

Ag 400 - Quiz 1
Makeup - Fall 2008

(20) A. Indicate whether each of the following statements is true (T) or false (F). If any part of a statement is untrue, it should be marked false (F). If you wish, write comments to clarify your answers.

- T 1. Samples are never larger than their corresponding populations.
- F 2. Examples of variables that could be measured by interval scales are height, weight, income (in dollars), student ID numbers, and number of hours a person watched television in a week.
- T 3. The mean, median, and mode are descriptive statistics that measure the central tendency of the distribution of an intervally scaled variable.
- F 4. The median is the midpoint of a distribution of scores. If the number of scores is "even" (such as 2, 4, 6, 8, 10, etc.), the median is a score that actually occurs in the distribution. If the number of scores is "odd" (1, 3, 5, 7, 9, etc.) the median is never a score that is in the distribution.
- F 5. The following is a unimodal frequency distribution that is skewed to the right.



- F 6. The variance is the square of the standard deviation. Both the variance and the standard deviation measure the central tendency and dispersion of an intervally scaled variable.
- T 7. The "sum of squares" (i.e. the sum of the squares of the deviations of X about \bar{X}) is never a negative number.
- F 8. $\Sigma (X - \bar{X})$ is large when the dispersion of X is large.
- T 9. If $\Sigma (X - \bar{X}) = 0$ and $\Sigma (X - \bar{X})^2 = 0$, X is a constant (i.e. X does not vary at all).
- T 10. In a grouped data distribution where one of the class intervals is open-ended, you cannot calculate the mean, range, standard deviation or variance.

- (20) B. The following table was compiled from a recent survey of a random sample of Centre County residents (those in the State College area and those living outside the State College area) concerning their views about Issue Z.

Attitude toward Issue Z	Area of Residence	
	State College Area	Outside State College Area
	-----number of residents-----	
Favor	60	30
Oppose	80	60
Undecided	60	10

Indicate whether each of the following statements is true (T) or false (F) in terms of these data. If any part of a statement is untrue, it should be marked false (F). Add comments if you wish to clarify your answers.

- T 1. The independent variable in the above table, area of residence, is measured by a two category nominal scale.
- F 2. The statistical unit is the area of residence (State College Area, Outside State College Area).
- T 3. The ratio of State College Area residents to residents of the County outside the State College area is 2.0.
- F 4. State College Area residents are more likely to favor Issue Z than are county residents living outside the State College area.
- F 4. State College Area residents are more likely to oppose Issue Z than are county residents living outside the State College area.
- F 5. There is no relationship in this sample between the person's area of residence and his/her attitude about Issue Z.
- T 6. The relationship between area of residence (State College Area and Outside State College Area) and attitude toward Issue Z can be described as follows:

Both residents of the State College Area and those living in the county outside the State College Area are more likely to "oppose" than to "favor" Issue Z..
- F 7. County residents living in the State College Area are six times as likely as those living outside the State College Area to be "undecided" about Issue Z.

- F 8. For the total sample, the modal attitude response is “undecided.”
- F 9. The mean number of residents opposing Issue Z is 70.
- F 10. The above table is a contingency table that can be used to describe the relationship between “Area of Residency” and “attitude toward Issue Z. Percentages calculated using these data are parameters.

- C. Suppose that you are interested in assessing the number of penalties amassed by the Nittany Lion football team during its last 30 games. You draw information from the records of those games and compile the following table:

Number of Penalties	midpt (x)	f Number of Games	xf	x ² f
10 - 12	11	1	11	121
7 - 9	8	6	48	384
4 - 6	5	15	75	375
1 - 3	2	8	16	32
		30	150	912

- (6) 1. Calculate the mean number of penalties per game from the above data or indicate why it cannot be calculated.

$$\bar{x} = \frac{150}{30} = 5.0$$

- (6) 2. What is the median number of penalties per game?

$$Md = 3.5 + \left(\frac{30/2 - 8}{15} \right) 3 = 4.9$$

- (6) 3. Calculate the variance for the above distribution.

$$S^2 = \frac{912 - \frac{(150)^2}{30}}{30} = 5.4$$

(12) 4. Indicate whether each of the following statements is true (T) or false (F) in terms of these data. If any part of a statement is untrue, it should be marked false (F). Add comments to clarify your answers if you wish.

F a. The above data are what we have termed "semi-grouped data."

F b. The statistical unit in the above problem is a penalty.

F c. The mode for the above distribution is 4.0.

F d. For the above data, the standard deviation is less than the variance. The standard deviation is always less than the variance.

F e. The above distribution of number of penalties is skewed to the left.

T f. The range in number of penalties for these data is greater than the standard deviation.

- D. Suppose you are interested in analyzing the number of hours of sleep that Penn State students get each day/night. You ask 100 students selected at random from University records to keep a log in which they record the number of hours they sleep each night for a week. Suppose further that the information from these logs is compiled into an SPSS data set (SLEEP.sav).

The data set has the following variables: (The SPSS variable names are indicated in parentheses)

Identification number (id)

Gender (SEX)

1 Male

2 Female

Hours slept on Sunday (SUNHRS)

Actual hours recorded

99 = Missing data

Hours slept on Monday (MONHRS)

Actual hours recorded

99 = Missing data

Hours slept on Tuesday (TUEHRS)

Actual hours recorded

99 = Missing data

Hours slept on Wednesday (WEDHRS)

Actual hours recorded

99 = Missing data

Hours slept on Thursday (THURHRS)

Actual hours recorded

99 = Missing data

Hours slept on Friday (FRIHRS)

Actual hours recorded

99 = Missing data

Hours slept on Saturday (SATHRS)

Actual hours recorded

99 = Missing data

The attached output was obtained using SPSS and this data set.

Indicate whether each of the following statements is true (T) or false (F) in terms of these data. If any part of a statement is untrue, it should be marked false (F). Write comments to clarify your responses if you wish.

- F 1. The statistical unit in the problem is an hour of sleep.
- F 2. The data set has at least 9 rows and 100 columns.
- F 3. You could calculate the total number of hours that each student slept during the week by running FREQUENCIES on the last seven variable and specifying that you want the "mean" and the "sum."
- F 4. COMPUTE could be used to calculate the mean number of hours that the total sample of 100 students slept on Friday night.
- F 5. The attached output was obtained using the data transformation procedure called FREQUENCIES
- F 6. There are two cases for which there is no data on subject's gender.
- T 7. There are more males than females in the sample.
- T 8. The distributions for SUNHRS, MONHRS, TUEHRS, WEDHRS, and THURHRS are skewed to the left. The distributions for FRIHRS and SATHRS are skewed to the right.
- F 9. The data presented on this output are parameters and describe the amount of sleep that Penn State students get.
10. For every day of the week half of the students in this data set slept 6 hours or less than 6 hours and half slept 6 hours or more than 6 hours.
- F 11. The mean of 1.47 for SEX tells you the SEX RATIO is 147.
- F 12. The number 588 in the "statistics" table on Page 1 of the output is the sum of the squares of the deviations of SUNHRS about the mean for SUNHRS.
- F 13. No student in the sample reported less than 2 hours of sleep and no student reported more than ten hours of sleep for a single day/night.
- ~~X~~ 14. Only one student in the sample failed to report information on how many hours he/she slept. If that single student were deleted from the data set, there would be no missing data in the data file for the variables describing hours of sleep.
- T 15. Overall, students slept more hours on Friday and Saturday than on the other days/nights, and there was more variation in hours of sleep among students on Saturday than on any other day/night.

Frequencies

Statistics

		SEX	SUNHRS	MONHRS	TUESHRS	WEDHRS	THURSHRS	FRIHRS	SATHRS
N	Valid	98	100	99	100	99	100	99	99
	Missing	2	0	1	0	1	0	1	1
Mean		1.46	5.88	5.17	5.76	5.86	5.85	6.21	6.21
Median		1.00	6.00	6.00	6.00	6.00	6.00	6.00	6.00
Std. Deviation		.577	1.343	1.229	1.379	1.414	1.410	1.662	2.782
Sum		143	588	512	576	580	585	615	615

Frequency Table

SEX

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1	56	56.0	57.1	57.1
	2	40	40.0	40.8	98.0
	3	1	1.0	1.0	99.0
	4	1	1.0	1.0	100.0
	Total	98	98.0	100.0	
Missing	System	2	2.0		
Total		100	100.0		

SUNHRS

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	3	5	5.0	5.0	5.0
	4	16	16.0	16.0	21.0
	5	10	10.0	10.0	31.0
	6	31	31.0	31.0	62.0
	7	32	32.0	32.0	94.0
	8	5	5.0	5.0	99.0
	9	1	1.0	1.0	100.0
	Total	100	100.0	100.0	

MONHRS

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 2	5	5.0	5.1	5.1
3	5	5.0	5.1	10.1
4	16	16.0	16.2	26.3
5	21	21.0	21.2	47.5
6	46	46.0	46.5	93.9
7	6	6.0	6.1	100.0
Total	99	99.0	100.0	
Missing 999	1	1.0		
Total	100	100.0		

TUESHRS

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 2	5	5.0	5.0	5.0
4	16	16.0	16.0	21.0
5	10	10.0	10.0	31.0
6	32	32.0	32.0	63.0
7	36	36.0	36.0	99.0
8	1	1.0	1.0	100.0
Total	100	100.0	100.0	

WEDHRS

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 2	2	2.0	2.0	2.0
3	8	8.0	8.1	10.1
4	8	8.0	8.1	18.2
5	8	8.0	8.1	26.3
6	37	37.0	37.4	63.6
7	30	30.0	30.3	93.9
8	6	6.0	6.1	100.0
Total	99	99.0	100.0	
Missing 999	1	1.0		
Total	100	100.0		

THURSHRS

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	3	9	9.0	9.0	9.0
	4	11	11.0	11.0	20.0
	5	13	13.0	13.0	33.0
	6	27	27.0	27.0	60.0
	7	33	33.0	33.0	93.0
	8	7	7.0	7.0	100.0
	Total	100	100.0	100.0	

FRIHRS

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	3	7	7.0	7.1	7.1
	4	10	10.0	10.1	17.2
	5	12	12.0	12.1	29.3
	6	30	30.0	30.3	59.6
	7	15	15.0	15.2	74.7
	8	16	16.0	16.2	90.9
	9	9	9.0	9.1	100.0
	Total	99	99.0	100.0	
Missing	999	1	1.0		
Total		100	100.0		

SATHRS

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	2	11	11.0	11.1	11.1
	3	11	11.0	11.1	22.2
	4	11	11.0	11.1	33.3
	5	10	10.0	10.1	43.4
	6	9	9.0	9.1	52.5
	7	10	10.0	10.1	62.6
	8	10	10.0	10.1	72.7
	9	12	12.0	12.1	84.8
	10	11	11.0	11.1	96.0
	11	4	4.0	4.0	100.0
	Total	99	99.0	100.0	
Missing	999	1	1.0		
Total		100	100.0		