Theory

Baseline Model: No Censoring, No Covariates, Single Sequence

Draw y_1, y_2, ... from $y_i \sim \text{Bernoulli}(\pi)$ until $y_i = 1$. This sampling procedure produces a seqence of $n-1$ zeros and a single one, where n is a random variable. In fact, $n \sim \text{geometric}(\pi)$.

Use $\hat{\pi} = \bar{y}$ to estimate π. Is $\hat{\pi}$ an unbiased estimator of π, so that $E(\bar{y}) = \pi$?

First, notice that, by construction, $\bar{y} = \frac{1}{n}$ (remember that n is a random variable).

But what is the distribution of $\frac{1}{n}$?

$$
\begin{aligned}
P\left(\frac{1}{n} = 1\right) &= \pi \\
P\left(\frac{1}{n} = \frac{1}{2}\right) &= (1-\pi)\pi \\
P\left(\frac{1}{n} = \frac{1}{3}\right) &= (1-\pi)^2\pi \\
&\vdots \\
P\left(\frac{1}{n} = \frac{1}{k}\right) &= (1-\pi)^{k-1}\pi \\
&\vdots
\end{aligned}
$$

Then

$$E\left(\frac{1}{n}\right) = \frac{\pi}{1-\pi}\sum_{i=1}^{\infty}\frac{1}{i}(1-\pi)^i$$

.

The series $\frac{1}{i}(1-\pi)^i$ converges because $(1-\pi) \leq 1 \leq |\frac{1}{i}|^{-\frac{1}{i}}$ for all i. (See radius of convergence for a power series.)