PURDUE UNIVERSITY MATH DEPARTMENT
PROBLEM OF THE WEEK
FALL 2014, PROBLEM 2

WILLIAM WU

Problem A standard six-sided die is rolled repeatedly and a running total is kept of all the numbers rolled. Which of 2, 6, 1006 is more likely to be one of these totals? Prove your answer.

Solution 1 Let $p(i,j)$ be the probability of generating a sum of i in j rolls. Then

$$p(i,1) = \begin{cases} q & \text{for } i \in \{1,2,3,4,5,6\} \\ 0 & \text{else} \end{cases}$$
$$p(i,j) = 0 \quad \text{for } i < j \text{ or } i < 1,$$

where $q := 1/6$. By conditioning on the first roll, we have the recurrence

$$p(i,j) = q \sum_{m=1}^{6} p(i-m, j-1).$$

Lastly, summing over all possible numbers of rolls, the probability of seeing i occur is the row sum $\sum_j p(i,j)$. Below is a short program to fill the table $p(i,j)$ using dynamic programming:

```
import numpy as np
n = 1006
p = np.zeros((n,n))
q = float(1)/6
for i in xrange(0,6): p[i,0] = q
for i in xrange(1,n):
    for j in xrange(1,n):
        for m in xrange(1,7):
            if i-m >= 0: p[i,j] += p[i-m,j-1]
        p[i,j] *= q
for t in [2,6,1006]:
    print "Probability of {0}: {1}".format(t, np.sum(p[t-1,:]))
```

```
Probability of 2: 0.194444444444
Probability of 6: 0.36023233882
Probability of 1006: 0.285714285714
```

Therefore, of 2, 6, and 1006, the most likely number to be seen amongst the totals is 6.

□

Date: September 23, 2014.

Solution 2 Let $p(i)$ be the probability of seeing a sum of i. Then, by conditioning on the first roll, we have the recurrence

$$p(i) = \frac{1}{6} \sum_{m=1}^{6} p(i-m)$$

with characteristic equation $6x^6 - x^5 - x^4 - x^3 - x^2 - x - 1 = 0$. The roots of this equation are

$$\begin{aligned}
r_1 &= -0.7666004408 \\
r_2 &= 0.9459814051 \\
r_3 &= -0.3206507125 - 0.6547779619 \,\mathbf{i} \\
r_4 &= -0.3206507125 + 0.6547779619 \,\mathbf{i} \\
r_5 &= 0.3142935637 - 0.5775691439 \,\mathbf{i} \\
r_6 &= 0.3142935637 + 0.5775691439 \,\mathbf{i}
\end{aligned}$$

and the general solution of the recurrence is $p(i) = \sum_{k=1}^{6} c_k r_k^i$. Starting with $p(1) = 1/6$ and evaluating the recurrence, we can easily compute the initial conditions $p(1) = 1/6$, $p(2) = 7/36$, $p(3) = 49/216$, $p(4) = 343/1296$, $p(5) = 2401/7776$ and $p(6) = 16807/46656$. Using these initial conditions we can then solve for the coefficients c_k, yielding a closed-form expression for $p(i)$, for every i. Direct computation then yields

$$\begin{aligned}
p(2) &= 0.194444444444 \\
p(6) &= 0.36023233882 \\
p(1006) &= 0.285714285714.
\end{aligned}$$

Therefore, of 2, 6, and 1006, the most likely number to be seen amongst the totals is 6.

□

E-mail address: w@qed.ai

Quantitative Engineering Design Inc.