Comparing Ratios

Statistical Literacy 2009

Chapter 6 Overview

by

Milo Schield

www.StatLit.org/pdf/...

2009StatLitTextOverviewCh6.ppt

2009StatLitTextOverviewCh6.pdf

Ch 1. Review

Statistics are generally used as evidence to support an argument.

The influences on a statistic are of four kinds: Context, Assembly, Randomness or Error.

The Point or the Target

The more disputable the point,
the stronger the evidence must be.

Statistic As Evidence

"All Statistics are Socially Constructed"
So, "Take CARE"!!
Statistics may be influenced by:

C	A	R	E
Confounding	Assembly	Randomness	Error

Context and Ratios

Context: Related factors taken into account; the confounders not taken into account.

The easiest way to take into account a related factor are to make a comparison or to form a ratio.

Making a comparison of ratios takes into account two factors: size of a relevant basis for comparison and the sizes of the groups.

The English grammar involved gets very complex.

Three Topics

Percent Attributable: A common almost undetectable, form of comparison. Examples: deaths from second-hand smoke, obesity and radon.

Ratio Comparisons: Distinct Part vs. Common Part

DP: Widows are more likely among suicides than widowers.

CP: Widows are more likely to suicide than widowers.

Comparing ratios using *Likely* grammar

Inverse Percentages

Inverse percentages: Swap part & whole. Examples:

- *Most A are B* vs. *Most B are A.*
  "Most CEOs had a pet as a kid" versus "most kids who had a pet become CEOs."

- "The percentage of A who are B" versus "the percentage of B who are A."
  "The percentage of high school dropouts who smoked" versus "the percentage of high school smokers who dropped out."

Three Factor Percentages

Percentages with more than two factors get confusing.

1. The percentage of AB who are C
2. The percentage of A who are BC

AB is the intersection of A and B (e.g., adjective noun).

If there are differences, AB is less than A and BC is less than C.

So, #1 is always bigger than #2

1. Percentage of U.S. kids 15-24 who died from suicide
2. Percentage of deaths of U.S. kids 15-24 that were suicides

Three Factor Percentages

Unemployed black males:

1. the percentage of black males who are unemployed
2. the percentage of males who are unemployed blacks?

Seniors in nursing homes with no family.

1. 30% of seniors are in nursing homes with no family.
2. 30% of seniors with no family are in nursing homes.

Could both of these be true at the same time and place?

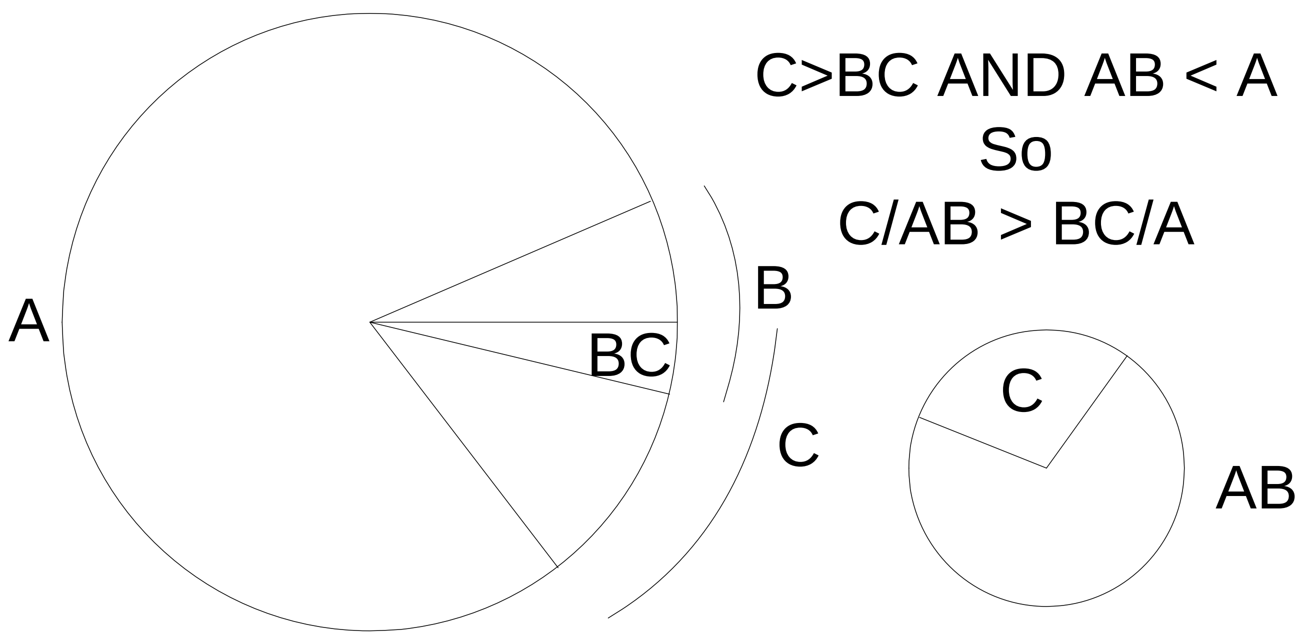
If one is wrong (percentage is too high) which one is it?

Three Factor Percentages

Graphical presentation:

Left circle: Percentage of A who are BC: (BC)/A.

Right circle: Percentage of AB who are C: C/(AB)



Assembly: Choice of Whole

Greenhouse Gas concentrations:

Atmospheric concentration (ppb) adjusted for heat capacity	Including Water Vapor	Excluding Water Vapor	Man-made as a % of each type	Type as a % of all manmade
Water Vapor	**95%**	-----	0.00%	0.40%
Carbon Dioxide (CO_2)	3.60%	**72.40%**	3.20%	**42.10%**
Methane (CH_4)	0.40%	7.10%	18.30%	23.70%
Nitrous Oxide (N_2O)	0.90%	19.00%	4.90%	16.90%
CFC's and misc. gases	0.10%	1.40%	**65.70%**	16.90%
ALL	100.00%	100.00%	0.28%	100.00%

Medical Tests

Rerun the O.J. Simpson trial: Suppose 1 chance in a million of an error -- a false match.

Quality/Accuracy: Percentage of outcome that test positive
Prediction: Percentage of positives that have outcome.

Adults in Area	Innocent	Guilty	Total
Test Negative	8,999,991	0	8,999,991
Test Positive	9	1	10
Total	9,000,000	1	9,000,001

Medical Tests: Four Outcomes

No Disease and negative test: OK, true negative

Diseased and positive test: OK, true positive.

No disease, positive test: False positive, false alarm.

Diseased, negative test: False negative, silent alarm.

	Diseased?	
Test Outcome	No	Yes
Negative	True negative	False Negative/Silence
Positive	False Positive/alarm	True Positive

Medical Test: Rare Disease

Given: 0.1% have disease and 95% test accuracy.

Population	DISEASED		
Test Result	No	Yes	Total
Negative	94,005	5	94,010
Positive	4,995 ③	95 ③	5,090 ④
Total	99,900	100 ②	100,000 ①

Medical Test: Common Disease

Given: 5% have disease and 95% test accuracy.
Confirmation: Test error = 100% - test accuracy.

Population	DISEASED		
Test Result	No	Yes	Total
Negative	90,250	250	90,500
Positive	4,750 ③	4,750 ③	9,500 ④
Total	95,000	5,000 ②	100,000 ①

If error = disease prevalence, prediction = 50%
If error > disease prevalence, prediction < 50%
If error < disease prevalence, prediction > 50%.

Summary

Context involves what is (not) taken into account.

What is taken into account can influence

- Counts or totals (by forming ratios)
- Averages (by selection or standardizing)

Part-whole ratios are one of the most common ways of taking into account a related factor.

Percentages have many uses – and misuses.