Experiment: Creating random sequences

…. sequence of 400 tosses of a fair coin.
Use H for heads and T for tails.

_ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _

_ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _

_ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _

_ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _

_ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _

_ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _

_ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _

_ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _

_ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _

_ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _

_ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _

_ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _

_ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _

Methodology: Sequence of coin tosses

Sequence #1

T HHHHT T T T HHHHT HHHHHHHHHT T THHT T HHHHHT T T T T HHT HHT HHH T
T T HT T HHHHT HT T T HT T T HHT T T T HHHHHHT T T HHT T HHHT HHHHHT T T T
T HT T T HHT T HT T HHT T T HHT T T HH THHT HHT T T T T HHT HHHHHHT HT HT T
HT HT T HHHT T HHT HT HHHHHHHHHT THT T HHHT HHT T HT T T T T T HHHT HHH

Sequence #2

T HT HT T T HT T T T T HT HT T T HT T HHHT HHT HT HT HT T T T HHT T HHT T HHH T
HHHT T HHHT T T HHHT HHHHT T T HT HT HHHHT HT T T HHHT HHT HT T T HHT H
HHT HHHHT T HT HHT HHHT T T HT HHHT HHT T T HHHT T T T HHHT HT HHHHT H
T T HHT T T T HT HT HT T HT HHT T HT T THT T T T HHHHT HT HHHT T HHHHHT HH

One of these sequences was generated by a sequence of coin tosses, the other one was generated by students being told to write down a sequence generated by coin tosses.

Which one is the sequence generated by coin tosses?

Methodology: Sequence of coin tosses

Sequence #1

T HHHHT T T T HHHHT HHHHHHHHHT T THHT T HHHHHT T T T T T HHT HHT HHH T
T T HT T HHHHT HT T T HT T T HHT T T T HHHHHHT T T HHT T HHHT HHHHHT T T T
T HT T T HHT T HT T HHT T T HHT T T HH THHT HHT T T T T HHT HHHHHHT HT HT T
HT HT T HHHT T HHT HT HHHHHHHHHT THT T HHHT HHT T HT T T T T T HHHT HHH

Sequence #2

T HT HT T T HT T T T T HT HT T T HT T HHHT HHT HT HT HT T T T HHT T HHT T HHH T
HHHT T HHHT T T HHHT HHHHT T T HT HT HHHHT HT T T HHHT HHT HT T T HHT H
HHT HHHHT T HT HHT HHHT T T HT HHHT HHT T T HHHT T T T HHHT HT HHHHT H
T T HHT T T T HT HT HT T HT HHT T HT T THT T T T HHHHT HT HHHT T HHHHHT HH

How can you tell? Which features can you look at?

- number of heads, number of tails
- number of alternations
- numbers and lengths of runs

Methodology: Sequence of coin tosses

Sequence #1

T HHHHT T T T HHHHT HHHHHHHHT T THHT T HHHHHT T T T T T HHT HHT HHH T
T T HT T HHHHT HT T T HT T T HHT T T T HHHHHHT T T HHT T HHHT HHHHHT T T T
T HT T T HHT T HT T HHT T T HHT T T HH THHT HHT T T T T HHT HHHHHHT HT HT T
HT HT T HHHHT T HHT HT HHHHHHHHT THT T HHHT HHT T HT T T T T T HHHT HHH

Sequence #2

T HT HT T T HT T T T T HT HT T T HT T HHHT HHT HT HT HT T T T HHT T HHT T HHH T
HHHT T HHHT T T HHHT HHHHT T T HT HT HHHHT HT T T HHHT HHT HT T T HHT H
HHT HHHHT T HT HHT HHHT T T HT HHHT HHT T T HHHT T T T HHHT HT HHHHT H
T T HHT T T T HT HT HT T HT HHT T HT T THT T T T HHHHT HT HHHT T HHHHHT HH

Runs of H of lengths r

H	3	4	5	6	7	8
#1	6	3	2	2	0	2
#2	11	5	1	0	0	0

Runs of T of lengths r

T	3	4	5	6	7	8
#1	6	1	2	2	0	0
#2	9	4	2	0	0	0

Methodology:
Expected number of runs in a sequence of coin tosses

$X_i \in \{0,1\}$ $(i = 1, 2, \ldots, N)$

independent identically distributed

$P(X_i = 0) = P(X_i = 1) = \frac{1}{2}$

$Z_r = \text{number of run of length } r$

Calculation (see blackboard) shows:

$$\boxed{E[Z_r] = \left(\frac{1}{2}\right)^{r-1} + (N - r - 1)\left(\frac{1}{2}\right)^{r+1}}$$

Explicit calculation for N=200:

```
R <- vector(length=10)
N=200
for (r in 1:10){
  R[r]=(N-r-1)*2^{-r-1}+2^{-r-1}
  }
round(R)
```

```
> round(R)
 [1] 50 25 12  6  3  2  1  0  0  0
```

Comparison

Theoretical formula for runs of length 3, 4, …, 8 in a sequence of 200

```
> round(R)[3:8]
[1] 12  6  3  2  1  0
```

N=200

Earlier observations (add up H and T runs tables):

H or T	3	4	5	6	7	8
#1	12	4	4	4	0	2
#2	20	9	3	0	0	0

Increase N to 400 and compare

Theoretical formula for runs of length 3, 4, … 8 in a sequence of 400

```
> round(R)[3:8]
[1] 25 12  6  3  2  1
```

N=400

Your sequence:

…. sequence of 400 tosses of a fair coin.
Use H for heads and T for tails.

In your sequence:
How many runs of
r= 5, 6, 7, 8?

_ _

_ _

_ _