CHAPTER 5: ALL MODELS ARE FALSE. BUT SOME ARE USEFUL ANYWAYS.

5

MODELS

A newspaper account of a murder omits many
of the grisly details, but it is still informative.

Introduction

A useful fact about models is that they are all wrong in a strict sense, if only because they are incomplete. If a model is examined closely enough, it will invariably be found to differ from what it represents. Thus, a newspaper article about a murder tells you the victim's name, age and sex. It fortunately omits the exact locations of the knife wounds and the total volume of blood spilt on the floor, and omits details deliberately kept secret to aid in identifying the murderer. The list of ketchup ingredients tells you that the ketchup contains more tomatoes than sugar, but it conveniently overlooks certain details, such as the miscellaneous insect parts and feces in the bottle.

Never be surprised to learn that a model of interest to you is incomplete, hence is "false." Furthermore, it often takes very little effort to determine how a particular model is false. The reasoning we applied to a newspaper article and the ketchup ingredients could easily be repeated for the models given in the last chapter. Even with no previous exposure to the scientific method, it is usually easy to identify several important ways that any given model differs from what it represents.

The fact that a model is wrong does not mean it is useless. We continually use false models, such as the newspaper article and the ketchup ingredients discussed above. Because no model is one hundred percent correct, refusing to use a model merely because it is false is tantamount to refusing to use any models at all. The objective in using the scientific method is to distinguish useful models from useless ones.

Models are useful because they simplify- they are false for the same reason

The main reason that all models are incomplete/false is that they are simplifications -- shortcuts. The ways in which they are simplifications may not be essential for certain purposes (the simplifications may in fact make the model useful). The budget for a corporation, for example, only approximates the actual expenditures, income and profits. Clearly, no matter how many accountants a corporation hires, and no matter how carefully these accountants work, it is impossible to prepare a budget that is completely correct. To predict income exactly, one would have to know exactly how many gizmos the company will make and sell in the next year, and what price each one would fetch. This prediction depends on details of the economy, on how each prospective customer will behave in the coming year, and so forth. Because these facts cannot be known when the budget is prepared, it is inaccurate, or false. Their faults notwithstanding, budgets are also universally used. A budget allows a corporation to make decisions and to plan, and thus to achieve higher profits than would be possible without the budget.

Although every corporation acknowledges that a budget plan will not predict exactly the future financial exchanges (hence the budget is a false model), there are limits to what kinds of false statements a corporation will tolerate in its budget. The omission of some events, such as a minor unexpected price increase of raw material may be of no consequence. But other false aspects create more serious problems, as with a serious underestimate of production costs, or a failure to pay employees enough to avoid a strike.

As a second example, consider a court trial over an automobile accident in which one car rear-ended another. The trial is a model of the accident itself. It is incomplete because eyewitnesses forget and make mistakes, because medical diagnoses of whiplash can be in error, and because a picture of an intersection will invariably differ from the intersection itself. These problems notwithstanding, the judge and jury, after listening to the evidence presented in a trial, will usually have a pretty good idea of what happened. Hence the trial is a useful model.

"False" models as an integral part of science

The models that scientists use are no different from the models you use in everyday life. They are simultaneously false and useful. Learning even a small amount about scientific models can be quite useful in detecting major limitations of scientific approaches. This knowledge enables one to pose relevant questions to those who developed the model.

The Harvard food pyramid mentioned previously is useful as a model of a health-effective diet even though it condenses thousands of scientific studies about diet and health into a single picture (it is a summary model). This reduction must have resulted in the loss of substantial information, considering all the words, data, nuances and caveats in the original papers. Despite this, the food pyramid is an effective tool for communicating the results of a wide range of scientific studies to large numbers of people with varying backgrounds and levels of scientific sophistication. In fact, it is much more effective at this task than are the original scientific papers.

Biologists use animal models in developing new medical treatments (a type of physical model). Pharmaceutical manufacturers test the safety of new drugs using rats and mice before giving the drugs to humans, and heart surgeons develop new surgical techniques on dogs before trying them on humans. These models, however useful in preventing humans from taking unsafe drugs or having untried surgery performed on them, are nonetheless imperfect. Animals do not respond to drugs in exactly the same way humans do (think of how cats and humans respond to catnip). And while it might be useful for a heart surgeon to practice on a dog, performing surgery on a dog is clearly not the same as performing surgery on a human.

Pieces and Parts as Models

It seems fairly straightforward to consider two Suburbans off the same assembly line as models of each other, or the renovated Ft. Davis Historic Site as a model of the cavalry fort of the 1800s. Many of us also accept without question that a picture of Abe Lincoln is a model of him. All of these models are obviously false: the picture of Abe is not the man (it is a cluster of silver oxide grains on paper); the modern Ft. Davis is not populated with cavalry soldiers, nor is it concerned with Indian raids. One Suburban is probably a good model of the other for many purposes, but there are countless differences between the two when it comes to how tightly bolts are fastened, which parts will fail first, inherent weaknesses in the materials, and in exact gas mileage.

Consider now, an airliner crash. TWA flight 800 exploded in mid-air just off Long Island only a few days before the Atlanta Olympics. Eyewitnesses reported seeing trails of orange behind the plane, suggestive of a missile. Early speculations focused on a bomb, both because airliners don't just blow up in mid-air by themselves, and because the Atlanta Olympics provided the kind of public focus that terrorists often target. Our airports switched to tightened security measures, and President Clinton and Congress responded by passing expensive legislation to increase airport security. In the long run, we don't know what caused the crash, but odds seem to favor an explosion caused by equipment malfunction rather than a bomb.

What, then, are models of the cause of this crash?

1. We would certainly want to include reconstructions of this crash -- the assembled wreckage and any computer simulations of how the same kind of plane explodes from a bomb.

2. We should also include as a model the deliberate detonation of another plane, which could be studied to understand how a plane breaks apart in mid-air (such a deliberate detonation has been contemplated).

3. A single piece of wreckage could give the valuable clue of bomb residue or metal twisted in a particular way, diagnostic of what went wrong.

4. Eyewitness accounts of the crash.

5. Data recorders from the plane.

6. Knowledge of the sources of baggage put on the plane (were bags from other flights transferred to TWA 800?).

7. Even the timing of the accident with the Olympic games are all pieces of information or pieces of physical evidence that could shed light on the cause of this crash.

They are thus all models of the cause, even though some seem to be more "obvious" models than others. It may seem strange to call a single piece of wreckage or information about baggage as a model, in that these particular models are clearly only portions of the entirety of the crash. Yet any model has countless differences from what it represents. A reconstruction of the whole plane from the recovered wreckage may seem more complete than all the pieces, and it may be more complete in many senses. Nonetheless, even the entire assembled wreck is a far cry from the actual accident -- it is not in the air, flying; there are no passengers, nor is there any way to retrieve the lost lives, and so on. So it is misleading to suppose that a single piece of wreckage or bit of information is too insignificant to be a model but that the sum of these insignificant parts is a legitimate model.

The important issue here is that some models are more USEFUL than others. One piece of wreckage may be more useful than a 1000 other pieces in understanding whether a bomb went off. What makes a model useful is explained next.

ACU: Why we use particular models (Accuracy, Convenience, Uniformity)

The goal determines model usefulness. Models vary in their usefulness. Some are so different from what they represent that we just refuse to use them. Yet there is no such thing as an intrinsically good or bad model without considering its context. A model is judged against the goal, and a model may be good for some purposes and bad for others. Consequently, the standards we use to decide the utility of false models are extremely diverse. An algebra problem that your math instructor assigns is a model of the problems you will be asked to solve in some careers. Because your future boss will never ask you to work a problem exactly like those at the end of the chapters in your algebra book, each problem you work in class is a false model of this future need. The relevant question is not whether the model is false, but whether the false aspects of the model seriously degrade its usefulness. The answer to this question will vary depending on how you use the model. Thus an algebra course might be usefully false model for accountants, business managers and engineers, but a hopelessly false model for artists.

More generally, the usefulness of a model depends on the problem to which it is applied (the goal of the work). Thus, any model may be useful for some purposes, and it will invariably be useless for other purposes. When considering the value of a model, it is therefore essential to know its application. In many cases, this point is obvious - a nuclear physicist would not be the least interested in using GM's annual budget model to predict the behavior of elementary particles. However, the match between model and goal applies on a much finer scale as well. For example, the file of previous exams owned by a fraternity may be very useful for some classes but not others.

The criteria for model acceptability can be classified in many ways. Here we will recognize 3 criteria: Accuracy, Convenience, and Uniformity, or ACU. Acceptance of a model depends on a combination of all 3 criteria, though there is no universal rule for assessing the relative benefits of one criterion versus the others. And we are not looking for one model that simultaneously best satisfies all three criteria. In fact, we often use several different models of any one thing to overcome the limitations of any single model -- there is NOT a most useful model for any particular goal.

Accuracy: This is the most obvious criterion to use in accepting or rejecting a model. After all, if we are trying to represent something, we hope that our model does a good job of actually representing what is intended. Accuracy is the measure of how well the results from the model will enable us to predict the real situation. This criterion is thus easy to grasp, and we will move on to the next criterion.

Convenience: This third criterion covers time, cost, ease of application, and ethics. In an ideal world, we might imagine that cost is no problem. Yet, budgets dictate that we make the best use of the money. And time constraints dictate that we get answers soon as opposed to later. We use mice instead of monkeys for initial tests of foods and drugs because mince are more convenient than monkeys (in cost, time to results, and ethics). Virtually all models used to test products on a large scale are chosen with a heavy emphasis on convenience. Some such models seem ridiculous because so much accuracy has been sacrificed in favor of convenience (e.g., condom testing). But at some level, most models sacrifice accuracy to achieve convenience.

Uniformity: This criterion is the consistency of the model -- is the model uniform from one use to the next? Uniformity is important mostly with physical models instead of abstract models. (It is very important in sampling models, because sampling models usually involve physical models). Inbred strains of mice offer good models to study cancer-causing agents in humans because they possess uniformity -- thousands of mice of the same genotype can be tested, so that results can be compared for different chemicals. Likewise, methods for industrial testing of products are geared toward uniformity, because the test will be performed many times and the outcomes from different trails need to be comparable across the trials.

These are not the only criteria that are relevant to a model. Another factor is the repeatability of a model (can it be applied multiple times?) Attempts to understand unique events after-the-fact are often based on models lacking repeatability. Wreckage of a plane crash provides various models of the crash -- pieces of the plane, for example -- but these models lack or are weak in repeatability. Eyewitness accounts also lack this property, and as such, are limited in an important respect. However, we will limit ourselves here to the three criteria of Accuracy, Convenience, and Uniformity.

An example: Consider a detailed example of the conflict between accuracy and convenience. If our goal is to understand cancer in humans, we might use genetic studies of of humans, monkeys, rodents, yeast, and/or bacteria. How do each of these models rank on the scales of accuracy and convenience?

MODEL	ACCURACY RANK	CONVENIENCE RANK
Humans	1	5
Monkeys	2	4
Rodents	3	3
Yeast	4	2
Bacteria	5	1

There is an inverse relationship between accuracy and convenience for these models. All of these model organisms are useful -- yeast and bacteria might be the most useful for some purposes because they are cheap, easy to manipulate, and do not raise ethical issues in experimentation (strong on convenience). The genetics of yeast is similar enough to that of humans (with respect to the control of cell division) that many breakthroughs in cancer research have come from them. Obviously, model accuracy is greatest with humans, monkeys next, and so on. So we can have very useful (convenient) models that are much less accurate than other models.

In research with plants, animals, and even yeast and bacteria, especially when treating them as models of humans, there is extraordinary emphasis on using strains. Why? The reason for doing an entire study with a single strain of an organism is to achieve uniformity. If we are investigating the effect of a chemical on the mutation rate of yeast, we want to minimize the variation due to causes other than from that chemical, so we use genetically uniform strains.

Model Incompleteness is the Basis for Improvement

Most models deemed useful at some point in history have only a temporary life, i.e., they are replaced by better ones after awhile (returning to the point that science does not prove models to be true). For example, today's technology enables companies to assess financial status faster and more accurately than in the past, so that their budgets incorporate different components now that 20 years ago. In science, most successful models are improved through time as well. The better models simply address and overcome some of the falsity or incompleteness in their predecessors.

One can often anticipate how a model may eventually be improved merely by contemplating how it is incomplete. The goal is not to eliminate all incompleteness in a model, but rather to correct the more serious limitations. It is often possible to anticipate how a model may ultimately be rendered obsolete merely by thinking about its limitations. In teaching you to think about models, we will therefore emphasize their limitations. Table 5.1 illustrates how some common models can be trivially wrong versus seriously wrong.

HOW SOME MODELS CAN BE FALSE (TABLE 5.1)		
MODEL	**MINOR FLAW**	**MAJOR FLAW**
Medical diagnosis	No two patients are identical, although most individual details do not affect treatment	Incorrect diagnosis results in patient death or malpractice
Credit rating	Small details of personal finances are omitted	Omission of major debts or credits that have a big impact on personal finances
Test for heroin	The test assays various chemicals other than heroin itself, but these compounds are minor constituents in the body	Eating poppy seeds before the test gives the impression that you have illegally taken drugs
Income tax return	Small transactions are overlooked	Omission of large sources of income which can result in a financial penalty or incarceration
College exam score	A lucky guess results in a few points on a subject that the student was not prepared for	Exam score is totaled incorrectly
New car	Slight differences exist between different cars of the same model	he car you purchase is a lemon
Space shuttle flight	Each flight faces different problems, which are usually fixable	The shuttle explodes

A Template for Models

To facilitate use of the information in this and the last chapter, we offer a template for models. This template is intended to help you think about the different aspects of models whenever you read (or think) about uses of the scientific method. Any news article describing medical research or a news article on studies into business practices can be analyzed from this perspective. In the next few chapters, we will describe some biological examples primarily from the perspective of models, and this template will be used in summarizing those presentations.

MODEL TEMPLATE
MODEL
KIND (Abstract, Physical, or Sampling)
APPLICATION (used as what?)
STATUS (Accepted, Rejected, or Uncertain)
LIMITATIONS

The kind of model is either abstract, sampling, or physical; if a model does not fall into any of these 3 classes, we will not worry about the class (there are too many ways to classify models for us to bother classifying all of them). The application is the problem for which the model is used. For example, a business (financial) plan is applied to managing company money, a monkey might be used as a model of humans in understanding AIDS, and so forth. The status of a model indicates whether it is currently regarded as useful (accepted), rejected, or is in dispute (undecided). For example, the model that X-rays directly cause bacterial mutation would be accepted for the set of experiments in which irradiation of bacteria leads to mutation but would be rejected for the experiments in which irradiation of Petri dishes alone leads to mutation.

The last item, limitations, is of interest only for models whose status is accepted or undecided. It is important to realize that all models have limitations and that these limitations may ultimately lead to the model's rejection when we have better data. By highlighting limitations of currently-accepted models, we should be constantly alert to possible revisions of the model that may be even more useful.

All Models Must be Refutable

One of the most widely publicized features of the scientific method is that scientific models must be refutable or falsifiable. By this it is meant that observations can be imagined that would cause the model to be rejected. The criterion of falsifiability differs from our claim that all models are false. Typical examples of unfalsifiable models are of the form that various demons or spirits control all events in the world, or that some person has mystical properties. These models are not falsifiable, because it is not possible to even imagine data which would call for their rejection.

How does falsifiability fit into our framework? The goal of any application of the scientific method invariably involves predicting the unknown (explaining future results) or manipulating the future (increasing profits). But because an unfalsifiable model admits all possible outcomes (since nothing is inconsistent with it), an unfalsifiable model cannot be improved upon. Consequently, unfalsifiable models are useless.

Unfalsifiable models are common outside of science. Some market analysts have a flair for "explaining" the ups and downs of the market after-the-fact. Regardless of the market changes, these nightly reports profess to account for all the ups and downs, and there is no pattern that can't be explained. Of course, as long as these reports don't attempt to forecast the directions in the future, they can never be challenged (of course, some analysts do predict future trends). Prophecies also tend to lack falsifiability. These statements are often couched in extremely vague terms, and only in retrospect are people able to "interpret" them to make sense. To be falsifiable, a prophecy needs to be specific enough that we know in advance what to expect (e.g., the world will collide with a large comet on the morning of your 3rd exam).

History as a Model of the Past

A senior (and now deceased) colleague of ours once commented that the subject of history became ever more interesting the closer one got to becoming a part of it. There seems to be some truth in that, at least as judged by the general lack of enthusiasm for the subject in high school and college. Nonetheless, there are several human endeavors that involve attempting to reconstruct the past. Aside from history per se, they include:

1. many sciences (evolution, astronomy, geology, anthropology)

2. criminal trials (reconstructing a crime to convince a jury that the defendant committed it)

3. religion (many teachings refer to events from the past)

4. medicine (who you contacted that gave you this disease, what you ate that led to your heart condition)

Reconstructing the past poses special problems to a scientific approach, because the past is gone and it is unique. This means we cannot study it directly, the same way we would study a blood sample or your computer software problem. However, we nonetheless reconstruct history, and a variety of models is used in doing so.

Consider "Billy the Kid." Virtually everyone knows this name – a notorious, young outlaw who killed several people in the late 1800s, mostly during the Lincoln County wars in New Mexico. Our model of Billy the Kid is a list of episodes and deeds, derived from legal documents, testimonies, newspaper accounts, diaries and the like from the 1800s: his name (William Bonney), birth (1859, New York City), death (1881, shot by Pat Garrett, buried in Ft. Sumner), along with twenty one murders, escapes from jails, associations with other outlaws, and raids he conducted.

Each of these descriptors is a potentially false model of Billy the Kid. Doubts have been raised that William Bonney was actually killed by Pat Garrett; rather it has been suggested that Pat Garrett shot someone else. And, it is of course possible that some of the murders attributed to Billy the Kid were committed by someone else. It is not uncommon that notoriety begets further notoriety, whether deserved or not. (The infamous "serial killer" Henry Lee Lucas confessed to over 600 murders and had law enforcement agents accepting over 100 of the confessions, but nearly all of those were later shown to be false – Lucas may have killed as few as 3 people – and his death sentence in Texas was commuted to life in prison because he could not possibly have committed the murder for which he was convicted.)

Interest in Billy the Kid is inspired by a current controversy. If you drive the 100+ miles north of Austin to the Hill Country town of Hico, you will find a quaint, aging downtown area of antique shops in sight of a stately but defunct stone cotton gin, harkening back to the days when cotton was a major crop in the area. (There is also a great chocolate shop nearby.) One of the old stores harbors a museum of sorts, dedicated to keeping alive the possibility that the real Billy the Kid died there in 1950 at the age of 90. The person's name was Brushy Bill Roberts, born in Buffalo Gap, TX in 1859. In his final year of life, he attempted to establish his identify as Billy the Kid, dictating the details of his acts and deeds, even requesting a pardon from the Governor of New Mexico (at which time he had suffered a stroke and could not communicate).

Brushy Bill Roberts and William H. Bonney were clearly different people. Who, then, was the real Billy the Kid? We will likely never have a satisfactory answer, at least not an answer that everyone can agree upon. There are many ways in which our Y2K (year-2000) model of Billy the Kid may be false – we may be associating the wrong name with the deeds, the deeds themselves may be different from legend, and the 21 murders may have been committed by more than one person. The legend of Billy the Kid is a model that may have little bearing on reality, or it may be accurate in some respects but not others.

The New Mexico town of Ft. Sumner, located at a crossroads along the Pecos River in eastern New Mexico, claims to be the final resting spot of Billy the Kid. Billy the Kid's grave is probably the main tourist attraction in this small town. At least a few people in this town are annoyed at Hico's insistence that the real Billy the Kid is buried nearby in Hamilton, TX. The dispute has escalated to the point that Ft. Sumner wants testing to establish that the bones in their grave have DNA related to the DNA in the bones of William Bonney's mother (whose grave is known). This test may indeed help establish that the body in the Ft. Sumner grave is that of William Bonney, but it will not resolve the more important issue of who did the killings.

Summary: Main Points About Models

To wrap up these two chapters on models, we offer a list of points that you should understand now. Refer back to earlier text for elaborations.

1. Models are shortcuts; they simplify

2. Models exist in at least 3 general classes: physical, abstract, and sampling

3. Any one thing can be represented many models; conversely, one model may represent many things

4. All models are false

5. Model usefulness is judged by accuracy, convenience, and uniformity

6. Pieces and parts are models of the whole

7. Models must be refutable