

# The Game of

P	E	R	D	I
P	E	R	D	I
P	E	R	D	I
P	E	R	D	I
P	E	R	D	I

Preparing for the A.P. Statistics Exam with Problems in

**Probability**

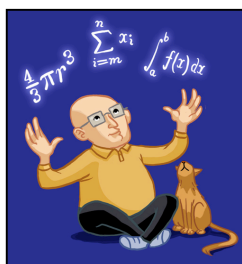
**Experimental Design**

**Regression**

**Descriptive Stats**

**Inference**

**Version 1**



# Acknowledgments

Graphics in this manual were produced from:

1. Texas Instruments TI-Smartview Emulator Software for the TI-84.  
©Texas Instruments, Inc.
2. InTaglio <sup>TM</sup> software. Intaglio@PurgatoryDesign.com.
3. Equation Editor Software for the Mac. © Design Science Corp.
4. Grapher Software: © Apple Computer Corp.
5. Fathom 2.0 Dynamic Data Software: © Key Curriculum Press.

Thanks to my partner Ted Tyree for his unwavering support in all of my projects and his monumental work in bringing my materials free to anyone who wants them through the Internet.

Full appreciation to Tim Greer for doing every problem checking for errors and looking at every possible way a student could misconstrue its meaning.

If you do find errors or have any questions or comments, please direct them to: [team@mastermathmentor.com](mailto:team@mastermathmentor.com).

# P – 1 – Version 1

80% of people buy PC's and the rest buying Macs. 70% of PC users buy laptops and the rest of PC users buy desktop machines. 60% of Mac users buy laptops and the rest of Mac users buy desktops. If a person buys a laptop, what is the probability that it is a Mac?

- A) .120
- B) .176
- C) .200
- D) .214
- E) .400

# P – 2 – Version 1

Suppose that the probability that a person over 50 years old chosen at random has high blood pressure is .39 and the probability that a person over 50 years old has high blood pressure and is overweight is .28. Given that a person over 50 chosen at random has high blood pressure, what is the probability that the person is also overweight?

- A) .109
- B) .280
- C) .670
- D) .718
- E) Cannot determine because having high blood pressure and being overweight are not independent.

# P – 3 – Version 1

A survey was done among people who live in the same community as to whether they pay their bills on time. Based on the table below, what value of  $n$  (nearest integer) results in people paying bills on time being independent of gender?

	Men	Women
Pay on time	$n$	48
Pay late	25	8

- A) 4
- B) 15
- C) 120
- D) 150
- E) 384

# P – 4 – Version 1

Following are parts of the probability distributions for the random variables  $x$  and  $y$ .

$x$	$P(x)$	$y$	$P(y)$
1	.2	1	.1
2	.5	2	?
3	?	3	?
		4	.3

If  $x$  and  $y$  are independent, and the probability that  $(x = 2 \text{ and } y = 2) = .2$ , find the probability that  $(x = 3 \text{ and } y = 3)$ .

- A) .06
- B) .12
- C) .2
- D) .24
- E) .25

# P – 5 – Version 1

The Apple Store sells four different types of Ipods. People purchasing the largest capacity Ipods from the store purchase the Ipods in the following proportions with the given prices (nearest \$10):

Type	Touch	Nano	Shuffle	Classic
Price	\$400	\$180	\$80	\$250
% Buying	45%	30%	20%	5%

Find the mean and standard deviation of iPod purchases from the Apple Store.

- A) \$227.50, \$232.97
- B) \$227.50, \$135.81
- C) \$262.50, \$131.22
- D) \$262.50, \$243.26
- E) \$262.50, \$121.63

# E – 6 – Version 1

Pasquale's Frozen Pizza develops a new type of better tasting pizza and wants to check out whether people like it so they design this taste-testing procedure:

The pizza will be available in both thin and thick crust. They are concerned that children and adults differ in their opinions, so both children and adults are divided into two groups, each getting a different type of pizza. This is an example of:

- A) an Observational Study
- B) a Voluntary Response Sample
- C) a Stratified Random Sample
- D) a Multistage Sample
- E) Blocking



# E – 7 – Version 1

A simple random sample of 20 people is taken from a population of 400 people to participate in a tasting experiment. The people who serve the food are not told how these 20 people are chosen. Which of the following is *false*?

- A) The sample is large enough to be considered an SRS.
- B) Looking at the people in the sample gives no indication whether or not it is an SRS.
- C) Attributes of this sample will be very similar to attributes of the population.
- D) The first 20 people alphabetically in the population have an equal chance of being chosen as any other group of 20 in the population.
- E) All of the above are true.

# E – 8 – Version 1

A school has a population that has 800 African-American students and 200 White Students. A study is being done on student perception of the education provided in this school and 80 students are chosen to be in the study, 40 African-Americans and 40 White students. Which term below best describes their sampling technique?

- A) Confounding
- B) Nonresponse
- C) Voluntary Response
- D) Biased
- E) None of the above

# E – 9 – Version 1

“Most Optimal Purchase”, an electronics store, has approximately 125 workers in each of its 1,110 stores. They plan to do an in-depth study on employee reactions to company policies. To accomplish this, they ask each store manager to randomly choose 3 men and 3 women employees to participate in the study and send those names to the corporate office where the names will be compiled into one big list. Which of the following is true about this type of sampling?

- I. While not an example of choosing a simple random sample, it ultimately provides the same information that an SRS would provide.
- II. It is cheaper and easier to choose by this method than choosing an SRS.
- III. This is an example of Multistage sampling.

- A) I only
- B) II only
- C) III only
- D) I and II
- E) I II, and III

# E – 10 – Version 1

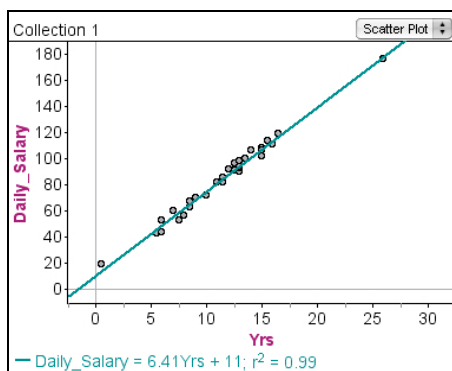
A nursing home is interested in improving the morale of its 80 patients. It uses a questionnaire to determine the average morale of its residents on a 1 – 10 basis, 1 being low, 10 being high. It then brings in dogs and cats for an hour a day for a week and allows the residents to interact with them. At the end of that time, it gives out the questionnaires again to each resident. For each resident, the difference in morale score is computed and then averaged for the entire nursing home. A difference of 3.82 with a standard deviation of 1.68 was computed.

Which of the following statements *cannot* be made based on the above?

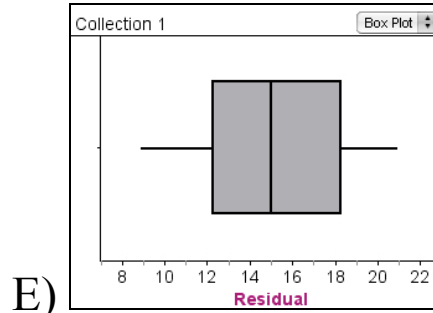
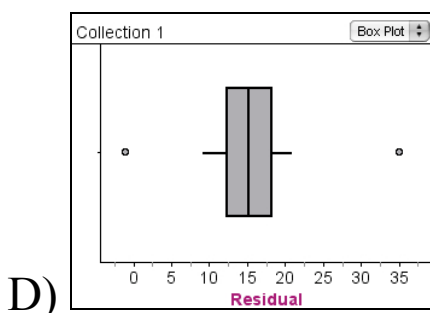
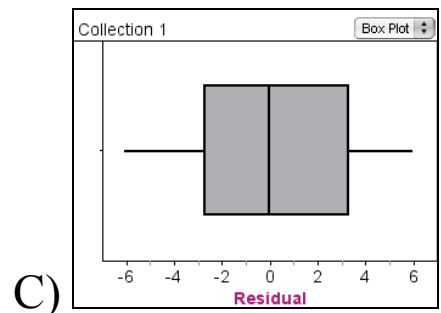
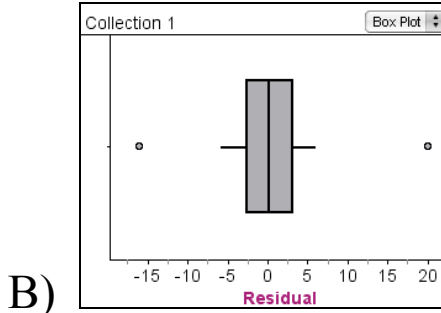
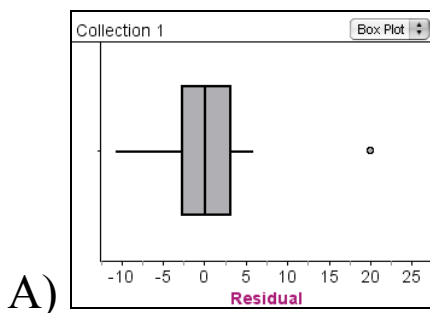
- A) Morale significantly increased.
- B) Bringing in pets increases the morale of the residents.
- C) There is one explanatory variable and one response variable.
- D) This is an example of a matched-pairs experiment.
- E) All statements can be made.

# R – 11 – Version 1

In a clothing store, employees are paid daily. The amount they are paid is based on the number of years they have been employed there and other factors such as punctuality, how much they sell, etc. Below is a scatterplot of the data (years employed vs. daily salary) along with the Least-Squares Regression Line for this data.



If the residuals were computed, which of the following would be a boxplot of the residual data?



# R – 12 – Version 1

Suppose the correlation between two variables  $x$  and  $y$  is  $r = -.24$ . If each  $x$  variable is doubled and each  $y$  variable is increased by  $.1$ , and the two variables are then interchanged, what is the new correlation?

- A) .24
- B) -.24
- C) -.48
- D) -.38
- E) .58

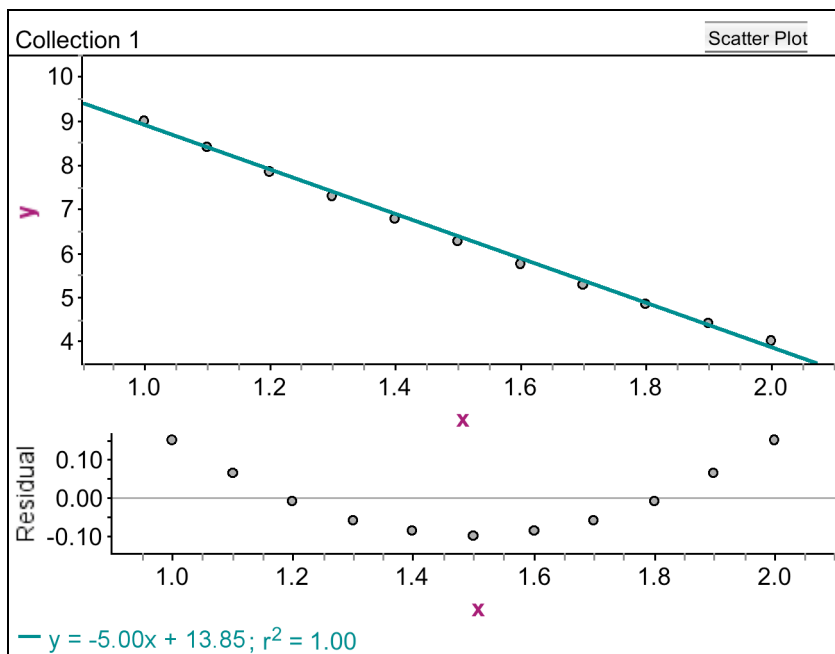
# R – 13 – Version 1

On a project where  $x$  is defined as the age of a particular type of car in years and  $y$  is the Kelly Blue Book value of the vehicle, it is found that the standard deviation of  $x$  is 3.5, the standard deviation of  $y$  is 4,650 and  $r = -.91$ . How can this be interpreted?

- A) For every year, the value of the car decreased by \$4,232.
- B) For every year, the value of the car decreased by \$1,460.
- C) For every year, the value of the car decreased by \$1,329.
- D) For every year, the value of the car decreased by \$1,333.
- E) For every year, the value of the car decreased by \$1,209.

# R – 14 – Version 1

Two variables  $x$  and  $y$  are related. Below are the data points along with the least-square regression line. A residual plot is also given.



Which of the following statements are true?

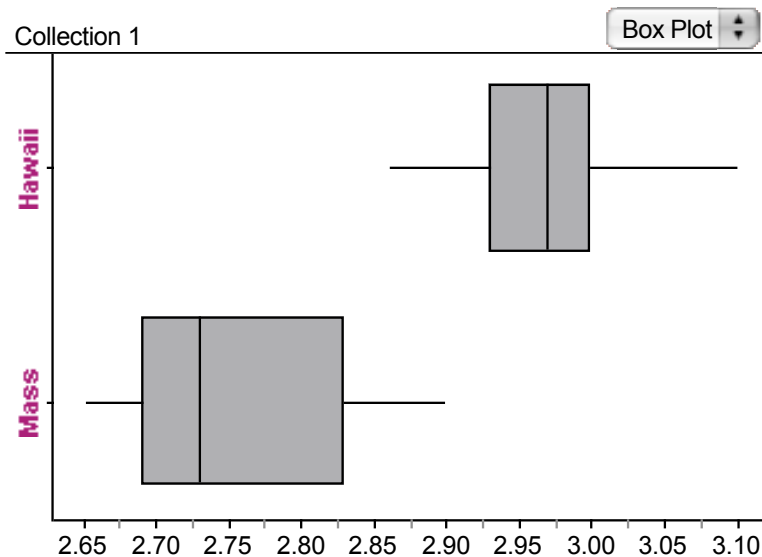
- I. There is a strong negative linear association between  $x$  and  $y$  which can be used to predict  $y$  when  $x = 10$ .
- II. The sum of the residuals will approximately equal zero.
- III. There exists a strong non-linear association between  $x$  and  $y$ .

- A) I only
- B) II only
- C) III only
- D) I and II
- E) II and III



# D – 15 – Version 1

The prices of gasoline was surveyed from a number of service stations in the states of Massachusetts and Hawaii. Their boxplots are shown below. Which of the following are necessarily true statements?



- I. There were more stations surveyed in Massachusetts than Hawaii.
- II. The range of both distributions are approximately the same.
- III. The standard deviation of both distributions are approximately the same.

- A) I only
- B) II only
- C) III only
- D) II and III only
- E) I, II and III

# D – 16 – Version 1

Two stores, Dan's Discount Store and Bill's Bargain Store sell some of the same items, none of which costs over \$10. An item whose price is in the 30<sup>th</sup> percentile for Dan's is at the 60<sup>th</sup> percentile for Bill's. Which of the following is the most reasonable conclusion?

- A) Items generally cost more at Dan's than at Bill's.
- B) Items generally cost more at Bill's than at Dan's.
- C) Bill carries twice as many items than Dan.
- D) Dan carries twice as many items than Bill.
- E) An item whose price is in the 40<sup>th</sup> percentile for Dan's is at the 80<sup>th</sup> percentile for Bill's.

# D – 17 – Version 1

It is said that people lose taste over time. A random sample of 50 people at age 30 were asked to taste chocolate cake and rate it from a scale of 1 to 100. Another random sample of 50 people at age 50 were asked to rate the same cake. The results are given in this back-to-back stemplot. What is the trend with regard to the mean and median of the two samples over time?

Age 30		Age 50
62	0	
87544	1	
96410	2	2
55321	3	2478
85432220	4	0111233333444455555668899
987765544332110	5	0013457789
98862	6	22679
873	7	237
8	8	0
1	9	8

- A) Both the mean and median decreased.
- B) Both the mean and the median increased.
- C) The mean increased while the median decreased.
- D) The mean decreased while the median increased.
- E) The mean remained the same while the median increased.

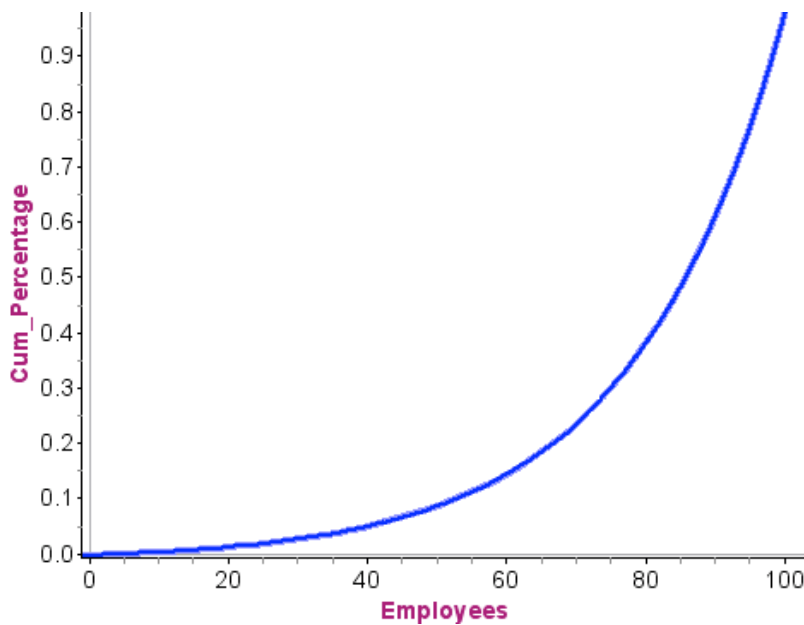
# D – 18 – Version 1

The income of 20,000 families living in Richland County has a standard deviation of \$4,200. What is the average income of families in Richland to the nearest \$10 if 25% of the people make more than \$50,000? Assume a normal distribution for salaries.

- A) \$45,800
- B) \$47,170
- C) \$48,950
- D) \$49,970
- E) \$52,830

# D – 19 – Version 1

A statistician does a survey of all supermarkets in a city to find the number of employees who are working at each market. The graph below illustrates the cumulative percentage plotted against the number of employees. Based on the graph, markets having below approximately what number of employees would be considered outliers?



- A) 50
- B) 39
- C) 25
- D) 12
- E) Impossible to determine

# I – 20 – Version 1

The number of hours to put a new roof on a particular type of house has a standard deviation of 3.57 hours. A sample of 25 houses shows an average roofing time of 10.5 hours. With what confidence can we assert that the average roofing time for this type of house is between 10 and 12 hours?

- A) 99.9%
- B) 95.0%
- C) 74.0%
- D) 71.4%
- E) 21.9%

# I – 21 – Version 1

In a simple random sample, 87 of 93 said they were happy with their Ipod Touch MP3 player while 48 of 62 said they were happy with their Zune MP3 player. What is a 95% confidence interval estimate for the difference between IPod and Zune owners who are satisfied with their MP3 player?

- A)  $.161 \pm .003$
- B)  $.161 \pm .006$
- C)  $.161 \pm .059$
- D)  $.161 \pm .116$
- E) Not enough information is given

# I – 22 – Version 1

It is a late Autumn day with changing weather, the car is not available, and you have to walk 2 miles to school. You are not sure whether to wear a heavy jacket. You make a judgment on the null hypothesis: the weather will be warm enough that you don't need the jacket. What are the ramifications of Type I and Type II errors?

- A) Type I error: Be cold  
Type II error: Needlessly carry around a heavy jacket
- B) Type I error: Needlessly carry around a heavy jacket  
Type II error: Be cold
- C) Type I error: Take the heavy jacket and it gets cold  
Type II error: Don't take the heavy jacket and it stays warm
- D) Type I error: Be cold  
Type II error: Don't carry the heavy jacket and it stays warm
- E) Type I error: Don't carry the heavy jacket and it stays warm  
Type II error: Be cold



# I – 23 – Version 1

A company that sells computers frequently has users call up for phone support. The company wants to make a claim about the average wait on the phone so it takes a sample of  $n$  callers, determines the average wait, and then creates a confidence interval. Which of the following would give the smallest confidence interval?

- A)  $n = 100$  and 90% confidence
- B)  $n = 100$  and 95% confidence
- C)  $n = 200$  and 90% confidence
- D)  $n = 200$  and 95% confidence
- E) depends on the standard deviation of the waiting times

# I – 24 – Version 1

A company claims that its brand of ink-jet printer can print out 75 four by six photos before it needs a new ink cartridge. A consumer's research company believes this is too high and does a study. It uses 20 of the same type printer and finds that the average number of photos that can be printed with one ink cartridge is 73.4 with a standard deviation of 3.3. What would be the conclusion of the research at the 95% confidence level?

- A) The printer company's claim should be rejected because the sample mean of 73.4 is less than the claimed value of 75 and using 20 printers is sufficient in the test.
- B) The printer company's claim should not be rejected because the  $p$ -value of .022 is too small.
- C) There is not sufficient evidence to reject the printer company's claim. 73.4 is too close to the claimed value of 75.
- D) The  $t$ -value of -2.168 with degrees of freedom 19 is sufficient evidence to reject the printer company's claim.
- E) The  $p$ -value of .022 with degrees of freedom 19 is sufficient evidence to prove that the printer company claim is false.