Beyond Expected Goals: A Probabilistic Framework for Shot Occurrences in Soccer

Tianshu Feng, Jonathan Pipping, & Paul Sabin

University of Pennsylvania

2025-07-29

What are Expected Goals (xG)?

- Expected Goals (xG) estimates the probability that a shot is scored
- Estimated by an XGBoost model trained on historical shot data
- Depends on factors like distance from goal, angle to goal, shot type, and player positions
- Often used to measure the quality of a chance or a team's performance across a game

Limitations of xG

- Selection bias: xG is only recorded for shots we observe!
- Better shooters are over-represented in the data
- Significant chances without a shot event aren't recorded
- Misses opportunities where players should have shot but didn't

Examples of xG Limitations

Video of big chance with no shot

Video of multiple shots on one attack

Methods

- Statistical approach
- Data processing steps
  - θ
- Model specifications

Results

- Key findings
- Statistical significance
- Practical implications

Conclusions

- Summary of main points
- Future work
- Questions and discussion

Thank You

Questions?