**A Case of Possible Discrimination (Spotlight Task)**

This activity is adapted by one of the authors, Christine Franklin, from *Navigating through Data Analysis in Grades 9-12*, Burrill, Gail, Christine Franklin, Landy Godbold, and Linda Young; p. 29-41, NCTM, Reston, VA. 2003.

Common Core Georgia Performance Standards Addressed:
- **MCC9-12.S.CP.1** Describe events as subsets of a sample space (the set of outcomes) using characteristics (or categories) of the outcomes, or as unions, intersections, or complements of other events (“or,” “and,” “not”).★
- **MCC9-12.S.CP.2** Understand that two events A and B are independent if the probability of A and B occurring together is the product of their probabilities, and use this characterization to determine if they are independent.★
- **MCC9-12.S.CP.3** Understand the conditional probability of A given B as P(A and B)/P(B), and interpret independence of A and B as saying that the conditional probability of A given B is the same as the probability of A, and the conditional probability of B given A is the same as the probability of B.★
- **MCC9-12.S.CP.4** Construct and interpret two-way frequency tables of data when two categories are associated with each object being classified. Use the two-way table as a sample space to decide if events are independent and to approximate conditional probabilities.★
- **MCC9-12.S.CP.5** Recognize and explain the concepts of conditional probability and independence in everyday language and everyday situations.★

Standards of Mathematical Practice Addressed:
1. Make sense of problems and persevere in solving them.
2. Reason abstractly and quantitatively.
3. Construct viable arguments and critique the reasoning of others.
4. Model with mathematics.
6. Attend to precision.
7. Look for and make use of structure.

**Spotlight Task: A Case of Possible Discrimination**

Statisticians are often asked to look at data from situations where an individual or individuals believe that discrimination has taken place. A well-known study of possible discrimination was reported in the *Journal of Applied Psychology*. The scenario of this study is given below. The study motivates statistical reasoning of a context that is consistently relevant (discrimination cases that often end up in court).

This particular study examined whether discrimination was being practiced against women by male supervisors in the banking industry.

**SCENARIO**

Researcher conducted a study at a banking conference where 48 male bank supervisors were each given the same personnel file and asked to judge whether the person should be promoted to a branch manager job that was described as ‘routine’ or whether the person’s file should be held and other applicants interviewed. The files
were identical except that half of the supervisors had files showing the person was male while the other half had files showing the person was female. Of the 48 files reviewed, 35 were promoted. (B. Rosen and T. Jerdee (1974), “Influence of sex role stereotypes on personnel decisions,” *J. Applied Psychology*, 59:9-14.)

**PRELIMINARY QUESTIONS**

1. What is the statistical question being asked by the researchers? What are the variables of interest that the researchers will measure to answer this question? Are these variables categorical or quantitative?

2. Why would this study be classified as an experiment and not an observational study?

What would need to be assumed about the manner in which the files were distributed to the banking supervisors in order to infer that gender is the cause of the apparent differences?

3. Suppose there was no discrimination involved in the promotions. Enter the expected numbers of males promoted and females promoted for this case in Table 1. Explain your choice of numbers.

<table>
<thead>
<tr>
<th></th>
<th>PROMOTED</th>
<th>NOT PROMOTED</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>MALE</td>
<td></td>
<td></td>
<td>24</td>
</tr>
<tr>
<td>FEMALE</td>
<td></td>
<td></td>
<td>24</td>
</tr>
<tr>
<td>TOTAL</td>
<td>35</td>
<td>13</td>
<td>48</td>
</tr>
</tbody>
</table>

Table 1 No discrimination

4. Suppose there was strong evidence of discrimination against the women in those recommended for promotion. Create a table that would show this case. Explain your choice of numbers.

<table>
<thead>
<tr>
<th></th>
<th>PROMOTED</th>
<th>NOT PROMOTED</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>MALE</td>
<td></td>
<td></td>
<td>24</td>
</tr>
<tr>
<td>FEMALE</td>
<td></td>
<td></td>
<td>24</td>
</tr>
<tr>
<td>TOTAL</td>
<td>35</td>
<td>13</td>
<td>48</td>
</tr>
</tbody>
</table>

Table 2 Strong case of discrimination against the women
5. Suppose the evidence of discrimination against the women falls into a ‘gray area’; i.e., a case where discrimination against the women is not clearly obvious without further investigation. Create a table that would show this case. Explain your choice of numbers.

<table>
<thead>
<tr>
<th></th>
<th>PROMOTED</th>
<th>NOT PROMOTED</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>MALE</td>
<td></td>
<td></td>
<td>24</td>
</tr>
<tr>
<td>FEMALE</td>
<td></td>
<td></td>
<td>24</td>
</tr>
<tr>
<td>TOTAL</td>
<td>35</td>
<td>13</td>
<td>48</td>
</tr>
</tbody>
</table>

Table 3. ‘Gray’ case of discrimination against the women

Returning to the study reported earlier in the activity scenario, the results were reported that of the 24 “male” files, 21 were recommended for promotion. Of the 24 “female” files, 14 were recommended for promotion.

6. Enter the data from the actual discrimination study in Table 4.

<table>
<thead>
<tr>
<th></th>
<th>PROMOTED</th>
<th>NOT PROMOTED</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>MALE</td>
<td></td>
<td></td>
<td>24</td>
</tr>
<tr>
<td>FEMALE</td>
<td></td>
<td></td>
<td>24</td>
</tr>
<tr>
<td>TOTAL</td>
<td>35</td>
<td>13</td>
<td>48</td>
</tr>
</tbody>
</table>

Table 4. Actual discrimination study

7. Given a file was recommended for promotion, what percentage was male? Given a file was recommended for promotion, what percentage was female?

8. Without exploring the data any further, do you think there was discrimination by the bank supervisors against the females? How certain are you?

9. Could the smaller number of recommended female applicants for promotion be attributed to variation due to randomness or chance? What is your sense of how likely the smaller number of recommended females could have occurred by chance? Your benchmark for comparing what was expected for the number of females is contained in Table 1.

10. Suppose that the bank supervisors looked at files of actual female and male applicants. Assume that all of the applicants were identical with regard to their qualifications and we use the same results as before (21 males and 14 females). If a lawyer retained by the female applicants hired you as a statistical consultant, how would you consider obtaining evidence to make a decision about whether the observed results were due to expected variation (random process) or if the observed results were due to an effect; i.e., discrimination against the women?
Statisticians would formalize the overriding question by giving two statements, a null hypothesis which represents no discrimination so that any departure from Table 1 is due solely to a chance process, and an alternative hypothesis which represents discrimination against the women.

SIMULATION OF THE DISCRIMINATION CASE

Using a deck of cards, let 24 black cards represent the males, and 24 red cards represent the females (remove 2 red cards and 2 black cards from the deck). This will simulate the 48 folders, half of which were labeled male and the other half female.

1: Shuffle the 48 cards at least 6 or 7 times to insure that the cards counted out are from a random process. You are simulating what can happen with random variation where no discrimination is being practiced.

2: Count out the top 35 cards. These cards represent the applicants recommended for promotion to bank manager. The simulation could be conducted more efficiently by dealing out 13 “not promoted” cards, which would be equivalent to dealing out 35 “promoted” cards.

3: Out of the 35 cards, count the number of black cards (representing the males).

4: On the number line provided, Figure 1, create a dot plot by placing a blackened circle or X above the number of black cards counted. The range of values for possible black cards is 11 to 24.

5: Repeat steps 1 – 4 nineteen more times for a total of 20 simulations.

Before you perform the 20 simulations, what would you expect for the shape of the distribution you create in the dot plot? At what value would you expect the distribution to be centered?
FIGURE 1

DOT PLOT TO BE USED TO GRAPH THE 20 SIMULATED RESULTS

<table>
<thead>
<tr>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
<th>17</th>
<th>18</th>
<th>19</th>
<th>20</th>
<th>21</th>
<th>22</th>
<th>23</th>
<th>24</th>
</tr>
</thead>
</table>

Number of Men Promoted

You have created a sampling distribution that is the distribution of possible values of a statistic for repeated samples of the same size from a population. For the scenario under consideration, the number of black cards (number of males promoted) from each of the simulations is the statistic.

6: Using the results (the counts) plotted on the number line, estimate the likelihood that 21 or more black cards (males) out of 35 will be selected if the selection process is random; that is, if there is no discrimination against the women in the selection process.

The probability of observing 21 or more black cards if the selection process is due to expected variation from a random process is called the **P-value**. Typically a P-value less than 0.05 is considered statistically significance to where a researcher would reject the assumption of the results being due to random variation and concluding there is strong evidence to support that the results indicate discrimination against women.

7: Based on you P-value, what would you conclude? How would you answer your original statistical question?

8. Look at the dot plot and comment on the shape, center, and variability of distribution of the counts by answering the following.
   (a) Is the distribution somewhat symmetric, pulled (skewed) to the right, or pulled to the left?
   (b) Do you observe any unusual observation(s)? Does 21 or higher appear to be unusual observations?
   (c) Where on the dot plot is the lower 50% of the observations?
   (d) Estimate the mean of the distribution representing the number of black cards
obtained out of the 20 simulations.

(e) Comment on the variability of the data.

9. Is the behavior of this distribution what you might expect? Compare your answers to question 8 to your answers to the questions earlier before performing the 20 simulations.