# Assessing Risks

Tom Carter

http://cogs.csustan.edu/~ tom/SFI-CSSS

Complex Systems

July, 2002

# Our general topics: $\leftarrow$

◎ Assessing risks

◎ Using Bayes' Theorem

◎ The 'doomsday argument'

◎ References

# The quotes  

- ◎ Science and wisdom
- ◎ Miracles

# Science and wisdom ←

"Science is organized knowledge. Wisdom is organized life."

- Immanuel Kant

"My own suspicion is that the universe is not only stranger than we suppose, but stranger than we can suppose."

- John Haldane

"Not everything that can be counted counts, and not everything that counts can be counted."

- Albert Einstein (1879-1955)

"The laws of probability, so true in general, so fallacious in particular ."

- Edward Gibbon

# Assessing risks ←

- The task of assessing risks in our lives is notoriously difficult. One thing we can try to do is calculate the probabilities of various events happening. Unfortunately, humans (even well trained scientists, mathematicians, and probabilists) often do a poor job of estimating probabilities.

  There is an apocryphal story of a statistician who always packed a bomb in his luggage when flying in a plane. When asked why, he explained that he knew that if the probability of there being one bomb on a plane was $\frac{1}{1000}$, then the probability of there being two bombs would be

  $$\frac{1}{1000} * \frac{1}{1000} = \frac{1}{1,000,000},$$

  and he felt much safer with the $\frac{1}{1,000,000}$ chance . . .

- Obviously there is something wrong with this statistician's calculation of probabilities.

  I sometimes wonder whether we in the US, holding on to our nuclear weapons, have fallen into a similar sort of confusion about calculating risks and "feeling safer."

- I won't go through much, but some probability basics, where $a$ and $b$ are events:
  $P(not\ a) = 1 - P(a)$.
  $P(a\ or\ b) = P(a) + P(b) - P(a\ \text{and}\ b)$.

  The probability of two events happening, $P(a\ \text{and}\ b)$ (often denoted by $P(a, b)$), can be quite difficult to calculate, since we often do not know how the events $a$ and $b$ are related to each other.

- Conditional probability:

  $P(a|b)$ is the probability of $a$, given that we know $b$. The joint probability of both $a$ and $b$ is given by:

  $$P(a, b) = P(a|b)P(b).$$

  Since $P(a, b) = P(b, a)$, we have Bayes' Theorem:

  $$P(a|b)P(b) = P(b|a)P(a),$$

  or

  $$P(a|b) = \frac{P(b|a)P(a)}{P(b)}.$$

- If two events $a$ and $b$ are such that

  $$P(a|b) = P(a),$$

  we say that the events $a$ and $b$ are *independent*.

- Note that if $a$ and $b$ are *independent*,

$$P(a|b) = P(a),$$

then from Bayes' Theorem, we will also have that

$$P(b|a) = P(b),$$

and therefore,

$$P(a, b) = P(a|b)P(b) = P(a)P(b).$$

This last equation is often taken as the definition of *independence*.


- We have in essence begun here the development of a mathematized methodology for drawing inferences about the world from uncertain knowledge.

# Miracles   ←

"The opposite of a correct statement is a false statement. The opposite of a profound truth may well be another profound truth."
- Niels Bohr (1885-1962)

"Groundless hope, like unconditional love, is the only kind worth having."
- John Perry Barlow

"There are only two ways to live your life. One is as though nothing is a miracle. The other is as though everything is a miracle."
- Albert Einstein (1879-1955)

"The Universe is full of magical things patiently waiting for our wits to grow sharper."
-Eden Phillpotts

"Nature uses only the longest threads to weave her patterns, so that each small piece of her fabric reveals the organization of the entire tapestry."

- Richard Feynman

# Using Bayes' Theorem ⟵

- A quick example:
  Suppose that you are asked by a friend to help them understand the results of a genetic screening test they have taken. They have been told that they have tested positive, and that the test is 99% accurate. What is the probability that they actually have the anomaly?

  You do some research, and find out that the test screens for a genetic anomaly that is believed to occur in one person out of 100,000 on average. The lab that does the tests guarantees that the test is 99% accurate. You push the question, and find that the lab says that one percent of the time, the test falsely reports the absence of the anomaly when it is there, and one percent of the time

the test falsely reports the presence of the anomaly when it is not there.

The test has come back positive for your friend. How worried should they be? Given this much information, what can you calculate as the probability they actually have the anomaly?

In general, there are four possible situations for an individual being tested:

1. Test positive (Tp), and have the anomaly (Ha).

2. Test negative (Tn), and don't have the anomaly (Na).

3. Test positive (Tp), and don't have the anomaly (Na).

4. Test negative (Tn), and have the anomaly (Ha).

We would like to calculate for our friend the probability they actually have the anomaly (Ha), given that they have tested positive (Tp):

$$P(Ha|Tp).$$

We can do this using Bayes' Theorem.

We can calculate:

$$P(Ha|Tp) \;=\; \frac{P(Tp|Ha) * P(Ha)}{P(Tp)}.$$

We need to figure out the three items on the right side of the equation. We can do this by using the information given.

Suppose the screening test was done on 10,000,000 people. Out of these $10^7$ people, we expect there to be $10^7/10^5 = 100$ people with the anomaly, and 9,999,900 people without the anomaly. According to the lab, we would expect the test results to be:

- Test positive (Tp), and have the anomaly (Ha):

$$0.99 * 100 = 99 \text{ people.}$$

- Test negative (Tn), and don't have the anomaly (Na):

$$0.99 * 9,999,900 = 9,899,901 \text{ people.}$$

- Test positive (Tp), and don't have the anomaly (Na):

$$0.01 * 9,999,900 = 99,999 \text{ people.}$$

- Test negative (Tn), and have the anomaly (Ha):

$$0.01 * 100 = 1 \text{ person.}$$

Now let's put the the pieces together:

$$P(Ha) = \frac{1}{100,000}$$

$$= 10^{-5}$$

$$P(Tp) = \frac{99 + 99,999}{10^7}$$

$$= \frac{100,098}{10^7}$$

$$= 0.0100098$$

$$P(Tp|Ha) = 0.99$$

Thus, our calculated probability that our friend actually has the anomaly is:

$$P(Ha|Tp) = \frac{P(Tp|Ha) * P(Ha)}{P(Tp)}$$

$$= \frac{0.99 * 10^{-5}}{0.0100098}$$

$$= \frac{9.9 * 10^{-6}}{1.00098 * 10^{-2}}$$

$$= 9.890307 * 10^{-4}$$

$$< 10^{-3}$$

In other words, our friend, who has tested *positive*, with a test that is 99% correct, has less that one chance in 1000 of actually having the anomaly!

- Some questions:

1. Are examples like this realistic? If not, why not?

2. What sorts of things could we do to improve our results?

3. Would it help to repeat the test? For example, if the probability of a false positive is 1 in 100, would that mean that the probability of two false positives on the same person would be 1 in 10,000 ($\frac{1}{100} * \frac{1}{100}$)? If not, why not?

4. In the case of a medical condition such as a genetic anomaly, it is likely that the test would not be applied randomly, but would only be ordered if there were other symptoms suggesting the anomaly. How would this affect the results?

# The 'doomsday argument'

←

- There is a line of reasoning called the "doomsday argument" (attributed to Brandon Carter in the 1980's) suggesting that we consistently underestimate the likelihood that the human race will end soon.

  What follows is a brief summary of the general theme of the argument.

- Suppose you have before you two urns, and are told that one contains ten balls labeled 1 through 10, and the other contains one thousand balls labeled 1 through 1000. You choose one of the urns at random. A ball is drawn at random from the urn you chose, and the ball drawn has on it the label '7'. What is

the probability that the urn you chose is the one with ten balls in it?

In the beginning, before drawing the ball labeled '7', we have

$$P(ten) = P(thousand) = \frac{1}{2}.$$

After drawing the ball, however, we can use Bayes' theorem to calculate:

$$
\begin{aligned}
P(ten \mid \text{draw '7'}) &= \frac{P(\text{draw '7'} \mid ten) * P(ten)}{P(\text{draw '7'})} \\
&= \frac{\frac{1}{10} * \frac{1}{2}}{\frac{1}{2} * \frac{1}{10} + \frac{1}{2} * \frac{1}{1000}} \\
&= \frac{\frac{1}{20}}{\frac{1}{20} + \frac{1}{2000}} \\
&= \frac{\frac{1}{20}}{\frac{101}{2000}} \\
&= \frac{2000}{2020} \\
&= 0.990099
\end{aligned}
$$

- Now the 'doomsday argument'
  . . . Consider the two possibilities:

  D: We humans destroy ourselves before we leave earth and colonize the universe.

  U: We colonize the universe.

  We now estimate: In the case 'D', no more than $100,000,000,000 = 10^{11}$ humans will ever live, and I am one of those $10^{11}$.

  In the case 'U', many more humans will live, say $10^{15}$.

  I know that I am among the first $10,000,000,000$ people to live, so in terms of human birth order, we can say that my 'label' is say $8,000,000,000$ (call this 'L8B').

We now use Bayes' theorem:

$$
\begin{aligned}
P(D \mid \text{L8B}) &= \frac{P(\text{L8B} \mid D) * P(D)}{P(\text{L8B})} \\
&= \frac{\frac{1}{10^{11}} * P(D)}{P(D) * 10^{-11} + (1 - P(D)) * 10^{-15}} \\
&= \frac{\frac{1}{10^{11}} * P(D)}{\frac{10^4 * P(D) + (1 - P(D))}{10^{15}}} \\
&= \frac{10^4 * P(D)}{10^4 * P(D) + (1 - P(D))}
\end{aligned}
$$

Now put in a value for $P(D)$, and see what happens:

If we start with an estimate $P(D) = \frac{1}{100}$, then

$$
\begin{aligned}
P(D \mid \text{L8B}) &= \frac{10^4 * 10^{-2}}{10^4 * 10^{-2} + (1 - 10^{-2})} \\
&= \frac{10^2}{10^2 + 0.99} \\
&= 0.990197
\end{aligned}
$$

- In other words, if our original estimate for the likelihood of 'doomsday' was $P(D) = \frac{1}{100}$, we should revise that estimate upward to 0.990197!

  You can try other values for the various pieces on your own . . .

# References

[1] Bostrom, Nick, *Existential Risks - Analyzing Human Extinction Scenarios and Related Hazards*, http://www.nickbostrom.com/existential/risks.html.

[2] Bostrom, Nick, *The Doomsday Argument: a Literature Review*, http://anthropic-principle.com/preprints/lit/index.html.

[3] Feller, W., *An Introduction to Probability Theory and Its Applications*, Wiley, New York,1957.

[4] Hamming, R. W., *Coding and information theory*, 2nd ed, Prentice-Hall, Englewood Cliffs, 1986.

[5] von Neumann, John, Probabilistic logic and the synthesis of reliable organisms from unreliable components, in *automata studies( Shanon,McCarthy eds)*, 1956 .

[6] Pierce, John R., *An Introduction to Information Theory − Symbols, Signals and Noise*, (second revised edition), Dover Publications, New York, 1980.