Fundamentals

Concepts for how evidence and knowledge increase

Version 0.1

Holger Sennhenn-Reulen

Forest Inventory, Informatics & Biometrics,
Department of Growth and Yield,
Northwest German Forest Research Institute

Göttingen, June 24, 2020

Contents

I should stress that knowledge isn't something out there for us to discover. Instead, knowledge is made.

Dammann (2018)

The art of probabilistic modeling is to describe in a mathematical form (model and prior distributions) what we already know and what we don't know.

(Vehtari, 2020, Slides `bayes_intro.pdf`)

Bayesian inference is the process of fitting a probability model to a set of data and summarizing the result by a probability distribution on the parameters of the model and on unobserved quantities such as predictions for new observations.

(Gelman et al., 2014, page 1)

[...] science appears to work in accordance with Bayesian principles. At each stage in the development of a scientific study new information is used to adjust old information. [...], this is how Bayesian modeling works. A posterior distribution created from the mixing of the model likelihood (derived from the model data) and a prior distribution (outside information we use to adjust the observed data) may itself be used as a prior for yet another enhanced model. New information is continually being used in models over time to advance yet newer models. This is the nature of scientific discovery. Yet, even if we think of a model in isolation from later models, scientists always bring their own perspectives into the creation of a model on the basis of previous studies or from their own experience in dealing with the study data. Models are not built independently of the context, so bringing in outside prior information to the study data is not unusual or overly subjective. Frequentist statisticians choose the data and predictors used to study some variable – most of the time based on their own backgrounds and external studies. Bayesians just make the process more explicit.

Hilbe et al. (2017, page xiii)

"Um Daten zu interpretieren, entwickelt man anhand des Vorwissens ..."

Sachs: Statistische Methoden 2: Planung und Auswertung, page 18

Knowledge through observation

Empirical research and *statistical reasoning* is based on the goal of accessing a (partially) unknown process by generating observations, and, by the use of these, gaining (further) insights into the process. Such a (partially) unknown process that can't be examined directly is called a *data-generating mechanism* and by using statistical methods we take an action that is very similar to *reverse development*, *reverse engineering* or the *reconstruction* of a production process.

> [Statistical] "Models produce values for things we can measure in the real world; we call such measured values *observations*."
>
> van de Meent et al. (2018)

Statistical reasoning is the craft of performing reverse engineering and is fundamentally different to *direct evidence*[1].

Direct evidence

Direct evidence by rules of logic is how philosophers, mathematicians, …, work in order to push the limits of their fields. Figure 1 shows an overly simple example for the application of these rules: Following two premises, the conclusion is a direct consequence of these in conjunction with the rules of logic.

Premise i: All trees grow towards the light. *Premise ii*: Oaks are trees.

Conclusion: All oaks grow towards the light.

Figure 1: Example for direct evidence.

In mathematics, some statement – for example *The number of prime numbers is infinite* – is proofed by rules of logic. Rules of logic also have a direct application in modeling, namely in the class of *forward models* (Gunawardena, 2014).

> "We can distinguish two kinds of modeling strategy in the current literature. We can call them *forward* and *reverse modeling*. Reverse modeling starts from experimental data and seeks potential causalities suggested by the correlations in the data, captured in the structure of a mathematical model. Forward modeling starts from known, or suspected, causalities, expressed in the form of a model, from which predictions are made about what to expect."
>
> Gunawardena (2014)

Gunawardena (2014) denotes the class of forward models in his remaining text as *mathematical model*:

[1] It hopefully makes sense to spend a couple of minutes on this alternative – in order to being able to differentiate the one the other – before we will finally begin on page 6 to concentrating on how we can learn by empirical evidence.

> "A mathematical model is a logical machine for converting assumptions into conclusions. If the model is correct and we believe its assumptions then we must, as a matter of logic, believe its conclusions."
>
> Gunawardena (2014)

Such a mathematical model is a *deterministic* system:

> "*Determinism* is usually shorthand for causal determinism: Every event has a sufficient cause. More precisely, given conditions $\{x_1, x_2, x_3, \ldots, x_i\}$ and event C, effect E must occur. Every event in a deterministic system is covered by such a conditional. Given the antecedent conditions and cause C, there is one unique E possible. Nothing is left to chance and nothing escapes the network of causes and effects, [...]."
>
> (Koperski, 2020, section 6.1)

The what-happens-next machine

Watch the following video (stop at 1 minute 52 seconds!) where *Kermit* presents his *what-happens-next* machine:

`https://www.youtube.com/watch?v=cog2a3YeDMM`

During these first 112 seconds of the video, Kermit guides us through the build-up of his machine and does *forward engineering*: He walks us through a sequential process which, as the final *outcome*, turns a switch that starts a radio playing music. Kermit is convinced by his invention. He says: "Now, I am turning on the radio using the principle of what-happens-next." He advertises his machine and makes us belief that it is able to fulfill a very complex task but is remains a completely *deterministic* procedure $f(x)$ for one *input* variable x measuring the state of the rope. Input variable x here measures the status of the event, it is an event indicator that can take on only two possible values which are the elements of set $\{\text{uncut}, \text{cut}\}$. What Kermit told us is that while $x = \{\text{uncut}\}$, the outcome Y will remain on the value $Y = \{\text{radio off}\}$ – this status of x and Y in combination with the whole build-up of the machine are the conditions. As soon as $x = \text{cut}$ – the event – the machine starts running and, by the principle of what-happens-next, will change the outcome to $Y = \{\text{radio on}\}$ – this is the effect. In mathematical notation we can define the following functional relationship:

$$Y = f(x) = \begin{cases} \{\text{radio off}\}, & \text{when } x = \{\text{uncut}\}, \\ \{\text{radio on}\}, & \text{when } x = \{\text{cut}\}. \end{cases}$$

When asking my 5 year old daughter what she think might happen when Kermit cuts the rope, she said: "The machine will work and the radio will begin to play."

Trees grow towards the sun

In the year 1999, artist-experimenter Natalie Jeremijenko installed six live trees bottom-up in her work *tree logic*.

> The familiar, almost iconic shape of the tree in nature is the result of gravitropic and phototropic responses: The tree grows away from the earth and towards the sun.
>
> Sullivan (2005)

Conditionals: Trees need light in order to live. The only reliable source of light is the sun. A tree can't move geographically.
Event: Surrounding vegetation start to build competitive pressure with respect to this resource.
Effect: A trees main direction of growth is towards the sun.

From direct to empirical evidence

Think of a situation where you have full confidence that a theory is true: there is not even the tiniest 'chance' that you might possibly be wrong. Such a situation would be equivalent to an event which you would give the attribute 'this is to 100% sure'. This is a colloquially formulated *probability statement* and from basic rules of probability, we know that the contrary here then must have the attribute 'there is no possibility of this contrary to happen', which is again a colloquial formulation for 'there is 0% probability for the contrary event to happen'. Now think that you are able to measure a process underlying your theory such that you can generate observations that base on your theory, and which can raise evidence that your theory might be false. For example the theory 'trees grow towards the sun': you could take a tree and measure it's height at two different points in time. This would generate an observation that links to your theory. Without any physical or other damage/illness and sufficient time – with respet to rather 'slow' tree growth – between the two time points, the value at the second time-point will be a larger tree height in comparison to the first measure. This means that these two measurements generate no evidence against your theory. Now think of the contrary to your theory, which is 'tree grows towards the earth or they might stop growing at some time or for some time'. Now think again of two measurements, but now you give yourself only a short time between the two measurements or you are able to measure only on a very coarse measurement scale, such as in steps of 0.1 meters. As an outcome, let's assume we that we get two identical height measurements. What does this mean to you? Is this again evidence for your theory? It is not! What if not? Is this evidence against your theory and for the contrary? Of course, we could make a personal judgement and state that 'well, there was obviously not enough time for the slow process to result in a measurable evidence for my theory'. But, the the simple algebraic answer is: it doesn't matter(!), given your 'I am a 100% sure' statement. As you have 0 probability that the contrary is right, any new evidence of any leverage to convince yourself about the opposite of your theory can't be strong enough to let you move away from 0 as 0 times anything will always stay 0. So you are trapped in your own personal too strong belief in your theory. This is *Cromwell's Rule* by Dennis Lindley and it is most beautifully described in Section 6.8 of (Lindley, 2006), a book I can't recommend enough. Think of another theory which might be just a little bit more controversial, such as 'a mixed stand with oak and beech is more productive than the same area with half of the stand stocking with pure beech, and the other with pure oak'. What would you do? Again make a 'I am a 100% sure' statement? The point I want to raise here: if there is the smallest possibility that your theory might be wrong[2] then you should never make a statement like 'it is for 100% sure that this theory is true'. Always leave some tiny little room such that any new and

[2] Which is the case in any scientific endeavour, as any different was a waste of (public) money.

very strong evidence is able to convince you from the opposite. However, with this you are leaving direct logic and direct evidence towards working with measurements in order to (re-)evaluate your beliefs about theories. And this means you need to embed your work in the environment of *empirical evidence* making.

Empirical evidence

> Data are used mainly as raw material for information generation. When these data are put into context, they yield information that may be useful as evidence. Based on such evidence, knowledge is generated. Knowledge is evidence-based belief that is predictive, testable, and consistently successful, as judged by consensus among stakeholders.
>
> Dammann (2018)

Contrary to a deterministic system – where effects directly follow from conditions and events by the the application of rules of logic – is a *probabilistic system* where the *empirical evidence* must be scraped out of an environment involving *chance*.

> There are some things that you, [...], know to be true, and others that you know to be false; yet, despite this extensive knowledge that you have, there remain many things whose truth or falsity is not known to you. We say that you are uncertain about them. You are uncertain, to varying degrees, about everything in the future; much of the past is hidden from you; and there is a lot of the present about which you do not have full information. Uncertainty is everywhere and you cannot escape from it.
>
> Lindley (2006)

Again, following the protocol of Dammann (2018), the idea is as follows: We generate one[3] to many observations of an outcome that we classify as a *random variable* and, based on those observations, we try to "yield information that may be useful as evidence" (Dammann, 2018) – i.e. we want to reduce our uncertainty which the effect or consequence of our lack of knowledge. *Reverse* (Gunawardena, 2014) or *statistical modeling* can be framed into this concept of *empirical learning*.

> Recall that there is a person "you", contemplating an "event", and it is desired to express your uncertainty about that event, which uncertainty is called your "belief" that the event is true.
>
> (Lindley, 2006)

Returning to the what-happens-next machine

Let's go back to Kermit's demonstration of his machine reducing the walking distance to the radio which is so shamelessly tiring his feet. Start where we've stopped before (at 1 minute and 52 seconds) and watch the next 5 seconds:

`https://www.youtube.com/watch?v=cog2a3YeDMM`

[3] https://statmodeling.stat.columbia.edu/2018/09/13/n1-survey-tells-me-cynthia-nixon-will-lose-by-a-lot-no-joke/

Obviously, the outcome of the machine was not *deterministic*. There have to be some inputs with unknown – but hopefully (indirectly) measurable – influences that lead to something statisticians call a *random variable*, here with two possible outcomes$\{$radio off, radio on$\}$. Now, with this new information we need to update our personal belief about the what-happens-next machine: Before we cut the rope, it is not as certain as Kermit pretended it to be that the outcome will be $Y = \{\text{radio on}\}$ conditional on $x = \{\text{cut}\}$.

Now get yourself back in the state of knowledge you had at 1 minute 52 seconds and let's change the storyline at this point: Highly esteemed scientist *Prof. Piggy* enters the stage and presents her valuable insights. Prof. Piggy explains:

> "By running a series of experiments, I found that conditional on:
>
> - sufficiently sharp scissors,
> - ordinarily lubricated deflection pulleys,
> - a sand bag weighting 1kg,
> - the seesaw always working but the velocity of it's vertical motion depending on the weight of the sand bag, and
> - the balloon being filled with some very light gas,
>
> it is my personal belief that the outcome is expected to take on $Y = \{\text{radio on}\}$ with a quantity I can describe with *three cases out of ten*. If I increase the weight of the sand bag to 2kg and hold all other inputs equal, this quantity rises to seven out of ten, if I decrease the weight of the sand bag to 0.1kg, the quantity decreases to one out of ten."

Prof. Piggy uses a language expressing *randomness*. She gives us values about with which *probabilities* ($\frac{3}{10}$, $\frac{7}{10}$ and $\frac{1}{10}$) we can *expect* a discrete random value to take on a specific value. She found this by running a series of appropriate experiments. Inspecting the data, she observes that the process leads to an increase in the probability for $Y = \{\text{radio on}\}$ when increasing the weight of the sand bag, ie. changing one of the conditionals.

Prof. Piggy concludes:

> "We got evidence for a change in the weight of the sand bag causing a relevant and positive *marginal effect* on the probability that Kermit's machine is switching on the radio."

But how did Prof. Piggy find out? The technique she applied was *reverse engineering*: Based on her generated experimental data, she observed what consequences a change of one input of $f(x)$ had on the output Y, and by this information she got evidence for the influence of this input on the sequential process constituting the what-happens-next machine.

> Reverse engineering, also called back engineering, is the process by which a man-made object is deconstructed to reveal its designs, architecture, or to extract knowledge from the object; similar to scientific research, the only difference being that scientific research is about a natural phenomenon.
>
> Wikipedia (27.05.2020):
> `https://en.wikipedia.org/wiki/Reverse_engineering`

Uncertainty

...

Tree height with stem diameter relationship

> The concept of measurement and [statistical] inference is often introduced through what are *supposed* to be simple experiments, such as inferring gravity from the time it takes objects to fall from various heights. The realizations of these experiments in practice, however, are often much more complex than their idealized designs imply, and any principled statistical analysis will have to go into much more detail than one might expect.
> `https://betanalpha.github.io/assets/case_studies/falling.html`

Such a *supposed to be simple* problem in forest sciences is the estimation of the relationship between tree height and stem diameter in order to, for example, replace missing tree height observations. Here, we're interested in the velocity – expressed in the unit m/cm – with which a tree grows towards the sun. After carefully recording the vertical position of tree tops – the *tree heights* – and their corresponding diameters, we want to identify a physical model that quantitatively describes this growing motion, as illustrated in Figure 2.

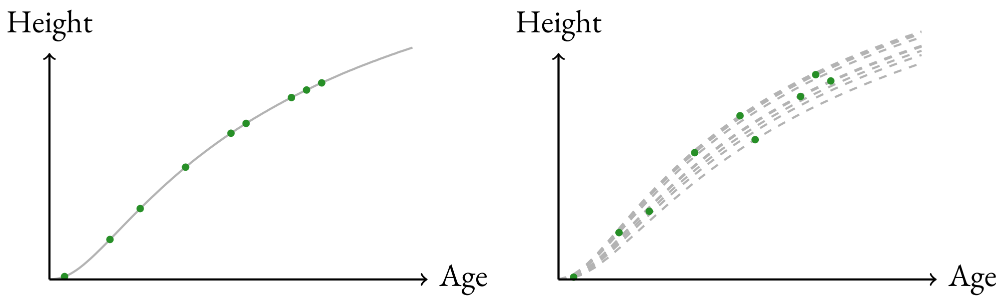

Figure 2: Left hand side: Perfect control over all possible influences on height, as well as measuring height and age, would strongly constraint the possible models, allowing us to exactly recover not only the trajectory taken by the tree top but also to make statements about any latent model parameter with absolute certainty.
Right hand side: In practice, however, measurements are inherently variable, which limits our ability to infer the exact trajectory and hence any model of how the trajectory itself.

Here and in all measurements, *uncertainty* is intrinsic to empirical learning:

$$\text{Learning by empirical evidence} \rightarrow \text{Knowledge with uncertainty}$$

Two types of randomness and two types of uncertainty

Philosophers of science differentiate between two types of randomness:

> Quantum mechanics forced physicists to reconsider the role of randomness in nature. To see why, let's distinguish two types of randomness. **Epistemic randomness** is what we (typically) find in classical mechanics. Games of chance use dice because only the probability of a given roll can be known in advance. But that isn't entirely true. A physicist with enough information about the linear and angular momentum of the dice, the coefficient of friction of the table, and a handful of other variables could predict the outcome with certainty. However, the dice end up, they had to have that result given the conditions and the laws of nature. To say this is a random event is merely a reflection of our ignorance. We don't have the relevant information and could not solve the equations quickly enough even if we did. In a practical sense then, one cannot predict the outcome of this "random" event even though it is just as determined by the laws of nature as the motion of a clock. The physicist Pierre-Simon Laplace extended this idea to the entire universe. As a thought experiment, he considered a superintelligence that knew the laws of nature and the precise state of all the particles in the universe at a particular instant. A so-called Laplacian demon could then calculate the state of the universe at any future time. Solving this set of equations is beyond the means of any computer, but there is nothing about the math- ematics to prevent it in principle. To a Laplacian demon in a Newtonian world, there would be no such thing as epistemic randomness and no need to resort to probabilities. Any given event would happen (or not) with absolute certainty. What about a Laplacian demon in a quantum world? That's an entirely different matter. Some events are indeterministic under the standard interpretation of quantum mechanics. Indeterminism produces a type of randomness that is not merely due to a lack of knowledge. Consider the radioactive decay of a specific uranium-232 atom. Such events are not physically determined by any prior cause. As far as nature is concerned, there is only a probability that a decay event will occur at any given time. A uranium atom has a 50 percent chance of decaying any time in the next 70 years. There is no hidden physics that triggers radioactive decay – no tiny fuse that determines when this event will happen. A Laplacian demon could only calculate the chance of such a decay at any point in time. This **ontological randomness** [...].
>
> (Koperski, 2020, p. 39-40)

Based on these two types of randomnes, we can distinguish two types of uncertainty when trying to increase our knowledge by empirical evidence:

> There are things that I am uncertain about simply because I lack knowledge, and in principle my uncertainty might be reduced by gathering more information. Others are subject to random variability, which is unpredictable no matter how much information I might get; these are the *unknowables*. The two kinds of uncertainty have been debated by philosophers, who have given them the names *epistemic uncertainty* (due to lack of knowledge) and *aleatory uncertainty* (due to randomness).
>
> (O'Hagan, 2004)

Here the etymology of *aleatory* is deriving from *alea*, which is Latin for 'dice'.

Aleatoric uncertainty due to randomness

- we are not able to obtain observations which could reduce this uncertainty

Epistemic uncertainty due to lack of knowledge

- we are able to obtain observations which can reduce this uncertainty
- two observers may have different epistemic uncertainty

(Vehtari, 2020, Slides `bayes_intro.pdf`)

Probability

> Take the familiar concept of a distance between two points, where a commonly used measure is the foot. What does it mean to say that the distance is one foot? All it means is that somewhere there is a metal bar with two thin marks on it. The distance between these two marks is called a foot, and to say that the width of a table is one foot means only that, were the table and the bar placed together, the former would sit exactly between the two marks. In other words, there is a standard, a metal bar, and all measurements of distance refer to a comparison with this standard. Nowadays the bar is not used, being replaced by the wavelength of krypton light and any distance is compared with the number of waves of krypton light it could contain. The key idea is that all measurements ultimately consist of comparison with a standard with the result that there are no absolutes in the world of measurement. Temperature was based on the twin standards of freezing and boiling water. Time is based on the oscillation of a crystal, and so on. Our first task is therefore to develop a standard for uncertainty.
>
> Lindley (2006, p. 31)

Our main tool to express evidence[4] is *probability*:

> In Bayesian statistics, probability is used as the fundamental measure or yardstick of uncertainty. Within this paradigm, it is equally legitimate to discuss the probability of 'rain tomorrow' or of a Brazilian victory in the soccer World Cup as it is to discuss the probability that a coin toss will land heads. Hence, it becomes as natural to consider the probability that an unknown estimand lies in a particular range of values as it is to consider the probability that the mean of a random sample of 10 items from a known fixed population of size 100 will lie in a certain range. The first of these two probabilities is of more interest after data have been acquired whereas the second is more relevant beforehand. Bayesian methods enable statements to be made about the partial knowledge available (based on data) concerning some situation or 'state of nature' (unobservable or as yet unobserved) in a systematic way, using probability as the yardstick. The guiding principle is that the state of knowledge about anything unknown is described by a probability distribution.
>
> Gelman et al. (2014, p. 11)

> If you think that a ball is to be withdrawn at random from an urn containing only red and white balls, then your probability that the [first] withdrawn ball will be red is defined

[4] Or knowledge?

> to be the fraction of all the balls in the urn that are red. [...] The first stage in the measurement of uncertainty has now been accomplished; we have a standard. The urn is our equivalent of the metal bar for distance, perhaps to be replaced by some improvement in the light of experience, as light is used for distance.
>
> Lindley (2006, p. 34/35)

The idea behind what is a *standard for uncertainty* is that to any event that is uncertain to us, for example 'the expected height E_h of the largest 100 trees per ha is below 25m' of a forest stand of some sort, we can make an analogy to another uncertain event from the urn example. We don't know what the probability is that this event will be true, but – as we can, to any flexibility, vary the numbers r and n of an urn containing n balls of which r red – there must be a fraction $\frac{r}{n}$ such that our beliefs in 'E_h is below 25m' and 'the first withdrawn ball will be red' are the same.

Systems

This section is a translation – with small changes from *hydro-* to *forest ecology* – of Petzoldt (2017, p. 1-2) `https://wwwpub.zih.tu-dresden.de/~petzoldt/modlim.pdf`

By a system we mean a part of the world – which consists of elements and relations ('connections') – that is delimited according to certain points of view. We call everything that does not belong to the system the *environment*, the border of which is the system's *edge*. The cut-off here is defined such that the relationships between system elements that are of interest from the respective point of view run, as far as possible, within the system and only a few relationships connect with the environment. This delimitation is done spatially, or by any other criterion. As an example, we can consider a spruce as an element in a spatial or three-dimensional forest system, or the 'system' spruce as a genetic or as a physiological or systematic unit (Piceoideae, Pinaceae, conifers, plants). In addition, the elements of any system can in turn be understood as systems (subsystems), e.g. the individual organs of the spruce. By consideration of a spruce as a wood production machine in a fixed location and given we are not interested in the physiology behind it, a clear system edge is arising. This or perhaps a slightly modified system view is part of almost every science and thus of course a general feature of forest science and forest ecology in general, i.e. forest ecology examines forest ecological systems (system ecology) and their subsystems. We do not consider entire ecosystems in every case, but often only individual communities, populations or individual organisms with their specific behavior, their physiological performance, their ontogenetic development or their evolution in a biotic and abiotic environment.

Models

> Models are approximations of the complex dynamics that drive the observable phenomena in the world around us. They provide the setting in which we can formalize learning and decision making and hence are a foundational aspect to any rigorous analysis.
>
> Betancourt (2019)

Construction of *models* is a pre-eminent strategy of how we, the human species, attain knowledge, learn, and understand most parts of the world around us. As we grow up[5], we become more experienced with the *construction of a model*, built in order to learn something about a phenomenon we seek to understand.

> Model-building starts early. Children build model airplanes then blow them up with firecrackers just to see what happens. Civil engineers build physical models of bridges and dams then see what happens inscale-model wave pools and wind tunnels. Disease researchers use mice as model organisms to simulate how cancer tumors might respond to different drug dosages in humans. These examples show exactly what a model is: a stand-in, an im-poster, an artificial construct designed to respond in the same way as the system you would like to understand. A mouse is not a human but it is often close enough to get a sense of what a particular drugwill do at particular concentrations in humans anyway. A scale-model of an earthen embankment dam has the wrong relative granularity of soil composition but studying overtopping in a wave pool still tells us something about how an actual dam might respond.
>
> van de Meent et al. (2018, page 10)

In our world of grown-ups, numerous modeling concepts exist making *model* an umbrella term collecting numerous manifestations. For example, in order to understand and / or predict tree growth and forest dynamics, a lot of different model types have been developed, from *yield tables* to *aggregated stand-growth simulators* to *physiologically based single-tree models*. For all, a model always is an artificial construct designed to respond in the same way as the system – a stand, two neighboring trees of different species, a single tree, one component of the tree-system, … – we would like to understand.

Working on an applied research project, we are interested in the dimension of trees of a certain species. This interest is already the product of the first step in the scientific process, which is the development of our subject-matter knowledge identifying and leading to this knowledge gap. In a third step, we make a travel to simplification kingdom and define more precisely what we understand under *dimension* here: so say we're interested in the *volume of the stem*. We chose the size of this tree organ as a *model* for the too complex term *tree dimension* identified in the second step: it has just too many aspects (leaf area/mass, fine and coarse root mass, branch/crown architecture/development/volume/vitality, ...) in order to deduce a quantity with which one can answer questions without academically working one's fingers to the bone just on the definition of *tree dimension*. So as scientists we need simplification just as much as we need it as normal citizens navigating through daily life: without the ability to focus our attention on the information that is relevant to the task by disregarding all the thousand distractions around us, we wouldn't even be able to make it to the office through morning's hectic traffic. But stem

[5] `https://bit.ly/36RlNlg`

volume is itself still quite complex to measure, at least for a sufficient number of trees to be measured such that we are able to come up with an answer to our applied research question *with desired precision*. So what we usually end up with is a measure of diameter at breast height (*dbh*) and an equation that – conditional on some auxiliary information – calculates stem volume based on this single input. Here, *pragmatism rather then detail* is the rule rather the exception, and directly after the first step in the scientific, it is using models as a pragmatic simplification all the way.

Subjective simplification

When we build a *model*, we have to make certain *assumptions* just by definition of the concept of a model – which is, as we have just seen, itself synonymous for *pragmatic simplification*.

Where there will be no debate about whether we need to simplify – the universe is just too complex! –, there can just never be a non-subjective consensus on the composition of pragmatism – different experts will come to different solutions.

> In order to come up with a *useful* model, we need to think through and justify our choices in the multi-dimensional and often un-measurable space of the nuts and bolts that drive the model's *performance*.

Here, the intrinsic variability of the real world is on the same the the basic motivation for – we can't write down a deterministic formula taking into account each influence –, but also the most significant complication of modeling – we don't know how the variability relates to the signal we wish to detect.

Cox and Donnelly (2011), Section 1.9:

> "Most of our later discussion centres on analyses based on probability models for the data, leading, it is hoped, to greater subject-matter understanding. Some probability models are essentially descriptions of commonly occurring patterns of variability and lead to methods of analysis that are widely used across many fields of study. Their very generality suggests that in most cases they have no very specific subject-matter interpretation as a description of a detailed data-generating process. Other probability models are much more specific and are essentially probabilistic theories of the system under investigation. Sometimes elements of the second type of representation are introduced to amend a more descriptive model. The choice of statistical model translates a subject-matter issue into a specific quantitative language, and the accuracy of that translation is crucial."

Subjectivity

> [...] somewhat arbitrary choices come into many aspects of statistical models, Bayesian and otherwise, and therefore we think that it is a mistake to consider the prior distribution as the exclusive gate at which subjectivity enters a statistical procedure.
>
> (Gelman and Hennig, 2017, p. 4-5)

A recent Twitter-'conversation' on (diverse) replies that can follow the ordinary 'attack' that priors depend on personal experiences:

`https://twitter.com/PhDemetri/status/1272318372076826626`

> Personal decision making cannot be avoided in statistical data analysis and, for want of approaches to justify such decisions, the pursuit of objectivity degenerates easily to a pursuit to merely *appear* objective. Scientists whose methods are branded as subjective have the awkward choice of either saying, 'No, we are really objective', or else embracing the subjective label and turning it into a principle, and the temptation is high to avoid this by hiding researcher degrees of freedom from the public unless they can be made to appear 'objective'. Such attitudes about objectivity and subjectivity can be an obstacle to good practice in data analysis and its communication, and we believe that researchers can be guided in a better way by a list of more specific scientific virtues when choosing and justifying their approaches.
>
> (Gelman and Hennig, 2017, p. 3)

Uncertainty is subjective. No algorithm is practically able to process all the effects that influence the throw of dice, even under the most controlled laboratory settings, there will always be to many influences coming in effect in too short a time. So if two different, independent groups or researchers would try, they would almost surely use different algorithms and different hardware – their approach on 'explaining' the underlying randomness would be *subjective*. In this sense, every scientific approach is subjective, even a microscope is operated by a human being and the details it shows must be put into context, which will again be done by a trained human being.

Measurements are subjective. We can as questions like *Why was a certain measurement device used?*, or*How were the observation units selected?*, which will find answers based on a researchers personal experiences.

Models are subjective, different experiences shape decisions by modelers, during review process reviewers contribute with their own subjective experiences.

Statistical Methods are subjective:

> For many statistical methods tuning constants need to be decided such as the proportion of trimmed observations when computing a trimmed mean or bandwidths for smoothing indensity estimation or non-parametric regression; one could also interpret the conventionaluse of the 0.05 level of significance as a kind of tuning parameter. In the statistical literature,methods are advertised by stating that they do not require any tuning decisions by the user.Often these choices are hidden so that users of statistics (particularly those without specificbackground knowledge in statistics) expect that for their data analysis task there is a uniquecorrect statistical method. [...] Although such an approach obviously tempts the user by its simplicity, it also appeals on thelevel of avoiding individual impact or subjectivity. [...] This is in stark contrast with the typical trial-and-error way of building one or more statistical models with plenty of subjective decisions starting from data preprocessing via data exploration and choice of method onto the selection of how to present which results. Realistically, even one-click methods require user choices on data coding and data exclusion, and these inputs canhave big influences on end results such p-values and confidence intervals.
>
> (Gelman and Hennig, 2017, p. 3)

Priors are subjective, but prior subjectivity is (usually) washing out rather quickly – in comparison to model mis-specificiations which never will – as more measurements become available. And moreover, priors is a self-contained concept creating an unbelievable rich set of possibilities, such as regularisation, that is usually hunted after in frequentist settings as well, but priors are just more elegant.

Beyond subjective and objective in statistics

The title of this section is the title of an article by Gelman and Hennig published in 2017 – see Gelman and Hennig (2017) in the References.

I want to sketch out their alternative framing *beyond subjective and objective*:

- ...

References

Betancourt, M. (2019). Probabilistic Modeling and Statistical Inference. *betanalpha.github.io*.

Cox, D. and Donnelly, C. (2011). *Principles of Applied Statistics*. Cambridge University Press.

Dammann, O. (2018). Data, information, evidence, and knowledge: A proposal for health informatics and data science. *Online journal of public health informatics*, 10.

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., and Rubin, D. B. (2014). *Bayesian Data Analysis*. Texts in Statistical Science Series. CRC Press, Boca Raton, FL, third edition.

Gelman, A. and Hennig, C. (2017). Beyond subjective and objective in statistics. *Journal of the Royal Statistical Society, Series A*.

Gunawardena, J. (2014). Models in biology: 'accurate descriptions of our pathetic thinking'. *BMC Biology*, 12(29).

Hilbe, J. M., de Souza, R. S., and Ishida, E. E. O. (2017). *Bayesian Models for Astrophysical Data: Using R, JAGS, Python, and Stan*. Cambridge University Press.

Koperski, J. (2020). *Divine Action, Determinism, and the Laws of Nature.* Routledge, London.

Lindley, D. V. (2006). *Understanding Uncertainty*. John Wiley & Sons, Inc., Hoboken, New Jersey.

O'Hagan, T. (2004). Dicing with the unknown. *Significance*, 1(3):132–133.

Petzoldt, T. (2017). Konstruktion ökologischer Modelle mit der Open Source Software R. *https://wwwpub.zih.tu-dresden.de/ petzoldt/*.

Sullivan, G. (2005). *Art Practice as Research: Inquiry in the Visual Arts*. SAGE.

van de Meent, J.-W., Paige, B., Yang, H., and Wood, F. (2018). *An Introduction to Probabilistic Programming*. Arxiv.

Vehtari, A. (2020). Bayesian data analysis course at Aalto University, Finland. `https://github.com/avehtari/BDA_course_Aalto`.