

A Stochastic Model of Sabre Fencing

Lynn Uemoto

Franklin School, 80 Town Square Place, Jersey City, NJ, 07310, USA; lynn.uemoto@gmail.com

ABSTRACT: This paper presents a statistical analysis and model framework for simulating and analyzing sabre fencing bouts using discrete stochastic processes. Sabre fencing is a ‘first-to-15’ sport without time constraints, which aligns well with event-based modeling. The model uses a Poisson race model, where each fencer scores touches based on their average scoring rate (λ). The parametric approach allows for the derivation of analytical win probabilities via the negative binomial distribution and enables comparison of fencers who have never competed against each other. To capture the real-world impact of penalties, a state-dependent Markov chain was used to model the FIE’s escalating card rules, moving beyond a simple memoryless model. Simulations of 10,000 bouts validate the framework against theoretical results and quantify the influence of scoring rates and penalties on bout outcomes. This study establishes a foundational, parametric model for sabre, demonstrating how classical probability models can effectively represent dynamic, rule-driven competition and provides a framework for understanding and predicting outcomes in sabre fencing.

KEYWORDS: Mathematics, Probability and Statistics, Stochastic Processes, Poisson Distribution, Sport Analytics.

■ Introduction

In recent years, the field of statistical modeling has become fundamental to advanced performance analysis across many sports. The rise of sports analytics, popularized by “sabermetrics” as seen in the movie *Moneyball*, signaled a major shift: the use of probabilistic models to uncover hidden values that traditional scouting methods could not.¹ This concept has since altered many other major sports. For instance, Poisson distribution-based models have become prominent in predicting goals scored in hockey and soccer. In contrast, Markov chain models have been used to evaluate possession value in basketball, and linear probability models for tennis.²⁻⁵ At the center of this revolution is the use of probability not just to describe what happened, but to forecast what will happen.

Despite this revolution, fencing remains largely unexplored. Existing quantitative research falls into three groups. The first focuses on physical and technical characteristics for technical and tactical analysis, notably biomechanical or motion analysis.^{6,7} While valuable, there is a gap, as strength in fencing is not only a product of isolated physical attributes. A fencer with high speed or reach may lose to an opponent with great and practical strategy or decision-making. A great example is Aron Szilagyi, the three-time individual Olympic sabre champion. While not being the fastest or tallest, his exceptional footwork and tactical intelligence have consistently allowed him to outperform athletes with more physical advantages. The second and third focus on visualization and deep learning tools for interpreting bout data and statistical models, as exemplified by Bradley-Terry to infer aggregate “player strength” from win-loss records.⁸⁻¹¹ Addressing the limitations in the first group, a recent and complementary research focuses on simulating bouts from video data by modeling sequences of discrete actions. Lin *et al.* (2025) developed a pipeline to extract fencer motion and positions from video, clustering them into discrete actions, and

then modeling bout progression using a probabilistic state machine.⁹ This data-driven model effectively captures the tactical interplay of action in a touch.

While these existing tools have advanced fencing analytics, two limitations remain. First, biomechanical and technical analyses describe physical attributes or movement patterns but are not designed to produce win probabilities or to isolate how a fencer’s ability translates to bout contexts. Second, approaches, such as Bradley-Terry ratings and video-based state machine simulators, can rank fencers or replicate actions, yet they do not provide a compact, parametric model that yields closed-form solutions and enables prediction for fencers who have never faced each other. Consequently, a gap exists for a foundational, interpretable probability model at the bout level. The approach of Lin *et al.* is complex, while foundational works in sports analytics have consistently shown the power of simple, parametric models, for example, the Poisson distribution, for general insights and closed-form solutions.^{2,3} This proposed Poisson race model addresses this gap, positioning as the “xG of fencing,” whereas the model of Lin *et al.* functions as a detailed “play-by-play” simulator that generates tactical sequences.⁹

The sport’s ‘first-to-15’, discrete-points structure naturally aligns with discrete-event models. Bouts are treated as a race between two independent Poisson processes, providing closed-form win probabilities via the negative binomial distribution. Penalties are incorporated through a discrete-time Markov chain that captures the FIE’s escalating penalty card rules. The research addresses two core questions: (1) How can discrete probability distributions simulate and analyze sabre bouts? (2) What insights can these simulations provide about match outcomes and the influence of penalties? By addressing these, the study aims to provide a simple, parametric framework that yields interpretable win probabilities and quantifies risk from

penalties, establishing a benchmark for future fencing analytics.

■ Methods

This study combines both theoretical probability models and computational simulations to model sabre fencing. A sabre bout is won by the first fencer to score 15 touches. As touches occur as separate, random events, the scoring process can be described using discrete probability distributions. The model uses a Poisson process and its exponential inter-arrival times, while penalties (yellow and red cards) are modeled as a discrete-time Markov chain. Analytical win probabilities are derived from the negative binomial distributions and simulations, and Python is used to validate and extend these analytical results.

Poisson Race Model and Closed-Form Win Probability:

A sabre bout is conceptualized as a race between two independent Poisson processes.¹² A scoring parameter characterizes each fencer, λ , which represents the average number of touches scored per bout. In this context, since a bout ends when reaching 15 points and has no time limit, the “time” variable is abstract. It represents the bout through scoring exchanges, not physical clock time. Therefore, the “time interval” is set to $t=1$ bout, simplifying the model. The number of touches, X , scored by a fencer thus follows a Poisson distribution.¹²

The Probability Mass Function (PMF) of a Poisson distribution gives the probability of how many times k events occur in a fixed interval.^{13,14}

$$P(X = k) = \frac{(\lambda t)^k e^{-\lambda t}}{k!}, \quad k = 0, 1, 2, \dots$$

This model relies on key assumptions that are reasonable approximations for sabre: (1) events occur independently of previous touches, (2) scoring rate λ is constant throughout the bout (no fatigue or momentum), and (3) in a single exchange, the probability of more than one touch is negligible. Strictly, only one scoring touch can occur per exchange, but a penalty point (awarded for a red card) may be added during the same exchange. The model, therefore, treats the scoring touch and penalty point as separate events rather than multiple touches.

Because the two Poisson processes are independent, the probability that the next touch in the bout is scored by Fencer A against Fencer B is constant and given by¹³

$$p_A = \frac{\lambda_A}{\lambda_A + \lambda_B}$$

where this probability does not depend on the state of previous touches; a memoryless property inherited from the Poisson processes. Although the number of exchanges between touches is random (and would follow an exponential distribution if we measured abstract “bout time”), only the order of touches matters to determine the winner.^{13,14} Thus the bout can be treated as a sequence of independent Bernoulli trials with success probability p_A .

A key property of the Poisson distribution is that its variance is equal to its mean (λ).¹²⁻¹⁴ This means that a fencer's scoring is unpredictable, where the number of touches they score in a bout will naturally vary from one bout to another,

with a spread defined by λ . This built-in variance is a strength of the model, as it captures the element of chance in any sport. Adding to the memoryless property of the exponential distribution, this ‘first-to-15’ race amplifies that underlying randomness, ensuring that even a significantly stronger fencer ($\lambda_A > \lambda_B$) has a non-zero probability of losing, which aligns with the real-world dynamics of sabre fencing, where upsets are possible, and no fencer is immune to it.

The number of touches scored by Fencer B before Fencer A reaches 15 therefore follows a negative binomial distribution. (see Appendix A). In general, for a sequence of independent trials with success probability p , the cumulative probability of at most k failures before the r -th success is:¹⁴

$$P(\text{at most } k \text{ failures}) = \sum_{k=0}^r \binom{k+r-1}{k} (1-p)^k p^r$$

Applying this to a first-to-15 sabre bout with $r=15$, $k=14$, and $p=p_A$, Fencer A wins if Fencer B scores at most 14 touches before A's 15th touch. Therefore:

$$P(A \text{ wins}) = \sum_{k=0}^{14} \binom{k+14}{k} (1-p_A)^k p_A^{15}$$

Here the binomial coefficient $\binom{k+14}{k}$ counts the number of sequences in which Fencer B scores exactly k touches before Fencer A's 15th touch.¹³ This closed-form theoretical probability serves as a benchmark to compare with simulation results and validate the simulation.

For example, with $\lambda_A=12$ and $\lambda_B=8$, we have $p_A=0.6$ and closed-form probability as follows:

$$P(A \text{ wins}) = \sum_{k=0}^{14} \binom{k+14}{k} 0.6^{15} 0.4^k \approx 0.864$$

Modeling Penalties with State Dependent Thresholds:

The initial baseline model for penalties treated each card as an independent event, modeled with a Bernoulli process. However, the FIE's penalty escalation rules are state-dependent.¹⁵ To accurately capture this, we extend the model to a discrete-time Markov chain, where the probability and consequence of a penalty depend on the fencer's current card state.

In sabre fencing, we consider 2 penalty types relevant to bout scoring: yellow cards and red cards. Black cards, which result in immediate disqualification from the entire competition, are excluded from this model due to their extreme rarity in bout-level outcomes. Under FIE rules, a fencer starts in a two-yellow threshold regime: two yellow cards accumulated before any red card result in a red card and a penalty point for the opponent. Once a fencer receives any red card (directly or through two yellow cards), they enter a one-yellow threshold regime: any subsequent yellow card immediately yields another red card (and another penalty point).

This system is naturally modeled as a discrete-time Markov chain, where the probability of transitioning between penalty states depends only on the current state.¹⁶ The model defines three distinct states based on the current threshold and the number of yellow cards a fencer has at touch k , where $S_{0,k}$ is the starting state for every bout. The states S_0 and S_1 belong to the initial two-yellow regime; T_0 is the post-red-one-yellow threshold regime.

- $S_{0,k}$: The probability that the fencer is in the initial “two-yellow threshold” regime and has zero yellow cards

- $S_{1,k}$: The probability that the fencer is in the initial “two-yellow threshold” regime and has one yellow card
- $T_{0,k}$: The probability that the fencer is in the post-red “one-yellow threshold” regime and has zero yellow cards

On any touch, a fencer may receive a yellow card with probability p_y , a direct red card with probability p_r , or no card ($p_N = 1 - p_y - p_r$). The state transitions are illustrated in Figure 1, which shows the three penalty states and probabilities of moving between them.

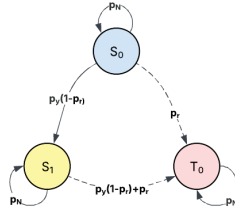


Figure 1: State transition diagram for the sabre fencing penalty model. This diagram illustrates the three penalty states and the probabilistic transitions between them through a Markov chain.

It is important to distinguish the triggering mechanism from the consequences of a card. While p_y and p_r are constant, memoryless probabilities, whether a yellow card awards a penalty point is state-dependent: in the two-yellow regime (S_0 , S_1), two yellows are needed for a red card; in the one-yellow regime (T_0), a single yellow immediately yields a red card. This hybrid model is set with Bernoulli card events with a state-dependent scoring impact, which is the key extension beyond a simple binomial penalty model.

The probability of a penalty point being awarded on touch k is therefore

$$a_k = S_{0,k}p_r + S_{1,k}(p_y + p_r) + T_{0,k}(p_y + p_r)$$

The state probabilities update as follows:

$$\begin{aligned} S_{0,k+1} &= S_{0,k}p_N \\ S_{1,k+1} &= S_{0,k}p_y + S_{1,k}p_N \\ T_{0,k+1} &= T_{0,k}p_N + a_k \end{aligned}$$

with $p_N = 1 - p_y - p_r$ and initial conditions $S_{0,0} = 1$, $S_{1,0} = T_{0,0} = 0$. This models the probability of a red card at each exchange, capturing the escalation rules compactly.

Integration of Models:

To unify the scoring and penalty processes, we frame the bout using joint probability distributions.¹⁶ The framework assumes that the scoring fencer is not the one receiving a card in the same exchange; situations where a touch is annulled because the scored committed an offence (e.g., guard punch, leaving piste) are excluded. This simplification keeps the joint distribution tractable and focuses on the more common case where the penalty goes against the non-scoring fencer.

Let X represent the scoring outcome and Y represent the penalty outcome for each exchange. Their joint probability mass function is $P(X=x, Y=y)$.¹⁶ This gives the probability that the outcome x for scoring and outcome y occur together on the same touch.

For each exchange, we consider six mutually exclusive joint events:

1. A scores, no red card for either $\rightarrow A + 1$
2. B scores, no red card for either $\rightarrow B + 1$

3. No score, A receives a red $\rightarrow B + 1$ (penalty point)
4. No score, B receives a red $\rightarrow A + 1$ (penalty point)
5. A scores and B receives a red card. $\rightarrow A + 2$ (touch + penalty point)
6. B scores and A receives a red card. $\rightarrow B + 2$ (touch + penalty point)

Let R_B denote the event that fencer B receives a red card. Then $P(A \text{ scores}, Y = R_B) = P(A \text{ scores}) \cdot P(R_B)$, assuming independence between the scoring outcome and card events as the bout progresses.

And similarly for the symmetric case. This reflects situations where a fencer scores a valid touch while their opponent gets a penalty card.

The overall bout outcome marginalizes over the joint distribution of penalty states:

$$P^*(A \text{ wins}) = \sum_{y_A} \sum_{y_B} P(A \text{ wins} | Y_A = y_A, Y_B = y_B) \cdot P(Y_A = y_A, Y_B = y_B)$$

where, $P(Y_A = y_A, Y_B = y_B)$ is the joint probability of penalty states from the Markov chain. This is estimated by running a Markov chain simulation and counting how often each state combination occurs. $P(A \text{ wins} | Y_A = y_A, Y_B = y_B)$ incorporates the adjusted probabilities for all six joint outcomes and is the conditional probability.

This comprehensive joint framework captures the full complexity of sabre fencing, where scoring and penalty can occur simultaneously, sometimes resulting in multiple point awards within a single exchange.

Simulation Design and Parameter Justification:

To operationalize the theoretical model and explore scenarios, a discrete-event simulation was developed in Python (v3.10). The simulations of 10,000 bouts were implemented using the NumPy library for random number generation and the Matplotlib and Seaborn libraries for visualizations. The Python code was developed with the assistance of large language models to generate code for the simulation and is publicly available in a GitHub repository: <https://github.com/LynnAchamo/sabre-fencing-model>.

The baseline matchup uses $\lambda_A = 12$ and $\lambda_B = 8$. The relative advantage of 50% creates a clear but not overwhelming skill gap. Analytically, Fencer A wins 86% of bouts, leaving a 14% upset rate. This reflects the reality that no fencer is unbeatable. However, the parameters are illustrative, and λ for each fencer could be estimated from historical bout data.

Yellow and direct-red card probabilities were set at $p_y = 0.10$ and $p_r = 0.02$ per touch. Public datasets of sabre penalty rates are not available: for this reason, values are synthetic. They are set somewhat higher than elite-level bouts so that the effect of the state-dependent escalation is observable across 10,000 bouts, while keeping cards rare.

Every figure in the Results section is generated from the same baseline parameter set ($\lambda_A = 12$, $\lambda_B = 8$, $p_y = 0.10$, $p_r = 0.02$) and 10,000 bout simulations unless explicitly noted otherwise.

Sensitivity Analysis:

A sensitivity analysis will be conducted to examine how changes in fencer strength affect win probabilities.^{14,18} We define fencer strength through a relative scoring rate or the ratio of the two Poisson rate parameters $\rho = \lambda_A/\lambda_B$. This ratio is different from the probability that Fencer A wins a single scoring exchange described as $p_A = \lambda_A/(\lambda_A + \lambda_B) = r/(1+r)$, from which the race-to-15 win probability $P(\rho)$ follows the negative-binomial cumulative distribution (Appendix A):

$$P(A \text{ wins}) = P(\rho) = \sum_{k=0}^{14} \binom{k+14}{k} (1-p_A)^k p_A^{15}, \quad p_A = \frac{\rho}{1+\rho}$$

This relationship is nonlinear, where small changes in r near parity (i.e., when $\rho \approx 1$, meaning the fencers are relatively even) produce larger changes in $P(\rho)$, while the same amount of change far from parity has a much smaller effect.

Sensitivity is measured by $dP/d\rho$. Applying the chain rule:

$$\frac{dP}{d\rho} = \frac{dP}{dp_A} \cdot \frac{dp_A}{d\rho}, \quad \frac{dp_A}{d\rho} = \frac{1}{(1+\rho)^2}$$

For two evenly matched fencers ($\rho = 1$, $p_A = 0.5$), we have $dp_A/d\rho = 1/4$.

The derivative dP/dp_A can be calculated from the PMF of the negative binomial distribution at $p_A = 0.5$, and is approximately 4.380, giving $dP/d\rho \approx 1.10$. Thus, a 10% relative improvement in ρ near parity translates into an expected increase of 11% in win probability, confirming the large impact of small skill improvements in a 1:1 skill matchup.

To reflect the inherent variability of one bout, we construct a 95% confidence interval around the theoretical win probability. Treating the bout as a sequence of independent point-level trials, the effective sample size is the expected number of total exchanges:^{13,14}

$$n_{eff} = \frac{15}{p_A} + \frac{15}{1-p_A} - 1$$

The standard error of the win probability is $SE = \sqrt{P(1-P)/n_{eff}}$, and the resulting 95% interval for a single bout is $P \pm 1.96 SE$ (reported in Table 1).

A note is that this interval quantifies the “bout-level” confidence interval, rather than the Wilson confidence interval used later in the simulation validation. The Wilson interval captures the faithful representation of the empirical win proportion across 10,000 independent simulations, whereas the interval here computed reflects the uncertainty of a single bout under the same fixed parameters.

Simulation Validation and Statistical Methods:

To assess whether the discrete-event simulation faithfully represents the closed-form Poisson race model, we will perform a hypothesis test and construct a confidence interval for the simulated win probability. The theoretical win probability for Fencer A is derived from the negative binomial distribution (see “Poisson Race Model and Closed-Form Win Probability”) and serves as the null-hypothesis value p_0 . The simulation will be run for $N=10,000$ independent bout simulations for the base scenario ($\lambda_A = 12$, $\lambda_B = 8$) and will record the number of bouts won by Fencer A. Let $\hat{p}$ denote the simulated win proportion.

For the one-sample proportion test, the null hypothesis $H_0: p = p_0$ is tested against the two-sided alternative $H_A: p \neq p_0$. Because the number of trials is large and p_0 is not extreme, the test statistic given by

$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$

approximately follows a standard normal distribution under H_0 .¹⁴ The corresponding two-tailed p -value is $P(|Z| > |z|)$ to determine whether the true probability is different from p_0 (either larger or smaller). We use a significance level of $\alpha=0.05$, where failure to reject H_0 indicates that the simulation and the theoretical model are statistically consistent.

To quantify the uncertainty around $\hat{p}$, we construct a 95% confidence interval using the Wilson score method. The Wilson interval was selected over the Wald (normal approximation) interval due to its improvement over the normal approximation interval for handling proportions near 0 or 1 better and does not rely on the assumption that $\hat{p}$ is normally distributed when probability is skewed.¹⁷ The interval is given as:

$$CI = \frac{\hat{p} + \frac{z_{\alpha/2}^2}{2n} \pm z \sqrt{\frac{\hat{p}(1-\hat{p})}{n} + \frac{z_{\alpha/2}^2}{4n^2}}}{1 + \frac{z_{\alpha/2}^2}{n}}$$

Here, $z_{\alpha/2} = 1.96$ for 95% confidence; that leaves the tail area of $\alpha/2$ on each side. An interval that contains the theoretical value p_0 provides further evidence that the simulation faithfully represents the theoretical model.

After running the simulation, the observed $\hat{p}$, z -score with its p -value, and Wilson confidence interval will be reported in the Results section. Additionally, we will compute Wilson intervals at the 90% and 99% level to examine robustness.

Model Assumptions and Limitations:

The model is built upon several key assumptions, where the consequences are important in interpreting results and limitations. These assumptions fall into two categories: those for the scoring process and those for the penalty events.

For the scoring process, the model assumes a fencer's scoring rate (λ) remains constant through the bout, without accounting for fatigue, momentum, or tactical changes. This means the model may underestimate the probability of comebacks and a performance decline that occur in real bouts. Additionally, each touch is treated as an independent event following a Poisson process, thereby ignoring psychological factors, namely confidence shifts after winning consecutive touches or frustration following losses.

The penalty model introduces a state-dependent logic through a discrete-time and discrete-state-space Markov process (i.e., Markov Chain), where the consequence of the card depends on the current penalty state.¹⁶ A key assumption of this Markov model is time-homogeneity, where the probabilities of a yellow card (p_y) and a direct red card (p_r) are constant for every touch. This modeling convenience provides a baseline but may understate real-world dynamics where a fencer, feeling pressured late in a bout, becomes more aggressive and prone to penalties, or a referee's card frequency changes.

However, this approach still assumes the fundamental probabilities of receiving yellow (p_y) and direct red (p_r) cards are constant and not influenced by scoring dynamics, fencer behavior, or referee biases. In reality, a fencer losing by a large margin may become more aggressive and susceptible to penalties, or a referee's card frequency may change as a bout progresses.

A fundamental limitation comes from the interaction between these two modeling approaches. The scoring model assumes memoryless, independent events, while the penalty model explicitly incorporates memory through state tracking. This hybrid provides a baseline representation but cannot capture more subtle interactions, notably how receiving a yellow card might affect a fencer's subsequent scoring probability.

Finally, the model treats the two fencers' processes as independent, neglecting adaptive strategies where a fencer may alter their tactics based on the opponent's habits. A high λ might be suppressed by effective defensive strategies from the opponent, and interaction is not accounted for here.

These limitations are not weaknesses but rather clear directions for future work, as discussed later in the conclusion, where paths for incorporating more dynamic, interactive elements into the modeling framework will be outlined based on these assumptions.

■ Results and Discussion

Validation of Poisson Race Model:

To confirm that the simulation faithfully implements the Poisson race model, we first compare its output to the theoretical win probability. For the baseline matchup ($\lambda_A = 12, \lambda_B = 8$), the theoretical win probability derived from the negative binomial model is 0.864 (see Methods). The simulation of 10,000 bouts produced an empirical win proportion $\hat{p} = 0.861$ for the stronger fencer. The close agreement between the theoretical and simulated win probabilities, a difference of 0.3 percentage points, validates that our Poisson race framework captures the core dynamics of a sabre bout.

To assess whether this discrepancy is statistically significant, a one-sample proportion test was performed as described in the Methods. The resulting z score was -0.758, yielding a two-tailed p-value of 0.448. At the standard significance level of $\alpha=0.05$, we fail to reject the null hypothesis that the true win probability equals the theoretical value. Therefore, the simulation and the analytical model are statistically consistent.

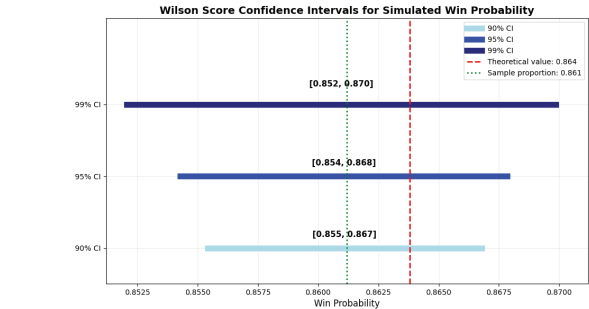


Figure 2: Wilson confidence intervals for simulated win probability ($\lambda_A = 12, \lambda_B = 8, 10,000$ bouts). The theoretical value 0.864 from the negative binomial formula falls inside all intervals, validating the simulation's consistency with the model.

Figure 2 displays Wilson confidence intervals for the simulated win proportion at 90%, 95%, and 99% confidence intervals. The 95% Wilson interval was [0.854,0.868], which contains the theoretical value of 0.864. The 90% and 99% confidence levels also include 0.864, showing consistency of simulation results across different statistical confidence levels.

These validation results demonstrate that a simple Poisson race model can serve as a reliable benchmark for sabre bout outcomes. If a fencer's scoring rate λ can be estimated from historical bout data, then the model offers a way to compare two fencers even when they have never faced each other. Because the win probability follows directly from the two λ values, the model provides a quantitative baseline against which real-world outcomes can be compared. While the current parameters are illustrative, the structure demonstrates how such a model could, with empirical data, help produce interpretable competitive bout expectations, where more complex psychological or tactical effects can later be measured.

Sensitivity Analysis:

Having established that the simulation faithfully reproduces the closed-form win probability, we next examine how that probability responds to changes in relative fencer strength. The table below reports the theoretical win probabilities and single-bout 95% confidence intervals for various ratios of ρ (derivation of sensitivity and confidence-interval setup is given in the Methods). Three patterns emerge from Table 1.

Table 1: Sensitivity analysis of win probabilities to scoring rate variations. A small increase in relative scoring rate yields substantial gains in win probability.

λ_A	λ_B	λ_A/λ_B	Advantage	Theoretical P(A wins)	Win Probability	95% CI
10	10	1.00	0%	0.500	50%	(0.422, 0.578)
11	10	1.10	+10%	0.602	60.2%	(0.527, 0.677)
12	10	1.20	+20%	0.690	69.0%	(0.619, 0.761)
12	8	1.50	+50%	0.864	86.4%	(0.812, 0.916)
14	10	1.40	+40%	0.819	81.9%	(0.760, 0.878)
15	9	1.67	+67%	0.916	91.6%	(0.876, 0.956)
8	12	0.67	-33%	0.136	13.6%	(0.078, 0.194)
9	11	0.82	-18%	0.293	29.3%	(0.210, 0.376)

First, the relationship between relative scoring rate and win probability is strongly nonlinear: a +10% advantage ($\rho=1.10$) increases win probability from 50% to 60.2%, while a +50% advantage ($\rho=1.50$) results in an 86.4% win probability. This nonlinear amplification effect, where small skill differences create substantial changes in win probability, occurs because the first-to-15 structure magnifies small per-exchange advantages.

Second, the most efficient range for competitive improvement is between parity and about a 20% advantage. The sensitivity of win probability to relative scoring rate at parity is $dP/d\rho \approx 1.10$ (Methods), meaning that near a 1:1 skill ratio, each 10% increase in ρ translates into an approximate 11% increase in win probability. Beyond a 50% advantage, the resulting increase in win probability diminishes sharply, creating a critical competitive threshold where further skill improvements add little to the probability of winning.

Third, the confidence intervals in Table 1 highlight the enormous bout-to-bout variability inherent in a race to 15. Even for evenly matched fencers ($P=0.500$), the 95% interval spans from 0.422 to 0.578, so a single bout is very noisy, indicating the underlying skill difference. This interval narrows as P approaches 0 or 1, but they remain wide enough that upsets are plausible even when there is a clear winner.

This analysis supports the paper's premise that a compact parametric model can capture a key competitive dynamic where small skill differences, when in the first-to-15 format, can create significant changes in expected outcomes. The model therefore provides a benchmark backing for discussing the importance of technical improvements, even if they are small, and explains why stronger fencers are not guaranteed victory in any bout.

Score Margin Distribution:

Sensitivity analysis reveals how win probability shifts with scoring rates, but it does not capture characteristics of the bout outcome. We therefore examine the distribution of winning and losing margins for the baseline matchup. We define a "decisive victory" as a win by four or more touches and a "close victory" as a win by three or fewer.

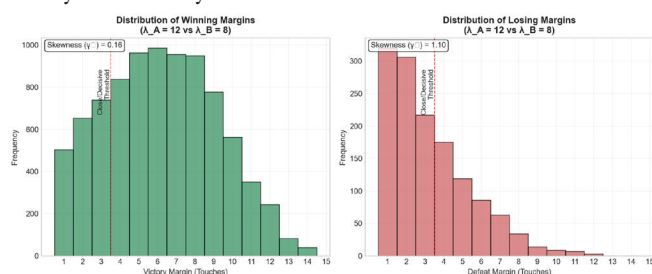


Figure 3: Distribution of victory and defeat margins for a stronger fencer ($\lambda_A = 12$ vs $\lambda_B = 8$) from 10,000 simulated bouts. Wins tend to be decisive (≥ 4 touches), while losses tend to be close (1-3 touches). This asymmetry shows that when a stronger fencer loses, it is most likely a narrow upset, while their victories are by a larger margin, underscoring how the first-to-15 format amplifies randomness.

Figure 3 displays the histograms of margins for wins and losses of the stronger fencer. When Fencer A won (85.83% of bouts), the victories were mostly decisive: 68% of wins were by a margin of four or more touches. In contrast, when the stronger fencer lost (14.17% of bouts), the defeats were close for an overwhelming number of times: 72% of losses came from a margin of one to three touches.

The sample skewness quantifies this contrast. For winning margins, Pearson's coefficient $\gamma=0.16$, indicating approximate symmetry. For losing margins, Pearson's coefficient $\gamma=1.1$, reflecting the right skewness visible in the distribution of Losing Margins (Figure 3), where most losses are narrow. The variance of the margin distribution is significant (35.340, see Table 2), a direct consequence of the Poisson assumption linking mean and variance and of the amplification introduced by the first-to-15 rule. It quantitatively demonstrates the high level of unpredictability in bout outcomes. Thus, even a decisive-looking 15 -10 win and a narrow 15 -14 loss are both plausible outcomes for the same skill matchup, purely due to random variation in the touches. This model can be used in the future

to examine whether an observed sequence of bout outcomes deviates from what randomness would produce, tracking progression of player growth or stagnation.

Penalty Analysis with State-Dependent Model:

Finally, we incorporate the penalty model to assess FIE's escalating card rules influences bout outcomes. All simulations in this section use the baseline scoring parameters and the penalty probabilities $p_y = 0.10$, $p_r = 0.02$. Although sabre bouts end when one fencer reaches 15 touches, the penalty-state model was simulated both under bounded (race-to-15) and extended (up to 30 total touches) conditions. Tables and figures are generated from 10,000 bounded bout simulations, unless noted. The unbounded results appear in Appendix B.

Table 2: Penalty Dispersion and Descriptive Statistics Bounded (race-to-15). Penalty events show high variance within the first penalty, typically occurring around touch 7.

Variable	Trials	Mean	Std. Dev.	Var
margin all	10000	4.496	5.945	35.340
margin penalty	9720	4.497	5.957	35.488
margin no penalty	280	4.458	5.504	30.293
yellow cards per bout	10000	2.909	1.696	2.963
red cards per bout	10000	0.438	0.586	0.368
first penalty touch	9720	7.337	3.112	11.972

Table 2 summarizes the number of penalties and the timing of the first penalty. Under the assumed baseline condition, an average bout contained 2.909 ± 1.696 yellow cards and 0.438 ± 0.586 red cards. The first penalty was observed, on average, at touch 7.337 ± 3.112 , indicating that most bouts saw a transition into a state with higher card-escalation risk (S_1 or T_0) within the first third of the bout. The high variance for both yellow cards ($\text{Var} = 2.963$) and red cards ($\text{Var} = 0.368$) indicates that penalties are not evenly distributed across bouts but instead are clustered randomly. This is a consequence of two factors: the underlying Bernoulli process that governs the issuing of cards, and the state-dependent escalation rules that transform the independent card events into a non-homogeneous process for scoring penalty points. While the probability of a referee giving a card on any single touch is constant (a Bernoulli trial), the probability that a card results in a point-winning red card is state-dependent. This combination of a memoryless trigger and memory-based consequence is what creates the high variance and random clustering observed in the data.

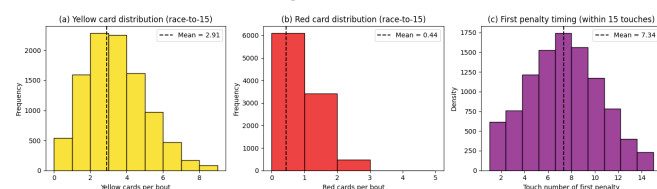


Figure 4: Marginal penalty distribution and first penalty timing under bounded (race-to-15) conditions ($\lambda_A = 12$, $\lambda_B = 8$, $p_y = 0.10$, $p_r = 0.02$, 10,000 bouts). Most first penalties occur in the middle stage of the bout, and red card distributions are heavily right-skewed. This non-uniformity is a direct consequence of combining the memoryless card issuing with state-dependent escalation.

Figure 4 shows the distribution of the first penalty touch and the yellow and red card count per bout. The first penalty timing distribution is unimodal with a peak in the middle phase of the bout. Red cards are heavily right-skewed: 62% of bouts have none, 25% have one, and 13% have two or more. This asymmetry comes naturally from the threshold rule.

Table 3: Penalty Win Probabilities Bounded (race-to-15). Penalties reduce the stronger fencer's win probability, especially in bouts where penalties occur, revealing high conditional impact despite low frequency.

Variable	Probability
win prob. (with penalties)	0.857
win prob. (no penalties)	0.864
win prob. (penalty bouts only)	0.835

Introducing penalties into the Poisson race model simulation lowered the stronger fencer's overall win probability from the penalty-free theoretical value of 0.864 to 0.857 (Table 3). While the reduction is modest in absolute terms (-0.8%), the impact becomes significant when conditioned on the presence of any penalty. In 97.2% of bouts where at least one yellow or red card was issued, Fencer A's win probability dropped to 0.835, a decrease of 3.4% relative to the penalty-free case. Thus, penalties function as high-impact, low-frequency events. While the probability of a penalty point due to a red card is rare on any single touch, when one does occur, it has a meaningful impact on the bout's trajectory, rather than being a negligible background factor.

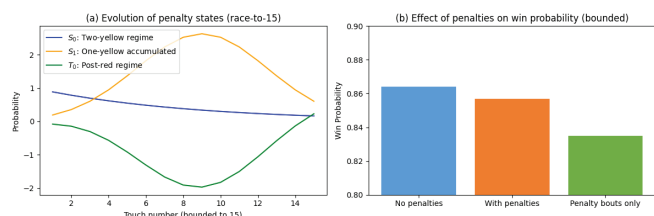


Figure 5: Markov penalty state probabilities over first 30 touches, computed analytically from the state transition equation with $p_y = 0.10$, $p_r = 0.02$. The initial state S_0 decays exponentially, while the post-red regime T_0 rises steadily, showing how penalty risk accumulates.

Figure 5 presents the evolution of Markov state probabilities across consecutive touches. The probability of remaining in the safe two-yellow regime (S_0) decreases exponentially from 1.0 to 0.04 by touch 30. The intermediate one-yellow state (S_1) peaks around touches 8-10 at approximately 0.38, and the post-red regime (T_0) rises steadily to 0.58. By the latter half of the bout, more than half of the fencers operate under the heightened one-yellow threshold, where any yellow card immediately results in a penalty point. This quantitatively captures the intuition that early warnings create long-lasting, escalating danger.

From a practical standpoint, the results could be used by coaches to underscore the importance of discipline from the first touch and through a bout. Additionally, the analysis supports the paper's central argument. By treating a bout as a stochastic race with state-dependent rules, we can move beyond simple outcome descriptions and begin to quantify how skill, risk, and rules interact to create the dynamics we see in competitive sabre fencing. From a modeling standpoint, these results demonstrate how integrating even a simple three-state chain can reproduce the cumulative escalation rules observed

in actual competition, providing a statistically interpretable bridge between rules and bout dynamics.

■ Conclusion

This paper presented an integrated probabilistic framework for modeling sabre fencing bouts using discrete stochastic processes. The Poisson race model, supported by negative binomial analytical results and large-scale Python simulations, demonstrated that fencers' scoring rates can accurately predict bout-level outcomes. Validation tests confirmed close agreement between theoretical and simulated win probabilities, with an absolute difference of less than 0.3 percentage points (0.864 theoretical vs 0.862 simulated). Incorporating the Markov-based penalty model extended this analysis to rule-dependent dynamics, revealing that although penalties are rare, their cumulative nature introduces asymmetry and a measurable reduction in win probability.

The primary contribution of this work is establishing a foundational, parametric model for Sabre, similar to how the early 'xG' models revolutionized soccer analytics. While this model excels at quantifying the baseline interplay of skill and discipline, its simplifying assumptions mean it does not directly capture psychological factors, namely, momentum or frustration. However, this is also its strength, as it provides a neutral expectation, where it becomes a benchmark against which these more complex effects can be identified and measured. For instance, a fencer who consistently outperforms their predicted win probability during a competition may have exceptional psychological resilience or tactical creativity under pressure.

The findings offer several practical insights for the fencing community. For the coaches and athletes, sensitivity analysis reveals a strong nonlinear relationship, where small, achievable improvements in scoring rates (e.g., 10-20%) can result in a large boost for win probability. This emphasizes the value of refining technique with a wide range of possibilities for development. Furthermore, the penalty model offers a concrete tool for teaching discipline management, visually showing how an early yellow card can significantly increase a fencer's vulnerability for the entire bout. For referees and officials, the state-dependent penalty model provides a quantitative perspective of the set rules of fencing. It underscores that a yellow card is not just an isolated warning but a fundamental shift in the bout's risk structure, potentially influencing the final bout outcome.

Looking forward, this framework opens many directions for future research. The most immediate step is adding empirical data: collecting historical bout data to estimate real-world scoring rates and penalty frequencies, then performing goodness-of-fit tests to compare simulated and actual victory margin distributions. The model can be extended to other fencing disciplines (foil and epee) by adjusting to time limits and incorporating discipline-specific rules. Further study could incorporate probabilities that change over time to reflect referee behavior or psychological effects (e.g., increased card frequency when a fencer is losing), and could also examine whether penalties disproportionately affect weaker fencers in

actual competition. Integrating this parametric approach with machine learning and motion analysis could also improve both accuracy and tactical insight.

The clearest takeaway from this work is that classical probability models, even with simple assumptions, can offer a useful framework for understanding strategy, risk, and outcome patterns in competitive sabre fencing. By treating a bout as a stochastic race with state-dependent rules, we can move beyond just outcome description to a deeper analysis of how skill, risk, and rules interact to shape bout dynamics, providing a foundation for the future of fencing analytics.

Acknowledgments

I wish to express my sincere gratitude and thanks to Adam Almakroudi of Imperial College London for his mentorship. Working on this research and learning was truly fun, and I am very happy to have the chance to explore and expand upon my passions.

References

- Baumer, B.; Zimbalist, A. *The Sabermetric Revolution: Assessing the Growth of Analytics in Baseball*; University Of Pennsylvania Press, 2014.
- Campbell, M.; Hong, S. Predicting the Outcome of Sports Competitions Using Poisson Distribution. *Issues in Information Systems* **2024**, *25* (1), 188–198. https://doi.org/10.48009/1_iis_2024_116.
- Everson, P.; Goldsmith-Pinkham, P. S. Composite Poisson Models for Goal Scoring. *Journal of Quantitative Analysis in Sports* **2008**, *4* (2). <https://doi.org/10.2202/1559-0410.1107>.
- Skinner, B.; Guy, S. J. A Method for Using Player Tracking Data in Basketball to Learn Player Skills and Predict Team Performance. *PLOS ONE* **2015**, *10* (9), e0136393. <https://doi.org/10.1371/journal.pone.0136393>.
- Shirley, K. *A Markov Model for Basketball*; New England Symposium for Statistics in Sports, 2007. https://nessis.org/nessis07/Kenny_Shirley.pdf (accessed 2025-10-03).
- Bottoms, L.; Tarragó, R.; Muñoz, D.; Chaverri, D.; Irurtia, A.; Castizo-Olier, J.; Carrasco, M.; Rodríguez, F. A.; Iglesias, X. Physiological Demands and Motion Analysis of Elite Foil Fencing. *PLOS ONE* **2023**, *18* (2), e0281600. <https://doi.org/10.1371/journal.pone.0281600>.
- Chen, T. L.-W.; Wong, D. W.-C.; Wang, Y.; Ren, S.; Yan, F.; Zhang, M. Biomechanics of Fencing Sport: A Scoping Review. *PLOS ONE* **2017**, *12* (2), e0171578. <https://doi.org/10.1371/journal.pone.0171578>.
- Zhang, M.; Chen, L.; Yuan, X.; Huang, R.; Liu, S.; Yong, J. Visualization of Technical and Tactical Characteristics in Fencing. *Journal of Visualization* **2018**, *22* (1), 109–124. <https://doi.org/10.1007/s12650-018-0521-3>.
- Lin, Z.; Goel, P.; Yun, J.; Liu, K.; Araujo, J. P. VirtualFencer: Generating Fencing Bouts Based on Strategies Extracted from In-The-Wild Videos. **2025**. <https://doi.org/10.48550/arXiv.2507.00261>.
- Chang, C.-C.; Lee, T.-L. Fencing Tactics Analysis in Broadcast Video: A Point-By-Point Analytical System. *IEEE Transactions on Circuits and Systems for Video Technology* **2019**, *29* (7), 2162–2175. <https://doi.org/10.1109/tcsvt.2018.2855098>.
- Zeng, Y. *Modeling Player Strength in Fencing*; The University of Iowa https://homepage.divms.uiowa.edu/~dzimmer/sports-statistics/slides_yiran.pdf (accessed 2025-10-03).

- Haight, F. A. *Handbook of the Poisson Distribution*; John Wiley & Sons. New York, NY, US, 1967.
- Ross, S. M. *The Exponential Distribution and the Poisson Process. In Introduction to Probability Models (Tenth Edition)*, Academic Press, 2010; pp. 291–370.
- Casella, G.; Berger, R. L. *Statistical Inference*, 2nd ed.; Thomson Learning: Belmont, CA, 2002; pp. 92–97.
- Fédération Internationale D'Esclime. *The Penalties and the Competent Juridical Authorities*. In Technical Rules; FIE, 2025; pp. 158–170.
- Ross, S. In Introduction to Probability Models, Eleventh Edition.; Academic Press, 2014.
- Wallis, S. Binomial Confidence Intervals and Contingency Tests: Mathematical Fundamentals and the Evaluation of Alternative Methods. *Journal of Quantitative Linguistics* **2013**, *20* (3), 178–208. <https://doi.org/10.1080/09296174.2013.799918>.
- Saltelli, A.; Tarantola, S.; Campolongo, F.; Ratto, M. *Sensitivity Analysis in Practice: A Guide to Assessing Scientific Models*; John Wiley & Sons, 2004; Vol. 1.

Author

Lynn Uemoto is a Junior at Franklin School (New Jersey, United States), with research interests in applied mathematics and statistics. She plans to pursue a degree in mathematics or statistics and continue exploring quantitative and risk analysis.

Appendix A: Derivation of the Negative Binomial Distribution

This appendix provides the mathematical derivation of the negative binomial distribution from first principles.

Imagine a sequence of independent Bernoulli trials (e.g., coin flips), each with a probability of success p and failure $1-p$. Let the random variable X represent the number of failures observed before the r -th success. The event $\{X=k\}$ occurs if and only if exactly $r-1$ successes occur in the first $k+r-1$ trials, and the $(k+r)$ -th trial is a success.

The number of ways to arrange $r-1$ successes and k failures in $k+r-1$ trials is given by the binomial coefficient:

$$\binom{k+r-1}{k} = \frac{(k+r-1)!}{k!(r-1)!}, \text{ or equivalently } \binom{k+r-1}{r-1}$$

The probability of any specific sequence with $r-1$ successes and k failures is:

$$p^{r-1}(1-p)^k$$

Therefore, the probability of getting the $(r-1)$ -th success exactly on the $(k+r-1)$ -th trial is:

$$\binom{k+r-1}{k} p^{r-1}(1-p)^k$$

We now need one more success on the next trial, or the final trial. The probability of this is p . Since the trials are independent, the joint probability is the product of the two probabilities above:

$$P(X=k) = \binom{k+r-1}{k} p^{r-1}(1-p)^k \cdot p$$

Simplifying the exponents of p gives us the classic PMF of the Negative Binomial distribution:

$$P(X=k) = \binom{k+r-1}{k} p^r (1-p)^k, \text{ for } k = 0, 1, 2, \dots$$

■ Appendix B: Unbounded Penalty Analysis

This appendix presents the results of the Markov state-dependent penalty model under unbounded conditions, where the simulation was run for a cumulative 29 touches (the maximum number of exchanges in a 15-14 bout). This is different from the bounded (race-to-15) model in the main paper. The unbounded model removes the variability in total touches, allowing us to observe the full evolution of the penalty state probabilities and get a smoother estimate of long-term penalty dynamics.

Results and Comparison with Bounded Model:

Under the same baseline conditions ($p_y = 0.10$, $p_r = 0.02$), the 10,000 unbounded simulations over 29 touches produced the following averages: 2.995 ± 1.722 yellow cards, 0.494 ± 0.609 red cards per bout (Table B1). The mean first penalty occurred at touch 7.78 ± 3.5 . As expected, the average number of yellow cards and red cards is higher than in the bounded model (2.909 ± 1.696 and 0.438 ± 0.586 , respectively), because bouts are artificially extended to their maximum possible length. The conditional win probabilities (Table B2) revealed a more significant effect in unbounded conditions, with Fencer A's win probability in penalty bouts at 0.832 compared to 0.864 in a case without penalties.

Table B1: Penalty Dispersion and Descriptive Statistics Unbounded (total 29 points). Unbounded simulations show higher average penalties than bounded bouts, confirming that longer exchanges increase the likelihood of penalty points.

Variable	Mean	Std. Dev.
yellow cards per bout	2.995	1.722
Red cards per bout	0.494	0.609
bouts with ≥ 1 penalty	97.14	N/A
touch of the first penalty	7.780	3.500

Table B2: Penalty Win Probabilities Unbounded (total 29 points). The win probability reduction due to penalties is slightly more pronounced in unbounded simulations, highlighting the increasing influence of penalty escalation in extended bouts.

Variable	Probability
win prob. (with penalties)	0.858
win prob. (no penalties)	0.864
win prob. (penalty bouts only)	0.832

The distribution of red cards (Figure B1) remains heavily right-skewed, but the proportion of bouts with two or more red cards increases compared to the bounded model. This demonstrates that if bouts were to consistently run longer, the cumulative penalty rules would lead to a higher frequency of point-awarding red cards. The state dynamics (Figure B2) show the clear, long-term trajectory that is truncated in a real bout. Compared to Figure 5 with the bounded model.

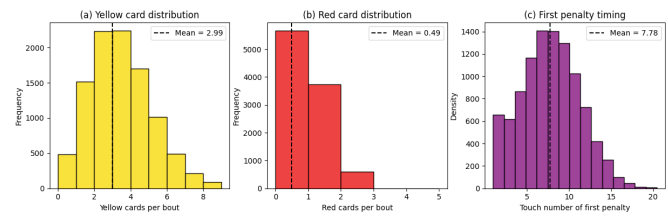


Figure B1: Marginal Penalty Distribution and Timing under unbounded conditions (total 29 points). Extending the bout length increases the frequency of red cards.

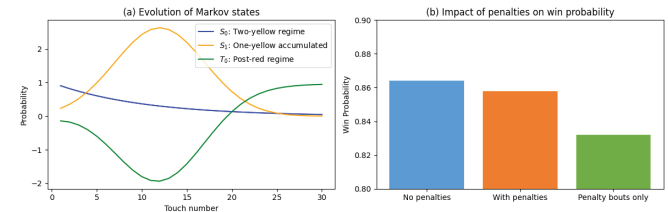


Figure B2: State Dynamics and Win Probability are unbounded (not depending on JPMF). Under unbounded conditions, penalty state evolution is smoother and more pronounced, illustrating the long-term trajectory of escalating card risk.