7/05/2006 1

Statistical Literacy Survey

MILO SCHIELD

Augsburg College Dept of Business Administration

Director,

W. M. Keck Statistical Literacy Project

www.StatLit.org **Schield@Augsburg.edu**

ICOTS-7

5 July 2006

2006SchieldICOTS6Up.pdf

7/05/2006 2

Conditional Probabilities

Thesis: It is **difficult** to decipher (much less to write) a description or comparison of rates or percentages as presented in tables and graphs.

Survey: To ascertain the ability of college-trained adults to decipher these statistics.

Subjects: college students (mostly working adults), professionals (mostly government data analysts) and college teachers (mostly IASE).

7/05/2006 3

Survey Subjects and Statistics Training

College students (85): Over half are working adults

Data Professionals (47): US Census Bureau and South African Statistical Service.

College Teachers (37): 14 US and 23 at ICOTS-6

STATISTICS TRAINING:

1+ courses: 78% of college teachers (87% of data analysts)

2+ courses: 29% of college teachers (34% of data analysts)

7/05/2006 4

Survey Subjects Comfort with Stats

"Very comfortable" dealing with *formal statistics*: sampling distributions, confidence intervals.

- 0% of students,
- 30% of data analysts and
- 57% of college professors.

"Very comfortable" dealing with *informal statistics*: rates and percentages in tables and graphs

- 7% of students,
- 62% of data analysts and
- 76% of college professors.

7/05/2006 5

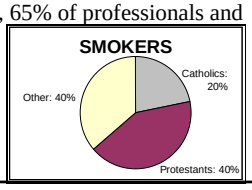
Pie Chart: Describe & Compare

20% of smokers are Catholic? [Yes]
19% student error rate ("No" or "I don't know").

Protestants (40%) are twice as likely ***to be smokers*** *as are Catholics (20%)*? [No, smoker is whole]

Error rate: 62% of students, 65% of professionals and 16% of college teachers.

P(Sm|Pr)/P(Sm|Ca) vs.
P(Pr|Sm)/P(Ca|Sm).

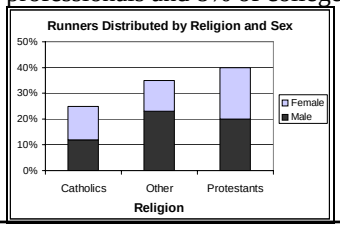


7/05/2006 6

100% Bar Chart: Describe & Compare

Q17. *20% of Protestant males are runners*? [No]

Error rates: 68% of college student, 53% of data professionals and 8% of college teachers.

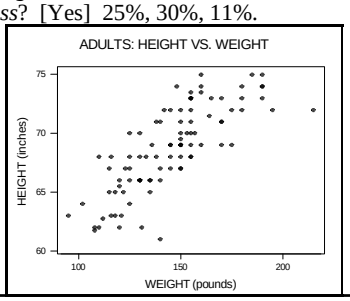


7/05/2006 7

Scatter Plot: Association/Causation

Q12. *Adults who weigh more tend to be taller than those who weigh less*? [Yes] 25%, 30%, 11%.

Q19. *If an adult increases their weight, they can expect to increase their height*? [No] 19%, 28%, 16%



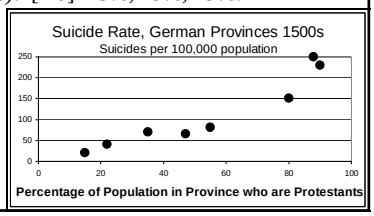
7/05/2006 8

Scatter Plot: the Ecological Fallacy

Q21. *As the percentage of Protestants increases, the suicide rate tends to increase*. [Yes] 34%, 21%, 32%

Q22. *Protestants are more likely to commit suicide than non-Protestants (are)*. [No] 45%, 49%, 49%.

Saying "Yes" to Q22 involves the 'ecological fallacy': going from groups to sub-groups.



7/05/2006 9

100% Row Table: Descriptions

Q23. *25% of females are blacks*? [No] 44%, 28%, 11%

Q25. *25% is the percentage of blacks among females*? [No] 38%, 28%, 16%.

	SEX		
RACE	Male	Female	TOTAL
Black	75%	25%	100%
White	50%	**50%**	100%
Other	40%	60%	100%
TOTAL	50%	50%	100%

7/05/2006 10

100% Row Table: Comparisons

Q27. *Whites are two times as likely to be female than are blacks*? [No] 60%, 53%, 57%.

Q28. *Females are two times as likely to be white as to be black*? [No] 44%, 38%, 19%.

Q29. *Whites are two times more likely to be female than are blacks*? [No] 65%, 49%, 46%.

	SEX		
RACE	Male	Female	TOTAL
Black	75%	25%	100%
White	50%	**50%**	100%
Other	40%	60%	100%
TOTAL	50%	50%	100%

7/05/2006 11

Two-Way Half Tables: Descriptions

20% of runners are female smokers? No: 55%, 53%, 30%

20% of females are runners who smoke? No: 53%, 55%, 32%

20% of female smokers are runners? Yes: 62%, 55%, 54%

20% of smokers are females who run? No: 42%, 36%, 27%

PERCENTAGE WHO ARE RUNNERS			
	Non-smoker	Smoker	Total
Female	50%	20%	40%
Male	25%	**10%**	20%
Total	37%	15%	30%

7/05/2006 12

Two-Way Half Tables: Comparisons

36. *The percentage of runners is twice as much among female smokers as among male smokers*? Yes: 42%, 47%, 46%

37. *The percentage of smokers who run is twice as much among females as among males*? Yes: 41%, 55%, 49%

PERCENTAGE WHO ARE RUNNERS			
	Non-smoker	Smoker	Total
Female	50%	20%	40%
Male	25%	**10%**	20%
Total	37%	15%	30%

7/05/2006 13

Simpson's Paradox

A research hospital had a higher death rate than a rural hospital. Each patient's condition was classified as either "poor" or "fair."

Q43. Is it possible that this research hospital had a lower death rate than this rural hospital for those patients in "poor" condition AND for those patients in "fair" condition?

Choice of answers: Yes, No, Don't know.

[Yes, possible.] Error rates: 44%, 68%, 41%.

7/05/2006 14

Multiple Half-Tables: Description

Assume, "In 1990" ahead of each statement:

Q45. *26.2% of blacks were smokers.* Yes: 60%, 43%, 19%

Q46. *26.2% of smokers were black.* No: 72%, 62%, 32%.

Table 3: Percentage of Smoking Prevalence

Year	All	Male	Female		White	Black
1955	--	56.9	28.4		--	--
1965	42.4	51.9	33.9		42.1	45.8
1980	33.2	37.6	29.3		32.9	36.9
1990	25.5	28.4	22.8		25.6	26.2

7/05/2006 15

Multiple Half-Tables: Description

Assume, "In US in 1996" ahead of each statement:

- *6% of low-weight births were in Calif.* No: 60%, 43%, 19%
- *6% of Calif. births were low-weight.* Yes: 39%, 36%, 11%

Percent of Births with Low Birth Weight			
State	1990	1995	1996
U.S.	7	7.3	7.4
Alabama (AL)	8.4	9	9.3
California (CA)	5.8	6.1	6

7/05/2006 16

Multiple Half-Tables: Comparison

Q52. *In the US in 1996, there were more low-weight births in Alabama (AL) than in California (CA).*

[No. No named ratio keyword] Errors: 66%, 68%, 30%.

Percent of Births with Low Birth Weight			
State	1990	1995	1996
U.S.	7	7.3	7.4
Alabama (AL)	8.4	9	9.3
California (CA)	5.8	6.1	6

7/05/2006 17

Multiple Half-Tables: Description

Q53. *10% of these women who received an HIV test were 40-44*? No: 78%, 55%, 19%.

Q54. *10% of these women 40 to 44 had an HIV test*?

Yes:
66%, 68%, 30%.

Table 5: Percent of Women, 15 to 44, who Received Selected Medical Services

Age	HIV	Pregnancy	Pap
15-19	14.6	16.1	33.5
20-24	20	27.4	68.7
25-29	25.6	25.3	70.9
30-34	18.5	17.4	69.5
35-39	14.2	8.1	62.9
40-44	10	4.3	62.7
ALL	17.3	16	61.9

7/05/2006 18

Multiple Half-Tables: Comparison

Q57. *Women 40-44 were twice as likely to have an HIV test as were women 20-24*? [Yes] 32%, 26%, 14%.

Q59. *Women 20-24 were two times more likely to have an HIV test than were women 40-44*?

[No]
82%, 60%, 81%.

Percent of Women, 15 to 44, who Received Selected Medical Services

Age	HIV	Pregnancy	Pap
15-19	14.6	16.1	33.5
20-24	20	27.4	68.7
25-29	25.6	25.3	70.9
30-34	18.5	17.4	69.5
35-39	14.2	8.1	62.9
40-44	10	4.3	62.7
ALL	17.3	16	61.9

7/05/2006 19

Survey Evaluation

Possible answers: Strongly agree, Moderately agree, Moderately disagree and Strongly disagree.

Q64. *This survey was much more difficult than I thought it would be*. 25%, **50%**, 20%, 5%

Q66. *This survey was unnecessarily tricky.* 24%, 27%, **36%**, 14%

Q68. *These tables and graphs are the kind I need or want to be able to read or understand.* **53%**, 37%, 7%, 4%.

Q69. *College students should be able to read these kinds of tables and graphs.* **63%**, 29%, 5%, 2%.

7/05/2006 20

Summary

92% of respondents *agreed* that "*College students should be able to read these tables and graphs.*"

75% of respondents *agreed* that "*This survey was much more difficult than I thought it would be.*"

In a pie chart,
- 19% of students misread a description
- 65% of professionals misread a "times as" compare.

In a 100% row table,
- 44% of students (28% of pros) misread a description.
- 46% of teachers misread "times more" compare

7/05/2006 21

StatLit Survey Error Rate

The average error rate was about
- 50% for college students,
- 45% for data analysts and
- 30% for college teachers.

Using data analysts' 80th percentile score (67% correct), the following reached that level:
- 5% of students,
- 20% of data analysts
- 45% of college teachers

7/05/2006 22

Conclusion

Describing and comparing rates & percentages is conditional probability in ordinary English.

Statistical educators will be seen as negligent if most of their students cannot read – much less write – descriptions & comparisons of rates & percentages as presented in tables and graphs.

Statistical educators should accept responsibility for teaching students how to read and write ordinary English descriptions and comparisons of rates and percentages as found in tables and graphs.

7/05/2006 23

Recommendations

Try out the simple 5-table survey on your students: www.StatLit.org/Survey. Paper copy available.

Try out the on-line grammar checker program. www.StatLit.org/RSVP.

Give your students a table or graph involving rates or percentages. Have them describe a single ratio (or compare two ratios) using ordinary English.

Try teaching this in your intro stats class.

7/05/2006 24

Related Articles at www.StatLit.org

Schield, Milo (2004). *Statistical Literacy and Liberal Education at Augsburg College*. AAC&U Peer Review. See www.StatLit.org/pdf/2004SchieldAACU.pdf.

Schield, Milo (2000). Difficulties in Describing and Comparing Rates and Percentages. *2000 ASA Section on Statistical Education*. P. 176. See www.StatLit.org/pdf/2000SchieldASA.pdf.

Schield, Milo (2001). Statistical Literacy: Reading Tables of Rates and Percentage. *ASA Proceedings of Statistical Education Section*. See www.StatLit.org/pdf/2001SchieldASA.pdf

Schield, Milo (2004). *Statistical Literacy Curriculum Design*. IASE. See www.StatLit.org/pdf/2004SchieldIASE.pdf.