Introduction to Probability and Information Theory
Homework #2

G&S pg. 154

#30 Of course, both managers are technically right, but factory A probably really is the better producer. If A switched to B's lot sizes, its overall bad bulb rate would fall to $2000 \times 0.002 + 1000 \times 0.0055 = 9.5$ per delivery, while if B switched to A's lot sizes its bad bulb rate would increase to $1000 \times 0.0025 + 2000 \times 0.006 = 14.5$ per delivery. This problem is an example of **Simpson's paradox**: factory A is better than B on every individual measure, but apparently worse overall (until we take into account the differences in lot sizes).

#46 Let X be the number of aces in a hand and Y be the number of kings. We want to know their joint distribution: the probability $P(X = a, Y = b)$ of getting a aces and b kings in a five card hand. There are $\binom{4}{a}$ ways of dealing out a aces and $\binom{4}{b}$ ways of dealing out b kings. Finally, to make a hand we'll need to deal out $5 - a - b$ additional non-aces and non-kings, and there are $\binom{44}{5-a-b}$ ways to do that. Since there are $\binom{52}{5}$ total hands, the joint probability is:

$$P(X = a, Y = b) = \frac{\binom{4}{a} \times \binom{4}{b} \times \binom{44}{5-a-b}}{\binom{52}{5}}$$

for $0 \le a, b \le 4$.

This problem is complicated enough that it's a good idea to experimentally validate our theoretical result. It would take a very long time to deal out enough poker hands to get a good sense of this distribution, but we can perform a simple simulation in R (in dealing out a hand with `sample(1:52,5)`, we'll call cards 1,2,3,4 aces and 5,6,7,8 kings).

```
> aces = 0
> kings = 0
> R = 100000
> for (i in 1:R) {
+   hand = sample(1:52,5)
+   aces[i] = sum(hand <= 4)
+   kings[i] = sum(hand >= 5 & hand <= 8)
+ }
> table(aces,kings)/R
    kings
aces       0       1       2       3       4
   0 0.41949 0.20911 0.02996 0.00137 0.00000
   1 0.20848 0.08090 0.00876 0.00032 0.00002
   2 0.03066 0.00863 0.00062 0.00001 0.00000
   3 0.00135 0.00026 0.00003 0.00000 0.00000
   4 0.00003 0.00000 0.00000 0.00000 0.00000
```

For comparison, we can also compute the predicted values given our model:

```
> p = function(a,b) {
+   (choose(4,a)*choose(4,b)*choose(44,5-a-b))/choose(52,5)
+ }
> outer(0:4,0:4,p)
             [,1]         [,2]         [,3]         [,4]
[1,] 4.178625e-01 2.089313e-01 3.057531e-02 1.455967e-03
[2,] 2.089313e-01 8.153415e-02 8.735802e-03 2.708776e-04
[3,] 3.057531e-02 8.735802e-03 6.094746e-04 9.234463e-06
[4,] 1.455967e-03 2.708776e-04 9.234463e-06 0.000000e+00
[5,] 1.692985e-05 1.539077e-06 0.000000e+00 0.000000e+00
             [,5]
[1,] 1.692985e-05
[2,] 1.539077e-06
[3,] 0.000000e+00
[4,] 0.000000e+00
[5,] 0.000000e+00
```

The simulated values are pretty close to the predicted values, so we can conclude that our model is probably reasonable. (Later in the course we'll get into measuring how close is close enough and how confident with should be about our model based on results like this).

#56 Suppose A attracts B. Assuming that $P(A) > 0$ and $P(B) > 0$, we have:

$$\begin{aligned} P(B|A) &> P(B) \\ \frac{P(AB)}{P(A)} &> P(B) \\ P(AB) &> P(A)\,P(B) \\ P(A|B)\,P(B) &> P(A)\,P(B) \\ P(A|B) &> P(A) \end{aligned}$$

which means that B attracts A. By the same reasoning, we can show that if B attracts A then A attracts B.

G&S pg. 197

#1 (a) yes

(b) no

(c) yes

(d) maybe (yes in the usual idealized probability-textbook case, no in reality)

(e) no

#2 There are 20! possible permutations of the numbers $1, \ldots 20$, so if X is uniformly distributed, then $P(X)$ must be $\frac{1}{20!}$ for every possible value of X. Take some permutation $\langle x_1 x_2 \ldots x_{20} \rangle$. The probability that the number x_1 is selected first is $\frac{1}{20}$, the probability that x_2 is selected second given that x_1 is selected first would be $\frac{1}{19}$, the probability of x_3 being selected third given x_1 and x_2 is $\frac{1}{18}$, etc. Then:

$$P(\langle x_1 x_2 \ldots x_{20} \rangle) = \frac{1}{20 \times 19 \times \cdots \times 1} = \frac{1}{20!}$$

G&S pg. 247

#1 There are five even cards and four odd cards, so the expected payoff is $1 \times \frac{4}{9} - 1 \times \frac{5}{9} = -\frac{1}{9}$.

#3 To make things easier, we'll calculate the expected value in inches: $66 \times \frac{3}{20} + 68 \times \frac{5}{20} + 70 \times \frac{4}{20} + 72 \times \frac{4}{20} + 74 \times \frac{4}{20} = 70.1$ or 5'10.1". There are lots of ways of doing this arithmetic, but I did it this way using R:

```
> sum(c(66*3,68*5,70*4,72*4,74*4)/20)
[1] 70.1
```