecting the performance of female all-rounders in Test cricket and their influence on bowling, batting, and wicket-keeping. It also evaluates the overall performance of women all-rounders and their impact on match outcomes in Test cricket.

3.1 Universe of the Study. This study encompasses all women's Test cricket matches played from 1934 to 2017, totaling 139 matches. Of these, 51 matches had definitive results, 88 were drawn, and none ended in a tie. All-rounders frequently earned "Player of the Match" honors. However, for matches where this wasn't identified, any bowler taking at least 3 wickets in a single inning or any batter scoring over 50 runs was considered "Player of the Match." The data, sourced from ESPN and Cricinfo, includes additional information such as pitch analysis, toss results, ball-by-ball commentary, match outcomes, victory margins, best players, and series conditions.

3.2 Sequential Sampling. The sample size will be determined using a sequential sampling technique, a complex model where the basic sample size evolves based on ongoing results. This approach offers numerous advantages, including flexibility in sample size selection and scheduling. It is not time-consuming, costly, or labor-intensive, and allows for adjustments during the initial study phase to improve the research method (Etikan, I et. al. 2016).

3.3 Study Design. Data was gathered from ESPN and Cricinfo, focusing on identifying factors influencing the performance of women all-rounders in Test cricket. The study employed a two-stage statistical analysis. First stage Logistic Regression was used to compare the overall performance of all-rounders against non-all-rounders, leveraging the binary nature of the response variable. The Chi-Square Test was applied with R^2 as the goodness-of-fit criterion to select the best model compared to a no-

intercept model. Second Categorization of All-rounders was divided into three categories: wicket-keeper all-rounders, bowling all-rounders, and batting all-rounders. Multinomial Logistic Regression was used to evaluate the performance of each category. The Chi-Square Technique was used as the goodness-of-fit criterion for model selection, compared to a no-intercept model
.

3.4.1 Statistical Techniques. To achieve the study's objectives, data on female Test cricket all-rounders was collected and analyzed using SPSS version 23. Initially, descriptive statistics were computed for all observed variables and visualized with graphs and diagrams, such as pie charts and bar charts. A binary logistic regression model was then fitted to evaluate the overall performance of both all-rounders and non-all-rounders. Subsequently, all-rounders were classified into three categories: bowling, batting, and wicket-keeping. To assess the performance of each category, a multinomial logistic regression model was applied. These statistical techniques were employed to thoroughly investigate the factors affecting the performance of female all-rounders in Test cricket.

3.4.2 Multinomial Logistic Regression. Multinomial logistic regression was employed to assess the performance of all-rounders in women's Test cricket, focusing on "Player of the Match" data from 139 matches. The study observed that all-rounders frequently received this accolade. The model categorized all-rounders into bowling, batting, and wicketkeeper categories, predicting their probabilities based on various explanatory variables. Unlike linear regression, which requires continuous variables, multinomial logistic regression handles discrete outcomes and does not assume linearity or normality. It involved simultaneous estimation of multiple binary logistic models, comparing bowling and batting all-rounders to wicketkeeper all-rounders, providing insights into factors influencing their performance in women's Test cricket.

$$logit(P) = \beta_0 + \beta_1 Y_1 + \beta_2 Y_2 + \cdots + \beta_k Y_k. \tag{1}$$

$$logit(P) = \beta_0 + \sum_{i=1}^{k} \beta_i Y_i \tag{2}$$

Where P shows the probability of the Occurrence of the characteristic of interest. The logit transformation is well-defined as the logged odds:

$$Odds = \frac{P}{1-P} \tag{3}$$

$$logit(P) = \ln\frac{P}{1-P} = e^{\beta_0 + \beta_1 Y_1 + \beta_2 Y_2 + \cdots + \beta_k Y_k} \tag{4}$$

It is clear that, when a variable Y_i enhances by 1 unit, with all the other variables holding fixed, then the odds will move by a factor $e^{\beta i}$. This factor e^b is the odds ratio for the regressor variable Y_i and it gives the relative amount by which the odds of the outcome increased or decreased when the value of the independent variable is increased by 1 unit. Nominal multinominal logistic regression models are described as follows:
The first binary logistic regression model for the probability of bowling all-rounder and wicketkeeper is given by

$$\log\left[\frac{P\left(Y=\frac{1}{x}\right)}{P\left(Y=\frac{2}{x}\right)}\right] = \log\left(\frac{p_1}{p_2}\right) = \alpha_1 + \beta_{11}x_1 + \beta_{12}x_2 + \ldots + \beta_{1k}x_k = g_1(x) \tag{5}$$

And

$$\log\left[\frac{P\left(Y=\frac{3}{x}\right)}{P\left(Y=\frac{2}{x}\right)}\right] = \log\left(\frac{p_3}{p_2}\right) = \alpha_2 + \beta_{21}x_1 + \beta_{22}x_2 + \ldots + \beta_{2k}x_k = g_2(x) \tag{6}$$

Then the probabilities of each category given the covariate vector are

$$p_1 = \frac{e^{g_1(x)}}{\left(1+e^{g_1(x)}+e^{g_2(x)}\right)},\ p_2 = \frac{1}{\left(1+e^{g_1(x)}+e^{g_2(x)}\right)},\ p_3 = \frac{e^{g_2(x)}}{\left(1+e^{g_1(x)}+e^{g_2(x)}\right)}.$$

Where p_1 represents the probability of batting all-rounder, p_2 represents the probability of bowling all-rounder and p_3 represents the probability of wicketkeeping an all-rounder. In this study, to assess model fit Akaike information criterion and Bayesian information criterion, the classification table for goodness of fit and explanatory power Nagelkerke R^2 will be used.
The Akaike information criterion can be defined as

$$AIC = 2K - 2\ln(L) \tag{7}$$

where K is the number of estimated parameters in the model and L is the maximum likelihood value of the estimated model. The Bayesian information criterion (Schwarz, 1978) is defined as

$$BIC = \ln(n)\,K - 2\ln(L) \tag{8}$$

Where "n" is the total sample size.
According to Claeskens and Hjort "For selecting a model among a list of candidates, AIC is among the most popular and versatile strategies". The lower value of AIC is better while the individual value of AIC is meaningless. We prefer to use the AIC and it will be adopted in this study. This is because the AIC is more associated with (average) prediction and precision

of estimates than BIC. To summarize the results of a fitted logistic regression model by a classification table. It shows the predictive success of the fitted logistic regression model by classifying correct and incorrect predictions. Some statistics can be used to measure the strength of the association between the outcome variable and the explanatory variables, and on the other hand for explanatory power of the covariates in the model, Nagelkerke R^2 will be used. Nagelkerke R^2 is given by

$$R^2 = 1 - exp\{(-2(L_1 - L_0)/n)\}/\{(1 - exp(2L_0)/n)\} \quad (9)$$

Where L_1 represents the log-likelihood for the model with covariates, L_0 represents the log-likelihood for the model with no covariates, and n is the number of observations; broadly speaking Nagelkerke R^2 for a generalized linear model is the percentage of variability in the outcome that is explained by the covariates in the model

4. Statistical Analysis

Statistical analysis of female all-rounder cricket players in Test matches utilized data from Cricinfo, employing descriptive statistics (tabulation, graphs, and Summary statistics) and inferential statistics. Logistic regression differentiated all-rounders from non-all-rounders, while multinomial logistic regression categorized all-rounders into batting, bowling, and wicket-keeping types.

4.1 Descriptive Statistics

In this section, descriptive statistics were employed to illustrate key variables related to the status of female Test cricket players. Specifically, the table highlights that out of 212 female players, 145 (68.4%) competed in Test matches held at their home ground, while 67 players participated in matches held at venues not associated with their team's home ground.

Table 4.1: Home Ground

	Frequency	Percent
Yes	145	68.4
No	67	31.6
Total	212	100.0

The table categorizes female players in Test cricket as either all-rounders or non-all-rounders. It shows that out of 212 female players, 137 (64.6%) are all-rounders, while 75 (35.4%) are non-all-rounders.

Table 4.2: Allrounder Status

	Frequency	Percent
Yes	137	64.6
No	75	35.4
Total	212	100.0

Table 4.3 describes the type of all-rounder female players in test cricket. The status of allrounder female players is categorized as Bowling allrounder, batting allrounder, and wicket keeping (WK) allrounder player in test cricket. The table defines that out of 212 female allrounders, there are 88 bowling allrounders, 118 batting allrounders, and 6 wicketkeeper allrounders. The percentages of the three categories were found as 41.5, 55.7, and 2.8 percent respectively.

Table 4.3: Allrounder type

	Frequency	Percent
Bowling	88	41.5
Batting	118	55.7
WK	6	2.8
Total	212	100.0

The table 4.3 depicted in the following sector diagram.

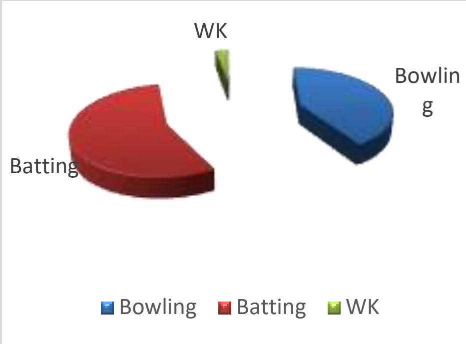


Fig 4.1: Sector Diagram for Allrounder status. The sector diagram presents a graphical presentation of the all-rounder type of female players in test cricket. The largest sector of the pie indicates that the largest portion of the sample data belongs to batting allrounder players in female test cricket. While the smallest sector of the pie defines the very few WK allrounders in female test cricket. The summary statistics for female test cricket all-rounders' scores across three innings are as follows: The mean scores were 224.04, 230.23, and 157.33 for the first, second, and third innings respectively. Median scores were 222.50, 206, and 164 for the same innings. Standard deviations were 82.89, 107.38, and 78.33, indicating variation in scores. The range of scores for each inning was 376, 534, and 344, highlighting the spread of scores across matches.

Table 4.4: Summary Statistics of scores in three innings

Statistics	1st inning result	2nd inning result	3rd inning result
	212	212	212
Mean	224.04	230.23	157.33
Median	222.50	206.00	164.00
Std. Deviation	82.896	107.389	78.331
Range	376	534	344
Minimum	38	35	50
Maximum	414	569	344

Table 4.5 provides a comprehensive overview of key performance metrics for female all-rounders in test cricket matches. It details several important factors: the average number of overs bowled by bowler all-rounders per match, calculated at 321.96 with a standard deviation of 49.791; the average number of maiden overs bowled, which stands at 104 per match with a standard deviation of 28.31; and the average number of extras conceded per match by female bowler all-rounders, reported as 33 with a standard deviation of 14.684. These statistics highlight the workload and effectiveness of bowler all-rounders in the context of test cricket, illustrating their contribution through overs bowled and additional deliveries such as maidens and extras. The data underscores the variability and consistency in their performance across different matches, essential for assessing their impact on game outcomes and team strategies.

Table 4.5: Summary Statistics of various factors

	Total number of overs bowled in a match	Number of medin overs in a match	Number of extras (no ball and wide ball) in the match	Total no: overs bowled in a match
N	211	211	211	211
Mean	321.96	104.01	33.08	321.96
Median	322.00	103.00	31.00	322.0
Std. Deviation	49.791	28.310	14.684	49.791
Range	302	137	67	302
Minimum	148	50	10	148
Maximum	450	187	77	450

To compare the batting average between the female test players i.e. allrounder and non-allrounder, a two-sample independent sample t-test was used. The following table 4.6 presents the two-sample t-test output.

Table 4.6; Two sample Independent Samples t-Test

		t-test for Equality of Means						
		T	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference Lower	Upper
Batting Average	Equal variances assumed	-2.275	210	.024	-4.37259	1.92175	-8.16099	-.58419

Table 4.7 presents descriptive statistics comparing the batting averages of three categories of female all-rounders in test cricket: bowler all-rounders, batting all-rounders, and wicketkeeper (WK) all-rounders. For bowler all-rounders, the batting average is 11.92 with a standard deviation of 4.98, and a 95% confidence interval ranging from 10.867 to 12.97. Batting all-rounders show an average of 33.402 with a standard deviation of 10.43, and a 95% confidence interval from 31.51 to 35.30. WK all-rounders have a batting average of 28.356 with a standard deviation of 7.3, and a 95% confidence interval ranging from 22.51 to 26.17. These statistics highlight the varying batting capabilities across different roles within female test cricket, emphasizing the significant differences in batting performance among bowler, batting, and WK all-rounders.

To

Descriptive Batting Average

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
Bowling	88	11.9229	4.98288	.53118	10.8671	12.9786	1.40	22.84
Batting	118	33.4026	10.43196	.96034	31.5007	35.3045	13.47	67.57
WK	6	28.3564	7.30042	2.98038	20.6951	36.0177	21.03	40.28
Total	212	24.3437	13.51069	.92792	22.5145	26.1728	1.40	67.57

compare the batting average among the allrounder type of female test cricketers i.e. bowling allrounder, batting allrounder and WK allrounder, one-way analysis of variance (ANOVA). The following table 4.8definde the output of one-way ANOVA.

Table 4.8; Analysis of variance (ANOVA) output Batting Average

	Sum of Squares	Df	Mean Square	F	Sig.
Between Groups	23356.480	2	11678.240	161.008	.000
Within Groups	15159.222	209	72.532		
Total	38515.702	211			

The above table defines the value of F-statistic as 161.008 with the p-value as 0.000. The p-value is smaller than 5% level of significance which rejects the null hypothesis of equal average batting average among the batting, bowling and WK allrounder of female test cricketers. It is concluded that there is a significant difference in the batting average of three categories of the female allrounder test match players at a 5% level of significance.

4.2 Logistic Regression Analysis

The logistic regression model was employed to assess the impact of factors on the performance of female Test cricket players, categorized as all-rounders and non-all-rounders. Table 4. provides a summary of the model's findings and results.

Table 4.9: Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	195.290[a]	.313	.431

a. Estimation terminated at iteration number 5 because parameter estimates changed by less than .001.

Table 4. presents the model summary of a logistic regression analysis where the success category is defined as the status of being an all-rounder player, with not being an all-rounder categorized as a failure. It includes Cox & Snell R Square and Nagelkerke R Square as measures of the model's goodness of fit. These statistics indicate that the variation in the status of players, ranging from 31.3% to 43.1%, can be explained by the variations in the explanatory variables included in the model. Table 4.b illustrates the results of a logistic regression model assessing the contribution and statistical significance of explanatory variables. The model demonstrated statistical significance with a chi-square statistic of 15.014 and 8 degrees of freedom, yielding a p-value of 0.00005, indicating significance at the 5% level. These findings confirm the overall relevance and impact of the explanatory variables on the logistic regression model.

Table 4.10: Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1[a]	Homeground	2.746	.379	2.532	1	.0000	1.574
	Batting_Average	.027	.018	2.153	1	.0142	1.027
	No_of_wkts	.130	.086	2.258	1	.0133	1.138
	Inning_1st	.004	.002	2.826	1	.0413	1.004
	Inning_2nd	.002	.002	.677	1	.4117	1.002
	No_of_extras	-.003	.012	.071	1	.7902	.997
	Constant	-6.559	1.095	35.867	1	.0000	.001

Variable(s) entered on step 1: Homeground, Batting_Average, No_of_wkts, Inning_1st, Inning_2nd, No_of_extras.
In Table 4. b, regression coefficients, standard errors (S.E.), Wald statistics, and significance (p-values) are reported for each variable in the logistic regression model. Significant variables such as home-ground effect, batting average, number of wickets taken by the player's team, and performance in the first inning of all-rounders are highlighted ($p < 0.05$). These variables' odds ratios (Exp(B)) indicate their impact on the likelihood of being an all-rounder compared to a non-all-rounder in women's Test cricket. For instance, the odds of a female all-rounder benefiting from home-ground advantage are 1.574 times higher than for non-all-rounders. A higher batting average increases the log odds of being an all-rounder by 0.027, indicating a 1.027 times greater likelihood for all-rounders. Similarly, a unit increase in wickets taken by the player's team correlates with 1.138 times higher odds of being an all-rounder. Additionally, a one-unit increase in first-inning score boosts the log odds of player performance by 0.004, suggesting a 1.004 times higher likelihood for all-rounders. These findings underscore the predictive power of these variables in determining the probability of female all-rounders' performance in Test cricket.

4.2 **Multinomial Regression**. In this section, a multinomial regression model was applied to analyze the performance characteristics of female players in Test cricket, focusing on their batting, bowling, and wicket-keeping abilities. The model categorized wicket-keeping ability as the reference category and predicted the probabilities of bowling and batting abilities among all-rounders. Table 4.c presents the goodness-of-fit results, indicating how well the model aligns with the data.

Table 4.11: Goodness-of-Fit

	Chi-Square	df	Sig.
Pearson	87.987	366	0.420
Deviance	102.952	366	0.021

Table 4.d provides model fitting information using Pearson chi-square and Deviance chi-square statistics to assess goodness of fit. The Pearson chi-square statistic of 87.987 with a p-value of 0.42 suggests the model fits the data well, as the p-value exceeds the 5% significance level.

Table 4.12: Model Fitting Information

	Model Fitting Criteria	Likelihood Ratio Tests		
Model	-2 Log Likelihood	Chi-Square	Df	Sig.
Intercept Only	317.257			
Final	54.338	262.919	14	.000

The significance of the regression coefficients in the final model is confirmed by a Chi-square test statistic of 256.919 with a significance value (sig) of 0.000, indicating statistical significance at the 5% level. The model predicts the response variable more effectively than an intercept-only model. The Pseudo R-

Table 4.13: Pseudo R-Square

Cox and Snell	.013
Nagelkerke	.038
McFadden	.025

square statistics (Cox and Snell, Nagelkerke, and McFadden) measure the amount of variation explained by the multinomial regression model. Values of 0.013, 0.038, and 0.025 respectively indicate the proportion of variability in the response variable that can be attributed to the independent variables included in the model.

The following table indicates the likelihood ratio table which explains the individual significance of explanatory variables included in the model.

Table 4.14: Likelihood Ratio Tests

	Model Fitting Criteria	Likelihood Ratio Tests		
Effect	-2 Log Likelihood of Reduced Model	Chi-Square	Df	Sig.
Intercept	54.338[a]	.000	0	.
Batting_Average	147.752	93.414	2	.000
no_of_catches	87.063	32.725	2	.000
Inning_1st	56.747	25.409	2	.000
match_runs	70.296	15.958	2	.000
medin_overs	56.686	2.348	2	.309
Wicket_get	59.443	5.105	2	.078
Homeground	62.934	8.595	2	.014

The chi-square statistic is the difference in -2 log-likelihoods between the final model and a reduced model. The reduced model is formed by omitting an effect from the final model. The null hypothesis is that all parameters of that effect are 0.
. This reduced model is equivalent to the final model because omitting the effect does not increase the degrees of freedom.

The table defined that batting average, no: of catches, playing in 1st inning, matches run and the home ground effect observed statistically significant as their p-value was found less than 5% while median overs and wickets found statistically insignificant with the p-value more than 5 % level of significance.

Table 4.15: Parameter Estimates

Alrounder_type[a]		B	Std. Error	Wald	df	Sig.	Exp(B)	95% Confidence Interval for Exp(B)	
								Lower Bound	Upper Bound
Bowling	Intercept	67.387	23.256	8.397	1	.004			
	`Batting_Average	.290	.321	3.816	1	.003	1.048	.399	1.403
	no_of_catches	-10.001	8.954	1.247	1	.264	4.537E-5	1.083E-12	1900.161
	Inning_1st	.028	.025	5.322	1	.002	.622	.926	1.020
	match_runs	-.104	.070	2.185	1	.139	.902	.786	1.034
	medin_overs	.109	.116	.883	1	.347	1.115	.889	1.398
	Wicket_get	.196	.204	5.920	1	.007	1.022	.551	1.227
	[Homeground=1]	-31.327	1.256	622.433	1	.000	1.236	.381	2.910
	[Homeground=2]	0[b]	.	.	0	.	.	.	.
Batting	Intercept	54.538	23.000	5.623	1	.018			
	Bowling_Average	.219	.309	5.503	1	.048	1.245	.680	2.279
	no_of_catches	-9.460	8.928	1.123	1	.289	7.792E-5	1.959E-12	3099.582
	Inning_1st	-.129	.024	5.413	1	.004	1.972	.927	2.019
	match_runs	-.061	.069	.782	1	.377	.941	.822	1.077
	medin_overs	.098	.116	.723	1	.395	1.103	.880	1.383
	Wicket_get	.206	.204	7.025	1	.451	.814	.546	1.213
	[Homeground=1]	-.234	.425	5.624	1	.004	1.258	.234	2.015
	[Homeground=2]	0[b]	.	.	0	.	.	.	.

a. The reference category is: WK.

a. This parameter is set to zero because it is redundant.

The multinomial regression analysis compared female all-rounders in Test cricket, focusing on bowling and batting categories against wicket-keeping as a reference. Significant variables included batting average, first-inning play, wickets taken by bowlers, and home ground effect, identified at a 5% significance level. Results indicated that female bowler all-rounders perform 1.048 times better in batting average and 1.022 times better in taking wickets compared to wicket-keeping all-rounders, and perform 1.236 times better in their home ground. Similarly, female batting all-rounders showed 1.245 times better batting average, 1.972 times better first-inning performance, and 1.258 times better home ground performance. These findings underscore the impact of these variables on female all-rounders' performance in Test cricket.

5. Discussion

The study focused on analyzing various characteristics of female test cricketers, categorizing them as all-rounders and non-allrounders. Data on batting, bowling, and wicket-keeping abilities were sourced from Cricinfo, and both descriptive and inferential statistics were employed to achieve the study's objectives. The analysis revealed that out of 212 female players, 145 (68.4%) played test matches on their home ground, suggesting a significant home-ground advantage. The players were categorized into all-rounders (137 players, 64.6%) and non-allrounders (75 players, 35.4%). Among the all-rounders, there were 88 bowling all-rounders (41.5%), 118 batting all-rounders (55.7%), and 6 wicket-keeping all-rounders (2.8%). A sector diagram visually represented these categories, showing batting all-rounders as the largest group and wicket-keeping all-rounders as the smallest.The study computed summary statistics for the scores of all-rounder players across three innings, with average scores of 224.04, 230.23, and 157.33, and median scores of 222.50, 206, and 164 for the first, second, and third innings, respectively. The standard deviations were 82.89, 107.38, and 78.33, while the score ranges were 376, 534, and 344 for the respective innings. For bowling all-rounders, the average number of overs bowled was 321.96 with a standard deviation of 49.791, average maiden overs per match were 104 with a standard deviation of 28.31, and average extras per match were 33 with a standard deviation of 14.684.

The study compared the batting averages of all-rounders and nonallrounders using a two-sample independent t-test, finding a significant difference (p-value = 0.024). A one-way ANOVA was conducted to compare the batting averages among the three types of all-rounders, showing significant differences (F-statistic = 161.008, p-value = 0.000). The batting averages for bowling all-rounders, batting all-rounders, and wicket-keeping all-rounders were 11.92, 33.402, and 28.356, respectively. A logistic regression model was used to determine the impact of various factors on a player's status as an all-rounder, considering the home-ground effect, batting average, number of wickets taken by the player's team, and first-inning performance. These

factors were found to significantly influence whether a player was categorized as an all-rounder, with the chi-square statistic being 15.014 (p-value < 0.00005). The odds ratios indicated that all-rounders had 1.574 times higher odds of playing on their home ground, 1.027 times higher odds of a better batting average, 1.138 times higher odds of taking more wickets, and 1.004 times higher odds of performing better in the first inning.
The multinomial regression model analyzed the performance characteristics of all-rounders, with wicket-keeping ability as the reference category. Key findings included that batting all-rounders had 1.048 times higher odds of a better batting average, 1.972 times better performance in the first inning, and 1.258 times better performance on their home ground compared to wicket-keeping all-rounders. Overall, the study provided a comprehensive analysis of the performance and characteristics of female test cricketers, highlighting significant differences between all-rounders and non-allrounders, and among different types of all-rounders.

6. Conclusion

This study analyzed data from Cricinfo on female all-rounders and non-all-rounders in Test cricket, focusing on their batting, bowling, and wicket-keeping abilities. It categorized players as all-rounders (further divided into batting, bowling, and wicket-keeping all-rounders) and non-all-rounders. Summary statistics, including mean, median, standard deviation, and range, were calculated for scores across three innings. Key performance factors such as total overs bowled, maiden overs, extras, and overall performance were assessed. Comparisons using ANOVA revealed significant differences in batting averages among the three all-rounder types, while logistic regression identified significant factors like home-ground effect and batting average impacting performance. Multinomial regression showed that batting and bowling all-rounders outperformed wicket-keeping all-rounders.

References

1. Albert, J, and R. H. (Eds. & Koning. 2007. *Statistical thinking in sports.* . CRC Press.
2. Albert, J; Glickman, et al. 2017. *Handbook of Statistical Methods and Analyses in Sports.* CRC Press.
3. Allsopp, P. E, and S. R. & Clarke. 2004. "Rating teams and analysing outcomes in one-day and test cricket. ." *Journal of the Royal Statistical Society: Series A (Statistics in Society),* 167(4), 657-667.
4. Crowe, S. M., & Middeldorp, J. 1996. "A Comparison of Leg Before Wicket Rates between Australians and their Visiting Teams for Test Cricket Series P layed in Australia, 1977–94. ." *Journal of the Royal Statistical Society: Series D (The Statistician)* 45(2), 255-262.
5. De Silva, B. M., et al. 2001. "Applications: Estimation of the Magnitude of Victory in One-day Cricket RMIT University, Mayo Clinic Rochester and Simon Fraser University." *Australian & New Zealand Journal of Statistics* 43(3), 259-26.
6. Etikan, I et. al. 2016. "Comparision of snowball sampling and sequential sampling technique. ." *Biometrics & Biostatistics International Journal* 3(1), 1-2.
7. Huck, S. W et. al. 1974. *Reading statistics and research (pp. 74-102).* New York: Harper & Row.
8. Khairullah, M. 2011. "A novel text steganography system in cricket match scorecard." *International Journal of Computer Applications* 21, 43-47.
9. Kimber, A, and & Hansford, A. R. C. 1993. "A Statistical Analysis of Batting in Cricket ." *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 156(3), 443-455.
10. Koning, R. H. 2000. "Balance in competition in Dutch soccer. ." *Journal of the Royal Statistical Society: Series D (The Statistician)* 49(3), 419-431.
11. MacDonald, D et al. 2013. " Wicket-keeping in cricket: A literature review. ." *International Journal of Sports Science & Coaching* 8(3), 531-542.
12. Malhotra, A; & Krishna,. 2017. "A Statistical Analysis of Bowling Performance in Cricket. ." arXiv preprint arXiv:1701.04438.
13. Munro, C. E; Christie,. 2018. "Research directions for the enhancement of women's cricket." *International Journal of Sports Science & Coaching,* 13(5), 708-712.
14. Nicholson, R. 2015. *Like a man trying to knit'?: Women's Cricket in Britain, 1945-2000* . Doctoral dissertation, Queen Mary University of London.
15. Saikia, H, and D. & Bhattacharjee. 2010. "A Bayesian Classification Model for Predicting the Performance of All-Rounders in the Indian Premier League. ." Available at SSRN 1622060.
16. Saikia, H; et. al. 2016. "A new model for player selection in cricket." *International Journal of Performance Analysis in Sport,* 16(1), 373-388.
17. Shah, P. 2017. "New performance measure in Cricket. ." *ISOR Journal of Sports and Physical Education,* 4(3), 28-30.
18. Shah, S et al. 2017. "A study on performance of cricket players using factor analysis approach." *International Journal of Advanced Research in Computer Science* 8(3). 178.
19. Singh, A; & Sharma, R. 2017. "Construction of Test to Judge Ball Sense Ability of Batsman in Cricket ." *Spor Bilimleri Araştırmaları Dergisi* 2(2), 48-58.
20. Sumner, J; & Mobley, M. 1981. "Are Cricket Umpires Biased." *New Scientist,* 91(1260), 29-31.
21. Wickramasinghe, Indika Pradeep. 2014. "Predicting the performance of batsmen in test cricket."
22. Willoughby, K. A. 2002. "Winning games in Canadian football: A logistic regression analysis. ." *The College Mathematics Journal* 33(3), 215-220.
23. Wormgoor, S et al. 2010. " Anthropometric, biomechanical, and isokinetic strength predictors of ball release speed in high-performance cricket fast bowlers. ." *Journal of Sports Sciences* 28(9), 957-965.