Mathematics 1550H – Probability I: Introduction to Probability
TRENT UNIVERSITY, Summer 2020 (S62)
Assignment #2
Cunning Experimental Design

In each of the following problems you are given a tool and asked to design an experiment using that tool that meets certain requirements.

1. You are given a coin, whether fair or not you do not know, and asked to use it to randomly generate a yes or no answer, with the stipulation that the probability of a yes should be equal to the probability of a no. How can you do the job without having to determine the possible bias of the given coin? Explain why your method works. *[3]*

SOLUTION. Toss the possibly biased coin twice. If the outcome is HT, record a "yes"; if the outcome is TH, record a "no"; and if the outcome is HH or TT, repeat the experiment as necessary until you get HT or TH.

Why does this work? First, if $P(H) = p$ and $P(T) = q = 1 - p$ for a single toss of the coin, then $P(HT) = pq = qp = P(TH)$, so the probability of "yes" is equal to the probability of "no". Second, this process has a probability of 1 of eventually giving you a "yes" or "no" answer instead of having the experiment repeat forever, at least as long as $0 < p < 1$. (So no two-headed or two-tailed coins allowed ... :-) The reason is that if $r = P(HH) + P(TT) = p^2 + q^2$ is the probability of having to repeat the experiment and $0 < p < 1$, then $r < 1$ (Why?), so the probability of having to repeat the experiment forever is $\lim_{n\to\infty} r^n = 0$. $\square$

2. You are given a fair coin and asked to use to to randomly generate a yes or no answer, with the stipulation that the probability of a yes should be exactly $\frac{8}{13}$ and the probability of a no should be exactly $\frac{5}{13}$. How can you do the job? Explain why your method works. *[3]*

SOLUTION. Toss the fair coin four times. The sixteen possible outcome are equally likely because the coin is fair. If the outcome is $HHHH$, $HHHT$, $HHTH$, $HTHH$, $THHH$, $HHTT$, $HTHT$, or $HTTH$, record a "yes"; if the outcome is $THTH$, $THHT$, $TTHH$, $HTTT$, or $THTT$, then record a "no"; and if the outcome is $TTTH$, $TTHT$, or $TTTT$, repeat the experiment until you get a "yes" or a "no".

Why does this work? First, the probability of getting a "yes" on any single run of the experiment is $\frac{8}{16}$ and the probability of getting a "no" on any single run of the experiment is $\frac{5}{16}$, so the odds of a "yes" *versus* a "no" are $8 : 5$. Since the probability of having to repeat the experiment forever is 0 – the probability of having to repeat it after any given run is $\frac{3}{16} < 1$, $\lim_{n\to\infty} \left(\frac{3}{16}\right)^n = 0$ – it follows that the probability of eventually getting a "yes" is $\frac{8}{8+5} = \frac{8}{13}$ and of eventually getting a "no" is $\frac{5}{8+5} = \frac{5}{13}$, as required. $\square$

3. You are given a fair coin and asked to use to to randomly generate a yes or no answer, with the stipulation that the probability of a yes should be exactly $\frac{1}{\sqrt{2}}$ and the probability of a no should be exactly $1 - \frac{1}{\sqrt{2}}$. How can you do the job? Explain why your method works. *[4]*

NOTE: **3** is different from **2** in that $\frac{1}{\sqrt{2}}$ is an irrational number, and so cannot be expressed precisely as a ratio of integers. This means that the methods most people find to solve **2** do not really work for **3**.

Hint: To solve **3** it helps to think of $\frac{1}{\sqrt{2}}$ in terms of its decimal or, even better, it's binary expansion.

SOLUTION. Just to keep as much familiar as possible, we'll use the decimal expansion of $\frac{1}{\sqrt{2}} = 0.707106781186\dots$ instead of the binary expansion. This decimal expansion goes on forever and does not eventually fall into a repeating pattern because $\frac{1}{\sqrt{2}}$ is irrational. In principle, we – or at least our computers – can work out as large a finite piece of this expansion as we like, so we'll assume that whenever we might need more of the expansion, we just compute it.

We will have a subsidiary experiment that uses the fair coin to generate one of the decimal digits 0 through 9 with equal probability. This will be done by tossing the coin four times, as in the solution to **2** above, and assigning one of the sixteen equally likely to each of the digits 0 through 9. If one of the remaining six outcomes occurs, repeat this subsidiary experiment as necessary until it generates a decimal digit. Reasoning similar to that in the solution to **2** tells us that each digit is equally likely to come up (so each has a probability of $\frac{1}{10} = 0.1$ of occurring) and that a digit will be generated eventually with probability 1.

Our main experiment works as follows. Use the subsidiary experiment to generate a decimal digit. If this digit is less than the first digit of the decimal expansion of $\frac{1}{\sqrt{2}}$, *i.e.* any one of 0 through 6, record a "yes", and if this digit is greater than the first digit of the decimal expansion of $\frac{1}{\sqrt{2}}$, *i.e.* either of 8 or 9, record a "no". If the digit is the same as the first digit of the decimal expansion of $\frac{1}{\sqrt{2}}$, *i.e.* 7, repeat the experiment except that the digit generated will be compared to the second digit in the decimal expansion of $\frac{1}{\sqrt{2}}$ instead. In general, if we have to perform the experiment for the nth time, we compare the digit generated to the nth digit of the decimal expansion of $\frac{1}{\sqrt{2}}$, recording a "yes" if the generated digit is less than the n digit of the decimal expansion of $\frac{1}{\sqrt{2}}$, a "no" if the generated digit is greater than the nth digit of the expansion of $\frac{1}{\sqrt{2}}$, and repeating the experiment if the generated digit is equal to the nth digit of the expansion of $\frac{1}{\sqrt{2}}$.

Why does this work? First, the process will eventually terminate unless the sequence of randomly generated digits is actually equal to the decimal expansion of $\frac{1}{\sqrt{2}}$, but this has probability $\lim_{n\to\infty} \left(\frac{1}{10}\right)^n = 0$ since the probability that each generated digit is equal to the one in the same place in the expansion of $\frac{1}{\sqrt{2}}$ is $\frac{1}{10}$.

Second, any real number in the interval $[0, 1]$ is as likely to be generated by randomly generating the digits of its decimal expansion as any other. (If we look at the process this way, note that we are only generating enough of the decimal expansion of that real number to tell if it is less than or greater than $\frac{1}{\sqrt{2}}$.) Thus the odds that the number will land in $\left[0, \frac{1}{\sqrt{2}}\right]$ rather than $\left[\frac{1}{\sqrt{2}}, 1\right]$ are proportional to the lengths of the intervals, namely $\frac{1}{\sqrt{2}} : \left(1 - \frac{1}{\sqrt{2}}\right)$. Since the two intervals have total length 1, it follows that the probability of getting a real number less than $\frac{1}{\sqrt{2}}$, *i.e.* a "yes", is $\frac{1}{\sqrt{2}}$, and the probability of getting a real number greater than $\frac{1}{\sqrt{2}}$, *i.e.* a "no", is $1 - \frac{1}{\sqrt{2}}$. (There is a somewhat similar discussion in Example 2.18 in Section 2.2 of the textbook for binary sequences.) ■