

08/13/2000

Ordered Relations Conditional Thinking **Foundational Problems in Statistical Literacy** 1

JSM 2000 ASA

August 13, 2000

MILO SCHIELD

Augsburg College

www.augsburg.edu/ppages/schild
schild@augsburg.edu

08/13/2000

Ordered Relations Conditional Thinking **Conditional Probability What is part? What is whole?** 2

Probability of being in poverty as IQ goes from low to high

08/13/2000

Ordered Relations Conditional Thinking **Difficulties Reading Rates and Percentages in Tables** 3

Percentage of Smoking Prevalence Among U.S. Adults, 18 Years of Age and Older:

Year	All	Males	Females	Whites	Blacks
1955	--	56.9	28.4	--	--
1965	42.4	51.9	33.9	42.1	45.8
1970	37.4	44.1	31.5	37.0	41.4
1980	33.2	37.6	29.3	32.9	36.9
1990	25.5	28.4	22.8	25.6	26.2

CHOICES:

- percentage of these adults who are black smokers
- percentage of these black adults who are smokers
- percentage of these adult smokers who are black.

08/13/2000

Ordered Relations Conditional Thinking **Reading Percentages: Ambiguity of 'with' and 'to'** 4

No. 103. Low Birth Weight and Births to Teenage Mothers and to Unmarried Women—States: 1990 to 1996

STATE	PERCENT OF BIRTHS WITH LOW BIRTH WEIGHT ^a			BIRTHS TO TEENAGE MOTHERS, PERCENT OF TOTAL		
	1990	1995	1996	1990	1995	1996
U.S.	7.0	7.3	7.4	12.8	13.1	12.9
AL.	8.4	9.0	9.3	18.2	18.5	18.3
AK.	4.8	5.3	5.5	9.7	11.2	11.2
AZ.	6.4	6.8	6.6	14.2	15.1	15.0
AR.	8.2	8.2	8.5	19.7	19.6	19.8
CA.	5.8	6.1	6.0	11.6	12.4	12.0

Source: 1998 US Statistical Abstract (Section on unmarried women omitted)

08/13/2000

Ordered Relations Conditional Thinking **Named Ratio Usage Varies by Source** 5

SOURCE	Ratio			Chance-Probability
	% of	Rate	Percentage	
1. Intro Statistics Text	5	5	0	90
2. Popular Essays	30	20	10	40
3. Data: 1998 U. S. Statistical Abstract	40	40	20	0

Percents are estimates at this time
Intro Statistics text: Anderson & Sweeney.

08/13/2000

Ordered Relations Conditional Thinking **Conditional Thinking: Students lack the basics** 6

Basic	Advanced
Tables Graph Series	Mean, Std.Deviation, Percentile, Z,
% , Rates, Percentages, Chance, Odds, Risk Bayes Rule (counts) Arithmetic Comparisons,	Bayes Rule (Algebra) Correlation, Linear Regression ANOVA
Likely, Attributable	Logistic Regression

Ordered Relations Conditional Thinking

Arithmetic Comparisons of Counts and Named Ratios

7

```

graph TD
    A[Arithmetic Comparisons Of Counts] --> B[Simple Difference [Test - Base]]
    A --> C[Simple Ratio [Test / Base]]
    A --> D[Relative Difference [(Test- Base) /Base]]
    C --> E[Named Ratios: Percent, Rate, Percentage, Chance, Risk, Probability]
    E --> F[Attributable To: Count or Percent]
    E --> G[Arithmetic Comparisons]
    E --> H[Likely Family: Risky, Probable]
    
```

Ordered Relations Conditional Thinking

Grammar Difference: Rates versus Percentages

8

1. **Adjectives:** "accident rate" or "accident percentage"
2. **'Of':** "Rate of inflation" or "Percentage of inflation"
3. **'Of' followed by a relative clause:**
"Rate of workers who are unemployed" or "Percentage of workers who are unemployed"
4. **'Of' and 'among':**
"Rate of unemployment among workers" or "Percentage of unemployment among workers"

Ordered Relations Conditional Thinking

Grammar of Rates Problems

9

1. The accidental death rate among *teenagers*
2. The *teenagers'* accidental death rate is ...
3. The accidental death rate of *teenagers* is ...

4. The *teenager* accidental death rate is ...
5. The rate of *teenager* deaths is ...

These are ambiguous; possessive is unstated.

Ordered Relations Conditional Thinking

Grammar of Rates

Part is underlined

10

The part can be indicated several ways:
Rate of: The rate of births is ...
Rate Modifier: The birth rate is ...
Verb: People died at a rate of
Predicate: People are dying at a rate of

The whole can be indicated several ways:
Rate of: The death rate of teenagers is ...
Possessive: The teenagers' death rate ...
Among: Among teens, the death rate ...

Ordered Relations Conditional Thinking

Statistical Literacy Focus on Ordered Relations

11

Students have difficulty with conditional probability.

- Hypothesis tests and p-values:
 $P(z > k | H_0 \text{ is true})$ with $P(H_0 \text{ is true} | z > k)$
- Confidence Intervals:
 $P(\text{sample mean will be in interval} | \mu)$ with
 $P(\mu \text{ will be in the interval} | \text{sample mean})$.

•David Moore "What is Statistics?" MAA Notes #21
 Garfield and Ahlgren, 1988. "Difficulties in Learning Basic Concepts in probability and Statistics..."

Ordered Relations Conditional Thinking

Conclusion for Statistical Literacy

12

**Greater focus on Named Ratios:
 Percents, Rates, Percentages,
 Chance, Risk, Odds and Probability.**

- Describing and comparing
- Separating association & causation,
- Separating spurious from biased,
- "Check your assumptions..."

08/13/2000

Ordered Relations Conditional Thinking

Grammar of Percentages

Part is underlined

13

"Percent(age) of" normally indicates a whole:
52% OF males are smokers
The percentage OF males who are smokers is 20%

"Percentage of" can indicate the part:
Among males, the percentage of smokers is 20%

Sometimes it is hard to tell.
Among men, the percentage OF smokers who run

08/13/2000

Ordered Relations Conditional Thinking

Rates

Ambiguity of 'by'

14

No. 149. Death Rates for Injury by Firearms, Sex, Race, and Age: 1995
[Death rate per 100,000 population. Deaths classified according to the ninth revision of the International Classification of Diseases]

ITEM	5-14 yrs. old	15-24 yrs. old	25-34 yrs. old	35-44 yrs. old	45-54 yrs. old	55-64 yrs. old	65-74 yrs. old	75-84 yrs. old
MALE								
Firearms: White	2.5	31.4	26.1	21.2	19.6	19.9	26.1	39.8
Black	5.5	140.2	94.4	46.6	32.1	24.3	22.0	20.9
Accidents: White	0.7	1.8	0.8	0.6	0.5	0.4	0.6	0.7
Black	0.8	4.3	1.5	(B)	(B)	(B)	(B)	(B)
Suicide: White	0.8	15.4	15.1	14.2	14.9	16.6	23.9	38.2
Black	(B)	13.2	11.9	7.6	6.9	7.5	10.2	13.9
Homicide: White	0.9	13.6	9.8	6.3	4.0	2.8	1.5	0.8
Black	4.1	121.0	80.7	38.3	24.6	15.9	10.8	(B)

'by means of' versus 'categorized by'
 Source: 1998 US Statistical Abstract (See Table 152 for a better title)

08/13/2000

Ordered Relations Conditional Thinking

Percent:

Ambiguous Phrase

15

No. 113. Percent Low Birthweight, by Smoking Status, Age, and Race of Mother: 1993
[Low birthweight is defined as weight of less than 2,500 grams (5 lb. 8 oz.). Excludes California, Indiana, New York, and South Dakota, which did not require reporting of tobacco use during pregnancy.]

SMOKING STATUS AND RACE OF MOTHER	All ages	AGE OF MOTHER									
		Under 15 years		15-19 years		20-24 years	25-29 years	30-34 years	35-39 years	40-49 years	
		Total	15-17 years	18-19 years							
All races ¹	7.4	13.8	9.6	10.5	9.0	7.5	6.5	6.8	8.0	9.0	
Smoker	11.8	14.7	10.8	11.4	10.5	10.4	11.5	13.6	16.1	17.8	
Nonsmoker	6.6	13.8	9.3	10.3	8.6	6.8	5.6	5.7	6.8	7.9	
Not stated	9.2	14.2	11.8	12.9	11.1	9.0	8.2	8.7	9.9	10.3	
White	6.1	10.8	7.9	8.6	7.5	6.1	5.4	5.7	6.8	7.7	
Smoker	10.1	14.0	10.3	11.0	9.9	9.2	9.4	10.9	13.3	14.7	
Nonsmoker	5.2	10.3	7.1	7.9	6.6	5.2	4.6	4.9	5.9	6.9	
Not stated	7.6	(B)	9.7	10.9	9.1	7.8	6.6	7.4	8.2	9.5	
Black	13.4	16.1	13.4	13.9	13.0	12.3	13.2	14.8	16.6	17.4	
Smoker	22.6	19.6	17.2	17.1	17.3	18.8	23.2	26.3	27.8	30.4	
Nonsmoker	12.0	15.9	13.1	13.7	12.6	11.4	11.2	11.8	13.5	14.6	
Not stated	16.0	(B)	17.5	17.5	17.4	13.9	16.9	18.4	23.7	22.7	

"by" means "among" -- not 'distributed by'
 Source: 1998 US Statistical Abstract

08/13/2000

Ordered Relations Conditional Thinking

Rates

Non-standard usage of 'by'

16

No. 139. Age-Adjusted Death Rates, by Selected Causes: 1990 to 1996
[Rates per 100,000 population. For explanation of age-adjustment, see text, Section 2. The standard population is the total population of the United States enumerated in 1940. See also headnote, Table 138]

CAUSE OF DEATH	1990	1991	1992	1993	1994	1995
All causes	520.2	513.7	504.5	513.3	507.4	503.9
Major cardiovascular diseases	189.8	185.0	180.4	181.8	176.8	174.9
Diseases of heart	152.0	149.2	144.3	145.3	140.4	138.3
Rheumatic fever and rheumatic heart disease	1.5	1.4	1.3	1.3	1.2	1.1
Hypertensive heart disease	4.8	4.7	4.8	4.9	5.0	5.1
Hypertensive heart and renal disease	0.5	0.5	0.5	0.5	0.5	0.4
Ischemic heart disease	102.6	99.1	95.7	94.9	91.4	89.5
Other diseases of endocardium	2.5	2.5	2.6	2.6	2.6	2.6
Acute myocardial infarction	53.7	51.5	49.1	47.5	45.6	43.8
Old myocardial infarction and other	47.8	46.6	45.7	46.5	45.0	44.9
Hypertension	1.9	1.9	2.0	2.2	2.2	2.3

Source: 1998 US Statistical Abstract. Data for 1996 omitted to improve visibility of title.

08/13/2000

Ordered Relations Conditional Thinking

Arithmetic Comparisons

17

Arithmetic Comparisons	
DIFFERENCE (MORE THAN)	RATIO (AS MUCH AS)
3 is 1 more than 2 3 is 50% more than 2 3 is 0.5 times more than 2	3 is 1.5 times [as much as] 2 3 is 150% times [as much as] 2
Common Errors	3 is 150% more than 2 3 is 1.5 times more than 2 2 is 1.5 times less than 3