Heuristics & Biases:

The law of small numbers

EE434 SEM 1/2022
Sunsiree Kosindesha

The weak law of large number

In statistics, the weak law of large numbers is often expressed as follows.

Let $\{X_n\}$ be a sequence of independently and identically distributed random variables, with a common mean μ and finite variance σ^2. Denote the sample mean based on n observations by $\bar{X}(n) = \sum_{i=1}^{n} X_i / n$. Then $\bar{X}(n)$ "converges in probability" to the population mean, μ, i.e.,

$$\lim_{n \to \infty} P[|\bar{X}(n) - \mu| \geq \epsilon] = 0 \ \ \forall \epsilon > 0$$

By contrast,

Individuals are said to subscribe to **the law of small numbers** if they believe that **for any sample size**, the sample mean is identical to the population mean.

In other words,

Such people believe that the sample proportions mimic the population proportions for any sample size.

A formal model of the law of small numbers (Rabin, 2002)

Rabin, M. (2002). Inference by believers in the law of small numbers. Quarterly Journal of Economics 117: 775-816.

1.

Move by Nature

Nature chooses a parameter $\theta \in \Theta$. Individuals have a priori distribution $\pi(\theta)$ over the set Θ.

2.

The *actual* data generating process:

Consider a Bernoulli trial with two possible outcomes:

$x(success)$ and $y(failure)$ that occur with respective probabilities $\theta \in [0,1]$ and $1 - \theta$.

Suppose that we independently repeat the Bernoulli trial n times, with replacement.

Denote by n_x, the number of times that the outcome equals x in the n trials.

2. (cont.)

The *actual* data generating process:

Then, the number of successes follows a binomial distribution; the probability that there are n_x successes out of n Bernoulli trials is given by:

$$p(n_x) = \binom{n}{n_x} \theta^{n_x} (1-\theta)^{n-n_x} \ ; n_x = 0,1, \dots, n.$$

2. (cont.)

The *actual* data generating process:

The true underlying process is one with replacement.

An individual who knows the true model of the world is known as a *Bayesian*.

3.

Misperception about the data generating process:

Suppose that individuals believe that the data is generated by the following process.

An urn has N balls.

θN balls denote outcome x and

the remaining $(1-\theta)N$ balls denote outcome y.

3.(cont.)

Misperception about the data generating process:

A ball is drawn with replacement at time t.

The probability of outcome x is $\frac{\theta N}{N} = \theta$.

Suppose that outcome x materializes at time t.

3.(cont.)

Misperception about the data generating process:

At time $t + 1$, the individual believes that there are $N - 1$ balls left in the urn.

The probability of outcome x is $\frac{\theta N - 1}{N - 1}$.

The probability of outcome y is $\frac{(1-\theta)N}{N-1}$.

3.(cont.)

Misperception about the data generating process:

A success today implies the probability of success tomorrow to be less than θ.

A success today implies the probability of failure tomorrow to be more than $1 - \theta$.

The proportion of successes/failures must balance out to the population rate for N signals observed.

3.(cont.)

Misperception about the data generating process:

- Such an individual, who believes in an incorrect model of the world, but is otherwise skilled in all aspects of classical statistical inference, is referred to as an ***N-Freddy***.

3.(cont.)

Misperception about the data generating process:

- If N is very large, then N-Freddy behaves very much like a Bayesian but for small N he can be very biased.