Math 20 – Problem Set 3 (due July 18)

This problem set is due at the *beginning* of class. This is just the problem list; please work out these problems on a different sheet of paper. Please write neatly, staple the pages together, and explain your work where appropriate. You do not need to simplify binomial coefficients $\binom{n}{k}$ for both which $k > 3$ and $n - k > 3$, or exponentials n^k where $n + k > 8$.

1. *Someone offers you the following game: You roll a fair six-sided die. If you roll a 1, you win \$25. If you roll a 2, you win \$5. If you roll a 3, you win nothing. If you roll a 4 or a 5, you lose \$10. If you roll a 6, you lose \$15. Should you play this game? Why or why not?*

   We'll compute the expected payoff of this game. Since each of the results (1 through 6) occurs with equal probability, we should add the results and divide by 6:

   $$E(\text{payoff}) = \frac{1}{6}\left(25 + 5 + 0 - 10 - 10 - 15\right) = -\frac{5}{6}.$$

   On average, you will lose \$0.83 playing this game – you shouldn't play it!

2. *Suppose n bar patrons each check their hats at the door. When they leave, intoxicated, they each take a hat at random. What is the expected number of patrons who leave with the same hat they were wearing when they arrived? How does this number depend on n? (There's an easy way to do this, and there's a hard way...)*

   Let X be a random variable whose value is the number of patrons who receive their own hats back upon leaving. It's possible to compute the probability distribution for X, but this is hard – we'd need to use the full strength of the principle of inclusion-exclusion. Instead, we'll use linearity of expectation!

   For each i between 1 and n, let X_i be a random variable taking the value 1 if the ith person gets his own hat back, or 0 otherwise (this is called an *indicator random variable*). Each X_i is a Bernoulli random variable with parameter $\frac{1}{n}$, since this is the probability that any single patron will get his or her hat back (notice that we don't need to consider the other $n - 1$ people in making this conclusion).

   Then X is a *count* of the number of X_i's that take the value 1:

   $$X = \sum_{i=1}^{n} X_i.$$

   Linearity of expectation then gives us

   $$E(X) = E\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} E(X_i) = n \cdot \frac{1}{n} = 1.$$

It turns out that, on average, only one person will get his or her hat back, and this result doesn't depend on the value of n! This might seem somewhat surprising as the events are clearly dependent, but this method hints at the power of the technique of linearity of expectation for solving problems involving many dependent events.

Note: it's incorrect to say that X has the binomial distribution with parameters n and $\frac{1}{n}$, since one person getting their hat back is not independent of another person getting their hat back (in fact, these events are positively correlated).

3. *If 20 indistinguishable balls are allocated randomly into 10 differently-colored boxes in such a way that all allocations are equally likely (eg. having 20 balls in the first box is just as likely as having 20 balls in the third box, and both are just as likely as having 2 in each box), what is the probability that no box is empty? Explain your answer.*
(This was problem #12 on your exam.)

This is a stars-and-bars problem! If all allocations are equally likely, then the number of allocations of 20 identical balls into 10 distinct boxes is $\binom{20+10-1}{20} = \binom{29}{20}$. Since all outcomes are equally likely, it remains only to count the number of favorable outcomes – events in which all boxes have at least one ball.

One way to do this is to fill each box with one of the 20 balls; all that's left after doing this is to allocate the remaining 10 balls into the 10 boxes with no restrictions. Thus the probability is

$$\frac{\binom{10+10-1}{10}}{\binom{29}{20}} = \frac{\binom{19}{10}}{\binom{29}{20}}.$$

4. *In class on Friday we (hopefully) discussed the triangular distribution that results from rolling 2 fair 6-sided dice and taking their sum. Write an R program to simulate rolling 4 fair 6-sided dice and taking their sum. Do this at least 10000 times and make a table of your results. Then visualize your results using R's* `plot()` *function. Explain in a sentence or two how this distribution compares to the triangular distribution. Please print your code and output and attach it with this problem set.*

The resulting plot is no longer triangular – it has more of a bell-shaped curve. The distribution is more centered around the mean of 14 than the triangular distribution, which decreases linearly away from its mean of 7. It will turn out that the more dice we add together, the more the result will look like a bell-shaped normal distribution – this is the content of the *central limit theorem*.

You can see my code next to the link to this solutions page on the course website.