Assignment 3 – Question 1

This question is similar to Summer Exam 2014, Question 1 (iii).

Suppose that a fair coin is tossed until a head is obtained, and that this entire experiment is then performed a second time. Find the probability that the second experiment requires more tosses than the first experiment.

We conduct a two-stage experiment.

Stage 1. A fair coin is tossed repeatedly until a head occurs. The sample space is the set $S_1 = \{\text{H}, \text{T H}, \text{T T H}, \text{T T T H}, \ldots\}$.

Stage 2. A fair coin is tossed repeatedly until a head occurs. The sample space is the set $S_2 = \{\text{H}, \text{T H}, \text{T T H}, \text{T T T H}, \ldots\}$.

The sample space S for the experiment is the Cartesian product of S_1 and S_2, i.e. $S = S_1 \times S_2 = \{(\text{H}, \text{H}), (\text{H}, \text{T H}), (\text{T H}, \text{H}), (\text{H}, \text{T T H}), (\text{H}, \text{T T T H}), (\text{T T H}, \text{H}), \ldots\}$.

An obvious question to ask is whether **Stage 1** terminates, so that we can move to **Stage 2**. Indeed, it terminates with the probability 1.

Stage 1. We conduct a sequence of independent Bernoulli trials with probability of success $\frac{1}{2}$. Let X_1 be a random variable on S_1 that represents the number of trials required to obtain the first success.
Model: $X_1 \sim \text{GEO}(\frac{1}{2})$.

Stage 2. We conduct a sequence of independent Bernoulli trials with probability of success $\frac{1}{2}$. Let X_2 be a random variable on S_2 that represents the number of trials required to obtain the first success.
Model: $X_2 \sim \text{GEO}(\frac{1}{2})$.

Problem: Calculate the probability of the event

$$\{X_2 > X_1\} = \{(e_1, e_2) \in S : X_2(e_2) > X_1(e_1)\}.$$

Approach: We apply the Law of Total Probability to calculate $P[X_2 > X_1]$. As a partition of the sample space, we choose the countably infinite set of the events

$$\{X_1 = k\} = \{(e_1, e_2) \in S : X_1(e_1) = k\}, \text{ where } k = 1, 2, 3, \ldots,$$

and get

$$\begin{aligned} P[X_2 > X_1] &= \sum_{k=1}^{\infty} P[X_2 > X_1 | X_1 = k] \cdot P[X_1 = k] \\ &= \sum_{k=1}^{\infty} P[X_2 > k | X_1 = k] \cdot P[X_1 = k]. \end{aligned}$$

But the random variables X_1 and X_2 are independent by the design of our experiment. Hence, $P[X_2 > k|X_1 = k] = P[X_2 > k]$. Our task reduces to calculating the infinite series $\sum_{k=1}^{\infty} P[X_2 > k] \cdot P[X_1 = k]$.

The following facts will be useful:

① If $X \sim \text{GEO}(p)$ then $P[X > k] = (1-p)^k$ for any integer $k \geq 1$.

② $\sum_{k=1}^{\infty} r^k = \left(\sum_{k=0}^{\infty} r^k\right) - 1 = \frac{1}{1-r} - 1 = \frac{r}{1-r}$ for any real number r with $|r| < 1$.

To prove the first fact, observe that $P[X > k] = 1 - P[X \leq k] = 1 - \sum_{\alpha=1}^{k} p(1-p)^{\alpha-1}$. So, we need to prove that

$$1 - \sum_{\alpha=1}^{k} p(1-p)^{\alpha-1} = (1-p)^k,$$

which can be done by induction on k.

We can write a little bit of code in R to verify the result numerically as follows: We repeat the experiment n times and observe whether or not a particular event E occurs during each experiment. Let X_n denote the number of occurrences of E we have observed in n runs. Then we can use the Law of Large Numbers to deduce that the relative frequency $\frac{1}{n}X_n$ converges stochastically to $P(E)$ as n goes to infinity – broadly speaking, as more replicates of the experiment are performed, the proportion of occurrences of E gets closer to the exact value of the probability of E.

Summer Exam 2009, Question 5 (iii) is based on this empirical phenomenon, which we will explain later – in the context of limiting distributions.

Consider a sequence of n Bernoulli trials, with probability of success denoted by p, and let X_n denote the number of successes. Show that the sequence (X_n/n) converges in probability to p.

```
run_count = 10000;
greater_count = 0;

for (i in 1:run_count) {
  x = rgeom(2, prob = 1/2);
  if (x[2] > x[1]) {greater_count = greater_count + 1}
}
eval(greater_count/run_count);
```

We set up our simulation to run for 10,000 times, in order to obtain a good estimate of the true value of $P[X_2 > X_1]$, which is $\frac{1}{3}$.

"How many times do we need to run the experiment?" is an obvious question to ask, which we will deal with later – in the context of confidence intervals.

The coolest thing about computer experiments is that they often suggest an approach for formal proof and not just make it possible to guess the exact values or confirm the theoretically derived results. In other words, we can use computer experiments to discover a proof that is later written down in a rigorous way. Because proofs don't fall upon us like rain from the sky!

Assignment 3 – Question 4

We will use this question to check our solution to Summer Exam 2014, Question 1 (iii).

> Suppose that k events $A_1, \ldots, A_k$ form a partition of the sample space S. For $i = 1, \ldots, k$, let $P(A_i)$ denote the prior probability of A_i. Also, for any event B such that $P(B) > 0$, let $P(A_i|B)$ denote the posterior probability of A_i given that the event B has occurred. Prove that if $P(A_1|B) < P(A_1)$ then $P(A_i|B) > P(A_i)$ for at least one value of $i = 2, \ldots, k$.

We need to prove the following statement: If $P(A_1|B) < P(A_1)$ then there exists an index i in the range $2 \leq i \leq k$ such that $P(A_i|B) > P(A_i)$. We will do this indirectly, by assuming that the hypothesis "$P(A_1|B) < P(A_1)$" is true and the conclusion "$P(A_i|B) > P(A_i)$ for some i in $\{2, \ldots, k\}$" is false. We will then try to come up with a chain of logical arguments until we arrive at a contradiction.

We start by negating the proposition "There exists an index i in the range $2 \leq i \leq k$ such that $P(A_i|B) > P(A_i)$". The negation is "$P(A_i|B) \leq P(A_i)$ holds for all indices i in the range $2 \leq i \leq k$".

We assume that $P(A_1|B) < P(A_1)$ and $P(A_i|B) \leq P(A_i)$ for all i in $\{2, \ldots, k\}$. We apply the Law of Total Probability to $P(B)$ and get

$$P(B) = \sum_{i=1}^{k} P(B|A_i)P(A_i) = \sum_{i=1}^{k} P(A_i|B)P(B).$$

Note that, since $P(B) \neq 0$, it is possible to divide both sides of the above equation by $P(B)$.

$\ldots$, which is a contradiction. Hence, our assumption "$P(A_i|B) \leq P(A_i)$ for all i in $\{2, \ldots, k\}$" was false.