Alice HH vs Bob HT using an Edgeworth Expansion

Mihai Nica, University of Guelph

June 11, 2024

Abstract

In the Alice HH vs Bob HT problem, Alice and Bob flip a sequence of n coins, with Alice scores a point whenever "HH" appears and Bob scoring a point whenever "HT" appear. In this note we calculate the first 3 moments of the random variable "Alice's Score minus Bob's Score" and then use an Edgeworth expansion approach to estimate that Bob's probability of winning exceeds Alice's probability of winning by $\frac{1}{2\sqrt{\pi n}}$.

1 Definitions

Symbol	Definition
$n \in \mathbb{N}$	Number of coinflips in the sequence
$C_1 C_2, \ldots, C_n$	Sequence of coinflips $C_i \in \{H, T\}$
$\Delta_i := \mathbb{1}_{C_i C_{i+1}=HH} - \mathbb{1}_{C_i C_{i+1}=HT}$	The difference (Alice score - Bob score) due to the coinflips C_i, C_{i+1}
$X_n = \sum_{i=1}^{n-1} \Delta_i$	Alices total Score minus Bob's total Score. The total difference in score after the n flips
$\mathbb{1}_{HH\bullet}$ and $\mathbb{1}_{\bullet HH}$	Used as a shorthand notation for the indicators of the events like $\mathbb{1}_{HH\bullet} := \mathbb{1}\{C_i C_{i+1} C_{i+2} = HH\bullet\}$ when the value of i is implicit. The bullet "$\bullet$" stands for a "wildcard" that could be either H or T, but makes calculations a bit easier to visualize when used in multiple places to keep i constant. For example: $\mathbb{1}_{HH\bullet} \cdot \mathbb{1}_{\bullet HH} = \mathbb{1}_{HHH}$ or $\mathbb{1}_{HH\bullet} \cdot \mathbb{1}_{\bullet HT} = \mathbb{1}_{HHT}$ or $\mathbb{1}_{HT\bullet} \cdot \mathbb{1}_{\bullet HT} = 0$ (since the middle coinflip cannot be both H and T simultaneously).

2 Moment Calculations

2.1 1st Moment

$$\mathbf{E}\left[X_n\right] = (n-1)\mathbf{E}\left[\Delta_i\right] = 0$$

2.2 2nd Moment

Note that Δ_i and Δ_j are independent unless $i = j$ or $i = j \pm 1$ (i.e. the "range" of the interactions is only 1 coinflip.) From this it follows that we can expand to get:

$$\mathbf{E}\left[X_n^2\right] = (n-1)\mathbf{E}\left[\Delta_i^2\right] + 2(n-2)\mathbf{E}\left[\Delta_i \Delta_{i+1}\right]$$

and we compuite:

$$\mathbf{E}\left[\Delta_i^2\right] = \mathbf{E}\left[\left(\mathbb{1}_{HH} - \mathbb{1}_{HT}\right)^2\right] = \mathbf{E}\left[\mathbb{1}_{HH} + \mathbb{1}_{HT} + 2 \cdot 0\right] = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$$

and:

$$\begin{aligned}
\mathbf{E}\left[\Delta_i \Delta_{i+1}\right] &= \mathbf{E}\left[\left(\mathbb{1}_{HH\bullet} - \mathbb{1}_{HT\bullet}\right)\left(\mathbb{1}_{\bullet HH} - \mathbb{1}_{\bullet HT}\right)\right] \\
&= \mathbf{E}\left[\mathbb{1}_{HHH} - \mathbb{1}_{HHT} - 0 + 0\right] \\
&= \frac{1}{16} - \frac{1}{16} = 0
\end{aligned}$$

so combining gives:

$$\mathbf{E}\left[X_n^2\right] = \frac{1}{2}(n-1)$$

2.3 3rd Moment

Again, keeping track of only interactions at range 1 away, we have the expansion:

$$\begin{aligned}\mathbf{E}\left[X_n^3\right] =&\mathbf{E}\left[\sum_{i,j,k=1}^{n}\Delta_i\Delta_j\Delta_k\right]\\ =&(n-1)\mathbf{E}\left[\Delta_i^3\right]+3(n-3)\cdot 3!\cdot\mathbf{E}\left[\Delta_i\Delta_{i+1}\Delta_{i+2}\right]\\ &+(n-2)\cdot\binom{3}{1}\mathbf{E}\left[\Delta_i^2\Delta_{i+1}\right]\\ &+(n-2)\cdot\binom{3}{1}\mathbf{E}\left[\Delta_i\Delta_{i+1}^2\right]\end{aligned}$$

Now we notice that any expectation of the form $\mathbf{E}\left[(\text{...product of stuff...})\cdot\Delta_{i+2}\right]$ (i.e. it ends with a Δ on its own and not squared) is exactly 0 because the the last Δ is equally likely to be $+1$ or -1 (either both 50% or both 0% depending on whether the product of stuff at the beginning end is a H or ends in a T.) So the only surviving term here is the term $\mathbf{E}\left[\Delta_i\Delta_{i+1}^2\right]$ which gives:

$$\begin{aligned}\mathbf{E}\left[\Delta_i\Delta_{i+1}^2\right] &= \mathbf{E}\left[\left(\mathtt{1}_{HH\bullet}-\mathtt{1}_{HT\bullet}\right)\left(\mathtt{1}_{\bullet HH}-\mathtt{1}_{\bullet HT}\right)^2\right]\\ &= \mathbf{E}\left[\left(\mathtt{1}_{HH\bullet}-\mathtt{1}_{HT\bullet}\right)\left(\mathtt{1}_{\bullet HH}+\mathtt{1}_{\bullet HT}+2\cdot 0\right)\right]\\ &= \mathbf{E}\left[\mathtt{1}_{HHH}+\mathtt{1}_{HHT}-0-0\right]\\ &= \frac{1}{8}+\frac{1}{8}=\frac{1}{4}\end{aligned}$$

So we have:

$$\mathbf{E}\left[X_n^3\right]=\frac{3}{4}(n-2)$$

3 Edgeworth Expansion

We use an expansion for the density function of the random variable of the form (see https://en.wikipedia.org/wiki/Edgeworth_se

$$\rho_X(x)\approx\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}\left(1+\frac{\kappa_3}{3!\sigma^3}He_3\left(\frac{x-\mu}{\sigma}\right)+\dots\right)$$

where $\mu=\mathbf{E}\left[X\right],\ \sigma=\sqrt{\mathbf{Var}\left[X\right]}$ and $\kappa_3=\mathbf{E}\left[(X-\mu)^3\right]$, and $He_3(x)=x^3-3x$ is the 3rd Hermite polynomial. In our case $\mu=0$ so this simplifies a bit to:

$$\rho_X(x)\approx\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{x^2}{2\sigma^2}}\left(1+\frac{\kappa_3}{3!\sigma^3}He_3\left(\frac{x}{\sigma}\right)+\dots\right)$$

3.1 A lemma about integrating He_3

Lemma 1. *Have that:*

$$\int_0^\infty\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{x^2}{2\sigma^2}}He_3\left(\frac{x}{\sigma}\right)dx=-\frac{1}{\sqrt{2\pi}}$$

Proof. Have:

$$\begin{aligned}\int_0^\infty \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}} He_3\left(\frac{x}{\sigma}\right) \mathrm{d}x &= \mathbf{E}_{Z\sim\mathcal{N}(0,\sigma^2)}\left[\mathbb{1}_{Z>0} He_3\left(\frac{Z}{\sigma}\right)\right] \\ &= \mathbf{E}_{Z\sim\mathcal{N}(0,1)}\left[\mathbb{1}_{Z>0} He_3(Z)\right] \\ &= \mathbf{E}_{Z\sim\mathcal{N}(0,1)}\left[\mathbb{1}_{Z>0} Z^3\right] - 3\mathbf{E}_{Z\sim\mathcal{N}(0,1)}\left[\mathbb{1}_{Z>0} Z\right] \\ &= \mathbf{E}_{Z\sim\mathcal{N}(0,1)}\left[\varphi(Z)^3\right] - 3\mathbf{E}_{Z\sim\mathcal{N}(0,1)}\left[\varphi(Z)\right] \\ &= (3-1)!!\frac{1}{\sqrt{2\pi}} - 3\frac{1}{\sqrt{2\pi}} \\ &= -\frac{1}{\sqrt{2\pi}}\end{aligned}$$

where we have used the relu function $\varphi(x) = x\mathbb{1}_{\{x>0\}}$ and the result for Gaussians that (which can be proved by a nice integration by parts induction)

$$\mathbf{E}\left[\varphi(Z)^k\right] = \begin{cases} \frac{(k-1)!!}{2} & k \text{ is even} \\ \frac{(k-1)!!}{\sqrt{2\pi}} & k \text{ is odd} \end{cases}$$

□

3.2 Alice - Bob using the Edgeworth Approximation.

Lemma 2. *If we use the Edgeworth approximation:*

$$\rho_X(x) \approx \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}} \left(1 + \frac{\kappa_3}{3!\sigma^3} He_3\left(\frac{x}{\sigma}\right) + \dots\right)$$

then we get:

$$\mathbf{P}(X>0) - \mathbf{P}(X<0) = -\frac{2\kappa_3}{3!\sigma^3}\frac{1}{\sqrt{2\pi}}$$

Proof. By making the chage of variable $x \to -x$ and using the fact that $He_3(x)$ is an odd polynomial:

$$\begin{aligned}\mathbf{P}(X<0) &= \int_{-\infty}^0 \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}} \left(1 + \frac{\kappa_3}{3!\sigma^3} He_3\left(\frac{x}{\sigma}\right) + \dots\right) \mathrm{d}x \\ &= \int_0^\infty \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}} \left(1 + \frac{\kappa_3}{3!\sigma^3} He_3\left(\frac{-x}{\sigma}\right) + \dots\right) \mathrm{d}x \\ &= \int_0^\infty \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}} \left(1 - \frac{\kappa_3}{3!\sigma^3} He_3\left(\frac{x}{\sigma}\right) + \dots\right) \mathrm{d}x\end{aligned}$$

Hence, when we subtract the probabilities we get:

$$\begin{aligned}\mathbf{P}(X>0) - \mathbf{P}(X<0) &= \int_0^\infty \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}} \left((1-1) + (1-(-1))\frac{\kappa_3}{3!\sigma^3} He_3\left(\frac{x}{\sigma}\right) + \dots\right) \mathrm{d}x \\ &= 2\frac{\kappa_3}{3!\sigma^3} \cdot \int_0^\infty \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}} He_3\left(\frac{x}{\sigma}\right) \mathrm{d}x + \dots \\ &= -\frac{2\kappa_3}{3!\sigma^3}\frac{1}{\sqrt{2\pi}}\end{aligned}$$

by the Lemma. □

Proposition 3. *Suppose that the Edgeworth approximation holds for the random variable X_n which is Alice's Score minus Bob's score. Then:*

$$\mathbf{P}\left(\textit{Alice wins}\right) - \mathbf{P}\left(\textit{Bob wins}\right) = -\frac{1}{2\sqrt{\pi n}} + \dots$$

Remark 4. Note that we know already from other methods that the result of the proposition is true. However, it is **NOT YET PROVEN** that the Edgeworth expansion actually holds (need to deal with the fact that the sequence Δ_i are not purely independent but are instead dependent). One relevant paper that seems to have the type of result needed to justify this Edgeworth expansion is MR0871198 > Heinrich, Lothar > Some remarks on asymptotic expansions in the central limit > theorem for > m-dependent random variables. > Math. Nachr. 122 (1985), 151–155.

Proof. By the lemmas we have:

$$\begin{aligned}
\mathbf{P}\left(\text{Alice wins}\right) - \mathbf{P}\left(\text{Bob wins}\right) &= \mathbf{P}\left(X_n > 0\right) - \mathbf{P}\left(X_n < 0\right) \\
&= -\frac{2\kappa_3}{3!\sigma^3}\frac{1}{\sqrt{2\pi}} + \ldots \\
&= -\frac{2\left(\frac{3}{4}(n-2)\right)}{3!\left(\frac{1}{2}(n-1)\right)^{3/2}}\frac{1}{\sqrt{2\pi}} + \ldots \\
&= -\frac{2\left(\frac{3}{4}\right)}{3!\left(\frac{1}{2}\right)\sqrt{\frac{1}{2}}}\frac{1}{\sqrt{2\pi}}\frac{(n-2)}{(n-1)^{3/2}} + \ldots \\
&= -\frac{1}{2\sqrt{\pi n}} + \ldots
\end{aligned}$$

where we have used $\frac{n-2}{(n-1)^{3/2}} = \frac{1}{\sqrt{n}} + \ldots.$ as $n \to \infty$. $\square$