WINNING AT CRAPS WITH A RANDOM DIE

by

Anirban DasGupta*
Purdue University

Technical Report # 95-24

Department of Statistics
Purdue University

May 1995

*Research supported by NSF Grant DMS-9307727

*

WINNING AT CRAPS WITH A RANDOM DIE

by

Anirban DasGupta*

Purdue University

Abstract

In the popular dice game of Craps, the house has a slight edge: the player's winning probability is .492929. How does the game turn out for the house or the player if the dice are not exactly balanced? We find that if the dice are constructed at random by choosing the probability vector according to a uniform distribution on the natural simplex, then the house loses its edge. Indeed, .492929 is only about the 31st percentile of the distribution of winning probabilities if a random die is used to play Craps. There is also some effect on how long it takes to finish a game. We find that both the winning probability and the expected duration can be well approximated by normal distributions. These and many other probabilistic questions related to Craps are studied in this note.

1. Introduction

Craps is a popular and common game with dice. A pair of dice is rolled. If the scores on the two dice add to 7 or 11, the player wins; if they add to 2, 3, or 12, the house wins. If they add to some other number, that number is called the "point" of the player, and the player continues rolling the two dice until either the scores add to 7 or the "point". If it is the point, the player wins and if it is 7, the house does. It is well known that craps is slightly favorable to the house with a balanced die; the probability of the player winning is .492929 (actually, the exact value is 244/495). See Isaac (1995). Thus, in the event of the dice being balanced, the house makes money in the long run, and keeps going. We will assume that the same die is rolled twice, which can be assumed if the two dice are identical. In practice, the two die may be (slightly) different.

*Research supported by NSF grant DMS 9307727

We intend to take a leisurely look at craps if the die is not necessarily balanced. Obvious questions that come to mind are what happens to the player's winning probability if the die is slightly unbalanced. Which faces have a greater bearing on the outcome? Which die makes the game exactly fair? If a die is constructed in a random way with six faces but with no particular attention to shape or balance, how does craps turn out for the house? What is the distribution and the mean of the winning probability with a random die? Is the distribution about a normal? Does the game become more favorable to the player the more unbalanced the die is? In other words, how careful does the house have to be in its choice of dice? How long does it take to finish a game? How is the duration of the game related to whether the die is balanced or not? In the next few sections, we take a look at all of these questions. Perhaps the gambling connoisseur or a casino may find some of our findings important.

2. Probability of Winning.

Consider a general six sided die with $P(\text{ith face}) = p_i, 1 \leq i \leq 6$. The player can win in one of the following ways:

a. The Player immediately scores 7;

b. The Player immediately scores 11;

c. The Player scores one of $i = 4, 5, 6, 8, 9, 10$ in the first try, then scores something different from i and 7 for $n-2$ consecutive trials for some $n \geq 2$ and ends with i on the nth roll.

Then easy algebra gives the following formula:

Proposition 1. The winning probability for the player equals

$$
\begin{aligned}
&2p_3p_4 + 2p_2p_5 + 2p_1p_6 + 2p_5p_6 + (p_2^2 + 2p_1p_3)^2/ \\
&\quad (p_2^2 + 2p_1p_3 + 2p_3p_4 + 2p_2p_5 + 2p_1p_6) + (2p_2p_3 + 2p_1p_4)^2/ \\
&\quad (2p_2p_3 + 2p_1p_4 + 2p_3p_4 + 2p_2p_5 + 2p_1p_6) + (p_3^2 + 2p_2p_4 + 2p_1p_5)^2/ \\
&\quad (p_3^2 + 2p_2p_4 + 2p_3p_4 + 2p_1p_5 + 2p_2p_5 + 2p_1p_6) + (p_4^2 + 2p_3p_5 + 2p_2p_6)^2/ \\
&\quad (2p_3p_4 + p_4^2 + 2p_2p_5 + 2p_3p_5 + 2p_1p_6 + 2p_2p_6) + (2p_4p_5 + 2p_3p_6)^2/ \\
&\quad (2p_3p_4 + 2p_2p_5 + 2p_4p_5 + 2p_1p_6 + 2p_3p_6) + (p_5^2 + 2p_4p_6)^2/ \\
&\quad (2p_3p_4 + 2p_2p_5 + p_5^2 + 2p_1p_6 + 2p_4p_6)
\end{aligned}
\tag{2.1}
$$

Corollary 1. For playing craps with a balanced die, the winning probability of the player equals $244/495 = .492929$.

Corollary 2. The minimum winning probability in craps is zero. In particular, a die with 1 on all faces gives a zero winning probability for the player.

Corollary 3. Among all dice with equal probability for the middle scores 3 and 4, and equal probability for the other four scores, the maximum and minimum winning probability are $2/3$ and $.458224$ respectively. In particular, the winning probability is $2/3$ for a die with 3 on three faces and 4 on the other three faces.

Corollary 4. Craps is fair to both the player and the house for a die with probability .154035 for the outer scores 1,2,5,6 and probability .19193 for the middle scores 3 and 4. In addition, this is the only die with equal probability for the outer scores and equal probability for the middle scores for which craps is fair.

Corollary 5. The partial derivatives of the winning probability at $p_1 = p_2 = \ldots = p_6 = 1/6$ are all positive. They are respectively equal to 17629/27225, 24004/27225, 29812/27225, 29812/27225, 33079/27225, and 26704/27225. In particular, deviation from the balanced case in the form of a slightly larger probability for 5 is the most damaging to the house.

Discussion. It follows from the formula for the winning probability in Proposition 1 that there are many other dice for which craps is fair to both parties. In Figure 1, the winning probability is plotted for the special dice with probability p for each of 1, 2, 3, 4 and probability $(1-4p)/2$ for 5, 6. Notice that after $p = .22$ (approximately), the winning probability comes back up again, indicating the complex nature of the game.

There is some natural curiosity and interest in knowing how the game turns out if one makes a six sided die at random. Mathematically, this corresponds to the behavior of the winning probability if the probability vector $\underline{p}$ for the sides $1, 2, \ldots, 6$ is chosen according to a Uniform distribution on the natural simplex

$$S_6 = \{\underline{p} : 0 \leq p_i \leq 1, \sum_{i=1}^{6} p_i = 1\}.$$

Since we have an explicit formula for the winning probability under any given $\underline{p}$, the mean winning probability for a random die is easily found. It turns out craps is in fact unfavorable to the house with a random die: the mean winning probability for the player is (approximately) .539. Although the mean winning probability is easily found, the distribution of the winning probability for $\underline{p}$ chosen at random from the simplex is a tremendously difficult object. Aside from the multivariate nature of the problem, a main factor contributing to the difficulty is the fact that the winning probability is unfortunately not a one-one function of the vector $\underline{p}$. We can, however, get sufficiently accurate information regarding the distribution of the winning probability by simulating from the Uniform distribution on the simplex, which is done very easily by simulating from Exponential distributions. The following information is obtained from a simulation of size 32500 from the simplex:

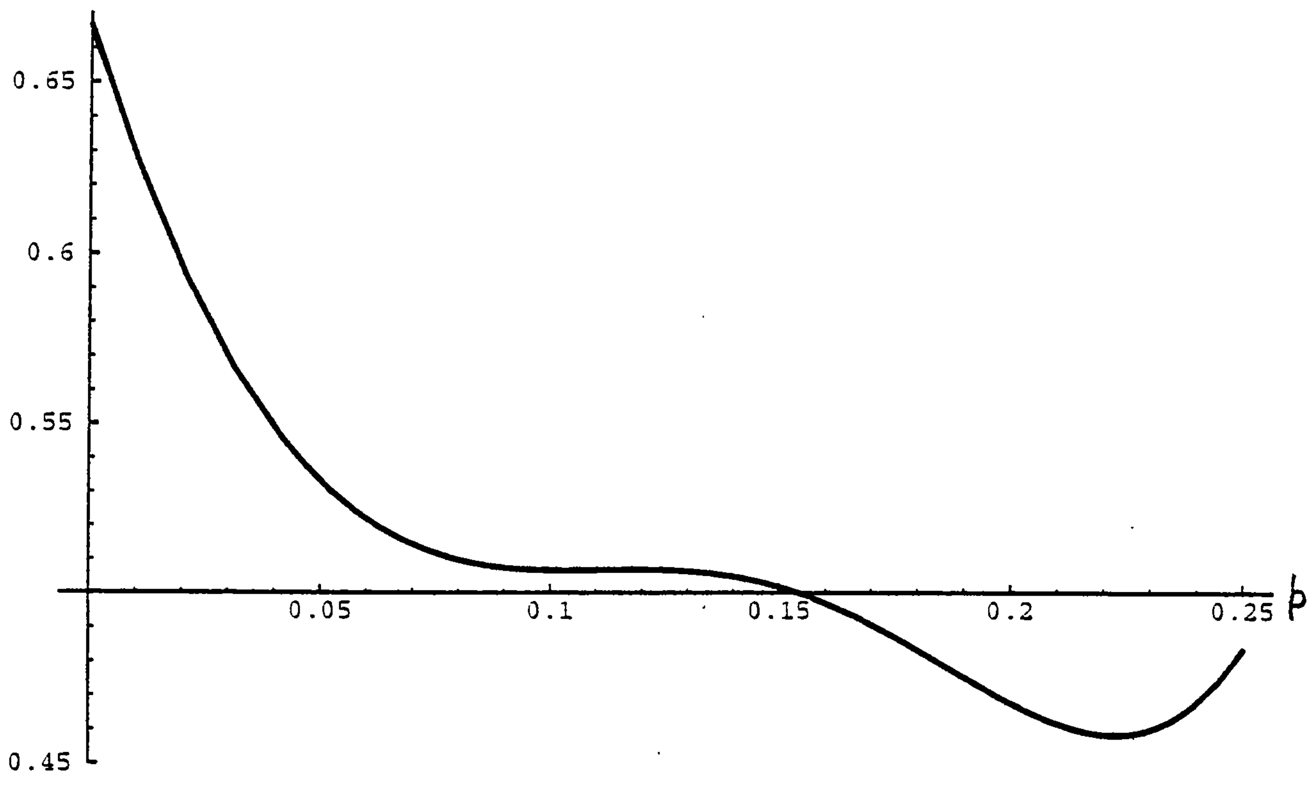
a. The median winning probability is .534.

b. The first and the third quartile are .477 and .597.

c. .492929, which is the winning probability for an exactly balanced die, is the 30.75th percentile of the distribution of winning probabilities. Thus, the house has plenty of reasons to be happy. In comparison to how bad it can get for them if a die was made by a novice on the spot, a balanced die places them in a relatively good position: the winning probability for a balanced die is just about only the 31st percentile of the distribution of winning probabilities.

Figure 1

Player's winning probability for die with probabilities $(p,p,(1-4p)/2,(1-4p)/2,p,p)$



The close proximity of the mean and the median hint at the possibility of a normal like approximation. Figure 2 is a normal $Q-Q$ plot of the simulated winning probabilities and Figure 3 is a histogram of the same simulated winning probabilities. Although some right skew is seen in the histogram, the following seems to be about right:

The player's winning probability in craps with a randomly constructed die is approximately normal with mean .539 and standard deviation .104.

Further discussion: Since in practice, the die being used is not going to be exactly balanced (for reasons in and out of the house's control), it is interesting to ask if deviation from the balanced case is good or bad for the house. In some sense, we already saw some evidence that relatively speaking, a balanced die is a good thing for the house. A reasonable way to further understand this is to measure deviation from the balanced case by considering the supnorm $\| \underset{\sim}{p} - 1/6 \ \underset{\sim}{1} \|_\infty = \max [\, | p_i - 1/6 | \,]$, and then investigating what happens to the player's winning probability if $\| \underset{\sim}{p} - 1/6 \ \underset{\sim}{1} \|_\infty$ increases. Of course, there is no functional relationship exactly. Figure 4 shows that there is no clear trend in either direction. A large deviation from the balanced case is neither necessarily bad or necessarily good for the house. But note the very interesting hole in the right of the plot. For dice that are badly unbalanced, the situation is black or white: it is either extremely bad or extremely good for the house. The next picture, Figure 5, shows an expected monotonicity. The further one is from the balanced case, the further is the winning probability from .492929, generally, although again there is no functional relationship obviously.

3. Duration of the game.

For the house, the duration of the game is clearly important. Even for the player, probably there is some psychological importance. Even for a fixed die, say the balanced die, the duration is a random variable: we define duration to be the number of rolls necessary to find a winner. If the die is not fixed, there is a marginal distribution for the duration implied by the Uniform distribution on the simplex. Alternatively, one can look at the expected duration as a function of the probability vector $\underset{\sim}{p}$, which itself has a distribution implied by the Uniform distribution on the simplex. We will study each of these a bit.

Figure 3

Histogram of simulated winning probabilities for a random die

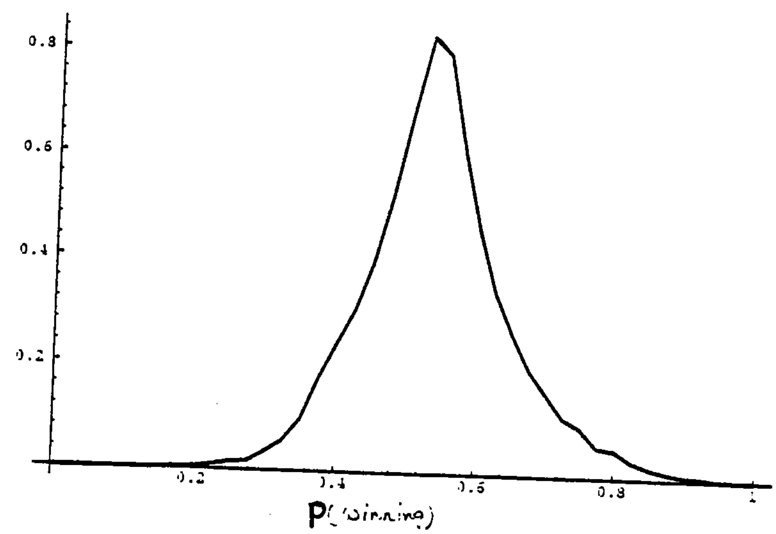


Figure 2

Normal Q-Q plot of simulated winning probabilities for a random die

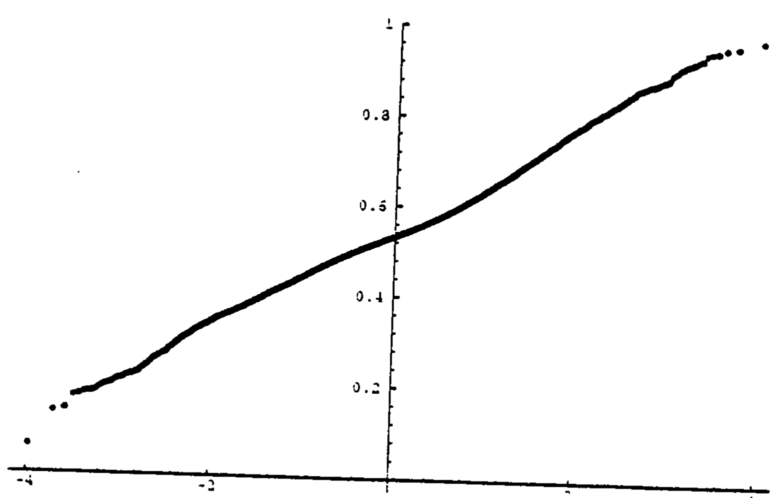


Figure 4

Player's winning probability vs. $\|\underset{\sim}{p} - 1/6\,\underset{\sim}{\mathbf{1}}\|_{\infty}$

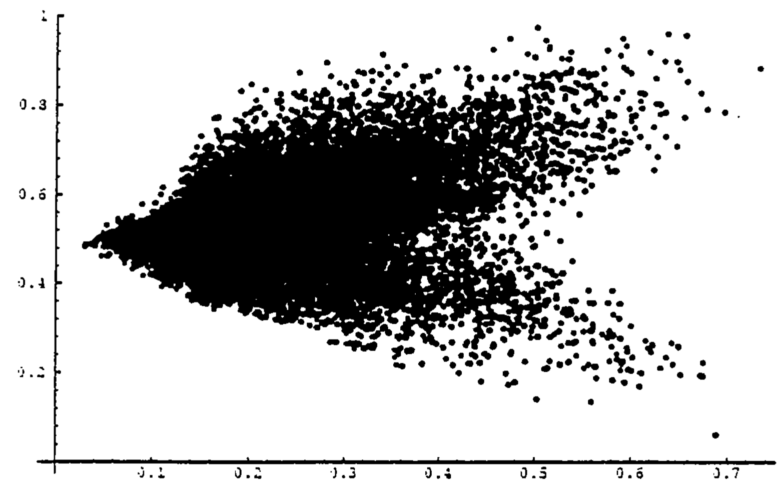
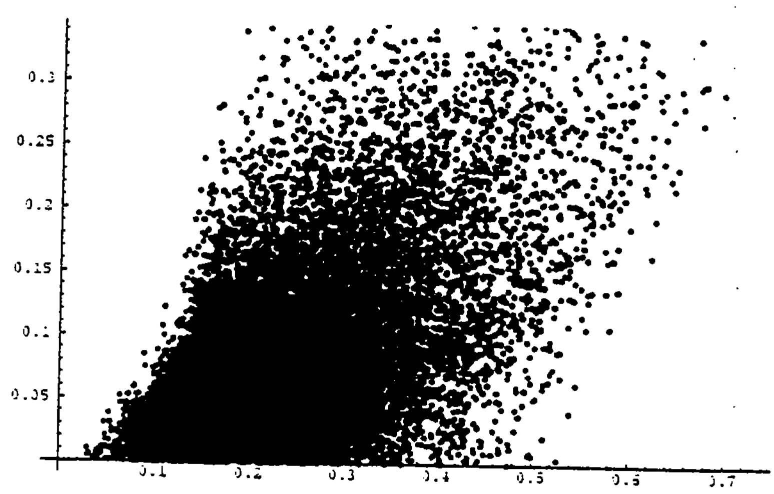


Figure 5

|Player's winning probability - winning probability for balance die|

vs. $\|p - 1/6\,\underset{\sim}{\mathbf{1}}\|_{\infty}$



3.1. Duration for a balanced die. With a balanced die, the duration is 1 (roll) with probability 1/3 since the game stops immediately if the player scores $2, 3, 7, 11$ or 12. In the next subsection, we will give the general formula for the probability mass function of the duration under a general probability vector $\underline{p}$. The following is obtained as a special case of that:

Duration:	1	2	3	4	5
Probability	1/3	.188272	.134774	.096567	.069257
	6	7	8	9	10
Probability	.049717	.035725	.025695	.018499	.013331
	11	12	13	14	15
Probability	.009617	.006944	.005019	.003631	.002629
	16	17	18	19	20
Probability	.001906	.001383	.001004	.000730	.000531

Thus, with a 99% probability, the game ends in 14 or less rolls. The expected duration is 3.37576. Thus, there is a very long right tail for the distribution of the duration with a balanced die. Figure 6 gives a plot of the probability mass function of the duration in this case.

3.2. Duration for a general die. The following formula is obtained on algebra by using the definition of the game:

Proposition 2. Consider a general six sided die with probabilities $p_1, p_2, \ldots, p_6$ for the scores $1, 2, \ldots, 6$. Then using the notation $p_1 = p$, $p_2 = q$, $p_3 = r$, $p_4 = s$, $p_5 = t$

Figure 6

Distribution of duration of a game for a balanced die

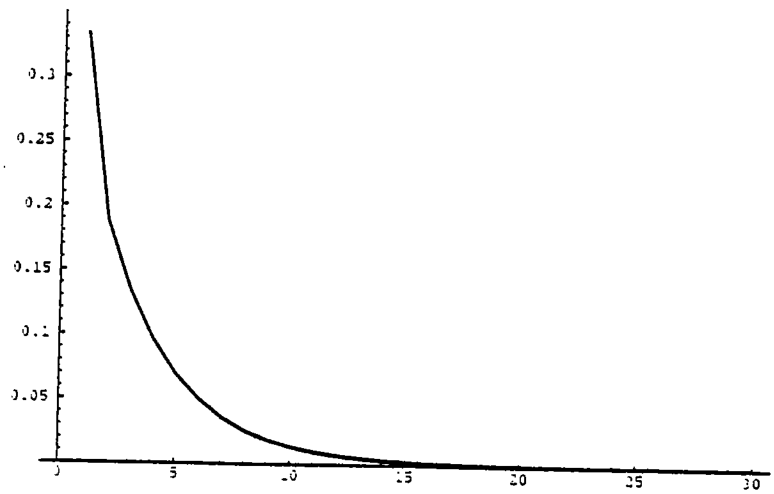
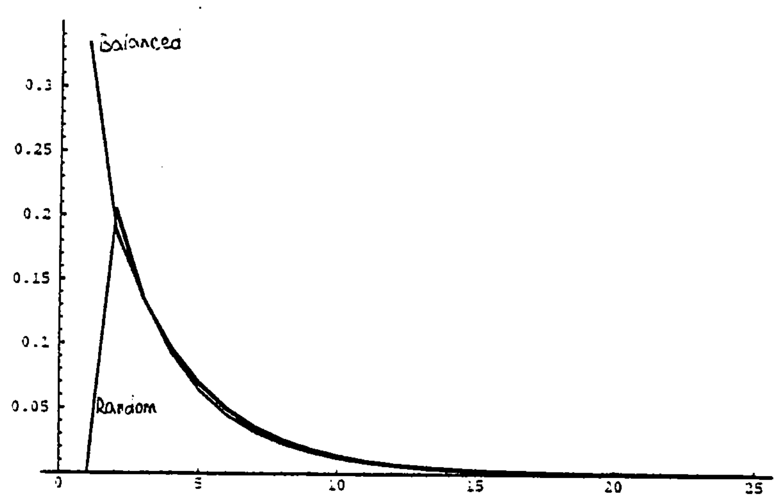


Figure 9

Distribution of duration of a craps game : Balanced and random die



and $p_6 = u$, one has

$$P(\text{Duration} = 1) = p^2 + u^2 + 2pq + 2qt + 2rs + 2pu + 2ut$$

For $n \geq 2$,

$$\begin{aligned}
P(\text{Duration} = n) &= (q^2 + 2pr)(1 - q^2 - 2pr - 2rs - 2qt - 2pu)^{(-2+n)} \\
&\quad (q^2 + 2pr + 2rs + 2qt + 2pu) + \\
&\quad (2qr + 2ps)(1 - 2qr - 2ps - 2rs - 2qt - 2pu)^{(-2+n)} \\
&\quad (2qr + 2ps + 2rs + 2qt + 2pu) + \\
&\quad (r^2 + 2qs + 2pt)(1 - r^2 - 2qs - 2pt - 2qt - 2pu)^{(-2+n)} \\
&\quad (r^2 + 2qs + 2rs + 2pt + 2qt + 2pu) + \\
&\quad (1 - 2rs - s^2 - 2qt - 2rt - 2pu - 2qu)^{(-2+n)} \\
&\quad (s^2 + 2rt + 2qu)(2rs + s^2 + 2qt + 2rt + 2pu + 2qu) + \\
&\quad (1 - 2rs - 2qt - 2st - 2pu - 2ru)^{(-2+n)}(2st + 2ru) \\
&\quad (2rs + 2qt + 2st + 2pu + 2ru) + \\
&\quad (1 - 2rs - 2qt - t^2 - 2pu - 2su)^{(-2+n)}(t^2 + 2su) \\
&\quad (2rs + 2qt + t^2 + 2pu + 2su)
\end{aligned} \tag{3.1}$$

In particular, consider the die with probability vector (.154035, .154035, .19193, .19193, .154035, .154035) that makes the game fair for both parties. For this die, the duration has the following probabilities for $1, 2, \ldots, 18$: .310942, .198999, .141142, .100212, .071228, .050682, .036103, 025747, .018383, .013140, .009403, .006737, .004833, .003471, .002495, .000934, .000675, .000488. The expected duration for this die is 3.37227, a trifle less than that for the balanced die.

3.3. Expected duration for a random die. As with the winning probability for the player, algebra also gives an exact formula for the expected duration under any given probability vector p. This formula can then be used to get information about the distribution of the expected duration if a die is constructed at random by choosing a vector p uniformly from the simplex.

Proposition 3. The expected duration of a craps game under a general probability vector $\underset{\sim}{p}$ equals

$$
\begin{aligned}
&2p_3p_4 + 2p_2p_5 + 2p_1p_6 + 2p_5p_6 + \\
&\quad ((p_2^2 + 2p_1p_3)^2(-1 + (p_2^2 + 2p_1p_3 + 2p_3p_4 + 2p_2p_5 + 2p_1p_6)^{-2})) \\
&\quad /(1 - p_2^2 - 2p_1p_3 - 2p_3p_4 - 2p_2p_5 - 2p_1p_6) \\
&\quad + ((2p_2p_3 + 2p_1p_4)^2 \\
&\quad (-1 + (2p_2p_3 + 2p_1p_4 + 2p_3p_4 + 2p_2p_5 + 2p_1p_6)^{-2}))/ \\
&\quad (1 - 2p_2p_3 - 2p_1p_4 - 2p_3p_4 - 2p_2p_5 - 2p_1p_6) \\
&\quad + ((p_3^2 + 2p_2p_4 + 2p_1p_5)^2 \\
&\quad (-1 + (p_3^2 + 2p_2p_4 + 2p_3p_4 + 2p_1p_5 + 2p_2p_5 + 2p_1p_6)^{-2}))/ \\
&\quad (1 - p_3^2 - 2p_2p_4 - 2p_3p_4 - 2p_1p_5 - 2p_2p_5 - 2p_1p_6) \\
&\quad + ((p_4^2 + 2p_3p_5 + 2p_2p_6)^2 \\
&\quad (-1 + (2p_3p_4 + p_4^2 + 2p_2p_5 + 2p_3p_5 + 2p_1p_6 + 2p_2p_6)^{-2}))/ \\
&\quad (1 - 2p_3p_4 - p_4^2 - 2p_2p_5 - 2p_3p_5 - 2p_1p_6 - 2p_2p_6) \\
&\quad + ((2p_4p_5 + 2p_3p_6)^2 \\
&\quad (-1 + (2p_3p_4 + 2p_2p_5 + 2p_4p_5 + 2p_1p_6 + 2p_3p_6)^{-2}))/ \\
&\quad (1 - 2p_3p_4 - 2p_2p_5 - 2p_4p_5 - 2p_1p_6 - 2p_3p_6) \\
&\quad + ((p_5^2 + 2p_4p_6)^2 \\
&\quad (-1 + (2p_3p_4 + 2p_2p_5 + p_5^2 + 2p_1p_6 + 2p_4p_6)^{-2}))/ \\
&\quad (1 - 2p_3p_4 - 2p_2p_5 - p_5^2 - 2p_1p_6 - 2p_4p_6)
\end{aligned}
\tag{3.2}
$$

Again, for exactly the same reasons as for the winning probability, inspite of an exact formula, the distribution of the expected duration is extremely difficult to obtain exactly. Notice also that the expected duration has a Lebesgue density on $[1, \infty)$; it is not discrete unlike the duration itself. The marginal mean (i.e., the expectation of the expected duration) can be obtained by integrating the formula (3.2) on the simplex. It equals 3.4458 (approximately). The following information is obtained from 15000 simulations from the

Uniform distribution on the simplex :

$$\text{Median } = 3.43; \text{ standard deviation } = .625.$$

Figure 7 and Figure 8 give the normal $Q-Q$ plot and a histogram of the simulated values. From these, it seems the following can be stated as a rough guide:

The distribution of the expected duration of a craps game for a random die is approximately normal with mean 3.446 and standard deviation .625.

3.4. Marginal distribution of duration. Since Proposition 2 gives the conditional distribution of the duration under any fixed vector $\underset{\sim}{p}$, the marginal distribution is in principle obtained by integrating that formula in Proposition 2 on the simplex. This is not possible in a closed form except for the marginal probability of the value 1; the marginal probability of 1 is $1/2880 = .000347$. However, simulation from the simplex can be used to get accurate approximations to the true values of the marginal probabilities. Figure 9 gives a plot of the marginal distribution, superimposed on the distribution for a balanced die. The approximate marginal probabilities are given below for $1, 2, \ldots, 25$: .000347, .204067, .135583, .092213, .063865, .044909, .032004, .023086, .016843, .012420, .009253, .006961, .005287, .004053, .003134, .002445, .001923, .001525, .001219, .000982, .000796, .000651, .000535, .000442, .000369.

The most striking difference from the balanced case is that 2 rolls is the most likely value for a random die; one can see from Figure 9 that barring 1 and 2, the probabilities of every other case are very close for a balanced and a random die.

4. Conclusion.

Thus, if one considers both the player's winning probability and the duration of a game, then a random die provides some interesting variations from a balanced one.

References

Hamming, Richard W. (1991). The Art of Probability, Addison-Wesley, New York.

Isaac, Richard (1995). The Pleasures of Probability, Springer-Verlag, New York.

Scarne, John (1978). Scarne's Guide to Casino Gambling, Simon and Schuster.

Figure 8

Histogram of simulated expected durations for random dice

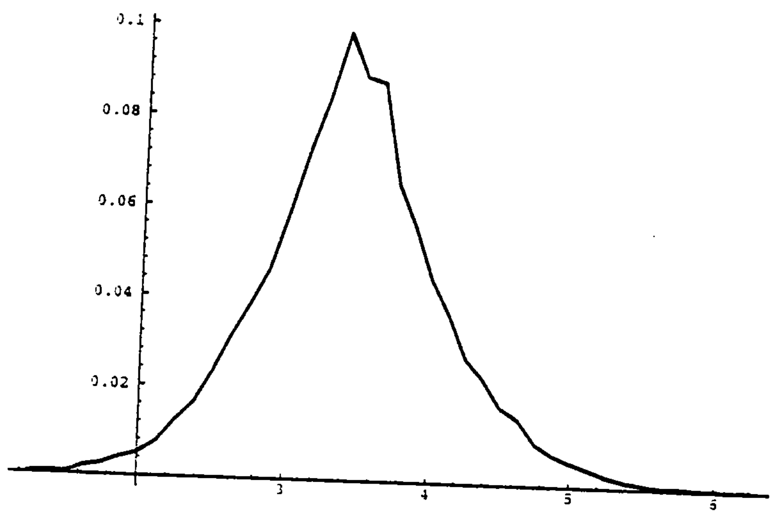


Figure 7

Normal Q-Q plot for simulated expected durations for random dice