Ten out of Nine Dentists Prefer Crest...

One of the benefits of a sound mathematical/statistical education is the ability to distinguish plausible quantitative information from that which is false or misleading. In today's information-overloaded world, information quantity often takes precedence over quality. As a result, unreliable and misinformed statements by the media, by politicians, and across the Internet are proliferating.

Sometimes claims are so patently bogus as to be quite humorous. Consider the following, taken from actual news accounts:

"Of those aged more than 60 living alone, 34% are women and only 15% are men."

"How much air is acceptable in ice cream? Rowcombe's is a lowish 35 percent; some manufacturers take theirs right up to 70 percent; a few even double the volume with 100 percent air."

"According to Lancaster Insurance, which arranges coverage for our club members, five out of four drivers between 17 and 21 have some sort of accident. The figure is correct because some have two accidents."

"Due to deforestation the rainfall (in the Peruvian rain forest) is now 120% less than 25 years ago."

"Only 25% of households consist of the classic couple with 2.4 children..."

"I think we finally have a poll without a margin of error." —Presidential candidate John McCain

Often the joke is only appreciated by the alert reader:

"Most women drive warily—thorough research shows that women have more accidents per mile (21.44 compared with 16.92 for men)."

"Tea accounts for 42 per cent of everything drunk in Britain, except tap water. Every man, woman and child over ten drinks 3.4 cups of tea per day."

MARK SCHILLING is Professor of Mathematics at California State University, Northridge.

Many errors undoubtedly slip by most mathematically unaware individuals:

"A total of 100 people were recruited during 1992, of which 19.8% were women."

"This ... is a busy motorway carrying 109,000 vehicles in a typical 24-hour period and 4,000 an hour during peak periods."

"Infant mortality (deaths at ages under one year, per 1000 births): 1948 — 26,766; 1994 — 3,979."

"Forty-two percent of all fatalities occurred on Friday, Saturday, and Sunday, apparently because of increased drinking on the weekends."

Sometimes one can only speculate about what the writer intended, as in the following cases:

"The figures were hailed by the Government as a major step towards its target that by 2002 some 80 percent of 11-year-olds should reach the average level in English and 75 percent in Mathematics."

Normally one expects approximately 50% of any group of test scores to fall above the average. It is possible to achieve the percentages given above, however, if the distributions of test scores in English and Mathematics have strong negative skewness—that is, if many of the low scores fall far below the average and most of the high scores fall just a little above average. Is that really what's happening here? It's much more likely that the "average level" refers to something other than the actual average score of 11-year-olds.

"In the study, men who began taking light exercise in their sixties reduced their chances of dying by about 45 per cent compared with those who stayed inactive."

Of course the chances of dying are the same for both groups—100%. The 45% reduction probably refers to a comparison of age-adjusted mortality rates. (The age-adjusted mortality rate is essentially a weighted average of the risk of death in one year, weighted by the number of individuals in each age group.)

Illustration by Loel Barr

"All over the world the key indicator to a man of a woman's fertility is the relationship of her hip measurement to that of her waist. A ratio of 0.7 is deemed ideal."

Certainly here the intended reference is to the ratio of waist circumference to hip circumference, not the implied inverse!

Some reporting errors are more subtle than others, and may require more careful thought or mathematical thinking to detect. Here are three cases for you to examine. Can you find the error(s) in each? (Answers below.)

(1) From a 1998 news article:

"It's Official: The Age of a Car Does Affect the Driver's Chance of Surviving a Serious Accident...",

which provided the following table as support:

"Ages of vehicles involved in fatal and serious accidents in the Newcastle area since 1992:

cars built:	
pre-1970	1.9%
1971–75	8.6%
1976–80	16.3%
1981–85	30.3%
1986–90	25.4%
1991–95	15.4%
1996 to date:	2.1%"

(2) And this one from the Reuters news agency:

"Republican George Ryan held on to a double-digit lead over Democratic Rep.Glenn Poshard in the Illinois governor's race, but his support had dwindled over the past two weeks...Ryan...led Poshard 48% to 37%...based on the Mason-Dixon Political/Media Research poll of 823 likely voters.... Two weeks ago, [the] poll showed Ryan leading Poshard 51% to 36%. The polls had a margin of error of 3.5 points."

(3) Finally, we have:

"Studies estimate that one in six adolescent girls has an eating disorder, but that adolescent gymnasts are more than twelve times as likely as non-gymnasts to have serious eating problems." [1]

(Assume that the entire statement refers to adolescent girls only and that the terms "eating disorder" and "serious eating problems" are synonymous.)

How did you do? The first case is fairly simple. The accident rates are low for cars built before 1981 because fewer of them remained on the road in the 1990s, while the drop in rates for the newest cars is explained by the fact that cars built subsequent to 1992 could not have had accidents during the first years of the study.

There are actually two errors in the second case, both stemming from the fact that the writer of the report has ignored his or her own last sentence. The percentages favoring each candidate are each subject to sampling error as well as to changes over time. That is, they represent only the views of those polled (probably around 1000 voters each time), not those of all likely voters. In each poll the two percentages given represent dependent binomial proportions, and the margin of error applies to each of these.

In particular the percentage of likely voters who favored Ryan in the first poll could easily have been anywhere in the range of 48% ± 3.5%, while Poshard was supported by around 37% ± 3.5% of likely voters. Furthermore, the 11% difference has an even greater margin of error, probably close

continued on p. 31

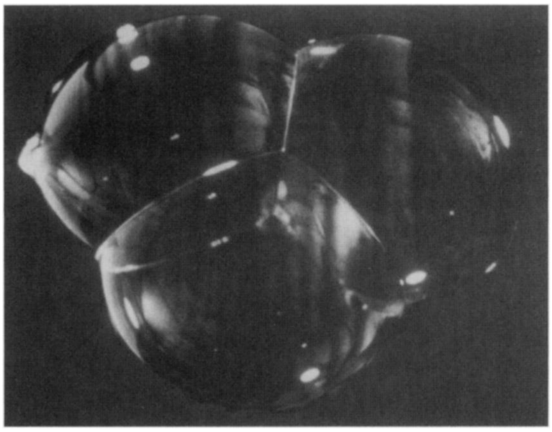

Figure 5. *It remains conjectural whether the standard triple bubble is the least area way to enclose and separate three given volumes. Photo by Jeremy Ackerman. Copyright Frank Morgan.*

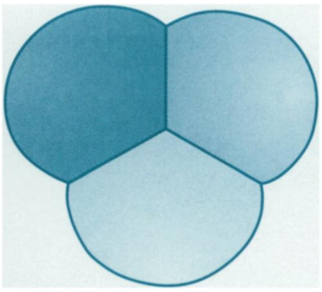

Figure 6. *Indeed, it remains conjectural whether the standard triple bubble is the least perimeter way to enclose and separate three given areas in R^2, even for equal areas.*

consisting of a central bubble with a symmetric toroidal innertube wrapped around it from top to bottom. For this case, rotate the left half of the bubble to the left and the right half to the right. The innertube gets fatter on the top and skinnier on the bottom. The total area goes down, proving that the original bubble was not area minimizing.

In general the argument is more difficult, as preserving the volume constraints requires two extra degrees of freedom, obtained by rotating four pieces as opposed to just two. The axis and the division into four pieces must be carefully chosen so that the rotations stretch the bubble rather than ripping it apart. One has to consider many cases, as suggested by Figure 4.

In an amazing postscript the 1999 Williams College SMALL Geometry Group consisting of undergraduates Ben Reichardt (Stanford), Cory Heilmann (Williams), Yuan Lai (MIT), and Anita Spielman (Williams), working with an early draft of the proof of the Double Bubble Conjecture in R^3, extended it to bubbles in R^4 and certain other higher dimensional cases. Faced with thousands of cases more complicated than Figure 4, they found a beautiful unified approach.

Just this summer, Andrew Cotton and David Freeman (Harvard) of the 2000 Geometry Group extended the proof to the case of equal volumes in hyperbolic space H^3 and to certain cases in the hypersphere S^3.

Looking for a challenge? It remains an open question whether the standard triple bubble of Figure 5 is the least area way to enclose and separate *three* given volumes in R^3. Even in R^2 it is not known whether the triple bubble of Figure 6 is the least-perimeter way to enclose three given areas, despite notable progress by a 1992 SMALL group of undergraduates: Christopher Cox, Lisa Harrison, Michael Hutchings, Susan Kim, Janette Light, Andrew Mauer, and Meg Tilton, who proved this in the category of *connected* regions.

For Further Reading

For more details, background, and references consult *Geometric Measure Theory: A Beginner's Guide, third ed.*, Frank Morgan, Academic Press, 2000, and watch for an upcoming article in *The American Mathematical Monthly*.

---

continued from p. 27

to 7%, due to the dependency between the proportions favoring each candidate. A difference in support of 11% ± 7% hardly supports a double-digit lead as a statement of fact.

The second error regarding the poll is the statement that Ryan's "support had dwindled over the past two weeks." Although more than likely that *is* the case, it could easily be true that his support had actually *increased* when one considers the margin of error.

The problem in the third situation is that while neither statement by itself is clearly erroneous, taken together they lead to an implausible conclusion. To see this, note that the proportion of girls who are gymnasts is certainly quite low. Therefore if the first claim above is true, approximately one in six adolescent girls who are *not* gymnasts has an eating disorder. Then the incidence among girls who *are* gymnasts must be about 200% or more in order for the second claim to hold.

The only way for both claims to hold is if an unacceptably large fraction of adolescent girls are gymnasts. Let p be the percentage of adolescent girls who are gymnasts, and let D_G and D_{NG} be the proportions of gymnasts and non-gymnasts with eating disorders, respectively. We have $pD_G + (1-p)D_{NG} = 1/6$ and $D_G = kD_{NG}$, with $k > 12$. Substituting D_G/k for D_{NG} in the first equation and solving for p yields

$$p = \frac{1}{6}\left(\frac{k/D_G - 6}{k-1}\right) \geq \frac{1}{6}\left(\frac{k-6}{k-1}\right)$$

since $D_G \leq 1$. From $k > 12$ it follows easily that $p > 1/11$. That is, more than one out of every eleven adolescent girls would have to be gymnasts.

Endnotes

Many of the excerpts above come (courtesy of the *Chance News* website at Dartmouth College, `http://www.dartmouth.edu/~chance`) from "Forsooth", a column in the Royal Statistical Society publication RSS News, which searches out such memorable quotes from the press.

References

1. Varda Burstyn, *The Rites of Men: Manhood, Politics, and the Culture of Sport*, Univ. of Toronto Press, 1999.