STAT 425 – HW12

Assigned on Wed Oct 19, due on Tues Oct 25

Section 3.1 Random Variables and Discrete Distributions

1. Reconsider the “composite blood testing” example from class: Suppose that in a group of 10 people, each person has a .1 probability of having a certain disease, independently from person to person. Now consider a slightly different plan: The ten people are first randomly divided into two groups of five people, and then the composite testing plan (described in class) is implemented within each group of five people. Let the random variable $X$ represent the total number of tests required with this plan.

a) Determine the probability distribution of $X$.

If nobody has the disease, then 2 tests are needed. The probability of this is $(.9)^{10} \approx .3487$.

If one group of 5 has someone with the disease (which has probability $1 - (.9)^5 \approx .4095$) and the other group of 5 has nobody with the disease (which has probability $(.9)^5 \approx .5905$), then 7 tests are needed. The probability of this is $2(.9)^5[1 - (.9)^5] \approx .4836$.

If both groups of 5 have someone with the disease, then 12 tests are needed. The probability of this is $[1 - (.9)^5]^2 \approx .1677$.

The probability distribution of $X$ is therefore:

<table>
<thead>
<tr>
<th>$x$</th>
<th>2</th>
<th>7</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p(x)$</td>
<td>$(.9)^{10} \approx .3487$</td>
<td>$2(.9)^5[1 - (.9)^5] \approx .4836$</td>
<td>$[1 - (.9)^5]^2 \approx .1677$</td>
</tr>
</tbody>
</table>

b) Determine the probability distribution of $X$ in the more general case where the probability of having the disease is $p$ for each person.

By the same analysis, the probability distribution of $X$ is now:

<table>
<thead>
<tr>
<th>$x$</th>
<th>2</th>
<th>7</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p(x)$</td>
<td>$(1-p)^{10}$</td>
<td>$2(1-p)^5[1 - (1-p)^5] \approx .4836$</td>
<td>$[1 - (1-p)^5]^2 \approx .1677$</td>
</tr>
</tbody>
</table>

c) Determine the probability distribution of $X$ in the still more general case where there are $2n$ people, where $n$ is a positive integer, and assuming that the plan first divides the $2n$ people into two groups of $n$ people.

By the same analysis, the probability distribution of $X$ is now:

<table>
<thead>
<tr>
<th>$x$</th>
<th>2</th>
<th>7</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p(x)$</td>
<td>$(1-p)^{2n}$</td>
<td>$2(1-p)^n[1 - (1-p)^n] \approx .4836$</td>
<td>$[1 - (1-p)^n]^2 \approx .1677$</td>
</tr>
</tbody>
</table>
2. a) Suppose that you choose a letter at random from the word “statistics” and that you continue to choose letters at random (without replacement) until you select a vowel. Let the random variable $Y$ represent the number of letters that you choose in order to get a vowel. Determine the probability distribution of $Y$.

Let $V_i$ be the event that the $i$th letter chosen is a vowel, and similarly define $C_i$ for a consonant.

\[
\Pr(Y = 1) = \Pr(V_1) = \frac{3}{10}
\]

\[
\Pr(Y = 2) = \Pr(C_1 \cap V_2) = \Pr(C_1) \Pr(V_2 | C_1) = \left(\frac{7}{10}\right)\left(\frac{3}{9}\right) = \frac{7}{30}
\]

\[
\Pr(Y = 3) = \Pr(C_1 \cap C_2 \cap V_3) = \left(\frac{7}{10}\right)\left(\frac{6}{9}\right)\left(\frac{3}{8}\right) = \frac{7}{40}
\]

\[
\Pr(Y = 4) = \Pr(C_1 \cap C_2 \cap C_3 \cap V_4) = \left(\frac{7}{10}\right)\left(\frac{6}{9}\right)\left(\frac{5}{8}\right)\left(\frac{3}{7}\right) = \frac{1}{8}
\]

\[
\Pr(Y = 5) = \Pr(C_1 \cap C_2 \cap C_3 \cap C_4 \cap V_5) = \left(\frac{7}{10}\right)\left(\frac{6}{9}\right)\left(\frac{5}{8}\right)\left(\frac{4}{7}\right)\left(\frac{3}{6}\right) = \frac{1}{12}
\]

\[
\Pr(Y = 6) = \Pr(C_1 \cap C_2 \cap C_3 \cap C_4 \cap C_5 \cap V_6) = \left(\frac{7}{10}\right)\left(\frac{6}{9}\right)\left(\frac{5}{8}\right)\left(\frac{4}{7}\right)\left(\frac{3}{6}\right)\left(\frac{3}{5}\right) = \frac{1}{20}
\]

\[
\Pr(Y = 7) = \Pr(C_1 \cap C_2 \cap C_3 \cap C_4 \cap C_5 \cap C_6 \cap V_7) = \left(\frac{7}{10}\right)\left(\frac{6}{9}\right)\left(\frac{5}{8}\right)\left(\frac{4}{7}\right)\left(\frac{3}{6}\right)\left(\frac{2}{5}\right)\left(\frac{3}{4}\right) = \frac{1}{40}
\]

\[
\Pr(Y = 8) = \Pr(C_1 \cap C_2 \cap C_3 \cap C_4 \cap C_5 \cap C_6 \cap C_7 \cap V_8) = \left(\frac{7}{10}\right)\left(\frac{6}{9}\right)\left(\frac{5}{8}\right)\left(\frac{4}{7}\right)\left(\frac{3}{6}\right)\left(\frac{2}{5}\right)\left(\frac{1}{4}\right)\left(\frac{3}{3}\right) = \frac{1}{120}
\]

b) Repeat a) in the case that the sampling of letters is conducted with replacement.

Now $\Pr(Y = y) = (0.3)^{y-1} \times (0.7)^{y-1}$ for $y = 1, 2, 3, \ldots$

3. Suppose that slips of paper containing integers 1-5 are randomly distributed to five people (call them A, B, C, D, and E). The game starts with A and B comparing their numbers. Whoever has the higher number is the winner, and that person proceeds to compare his/her number with C’s number. The game proceeds in this manner until the fourth comparison (involving E and the winner of the previous comparison) takes place. Let the random variable $W$ represent the number of comparisons for which A is the winner. Determine the probability distribution of $W$.

$W = 0$ when A draws a smaller number than B. There are $5 \times 4 = 20$ ways for A and B to choose their two numbers, and there are 10 ways for A to draw a smaller number than B: $(1,2), (1,3), (1,4), (1,5), (2,3), (2,4), (2,5), (3,4), (3,5), (4,5)$. So, $\Pr(W = 0) = 10/20 = 1/2$.

$W = 1$ when A draws a larger number than B and a smaller number than C. There are $5 \times 4 \times 3 = 60$ ways for A and B and C to choose their three numbers. If A selects number 2, then there is only 1 choice for B and 3 choices for C so that A wins only one comparison. If A selects number 3, then there are 2 choices for B and 2 choices for C, so 4 choices so that A wins one
comparison. If A selects number 4, then there are 3 choices for B and 1 choice for C so that A wins one comparison. Thus, there are $3 + 4 + 3 = 10$ ways for A to win one comparison. So, $\Pr(W = 1) = 10/60 = 1/6$.

$W = 2$ when A draws a larger number than B, a larger number than C, and a smaller number than D. There are $5 \times 4 \times 3 \times 2 = 120$ ways for A and B and C and D to choose their three numbers. If A selects number 3, then there are 2 choices for B and then 1 choice for C and 2 choices for D, so 4 choices altogether, so that A wins exactly two comparisons. If A selects number 4, then there are 3 choices for B and then 2 choices for C and 1 choice for D, so 6 choices altogether so that A wins two comparisons. Thus, there are $4 + 6 = 10$ ways for A to win two comparison. So, $\Pr(W = 2) = 10/120 = 1/12$.

$W = 3$ when A draws a larger number than B, C, and D but a smaller number than E. This can only happen when A draws 4 and E draws 5, but B, C, and D have $3! = 6$ choices for their numbers. So, $\Pr(W = 4) = 6/120 = 1/20$.

$W = 4$ happens exactly when A draws number 5, which has probability $1/5$.

Thus, the probability distribution of $W$ is:

<table>
<thead>
<tr>
<th>$w$</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p(w)$</td>
<td>1/2</td>
<td>1/6</td>
<td>1/12</td>
<td>1/20</td>
<td>1/5</td>
</tr>
</tbody>
</table>

Not to turn in: Consider working on exercises 1-10 at the end of section 3.1 in the text.