

### Problems from Chapter 3 Handout:

13. A die is rolled three times. What is the probability that you get a larger number each time?

Since only one permutation will be strictly increasing of the 6 possible from 3 numbers, this means that:

$$P = \frac{1}{3!} \cdot (6)_3 = \frac{\binom{6}{3}}{6^3} = \frac{20}{216} = \frac{5}{54}$$

16. Four shoes are taken at random from five different pairs. What is the probability that there is at least one pair among them?

The probability that there is at least one pair among the four shoes will be complementary to the probability that there are *no* pairs among them.

The probability that there are no pairs is given by the ratio of the number of instances with no pairs to the total number of possible choices of 4 shoes.

The number of instances with no pairs follows by the fundamental rule as the product of the number of possibilities for the first shoe (10), the next shoe (8 – since it excludes the rest of the pair), and so forth – divided by the number of duplicate instances (since order doesn't matter).

Thus, by this reasoning, it follows that the probability that there is at least one pair is given by:

$$P = 1 - \frac{(10 \cdot 8 \cdot 6 \cdot 4) / 4!}{\binom{10}{4}} = 1 - \frac{80}{210} = \frac{13}{21}$$

21. Ten books are made into 2 piles. In how many ways can this be done if books as well as piles may or may not be distinguishable? Treat all four hypotheses and require that neither pile be empty.

I. Piles but not books are distinguishable

This can be thought of as a 'stars and bars' problem with 10 stars and 1 bar – in which case the two instances in which there are an empty pile must be excluded:

The number of arrangements is given by:  $\binom{11}{1} - 2 = 9$

II. Both books and piles are not distinguishable

This case follows from I, wherein if the case where the two piles are even is duplicated, then each situation has been counted exactly twice. Then, the empty case is excluded:

The number of ways is given by:  $\frac{\binom{11}{1} + 1}{2} - 1 = 5$

III. Both books and piles are distinguishable

Given that the order of the books is immaterial within each pile, then this case follows directly from the fundamental rule, whereby each book may be placed into either pile – and the two cases in which piles are empty must be excluded:

The number of ways is given by:  $2^{10} - 2 = 1022$

IV. Books, but not piles, are distinguishable

The number of ways for this case follows directly from III, whereby each instance is counted exactly twice:

The number of ways is given by:  $\frac{2^{10} - 2}{2} = 511$

23. In a draft lottery containing the 366 days of the year, the probability that the first 180 days drawn (without replacement, of course) are evenly distributed among the twelve months is given by the ratio of the number of possible ways in which the days could be evenly distributed among twelve months to the total number of ways that these 180 days can be chosen. Thus, by the Fundamental Rule, the total number of ways these days can be chosen evenly is a free combination of 15 days chosen from each month:

$$P = \frac{\binom{31}{15} \binom{29}{15} \binom{31}{15} \binom{30}{15} \binom{31}{15} \binom{30}{15} \binom{31}{15} \binom{30}{15} \binom{31}{15} \binom{30}{15} \binom{31}{15} \binom{31}{15}}{\binom{366}{180}} = \frac{\binom{31}{15}^7 \binom{30}{15}^4 \binom{29}{15}}{\binom{366}{180}}$$

The probability that the first 30 days chosen contain none from August or September is given by:

$$P = \frac{\binom{366 - 31 - 30}{30}}{\binom{366}{30}} = \frac{\binom{305}{30}}{\binom{366}{30}}$$

25. One hundred trout are caught in a little lake and returned after they are tagged. Later another 100 are caught and found to contain 7 tagged ones.

a) The probability of this if the lake contains  $n$  trout is given as a ratio between the number of possible ways in which exactly 7 tagged trout could have been sampled from the lake along with 93 untagged trout to the total number of ways that 100 trout can be sampled from the population. This numerator, thus, is given as a free combination of the number of ways of sampling 7 trout from the 100 in the tagged population (given that none have been caught or died, etc.) and the number of ways of sampling 93 trout from the rest of the population:

$$P(n) = \frac{\binom{100}{7} \binom{n-100}{93}}{\binom{n}{100}}$$

b) It seems reasonable that in order to deduce a value for  $n$  based on the number of tagged trout sampled from the lake that the greatest likelihood would occur if the proportion of tagged trout in the sample population is close to the proportion of the tagged trout within the greater lake population, that is:

$$\frac{7}{100} \sim \frac{100}{n} \text{ and } n \approx \frac{100 \cdot 100}{7} \approx 1428.571429 \rightarrow 1428 \text{ or } 1429$$

\*As shown by Professor Doyle in class on 4/2/10, by modeling the original probability function in (a) in MATLAB across integer values of  $n$  from 1420 to 1440 – we find that 1428 is slightly more likely than 1429.

27c. In a poker hand assume all “hands” are equally likely as under Sampling Case III.

The probability of a straight flush is given as the ratio of the number of instances of straight flushes to the total number of hands drawn. According to poker rules, an Ace can play either high or low in a straight flush, meaning that if numeric values are given to the top card in the hand, the number of straight flushes within a suit is given as:  $(14 - 5) + 1 = 10$  and there are 4 suits.

Assuming a “fresh” deck of 52 cards then the probability of the hand is given as:

$$P = \frac{10 \cdot 4}{\binom{52}{5}} = \frac{40}{2598960}$$

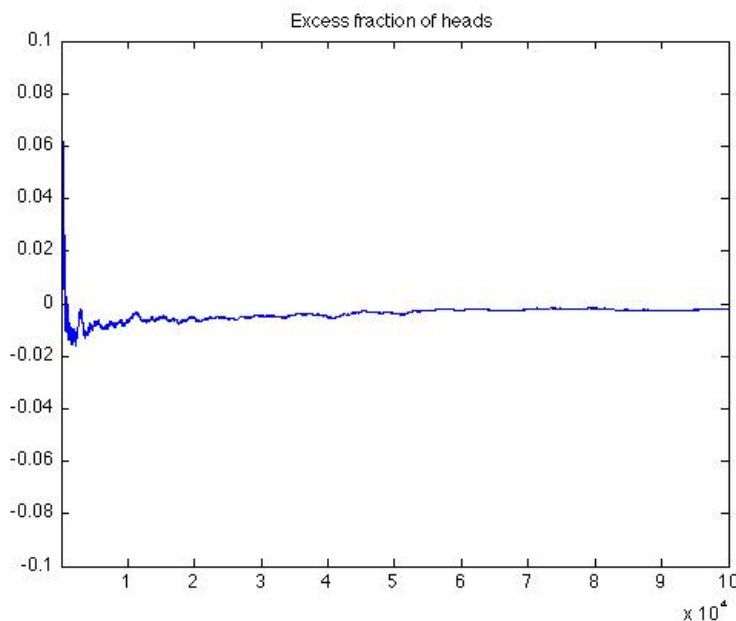
(this includes the 4 “royal” flushes which are ace-high straight flushes).

\*This result was discussed in class on 4/2/10.

### Selected Exercises from G&S Section 1.1:

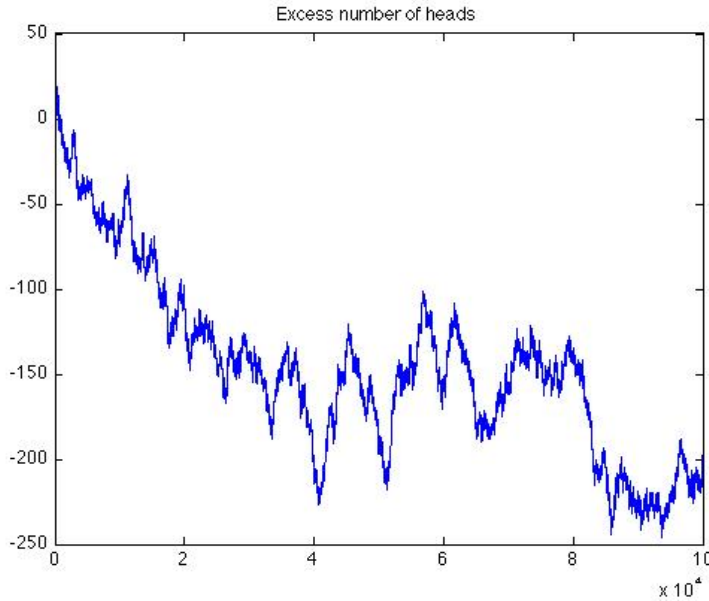
\*Problems 1-5 refer to programs authored by Professor Doyle as posted in <http://www.math.dartmouth.edu/~doyle/docs/60/prog/prog.html>  
Some parameters have been readjusted within the programs.

1. Running a program that prints the “*excess fraction*” of heads, the proportion of heads minus the probability 1/2 might return a plot as such:



The program used to create this plot and the next is unaltered from that provided, which prints after each “toss” rather than after every 100 tosses (thus is the advantage of computing power!). As the number of tosses increases, the excess fraction of heads does appear to approach 0. This is to say, that the average proportion seems to converge to the expected value (1/2).

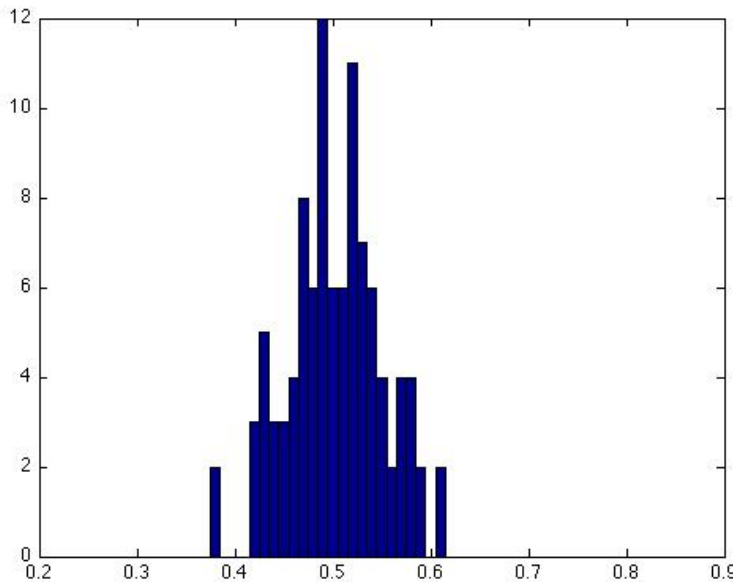
For the same trial, however, the “*excess number*” of heads – the number of heads minus half the number of tosses, a plot looks like this:



Evidently, the excess number of heads does not appear to approach 0 as  $n$  increases. In fact – anecdotally, running this experiment a number of times results in a number of graphs similar in appearance for the convergence of the proportion – but the excess number graphs can vary in general character quite wildly.

2. If the parameter of repetitions on the experiment is set to 100 within the program provided (as Grinstead and Snell state), trials with 95 out of the 100 experiments having a proportion of heads between 0.4 and 0.6 start appearing with frequency (anecdotally) when  $n \approx 100$ .

A histogram of these results might look like this:



The fraction of success (within the tolerance) for this trial was 0.9600.

However, with significant computing power – one can repeat this experiment many more times than 100, over many values of  $n$ . Tweaking the original program as such:

% Exercise 1.1.2

% Flip a coin n times, see how close we come to the expected fraction of heads.

```
reps=10000 % Number of repetitions of the experiment
```

```
p=.5 % Probability of heads
```

```
lower=50;
```

```
upper=150;
```

```
results = NaN(1,upper-lower);
```

```
for n=lower:upper %Cycle through different values of n
```

```
headfrac=NaN(reps,1); % Get ready to record the results
```

```
for k=1:reps
```

```
    flips=rand(1,n)<p;
```

```
    total=sum(flips);
```

```
    headfrac(k)=total/n;
```

```
end
```

```
tolerance = .1;
```

```
successes = sum(headfrac>=(p-tolerance) & headfrac<p+tolerance);
```

```
successfrac= successes / reps;
```

```
results(1,n+1-lower)=successfrac;
```

```
end
```

```
results %Report the array with the fractions of success from the lower n to the upper n.
```

Increasing the number of repetitions of the experiment from 100 to 10,000 can be thought of as running that experiment many times over and averaging the results. (to give a slightly better idea). The results of running this program are:

Columns 1 through 10

0.8373 0.8357 0.8734 0.8274 0.8679 0.8550 0.8597 0.8946 0.8506 0.8821

Columns 11 through 20

0.8787 0.8749 0.9070 0.8702 0.8979 0.8903 0.8925 0.9159 0.8794 0.9095

Columns 21 through 30

0.9059 0.9013 0.9235 0.9008 0.9175 0.9119 0.9116 0.9361 0.9143 0.9297

Columns 31 through 40

0.9259 0.9261 0.9370 0.9207 0.9380 0.9340 0.9320 0.9471 0.9297 0.9408

Columns 41 through 50

0.9394 0.9389 0.9526 0.9371 0.9519 0.9478 0.9500 0.9593 0.9459 0.9577

Columns 51 through 60

0.9524 0.9559 0.9659 0.9527 0.9625 0.9571 0.9594 0.9669 0.9632 0.9668

Columns 61 through 70

0.9608 0.9632 0.9727 0.9604 0.9706 0.9645 0.9649 0.9733 0.9664 0.9732

Columns 71 through 80

0.9690 0.9707 0.9794 0.9704 0.9731 0.9743 0.9748 0.9796 0.9732 0.9782

Columns 81 through 90

0.9736 0.9818 0.9804 0.9754 0.9821 0.9798 0.9776 0.9846 0.9803 0.9813

Columns 91 through 100

0.9809 0.9821 0.9864 0.9826 0.9847 0.9850 0.9824 0.9864 0.9834 0.9858

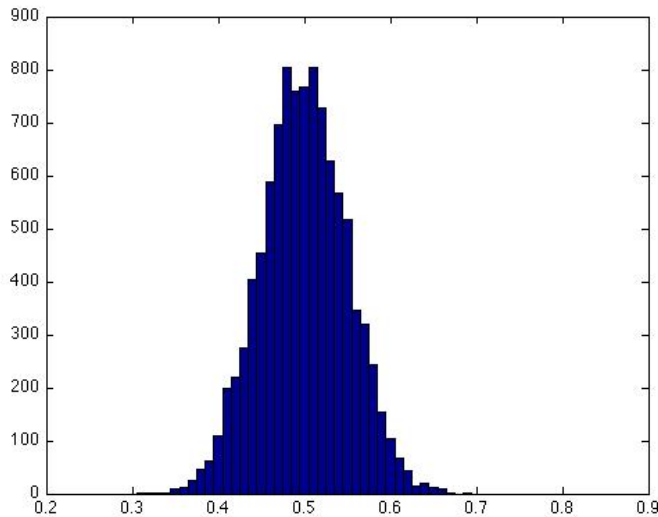
Column 101

0.9850

(where column 1 corresponds to n = 50 and column 101 corresponds to n = 150).

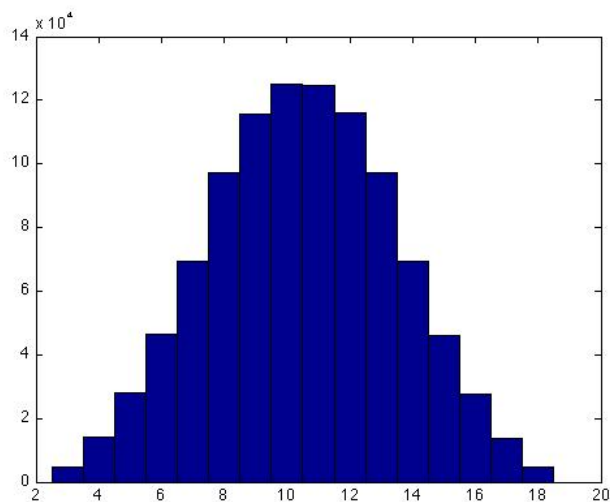
From these results, we see that about 95% of the results are within the tolerance beginning first at  $n = 92$ , but regularly when  $n \approx 100$ .

Running the original program authored by Professor Doyle provides reassurance of this result. This cumulative histogram produced by that program illustrates a tendency towards a normal distribution from 0 to 1 symmetric about 0.5. With  $n = 100$  this produces a “success fraction” of 0.9573:



In conclusion,  $n$  must be about 100 for 95% of the proportion of heads to be within a tolerance of 0.1 about 0.5.

3. In the early 1600s, Galileo was asked to explain the fact that, although the number of triples of integers from 1 to 6 with sum 9 is the same as the number of such triples with sum 10, when three dice are rolled, a 9 seemed to come up less often than a 10 – supposedly in the experience of gamblers. The program provided creates a histogram of the sums of rolling three dice. A graph of the sums of three dice rolled 1000000 times (I adjusted the parameter for a larger sample size) might look like this:



From this graph of a large number of dice rolls, it seems reasonable to conclude, that the gamblers were correct and that the probability of rolling a 10 is greater than that of rolling a 9.

4. In racquetball, a player continues to serve as long as she is winning; a point is scored only when a player is serving and wins the volley. The first player to win 21 points wins the game. Assuming that you serve first and have a probability of 0.6 of winning a volley when you serve and probability of 0.5 when your opponent serves, then by using the simulation as given:

After changing the “repetitions” parameter on the program provided, when running the simulation with 100000 repetitions, the fraction of wins that resulted as 0.8252. From the results of this experiment, it seems reasonable that the probability that you will win a game is likely given as very near 0.825.

5. Consider the bet that all three dice will turn up sizes at least once in  $n$  rolls of three dice. Then the probability of at least one triple-six when the dice are rolled  $n$  times is given as:

$$P_n = f(n) = 1 - \left(\frac{6^3 - 1}{6^3}\right)^n = 1 - \left(\frac{215}{216}\right)^n$$

Given the program provided by Professor Doyle, which conveniently focuses on the interval on which DeMoivre would search, from 145 rolls of three dice to 155 rolls of three dice, it becomes clear that the probabilities over this range are given as:

<i>Rolls</i>	<i>P</i>
145	0.4898
146	0.4921
147	0.4945
148	0.4968
149	0.4991
150	0.5015
151	0.5038
152	0.5061
153	0.5083
154	0.5106
155	0.5129

Thus, the minimum number of rolls at which this is a favorable bet is 150 (and DeMoivre would have been correct with the calculation  $216 \log 2 = 149.7$ ).

## 10. *The Martingale Doubling System*

Using a program in MATLAB as such,

`%Exercise 1.1.10`



```
%An investigation into the Martingale Betting System
%As applied when betting on roulette:
```

```
%Rules
```

```
%(1) Bet one dollar the first time
%(2) Every time you win, bet 1 dollar the next time.
%(3) Every time you lose,double your previous bet.
%(4) Continue to play until you have won at least 5
%dollars or you have lost more than 100 dollars.
```

```
%The probability of landing on red in American roulette is 18/38
P=18/38
```

```
trips=1000 %Number of times playing the game as stated by the rules above
totals = zeros(1,trips);
```

```
for i=1:trips
fortune = 0;
bet = 1; % The initial bet is 1
```

```
while((fortune < 5) && (fortune > -100))
winlose = rand < P;
if(winlose == 1)
fortune = fortune+bet;
bet=1;
else
fortune = fortune-bet;
bet=bet*2;
end
end
```

```
%Just by executing the "while" statement above which simulates one trip to
%the casino - it might seem like a good idea to use the Martingale system.
%HOWEVER - over a number of "trips" to the casino - we see that because
%losses add up exponentially - over several trips the 'bad' can quickly
%outweigh the 'good' -- to drastic effect on one's fortune!
```

```
totals(1,i) = fortune;
end
```

```
sum(totals)
totals
```

```
%In the long run -- it isn't a winning strategy (especially given that no one has
%infinite wealth to risk on roulette).
```

This program is designed to investigate the long-term efficacy of the martingale doubling system on this bet. Running just a small number of trials – it might seem as if the system is effective (it most often ends up winning \$5). Unfortunately, this might even convince amateur gamblers that it is a good system. However, when the

system fails – it fails to such a large degree that it could bankrupt the gambler (under these rules, failure tends to be around -\$125, or 25 winning trips! – but the parameters of the game can be tweaked to see different results). By summing the total winnings (loses) from a large number of trips to the casino (theoretically, either by one person or many) we see that it very often loses more than it wins in the long run. An unfair game remains an unfair game despite the betting system.

The program above returns a large array filled with mostly 5s and scattered numbers about equal to -127 to -123. The total might return -1634.