Appendix D

Class Assignments
This page intentionally left blank
Ages of Presidents when they were inaugurated.

57  61  57  57  58  57  61  54  68  51  49  64  50  48  65  
52  56  46  54  49  51  47  55  55  54  42  51  56  55  51  
54  51  60  62  43  55  56  61  52  69  64  46  54  47  70

Step 1: Choose number of classes.

Step 2: \[ \text{Class Width} \approx \frac{\text{highest value} - \text{lowest value}}{\text{number of classes}} \] Round up

Step 3: Choose lower class limit.

Step 4: Use calculator procedures to construct histogram

Xmin = __________ lower class limit

Xmax = __________ largest upper class limit or higher

Xscl = __________ class width

Ymin = -5

Ymax = ________ a little higher than largest frequency

Step 5: Create frequency table

<table>
<thead>
<tr>
<th>Limits</th>
<th>( f )</th>
</tr>
</thead>
</table>

Sum of \( f \)
Ages of Presidents when they were inaugurated.

Complete the table

<table>
<thead>
<tr>
<th>midpts (x)</th>
<th>Limits</th>
<th>f</th>
<th>r f</th>
<th>c f</th>
<th>c r f</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Construct a Frequency Histogram and Frequency Polygon
• Construct a Cumulative Frequency Histogram and Cumulative Frequency Polygon (Ogive)
Ages of Presidents when they were inaugurated.

Analysis:

1. Use the frequency histogram to comment on the center, variation and shape of the data.

Calculations (To be completed after covering section 2.4)

2. Find the following from the raw data.
   a) Mean =
   b) Median =
   c) Mode =
   d) Midrange =
   e) Range =
   f) Standard Deviation =
   g) Variance =

3. Find the following from the frequency table
   a) Mean =
   b) Standard Deviation =

Comment on why there is a difference in mean and standard deviation when using the raw data vs. a frequency table.

4. Use the mean and standard deviation of raw data to find the boundaries for unusual boundaries then determine whether any ages of presidents at the time of inauguration is unusual.
Ages of Presidents when they were inaugurated.

Calculations (To be completed after covering section 2.5)
4. Find the following from the raw data.
   a) $P_{25} =$
   
   b) $P_k$ for President Trump
   
   c) z Score for President Trump

Interpret the z-score
This page intentionally left blank
Class Assignment #2
You have an URN with 3 red marbles, 7 white marbles, and 1 green marble

Let $A = \text{choose red}$
$B = \text{choose white}$
$C = \text{choose green}$

Find the following:

1. $P(A^c)$, that is find $P(\text{not red})$
2. $P(A \text{ or } B)$
3. If two marbles chosen what is the probability that you choose a white marble 2$^{nd}$ when a red marble was chosen first. This is, find $P(B \mid A)$
4. If two marbles are chosen, find $P(A \text{ and } C)$ with replacement
5. If two marbles are chosen, find $P(A \text{ and } C)$ without replacement

Very Similar to test question #1
### Extra Example

#### Test Questions

<table>
<thead>
<tr>
<th></th>
<th>Positive</th>
<th>Negative</th>
<th>Totals</th>
</tr>
</thead>
<tbody>
<tr>
<td>HIV</td>
<td>57</td>
<td>3</td>
<td>60</td>
</tr>
<tr>
<td>Not HIV</td>
<td>997</td>
<td>18,943</td>
<td>19,940</td>
</tr>
<tr>
<td><strong>Totals</strong></td>
<td>1,054</td>
<td>18,946</td>
<td>20,000</td>
</tr>
</tbody>
</table>

Let

- \( A = \text{Positive} \)
- \( B = \text{Negative} \)
- \( C = \text{HIV} \)
- \( D = \text{Not HIV (healthy)} \)

If one chosen (addition rule):

- \( P(A \text{ or } C) \)
- \( P(B \text{ or } C) \)
- \( P(A \text{ or } B) \)

If two are chosen (multiplication rule):

- \( P(\text{Both positive}) \)
- \( P(\text{Both HIV}) \)

Reading Table Directly OR multiplication rule:

- \( P(A \text{ and } C) \)
- \( P(B \text{ and } C) \)
Class Assignment #3 (Part I)
Introduction to Probability Distributions

In chapter 3 we calculated a probability for a single event. For example, if we tossed a coin 4 times we might find the probability of getting 2 heads. Using the counting rules

\[
\begin{align*}
\text{Experiment: Toss a coin 4 times} \\
P(\text{getting 1 head}) &= \frac{\binom{4}{1}}{2^4} = \frac{4}{16} = \frac{1}{4}
\end{align*}
\]

In chapter 4 we will calculate the probabilities for all possible events of a certain experiment and put the results in a table called a probability distribution. This table will look very similar to the frequency distribution we did in chapter 2 but instead of a midpoint we will define a random variable and instead of a frequency we will calculate a probability.

**Exercise:** Let’s define the same experiment as above. In addition let’s define a random variable \(x\). Let \(x = \) number of heads observed. Using the counting rules, fill in the table for all possible values of the random variable then calculate the probabilities for each. I filled in the probability for getting 1 head for you. This was calculated above.

Probability Distribution:

<table>
<thead>
<tr>
<th>Random Variable (x)</th>
<th>Probability P(x)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.25</td>
</tr>
</tbody>
</table>

What is the sum of all possible probabilities? ____________

What would be the shape of the distribution? ______________

Would it be unlikely to get 4 heads? Why or Why not?
Class Assignment #3 (Part II)  
Experiment: Roll 2 dice  
Let random variable $x =$ sum of dice  

$P(X=x) = \frac{\text{Total outcomes favorable to } x}{\text{Total possible outcomes}}$  

<table>
<thead>
<tr>
<th>$x$</th>
<th>$P(x)$ - Fractions</th>
<th>$P(x)$ - Decimals</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Mean = ________  
Questions:  
Variance = ________________  
$\Sigma p(x) = ________$  
Expected Value = ________________  
Minimum Unusual Value = ____________  
P(2 or 12) = ________________  
Maximum Unusual Value = ____________  
P(at least 3) = ________________  
Which values of $x$ are Unusual? = ____________  
Which probabilities are Unlikely? = ____________
A multiple choice test has 6 questions with 4 possible answers, one of which is correct. For all students that guess the answers, respond to the following.

**Use a separate piece of paper.**

W = wrong answer  C = correct answer

a)  \( p = \quad \)  \( q = \quad \)  \( n = \quad \)

b)  Find \( P(WWCCCC) \) using multiplication rule

c)  How many ways are there to get any 4 correct answers?

d)  Use b) and c) above to calculate the probability of getting any 4 answers correct.

e)  Let \( X = \# \) of correct answers. Construct a binomial probability distribution (table).

\[
\begin{array}{c|c}
 x & P(x) \\
\end{array}
\]

f)  Find the probability of getting 0 correct answers

g)  Find the probability of getting at least one correct answer

h)  Would getting 4 answers correct be unlikely?

i)  Calculate the mean, variance and standard deviation for the probability distribution in e) above

1. Using the formulas or calculator (1-var stats) for any probability distribution.
   \[
   \mu = \sum xP(x) \quad ; \quad \sigma^2 = \left( \sum x^2 P(x) \right) - \mu^2 \quad ; \quad \sigma = \sqrt{\sigma^2}
   \]

2. Using the formulas special to a binomial distribution.
   \[
   \mu = np \quad ; \quad \sigma^2 = npq \quad ; \quad \sigma = \sqrt{\sigma^2}
   \]

3. Compare your results from 1 and 2.

j)  What values of the random variable are unusual? Does this validate your conclusion in h) above? Use the \( 2\sigma \) rule.
This page intentionally left blank
1. A company manufactures calculators in batches of 15 and there is a 5% rate of defects. Write out the probability statements.
   a. Define the random variable
   b. Find the probability of getting exactly 1 defect in a batch.
   c. Find the probability of getting 0 or 1 defect in a batch.
   d. Find the probability of getting at least one defect in a batch.
   e. At what point does the probability for the number of defects in a batch become unlikely?
   f. Find the mean and standard deviation of this random variable.
   g. What values (if any) of the random variable are unusual?

2. Given the following probability distribution.

<table>
<thead>
<tr>
<th>x</th>
<th>p(x)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.112</td>
</tr>
<tr>
<td>1</td>
<td>0.234</td>
</tr>
<tr>
<td>2</td>
<td>0.387</td>
</tr>
<tr>
<td>3</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0.162</td>
</tr>
<tr>
<td>---</td>
<td>------</td>
</tr>
</tbody>
</table>

   A) Fill in the missing probability
   B) Find the mean
   C) Find the standard deviation
   D) Find P(3 or more). Write out your probability statement.
   E) Is P(4) unlikely? Why?
   F) Are any values of the random variable unusual? Calculate the boundaries for unusual values.
This page intentionally left blank
CLASS ASSIGNMENT #6
Both are very similar to test questions.

1. The weekly salaries of teachers in one state are normally distributed with a mean of $500 and a standard deviation of $50.
   a. Find the probability a randomly selected teacher earns more than $525. Write out the probability statement and show a sketch.
   b. In a district of 2000 teachers, how many would you expect to earn more than $525?
   c. Find the salary that separates the top 25% from the bottom 75%, that is find Q3. Show a sketch.

2. A study of the amount of time it takes to repair a mechanical device shows that the mean is 6 hours and the standard deviation is 2 hours.

   A) If one mechanic is randomly selected, find the probability that his repair time exceeds 6.8 hours. Assume normally distributed.
   B) If 64 mechanics are randomly selected, find the probability that their mean repair exceeds 6.8 hours.
   C) Based on the result from part B, if the mechanics take 6.8 hours does it appear that the mechanics are slower, faster or about average? Explain. This is a little tricky.
This page intentionally left blank
Class Assignment #7

1. Use OLDFAITHFUL Data in Datasets File
2. Construct a 95% and 90% confidence interval for the mean eruption duration. Write a conclusion for the 95% interval. Assume $\sigma$ to be 58 seconds
3. Compare the 2 confidence intervals. What can you conclude?
4. How large a sample must you choose to be 99% confident the sample mean eruption duration is within 10 seconds of the true mean

Guidelines:
1. Choose a partner
2. Suggest having one person working the calculator and one writing
3. Due at the end of class (5 HW points)
4. Each person must turn in a paper
### Class Assignment #7 Data

<table>
<thead>
<tr>
<th>OLDFAITHFUL</th>
<th>ERUPTION DURATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>240</td>
<td>243</td>
</tr>
<tr>
<td>237</td>
<td>241</td>
</tr>
<tr>
<td>122</td>
<td>214</td>
</tr>
<tr>
<td>267</td>
<td>114</td>
</tr>
<tr>
<td>113</td>
<td>272</td>
</tr>
<tr>
<td>258</td>
<td>227</td>
</tr>
<tr>
<td>232</td>
<td>237</td>
</tr>
<tr>
<td>105</td>
<td>238</td>
</tr>
<tr>
<td>276</td>
<td>203</td>
</tr>
<tr>
<td>248</td>
<td>270</td>
</tr>
</tbody>
</table>

n = 50
1. The health of employees is monitored by periodically weighing them in. A sample of 54 employees has a mean weight of 183.9 lb. Assuming that $\sigma$ is known to be 121.2 lb, use a 0.10 significance level to test the claim that the population mean of all such employees weights is less than 200.

\[
\text{TEST OF } \mu = 200 \text{ VS } \mu < 200
\]

<table>
<thead>
<tr>
<th>N</th>
<th>MEAN</th>
<th>STDEV</th>
<th>SE MEAN</th>
<th>Z*</th>
<th>P VALUE</th>
</tr>
</thead>
<tbody>
<tr>
<td>54</td>
<td>183.9</td>
<td>121.2</td>
<td>??????</td>
<td>?</td>
<td>.1645</td>
</tr>
</tbody>
</table>

SE MEAN = _____________ Z* = ________________

Conclusion (justify)

2. A cereal company claims that the mean weight of the cereal in its packets is at least 14 oz. The weights (in ounces) of the cereal in a random sample of 8 of its cereal packets are listed below.

14.6 13.8 14.1 13.7 14.0 14.4 13.6 14.2

At the 0.01 significance level, test the claim that the mean weight is at least 14 ounces. First test for normality. Why is a two-tailed test best for this analysis?
This page intentionally left blank