Example of Experimental Design in Biology: Deciding the Sample Size Needed

This example concerns a Biology student who took STAT 3332 and then became involved in a research project during the summer. In planning the experiment, a key question was how big a sample was needed. This question had not been covered in the course, so the student quite naturally and appropriately consulted a statistician (myself, in this case).

The Experiment

The project involved nematodes (round worms), in particular a hermaphrodite/male species, *Pristionchus pacificus*. The initial question was what amount of data would be needed in order to have significant results.

The males are produced when a hermaphrodite and a male mate, and also males are produced due to non-disjunction when the hermaphrodite self-fertilizes. The goal of the experiment was to determine the rate of non-disjunction in the hermaphrodite self-progeny. Hermaphrodites are XX and males are XO (only one X chromosome). When the hermaphrodite sperm fertilizes the oocyte during meiosis, both of the X chromosomes of the sperm forms a XXX gamete, and so there is no X chromosome donated to the other gamete, which then remains XO (male). Thus the goal was determine the rate of XO (males) in the self progeny of the *Pristionchus pacificus* hermaphrodites.

To do this, 1 hermaphrodite will be placed onto a petri dish for a total of N plates. The hermaphrodite will be allowed to lay eggs the first day, and the 2nd day she will be transferred to another clean petri dish where she will continue laying eggs. This process is continued until the hermaphrodite stops laying eggs. The eggs will be counted daily, and when the progeny are old enough, they will be sexed as male or hermaphrodite. The worms are about 1 mm in length and are transparent, so it is easy to determine the sex of the progeny by looking at their phenotype. Each hermaphrodite will lay about 100 eggs throughout her lifespan (about 4 days). If, as a group, there are N plates of 1 hermaphrodite each, and they all lay 100 eggs each, the data will show male vs. hermaphrodite for 100N self progeny at the end of the experiment. In another similar species, *Caenorhabditis elegans*, research has shown that the rate of non-disjunction in hermaphrodite self-progeny is approximately 1 male for every 1000 self-progeny.

Assuming that the rate of non-disjunction in *Pristionchus pacificus* is probably more or less near that number, how many total self-progeny (assuming that each hermaphrodite will lay 100 eggs) would be needed to determine the rate of non-disjunction in *Pristionchus pacificus* hermaphrodites within acceptable error limits?

Remarks

- It became planned to acquire 9000 progeny. Before the Statistician became involved, the experimenters were thinking of only 3000 progeny, which would have been insufficient to obtain sufficiently accurate estimates.
- With 9000 progeny, the experiment successfully estimated the rate within desired confidence, and the research results became published with the student as co-author.
- The student also wrote this project up as a Senior Honors Thesis.
- The student became accepted to medical school.