CHAPTER 2: LEVELS OF MEASUREMENT AND AGGREGATION

DISCUSSION GROUP QUESTIONS

1. What is the level of measurement for each of the following variables? Why?

   a) Length of incarceration measured in months (with 0 meaning no incarceration)

   b) Number of self-reported deviant acts measured in categories ("0-4"; "5-9"; "10-14"; and "15 or more")

   c) Types of violent crimes (measured in categories of "aggravated assault", "rape", "robbery" and "homicide")

2. In a study examining the effect of child abuse on subsequent juvenile delinquency, what is the independent variable and what is the dependent variable?

3. In 2001, Maryland had the highest robbery rate of any state in the country. If its population was 5,575,156 people, and there were 13,125 total robberies, what was the robbery rate for the state per 100,000 residents in the state?

4. In 2001, there were 3,840 robbery offenses and 208 rape offenses known to the police in the city of Washington D.C. In that year what was the ratio of robberies to rapes?

For the following questions (5-8), remember that the independent variable is the variable thought to affect or in some way contribute to fluctuations or changes in the dependent variable.

5. Suppose that you as a researcher wanted to examine the impact of police officer rank on perceptions of friendliness by citizens in College Park, MD. Rank is measured as patrol officer, sergeant or captain, and citizen perception is captured by a scale ranging from -50 to 50 and measured in a survey given to 1,000 randomly selected people after an interaction with a police officer. 500 people talked to a patrol officer, 300 talked to a sergeant and 200 people talked to a captain.

   a) What is the unit of observation and reasonable population for this sample?

   b) Identify the independent and dependent variable.

   c) Identify what kind of variables citizen perception and officer rank are? (nominal, ordinal, interval, or ratio).

6. A researcher is interested in the relationship between attention deficit hyperactivity disorder (ADHD) and conduct disorder in middle school classes. Her hypothesis is that students with ADHD are more likely to experience conduct disorders at school. She obtained a random sample of seventh and eighth graders from a suburban middle school in Washington D.C. ADHD was measured using a self administered test that ranked the student on the following order: no

attention deficit; some attention deficit; high attention deficit, and conduct disorder was a dichotomous (i.e. two category) variable consisting of conduct disorder/no conduct disorder.

a) What is the unit of observation and population for this sample?

b) Identify the independent and dependent variable.

c) For each variable, name what kind of variable it is (nominal, ordinal, interval, or ratio).

7. A researcher is interested in whether a rehabilitation program reduces the likelihood that prisoners will re-offend after they are released. Two hundred inmates at Maryland Correctional Institution in Jessup are randomly assigned to participate in the program while another two hundred inmates are kept in prison without participating in any rehabilitative programs. All of the inmates were then studied for two years after their release to see how many times they were arrested again.

a) What is the unit of observation and population for this sample?

b) Identify the independent and dependent variable.

c) For each variable, name what kind of variable it is (nominal, ordinal, interval, or ratio).

8. The Women's Resource Center of Maryland (WRCM) has just hired you as a statistical consultant. Your first task is to investigate whether or not rape reporting legislation (measured as a dichotomy: rape legislation; no rape legislation) affects the reporting rate for sexual assaults in Maryland counties (measured as the number of rapes per 100,000 residents). You take a random sample of 4 Maryland counties, 2 of which have the legislation (Baltimore and Frederick) and two of which do not (Prince Georges and Montgomery) Below are your data.

Rape Rate Per 100,000 residents in a Sample of Maryland Counties

County	# of Rapes Reported	County Population	Rape Reporting Rate
Prince George	517	533,398	
Montgomery	371	422,384	
Baltimore	2,124	1,042,350	
Frederick	397	209,751	

a) Calculate the rape rate per 100,000 residents for each county in the table above.

b) What are the unit of observation and population for this sample?

c) Identify the independent and dependent variable and name the level of measurement for each (nominal, ordinal, interval, or ratio).

CHAPTER 2: LEVELS OF MEASUREMENT AND AGGREGATION

DISCUSSION GROUP ANSWERS

1. What is the level of measurement for each of the following variables? Why?

a) Length of incarceration in months (starting at 0 meaning no offenses)

Ratio - non-arbitrary zero

b) Number of self-reported deviant acts (measured in categories of "0-4", "5-9", "10-14", and "15 or more")

Ordinal - rank ordered data

c) Types of violent crimes (measured in categories of "aggravated assault", "rape", "robbery" and "homicide")

Nominal - because it is categorical

2. In a study examining the effect of child abuse on subsequent juvenile delinquency, what is the independent variable and what is the dependent variable?

Independent Variable (IV): ***Child Abuse***
Dependent Variable (DV): ***Subsequent Juvenile Delinquency***

3. In 2001, Maryland had the highest robbery rate of any state in the country. If its population was 5,575,156 people, and there were 13,125 total robberies, what was the robbery rate for the state per 100,000 residents in the state?

(13,125 / 5,575,156) * 100,000 = 235.42 per 100,000

4. In 2001, there were 3,840 robbery offenses and 208 rape offenses known to the police in the city of Washington D.C. In that year what was the ratio of robberies to rapes?

3,840 to 208 = 3,840 /208 = 18.46 robberies to each rape

For the following questions (5-8), remember that the independent variable is the variable thought to affect or in some way contribute to fluctuations or changes in the dependent variable.

5. Suppose that you as a researcher wanted to examine the impact of police officer rank on perceptions of friendliness by citizens in College Park, MD. Rank is measured as patrol officer, sergeant or captain, and citizen perception is captured by a scale ranging from -50 to 50 and measured in a survey given to 1,000 randomly selected people after an interaction with a police

officer. 500 people talked to a patrol officer, 300 talked to a sergeant and 200 people talked to a captain.

a) What is the unit of observation and reasonable population for this sample?
Unit of Observation: Citizen
Population: Citizens of College Park

b) Identify the independent and dependent variables.
IV: Officer Rank
DV: Citizens' Perceptions of Friendliness

c) For each variable, name what kind of variable it is (nominal, ordinal, interval, or ratio).
IV: Ordinal
DV: Interval

6. A researcher is interested in the relationship between attention deficit hyperactivity disorder (ADHD) and conduct disorder in middle school classes. Her hypothesis is that students with ADHD are more likely to experience conduct disorders at school. She obtained a random sample of seventh and eighth graders from a suburban middle school in Washington D.C. ADHD was measured using a self administered test that ranked the student on the following order: no attention deficit; some attention deficit; high attention deficit, and conduct disorder was a dichotomous (i.e. two category) variable consisting of conduct disorder or no conduct disorder.

a) What is the unit of observation and population for this sample?
Unit of Observation: students
Population: 7th and 8th grade suburban middle school students in Washington, D.C.

b) Identify the independent and dependent variable.
IV: ADHD
DV: Conduct disorder

c) For each variable, name what kind of variable it is (nominal, ordinal, interval, or ratio).
IV: Ordinal
DV: Nominal

7. A researcher is interested in whether a rehabilitation program reduces how often prisoners will re-offend after they are released. Two hundred inmates at Maryland Correctional Institution in Jessup are randomly assigned to participate in the program while another two hundred inmates are kept in prison without participating in any rehabilitative programs. All of the inmates were then studied for two years after their release to see how many times they were arrested again.

a) What is the unit of observation and population for this sample?
Unit of Observation : Inmates
Population: Inmates at the MD Correctional Institution in Jessup

b) Identify the Independent and Dependent Variable.
IV: Rehabilitation program participation
DV: Re-arrest upon release from prison

c) For each variable, name what kind of variable it is (nominal, ordinal, interval, or ratio).
IV: Nominal
DV: Ratio

8. The Women's Resource Center of Maryland (WRCM) has just hired you as a statistical consultant. Your first task is to investigate whether or not rape reporting legislation (measured as a dichotomy: rape legislation; no rape legislation) affects the reporting rate for sexual assaults in Maryland counties (measured as the number of rapes per 100,000 residents). You take a random sample of 4 Maryland counties, 2 of which have the legislation (Baltimore and Frederick) and two of which do not (Prince Georges and Montgomery) Below are your data.

Rape Rate Per 100,000 residents in a Sample of Maryland Counties

County	# of Rapes Reported	County Population	Rape Reporting Rate
Prince George	517	533,398	**96.9**
Montgomery	371	422,384	**87.8**
Baltimore	2,124	1,042,350	**203.8**
Frederick	397	209,751	**189.3**

a) Calculate the rape rate per 100,000 residents for each county in the table above.

b) What are the unit of observation and population for this sample?

Unit of Observation: County
Population: All MD Counties

c) Identify the independent and dependent variable and name the level of measurement for each (nominal, ordinal, interval, or ratio).

IV: Rape Reporting Legislation (nominal)
DV: Rape Reporting Rate (ratio)

CHAPTER 2: LEVELS OF MEASUREMENT AND AGGREGATION

SOLUTIONS

1. a. variable b. variable c. constant

2. IQ is the independent variable and delinquency is the dependent variable.

3. Age is the independent variable and delinquency is the dependent variable.

4. a. ratio/interval b. ordinal c. nominal d. ratio/interval
   e. ordinal f. ratio/intervalg. nominal

5. Boston 3.63
   Columbus 9.89
   Montgomery 23.52
   L.A. 2.34
   Chicago 2.12
   NY 2.44

6. Montgomery, AL is the most likely and Chicago is the least likely for becoming the victim of a hate crime.

7.

State	# of executions	p	%
Alabama	6	.06	6%
Virginia	13	.13	13%
Texas	52	.50	50%
Maryland	1	.01	1%
Louisiana	8	.08	8%
Oklahoma	24	.23	23%
Total	104		

Note: proportions do not sum to 1 and percentages do not sum to 100% due to rounding.

8. 50%

9. .13

10. 24

CHAPTER 2: LEVELS OF MEASUREMENT AND AGGREGATION

PRACTICE PROBLEMS

1. If I measured the age, IQ, and number of prior delinquent acts committed within a sample of 200 boys sent to juvenile court:
   a. Would age be a variable or a constant?
   b. Would IQ be a variable or a constant?
   c. Would sex (gender) be a variable or a constant?

2. If I was interested in knowing whether or not boys who had low IQ scores were more delinquent than boys with higher IQ scores, what would be my independent variable and what would be my dependent variable?

3. If I was interested in knowing if older boys were more likely to be delinquent than younger boys, what would be my independent variable and what would be my dependent variable?

4. What is the level of measurement for the following variables:
   a. Age measured as the number of years old someone is.
   b. Age measured as: under 10 years old or younger
      10-29 years old
      30-59 years old
      60 years old or older
   c. Whether someone went to a public, private or parochial elementary school.
   d. Number of prior convictions measured as the actual number of convictions.
   e. Number of prior convictions measured as: None
      1-2 convictions
      3-4 convictions
      5-6 convictions
      7 or more convictions
   f. Number of months sentenced to prison in actual months.
   g. Type of weapon used in committing a murder: gun, knife, club or other blunt instrument, hands, some other means.

5. Calculate the rate of hate crimes for the following cities:

	# of Hate Crimes	Population	Rate per 100,000
Boston, MA	234	6,453,442	
Columbus, OH	112	1,132,475	
Montgomery, AL	100	425,185	
Los Angeles, CA	198	8,453,809	
Chicago, IL	202	9,535,000	
New York, NY	255	10,435,333	

6. In which city are you most likely to become the victim of a hate crime? In which city are you least likely to become the victim of a hate crime?

7. Here is a frequency distribution of the number of executions in several U.S. states in the past five years.

State	# of executions	p	%
Alabama	6		
Virginia	13		
Texas	52		
Maryland	1		
Louisiana	8		
Oklahoma	24		
Total			

Calculate the total number of executions in these states, and fill in the column of proportions and percentages.

8. What percent of the total number of executions were done in Texas?

9. What proportion of the total number of executions were done in Virginia?

10. How many executions were done in Oklahoma?

STATISTICS FOR CRIMINOLOGY & CRIMINAL JUSTICE
EXAM 2

NAME _____ SECTION_____

1. Social Control Theory argues that individuals with strong ties to conventional institutions will tend to commit fewer criminal acts. The following contingency table presents information regarding the strength of conventional ties and number of criminal acts for 89 individuals.

Contingency Table Between Convential Ties and Criminal Acts

	Criminal Acts			
Convential Ties	0-2	3-5	6 or more	TOTAL
Strong	25	13	6	44
Weak	16	13	16	45
TOTAL	41	26	22	89

a) What is the probability that someone has strong conventional ties?

b) What is the probability that someone has committed 6 or more criminal acts?

c) What is the probability that someone had 5 or less criminal acts?

d) What is the probability that someone had weak ties or 3-5 criminal acts? Are these mutually exclusive events?

e) What is the conditional probability that someone had 6 or more criminal acts given that they had strong ties?

f) What is the probability that someone had weak ties and 0-2 criminal acts?

e) Are conventional ties and criminal acts statistically independent or statistically dependent? Explain.

f) What level of measurement is “conventional ties”?

2. You are the Colonel of the Maryland State Police, and you are concerned about drinking among state troopers. You take a random sample of 150 police offers and find that the median number of drinks per week is 2 and the mean number of drinks per week is 4.7 with a standard deviation of 8.

a. Is your sample skewed, how can you tell, and if so, in what direction?

b. Construct a 90% confidence interval around the appropriate point estimate. Interpret your results.

c. Construct a 93% confidence interval around your point estimate. Interpret your results. Why did the interval change?

d. What would happen if we increased our confidence level to 99% (You do not need to recalculate the CI)?

e. What would happen to the size of your 90% confidence interval if you increased your sample size from 150 to 500? (You do not need to recalculate the CI).

f. Recalculate a 90% Confidence Interval but assume that now you only have a sample size of 16 officers (your sample standard deviation is still 8). Interpret this new interval and explain why it is different from the 90% Confidence Interval you calculated in part b) above? Be specific.

.

3. According to the Gallup Polling Organization, 46% of the United States population thinks that the criminal justice system is too lenient. You take a sample of 200 University of Maryland students and find that 80 of them think that the criminal justice system is too lenient.

a. Construct a 95% confidence interval around your point estimate.

b. In a hypothesis test, where the null hypothesis is that the percentage of Maryland students who think that the CJ System is too lenient is no different than the

national average; the alternative hypothesis is that Maryland students are different than the national average. Based on the results from your confidence interval above (i.e. with an alpha of .05), would you reject or fail to reject the null hypothesis? You *do not* need to do the steps of the test.

c. What would happen to the size of your interval if you increased your confidence from 95% to 99%?

d. What would happen to the size of your 95% confidence interval if you increased your sample size from 200 to 500?

e. What would happen to the size of your 95% confidence interval if you decreased your sample size from 200 to 50?

4. Your job as the research director in the Maryland Department of Youth Services is to advise the Director which policies to follow. You know that, on average, there are 35 "disturbances" (riots, stabbings, fights, etc.) per year per institution, with a standard deviation of 7.0.

a. What is the z-score for an institution that had 28 disturbances in one year? What does this z-score indicate?

b. What is the z-score for an institution that had 45 disturbances in one year?

c. You want to give raises to the wardens in the prisons that have disturbances in the bottom 10% of the distribution, and fire the wardens in prisons in the top 5%. What number of disturbances will be the cut-off for each of these decisions?

d. What is the probability that an institution would have 40 or more disturbances?

e. What is the probability that an institution would have 32 or fewer disturbances?

f. Suppose the Director wanted to identify the top 20% of institutions for reprimand. How many disturbances would qualify as the cut-off?

g. Suppose the Director wanted to identify the bottom 5% of institutions for praise. How many disturbances would qualify as the cut-off?

5. Suppose you take a sample of 81 offenders from Maryland state prisons and you find that the average number of tattoos is 4.4 with a standard deviation of 1.9. According to *Tattoo Me* magazine, the average number of tattoos for all incarcerated offenders is 2.9.

a) Test the null hypothesis that Maryland offenders have the same mean number of tattoos as other offenders in the general incarcerated population against the alternative that they have a significantly different mean number of tattoos. Use an alpha level of .05. Interpret your results.

Step 1:

Step 2:

Step 3:

Step 4:

Step 5:

b) Suppose your sample size was only 40 prisoners. Can you still conclude that there is a significant difference in the mean number of tattoos? Conduct the appropriate hypothesis test to answer this question. Interpret your results.

Step 1:

Step 2:

Step 3:

Step 4:

Step 5:

c) You take a new sample of 100 prisoners who have tattoos and find that 42% of their tattoos are located on their arms. *Tattoo Me* magazine reports that 38 percent of all individuals with tattoos have them on their arms. Is the proportion of tattoos on prisoners' arms significantly greater than the proportion in the general public? Using an alpha of .01, conduct a hypothesis test to answer this question. Interpret your results.

Step 1:

Step 2:

Step 3:

Step 4:

Step 5:

6. You are interested in examining if the gender of the offender is related to whether or not armed robbers are arrested. You know that the overall probability of arrest for armed robbers is .7. You obtain a random sample of 10 female armed robbers and observe that only 2 of them were arrested. Assuming we can use a binomial to answer this question, calculate the probability of observing exactly 2 arrests out of 10 robberies in your data.

STATISTICS FOR CRIMINOLOGY & CRIMINAL JUSTICE
EXAM 2

NAME _____ SECTION_____

1. Social Control Theory argues that individuals with strong ties to conventional institutions will tend to commit fewer criminal acts. The following contingency table presents information regarding the strength of conventional ties and number of criminal acts for 89 individuals. Use this information to answer the following probability questions.

Contingency Table Between Convential Ties and Criminal Acts

	Criminal Acts			
Convential Ties	0-2	3-5	6 or more	TOTAL
Strong	25	13	6	44
Weak	16	13	16	45
TOTAL	41	26	22	89

a) What is the probability that someone has strong conventional ties?
44/89 = .494

b) What is the probability that someone has committed more than 5 criminal acts?
22/89 = .247

c) What is the probability that someone had 5 or less criminal acts?
67/89 = .753

d) What is the probability that someone had weak ties or 3-5 criminal acts? Are these mutually exclusive events? Explain.

P(weak ties or 3-5 acts) = (45+26-13)/89 = .652

OR

P(weak ties) = 45/89 = .506 P(3-5 acts) = 26/89 = .292
P(weak ties and 3-5 acts) = 13/89 = .146
P(weak ties or 3-5 acts) = (.506 + .292) - .146 = .652

These are not mutually exclusive events because an individual can be in both categories. That is, someone can have weak ties and commit 3-5 criminal acts (this is why we need to subtract the joint probability of weak ties and 3-5 acts = .146).

e) What is the conditional probability that someone had 6 or more criminal acts given that they had strong ties?
6/44 = .136

f) What is the probability that someone had weak ties and 0-2 criminal acts?
16/89 = .180

e) Are conventional ties and criminal acts statistically independent or statistically dependent? Explain.

Test if P(A) = P(A | B)
P(strong ties) = 44/89 = .494
P(ST | 0-2) = 25/41= .610
P(ST | 3-5) = 13/26 = .500
P(ST | 6 or more) = 6/22 = .273
Because P(A) ≠ P(A | B), the two variables are Statistically Dependent

[You only need to calculate one of the conditionals to conclude if it's independent or dependent.]

This means that the level of conventional ties (i.e. having strong ties) tells us something about the probability of criminal acts. As you can see, the probability of strong ties decreases as the number of criminal acts increases. Individuals with strong ties are less likely to commit higher numbers of criminal acts.

f) What is the probability that someone did not have strong ties and did not commit 6 or more criminal acts?

P(Weak AND not 6 or more) = 16/89 + 13/89 = 29/89 = .326

If you want to do it the hard way, you can use the multiplication rule…
Find P(A)*P(B|A)
Not having strong ties means having weak ties. P(weak ties) = 45/89 = .506
Not committing 5 or more acts means committing either 0-2 or 3-5 acts.
P(0-2 or 3-5 | weak ties) = 29/45 = .644
The joint probability of weak ties and 0-5 acts = .506*.644 = .326

2. You are the Colonel of the Maryland State Police, and you are concerned about drinking among state troopers. You take a random sample of 150 police offers and find that the median number of drinks per week is 2 per week and the mean number of drinks per week is 4.7 with a standard deviation of 8.

a. Is your sample skewed, how can you tell, and if so, in what direction?

Yes. Positively skewed. The mean is greater than the median.

b. Construct a 90% confidence interval around the appropriate point estimate. Interpret your results. (5 pts)

$$90\%C.I. = \bar{X} \pm z_{\alpha}\left(\frac{s}{\sqrt{n}}\right) = 4.7 \pm 1.65(8/12.25) = 4.7 \pm 1.08$$

$$= 3.62 \leq \mu \leq 5.78$$

You can be 90% confident that the true population mean number of drinks among Maryland state troopers is between 3.62 and 5.78.

c. Construct a 93% confidence interval around your point estimate. Interpret your results. Why did the interval change? (5 pts)

With a 93% C.I. 7% of the distribution is split between the tails, leaving .035 in each tail. .5 - .035 = .465 – look this up in the Z Table and you get a z_α of 1.82.

$$93\%C.I. = \bar{X} \pm z_\alpha\left(\frac{s}{\sqrt{n}}\right) = 4.7 \pm 1.82(8/12.25) = 4.7 \pm 1.19$$

$$= 3.51 \le \mu \le 5.89$$

You can be 93% confident that the true population mean number of drinks among Maryland state troopers is between 3.51 and 5.89. Increasing our confidence (from 90% to 93%) results in a wider range. If you want to be more sure of including the true population mean in your confidence interval, you must include more values in your interval.

d. What would happen if we increased our confidence level to 99% (You do not need to recalculate the CI)? (2 pts)

Increasing the confidence level 99% would result in an even wider interval.

e. What would happen to the size of your 90% confidence interval if you increased your sample size from 150 to 500? (You do not need to recalculate the CI). (2 pts)

Increasing the sample size shrinks the standard error (by increasing the denominator of the fraction) resulting in a narrower interval. Also, as you include more people in your sample, it becomes more and more likely that the sample mean reflects the true population mean.

f. Recalculate a 90% Confidence Interval but assume that now you only have a sample size of 16 officers (your sample standard deviation is still 8). Interpret this new interval and explain why it is different from the 90% Confidence Interval you calculated in part b) above? Be specific.

.

$$90\%C.I. = \bar{X} \pm t_\alpha\left(\frac{s}{\sqrt{n}}\right) = 4.7 \pm 1.753(8/4) = 4.7 \pm 3.51$$

$$= 1.19 \le \mu \le 8.21$$

You can be 90% confident that the true mean number of drinks among MD state troopers is between 1.19 and 8.21. This interval is considerably wider than the 90% confidence interval calculate in part b) above because with a smaller sample size it is necessary to use the t distribution. The shape of the t distribution is a function of the sample size n (incorporated in the formula through the degrees of freedom (df)), and with small sample sizes (such as 16 police officers) the t distribution is flatter and more spread out than the z distribution resulting in a wider confidence interval at the same alpha level.

3. According to the Gallup Polling Organization, 46% of the United States population thinks that the criminal justice system is too lenient. You take a sample of 200 University of Maryland students and find that 80 of them think that the criminal justice system is too lenient.

a. Construct a 95% confidence interval around your point estimate. **(5 points)**

$$95\%C.I. = \hat{p} \pm z_{\alpha}\left(\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right) = .40 \pm 1.96\sqrt{\frac{(.40)(.60)}{200}} = .40 \pm .068$$

$$= .332 \le P \le .468$$

We can be 95% confident that the true population proportion of Maryland students who think the CJ system is too lenient is between .332 and .468

b. In a hypothesis test, where the null hypothesis is that the percentage of Maryland students who think that the CJ System is too lenient is no different than the national average; the alternative hypothesis is that Maryland students are different than the national average. Based on the results from your confidence interval above (i.e. with an alpha of .05), would you reject or fail to reject the null hypothesis? You DO NOT need to do the steps of the test…

Fail to Reject the null -- .46 falls within the confidence interval:
.332 < .46 < .468 so there is no evidence of a significant difference

c. What would happen to the size of your interval if you increased your confidence from 95% to 99%? **(2 points)**

Increasing your level confidence results in widening your interval.

d. What would happen to the size of your 95% confidence interval if you increased your sample size from 200 to 500? **(2 points)**

Increasing your sample size reduces your standard error and results in a narrower, more precise, confidence interval.

e. What would happen to the size of your 95% confidence interval if you decreased your sample size from 200 to 50? **(2 points)**

Conversely, shrinking your sample size increases your standard error, resulting in a wider and less precise confidence interval.

4. Your job as the research director in the Maryland Department of Youth Services is to advise the Director which policies to follow. In the past, you have heard that some institutions have had as many as 80 disturbances in a year. The average number of disturbances (riots, stabbings, fights, etc.) per year per institution, though, is 35, with a standard deviation of 7.0. The distribution is normal. Answer the following questions.

a. What is the z-score for an institution that had 28 disturbances in one year? What does this z-score indicate?

$$z = \frac{x - \bar{X}}{s} = \frac{28 - 35}{7} = -1$$

28 disturbances is exactly 1 standard deviation below the mean of 35

b. What is the z-score for an institution that had 45 disturbances in one year?

$$z = \frac{x - \bar{X}}{s} = \frac{45 - 35}{7} = 1.43$$

c. You want to give raises to the wardens in the prisons that have disturbances in the bottom 10% of the distribution, and fire the wardens in prisons in the top 5%. What number of disturbances will be the cut-off for each of these decisions?

For the bottom 10%, we have .10 in the left tail of the distribution, and (.50 - .10) or .40 between the mean and our cutoff. We find the Z-score associated with a probability of .40, which is -1.28. Now we will solve for -1.28

$$-1.28 = \frac{x - \bar{X}}{s} = \frac{x - 35}{7} = -1.28(7) + 35 = 26.04$$

For the top 5%, we have .05 in the right tail of the distribution, and (.50 - .05) between the mean and our cutoff. We find the Z-score associated with a probability of .45, which is 1.65. Now we will solve for 1.65

$$1.65 = \frac{x - \bar{X}}{s} = \frac{x - 35}{7} = 1.65(7) + 35 = 46.55$$

Wardens who have 26.04 disturbances or fewer should get a raise, and wardens with 46.55 disturbances or more should be fired.

d. What is the probability that an institution would have 40 or more disturbances?

$$z = \frac{x - \bar{X}}{s} = \frac{40 - 35}{7} = .714$$

Find z-score of .71 in Z-Table = .2611
(.5 - .2611) = .239 or about a 23.9% probability of observing 40 or more disturbances

e. What is the probability that an institution would have 32 or fewer disturbances?

$$z = \frac{x - \bar{X}}{s} = \frac{32 - 35}{7} = -.429$$

Find z-score of .43 = .1664; (.5-.1664) = .334
Or 33.4 % probability of 32 or fewer disturbances.

f. Suppose the Director wanted to identify the top 20% of institutions for reprimand. How many disturbances would qualify as the cut-off?

Identify Z-score that corresponds to .20 in the right tail of the distribution.
.5 - .2 = .3 Look up .3 in body of Z Table
Z = .84 Then solve the Z equation for x

$$.84 = \frac{x - \bar{X}}{s} = \frac{x - 35}{7} = .84(7) + 35 = 40.88 \sim 41 \text{ disturbances}$$

g. Suppose the Director wanted to identify the bottom 5% of institutions for praise. How many disturbances would qualify as the cut-off?

See above explanation (be sure to use -1.65 not 1.65).

$$-1.65 = \frac{x - \bar{X}}{s} = \frac{x - 35}{7} = -1.65(7) + 35 = 23.45 \text{ disturbances}$$

5. Suppose you take a sample of 81 offenders from Maryland state prisons and you find that the average number of tattoos is 4.4 with a standard deviation of 1.9. According to *Tattoo Me* magazine, the average number of tattoos for all incarcerated offenders is 2.9.

a) Test the null hypothesis that Maryland offenders have the same mean number of tattoos as other offenders in the general incarcerated population against the alternative that they have a significantly ***different*** mean number of tattoos. Use an alpha level of .05. Interpret your results.

Step 1: $H_0: \mu_{MD\ tattoos} = 2.9$
$H_1: \mu_{MD\ tattoos} \neq 2.9$

Step 2: **Z-Distribution Two Tailed**

Step 3: **Alpha = .05**
$z_{crit} = \pm 1.96$ **Reject if $z_{obt} > 1.96$ or $z_{obt} < -1.96$**
[or Reject of $|z_{obt}| > 1.96$]

Step 4: $$z_{obt} = \frac{\bar{X} - \mu}{\left(\frac{s}{\sqrt{n}}\right)} = \frac{4.4 - 2.9}{\left(\frac{1.9}{\sqrt{81}}\right)} = \frac{1.5}{.211} = 7.11$$

Step 5: **7.11>1.96 so we reject the null hypothesis of no difference Maryland prisoners have significantly more tattoos than the general population of incarcerated individuals**

b) Suppose your sample size was only 40 prisoners. Can you still conclude that there is a significant difference in the mean number of tattoos? Conduct the appropriate hypothesis test to answer this question. Interpret your results.

N=40

Step 1: $H_0: \mu_{MD\ tattoos} = 2.9$
$H_1: \mu_{MD\ tattoos} \neq 2.9$

Step 2: **t-Distribution Two Tailed**

Step 3: **Alpha = .05**
$t_{crit} = \pm 2.021$ **Reject if $t_{obt} > 2.021$ or < -2.021**
[or Reject of $|t_{obt}| > 2.021$]

Step 4: $$t_{obt} = \frac{\bar{X} - \mu}{\left(\frac{s}{\sqrt{n}}\right)} = \frac{4.4 - 2.9}{\left(\frac{1.9}{\sqrt{40}}\right)} = \frac{1.5}{.300} = 5.00$$

Step 5: **5.00>2.021 so we still reject the null hypothesis of no difference Even with a sample of only 42 prisoners we still conclude that Maryland prisoners have significantly more tattoos than the general population of incarcerated individuals**

c) You take a new sample of 100 prisoners who have tattoos and find that 42% of their tattoos are located on their arms. *Tattoo Me* magazine reports that 38 percent of all individuals with tattoos have them on their arms. Is the proportion of tattoos on prisoners' arms significantly ***greater*** than the proportion in the general public? Using an alpha of .01, conduct a hypothesis test to answer this question. Interpret your results.

Step 1: **H_0: $p_{MD\ arm\ tattoos} = .38$**
H_1: $p_{MD\ arm\ tattoos} > .38$

Step 2: **Np, Nq > 5 (100*.38=38; 100*.62=62)**
Z-Distribution One Tailed

Step 3: **Alpha = .01**
$z_{crit} = 2.33$ **Reject if $z_{obt} > 2.33$**

Step 4:

$$z_{obt} = \frac{\hat{p} - p}{\left(\sqrt{\frac{p(1-p)}{n}}\right)} = \frac{.42 - .38}{\left(\sqrt{\frac{(.38)(.62)}{100}}\right)} = \frac{.04}{.049} = .82$$

Step 5: **.82<2.33 so we fail to reject the null hypothesis.**
There is no evidence that Maryland prisoners are more likely to have tattoos on their arms than the general public

6. You are interested in examining if the gender of the offender is related to whether or not armed robbers are arrested. You know that the overall probability of arrest for armed robbers is .7. You obtain a random sample of 10 female armed robbers and observe that only 2 of them were arrested. Assuming we can use a binomial to answer this question, calculate the probability of observing exactly 2 arrests out of 10 robberies in your data.

$$P(r) = \left(\frac{n!}{r!(n-r)!} \right) p^r q^{n-r} \quad = \quad P(2) = \left(\frac{10!}{2!(10-2)!} \right) .7^2 .3^{10-2} \quad = \quad \textbf{(45)(.49)(.00006561)}$$

= .0014

STATISTICS for Criminology and Criminal Justice

FOURTH EDITION

Ronet D. Bachman • Raymond R. Paternoster

Chapter 2: Levels Of Measurement And Aggregation

Types of Data and Levels of Measurement

Two Major "Types" of Data

1. Variable: A property or characteristic of the sample or population that takes on more than one value.

Examples: In this class, is

Age a variable?	Values?
Sex a variable?	Values?
G.P.A. a variable?	Values?
Year in College?	Values?
High School Graduate?	Values?

Types of Data and Levels of Measurement

2. Constant: A property or characteristic of the sample or population that takes on one and only one value.

Examples:

Sex in an all male high school? A college football team?

Values?

A VARIABLE HAS VALUES THAT DIFFER BUT A CONSTANT DOES NOT!!

A CONSTANT HAS ONLY ONE VALUE.

Why would we be interested in variables?

Point Estimation: We want to know or to estimate the value of a particular variable. Estimate values of μ and P.

What are relationships?

Examining Relationships: We want to know if one variable is related to/correlated or associated with another.

Concern whether two variables are related to each other – X & Y. Two variables are said to be related or associated or correlated when variation in X is related to variation in Y. They change together – THIS IS CALLED COVARIATION. Covariation tells us about association.

Levels of Measurement for Variables

- Nominal – categories that provide distinctions of kind
- Ordinal- categories that provide rank ordered distinctions (more or less than)
- Interval
- Ratio

} both are continuous measures with the property of equidistant intervals – the distance between any two adjacent intervals is the same.

- We will use interval/ratio as the same level.

Examples: What is the Level of Measurement for these Variables?

Gender

1. Male
2. Female

Religion

1. Protestant
2. Judaism
3. Muslim
4. Roman Catholic
5. Evangelical
6. Hinduism
7. Agnostic/Atheist
8. Other

Political Views

1. Liberal
2. Conservative
3. Moderate/Independent

Income

In Dollars $ _____

Income

_____ 0 - $29,999
_____ $30,000 – 59,999
_____ $60,000 - 89,999
_____ $90,000 - 114,999
_____ $115,000 and over

Year in High School

_____ Freshman (9th)
_____ Sophomore
_____ Junior
_____ Senior

Year in High School

_____ years

Years in College

_____ years

Number of Prior Convictions

_____ Prior Convictions

Number of prior convictions

- 0
- 1-3
- 4-6
- 6-8
- 8 or more

Type of Crime Committed

- Property
- Violent
- Drug Possession

Type of Sentence Received after Conviction

- Fine
- Probation
- Jail
- Prison
- Death

Simple Descriptive Statistics

1. Descriptive Statistics

- Used to describe, summarize, or inform about the variables in a dataset.

Let's say I wanted to describe some characteristics of the class I had last semester.

Frequency or Count

Gender	f
Males	66
Females	75
Total	141

Race/Ethnicity	
Caucasian	84
African-American	24
Hispanic/Latino	15
Asian-American	2
Other	8
Total	133

# of Previous Semesters in College	f
1	1
2	36
3	17
4	31
5	21
6	14
7	11
8	6
10	2
13	2
16	1
TOTAL	141

# of Siblings At Home	f
0	24
1	47
2	31
3	20
4 or more	19
Total	141

2013 Homicides	f
New York	333
Chicago	412
Detroit	332
Baltimore	234
Philadelphia	246
New Orleans	155

2. Rate – frequency or count per some standard unit of measurement

Rate = {f/population} x standard unit

2013 Homicides	f	Pop.	Rate per 100,000
New York	333	8,405,837	3.96 per 100,000

NY = 333/8,405,837 x 100,000 = 7.30

Chicago	412	2,718,782	15.15 per 100.000
Detroit	332	668,701	49.65 per
Baltimore	234	622,104	37.61 per
Philadelphia	246	1,553,165	15.84 per
New Orleans	155	378,715	40.93 per

San Pedro Sula, Honduras	1,319	769.025	171.20 per 100,000
Acapulco,MX	883	847,735	104.16

When do you want to calculate a rate? When the things you are counting can vary in their opportunity to take place.

Rate can also be measured at the individual level– frequency or count per some standard unit of measurement

Data are the number of new offenses committed by five people released from prison.

Person	f
# 1	5
2	11
3	4
4	16
5	5

Rate = f/time exposed x standard unit of time

Person	f	Months Out	Monthly Rate	Annual Rate
# 1	5	18	.278	3.33
2	11	24	.458	5.50
3	4	8	.500	6.00
4	16	42	.381	4.57
5	5	27	.185	2.22

5/18 x 12 MONTHS = 3.33 crimes per month

3. Proportion

- A relative frequency – it's the frequency of some value relative to the total number of events.

P = fi /N where fi = number of cases at value i, N = total number of cases

Race of Victim/Offender in 350 Maryland Death Cases

Type	f	p	
White Kills white	114	114/350	= .33
White Kills black	23	23/350	= .06
Black Kills white	28	28/350	= .08
Black Kills black	185	185/350	= .53
	350	1.00	
	Σ f = N	Σ p = 1.0	

Proportions are bound by 0.0 and 1.0

Proportion can also be understood as a probability.

4. Percent

- A standardized proportion – it's what the frequency would be if N = 100.

$\% = (f_i / N) \times 100$

where f_i = number of cases at value I, N = total number of cases

Race of Victim/Offender in 350 Maryland Death Cases

Type	f		p	%
White Kills white	114	114/350	.33	33%
White Kills black	23	23/350	.06	6%
Black Kills white	28/350		.08	8%
Black Kills black	185/350		.53	53%
	350		1.00	100%
	Σ % = 100			

Percentages are bound by 0.0 and 100.0

5. Percent Change

- Measures the percent change in some event from one time to another.

# of Violent Crimes

1980 1,241 per 100,000

1990 1,789

2000 1,137

% change measured as:

[(ending value – starting value)/starting value] x 100

{[(EV – SV)/SV] x 100}

1980 -1990 Change in Violent Crime

(1,789 – 1,241)/1,241 x 100 = (548/1,241) x 100 = .442 x 100 = 44.2 or 44% increase in violent crime.

1990 – 2000 Change in Violent Crime2

(1,137 – 1,789)/1,789 x 100 = (-652/1,789) x 100 = -.364 x 100 = -36.4 or 36%
decline in violent crime.

1980 – 2000 Change in Violent Crime

(1,137-1,241)/1,241 x 100 = (-104/1,241) x 100 = -.083 x 100 = -8.3 or 8% decline in violent crime.

Open-Access Student Resources

- Action Plan
- Quizzes
- eFlashcards
- SAGE Journal Articles
- Web Resources
- Additional Practice Problems & Solutions
- Datasets

edge.sagepub.com/bachmansccj4e