Two Simple Putting Models in Golf

Mark Broadie, Columbia Business School

This version: October 2018
• PGA Tour data from 2016-2018 represents more than 1.2 million putts
• Standard errors range from 0.01% to 0.3% (not shown for clarity)
• Horizontal axis (initial putt distance, in feet) shown in log-scale for clarity
One-putt: $|\alpha| \leq \alpha_c$, where $d$ is the distance to the hole, $r$ is the radius of the hole, and $\alpha_c = \tan^{-1}(r/d)$.

Suppose $\alpha \sim N(0, \sigma^2_\alpha)$. Then

$$P(\text{One-putt}) = P(|\alpha| \leq \alpha_c) = P(|Z| \leq \alpha_c / \sigma_\alpha)$$
$$= \Phi(\alpha_c / \sigma_\alpha) - \Phi(-\alpha_c / \sigma_\alpha) = 2\Phi(\alpha_c / \sigma_\alpha) - 1$$

Gelman and Nolan (2002)
Gelman and Nolan Model: Fit to PGA Tour Data

Model: $\alpha \sim N(0, \sigma_\alpha^2)$. Choose $\sigma_\alpha$ to minimize the sum of squared differences between the model and the data. Optimal: $\sigma_\alpha = 2.00^\circ$ (RMSE: 4.6%)

- Model probability $\to 1$ as $d \to 0$ and $\to 0$ as $d \to \infty$
- Model is biased low for $d < 8$ and biased high for $d > 8$
- See pga_tour_putt_data_models.xlsx for details
One-putt if endpoint in the **hole out region**: $|\alpha| \leq \alpha_c$ and the putt distance, $l$, satisfies $d \leq l \leq d + 3$.

Suppose $l = (d + 1)(1 + \sigma_d Z)$, $Z \sim N(0, 1)$, i.e., the target is one foot beyond the hole, $\sigma_d$ is the fractional distance error and $Z$ is independent of $\alpha$.

$$P(\text{One-putt}) = P(|\alpha| \leq \alpha_c)P(d \leq l \leq d + 3)$$

$$= P(|\alpha| \leq \alpha_c)P\left(\frac{-1}{\sigma_d(d+1)} \leq Z \leq \frac{2}{\sigma_d(d+1)}\right)$$

$$= \left(2\Phi\left(\frac{\alpha_c}{\sigma_\alpha}\right) - 1\right)\left(\Phi\left(\frac{2}{\sigma_d(d+1)}\right) - \Phi\left(\frac{-1}{\sigma_d(d+1)}\right)\right)$$
Random Dis and Dir Model: Fit to PGA Tour Data

- Optimal parameters: $\sigma_\alpha = 1.69^\circ$ and $\sigma_d = 7.96\%$ (RMSE: 0.3%)
- Model fits well for all distances
- See pga_tour_putt_data_models.xlsx for details