Practice Exam 1 - Answers

1. 

\[
\begin{array}{ccc}
\text{(D+)} & \text{(D-)} \\
\text{(T+)} & \text{(T-)} \\
.080 & .020 & .135 & .765 \\
.8 / & \text{.2} & .15 / & \text{.85} \\
\end{array}
\]

\(1b) \quad .080 + .135 = .215\)

\(1c) \quad .080/.215 = .372\)

2. The data shown in the scatter plot is the distance traveled and the airfare for 12 flights on Delta Airlines:
   a. Which of the following is a reasonable estimate of the correlation coefficient? (Circle one answer)
      \(1.0 \quad (0.8) \quad 0.0 \quad -0.5\)
   b. What does this graph tell us about distance and airfare
      \text{As distance goes up price goes up.}
   c. What is the type and level of distance traveled?
      \text{Type (Circle One) Categorical Discrete (Continuous)}
      \text{Level (Circle One) Nominal Ordinal Interval (Ratio)}

3. You have a 70% chance of being on time to class today and a 80% chance of being on time to class tomorrow. Assume these two days are independent events.
   a. Find the probability of being on time to class both today \textbf{and} tomorrow.
      \(0.7 \times 0.8 = 0.56\)
   b. Find the probability of being on time to class at least once today \textbf{or} tomorrow.
      \(0.7 + 0.8 - 0.56 = 0.94\)

4. The following data represent the daily births at a hospital for 20 days

\[
\begin{array}{ccccccccccccccc}
14 & 15 & 15 & 17 & 17 & 19 & 19 & 21 & 23 & 25 \\
26 & 27 & 31 & 35 & 36 & 47 & 48 & 59 & 70 & 99 \\
\end{array}
\]

a) Construct a stem and leaf diagram of the data

\[
\begin{array}{c}
1 | 4 5 5 7 7 9 9 \\
2 | 1 3 5 6 7 \\
3 | 1 5 6 \\
4 | 7 8 \\
5 | 9 \\
6 | \\
7 | 0 \\
8 | \\
9 | 9 \\
\end{array}
\]

b) Calculate the interquartile range for this data set.
   \(Q1 = 18, \quad Q3 = 41.5 \quad IQR = 41.5 - 18 = 23.5\)

c) Calculate the median for this data set.
   \text{Median} = 25.5

d) \textbf{Without calculating}, what can you say about the mean births for this Hospital.\footnote{\textbf{check one answer below}}
   \[\checkmark\text{ The mean is greater than the median. Data is skewed right}\]
   \[\text{x} \text{ The mean is less than the median.}\]
   \[\xmark\text{ The mean is about the same as the median.}\]
5. The following data represents the hours per week worked outside of school by 200 randomly selected night students at a community college:

<table>
<thead>
<tr>
<th>Hours</th>
<th>Frequency</th>
<th>Relative Freq</th>
<th>C.R.Freq</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-8</td>
<td>20</td>
<td>0.10</td>
<td>0.10</td>
</tr>
<tr>
<td>9-16</td>
<td>30</td>
<td>0.15</td>
<td>0.25</td>
</tr>
<tr>
<td>17-24</td>
<td>50</td>
<td>0.25</td>
<td>0.50</td>
</tr>
<tr>
<td>25-32</td>
<td>60</td>
<td>0.30</td>
<td>0.80</td>
</tr>
<tr>
<td>33-40</td>
<td>30</td>
<td>0.15</td>
<td>0.95</td>
</tr>
<tr>
<td>41-48</td>
<td>10</td>
<td>0.05</td>
<td>1.00</td>
</tr>
<tr>
<td>Total</td>
<td>200</td>
<td>1.00</td>
<td></td>
</tr>
</tbody>
</table>

a) In the space above, determine the relative frequencies and cumulative relative frequencies.

b) Sketch a relative frequency histogram, showing all horizontal and vertical labels.

c) Sketch a cumulative relative frequency ogive, showing all horizontal and vertical labels.

d) Estimate the median from the graph.
**median = 24.5**

e) What percentage of the night students work 32 hours per week or less?
**0.80 (crf)**

Without calculating but explaining your reasoning, which of the following is a reasonable estimate for the standard deviation?  

a) 0.5  b) 1  c) 10  d) 50

10 is the only answer that makes sense since the range is 48 and the range is between 4s and 6s.

6. Determine if each of the following data are categorical, continuous or discrete (circle one for each)

a. Number of fatalities from a tsunami: categorical
   b. Time spent in traffic: continuous
   c. Number of Songs on your I-pod: categorical
   d. Your student number: discrete
   e. Names of cities in California with a Walmart: categorical
   f. Price per gallon of gasoline: continuous
   g. Number of Courses taken in a year: continuous
   h. Tons of steel used by a manufacturer: categorical
7. 1000 students (500 morning, 300 afternoon, 200 night) were asked how often they use the campus library. The results are summarized in the table below:

<table>
<thead>
<tr>
<th></th>
<th>Never uses library</th>
<th>Sometimes uses library</th>
<th>Frequently uses library</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Morning</td>
<td>200</td>
<td>250</td>
<td>150</td>
<td>600</td>
</tr>
<tr>
<td>Afternoon</td>
<td>80</td>
<td>145</td>
<td>75</td>
<td>300</td>
</tr>
<tr>
<td>Night</td>
<td>80</td>
<td>10</td>
<td>10</td>
<td>100</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>360</strong></td>
<td><strong>405</strong></td>
<td><strong>235</strong></td>
<td><strong>1000</strong></td>
</tr>
</tbody>
</table>

a. Find the following probabilities:
   i) A randomly selected student never uses the library.
      \[\frac{360}{1000} = 0.360\]
   ii) A randomly selected student is a night student and frequently uses the library.
      \[\frac{10}{100} = 0.10\]
   iii) Given the student is an afternoon student, the student never uses the library.
      \[\frac{80}{300} = 0.267\]

b. Are “Afternoon Student” and “Never uses library” Independent Events? **Justify and explain your answer.**
   No. \(P(\text{Never}) \neq P(\text{Never}|\text{Afternoon})\) (0.360 \(\neq 0.267\))

c. Would the probabilities generated from this data be classical, empirical or subjective probability?
   **Empirical – based on data**

8. These descriptive statistics and boxplots were generated from data representing **calories per serving** for three types of hotdogs: All Beef, Mixed Meat and Poultry.

a. Compare the mean to the median calories for the Meat group. Is the result consistent with the shape of the box plot? Explain your answer.
   **Mean > median, but it is difficult to read the skewness from graph so its unclear.**

b. If the data is approximately bell shaped, between what two values of calories would you expect to find about 95% of the Beef data?
   \((111.57, 202.13)\)

c. Which of the three groups has the most variability in calories per serving? Explain your answer.
   **Meat – highest Standard deviation**

d. Hebrew National All Beef Hotdogs had 190 calories per serving. Calculate and interpret the z-score for Hebrew National Hotdogs using the Beef Category data.
   \([Z = 1.46\] Hebrew national calories are above average.\]

e. Determine the probability a randomly selected Poultry Hot Dog exceeds 113 calories.
   **50% (half the data is above the median)**

f. Compare the three groups and draw at least two conclusions from the results.
   **Chicken dogs are lower in calories. Meat and Beef are about the same. (other answers ok).**
9. 70% of students at a large New York University receive some financial aid. (use binomial table n=4, p=.7)
   a. If 4 students are randomly selected, determine the probability that exactly 2 students in the sample receive some financial aid.  \( P(X=2) = .265 \)
   b. If 4 students are randomly selected, determine the probability that less than 2 students in the sample receive some financial aid.  \( P(X<2) = P(0) + P(1) = .084 \)

10. Serious accidents occur at an oil refinery at a Poisson rate of 1 per year.
    a. Find the probability of exactly 2 serious accidents occurring in the next year.  0.1839
    b. Find the probability of at least 1 serious accident occurring in the next 3 years.  \( 1 - 0.0498 = 0.9502 \) (use \( \mu=3 \))

11. The random variable X follows the probability distribution function as shown to the right:
    a. Determine \( P(X=3) \)  \( P(3)=0.3 \)
    b. Determine the population mean.  \( \mu=2 \)
    c. Determine the population variance  \( \sigma^2=1.4 \)

<table>
<thead>
<tr>
<th>x</th>
<th>P(x)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.1</td>
</tr>
<tr>
<td>1</td>
<td>0.3</td>
</tr>
<tr>
<td>2</td>
<td>0.2</td>
</tr>
<tr>
<td>3</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0.1</td>
</tr>
</tbody>
</table>

12. 40% of students at a college use the cafeteria.
    a. If 9 students are randomly sampled, determine the probability that less than 3 use the cafeteria.  0.231
    b. If 9 students are randomly sampled and X represents the number of students in the sample who use the cafeteria, find the mean and standard deviation of X.
       \( \mu=3.6 \) students     \( \sigma=1.469 \)

13. Accidents in a power plant occur at a Poisson rate of 1.39 per year.
    a. Find the probability of at least 2 accidents occurring at the plant in the next year.
       \( 1 - 0.24908 - 0.34621 = 0.40471 \)
    b. Find the probability that the plant has zero accidents in two years
       \( 0.24908^2 = 0.06204 \)

14. Find the 30th percentile for each of the following populations:
    a. A Normal Distribution with a mean of 4 and a standard deviation of 3
       \( 4 - 0.52(3) = 2.44 \)
    b. A Uniform Distribution with a minimum of 3 and a maximum of 23
       \( 3 + 0.3(23-3) = 9 \)
    c. An Exponential Distribution with a mean of 10
       \( -10\ln(1-.3) = 3.567 \)

15. Students’ exam scores for a course follow a Normal Distribution with \( \mu=70 \) and \( \sigma=10 \).
    a. Find the probability a randomly selected student scores a 75 or more.
P(X > 75) = P[Z > (75-70)/10 ] = P(Z > 0.5) = 0.3085

b. Find the exam score which is the 25th percentile of this distribution.: 70 – 0.67(10) = 63.3

c. You take a random sample of 40 students. Find the probability the sample mean is between 68 and 72.
\[
P(68 < \bar{X} < 72) = P\left(\frac{68 - 70}{10\sqrt{40}} < \frac{Z}{10}\sqrt{40} < \frac{72 - 70}{10\sqrt{40}}\right) = P(-1.26 < Z < 1.26) = 0.7924
\]
d. Would your answer for part c be different if the probability distribution of “exam scores” did not follow a Normal distribution? Explain your answer.
No, because of the CLT.

16. The waiting time until a subway train arrives follows a uniform distribution where the minimum wait is 0 minutes and the maximum wait is 10 minutes.

a. Find the probability of waiting less than 3 minutes for the next train. P(X<3) = (3-0)/(10-0) = 0.3

b. Find the median waiting time. Med = 5 minutes

c. Find the probability of waiting exactly 2 minutes for a train. 0, makes no sense

d. Given the passenger has already waited 2 minutes, find the probability of waiting a total of 5 or more minutes for the train. P(X>5|X>2) = (10-5)/(10-2) = 0.625

17. The age of a grove of walnut trees follow a Normal Distribution with \( \mu=50 \text{ years} \) and \( \sigma=15 \text{ years} \).

a. Find the probability that the age of a randomly selected tree is between 40 and 70 years.
\[
P(40 < X < 70) = P\left(\frac{40 - 50}{15} < Z < \frac{70 - 50}{15}\right) = P(-0.67 < Z < 1.33) = 0.6568
\]
b. Find the probability of a randomly selected tree has lived exactly 45.231789 years. 0, makes no sense

c. Find the 30th percentile of this distribution. 50 - 0.52(15) = 42.2 years

18. The waiting time for a train follows a uniform distribution where the minimum time is 10 minutes and the maximum time is 50 minutes. For this distribution \( \mu = 30 \text{ minutes} \) and \( \sigma = 14.14 \).

a. Find the probability the wait for a randomly selected passenger is between 25 and 37 minutes. (37-25)/(50-10) = 0.3

b. Find the 55th percentile of this distribution. 10 + 0.55(50-10) = 32 minutes

c. A random sample of 36 passengers is taken. Find the probability the sample mean exceeds 33 minutes. Use CLT
\[
P(X > 33) = P\left(Z > \frac{33 - 30}{14.14}/\sqrt{36}\right) = P(Z > 1.27) = 0.1020
\]

19. Accidents occur at an oil refinery on the average once every 3 years. The waiting time until the next accident follows an Exponential Distribution.

a. Find the probability of waiting at least 5 years until the accident. \( P(X>5) = e^{-(5/3)} = 0.1889 \)

b. Find the Median waiting time until the next accident. -3ln(0.5) = 2.079 years

20. The random variable Y represents the daily car sales at an auto dealership. The population mean is 49 sales per day and the standard deviation is 7 sales per day. The probability distribution shape of Y is Poisson and a sample of 100 days is randomly selected.
\[\mu_Y = 49 \quad \sigma_Y = 0.7 \quad \text{Distribution is Normal (Three parts of the Central Limit Theorem)}\]