Conditional Laws for the Maximal Values of Anchored Random Walk Functionals

Jonathan Pipping Jiaoyang Huang Abraham J. Wyner

University of Pennsylvania, Department of Statistics & Data Science

Problem Setup

Setup: Let $(X_k)_{k=0}^N$ be a random walk with $X_0 = 0$ and i.i.d. increments ΔX_k of mean μ and variance σ^2, adapted to $(\mathcal{F}_k)$. For a position d at time k, define the functional

$$f(k, X_k) = \mathbb{P}(X_N > 0 \mid X_k = d)$$

Then the maximum of f over the time interval $[0, N]$ is

$$M = \max_{0 \le k \le N} f(k, X_k).$$

Question: What is the law of M conditional on the event $X_N < 0$? A closed-form solution is difficult to obtain, so we consider the limiting case.

Brownian Limit: By Donsker's invariance principle, as $N \to \infty$,

$$\frac{X_{\lfloor Nt \rfloor}}{\sigma\sqrt{N}} \xrightarrow{d} B_t + \mu^* t, \qquad \mu^* = \frac{\sqrt{N}\mu}{\sigma},$$

In the limit,

$$f(t, B_t) = \mathbb{P}(B_1 \ge -\mu^* \mid \mathcal{F}_t) = \Phi\left(\frac{B_t + \mu^*(1-t)}{\sqrt{1-t}}\right),$$

This represents a bounded martingale with $f(0, B_0) = \Phi(\mu^*)$ whose terminal value is in $\{0, 1\}$. The analogous maximum over the time interval $[0, 1]$ is

$$M = \sup_{0 \le t < 1} f(t, B_t),$$

and we study the law of M conditional on the event $B_1 < -\mu^*$.

Symmetric Random Walks

Functional: Since $\mu = 0$, the functional simplifies to

$$f(t, B_t) = \mathbb{P}(B_1 \ge 0 \mid \mathcal{F}_t) = \Phi\left(\frac{B_t}{\sqrt{1-t}}\right),$$

Unconditional Distribution: Optional stopping at the first hitting time of level x gives the unconditional distribution of M:

$$F_M(x) = \begin{cases} 0, & x < \frac{1}{2} \\ 1 - \frac{1}{2x}, & \frac{1}{2} \le x < 1 \\ 1, & x = 1 \end{cases}$$

Conditional Distribution: The conditional distribution of M given $B_1 < 0$ is

$$F_{M|B_1<0}(x) = 2 - \frac{1}{x}, \quad x \in [\tfrac{1}{2}, 1]$$

Density Function: Its corresponding density function is

$$f_{M|B_1<0}(x) = \frac{1}{x^2}, \quad x \in (\tfrac{1}{2}, 1)$$

Asymmetric Random Walks

Functional: In the presence of drift ($\mu^* \neq 0$), the functional

$$f(t, B_t) = \mathbb{P}(B_1 \ge -\mu^* \mid \mathcal{F}_t) = \Phi\left(\frac{B_t + \mu^*(1-t)}{\sqrt{1-t}}\right),$$

is a bounded martingale with $f(0, B_0) = \Phi(\mu^*)$.

Unconditional Distribution: Optional stopping at the first hitting time of level x gives the distribution of M:

$$F_M(x) = \begin{cases} 0, & x < \Phi(\mu^*) \\ 1 - \frac{\Phi(\mu^*)}{x}, & \Phi(\mu^*) \le x < 1 \\ 1, & x = 1 \end{cases}$$

Conditional Distribution: Conditional on $B_1 < -\mu^*$,

$$F_{M|B_1<-\mu^*}(x) = 1 - \frac{\Phi(\mu^*)}{\Phi(-\mu^*)} \cdot \frac{1-x}{x}, \quad x \in [\Phi(\mu^*), 1]$$

Density Function:

$$f_{M|B_1<-\mu^*}(x) = \frac{\Phi(\mu^*)}{\Phi(-\mu^*)} \cdot \frac{1}{x^2}, \quad x \in (\Phi(\mu^*), 1).$$

Complementary Processes

Definition: Consider the complementary martingales

$$p_t = \mathbb{P}(B_1 \ge -\mu^* \mid \mathcal{F}_t), \qquad q_t = 1 - p_t = \mathbb{P}(B_1 < -\mu^* \mid \mathcal{F}_t).$$

Terminal Behavior: At time 1, exactly one of $\{p_t, q_t\}$ converges to 1, while the other converges to 0 (a.s.). Our object of interest is defined as

$$M = \sup_{0 \le t < 1} \left(p_t \mathbf{1}_{\{p_1=0\}} + q_t \mathbf{1}_{\{q_1=0\}}\right),$$

i.e. the running supremum of whichever process ultimately vanishes.

Branch Distributions: For a process that starts at $s \in (0,1)$ and is conditioned to terminate at 0, denote by $F_s(x)$ the distribution function of its maximum:

$$F_s(x) = \mathbb{P}(\sup_{0 \le t < 1} p_t \le x \mid p_0 = s, p_1 = 0)$$

Distribution of the Eventual Vanishing Process: Let $p_0 := \Phi(\mu^*)$. Then the law of the eventual vanishing process is the mixture

$$F_{\text{vanish}}(x) = \Phi(\mu^*)\, F_{1-\Phi(\mu^*)}(x) + (1 - \Phi(\mu^*))\, F_{\Phi(\mu^*)}(x)$$

with support $x \in [\min\{\Phi(\mu^*), 1 - \Phi(\mu^*)\}, 1]$.

Monte Carlo Simulation

Validation: We validate our theoretical results with Monte Carlo simulations. The symmetric case is shown in Figure 1.

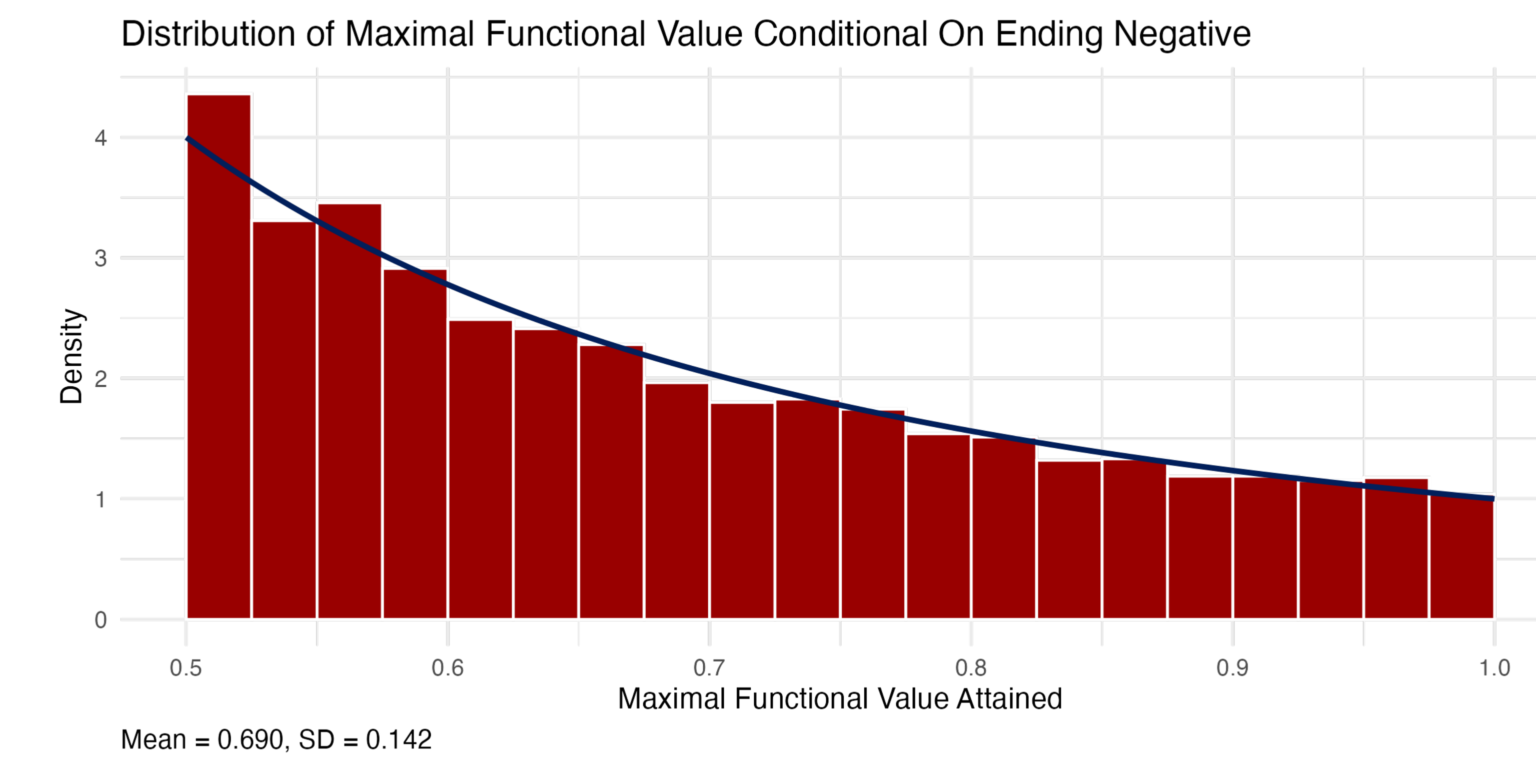

Figure 1. Symmetric case: simulated distribution (red) vs. theoretical density (blue).

Application to In-Game Win Probabilities

NFL Win Probabilities: Win probability trajectories from closely-matched games approximately follow our theoretical symmetric distribution.

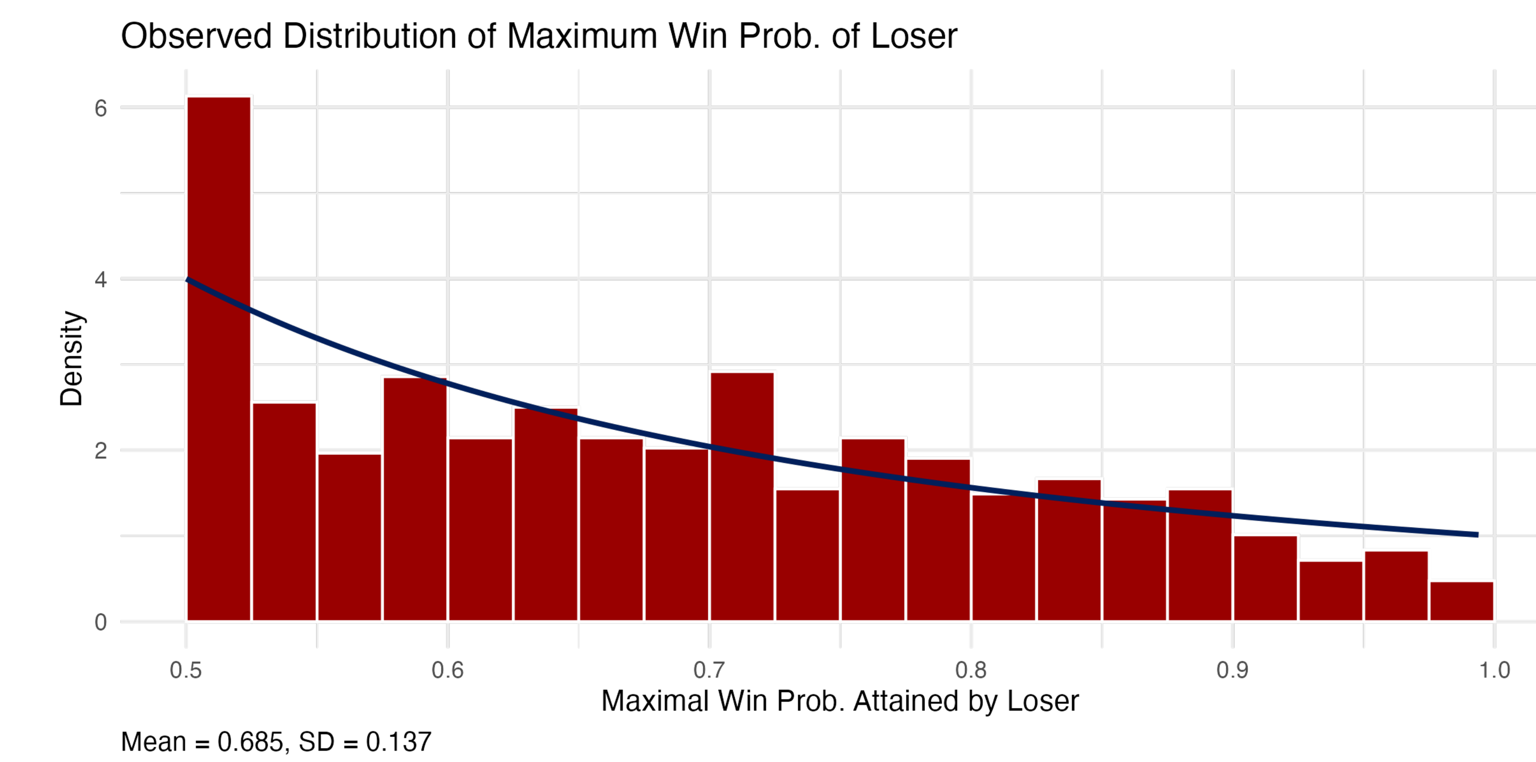

Figure 2. Distribution of maximal win probabilities of eventual NFL losers (2002-2024).

References

[1] Carl, S. and Baldwin, B. (2025). nflfastr: Functions to efficiently access nfl play by play data. R package version 5.1.0.9000.

[2] Donsker, M. D. (1951). An invariance principle for certain probability limit theorems. *Memoirs of the American Mathematical Society*, 6.

[3] Doob, J. L. (1953). *Stochastic Processes*. Wiley Publications in Statistics. John Wiley & Sons, New York.