**Prompt:**

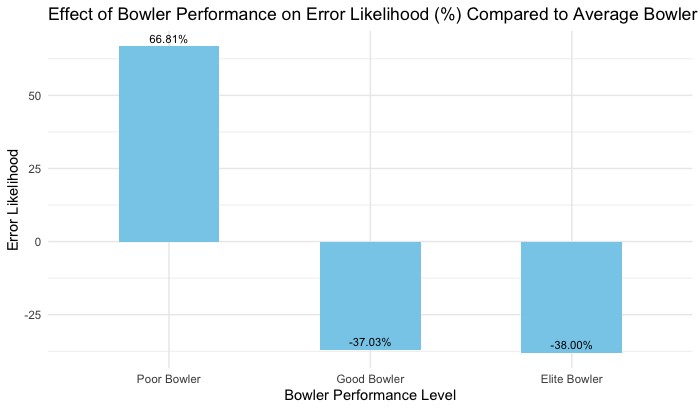
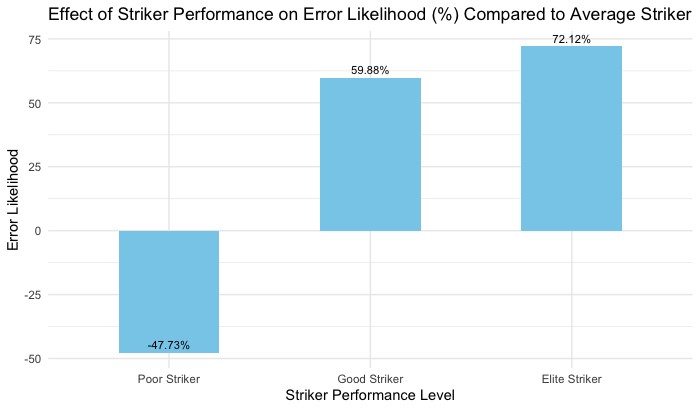
**Coach Ella leads a competitive cricket team. Through experience, Coach Ella has determined that teams which have fewer unforced errors, specifically the number of “wides” and “no-balls” than their opponent tend to win more often. This year, the team has set a goal to win the “Error Margin,” the term they use for the number of wides and no-balls for their team minus the number for their opponent. Since the team has limited practice time, Coach Ella wants to know where to focus the team’s training.**

1. **The team has compiled data from various cricket matches over several years. Coach Ella wants you to use the data to determine whether the offense (the identity of the batter) or the defense (the identity of the bowler) is more important in determining the likelihood of an error (a wide or no-ball). In other words, does the batter put pressure on their opponent and cause them to make a mistake, or do bowlers commit errors independent of the offense? Depending on your results, Coach Ella will focus more time on practicing batting or bowling to improve the team’s performance.**
2. **Coach Ella also wants to understand how fatigue changes the likelihood of making a mistake. Use the data to determine whether this is something the team should be worried about. If so, Coach Ella will focus on conditioning during practice to ensure the players do not fatigue during the match. Please include a visual aide related to your response that would help explain your findings to Coach Ella.**

**Striker Focus: The Key to Winning the Error Margin in Cricket**

Today, I'll present an analysis focused on the relationship between player performance and the occurrence of wides and no-balls in cricket matches. I will aim to identify whether the offense (striker) or the defense (bowler) holds greater influence over these errors, and I hope my findings will serve as a guide to your training strategy and help you make the most of your limited practice time for maximum impact on the 'Error Margin.’

In doing analysis on the data given to me, I was able to combine all of the datasets and trim them down into only necessary parts. From this simplified data, I created some new variables such as “error” that combined the “noball” and “wide” variables to see whenever an error was committed, as well as a “performance” variable that assigned a performance level to each player based off their runs per over. In assigning my categories of performance for bowlers, I labeled below 2 runs per over allowed as elite, between 2 and 3 as good, between 3 and 4 as average, and above 4 as poor. For the categories for strikers, I categorized above 5 runs scored per over as elite, between 3 and 5 as good, between 2 and 3 as average, and below 2 as poor. Creating this “performance” variable enabled me to fit a model that computed the change in probability of an error happening across each performance level. In conducting this data analysis, it became evident that the performance of the players, particularly the strikers, played a huge part in the likelihood of errors in a cricket match. I was able to get statistically significant results that showed me the log odds of each level of performance on the error likelihood, which I then converted into percentages. I captured this data in bar plots that show the effect of strikers and bowlers performance on the likelihood of errors compared to an average player:



To interpret the data above, let’s first look at the bar graph on the left. This bar graph is showing the effect of striker performance on error likelihood when comparing to an average striker. As we can see on the graph, our values are -47.73% for poor strikers, 59.88% for good strikers, and 72.12% for elite strikers. The interpretation of these values is that poor strikers show a negative impact of -47.73%, indicating a decreased likelihood of 47.73% of errors compared to an average striker. Conversely, good strikers demonstrate a 59.88% increase, while elite strikers exhibit a striking 72.12% rise in error likelihood. Obviously, we can see that the performance of strikers has a significant impact on errors, but let’s also look at the bar graph for bowlers before coming to any conclusions. Interpreting the same way as we did above, we see that poor bowlers show a positive impact of 66.81%, which means that when a poor bowler is bowling, there is an increase of 66.81% in error likelihood compared to average bowlers. On the other end, good bowlers show a 37.03% decrease in error likelihood and 38% decrease for elite bowlers.

Comparing the impacts of striker and bowler performance on error likelihood, we can see that both are significant and undoubtedly impact the game’s outcome. I think that the disparities between the two groups can be seen in the consistent change in percentage for the strikers and comparing the total change. Looking at the percentage changes observed for strikers, it ranges from -47.73% to 72.12% while for bowlers it goes from -38% to 66.81%. That is a 119.85% change in error likelihood when going across all the levels of performance for strikers, compared to a 104.81% change for bowlers. I also think that the consistent change across performance levels for strikers is significant to look at compared to bowlers. The significant change from good to elite performance for strikers shows that working on improving your skills at striking will pay off no matter what level of performance you are at, which is not the same for bowlers.

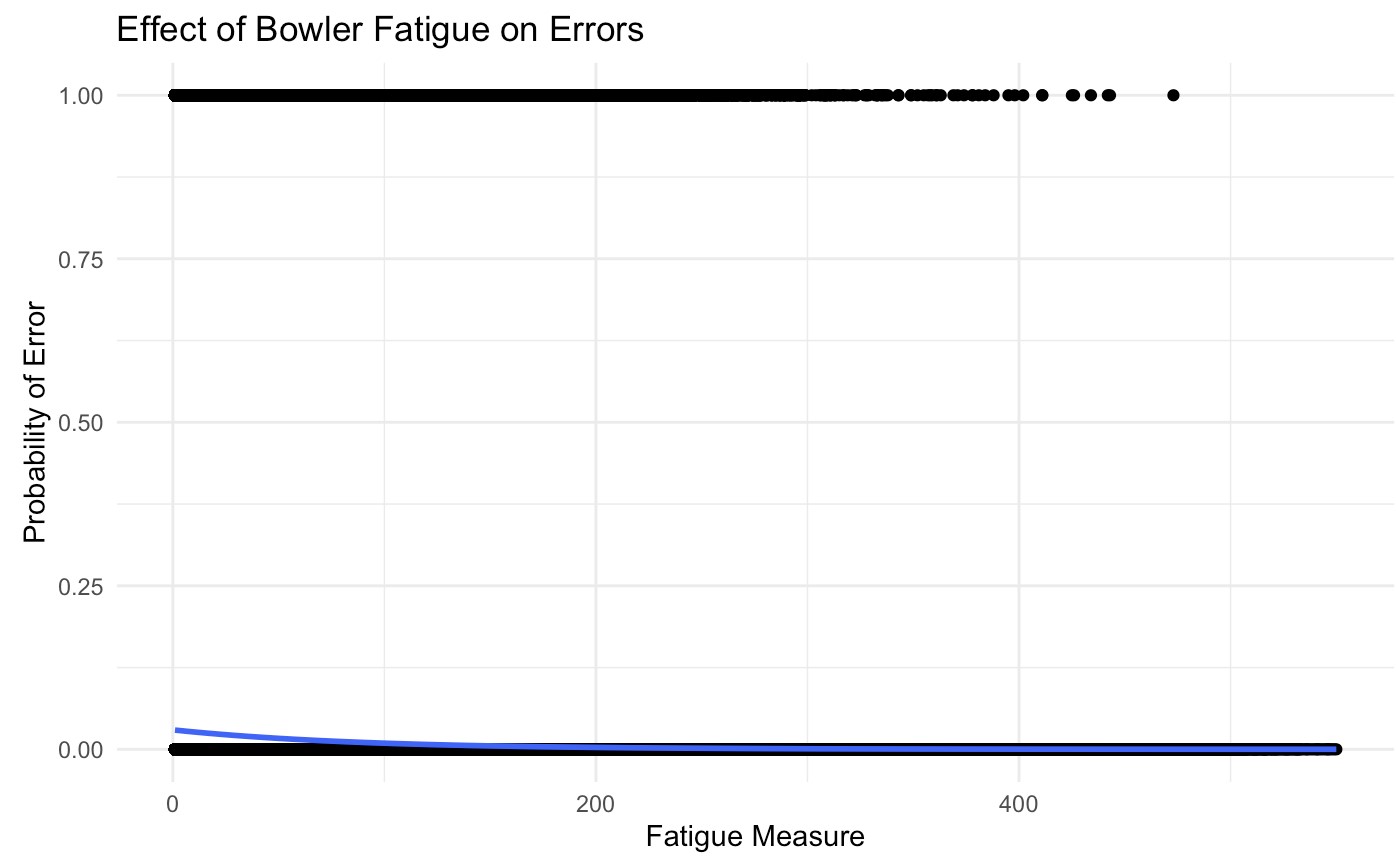
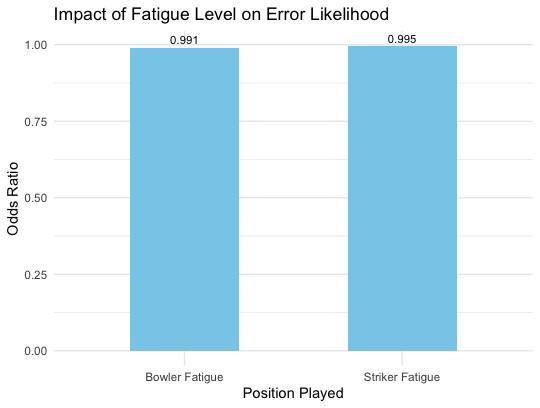
In conclusion, while both aspects — striker and bowler performance — hold significance in influencing error likelihood, the contrast in the consistent and impactful changes across performance levels suggests a stronger influence of striker performance. This emphasizes the benefits of focusing on striker training to enhance the team's overall performance maximize the impact on error margins in cricket matches.

**Why Conditioning is Not as Big of a Deal as People Might Think**

In this analysis, I'll dive into examining the correlation between player fatigue and the occurrence of wides and no-balls in matches. Using this data analysis, my goal is to decipher whether effects of fatigue are significant in shaping the 'Error Margin' during gameplay. This exploration aims to guide strategic decision-making, offering insights into whether conditioning practices could reduce errors and enhance overall team performance on the field.

Using the same set of simplified data from the previous analysis, this time I created a “fatigue” variable for the striker and bowler. To create this variable, I coded it so that for every ball that a player was on the field for in a game, their fatigue variable in the position they were playing would go up by 1 and would reset at the end of a game. This means that at the end of a

game, a player’s fatigue will be the amount of balls they played at for each position. Creating this fatigue variable allowed me to create a model that computed the change in probability of an error happening as a player’s fatigue increased. After creating this model, the analyzation of the results helped me come to the conclusion that conditioning might not play as big of a part in reducing errors as we thought it might. From the model, I was able to get statistically significant results that concluded as a player became more fatigued, their chance of committing an error minimally decreases. This conclusion can be seen in the bar plot and linear regression graph shown below:



Let’s first interpret the bar plot. The odds ratio of the positions is .991 for bowlers and .995 for strikers. The interpretations of the bowler value is that for each additional ball played for bowlers, the probability of them committing an error drops by .09%. Likewise, for strikers, the probability of them committing an error drops by .05% for each additional ball played. This notion can also be seen in the linear regression graph. In this graph, I have plotted the effect of bowler fatigue on errors. The blue line on the graph shows us the probability of an error as the fatigue of a bowler increases. We can see that the error probability starts at around .04 and slowly goes down as the fatigue increases.

However, these minute changes in error likelihood based on fatigue levels for bowlers and strikers don't warrant a drastic reaction as to not condition players at all. The marginal impact of .09% decrease in error likelihood for bowlers and .05% for strikers with each ball played showcases a negligible shift. Moreover, when we visualize the effect of bowler fatigue on errors through the linear regression graph, the gradual decrease in error probability from .04 as fatigue increases doesn't portray a drastic or crucial relationship either. This indicates that while fatigue may play a role, its impact on error occurrence remains relatively minimal. Therefore, I believe changing training efforts towards conditioning would not be helpful in trying to win the ‘error margin,’ and that coaches should continue to condition their players how they see fit.