Due: Friday, Jan. 30, 2015.

# STATISTICS #1

**1. Birthdays:** If there are 20 people in Advanced lab. What is the probability that ***at least*** 2 people share the same birthday? What is the probability that another person shares your birthday?

**2. The Monty Hall Problem:** In a famous game show Monty Hall has three doors with a gift behind one of them. The contestant picks a door. Then Monty Hall, who knows where the gift is, opens one of the other doors to show that the gift is not behind it. Now Monty gives the contestant a choice does they can stay with the original door choice or now pick the other unopened door. What should the contestant do to maximize their chances at winning, hold or switch? Seems simple enough, there is equal; probability to be behind either remaining door so switching should make no difference. But, alas, it is not so simple. In this exercise we are going to empirically test which strategy is best for the Monty Hall problem, and based on the conclusions we will re-think the problem.

**Empirical test:** One partner will be Monty the other the contestant. The Monty should secretly shake a die to decide where the “prize” is. (use 1 and 2 for door 1, 3 and 4 for door 2, etc.). The contestant should choose a door, and in response Monty should next reveal that the gift is not behind one of the doors. Now we can test the two strategies by revealing where the prize is and keeping a tally of the wins and losses.
Repeat this, recording the results, until your conclusions are statistically meaningful. *i.e.* give some thought as to what is statistically meaningful in this case.
**Re-think the problem:** The situation is more complicated than expected. To hone in on an explanation think of related problems where, say there are 10 doors with one gift. After choosing a door Monty now opens all doors but two, yours and another.
Now does it make sense to switch?

# STATISTICS #1 (cont'd)

**3. Fair and Loaded Dice:** Each laboratory pair of experimenters is to take two groups of 600 die rolls using a fair die and one group of 600 die rolls of the shaved and loaded die. Keep track of each roll in an ordered vertical column so that we keep the original set of data.

a) Histogram the die-roll numbers in unit bins (1 to 6) separately for the fair, loaded and shaved dice. Test the fairness of these histograms using simple counting statistics, $\sigma_\iota$ = sqrt ($N_i$ ) and using the "$\chi^2$ goodness of fit test". Are the shaved and loaded dice fair or not?

b) Now histogram the sum of independent die-roll pairs; *i.e.,* adjacent pair and histogram the sums. What is the expected fair histogram. Test the fairness of these histograms as above.

c) Finally, consider the length of intervals between ones for each of the dice. From the raw data determine the intervals and histogram them. Compare this with what is theoretically expected. Again test the fairness using "$\chi^2$ goodness of fit test".

**<u>Dice Data</u>**

| **Die Name** | **Mass (grams)** | **Face-to-Face Separation (in.) (1:6):(2:5):(3:4)** | **Other** |
|---|---|---|---|
| Normal 1 (N1) | 5.53 | 0.633:0.624:0.623 | |
| Normal 2 (N2) | 4.41 | 0.608:0.607:0.605 | |
| Milled (M)<br>*flat or brick* | 5.23 | 0.592:0.626:0.628 | "1" milled |
| Weighted (P)<br>*loaded* | $m_{die}$=6.45<br>$m_{Pb}$=1.18 | 0.632:0.623:0.622 | "4" weighted cylinder: dia. 0.188", L 0.239" |
| Magnetic (S)<br>*loaded* | $m_{die}$=5.85<br>$m_{Steel}$=0.778 | 0.619:0.613:0.623 | "4" weighted cylinder: dia. 0.213", L 0.187" |
| Beveled (B) | | 0.615:0.616:0.614 | $A_1$= 0.617" x 0.619";<br>$A_2$= 0.621" x 0.515"<br>$A_3$= 0.615" x 0.541";<br>$A_4$= 0.619" x 0.541"<br>$A_5$= 0.618" x 0.536";<br>$A_6$= 0.442" x 0.458" |