Caution: These lecture notes are under construction. You may find parts that are incomplete.

1 Pythagorean Formula

1.1 When do we switch to preferring actual winning percentage?

n_i is the number of games played by team i
X_i is the Pythag W% of team i
Y_i is the actual W% of team i
Z_i is the residual W% of team i ($Y_i = X_i + Z_i$)

Intuition

$$\begin{aligned}\text{Actual W\%} &= \text{Pythag W\%} + \text{Residual}\\ \text{Outcome} &= \text{Skill} + \text{Luck}\\ Y_i &= X_i + Z_i\end{aligned}$$

Model

$$\begin{aligned}X_i &\sim \text{ind. Normal}(\mu_i,\, \sigma_X^2/n_i)\\ \mu_i &\sim \text{i.i.d. Normal}(\mu_0,\, \sigma_\mu^2)\end{aligned}$$

Option #1 (Z_i is all luck):

$$Z_i \sim \text{i.i.d. Normal}(0,\, \sigma_Z^2/n_i)$$

Option #2 (Z_i is not purely luck):

$$\begin{aligned}Z_i &\sim \text{ind. Normal}(\eta_i,\, \sigma_Z^2/n_i)\\ \eta_i &\sim \text{i.i.d. Normal}(0,\, \sigma_\eta^2)\end{aligned}$$

One can show that X_i does better than Y_i at estimating $E[Y_i]$ if $\sigma_\eta^2 < \sigma_Z^2/n$, i.e. $n < \sigma_Z^2/\sigma_\eta^2$.

1.2 Exercises

1. Show that X_i does better than Y_i at estimating $E[Y_i]$ if $n < \sigma_Z^2/\sigma_\eta^2$.