SEPTEMBER 16, 2008

The Probability of Runs of K Consecutive Heads in N Coin Tosses

Problem

What's the probability that at least one run of k consecutive heads occurs in n coin tosses?

Method One

Definitions:

$R\{k, n\}$
: The fact that at least one run of k consecutive heads occurs in n coin tosses.

$\sim E$
: The fact that event E does not occur.

$A\&B$
: The fact that event A and event B occur simultaneously.

$P(E)$
: The number of permutations that cause event E to occur.

$H(k, n)$
: The number of head-or-tail permutations for n coins that contain at

least one run of k consecutive heads; the same as $P(R\{k, n\})$.

Theorem: $P(A\&B) + P(A\&{\sim}B) = P(A)$

Analysis:

- If $n = k$, $R\{k, n\}$ occurs in exactly one case, so $H(k, n) = 1$.
- If $n < k$, $R\{k, n\}$ is impossible, so $H(k, n) = 0$.
- Otherwise($n > k$), $H(k, n)$ permutations can be divided into two groups: $R\{k, n-1\}$ and ~$R\{k, n-1\}$.

  1. $R\{k, n-1\}$
     $R\{k, n\}$ follows necessarily. There are $2H(k, n-1)$ permutations of this kind.

  2. ~$R\{k, n-1\}$
     $R\{k, n\}$ occurs only if the last $k-1$ of the first $n-1$ toss are all heads and the nth toss is head. Then the kth last toss of the first $n-1$ tosses must be tail, otherwise the last k of the first $n-1$ tosses are all heads, which contradicts ~$R\{k, n-1\}$. Hence, the last $k+1$ tosses are fixed as $[T, H, H, \cdots, H]$.

     Define S as the fact that the last $k+1$ tosses of n tosses are $[T, H, H, \cdots, H]$. The condition now becomes ~$R\{k, n-1\}\&S$, which is equivalent to ~$R\{k, n-k-1\}\&S$. Thus the permutation number is:

     $$\begin{aligned} &P({\sim}R\{k, n-k-1\}\&S) \\ &= P(S) - P(R\{k, n-k-1\}\&S) \\ &= 2^{n-k-1} - H(k, n-k-1). \end{aligned}$$

Thus,

- $H(k, n) = 2H(k, n-1) + 2^{n-k-1} - H(k, n-k-1)$, for $n > k$;
- $H(k, n) = 1$, for $n = k$;
- $H(k, n) = 0$, for $n < k$.

Method Two

A cool method using probability distribution vector and probability distribution transition matrix. The original post is in Chinese. I am trying to translate it into English below. The author even proved the property used in Method Three along the line; however, this part is beyond my knowledge.

The states during the process of coin tossing is defined as follows:

- $S_t (0 \le t < k)$: no runs of k consecutive heads have occurred, and t heads have accumulated in the last run.
- S_k: at least one run of k consecutive heads has occurred.

Lemmas:

- The initial state is S_0.
- If current state is $S_t (0 \le t < k)$, next state has equal opportunity, i.e., $\frac{1}{2}$, to be S_{t+1} or S_0.
- Once state becomes S_k, it will never change again.

Definition:

d_i

The probability distribution vector after the ith toss is a column vector of length $k + 1$, whose hth element is the probability that current state is S_{h-1}.

Initially, $d_0 = [1, 0, 0, \cdots, 0]^T$; $d_{i+1} = M \times d_i$, wherein M is the probability distribution transition matrix.

M can be reduced from lemmas above:

$$M = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \cdots & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & 0 & \cdots & 0 & 0 \\ 0 & \frac{1}{2} & 0 & \cdots & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & 0 & 0 \\ 0 & 0 & 0 & \cdots & \frac{1}{2} & 1 \end{bmatrix}$$

Thus, the last element of d_n is the probability desired, denoted as P:

$$P = [0, 0, \cdots, 0, 1] \times d_n = [0, 0, \cdots, 0, 1] \times M^n \times d_0.$$

Method Three

Use the fact "the probability that no runs of k consecutive tails will occur in n coin tosses is given by $F_{n+2}^{(k)} / 2^n$, where $F_l^{(k)}$ is a Fibonacci k-step number" from Wolfram MathWorld.

math Tweet