IEOR 3106, Fall 2013, Professor Whitt

Topics for Discussion: Thursday, November 21

Renewal Theory: Patterns

1. Patterns: see §3.6.4 and §7.9

Consider successive independent flips of a biased coin. On each flip, the coin comes up heads (H) with probability p or tails (T) with probability $q = 1 - p$, where $0 < p < 1$. A given segment of finitely many consecutive outcomes is called a *pattern*. The pattern is said to occur at flip n if the pattern is completed at flip n. For example, the pattern $A \equiv HTHTHT$ occurs at flips 8 and 10 in the sequence $TTHTHTHTHTTTTHHHT\ldots$ and at no other times among the first 17 flips.

WARMUP

For parts (a) and (b) below, assume that $p = 1/2$, but for later parts do not make that assumption.

(a) Which pattern occurs more frequently in the long run: $A \equiv HHH$ or $B \equiv HTH$?

(b) For patterns A and B in part (a), let N_A and N_B be the numbers of flips until the patterns A and B, respectively, first occur. Is $E[N_A] = E[N_B]$?

MAIN PROBLEM

Now we revert to general probabilities p and $q = 1 - p$.

(c) What is the probability that pattern $A \equiv HTHTHT$ occurs at flip 72?

(d) Suppose that pattern A from part (c) does indeed occur at flip 72. What is the expected number of flips until pattern A occurs again?

(e) Let $N_A(n)$ be the number of occurrences of pattern A in the first n flips, where A is again the pattern in part (c). Does

$$\frac{N_A(n)}{n} \to x \quad \text{as} \quad n \to \infty \quad \text{w.p.1?}$$

If so, what is the limit x?

(f) What is $E[N_A]$, the expected number of flips until pattern $A \equiv HTHTHT$ *first* occurs?

(g) What is the probability that pattern A occurs before pattern $B \equiv TTH$? That is, what is $P(N_A < N_B)$?

2. Answers

(a) Which pattern occurs more frequently in the long run: $A \equiv HHH$ or $B \equiv HTH$?

Since $p = q = 1/2$, we have $P(A(n)) = P(B(n)) = 1/8$ for all $n \geq 3$. Thus the two patterns occur equally often in the long run.

(b) For patterns A and B in part (a), let N_A and N_B be the numbers of flips until the pattern first occurs. Is $E[N_A] = E[N_B]$?

No, we do *not* have $E[N_A] = E[N_B]$. See below and at the very end.

(c) What is the probability that pattern $A \equiv HTHTHT$ occurs at flip 72?

For any pattern C, let $C(n)$ be the event that pattern C occurs at time (flip) n. Then $P(C(n))$ is the probability of event $C(n)$, i.e., the probability that pattern C occurs at flip n. This question is very easy to answer: With general probabilities p and $q \equiv 1 - p$,

$$P(A(n)) = p^3 q^3, \quad n \geq 6 .$$

That is because the specified outcomes must occur at flips n, $n-1$, $n-2$, $n-3$, $n-4$ and $n-5$. We simply multiply the probabilities for independent events. We require $n \geq 6$, because this pattern is of length 6; it cannot occur before flip 6. Observe that the limiting value as $n \to \infty$ already occurs at $n = 6$; we have a common value for all $n \geq 6$. The limit is attained at a finite value of n.

(d) Suppose that pattern A does indeed occur at flip 72. What is the expected number of flips until pattern A occurs again?

We invoke renewal theory. We observe that **the times (flips) when the event occurs are renewals.** (Of course that is why we are discussing this problem while we are reading Chapter 7.) Note that here we have a **delayed renewal process**. The times between successive renewals are IID. We have a *delayed renewal process* because the time until the first pattern occurrence in general has a distribution that is *different* from the distribution of the number of flips between renewals. Let $N_A(n)$ be the number of times pattern A has occurred in the first n flips.

First we observe that

$$E[N_A(n)] = \sum_{k=1}^{n} P(A(n)) ,$$

so that, by the reasoning above for part (c),

$$\frac{E[N_A(n)]}{n} \to p^3 q^3 \quad \text{as} \quad n \to \infty .$$

Let T_A be the time between successive occurrences of event A. By Theorem 7.1 of Ross, which extends to delayed renewal processes,

$$\frac{E[N_A(n)]}{n} \to \frac{1}{E[T_A]} \quad \text{as} \quad n \to \infty .$$

Moreover, by the LLN (law of large numbers) for delayed renewal processes, we have

$$\frac{N_A(n)}{n} \to \frac{1}{E[T_A]} .$$

see Proposition 7.1 in Ross. As a consequence, we must have

$$E[T_A] = \frac{1}{P(A(n))} \quad \text{for} \quad n \quad \text{suitably large} .$$

Here, in our specific context,

$$E[T_A] = p^{-3}q^{-3}$$

.

(e) Let $N_A(n)$ be the number of occurrences of pattern A in the first n flips, where A is the pattern in part (c). Does

$$\frac{N_A(n)}{n} \to x \quad \text{as} \quad n \to \infty \quad \text{w.p.1?}$$

If so, what is the limit x?

We already used this result to answer the last question.

$$\frac{N_A(n)}{n} \to \frac{1}{E[T_A]} = p^3q^3 \quad \text{as} \quad n \to \infty \quad \text{w.p.1}$$

by the LLN for delayed renewal processes; Proposition 7.1 of Ross.

(f) What is $E[N_A]$, the expected number of flips until pattern $A \equiv HTHTHT$ first occurs?

Like question (b), this is a tricky question. To understand this, it is useful to reconsider the mean of T_A. When we consider $E[T_A]$, the time between occurrences of $A \equiv HTHTHT$, we do not start with nothing, but we start already having had the partial pattern $HTHT$. Let $N_{C \to D}$ be the number of flips to get pattern D after observing pattern C. (Our notation $N_{C \to D}$ corresponds to $N_{D|C}$ in Ross; we use the arrow to emphasize which pattern comes first.)

We relate $E[N_A]$ to $E[T_C]$ for various patterns C.

$$\begin{aligned} E[N_A] &= E[N_{HT}] + E[N_{HT \to HTHT}] + E[N_{HTHT \to HTHTHT}] \\ &= E[T_{HT}] + E[T_{HTHT}] + E[T_{HTHTHT}] = \frac{1}{pq} + \frac{1}{p^2q^2} + \frac{1}{p^3q^3} \; . \end{aligned}$$

(g) What is the probability that pattern A occurs before pattern $B \equiv TTH$?

This is another tricky question; see page 127 of Ross for a detailed explanation. We set up two equations in two unknowns and solve them. One unknown is the probability $P_A \equiv P(N_A < N_B)$ that A occurs before B. The other unknown is $E[M_{A,B}]$, where $M_{A,B} \equiv \min\{N_A, N_B\}$ is the first time that one of the patterns A or B first occurs. These variables are expressed in terms of four computable means:

$$E[N_A], \quad E[N_B], \quad E[N_{A \to B}] \quad \text{and} \quad E[N_{B \to A}] \; .$$

We have seen how to derive $E[N_A]$ and $E[N_B]$. From part (f),

$$E[N_A] = E[T_{HT}] + E[T_{HTHT}] + E[T_{HTHTHT}] = \frac{1}{pq} + \frac{1}{p^2q^2} + \frac{1}{p^3q^3} \; .$$

On the other hand, the occurrence of B gives no head start toward having B occur again; i.e., we have

$$N_B \stackrel{\mathrm{d}}{=} T_B \quad \text{and} \quad E[N_B] = E[T_B] = \frac{1}{pq^2} \ .$$

So now we are ready to consider $E[N_{A\to B}]$ and $E[N_{B\to A}]$. Note that $N_{A\to B} \stackrel{\mathrm{d}}{=} N_{T\to TTH}$ and

$$E[N_{TTH}] = E[N_T] + E[N_{T\to TTH}] \ ,$$

so that

$$E[N_{T\to TTH}] = E[N_{TTH}] - E[N_T] = E[T_{TTH}] - E[T_T] = \frac{1}{pq^2} - \frac{1}{q} \ .$$

Next note that $N_{B\to A} \stackrel{\mathrm{d}}{=} N_{H\to HTHTHT}$ and

$$E[N_{HTHTHT}] = E[N_H] + E[N_{H\to HTHTHT}] \ ,$$

so that

$$\begin{aligned} E[N_{H\to HTHTHT}] &= E[N_{HTHTHT}] - E[N_H] = E[T_{HT}] + E[T_{HTHT}] + E[T_{HTHTHT}] - E[T_H] \\ &= \frac{1}{pq} + \frac{1}{p^2q^2} + \frac{1}{p^3q^3} - \frac{1}{p} \ . \end{aligned}$$

Now, following Ross, we have

$$\begin{aligned} E[N_A] &= E[M_{A,B}] + E[N_A - M_{A,B}] \\ &= E[M_{A,B}] + E[N_A - M_{A,B} | B \text{ before } A](1 - P_A) \\ &= E[M_{A,B}] + E[N_{B\to A}](1 - P_A) \ . \end{aligned}$$

Similarly,

$$E[N_B] = E[M_{A,B}] + E[N_{A\to B}] P_A \ .$$

Solving these two equations, we obtain

$$P_A = \frac{E[N_B] + E[N_{B\to A}] - E[N_A]}{E[N_{B\to A}] + E[N_{A\to B}]}$$

and

$$E[M_{A,B}] = E[N_B] - E[N_{A\to B}] P_A \ .$$

Summary of the notation defined above:

$$\begin{gathered} \text{pattern } A, \quad A(n), \quad P(A(n)), \quad N_A, \quad T_A, \quad N_A(n), \\ N_{A\to B}, \quad M_{A,B} \equiv \min\{N_A, N_B\}, \quad P_A, \end{gathered}$$