1 (8 Points)

A researcher wants to learn whether regularly taking zinc supplements may reduce the risk of getting a cold. She assigns individuals to two treatment groups, one receiving a placebo, the other the zinc supplement. The two groups are compared with regards to the proportion of individuals who got a cold.

What type of data collection method is this?

[1] sample survey

[2] census

[3] comparative randomized experiment

[4] observational study

[0] no answer or skip this item

Submit Answer

2 (8 Points)

An example of a cluster sample is:

[1] From a list containing all students at all of the campuses in the Utah State University system a sample is drawn in the following manner: A student is chosen from the first 125 on the list, then every 125th student from that point forward is selected.

[2] We have a list containing all classes at Utah State University. Classes are divided into the times when they are offered. A simple random sample of the times is chosen and all students in the class occurring at that time are administered a survey.

[3] From a list containing all students at the main campus of Utah State University a sample is drawn in the following manner: Three-hundred twenty students are randomly chosen from the list.

[4] We have a list containing all students at the main campus of Utah State University. The list is divided into first year, second year, third year, etc. students. Thirty-five students are randomly selected from each year.

[0] no answer or skip this item

Submit Answer

3 (8 Points)

What are the benefits of randomization?

[1] Doing a randomized experiment reduces the risk of confounding.

[2] Random assignment to treatment groups makes it likely that the groups will be similar with regard to important extraneous factors.

[3] Both 1 and 2 above.

[4] None of the above.

[0] no answer or skip this item

Submit Answer

4 (8 Points)

A study is done to compare the extent of heart disease in people who drink 1 to 2 alcoholic drinks per day to the extent of heart disease in non- drinkers. The researcher is able to study 200 individuals of each type.

Other factors that might affect the extent of heart disease are smoking habits and exercise habits. The smoking habits of the two groups of people are similar, but those who drank generally exercised less than the non-drinkers.

In this study, the response variable is:

[1] exercise

[2] heart disease

[3] smoking

[4] drinking status

[0] no answer or skip this item

Submit Answer

5 (8 Points)

The National Personal Transportation Survey is conducted every few years. One question of interest is the percentage of households owning more than one automobile. For those households included in the NPTS, it was found that 78.3 percent of the households owned more than one automobile.

The 78.3 percent of the households owning more than one automobile represents ...?

[1] A population parameter, since they represent numbers derived from the population and not the sample.

[2] A population statistic, since they represent numbers derived from the population and not the sample.

[3] A sample statistic, since they represent numbers derived from the sample and not the population.

[4] A sample parameter, since they represent numbers derived from the sample and not the population.

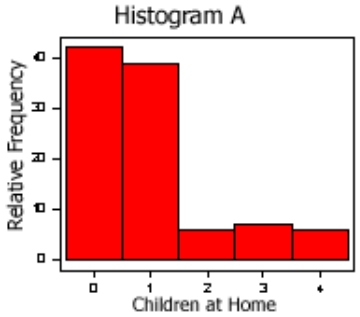
[0] no answer or skip this item

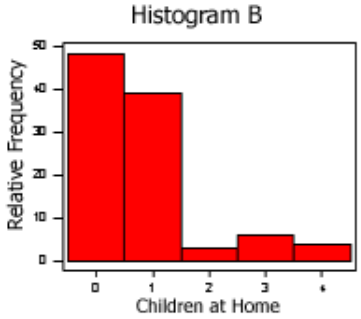
Submit Answer

6 (8 Points)

The following table is based on the numbers of children living at home, for 100 statistics students:

# Children	Number of Students
0	48
1	39
2	3
3	6
4	4





Which histogram corresponds to the data shown in the table?

[1] Histogram A
[2] Histogram B
[3] Both - they are identical
[4] None of them
[0] no answer or skip this item

Submit Answer

7 (8 Points)

Here are the tuitions paid by nonresidents (i.e., students from out of state) in thousands of dollars, arranged in order from smallest to largest for the Big Ten Universities:

7.6 8.6 8.7 9.1 9.3 9.6 10.2 10.5 10.7 15.9 16.4

The median of these 11 numbers is closest to:

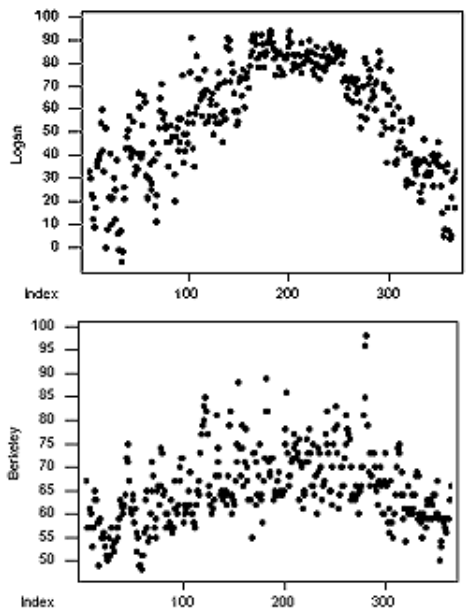
[1] 9.3
[2] 9.45
[3] 9.6
[4] 9.9
[0] no answer or skip this item

Submit Answer

8 (8 Points)

Here are scatter plots for Berkeley, California, and Logan, Utah, in 1996. For both, X = day of the year and Y = high temperature for that day (in degrees Fahrenheit).

Consider the spread in high temperatures for the two locations at the start of the year only, i.e., from day 1 to day 50.



[1] During this time period, Berkeley had the greater spread.
[2] During this time period, Logan had the greater spread.
[3] During this time period, both had the same spread.
[4] There is not enough information to answer the question.
[0] no answer or skip this item

Submit Answer

9 (8 Points)

A family has three children. (G = girl, B = boy)

If we consider gender of the children, what is the sample space for the gender of the first three children?

[1] The children.
[2] {G,B}
[3] {G,B,G}
[4] {GGG,GGB,GBG,GBB,BGG,BGB,BBG,BBB}
[0] no answer or skip this item

Submit Answer

10 (8 Points)

Use WebStat. Load from "Data > Sample data" the data set SMSA_data-59metro_areas.dat and compute the correlation coefficient for the variables "Humidity" and "Rain".

The correlation is closest to

[1] 0.9
[2] 0.5
[3] 0.0
[4] -0.1
[5] -0.5
[0] no answer or skip this item

Submit Answer

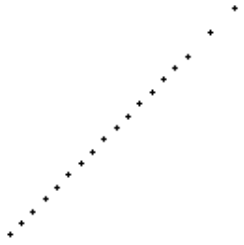
11 (8 Points)

Next are three plots based on the heights and weights of 126 college students. There is one plot each with

X = Weight in pounds and Y = Height in inches for the 126 people (Plot A);

X = Age in years and Y = Average Weight in pounds for everyone of Age X (Plot B);

X = Weight in pounds and Y = Weight in kilograms for the 126 people (Plot C).

Plot A	Plot B	Plot C
		

Which plot's correlation coefficient is exactly equal to 1.0?

[1] Plot A
[2] Plot B
[3] Plot C
[4] All 3 plots
[0] no answer or skip this item

Submit Answer

12 (8 Points)

280 manufactured items are randomly selected from a production run and 15 of them are defective.

The best guess for the probability that a randomly chosen item from this production run is defective is:

[1] about 0.6666
[2] about 0.5357
[3] about 0.0666
[4] about 0.0535
[5] Insufficient information to answer this question is given.
[0] no answer or skip this item

Submit Answer

13 (8 Points)

Suppose that we know that two events, A and B, are independent, and that P(A) = 0.40 and P(B) = 0.25.

What is the probability of the event "A and B"?

[1] 0.00
[2] 0.10
[3] 0.15
[4] 0.25
[5] 0.65
[0] no answer or skip this item

Submit Answer

14 (8 Points)

Here are the results from a group of 235 freshmen who took a Math Anxiety survey.

Math Anxiety	**Yes**	**No**	**Total**
Male	68	50	118
Female	61	56	117
Total	129	106	235

If a student is randomly selected from this group, P(No Math Anxiety) is:

[1] 61/117
[2] 68/129
[3] 106/235
[4] 129/235
[0] no answer or skip this item

Submit Answer

15 (8 Points)

In a partnership, two members decide to make business decisions independently of each other, then compare their conclusions. If they agree, the decision is made to proceed. Partner A makes the right decision 70% of the time and partner B make the right decision 65% of the time.

What is the probability that both make the wrong decision?

[1] 10.5%
[2] 30%
[3] 35%
[4] 65%
[0] no answer or skip this item

Submit Answer

16 (8 Points)

A business currently has 5 telephone lines. Suppose the number of lines in use at any one time has the following distribution:

Phone lines	**Probability**
0 lines	.10
1 line	.28
2 lines	.44
3 lines	.14
4 lines	.02
5 lines	.02

What is the probability that *less than* 3 lines are in use?

[1] 0.82
[2] 0.62
[3] 0.38
[4] 0.18
[0] no answer or skip this item

Submit Answer

17 (8 Points)

Suppose the amount of time that a customer waits to be admitted to a popular restaurant is uniformly distributed. The probability density function is $f(x) = 0.5$ for $0 < x < 2$, $f(x) = 0$ otherwise (where x represents time in hours).

What is the probability that a customer will have to wait less than 90 minutes (1.5 hours)?

[1] 1/3
[2] 1/2
[3] 3/4
[4] 1/4
[0] no answer or skip this item

Submit Answer

18 (8 Points)

You are given the following probability distribution where the expected value of x is 5:

x	4	5	6
Prob(x)	0.1	0.8	0.1

What is the standard deviation of X?

[1] about 0.632
[2] about 0.447
[3] about 0.400
[4] about 0.200
[0] no answer or skip this item

Submit Answer

19 (8 Points)

In a state lottery game, a player can win either $20,000, $1000, or nothing. The probability is 1/10000 that the player wins the $20,000 prize; the probability is 1/500 that the player wins the $1000 prize.

From the state's point of view, what is the mean payout per player?

[1] $0
[2] $2
[3] $3
[4] $4
[5] $510
[0] no answer or skip this item

Submit Answer

20 (8 Points)

Use WebStat. Load from "Data > Sample data" the data set Pollution_60Metro_areas.dat and compute summary statistics for the variables "NO" and "SO2".

Which of the following statements is NOT correct?

[1] The minimum is the same for both variables.
[2] The median for SO2 is bigger than the median for NO2.
[3] The third quartile (Q3) for SO2 is bigger than the third quartile (Q3) for NO2.
[4] The maximum for SO2 is bigger than the maximum for NO2.
[5] All four statements are correct.
[6] All four statements are incorrect.
[0] no answer or skip this item

Submit Answer

21 (12 Points)

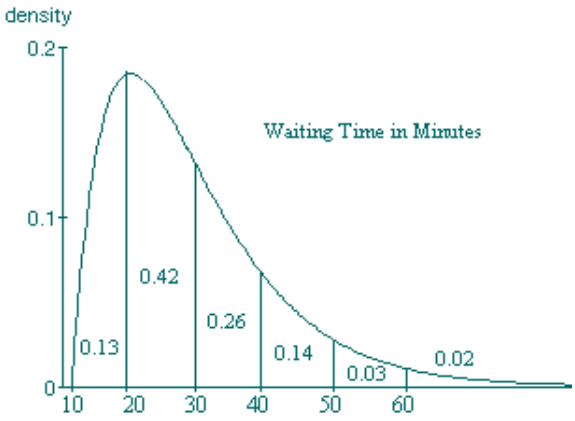
There is a .10 probability that an individual susceptible to tuberculosis(TB), will contract the disease if he or she comes into contact with a carrier of TB. Suppose that 25 susceptible individuals independently come into contact with a TB carrier.

What is the expected (mean) number who would contact the disease and what is the corresponding variance?

Submit Answer

22 (12 Points)

The waiting time in minutes of passengers for a domestic flight is described in this graph:



What is the probability that the waiting time is more than 40 minutes?

Submit Answer

23 (12 Points)

Here are summary statistics for the average daily temperatures for a number of U.S. cities in January and July, for 1996:

	January	**July**
Minimum	12	63
Lower quartile	27	72
Median	31	74
Upper quartile	40	77
Maximum	67	85

Indicate (i) the interquartile range for January and (ii) the interquartile range for July.

Submit Answer

24 (12 Points)

Obesity was not recognized as a health risk for many years because lurking variables were not accounted for in the studies. Here are typical data:

	Over-weight	NormalWeight
Early Death	50	60
Not Early	450	440

What is the probability of Normal Weight, given Early Death?

Submit Answer

25 (12 Points)

The following table shows a probability distribution for *x* = the number of days per week that students go to the college bookstore. The table is based on a survey of 230 students, but we will act as if this is a population distribution.

x	0	1	2	3	4	5	6
Prob(x)	0.60	0.20	0.08	0.04	0.04	0.03	0.01

What is the expected value of the number of days per week that students go to the college bookstore?

Submit Answer

26 (13 Points)

The next interactivity creates scatter plots using data from a survey of a class of 120 students in a college statistics course. Create the plot with X = Sex (1=F, 2=M) and Y = Weight (in pounds).

Summarize and compare the data for the two genders.

CLICK HERE FOR Interactive Tool!

Submit Answer

27 (13 Points)

Use the Venn diagram in the next Interactive to investigate how the conditional probability of two events changes. **Fully drag A inside B.**

What are (i) P(B) and (ii) P(A or B) when A is a subset of B?

CLICK HERE FOR Interactive Tool!

Submit Answer

28 (13 Points)

Look at the interactive boxplots for the athletes from two sports separately.

What are the maximum BMI scores for the two groups?

CLICK HERE FOR Interactive Tool!

Submit Answer

29 (13 Points)

What is your chance of getting a cold or the flu? Annually, about 25% of the U.S. population gets a cold and between 35 -50% gets the flu, but both illnesses are highly age-related. In 1994, for the 5-24 age group, 32.7 million of 74.8 million got the flu. For people 45 years of age and older, 18.8 million of 81.7 million got the flu.

What is the probability an older person (45 years of age and older) will get the flu?

Submit Answer

30 (13 Points)

Take the daily high temperatures for the first week in January, 1996, at Berkeley, but pretend that one was missing and recorded as "-9999" (i.e., with a minus sign in front of the number 9999), so that the data are given as 67, 67, 57, 57, 57, 61, -9999.

Find the average of these seven numbers (that is, pretend you do not realize the -9999 is not a true temperature). Compare this to the correct average of 60.8571. Is it very far off?

Submit Answer

31 (13 Points)

Suppose that only 21% of all drivers come to a complete stop at an intersection having flashing red lights in all directions when no other cars are visible.

For a simple random sample of 200 such drivers, find the expected number of drivers who will *NOT* come to a complete stop.

Submit Answer

32 (13 Points)

A random number generator produces the digits 0 through 9 randomly.

What is the probability that a random digit will be *at least* 7?

Submit Answer

33 (13 Points)

A gene is called 'lethal' (L) if offsprings who receive 2 copies of the gene fail to develop. A cross of parents with one copy each of a dominant 'lethal' gene results in 2/3 of the offspring being carriers of the gene. The combination LL is never born, as illustrated in this table, with *l* being recessive:

Lethal Gene	**father L**	***l***	
Mother L	(L, L)	(L, *l*)	P(L, L) = 1/4 embryo fails to form
l	(L, *l*)	(*l*, *l*)	P(carrier) = P(L, *l*) = 2/3 of children born

You examine 300 offspring from this cross and count the number of offspring *x* who are carriers of the lethal trait.

Find the mean and the standard deviation of *x*.

Submit Answer

34 (13 Points)

A university realized that about 17% of its students fail the introductory statistics class. Assume that the binomial settings hold.

In a class of 90 students, how many students do we expect to fail? And what is the probability that at most 10 students from this class fail?

Use a binomial probability calculator of your choice to answer this question.

Submit Answer

35 (13 Points)

Use WebStat. Load from "Data > Sample data" the data set "Chicago_5_tallest_buildings.dat" and compute the correlation coefficient for the variables "Height" and "Stories".

Report the correlation coefficient and make a statement of the form whether this correlation coefficient indicates whether we have a positive or a negative association. So, is this association strong?

Submit Answer

When you are done answering all questions above to your satisfaction, press the button below to complete your test.

Mark Test Completed