Nonlinear regression for Golf Putting

Mark Andrews
Psychology Department, Nottingham Trent University

✉ mark.andrews@ntu.ac.uk

Golf Putting Data

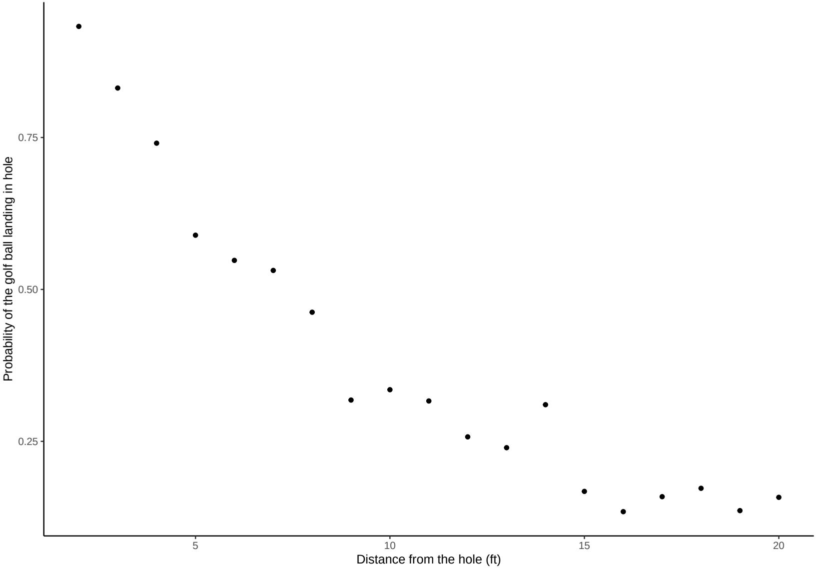
- This data provides the number of putting attempts and the number of successful putts at various distances (in feet) from the hole.

```
golf_df <- read_csv('golf_putts.csv')
```

- The absolute number of attempts and successes at each distance is vital information and so ideally we should base our analysis on this data, using a binomial logistic regression or a related model.
- However, for simplicity here, we will just use the relative frequencies of successes at each distance.

```
golf_df %<>% mutate(prob = success/attempts)
```

Golf Putting Data



A model

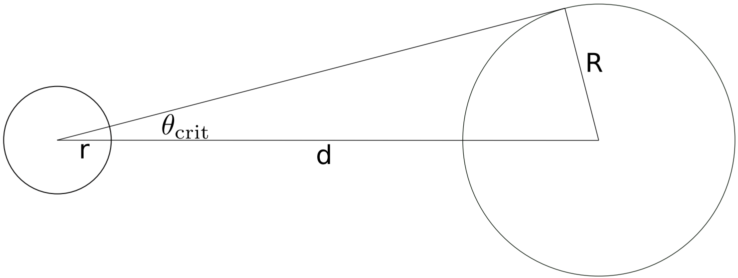


Figure 1: A golf ball of radius r (left) and the golf hole of radius R (right). The centres of these two circles are d apart. If the golf ball travels in a straight vertical line to the hole, it will fall in. If its trajectory deviates, either to the right or to the line, greater than an angle of θ_{crit}, it will miss. The angle θ_{crit} is the angle between the vertical line of length d and the tangent line from the centre of the ball to the hole. The line from the centre of the hole meets the tangent line at a right angle. As such, $\theta_{\text{crit}} = \sin^{-1}\left(\frac{R}{d}\right)$.

A physical model

The probability that the angle of their putt will be between 0 and θ_{crit} is

$$P(0 < \theta \leqslant \theta_{crit}) = \Phi(\theta_{crit}|0, \sigma^2) - \tfrac{1}{2},$$

where

$$\Phi(\theta_{crit}|0, \sigma^2) \triangleq \int_{-\infty}^{\theta_{crit}} N(\theta|0, \sigma^2),$$

which is the value at θ_{crit} of the cumulative distribution function of a normal distribution of mean 0 and standard deviation σ. We simply double the quantity $\Phi(\theta_{crit}|0, \sigma^2) - \frac{1}{2}$ to get $P(\theta_{crit} < \theta \leqslant \theta_{crit})$.

Therefore, the probability of a sucessful putt is

$$2\Phi\left(\sin^{-1}\left(\tfrac{R}{d}\right)|0, \sigma^2\right) - 1.$$

Implementation

- This is a nonlinear parametric function of distance d, where R is known to have a value of 53.975mm, and σ is the single unknown parameter.
- This nonlinear function is easily implemented as follows.

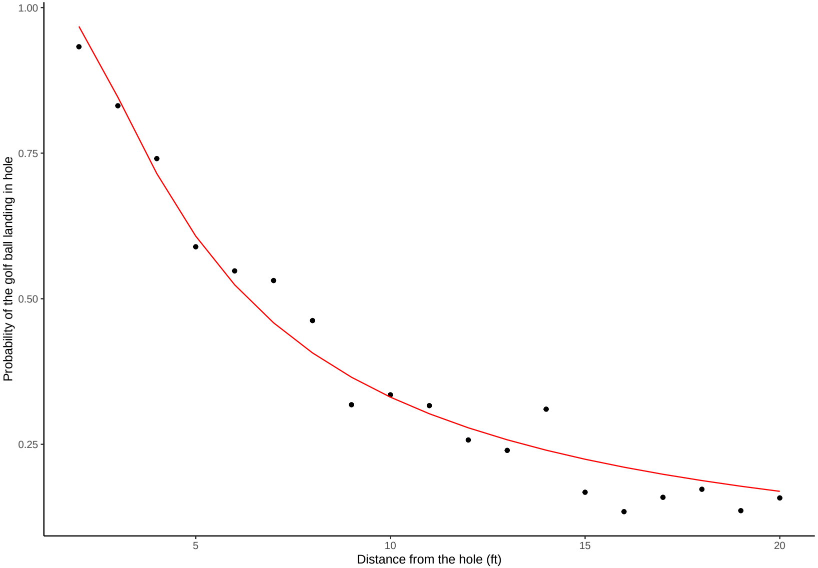
```
successful_putt_f <- function(d, sigma){
  R <- 0.17708333 # 53.975mm in feet
  2 * pnorm(asin(R/d), mean=0, sd=abs(sigma)) -1
}
```

Using `nls`

The `nls` based model using this `successful_putt_f` function is as follows.

```
M_putt <- nls(prob ~ successful_putt_f(distance, sigma),
              data = golf_df,
              start = list(sigma = 0.1)
)
```

The predictions



An simpler alternative?

```
M_putt_exp <- nls(prob ~ a + b * exp(-beta * distance),
                  data = golf_df,
                  start = list(a = 0, b = 1, beta = 1)
)
```