



# Glossary of Mathematical Mistakes

By Paul Cox

This is a list of mathematical mistakes made over and over by advertisers, the media, reporters, politicians, activists, and in general many non-math people. These come from many sources, which will appear in parenthesis. I will try to find an actual example of each for learning purposes. Note: In this document, I attack errors made by popular social organizations. I am not attacking their important causes, only their mathematical errors. I try to find errors from all political and social views. Any suggestions for better examples and new topics can be e-mailed to me: [paul@mathmistakes.com](mailto:paul@mathmistakes.com)

**Aftermath Counting** (A. K. Dewdney)- What the press has a tendency to do after a major disaster. It is the counting of known casualties from the police, paramedics, hospitals, and morgues, without considering duplication. Example: After the San Francisco Earthquake in 1989, reported deaths skyrocketed to 255 before finally settling on the number 64 dead. Interestingly, the Oklahoma City bombing of 1995 kept the body count at only the number of bodies actually recovered, though the press reported as many as 300 missing (the final tally was 163).

**"All Disasters Come in Threes" Conjecture"** - Also called the "All Celebrities Die in Threes" Conjecture. Essentially it is the mistake of grouping what is essentially a random occurrence. See Cancer Cluster Syndrome, "Shooting the Barn" Statistics.

**Astrology Amnesia** - When your astrological forecast comes true one day, you forget about the last three weeks when the forecast failed. Similar to Sample Trashing.

**Cancer Cluster Syndrome** - Making a lot of fuss over an above average number of cancer cases in a confined region. Note that all random functions have a tendency to cluster. For every reported cancer cluster there is also a cancer deficit region that does not get reported at all. This is not to say that all "cancer clusters" are just statistical abnormalities, there may be some toxic pollutant present in the area, but false reporting of cancer clusters is very common. See "Shooting the Barn" Statistics.

including interest for that same computer. That is over sixteen dollars more! If you transfer your credit card balance to a lower interest credit card there usually is this kind of surcharge, so run the numbers before you do it. Better yet, do not use a credit card if you can help it. A \$2500 computer only costs \$2500, if you pay it all up front. Watch out for "0% interest for 6 months" deals, also. Unless you can pay it all off within those 6 months, they will charge you back interest accumulated during those 6 months, usually at a high interest rate. You are better off charging with a bank card that charges interest from the start, than to go with one of these deals. Compare Ratiocinitis.

**Definition Errors-** A category of errors in which a mathematical term is misused or confused, possibly to mislead. The three most common definition errors are:

*Profit (or earnings)* - Could mean either gross profit or net profit, this is how "5 billion dollar" companies still have financial difficulty and have to reorganize, such as a former employer who shall remain nameless.

*Average* - There are three different kinds of averages. Given a sample set: 2,2,2,3,5,6,8: The mode average (the most common answer) is 2; the median average (the one in the middle) is 3; and the mean average (the sum of all samples divided by the number of samples) is 4. Any one of which can legitimately be the average.

Odds vs. Probability - Given a list of outcomes (rolling two dice) and a list of good outcomes (rolling a 7), you can calculate either the probability (good outcomes / all outcomes) to get 1/6. Or, you can calculate the odds (bad outcomes / good outcomes) to get 5:1. It is easy to get these two confused.

**Dimensional Demensia** - 1. The confusion of the significance of dimensions. 2. The attempt to compare two objects of different dimensions. A foot has 12 inches, but a square foot has 144 square inches.

Example: Here is a comparison of the world population in three different ways:

\* If every person in the world was lined up end to end, we would stretch four times longer than the orbit of the Moon around the Earth.

\* If we all lived in a city with the population density of New York City, that city would cover the state of Texas.

\* If every person was given a 20'x20'x20' apartment, the total volume would only fill the Grand Canyon half way.

Note that with each dimension added, the world population seems less significant. (Examples

predict past occurrences. It has been determined that these calculations failed to take into account the effect of sulfur particulates and other pollutants on global temperature. While the old factored models predict global warming of up to 8°F by 2040, these new predictive models show the globe heating up no more than 1°F by the year 2040. Turn on your Air Conditioners! (Note: The February 1994 Scientific American has an excellent article by Robert Charlson and Tom Wigley on this topic. Also, some have pointed out that I am oversimplifying the global warming phenomenon. Of course, I am. If you want the complete weather picture, see the May 1997 Scientific American.) Compare Factoritis.

**Factoritis - 1.** Taking into account factors that are not really relevant in order to inflate the numbers. AIDS activist groups regularly report that there are over a million Americans with HIV. Current statistics show that it is around 740,000 - including the approximate number that have never been tested, but are positive any ways. The only way to justify over one million people is to include those that died from the AIDS since 1980 when it was first discovered. While some may call this valid, it only looks impressive as a Raw Number. To use this number as a comparison figure, you have to include all of the people that have died from other diseases during the same time period. Stated another way: To figure out the percentage of the population with HIV we divide 740,000 (the number of HIV positives still alive) by 260,000,000 (the number of Americans still alive), resulting in 0.29% or about 1 in every 340 people. If we include the 360,000 people who have died of AIDS in the numerator, we have to include the tens of millions of people who have died from all causes during the same time period in the denominator. The resulting percent is actually lower when calculated this way.

2. Taking into account factors that were already calculated earlier. In 1990, the Department of Education reported that school expenditures have more than tripled since 1960. Some education lobbyists produced the figure that education spending has actually gone down on a per pupil basis when you figure in inflation. The problem is, the Department of Education already figured in inflation in their figures (but they left out per pupil figures, see the double exposure graph example below for details). The declining education statistic is the result of inflation being taken into account twice. Compare Factoredomy.

**False Positive Conjecture** - Tests with very high, but not perfect, accuracy may actually produce more false positives than true positives. Let us suppose that 3% of the population uses illegal drug X, Let us also suppose that a test has been developed that is 95% accurate in determining who has been using drug X. Say 1000 people took this test, 30 of them being users, Since the test is 95% accurate, 29 of these users gets caught (the other one gets a false negative). At the same time, of the 970 remaining, 48 also show positives, even though they are false. In other words, 78 people tested positive for use of drug X, but only 30 were true positives. Try this out on AIDS tests that are 99.7% accurate, or for more fun try polygraph (lie detector) tests that are only 80% accurate!

**Filter Counting** - An underestimate of reality caused by the deletion of important data. Similar

degrees of separation. A movie with that title has inspired a game called the Kevin Bacon game. Take any famous person, and you should be able connect them to actor Kevin Bacon in less than six moves. OK, lets try Elvis Presley, who was married to Priscilla Presley, who was in Naked Gun with Leslie Nielsen, who was in Airplane! with Lloyd Bridges, who was in Blown Away with Tommy Lee Jones, who was in JFK with Kevin Bacon. So, the next time the complete stranger sitting next to you turns out to have gone to high school with your brother in law's boss, it really is not that big of a coincidence. Compare Post Occurrence Miracle. The Readers Submitted Examples page has more on this topic.

**The Law of Averages thinking** - A belief by gamblers that the more often you win or lose the more likely your luck will change on the next try. If you flip a coin and it lands heads 10 times in a row, what are the odds that it will land heads on the 11th try? Answer 1:1. What about after 100 times in a row? Again it is 1:1. The odds are the same on each toss! The Readers Submitted Examples page has more on this topic.

**The Law of Zero Return** -  $\text{Return on Investment} = \text{Inflation} + \text{Taxes}$ . This topic was featured as a Mistake of the Month.

**Loaded Questionnaire** - Asking questions in such a way as to make the respondees feel foolish if they do not answer the questions the way the surveyor wants. Example: In 1995, the Corporation for Public Broadcasting surveyed the viewership of TV. The CPB had a certain interest in getting responses that favored their programming format so as to avoid federal budget cutters. One of many example questions:

"A recent study by a psychology professor at a leading university concluded that the amount of violence children see on television has an effect on their likelihood of being aggressive and committing crimes. From what you have seen or heard about this subject, do you agree strongly to that conclusion, agree somewhat, disagree somewhat, or disagree strongly?"

It is difficult to listen to such a question and not agree strongly. A more balanced question would include a contrary opinion, or better yet offer no opinions at all.

**Logical Fallacies** - If you want to deceive the majority of the people, use some of these in your arguments. (Note that these are used consistently by politicians, lawyers, and advertisers).

\* *Fallacy of Ambiguity* - Occurs when a word or phrase is used with one meaning in one premise, and with another meaning in another premise or conclusion. Example: People should do what is right + people have the right to disregard good advice = People should disregard good advice.

\* *Fallacy of Attacking the Messenger* - Attempting to discredit a message by discrediting the messenger. Example: He is innocent, the cop who arrested him is a racist and therefore must

**Lottery** - see Sucker Bet.

**Multiple Comparisons Fallacy** - (statistical epidemiology) Risk factor studies have a 5% chance of being too high and a 5% chance of being too low. Lets say a pre-election poll of 1,000 people shows candidate Smith with a 8% lead over Jones. If we did instead 20 polls of 50 people each, chances are at least one of those studies would show a slight lead in Jones' favor. In 1992, a Swedish study on the effects of power line radiation showed that children living close to power lines have a nearly four fold risk of childhood leukemia. But, upon closer examination, the Swedes did nearly 800 different studies in one. Other studies in the same report actually show a decrease of childhood leukemia from power line radiation. Studies with this many comparisons are not good for concrete results, they are best used to point out directions where future research should be done. (Frontline, "Currents of Fear")

**Num** (Dewdney)- A reported number with too few significant digits to be useful. These usually are round numbers like 1000 or 100,000. The term "six-figured salary" is an example, meaning any number between 100,000 and 999,999. Opposite of Dramadigit.

**Number Inflation** - 1. A gross overestimation or underestimation of reality. 2. A reporting of a statistic that is just not true. I used to call it "The Law of Five Times Reality" because of the tendency of political advocacy groups to over inflate their numbers by five times what is reality. Take the following examples:

**Situation Reported Actual**

AIDS Cases (1990) 1 million (ACT UP!) 200,000 (CDC)

Homeless (1990) 3 million (Mitch Snyder) 600,000 (Urban Institute)

Right to Life March (1989) 100,000 (Right to Life) 20,000 (Federal Park Police)

Spousal Abuse (1994) 6 million cases (NOW) 1.2 million cases (FBI)

Homosexuals (%) 10% (The Advocate) 2% (1993 sex survey)

It does not necessarily have to be five times reality. Besides reality is defined by the person doing the reporting. Some of these groups try to justify their numbers by taking in to account factors that should not be included (see Factoritis), or try to include numbers that we do not have enough information to make a good estimate (see Statistical Brick Wall). The Readers Submitted Examples page has more on this topic.

**Number Numbness** - The inability to fathom, compare, or appreciate really big numbers or really small numbers. Such as the difference between a million and a billion and a trillion. Politicians seem to make this mistake the most, noting the need to cut \$164 million from the National Endowment of the Arts, while insisting that it is important to spend 'a mere \$60 billion' on a missile defense system. The Readers Submitted Examples page has more on this topic. This topic was featured as a Mistake of the Month.

his wallet he says to himself, "Hey, I just increased my money by 50%." Later he discovers that he had a hole in his pocket where his \$5 bill was lost but thinks, "That's OK, first I gained 50% by finding the five dollars, and now I lost 33% by losing the five dollars, I am still 17% ahead." The Readers Submitted Examples page has more on this topic.

**Raw Number** - The reporting of an 'impressive' number that is meaningless without something to compare it to. Example: Everyday cars in America produce over a billion tons of pollution. Impressive? yes!, but let me state this statistic another way: Every day cars in America pollute 0.000000001% of our atmosphere. Impressive?, not really. Now this is a poor way to think about pollution, also. But, that is the problem with raw numbers: you have to compare them to something else to be meaningful. A better statistic: Cars are responsible for more than half of the carbon monoxide pollution.

**Recursive Arguments** - see Circular Reasoning

**Regression Jinx** - You probably have heard of the "sophomore jinx" popular in sports and music, a rookie athlete or musician performs poorly in their second year after an outstanding first year. Believe it or not, there is a mathematical explanation. When a star gains superstardom, it is partly because of talent, partly because of luck, or outside forces the star cannot control. When luck is extremely in the stars favor, it is only a matter of time that luck changes. Thus the "sophomore jinx" is not a jinx at all, but rather an expected outcome of statistical regression (the tendency of luck to move toward the norm). Alanis Morissette's first CD released in the US, sold 15 million copies. Despite critical acclaim, it is highly unlikely her second CD will sell even half that. An article in the March/April 1999 Skeptical Inquirer has more info and examples.

**Sample Occulting** - A disregard for an enormous sample resulting in coincidences seeming "supernatural", requiring an "occult" explanation when there really isn't a need. Example: An Advertisement for a TIME-LIFE book on unexplained phenomenon mentions a daughter who touches a hot stove and a mother 3000 miles away feeling pain in her arm at the same instance. Is the mother tele-empathic? Probably not! Consider how many times people touch hot surfaces, consider also how many times older people get mysterious pains. The likelihood of these two things happening at the same time to two people who are related are very good. Now if these events happened consistently to the same two people, and this could be replicated in a scientific experiment, then one might look to the supernatural. See also Post Occurrence Miracle.

**Sample Trashing** - Throwing out perfectly good data as "unreliable" because it goes against what the statistics are trying to prove. Popular with ESP believers who point to a few studies with positive results, and ignore the majority of the studies with negative results. The Readers Submitted Examples page has more on this topic.

**Selective Endpoints** - The reporting or graphing of a change in naturally random functions

hired compared to 65% of the male applicants. Who could argue with that? A week later ACME gets hit with a sex discrimination suit by one of the women that did not get hired, because overall 56% of the male applicants (14 of 25) got hired compared to 50% of the female applicants (3 of 6). See Ratiocinitis

**Statistical Brick Wall** -A number that cannot be verified, or accurately estimated, because the statistical data does not exist. A good example is the statistic of endangered species. Some biologists have estimated that over 10,000 species go extinct every year. Actual verified extinctions are around one species a year, including insects. These highly reported statistic tries to take into account the number of undiscovered species that go extinct, a number that is impossible to calculate. Sure there may be species that go extinct without anyone noticing, but the statistical data does not exist due to the fact that it is impossible to obtain. It could be anywhere from two to 10,000. No one can prove otherwise. Other major victims of the Statistical Brick Wall are studies that involve "ruling out" all possible factors. In 1994, a study was done to show how dangerous particulate pollution is, the result is that people who live in cities with high particulate pollution shorten their average life span by about 2 years. The study compared life spans in non-polluted regions (rural small towns) with high polluted regions (big cities), then they had to "rule out" other factors that might contribute to shorter life spans. They eliminated dozens of factors where statistics exist (i.e. violent crime), but then they were unable to rule out dozens of other factors (i.e. lifestyle differences, eating habits, exercise) because these statistics do not exist. In such cases statisticians either assume all remaining factors add up to zero, or they make an educated guess base on trends. While this does not invalidate the study, it makes such studies less than reliable. The Readers Submitted Examples page has more on this topic. This topic was featured as a Mistake of the Month.

**Statistical Rash** - A judgment based on statistical data that does not take into account all of the factors that cause the data to result as it did. Example: Here are some actual statistics on accident rates based on the speed at which the cars were driving:

20 mph or less	2.0%
20 to 30 mph	29.7%
30 to 40 mph	30.4%
40 to 50 mph	16.5%
50 to 60 mph	19.2%
over 60 mph	2.2%

It would seem to the casual observer that it is safer to speed than to travel at the speed limit. In fact, the reason that only 2.2% of the accidents happen at over 60 mph is because at any given time only about 2.2% of the cars on the road are traveling at over 60 mph. These statistics do not say anything about speed and accident rates, only about how fast the average car is traveling. (Example from Marilyn Vos Savant) The Readers Submitted Examples page has more on this