Solutions to Assignment #11

1. *[Problem 6.2.5 on page 238 in Allman and Rhodes].* Explain the following results by not referring to the formula $\binom{n}{k} = \frac{n!}{k!(n-k)!}$, but in terms of "choosing objects."

   (a) $\binom{n}{1} = n$ and $\binom{n}{n-1} = n$ for any n.

   Solution: $\binom{n}{1}$ is the number of distinct ways of choosing 1 object out of n. This can be done in n ways.

   $\binom{n}{n-1}$ is the number of way of choosing $n-1$ objects out of n. Since in each of these choices we would be missing exactly one of the objects, there are n ways of making the choices. □

   (b) $\binom{n}{0} = 1$ and $\binom{n}{n} = 1$ for any n.

   Solution: There is only one way of choosing no objects at all, and only one way of choosing all the n objects. □

2. *[Problem 6.2.6 on page 238 in Allman and Rhodes].* Suppose a family has six children.

   (a) What is the probability that four are boys and two are girls?
   Solution: Assume that having a boy or having a girl are equally likely events for this family. It then follows that the probability that a child is a boy is $p = \frac{1}{2}$. If X is the number of boys in the family, then X has a binomial distribution with parameters $p = \frac{1}{2}$ and $N = 6$. Hence,

$$P[X = n] = \binom{N}{n} p^n (1-p)^{N-n} = \binom{6}{n} \left(\frac{1}{2}\right)^6 \qquad (1)$$

   for $n = 0, 1, \ldots, 6$.

The probability of four boys and two girls is then

$$P[X = 4] = \binom{6}{4}\frac{1}{2^6} = \frac{15}{64}. \qquad \square$$

(b) Give the probability distribution for X.

Solution: This was given in equation (1). $\square$

(c) What is the expected number of boys in the family?

Solution: $E(X) = pN = \frac{1}{2} \cdot 6 = 3.$ $\square$

(d) What is the probability that the family has four or more girls?

Solution: This is the same as the probability that the family two or fewer than two boys:

$$P[X \leq 2] = P[X = 0] + P[X = 1] + P[X = 2].$$

Thus, by the probability distribution function of X in (1),

$$\begin{aligned} P[X \leq 2] &= \binom{6}{0}\frac{1}{2^6} + \binom{6}{1}\frac{1}{2^6} + \binom{6}{2}\frac{1}{2^6} \\ &= \frac{1}{64}[1 + 6 + 15] \\ &= \frac{22}{64} \\ &= \frac{11}{32} \end{aligned}$$

3. *[Problem 6.2.16 on page 240 in Allman and Rhodes].* In humans, the hereditary Huntington disease is caused by a dominant mutation. Onset of Huntington disease occurs in midlife, between 35 and 44 years of age typically, and the progressive disorder leads eventually to death. Suppose that in a married couple one individual carries the allele for Huntington disease, and that the couple has four children.

(a) What is the probability that none of the children will develop Huntington disease?

Solution: To do this problem first we model the number of children, X, that develops the disease by a binomial random variable with parameters

p and $N = 4$, where p is the probability that an offspring carries the Huntington allele, H. If we assume the genotype of the parents are Hh and hh, respectively, then p is the probability that an offspring has the genotype Hh. Thus,

$$p = P[\text{allele } H \text{ from } 1^{\text{st}} \text{ parent}] \cdot P[\text{allele } h \text{ from } 2^{\text{nd}} \text{ parent}] = \frac{1}{2} \cdot 1 = \frac{1}{2}.$$

It then follows that

$$P[X = n] = \binom{4}{n} \left(\frac{1}{2}\right)^4 \tag{2}$$

for $n = 0, 1, 2, 3, 4$.

Thus, the probability that none of the children will develop Huntington disease is

$$P[X = 0] = \binom{4}{0} \frac{1}{2^4} = \frac{1}{16}. \qquad \square$$

(b) What is the probability that at least one of the children will develop Huntington disease?

Solution: The event [at least one of the children will develop Huntington] is the complement of the event [none of the children will develop Huntington]. Thus, by the previous part,

$$P[X \geq 1] = 1 - P[X = 0] = 1 - \frac{1}{16} = \frac{15}{16} \quad \text{or} \quad 93.75\% \qquad \square.$$

(c) What is the probability that three or more of the children will develop Huntington disease?

Solution: This is $P[X \geq 3] = P[X = 3] + P[X = 4]$. Thus, by (2),

$$P[X \geq 3] = \binom{4}{3} \frac{1}{16} + \binom{4}{4} \frac{1}{16} = 4 \cdot \frac{1}{16} + \frac{1}{16} = \frac{5}{16} \quad \text{or} \quad 31.25\% \qquad \square.$$

4. *[Problem 6.2.18 on pages 240 and 241 in Allman and Rhodes].* The goal of this exercise is to derive the formula

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} \tag{3}$$

for counting combinations of n things take k at a time.

Formally, a combination of n things taken k at a time is an *unordered* subset of the set of n objects consisting of k elements. However, it is better to think of it more concretely as follows. Imagine a box containing n balls with the numbers $1, 2, 3, \ldots, n$ printed on them. Pick k of the balls and place them in a row in the order in which you picked them. Then, since the *order* in which you picked them does not matter, put them in a bag. What you end up with in the bag is a combination. The number of different bags of balls is $\binom{n}{k}$.

(a) When you pick the first ball out of the box, how many different choices could you make for it? When you pick the second ball, why are there $n-1$ different choices for it? For the l^{th} ball, why are there $n-l+1$ choices?

Solution: Since there are n different balls in the box to begin with, there are n different choices for the first pick. Once this one is chosen and taken out of the box, there remain $n-1$ different choices in it, and so there are $n-1$ different choices for the second pick. Similarly, for the third pick there would me $n-2$ choices; for the fourth pick, $n-3$ choices. Continuing in this fashion, we see that for the ℓ^{th} pick there would be $n-(\ell-1)$ or $n-\ell+1$ choices. $\square$

(b) Why does part (a) indicate that, when the k balls are all in a row, there are $n(n-1)(n-2)\cdots(n-k+1)$ possible choices you might have made? (The count of these *ordered* choices is sometimes called a *permutation.*)

Solution: For each choice in the first pick there are $n-1$ choices in the second pick; thus, there are $n(n-1)$ choices for the first two balls. For each one of these there are $n-2$ choices for the third pick, and so there are $n(n-1)(n-2)$ possibilities for the first three balls. Continuous in this fashion we then see that there are $n(n-1)(n-2)\cdots(n-k+1)$ for choosing k balls. $\square$

(c) Several different *ordered choices* might lead to the same collections of balls in the bag (i.e., to the same *combination*), so the answer to part (b) is bigger than the number of combinations. To see how much bigger, it's easiest to imagine having the balls in the bag, and (going backwards in time) putting them back in order in a row. Using reasoning similar to that used in parts (a) and (b), explain why there are $k!$ choices of ways this could be done.

Solution: When there are k balls in the bag, there are k different choices for the first ball in the row. Once this first ball is picked, for each of these choices, there are $k-1$ different ways of picking the second ball put of the bag; thus, there are $k(k-1)$ choices for the first to balls in the row. Similarly there are $k(k-1)(k-2)$ for the first three balls in the row. Hence, continuing in this fashion, there are $k(k-1)(k-2)\cdots 2\cdot 1 = k!$ choices for putting back the balls from the bag in a row. □

(d) Using parts (b) and (c), conclude that

$$\binom{n}{k} = \frac{n(n-1)(n-2)\cdots(n-k+1)}{k!}. \tag{4}$$

Solution: The total number of different bags of k balls chosen out of n balls in the box is equal to the number of ways of getting the k balls out of the box in the first place and putting them in the order of pick in a row divided by the different ways of ordering the k balls in the row, namely the number of permutations of the k balls in the row. Since there are $n(n-1)(n-2)\cdots(n-k+1)$ of picking k balls out of the box and putting them in row by part (b), the result follows. □

(e) Explain why the formula in (4) can be written as the formula in (3).

Solution: Observe that $n! = n(n-1)(n-2)\cdots(n-k+1)\cdot(n-k)!$, so that

$$\frac{n!}{(n-k)!} = n(n-1)(n-2)\cdots(n-k+1).$$

It then follows that

$$\frac{n(n-1)(n-2)\cdots(n-k+1)}{k!} = \frac{n!}{k!(n-k)!}. \qquad \square$$

5. The data in Table 1 were taken from page 504 of the Luria and Delbrück 1943 paper.

Table 1: Number of resistant bacteria in a series of similar cultures

Test–tube #	1	2	3	4	5	6	7	8	9	10	11	12
# of Mutants	1	0	0	7	0	303	0	0	3	48	1	4

For the data in Table 1:

(a) Estimate the average number of resistant bacteria right before the plating was made.

Solution: This is estimated by taking the average of the numbers of resistant bacteria in Table 1; namely, $\overline{r} \approx 30.58$. □

(b) Use the *sample–variance* formula $s^2 = \dfrac{\displaystyle\sum_{i=1}^{n}(r_i - \overline{r})^2}{n-1}$, where r_i denotes the number of resistant cells in test–tube i and $\overline{r}$ is the average number of resistant bacteria, to estimate the variance of the distribution.

Solution: $\text{var}(R) \approx s^2 \doteq 6913.74$. □

(c) Based on your results in the previous part and what you know about the Poisson process, would you say that the number of resistant bacteria follows a Poisson process? Justify your answer.

Solution: If the number of resistant bacteria truly follows a Poisson distribution, then its mean and variance should be very close. Thus, in a random sample, the sample mean and variance should be very close also. This is not the case for the data in Table 1. Thus, the number of the resistant bacteria does not follow a Poisson distribution. □