Stats II - Course Notebook

Bill Foote

2026-01-11

Contents

Preamble

This book is a compilation of many years of experience with replying to clients (CFO, CEO, business unit leader, project manager, etc., board members) and their exclamations and desire for the number. The number might be the next quarter's earnings per share after a failed product launch, customer switching cost and churn with the entry of a new competitor, stock return volatility and attracting managerial talent, fuel switch timing in a fleet of power plants, fastest route to ship soy from Sao Paolo to Lianyungang, launch date for refractive surgery services, relocation for end state renal care facilities, and so on, and so much so forth.

Here is a highly redacted scene of a segment of a board meeting of a medical devices company in the not so distant past. All of the characters are of course fictions of a vivid imagination. The scene and the company are all too real.

Sara, a board member, has the floor.

> Sara: Bill, just give us the number!

Bill (no relation) is the consultant that the CFO hired to provide some insight into a business issue: how many product units can be shipped in 45 days from a port on the Gulf of Mexico, USA to a port in Jiangsu Provice, PRC.

> Bill: Sara, the most plausible number is 42 units through the Canal in 30 days. However, …

Sara interjects.

> ?Sara: Finally! We have a number …

Harrumphing a bit, but quietly and thinking "Not so fast!",

> Bill: Yes, 42 is plausible, but so are 36 and 45, in fact the odds are that these results, given the data and assumptions you provided our

> team, as well as your own beliefs about the possibilities in the first place, are even. You could bet on any of them and be consistent with the data.

Jeri the CFO grins, and says to George the CEO in a loud enough whisper so that other board members perked up from their respective trances,

> Jeri, CFO: That's why we hired Bill and his team. He is giving us the range of the plausible. Now it's up to us to pull the trigger.

Sara winces, yields the floor to Jeri, who gently explains to the board,

> Jeri: Thank you Bill and give our regards to your team. Excellent and extensive work in a very short time frame!

Bill leaves the board room, thinking "Phew!, made through that gauntlet, again."

> Jeri: Now our work really begins. George, please lead us in our deliberations.

Pan through the closed board room door outside to the waiting room and watch Bill, wiping the sweat from his brow, prepare the invoice on his laptop.

A watchful administrative assistant nods his approval.

Why this book

This book is an analytical palimpsest which attempts to recast and reimagine the Bayesian Stan-based analytics presented by Richard **?** for the data analytics that help business managers and executives make decisions. Richard McElreath's vision and examples, though reworked, are nearly immediately evident throughout. But again this work is a palimpsest with its own quirks, examples, but surely built on McElreath's analytical architecture.

The data for the models is wrangled through the R tidyverse ecosystem and approaches that Solomon Kurz uses in recasting Richard McElreath's text with the brms R package developed by Paul Bürkner's brms package. The concepts, even the *zeitgeist* of a Danielle Navarro might be evident along with examples from her work on learning and reasoning in the behavioral sciences. And if there was ever a pool of behavior to dip into it is in the business domain.

Premise(s)

The one premise of this book is that *Learning is inference.*

By inference we mean reaching a conclusion. Conclusions can be either true or false. They are reached by a process of reflection on what is understood and by what is experienced.

Let's clarify these terms in this journey courtesy of Potter (1994):

- Ignorance is simply lack of knowledge.
- Opinion is a very tentative, if not also hesitant, subject to change, assent to a conclusion.
- Belief is a firm conviction.
- Knowledge is belief with sufficient evidence to justify assent.

= Doubt suspends assent when evidence is lacking or just too weak.

- Learning can only occur with doubt or ignorance or even opinion; learning generates new knowledge.

? develops the heuristic structure of human knowing about three successive movements from experience, through understanding, onto to reflection about what is and is not (plausibly of course!). The terms of each movement wind their way into the methodology of any excursus that implements human knowing about anything, including the knower, whom we often tag the analyst or decision maker. The knower observes and experiences, yes, with a bias.

The course will boil down to the statistics (minimum, maximum, mean, quantiles, deviations, skewness, kurtosis) and the probability that the evidence we have to support any proposition(s) we claim. The evidence is the strength (for example in decibels, base 10) of our hypothesis or claim. The measure of evidence is the measure of surprise and its complement informativeness of the data, current and underlying, inherent in the claim. In the interests of putting the bottom line up front Here is the formula for our measure of evidence e of a hypothesis H comes from Edwin T. **?**.

$$e(H \mid DX) = e(H \mid X) + 10lpg_{10}\left[\frac{Pr(D \mid HX)}{Pr(D \mid \overline{H}X)}\right] \tag{1}$$

Let's parse this relationship. First H is the claim, say for example, the typical null hypothesis H_0 that the status quo is true. If $H = H_0$ then $\overline{H}$ (literally not H) must be the alternative hypothesis logically so that $\overline{H} = H_1$.

$= D$ is the data at hand we observe. X is the data, including data about beliefs, we already have at our disposal before we observe D.

Now we look at the combinations. DX is the logical conjunction of the two data sets. This conjunction represents the proposition that both the new data D and the old data exist X and are true. HX is the phrase both the claim H is true and the available data X is true. $\overline{H}X$ is the phrase: both the claim is false and the available data X are true. Here are the conditions of the contract.

- We look for evidence $e(H \mid DX)$ that H is true given the existence both of new data and available data X, that is $H \mid DX$ where the $A \mid B$ reads A *given* B, a conditional statement.
- This compound (DX) depends on the evidence $e(H \mid X)$ that H is true given knowledge of available data X.
- The evidence impounded in the odds of finding data when the hypothesis is true relative to when the hypothesis is not true .

Everything is in log_{10} measures to allow us the luxury of simply adding up evidence scores. What is H? These are the hypotheses, ideas, explanations, parameters (like intercept and slope, means, standard deviations, you name it) we think when we work through answers to questions.

So many questions and too little time Where in the business domain do questions arise? The usual suspects are the components of any organization's, stakeholders', value chain. This chain extends from questions about the who, when, where, how, how long, how much, how often, how many, for whom, for what, and why of each component. There are decisions in each link. In planning an organization's strategy the decisions might be to examine new markets or not, invest in assets or retire them, fund innovation with shareholder equity or stand fast. In each of these decisions are questions about the size and composition of markets and competitors, the generation of cash flow from operations, the capture of value for stakeholders. It is this font from which questions will arise and analysis will attempt to guide.

Don't we know everything we need to know?

We don't seem to know everything all of the time, although we certainly have both many doubts and opinions. These hurdles are often managed through the collection of data against hypotheses, explanations, theories, and ideas. There are two bookend inferences we can make, that is, deduce plausibly.

The first is the glass is half-full approach, *modus ponens*. If A is true, then B is true. We then observe that B is true. We conclude that A becomes more plausible.

Then there is the glass is half-empty view, *modus tollens* We begin again with the implication that if A is true, then B is true. We then add data: A is false. The inference then concludes that B becomes less plausible.

There is no guarantee here, just plausible, but justified, belief. We will call plausibility a measure of belief, also known as probability.

The first chapter will detail the inner workings and cognitional operations at work in these inferences. But let us remember that learning is inference.

What we desire

These two ways of inferring plausible, justified, true belief will be the bedrock of this book. They imply three desiderata of our methods:

1. Include both the data of sensible observation and the data of assumptions and beliefs.
2. Condition ideas, hypotheses, theories, explanations with the data of experience and belief.
3. Measure the impact of data on hypotheses using a measure of plausibility.

What we will call rational will be the consistency, the compatibility, of data with hypotheses measured by plausibility, to be called posterior probability. The data of beliefs will be contained in what we will call prior probability. The data of beliefs include our assumptions about the distribution of hypotheses.

The conditioning of hypotheses with data is what we will call likelihood. Ultimately what we call uncertainty will be the range and plausibility of the impact of data on hypotheses. What we will solve for is the most plausible explanation given data and belief. Frequentist or probabilistic? Both.

The moving parts of our reasoning demand at least a probabilistic approach if we are to summarily, if not completely ever, deal with uncertainty in our descriptions and explanations of reality. For the business decision maker this mindset becomes critical as the way in which any decision must be made if it is to be compatible with the data of experience, understanding and reflection. Practically speaking the probabilistic approach directly aligns with the heuristic structure of human knowing: experience, understanding, reflection. All are fulfilled virtually (we don't know everything through bias, omission, ignorance, malice) a priori to the decision. Because our understanding of the future is imperfect we insert the word and idea of plausibility into every statement.

Frequentism is a quasi-objective approach to analysis. It is objective in that it focuses on only the data. It is quasi-objective in that the data is collected by human knowers and doers. An assumption of the analysis is always that the data are equally likely to occur. These beings do not always have access to or even want to collect all or the right data for further deliberation. That is an interesting historical, and a priori, fact of human existence. The collection itself rails against the time, space, and resources needed for collection and thus the epithet of garbage-in and garbage-out.

However, frequentist approaches contradict their supposed objectivism in that both the data selected, collected, and deliberated up are subject (yes the subjective begins to enter) to, conditional on, the collector and the deliberator. Both frequentist and probabilistic reasoning seem to intersect when prior knowledge is all but unknown (the uninformative or diffuse prior) or might as well assign equally plausible weights (probabilities) to any hypothesis the analyst might propose about the value of a moment of a distribution. They all but diverge when uneven, lumpy, step function priors on the values of the supposed estimators as hypotheses collide with the likelihood that the data is compatible with any of these hypotheses. Such divergence is not the destruction of objectivity, rather the transparent inclusion into a complete objective description and explanation of a tradition, a set of assumptions, a font of knowledge to date. [1]

A work in progress

This book will expand and contract over the days ahead. Email me with comments, errors, omissions, etc.

Thanks!

Bill Foote, Fordham University, Bronx, NY

[1] Charles Sanders Pierce (**?**) is perhaps the seminal study that incorporates prior knowledge into what would appear to be a frequentist null hypothesis approach, also known as *modus tollens*.

Topic 1 – The Basics

? sends a man through a labyrinth of a garden of "forked paths." By analogy Bayesian inference is really just counting and comparing of possibilities, the *forked paths*. Like Borges' character, we encounter numerous logical, causal, practical paradoxes and contraditions, often galore. FOr each possibility we decide to go left, go right, ah, we choose left. We reach another contradition, go left again, right or stand still? As we move outwards from any position, any possibility we continue to branch into an ever expanding, emergent, garden of forked paths.

Well, we are here to continue to learn to make reasonable inferences about what might happen given what has already been revealed as happening. In this way Bayesian analysis helps us structure our understanding about what could have happened, and might happen again. Borges' forks are alternative sequences of events in schemes of recurrence. It is interesting that as we proceed through the garden we discard paths along the way. We go left, forever never taking that particular right path. We keep going until we reach a condition which Bernard **?** calls the virtually unconditioned judgment of what is logically at the end of a seemingly endless sequence of forks. We arrive at the end of the current journey where what might be is what is logically consistent with what we observed. The result is virtual so there is no guarantee we have an absolutely correct, right answer. We have the best answer reasonably available in the here and now which can be derived from the data and the conjectures fed into our understanding.

We explore these areas from our past and extend them into our current work.

- Basic logic to help us form consistent arguments, conjectures, hypotheses, and interpretations of analytical results.
- The BayesBox, the workhorse of our statistical inference engine, which mashes together data with hypotheses. This device, a true robot, emerges from our work with contingency tables.
- Measures of probabilistic expectations. This handle some our summaries of risk. But we will need to wait a bit to handle our view of uncertainty, the known unknown.

Chapter 1

Counting the Ways

1.1 Plausibility, probability and information

According to Aristotle, if two claims are well-founded, their truth values can be ascertained. A plate is either green, or it is not. When thrown, the plate will land without a chip or it will break. If I claim that when I throw the plate, it will land with out a chip and you disagree, I can simply throw the plate to find out who was correct. One part of this is a mind game, a thought experiment with a potential outcome. The other is the reality of actually throwing the plate and observing its status on landing. We will eventually call the mind-game a hypothesis and the reality a datum. Actually both are data: the hypothesis is unobserved data, and the so-called observations are observed data.

It is in disagreement that logical deduction might (plausibly) break down. There is no guarantee that the plate will break, or, for that matter, that it will chip. We must simply experiment with plate(s), green, blue or otherwise, to support the claim (or not). These claims arise in everyday life. For example, despite my poor performance in plate throwing in the past, there is no cogent reason to believe that it is absolutely, positively false that the plate I throw would land without a chip. There is a degree of acceptance, of plausibility, in one side of the claim, and on the other as well. Certainly it is not as false as the claim that $2 + 2 = 5$ in base 10 arithmetic, or the patently spurious claim that true is false or my cat is a dog.

Claims about things that are neither definitely true nor definitely false arise in matters both mundane and consequential: producing weather reports, catching the bus, predicting the outcomes of elections, interpreting experimental vaccine results, and betting on sports games, throwing the plate, to name just a few.

So we would benefit from a method of comparing claims in these situations – which atmospheric model produces better predictions? What is the best source for predicting elections? Should I blow three times on my lucky dice, or is this all just a figment of my denial based imagination?

1.2 Some Surprise

The goal, in all the cases above, is to guess about something that we don't or can't know directly, like the future, or the fundamental structure of the economy, or reasons why customer preferences change, on the basis of things we do know, like the present and the past, or the results of an existing or past experiment.

Mostly we guess. Some us try to systematically consider and attempt to support with evidence the guess. Lacking precision, and sometimes even accuracy, we try to avoid bad surprises. Goods ones are often welcome. If I use a government weather application on my smart phone it might not surprise me to see rain pelting down at 1630 (this afternoon). After all the app indicated as much. In advance of the rain I brought in chair pads and anything else that might get ruined with rain. An airline pilot knows all the defects of her aircraft. That knowledge saves lives.

Our mortal inferences, clever or dumb as they are, must have a surprise somewhere between *totally expected*, or zero surprises and thus certain 100% of the ways to make the statement, and *totally surprising* and 0% chance of anticipation. We will generally be making statements like: *it will probably rain tomorrow*, or *nine times out of ten, the team with a better defense wins*. This motivates us to express our surprise in terms of plausibility and we hanker for more precision with probability.

1.3 How many ways?

Let's use a simple example. We have four voters in an upcoming election. They may be red or blue voters. Three of us go out and talk to three voters at random, that is, indiscriminately. One of us happens to come upon a blue voter, another of us, independently, happens to find a red voter, and the other separately finds a blue voter. This is the very definition of a random sample. Each of the finders does not know what the other is doing, all three do know that there are four voters out there and they happened to have independently talked to two blue and one red voter. How many red voters and how many blue voters are there?

Here are all of the possible conjectures we can make for $blue = \square$ and $red = \square$ voters.

Reading this we see that there are 4 voters and 5 different voter compositions ranging from all red to all blue. Our sample is 2 blue and 1 red voter, so we can very safely eliminate the first and fifth conjectures from our analysis, but for the moment just keep them for completeness sake.

Table 1.1: voter conjectures

1	2	3	4
$\color{red}{\square}$	$\color{red}{\square}$	$\color{red}{\square}$	$\color{red}{\square}$
$\color{blue}{\square}$	$\color{red}{\square}$	$\color{red}{\square}$	$\color{red}{\square}$
$\color{blue}{\square}$	$\color{blue}{\square}$	$\color{red}{\square}$	$\color{red}{\square}$
$\color{blue}{\square}$	$\color{blue}{\square}$	$\color{blue}{\square}$	$\color{red}{\square}$
$\color{blue}{\square}$	$\color{blue}{\square}$	$\color{blue}{\square}$	$\color{blue}{\square}$

Table 1.2: ways voter conjectures turn out

1	2	3	4	proportion	ways	plausibility
$\color{red}{\square}$	$\color{red}{\square}$	$\color{red}{\square}$	$\color{red}{\square}$	0.00	0 x 4 x 0 = 0	0.00
$\color{blue}{\square}$	$\color{red}{\square}$	$\color{red}{\square}$	$\color{red}{\square}$	0.25	1 x 3 x 1 = 3	0.15
$\color{blue}{\square}$	$\color{blue}{\square}$	$\color{red}{\square}$	$\color{red}{\square}$	0.50	2 x 2 x 2 = 8	0.40
$\color{blue}{\square}$	$\color{blue}{\square}$	$\color{blue}{\square}$	$\color{red}{\square}$	0.75	3 x 1 x 3 = 9	0.45
$\color{blue}{\square}$	$\color{blue}{\square}$	$\color{blue}{\square}$	$\color{blue}{\square}$	1.00	4 x 0 x 4 = 0	0.00

For each of the three remaining conjectures we may ask how many ways is the conjecture *consistent* with the collected data. For this task a tree is very helpful. Let's take the first realistic conjecture the $\color{blue}{\square}$, $\color{red}{\square}$, $\color{red}{\square}$, $\color{red}{\square}$ hypothesis and check if, when we sample all of the four voters, what are all of the ways this conjecture fans out. So here we go.

1. We sampled a $\color{blue}{\square}$ first. How many $\color{blue}{\square}$'s are in this version of the composition of voters? only 1.

2. We then sampled independently a $\color{red}{\square}$. How many $\color{red}{\square}$'s are in this conjecture? Quite a few, 3.

3. Finally we sampled a $\color{blue}{\square}$ at random. We know there is only one $\color{blue}{\square}$ in this version of the truth.

So, it is just counting the ways: 1 $\color{blue}{\square}$ way $\times$ 3 $\color{red}{\square}$ ways $\times$ 1 $\color{blue}{\square}$ way $= 1 \times 3 \times 1 = 3$ ways altogether.

When asked, many surmise that the 2 blue and 3 red conjecture is the right one. Are they right? Here is a table of the ways each conjecture pans out. We then in a separate column compute the contribution of each conjecture to the total number of ways across the conjectures, which is $3 + 8 + 9 = 20$ ways. Also each of the conjecture propose a proportion p of the successes, that is, the blue voters in this context.

We cannot help but note that the proportion of ways for each conjecture can range from 0, perhaps to 1, since the proportions add up to 1. The number of ways also expresses the number of true consistencies of the data with the conjecture, an enumeration of the quality of the logical compatibility of conjectures with what we observe.

We might now revise our commonsense surmise that 2 blue and 2 red is the better conjecture. However, if we use the criterion that the conjecture with the most ways consistent with the data is the best choice for a conjecture, then clearly here we would say that there are 3 blues and 1 red. Perhaps we have a better criterion that would choose our equinanimous choice of 2 blues and 2 reds? It does not appear to be so.

Ways are the count, the frequencies of logical occurrence of a hypothesis *given* the data. The data includes the knowledge that there are possibly blues and reds, that there are 4 voters, and that we sampled 2 blues and 1 red. The relative frequency of the ways in which our conjectures are consistent with the data is what we will finally call *probability*.

1. The plausibility is a measure between and including 0 and 1.
2. The sum of all plausibilities is 1.
3. The plausibility must be compatible with common sense.

That last requirement, in turn, requires a sketch of why it is important.

- Let's suppose the plausibility of a proposition A (Jaynie is a runner, for example) is changed by another proposition B (Jaynie participated in a marathon).
- Let's mix this up with the plausibility that proposition C (Jaynie is a business analyst, for example) has nothing to do with proposition B.
- Then common sense (as opposed to common non-sense) dictates that the plausibility of A and C together (Jaynie is a business analyst runner) is also increased by proposition B.

We have just quantified the *plausibility* of logical truth values. We have also found a very compelling criterion for the choice of a conjecture given the data and circumstances surrounding our inquiry.

1.4 Back to data

The location and scale parameters μ and σ (e.g., mean and standard deviation) we often calculate, and have so since middle school, are not the mean and standard deviation of data x. They are not properties of the physical representation of events called **observed data**. These measures use counts of events, measures of, perhaps physical properties like heat, size, time, but in the end they are abstractions and summarizations about the events.

These parameters do carry information about the probability distribution of the representation of physical reality in data. To say μ is the mean of the data is to ascribe a mind's eye idea to the physical reality, to invest in physical, objective reality, a property that only exists in the mind, but is in fact **unobserved data**.

This is an example of the *mind projection* or *reification* fallacy much in vogue in the circles of *fake*, or better yet, *chop logic*. In the same way, probabilities only exist in our minds: there is no physical reality that is a probability, just a mental construct that helps us think through potential outcomes. Is information just such a fiction?

1.5 Exercising our grip on reality

1. Suppose there are 5 voters of unknown affiliation. What is the probability that 60% are red affiliates?

2. So one of 5 moved out and we are back to 4 voters. What is the most plausible proportion of blue voters if we sample the district and find 2 blue and 3 red voters?

Chapter 2

Probability for Real People

2.1 Can we rationally reason?

Some might wonder if this section header even makes sense! A decision maker, such as as CFO, looks at it and knows everyone has a reason, some good and some not so good. But sometimes our reasons are not founded in any data that can be observed either by ourselves or by others. Our reasoning can be founded on falsified data, delusions, unfounded opinions, beliefs with no authoritative grounds. Her questions move from the various relationships in production planning in the previous part of this project to the forecasting of what might happen at all. What would be the rationale for a forecast? So what does our CFO decision maker do? She begins her thought experiment with the weather in her hometown of Albany, New York.

Rationality in the context we have been discussing at least means that we, as decision makers, would tend to act based on the consistency of observed reality with imagined models of the real world and through ideas about the world in which data are collected. We attempt to infer claims about the world based on our beliefs about the world. When confronting ideas, imbued with beliefs, with observed reality we might find ourselves in the position to *update our beliefs*, even those, and sometimes especially those, we so dearly hold.

In our thinking about anything we would venture candidate hypotheses h about the world, say the world of meeting demand, providing services, marshalling resources in specific markets. Of course the whole point is that we do not know which hypothesis h is more plausible, or not. We then collect some data d. When we perform this task, we move from the mental realm of the possibiity of hypotheses, theories, surmises, and model to the realm of observed reality. We may well have to revise our original beliefs about the data.

To implement our maintained hypothesis of rationality, we begin our search for **potential consistencies of the collected data with our hypotheses** that are fed by the data. In our quest we might find that some one of the hypotheses has more ways of being consistent with the data than others. When the data is consistent with a hypothesis, that is, when the hypothesis is reasonable logically, then our belief in that hypothesis strengthens,[1] and becomes more plausible. If the data is less consistent with the hypothesis, our belief in that hypothesis weakens. So far we have performed this set of tasks with conjectures about virus testing and voter alliance in zip codes. Let's switch up our program and consider the following very simplified question about the weather.

2.2 Before we get to the shovels!

Let's apply a probability analysis to understanding the plausibility of the demand for pies at one of the smaller restaurants the vegan pie maker Make-A-Pie LLC vends to. Here we query a day by day count of the number of pies sold (the *pies* variable) and the weather (clear coded as C or rainy coded as R). The original data only has the date, count, and the weather. The rest is analysis based on this data.

For our purposes we would like to know the conditional probability of selling more than 10 pies when the weather is clear. We can answer this query by making a 2x2 contingency table of the conditional data counts of pie demand. This panel displays the counts of pies sold.

pies weather	rain	clear	total
≥ 10	2	8	10
≤ 10	3	6	9
total	5	14	19

How likely is it to sell more than 10 pies when the weather is clear? When the weather is clear either we sell more than 10 pies or we don't. So we look at the clear weather column of counts of pies sold for our answer: $Pr(> 10 \mid clear) = 8/14$. We can also see in this column that there are 14 pies sold. So we can also find our that the likelihood of clear weather is $Pr(clear) = 14/19$. Also the >10 row gives us the likelihood of selling more than 10 pies, rain or shine, that is $Pr(> 10) = 10/19$. But it does not make much sense to ask if we sell more than 10 pies (or not for that matter) that it will rain. Unless of course the weather somehow depends on how many pies we sell. Rule 3: use common sense!.

The probability of selling more than 10 pies when the weather is clear is about $8/14 \times 100 = 57\%$. We sold 8 pies greater than the threshold and 6 pies less than the threshold when the weather was clear. The ratio of $8 : 6 = 1.33$ reports the

[1]The core idea of *strengthen* is to take us from a more vulnerable to a less vulnerable place or state. Synonyms for strength include *confirm* and *validate*.

odds of selling greater than 10 pies. We interpret this number with the phrase: it is 1.33x more likely to sell greater than 10 pies, than not, when the weather is clear.[2]

2.3 Shovels in Albany NY

The CFO's ask of the analysts is this question about winter consumer behavior in the Albany-area market:

> *We see people carrying snow shovels. Will it snow?*

What is the data d? We have recorded a simple observation about the state of the weather so that single piece of data ($d =$ We see people carrying snow shovels). Here is where our beliefs enter. We have two ***hypotheses***, h: either it snows today or it does not.

Let's figure out how to solve this problem? We have three *desiderata*:

1. We should include our experiences with snow in our analysis.
2. We should collect data about carrying snow shovels in January as well.
3. We prefer more consistency of data with hypotheses to less consistency.

Here we go, let's strap ourselves in.

2.3.1 Priors: what we think might happen

Our observation is about the weather: clouds, wind, cold. But we want to know about the snow! That is our objective and we have definite ideas about whether (don't pardon the pun!) it will snow or not. We will identify our beliefs, ever before we make our observations, about snow. The analytical profession and custom is to label these beliefs as *a priori*,[3] and thus the ellipsis *prior*,

[2] Some folks might want to aggrandize this statement by saying that is is 33% more likely to sell more than 10 pies when clear. We might caution ourselves as we might be falling into a distortion by magnifying an apparent difference. The probability of selling more than 10 pies when clear is $8/14 = 4/7$. on the other hand the complementary probability of selling up to 10 pies when clear is $6/14 = 3/7$. We get to sell only 1 pie in crossing the 10 pie threshold given this data set. An inflated way of expressing the results of this analysis? Isocrates would agree.

[3] The *a priori* elements of any argument include just about everything you and I know, including the kitchen sink! We can't help but to have these antecedent thoughts, experiences, shared and not-so-shared histories. They tend to persist in most humans, including us. At least that is what we will maintain. Thus it is a *necessity* to include these beliefs in our discussion. Without their consideration we most plausibly will introduce unsaid and denied

contentions we hold when we walk into the data story we create with the question of *will it snow?*

Our prior contentions are just the plausibilities of each hypothesis whatever data we eventually collect. After all we have to admit to everyone what we believe to be true as the antecedent to the consequent of observations and the plausibility of snow. This move allows us to learn, to revise, to update our dearly held beliefs. We thus can grow and develop. This is in a phrase a *sine qua non*, a *categorial imperative*, a *virtually unconditioned* requirement for change.

What might we believe about whether it will snow (today)? If you come from Malone, New York, north of the Adirondack mountains, you will have a different belief than if you come from Daytona, Florida, on the matter of how many ways snow might happen in a given month. So let's take as our benchmark Albany, the capital of the state of New York.

We will refer to some data to form hypotheses and their plausibility.using this weather statistics site. The site reports the average number of days of snowfall in January, when there is at least a 0.25 cm accumulation in a day. It is 10.3 days. These are the number of ways (days) in January, in Albany, NY, that it is true, on average and thus some notion of expected, or believed to be, that it snows. The total number of ways snow could possibly fall in any January (defined by calendar standards) is 31. While the formation of the hypotheses *snowy* and *nice* days is informed by data, we are asking a question about snow because we have yet to observe if will snow. We cannot observe something that has not yet happened. We can thus characterize hypotheses and conjectures as **unobserved data**.

Thus we might conclude that we believe that is it plausible (probable) that snow *can* fall $10.3/31 = 30\%$ of the different ways snow can fall. Note very well we will talk about *priors* as *potentials* and *conjectures* and *hypotheticals*, and thus used the modal verbs *can* or *might*. Thus we believe it might not snow, because it is possible, with plausibility $1 - 0.30 = 0.70$, or, multiplying by 100, 70%, according to the law of total probability of all supposed (hypothesized) events. We only have two such events: *snow* and *not snow*. Probabilities must, by definition, add up to 1 and must, again by definition be a number between 0 and 1.

Nice ideas, nice beliefs, are our as yet to be observed, but projected notions of a snowy day. But how real, how plausible, how rational, that is, how consistent are they with any *observed data*? Is there any *observed data* we can use to help us project which of our unobserved data, our hypotheses, is more or less reasonable?

---

bias, let blindspots have the same focus as clearly understood experiences, and produce time and resource consuming blind alleys. But we should hang on here: even blind alleys and blind spots are extremely important bits of knowledge that help us understand what does not work, an *inverse insight* as exposed by Bernard **?**.

Table 2.2: Priors by hypotheses

hypotheses	priors
snowy day	0.3
nice day	0.7

Table 2.3: data meets hypotheses

hypotheses	shovels	hands
snow day	0.7	0.3
nice day	0.1	0.9

2.3.2 Likelihoods: thinking about the data

Life in the Northeast United States in January much revolves around the number of snow days, also known as days off from school. A prediction of snow meets with overtime for snow plow drivers, school shut downs, kids at home when they normally are in school. On some snowy days we see people carrying snow shovels, on others we don't. On some nice days we see people with snow shovels, on others we don't. Confusing? Confounding? A bit.

Now we link our observations of shovels with our unobserved, but through about and hypothesized, prediction of snow. We then suppose we observe that people carry snow shovels about 7 of the 10 snowy days in January or about 70%. On nice days we observe that people carry shovels at most 2 days in the 21 nice days or about 10%.

This table records our thinking using data we observe in Januaries about weather conditions.

First of all these probabilities register yet another set of beliefs, this time about whether we see shovels or not, *given*, *conditioned by*, the truth of each hypothesis h. We write the conditional probability $\Pr(d \mid h)$, which you can read as "the probability of d given h". Also here we will follow the convention that this set of results of our assessment of the relationship of shovels to snowy days as a likelihood .[4]

2.3.3 Altogether now

Do we have everything to fulfill our *desiderata*? Let's check where we are now.

1. We should include our experiences with snow in our analysis.

[4] For Pierre Simon **?**, likelihood also has the idea of $\Pr(h \mid d)$. Let's stick to our knitting, and tolerance for ambiguity, with using the rows of this table as our entries for likelihood.

Yes! We put our best beliefs forward. We even (sometimes this is a courageous analytical step) quantified teh ways in which snow and not snow would occur, we believe, in Albany NY in an average January.[5]

2. We should collect data about carrying snow shovels in January as well.

   Yes we did! Again we elicited yet another opinion, belief, whatever we want to colloguially call it. That belief if what we register and document based on observation of shovels and just hands in the presence of snowy and nice days in a January.

3. We prefer more consistency of data with hypotheses to less consistency.

   Not yet! We will impose our definition of rationality here.

Let's start out with one of the rules of probability theory. The rule in question is the one that talks about the probability that *two* things are true. In our example, we will calculate the probability that today is snowy (i.e., hypothesis h is true) *and* people carry shovels (i.e., data d is observed). The ***joint probability*** of the hypothesis and the data is written $\Pr(d, h) - \Pr(d \wedge h$, and you can calculate it by multiplying the prior $\Pr(h)$ by the likelihood $\Pr(d \mid h)$. The conjunction is a *both-and* statement. We express conjunctions using the wedge $\wedge$ symbol. Logically, when the statement that both d and h is true, then the plausibility, now grown into probability is:

$$\Pr(d \wedge h) = \Pr(d \mid h) \Pr(h)$$

When we divide both sides by $\Pr(h)$ we get the definition, some say derivation, of condition probability. If we count $\#()$ the ways $d \wedge h$ are true and the ways that h are true then

$$\#(d \mid h) = \frac{\#(d \wedge h)}{\#(h)}$$

Then the number of ways the data d are true, given h is true, equals the total number of ways that d and h per each way that h is true. We have thus normed our approach to understanding a conditional statement like if h, then d. Even more so, when we combine the law of conditional probability with the law of total probability we get Bayes Theorem. This allows us to recognize the dialectical

[5] we really need to think further about our notions of an average or centrally located anything. This means more consideration later, including deviations from these locations measured by scale.

principle that, yes, we recognize $h = \textit{snowy}$, but we also know that every cloud has its silver lining and that there is a non-snowy day and thus a

$$\textit{not}\ h = \neg h = \textit{nice}$$

lurking in our analysis.

Here it in in all its glory.

$$\Pr(h \mid d) = \frac{\Pr(d \mid h)\ \Pr(h)}{\Pr(d \mid h)\ \Pr(h) + \Pr(d \mid \neg h)\ \Pr(\neg h)} \tag{2.1}$$

$$== \frac{\Pr(d \wedge h)}{\Pr(d \mid h)\ \Pr(h) + \Pr(d \mid \neg h)\ \Pr(\neg h)} \tag{2.2}$$

The numerator is the same as the conjunction both d and \$h. The denominator is the probability that either both d and h or both d and h are true. While the build up to this point is both instructive, and thus may at first be *confusing*, it is useful as it will highlight the roles these probabilities perform in the drama that is our analysis.

We had better get back to the data or get lost in the weeds of the maths. So, what is the probability it is true that today is a snowy day *and* we observed people to bring a shovel?

Let's see what we already have. Our prior tells us that the probability of a snowy day in any January is about 30%. Thus $\Pr(h) = 0.30$. The probability that we observe people carrying shovels is true given it is a snowy day is 70%. So the probability that both of these things are true is calculated by multiplying the two to get 0.21. We can make this

$$l\,\Pr(\textit{snowy},\ \textit{shovels}) = \qquad \Pr(\textit{shovels} \mid \textit{snowy}) \times \Pr(\textit{snowy}) \tag{2.3}$$

$$= \qquad 0.70 \times 0.30 \tag{2.4}$$

$$= \qquad 0.21 \tag{2.5}$$

This is an interesting result, something odds makers intuitively know when punters put skin in the game. There will be a 21% chance of a snowy day when we see shovels in people's hands. However, there are of course *four* possible pairings of hypotheses and data that could happen. We then repeatthis calculation for all four possibilities. We then have the following table.

Just to put this into perspective, we have for the 31 days in a January this table.

We have four logical possibilities for the interaction of observed data and unobserved hypotheses. We arrange these possibilities in two stacked rows. We recall that visualizatiton is everything, even in tables! Here is the first row.

Table 2.4: Both data and hypotheses

hypotheses	shovels	hands	sum
snow day	0.21	0.09	0.3
nice day	0.07	0.63	0.7
sum	0.28	0.72	1.0

Table 2.5: both data and hypotheses in days in January

hypotheses	shovels	hands	sum
snowy day	6.5	2.8	9.3
nice day	2.2	19.5	21.7
sum	8.7	22.3	31.0

1. Snowy and shovels

$$
\begin{aligned}
l\,\Pr(snowy,\, shovels) = & \quad \Pr(shovels \mid snowy) \times \Pr(snowy) && (2.6)\\
= & \quad 0.70 \times 0.30 && (2.7)\\
= & \quad 0.21 && (2.8)
\end{aligned}
$$

2. Snowy and just hands

$$
\begin{aligned}
l\,\Pr(snowy,\, hands) = & \quad \Pr(hands \mid snowy) \times \Pr(snowy) && (2.9)\\
= & \quad 0.30 \times 0.30 && (2.10)\\
= & \quad 0.09 && (2.11)
\end{aligned}
$$

In this row the prior probability about snow is 0.30.

Here is the second row with its separate calculations.

1. Nice and shovels

$$
\begin{aligned}
l\,\Pr(nice,\, shovels) = & \quad \Pr(shovels \mid nice) \times \Pr(nice) && (2.12)\\
= & \quad 0.10 \times 0.70 && (2.13)\\
= & \quad 0.07 && (2.14)
\end{aligned}
$$

2. Nice and just hands

$$l\,\Pr(nice,\,hands) = \Pr(hands \mid nice) \times \Pr(nice) \quad (2.15)$$
$$= 0.90 \times 0.70 \quad (2.16)$$
$$= 0.63 \quad (2.17)$$

In this row the prior probability about nice days is 0.70.

An insightful exercise is to carry these calculations from the number of ways snow with and without shovels occurs given we think we know something about snow. The same with the number of ways a nice day might occur with and without shovels, given what we think about nice days.

Let's put one calculation together with a not so surprising requirement. When we conjoin snow with shovels, how many possible ways can these logical statements occur? It is just the 31 days.

We now have all of the derived information to carry our investigation further. We also total the rows and, of course, the columns. We will see why very soon.

The row sums just tell us as a check that we got all of the ways in which snow occurs in 31 days. What is brand new are the column sums. They add up the ways that data occurs across the two ways we hypothesize that data can occur: snow, no snow (nice day). They tell us the probability of carrying a shovel or not, across the two hypotheses. Another way of thinking about the $p(d)$ column sums is that they are the expectation of finding snow or hands in the data. The consistency of all of these calculations is that column sums equal row sums, 100%. All regular, all present and correct, probability-wise.

2.3.4 Updating beliefs

The table lays out each of the four logically possible combinations of data and hypotheses. So what happens to our beliefs when they confront data? In the problem, we are told that we really see shovels, just like the picture from Albany, NY at the turn of the 20th century. Is surprising? Not necessarily in Albany and in January, so you might expect this behavior out of habit during a rough Winter. The point is that whatever our beliefs have been about shovel behavior, we should still subject them to the possibility of accomodating the fact of seeing shovels in hands in Albany in January, a winter month in the Northern Hemisphere.

We should recall this formula about the probability of seeing both an hypothesis and data:

$$\Pr(h \mid d) = \frac{\Pr(d \wedge h)}{\Pr(d)} = \frac{\Pr(d \mid h) \Pr(h)}{\Pr(d)}$$

Now we can trawl through about our intuitions and some arithmetic. We worked out that the joint probability of *both* ***snowy day*** *and* ***shovel*** is 21%, a rate reasonable given the circumstances. In our formula, this is the product of the likelihood $\Pr(d = \textit{shovels} \mid h = \textit{snow}) = 0.70$ and the prior probability we registered that snow might occur $\Pr(h = \textit{snow}) = 0.30$.

Relative to the product of the likelihood of shovels given a nice day and the chance that snow might occur is the the joint probability of *both* ***nice day*** *and* ***shovel*** at 10%, or $\Pr(d = \textit{shovels} \mid h = \textit{nice}) \Pr(h = \textit{nice}) = 0.10 \times 0.70 = 0.07$, again a reasonable idea, since we plausibly wouldn't see much shovel handling on that nice day in January..

Both of these estimates are consistent with actually seeing shovels in people's hands. But what are the chances of just seeing shovels at all? This is an ***either or*** question. We see shovels 21% of the time on snowy days or we see shovels 7% of the total days in January on nice days. We then add them up to get 28% of the time we see shovels in all of January, whether it snows or not.

So back to the question: if we do see shovels in the hands of those folk, will it snow? The hypothesis is $h = \textit{snow}$ and the data is $d = \textit{shovels}$. The joint probability of both snow and shovels is $\Pr(d, h) = 0.21$. But just focusing on the data we just observed, namely that we see shovels, we now know that the chances of seeing shovels on any day in January in Albany, NY is $\Pr(d) = 0.27$. Out of all of the ways that shovels can be seen in January then we would anticipate that the probability of snow, upon seeing shovels, must be $\Pr(h \mid d) = \Pr(d, h) / \Pr(d) = 0.21/0.28 = 0.75$.

What is the chance of a nice day given we see shovels? It would be again likelihood times prior or $0.10 \times 0.7 = 0.07$ divided by the probability of seeing shovels any day in January 28%. We then calculate $0.07/0.28 = 0.25$. We now have the posterior distribution of the two hypotheses, snow or nice, in the face of data, shovels. So what are the odds in favor of snow when we see shovels?

$$OR(h \mid d) = \frac{\Pr(h = \textit{snow} \mid d = \textit{shovels})}{\Pr(h = \textit{nice} \mid d = \textit{shovels})} \tag{2.18}$$

$$= \frac{0.75}{0.25} \tag{2.19}$$

$$= 3 \tag{2.20}$$

We can read this as: when we see people with shovels in January in Albany, NY, then it is 3 times more plausible to have a snowy day than a nice day. The ratio of two posteriors gives us some notion of the plausible divergence in likely

Table 2.6: Unobserved belief tempered by observed data = posteriors.

hypotheses	shovels	hands	priors	posterior shovels	posterior hands
snow day	0.7	0.3	0.3	0.75	0.13
nice day	0.1	0.9	0.7	0.25	0.88
sum	0.8	0.2	1.0	1.00	1.00

outcomes of snowy versus nice days. Again we must append the circumstances of time and place: in a January and in Albany, NY.

Here is table that summarizes all of our work to date.

2.4 What did we accomplish?

We have travelled through the complete model of probabilistic reasoning.

1. We started with a question. The question at least bifurcates into the dialectical ***is it?*** or ***is it not?***.
2. We then began to think about beliefs inherent in the question for each of the hypotheses buried in the question.
3. We then collected data that is relevant to attempting an answer to the question relative to each hypothesis.
4. Then we conditioned the data with the hypotheses inside the question. It is always about the question!
5. Finally we derived plausible answers to the question.
6. We then illustrated the example with data collected, threshold set, table constructed, odds computed.

What is next? We continue to use this recurring scheme of heuristic thinking, sometimes using algorithms to count more efficiently, applied to questions of ever greater complexity. In the end our goal will be to learn, and learning is inference.

2.5 Some exercise is in order

Exercise 2.1. Start with a question for analysis using a indicative-interrogative statement format, for example "We observe X. Will Y occur?" Based on this statement identify the unobserved data of the hypothesis and the observed data. Use binary hypotheses and observations.

Exercise 2.2. Rework the Albany NY example using your hometown or city. Develop initial distribution of hypotheses, distributions of data given a hypothesis, joint distributions of hypotheses and data. Find the probability that a particular hypothesis might occur given a specific piece of data.

Exercise 2.3. Rework the example using various thresholds. Use the Data Table sensitivity analysis to calculate results. Plot the Data Table. Lead a discussion with some skeptical people about your findings.

Chapter 3

Foundations

Here is an architectonic approach ready to drop straight into any of our analytics design and product.

- Objections (with references for further reading)
- On the Contrary (with classical anchors)
- I Respond That (your affirmative position, fully argued)
- Replies to objections
- Further reading list and considerations.

A *quaestio* is a consideration, in question form, for discussion, debate, formation, development of principles for analytics design.

3.1 QUAESTIO I. Whether the Whole Is Greater Than the Sum of the Parts

3.1.0.1 Objections

3.1.0.1.1 Objection 1 — Baconian Empiricism

One reading of Francis Bacon holds that knowledge arises from the accumulation of discrete observations, and that wholes are nothing but aggregates of parts. Emergent form is dismissed as a projection of the mind.

Reference: **?**, *Novum Organum*, Book I.

3.1.0.1.2 Objection 2 — Scientistic Reductionism

Contemporary scientism asserts that all phenomena are reducible to physical components and their interactions. Consciousness, meaning, and value are epiphenomena of chemistry.

Reference: Alex **?**, *The Atheist's Guide to Reality*.

3.1.0.1.3 Objection 3 — Neuroscientific Materialism

Some neuroscientists claim that thought is nothing but neural firing patterns, and thus any "whole" is merely shorthand for chemical processes.

Reference: Patricia **?**, *Neurophilosophy*.

3.1.0.1.4 Objection 4 — Mechanistic Physics

If physics reduces systems to fundamental particles and forces, then any "whole" is simply the sum of microphysical states.

Reference: Steven **?**, *Dreams of a Final Theory*

3.1.1 On the Contrary

Aristotle teaches that a whole possesses a *form* that organizes the parts and gives them their intelligibility.

Reference: **?**, *Metaphysics* VIII.6.

Aquinas affirms that form is "that by which a thing is what it is," and that wholes possess properties irreducible to their components.

Reference: **?**, *Summa Theologiae* I, q.76, a.8.

Lonergan argues that intelligibility is grasped in insight, not reducible to data alone.

Reference: Bernard **?**, *Insight*, ch. 1–3.

3.1.2 I Respond That

I answer that the whole is indeed greater than the sum of the parts, not by addition but by emergent intelligibility. A whole is constituted by:

- Form (Aristotle, Aquinas)
- Higher-order relations (Lonergan's "schemes of recurrence")

- Contextual unity (Polanyi's "tacit integration")
- Functional organization (modern complexity theory)

The parts do not explain the whole; rather, the whole explains the parts.

Reductionism fails because:

1. It cannot account for emergence (life, consciousness, culture).
2. It cannot account for meaning (semantics is not chemistry).
3. It cannot account for value (ethics is not physics).
4. It cannot account for intelligibility (insight is not reducible to data).

Thus, the whole is ontologically and epistemically prior to the parts. This is the foundation for any humane analytics: data points are not reality; they are fragments awaiting intelligible integration.

3.1.3 Replies to Objections

3.1.3.1 Reply to Objection 1 (Bacon)

Bacon's empiricism mistakes accumulation for understanding. Insight is not a sum of observations but a grasp of form.

3.1.3.2 Reply to Objection 2 (Scientism)

Scientism is self-refuting: the claim "only physical explanations are valid" is not itself a physical explanation.

3.1.3.3 Reply to Objection 3 (Neuroscience)

Neural correlates do not exhaust mental acts. Correlation is not identity. Thought transcends its material substrate.

3.1.3.4 Reply to Objection 4 (Physics)

Physics describes interactions, not intelligibility. Meaning, purpose, and organization are not derivable from microphysical laws.

3.2 QUAESTIO II. Whether Deduction Produces New Knowledge

3.2.0.1 Objections

3.2.0.1.1 Objection 1 — Humean Skepticism

Hume argues that deduction is analytic and tautological, adding nothing to knowledge.

Reference: David **?**, *An Enquiry Concerning Human Understanding*, Section IV.

3.2.0.1.2 Objection 2 — Locke's Passivity of Mind

Locke holds that the mind merely rearranges simple ideas; deduction is recombination, not discovery.

Reference: John **?**, *Essay Concerning Human Understanding*, Book II.

3.2.0.1.3 Objection 3 — Kantian Formalism

Kant claims deduction structures appearances but does not give knowledge of things-in-themselves.

Reference: Immanuel **?**, *Critique of Pure Reason*, A6–7.

3.2.0.1.4 Objection 4 — Nihilistic Power-Epistemologies

Callicles, an early Nietzsche, and Sartre reduce knowledge to power, rendering deduction a tool of domination rather than discovery.

References: **?**, *Gorgias*; Friedrich **?**, *Genealogy of Morals*; Jean-Paul **?**, *Being and Nothingness*.

3.2.1 On the Contrary

Lonergan teaches that understanding is a dynamic, self-transcending process: insight $\rightarrow$ formulation $\rightarrow$ verification. Deduction unfolds the intelligibility grasped in insight.

Reference: Bernard **?**, *Insight*, ch. 2–4.

Aquinas affirms that reasoning extends knowledge by drawing out implications latent in principles.

Reference: **?**, *Summa Theologiae* I, q.79, a.8.

3.2.2 I Respond That

I answer that deduction does indeed produce new knowledge, because:

1. Insight grasps a form, but deduction articulates its implications.
2. Judgment affirms reality, and deduction extends that affirmation.
3. Understanding is developmental, not static.
4. Implications are not present to consciousness until reason unfolds them.

Thus, deduction is not mechanical manipulation but the expansion of intelligibility. It reveals what was implicit but not yet known.

This is essential for analytics: models, regressions, and forecasts are not mere computations — they are the unfolding of intelligibility grasped in data.

3.2.3 Replies to Objections

3.2.3.1 Reply to Objection 1 (Hume)

Hume's empiricism cannot account for mathematics, science, or even his own argument. Deduction is not tautology; it is the unfolding of insight.

3.2.3.2 Reply to Objection 2 (Locke)

The mind is not passive. Insight is an active grasp of form, and deduction extends that grasp.

3.2.3.3 Reply to Objection 3 (Kant)

Kant's restriction of knowledge to appearances is a metaphysical decision, not a discovery. Deduction reveals real intelligibility.

3.2.3.4 Reply to Objection 4 (Nihilists)

If knowledge is only power, then the claim "knowledge is power" is itself a power-play and not knowledge. This is self-referentially incoherent.

The third and next *quaestio* is the keystone — the one that makes the first two *operative* in a Stats/Analytics classroom.

3.3 QUAESTIO III. Whether Philosophical and Theological Considerations Should Guide an Empirically Oriented Management Analytics Course

3.3.1 Objections

3.3.1.1 Objection 1 — Milton Friedman's Value-Neutral Economics

Friedman argues that the sole responsibility of business is to increase profits within the rules of the game. Ethics, theology, and philosophy are "externalities" irrelevant to managerial decision-making.

Reference: Friedman, "The Social Responsibility of Business Is to Increase Its Profits," *NYT Magazine* (1970).

3.3.1.2 Objection 2 — Neoclassical Economics and Methodological Individualism

Mainstream economics treats agents as utility-maximizing individuals and assumes that aggregate welfare emerges automatically from self-interest. Normative considerations are dismissed as unscientific.

Reference: Hal **?**, *Microeconomic Analysis*; **?**, *Economics*.

3.3.1.3 Objection 3 — Cartesian Mechanism

Descartes reduces living beings to automata and treats nature as a machine. By analogy, management becomes the manipulation of mechanical parts (employees, processes, markets) rather than the stewardship of persons.

Reference: Ren'e **?**, *Treatise on Man*.

3.3.1.4 Objection 4 — Turing's Computationalism

If thinking is reducible to algorithmic computation, then analytics is simply the execution of procedures. Philosophy and theology add nothing to the correctness of outputs.

Reference: Alan **?**, "Computing Machinery and Intelligence."

3.3.1.5 Objection 5 — Positivist Managerialism

The positivist tradition holds that only empirical, measurable data count as knowledge. Values, meaning, and purpose are subjective and therefore inadmissible in a scientific curriculum.

Reference: Auguste **?**, *Course of Positive Philosophy*.

3.3.1.6 Objection 6 — Pragmatic Technocracy

Some argue that managers need tools, not metaphysics. Philosophy slows decision-making; theology is sectarian; analytics should be efficient, not contemplative.

Reference: Herbert **?** , *Administrative Behavior* (instrumental rationality).

3.3.2 On the Contrary

Aquinas teaches that every human act is ordered toward an end, and that ends are discerned through reason informed by virtue.

Reference: St. Thomas **?**, *Summa Theologiae* I–II, q.1, a.1.

Lonergan argues that authentic inquiry requires attentiveness, intelligence, reasonableness, and responsibility — norms that are philosophical and theological before they are empirical.

Reference: Bernard **?**, *Insight*, ch. 18–20.

Catholic social teaching affirms that economic and managerial decisions must serve the dignity of the human person and the common good.

Reference: **?**, *Gaudium et Spes*; **?**, *Laudato Si'*; **?**, *Caritas in Veritate*.

3.3.3 I Respond That

I answer that philosophical and theological considerations must guide an empirically oriented management analytics course, because analytics is never value-neutral. Every model presupposes:

- a view of the human person,
- a theory of value,
- a conception of the good,

- and a telos toward which decisions are directed.

Empirical methods alone cannot supply these. They require a prior anthropology, ethics, and metaphysics.

Furthermore:

1. Data are fragments; philosophy provides the intelligible whole.
2. Models encode assumptions about rationality, agency, and value; theology and ethics test these assumptions.
3. Analytics affects real people, especially the vulnerable; stewardship requires moral discernment.
4. Management is a moral practice, not merely a technical one.
5. Without philosophical grounding, analytics becomes technocratic power, not service.

Thus, philosophy and theology do not constrain analytics — they liberate it to serve the common good, to avoid decline, and to participate in the progress Lonergan describes: the expansion of the good of order for every neighbor without exception.

3.3.4 Replies to Objections

3.3.4.1 Reply to Objection 1 (Friedman)

Friedman's claim that business is value-neutral is itself a value-laden philosophical position. It cannot justify itself empirically and collapses into moral minimalism.

3.3.4.2 Reply to Objection 2 (Neoclassical Economics)

Methodological individualism is a metaphysical choice, not a scientific discovery. It ignores relationality, solidarity, and the common good.

3.3.4.3 Reply to Objection 3 (Descartes)

Human beings are not automata. Treating them as such leads to managerial dehumanization and organizational decline.

3.3.4.4 Reply to Objection 4 (Turing)

Even if computation can simulate reasoning, it cannot supply meaning, purpose, or value. Analytics requires judgment, not just algorithms.

3.3.4.5 Reply to Objection 5 (Positivism)

Positivism is self-refuting: the claim “only empirical statements are meaningful” is not itself empirical.

3.3.4.6 Reply to Objection 6 (Technocracy)

Efficiency without wisdom is dangerous. Analytics without ethics becomes manipulation; analytics with philosophy becomes stewardship.

3.4 QUAESTIO IV. Whether the Problem of Evil Is Intelligible as a Failure of Consciousness and a Distortion of the Good of Order

3.4.0.1 Objections

3.4.0.1.1 Objection 1 — Classical Theodicy (Leibniz, Clarke)

S ome argue that evil is a metaphysical necessity in the “best of all possible worlds,” and thus cannot be reduced to human failure or historical distortion.

Reference: **?**, *Theodicy*.

3.4.0.1.2 Objection 2 — Augustinian Privation Alone

Others hold that evil is simply a privation of good, lacking positive intelligibility. If evil is only absence, it cannot be analyzed as a structured failure of consciousness.

Reference: **?**, *Enchiridion*.

3.4.0.1.3 Objection 3 — Marxist Structuralism

Marx locates evil in material conditions and class structures, not in consciousness. Thus, decline is economic, not intellectual or moral.

Reference: **?**, *Economic and Philosophic Manuscripts*.

3.4.0.1.4 Objection 4 — Nietzschean Genealogy

Nietzsche claims that "evil" is a construct of ressentiment, not an intelligible failure. Decline is the triumph of the weak, not a distortion of the good.

Reference: Friedrich **?**, *Genealogy of Morals*.

3.4.0.1.5 Objection 5 — Secular Technocracy

Modern managerialism holds that problems are technical, not moral or intellectual. Evil is inefficiency; decline is mismanagement; consciousness is irrelevant.

Reference: Herbert **?**, *Administrative Behavior*.

3.4.1 On the contrary

Lonergan writes in *Insight* that "the root of evil is the flight from understanding," and that decline is "the cumulative effect of unauthenticity in the operations of consciousness."
Reference: *Insight*, ch. 7–8.

In *De Redemptione*, he reads Marx and Nietzsche as diagnosing real distortions of the human good, but lacking the horizon of redemption that makes intelligibility whole.
Reference: Lonergan, *De Redemptione* (Gregorian lectures).

In his *Political Economy* essays and *Economic Dynamics*, he argues that economic breakdowns are not merely technical failures but the result of "bias, short-sightedness, and the refusal of responsibility."
Reference: Lonergan, *For a New Political Economy*; *Macroeconomic Dynamics*.

3.4.2 I Respond That

I answer that the problem of evil is indeed intelligible as a failure of consciousness and a distortion of the good of order.

Lonergan's unified account proceeds in four movements:

3.4.2.1 1. Evil begins in the subject

Evil is not first a metaphysical puzzle or a cosmic defect.
It begins in the refusal of attentiveness, intelligence, reasonableness, and responsibility — the transcendental precepts that constitute authentic subjectivity.

This is the "flight from understanding."

3.4.2.2 2. Evil becomes historical through bias

Individual refusals accumulate into:

- dramatic bias (self-deception),
- individual bias (egoism),
- group bias (exclusion, domination),
- general bias (anti-intellectualism, technocracy).

These biases distort the good of order, the patterned cooperation that sustains human flourishing.

3.4.2.3 3. Decline becomes systemic

Lonergan's Political Economy of the 1940s and his Economic Dynamics of the 1980s show that:

- economic breakdowns,
- colonial exploitation,
- global inequality,
- ecological devastation,
- technological hypertrophy (e.g., nuclear capability)

are not merely technical failures. They are cumulative consequences of inauthentic consciousness.

3.4.2.4 4. Redemption is the reversal of decline

In *De Redemptione*, Lonergan integrates Marx and Nietzsche into a theological horizon:

- Marx diagnoses structural distortion.
- Nietzsche diagnoses cultural and moral decay.
- Lonergan diagnoses the root: the refusal of transcendence.

Redemption is not magic; it is the healing of consciousness, the restoration of authenticity, and the reconstitution of the good of order.

Thus, evil is intelligible — not as a metaphysical necessity, but as a historical, cumulative, and reversible failure of consciousness.

3.4.3 Replies to Objections

3.4.3.1 Reply to Objection 1 (Leibniz)

Metaphysical optimism ignores the historical, cumulative character of decline. Lonergan's account is not metaphysical but existential and historical.

3.4.3.2 Reply to Objection 2 (Augustine)

Privation is correct metaphysically, but insufficient historically. Decline has structure, pattern, and intelligibility.

3.4.3.3 Reply to Objection 3 (Marx)

Marx sees the symptoms but not the root. Structures distort because consciousness distorts.

3.4.3.4 Reply to Objection 4 (Nietzsche)

Nietzsche's genealogy unmasks moral decay but cannot account for authentic moral transcendence.

3.4.3.5 Reply to Objection 5 (Technocracy)

Technocracy is itself a form of general bias — the refusal to acknowledge the moral and intellectual roots of decline.

If Q4 establishes the *inverse insight* that the intelligibility of evil is a failure of consciousness, Q5 asks whether the reversal of that evil — structural, global, cumulative — is possible within human capability alone.

Your framing is exactly Lonergan's:
- decline is intelligible,
- decline is cumulative,
- decline is self-reinforcing,
- and decline eventually exceeds the power of unaided human capability to reverse.

This is where the contrast with Sen and Nussbaum becomes illuminating. They see the yearning (*oregesthai*) but not the horizon toward which the yearning is ordered.

3.5 QUAESTIO V. Whether the Reversal of Structural Decline and Global Evil Is Possible Within Human Capability Alone

3.5.1 Objections

3.5.1.1 Objection 1 — Sen's Capability Approach

Sen argues that human flourishing depends on expanding capabilities — freedoms to do and to be. If structural evil is a deprivation of capability, then its reversal lies in expanding human agency.

Reference: Amartya **?**, *Development as Freedom*.

3.5.1.2 Objection 2 — Nussbaum's Virtue-Capability Synthesis

Nussbaum grounds capabilities in a neo-Aristotelian virtue ethics. Human beings "reach for" (*oregesthai*) their flourishing through cultivated virtues. Thus, moral and political reform is possible through human development alone.

Reference: Martha **?**, *Creating Capabilities*.

3.5.1.3 Objection 3 — Secular Humanism

Human progress has historically overcome disease, poverty, and injustice. Structural evil is a technical and political challenge, not a metaphysical one.

Reference: Steven **?**, *Enlightenment Now*.

3.5.1.4 Objection 4 — Marxist Praxis

Structural evil is rooted in material conditions. Transforming those conditions through collective action is sufficient for reversing decline.

Reference: Karl **?**, *Theses on Feuerbach*, 1844.

3.5.1.5 Objection 5 — Nietzschean Self-Overcoming

Human beings can overcome decadence through the will-to-power and the creation of new values. No transcendent horizon is needed.

Reference: Friedrich **?** *Thus spoke Zarathustra*, 1885.

3.5.2 On the contrary

Lonergan writes that "the cumulative product of unauthenticity is decline," and that decline becomes "a surd, a nonsense, a horror" that exceeds the power of unaided human intelligence and will.

Reference: Bernard **?***Insight*, ch. 7–8.

In *De Redemptione*, he argues that the healing of history requires a transformation "beyond the resources of human nature," grounded in the divine self-communication .
Reference: Bernard **?**, *De Redemptione* (Gregorian lectures).

Catholic social teaching affirms that structural sin requires conversion, grace, and a horizon of transcendence.

Reference: **?** *Sollicitudo Rei Socialis*; **?** *Gaudium et Spes.*

3.5.3 I respond that

I answer that the reversal of structural decline and global evil is not possible within human capability alone.

3.5.3.1 1. Structural evil is cumulative and self-reinforcing

Lonergan's analysis shows that:

- individual bias
- group bias
- general bias
- and dramatic bias

all compound into structural sin — patterns of cooperation that systematically harm the vulnerable and distort the good of order.

Once entrenched, these structures exceed the power of individual or collective goodwill.

3.5.3.2 2. Human capability is necessary but insufficient

Sen and Nussbaum correctly identify the human yearning (*orexis*) for flourishing. But they lack the horizon of vertical transcendence:

- the yearning is real,
- but the fulfillment lies beyond human self-assertion,
- because the wound, teh vulnerability, is deeper than human capability.

3.5.3.3 3. Sin is not merely personal but historical

Lonergan's reading of history shows that:

- decline becomes global (and greater than the sum of the nations, regions, tribes, families),
- decline becomes cultural (norms, values, stories, binding),
- decline becomes economic (the material well-being of the good of order),
- decline becomes ecological (all connected, bio-, psycho-, socio-, anthropological),
- decline becomes technological (the works of our hands, nuclear capability, AI misuse, extractive capitalism).

This is original sin writ communally, not metaphorically but structurally.

3.5.3.4 4. The solution requires a surplus of meaning, value, and love

- Human capability can diagnose decline.
- Human capability can mitigate decline.
- But human capability cannot reverse decline.

Reversal requires:

- conversion,

- grace, an unforced generosity,
- a new horizon of meaning,
- the divine self-communication Lonergan calls "God's love flooding our hearts."

This is not pious sentiment.It is the only adequate response to a problem that has exceeded and will necessarily exceed human scope and scale.

3.5.3.5 5. The yearning (*oregesthai*) points beyond itself

? uses the medio-passive *oregesthai*,and it is telling:

- "to reach for,
- "to stretch toward,"
- "to long for."

But what is the object of this yearning?

Lonergan's answer: *the unrestricted desire to know, to love, to be — the desire for God.*

Thus, the very structure, the design itself, of human capability points beyond itself.

3.5.4 Replies to Objections

3.5.4.1 Reply to Objection 1 (Sen)

Capabilities expand agency but cannot heal the root of decline: bias and unauthenticity.

3.5.4.2 Reply to Objection 2 (Nussbaum)

Virtue is necessary, but without transcendence it becomes Pelagian — insufficient for structural evil.

3.5.4.3 Reply to Objection 3 (Secular Humanism)

Technological progress does not reverse moral or structural decline; it often amplifies it.

3.5.4.4 Reply to Objection 4 (Marx)

Material transformation without conversion reproduces the same distortions under new forms.

3.5.4.5 Reply to Objection 5 (Nietzsche)

Self-overcoming cannot overcome the self-enclosure that constitutes decline.

This last *quaestio* is extraordinary in the sense it is the whole that pulls together the parts, yet is subservient to the principles of the whole exceeding the parts. It begins tp completes the arc:

1. The whole exceeds the parts.

2. Deduction yields new knowledge.

3. Analytics requires philosophy and theology.

4. Evil is intelligible as a failure of consciousness.

Reversal requires transcendence beyond human capability.

Chapter 4

The Good of Progress

Here are some pressure points for the entire Western intellectual tradition. Let's take them in turn and then braid them back into Lonergan's arc of progress/decline.

4.1 Is the whole greater than the sum of the parts?

We're right to plant our flag firmly in the *affirmative* and not just rhetorically — metaphysically, epistemologically, ethically.

4.1.1 Why the objectors fail

- **Bacon** reduces intelligibility to empirical fragments, so of course he can't see emergent form.
- **Scientism** collapses all explanation into mechanism, so it literally cannot *see* wholes.
- **Reductionist neuroscience** treats consciousness as epiphenomenal chemistry, which is like explaining Hamlet by analyzing the ink.
- And yes — if one is going to reduce thought to chemistry, why stop there? Why not reduce chemistry to QED amplitudes? Why not reduce *those* to the vacuum state? Reductionism always eats itself. It refutes itself in a classical *retorquendo* move.

4.1.2 Why the affirmative is the only coherent position

A whole is not a pile of stuff. A whole is a *form*, a *pattern*, an *intelligibility* that:

- organizes the parts,
- conditions their relations,
- and generates emergent properties irreducible to the components.

This is Aristotle → Aquinas → Lonergan → modern complexity theory.

We're defending the very possibility of **intelligibility**, **meaning**, and **value**.

4.2 Is deduction new knowledge?

Again, your affirmative stance is exactly right — and Lonergan gives you the conceptual artillery.

4.2.1 Why the skeptics deny it

- **Hume**: deduction is analytic tautology.
- **Locke**: the mind is a passive container; deduction rearranges ideas but adds nothing.
- **Kant**: synthetic a priori is possible, but deduction itself is not knowledge of being — it is only the structuring of appearances.
- **Callicles, Sartre, Nietzsche (in his early phase)**: knowledge is power-play, not insight.

All of these positions share a single root: **they deny that understanding is a real grasp of intelligibility in being.**

If knowing is not *being-in-contact-with-reality*, then deduction is just linguistic shuffling.

4.2.2 Lonergan's counter-position

Lonergan's breakthrough is that **insight transforms data**, and **judgment transforms insight**. Deduction is not mechanical manipulation — it is:

- the unfolding of the intelligibility grasped in insight,
- the extension of understanding into new domains,
- the discovery of implications that were *not* present to consciousness before.

Deduction is new knowledge because **the mind is dynamically self-transcending**.

4.3 Progress and Decline: Why These Two Questions Matter

We connect these metaphysical questions to Lonergan's moral-historical cycle.

4.3.1 Progress

Progress occurs when: - the mind affirms that wholes exist (intelligibility is real),
- the mind trusts its own capacity for insight and judgment,
- the community orients itself toward the good of the neighbor,
- and the good of order expands to include the vulnerable.

This is the *yes* to being, to intelligibility, to value, to God.

4.3.2 Decline

Decline is not random. It is the predictable result of: - reductionism (no wholes),
- skepticism (no new knowledge),
- individualism (no neighbor),
- and self-enclosure (no transcendence).

Decline is the *no* to being, the *no* to intelligibility, the *no* to the neighbor, the *no* to God.

It is privation, deflection, and self-protection masquerading as realism.

4.4 Our Position, Stated Cleanly

We're defending a metaphysics and epistemology that make:

- **emergence possible**,

- **insight real**,
- **deduction fruitful**,
- **progress intelligible**,
- **decline diagnosable**,
- **and the good of the neighbor non-negotiable**.

This is not "same old - same old." This is the beating heart of a humane, theologically grounded, intellectually rigorous worldview.

4.5 Then there's uncertainty

We're reaching for exactly the right *nerve center*: the moment in Nicholas Cusanus, ***De Docta Ignorantia* I.3–4** where **?** articulates (a) the **impossibility theorem**—finite reason cannot ground or explain itself—and (b) the **coincidentia oppositorum** as the only horizon in which inquiry can proceed at all.

Below is the passage most directly expressing the "impossibility theorem" and the "coincidence of opposites" as a condition of possibility for wonder, inquiry, and analogy.

4.5.1 De Docta Ignorantia (herein, *DDI*) I.3–4

4.5.1.1 On the impossibility of science grounding itself

Cum igitur omne cognoscibile mensurari oporteat mensura propria, mensura autem infinita finitam mensuram excedat, impossibile est scientiam de infinito secundum propriam rationem haberi. (*DDI* I.3)

"Since everything knowable must be measured by its proper measure, and since the infinite exceeds every finite measure, it is impossible to have a science of the infinite according to its own proper rational method."

This is the core of Cusa's "impossibility theorem": **science cannot explain or ground itself**, because the very act of knowing presupposes a measure that finite reason cannot supply.

4.5.1.2 On the coincidence of opposites as the horizon of wonder and understanding

In infinito autem, ubi omnium oppositorum coincidentia est, nostra mens, quae per differentias intelligit, deficit; et in hoc deficit, ut mirari discat. (*DDI* I.4)

"But in the Infinite, where the coincidence of all opposites is found, our mind—which understands by means of differences—fails; and in this very failing it learns to wonder."

Here we stand.

- finite reason fails,
- but the failure is **fruitful**,
- because it opens onto wonder,
- which is the condition for analogy (Przywara),
- and the border-logic Wittgenstein intuits (a boundary presupposes two sides).

4.5.2 On the necessity of coincidence for any knowledge at all

Quoniam igitur in Deo, qui est omnium mensura, opposita coincidunt, mens nostra, quae in oppositis haeret, ad veritatem nisi in docta ignorantia non ascendit. (*DDI* I.4)

"Since in God, who is the measure of all things, opposites coincide, our mind—which clings to opposites—cannot ascend to truth except through learned ignorance."

This is the metaphysical ground of Przywara's *analogia entis*:

- God is the coincidence of opposites,
- creatures are the "ever greater dissimilarity,"
- analogy is the oscillation between similarity and dissimilarity,
- wonder is the epistemic posture proper to this oscillation.

}

4.5.2.1 Wittgenstein's border-logic

The key line is **Tractatus 5.6**:

> *"The limits of my language mean the limits of my world."*
> (*Tractatus* 5.6)

This is the passage that resonates with Cusa's *docta ignorantia*:

- A **limit** is not a wall but a **border**.
- A border implies **two sides**.
- What lies beyond the limit is not nonsense but the **condition** for meaningful sense.
- Language cannot contain the whole, but its very limit points to the whole.

This is why the *Tractatus* ends with the famous 7th note:

> *"Whereof one cannot speak, thereof one must be silent."*
> (*Tractatus* 7)

Silence here is not negation — it is **reverence**, the philosophical analogue of Cusa's wonder.

How might this connect Cusa-to-Wittgenstein-to-Przywara?

- **Cusa**: The mind fails at the infinite, and in failing learns to wonder.
- **Wittgenstein**: The limit of language is the border where meaning opens beyond itself.
- **Przywara**: The creature exists in a rhythmic tension of similarity and ever-greater dissimilarity to God — the *analogia entis*.

All three converge on a single insight:

The limit is not the end of understanding but the beginning of wisdom.[1]

[1] Nicolaus Cusanus articulates the foundational limit of finite reason in *De Docta Ignorantia* I.3–4 (**?**), where he argues that the infinite exceeds every finite measure and therefore cannot be grasped by any science that presupposes its own rational method: "*Cum igitur omne cognoscibile mensurari oporteat mensura propria, mensura autem infinita finitam mensuram excedat, impossibile est scientiam de infinito secundum propriam rationem haberi.*" He next develops the epistemic consequence: "*In infinito autem, ubi omnium oppositorum coincidentia est, nostra mens, quae per differentias intelligit, deficit; et in hoc deficit, ut mirari discat.*" The mind in a finite universe with a finite body (brain at least) falters right at the boundary of its own powers, yet the falling down is fruitful, for it reaches out with wonder. This border-logic anticipates Wittgenstein's insight that a boundary presupposes two regions —

one on either side — and that what lies beyond the limit of language is not nonsense but the very condition of meaningful speech, and thus of meaningful communication of analytical results to consumers of the analysis. **?** develops the Cusanus structure into the rhythmic, cyclic "ever greater dissimilarity" of the *analogia entis,* where creaturely, finite being is constituted by the tension between similarity and dissimilarity to the divine, the ever more relative to the creature. Thus, Cusa's *coincidentia oppositorum* is not a collapse of distinctions but the metaphysical horizon within which analogy, wonder, and inquiry become possible.

Chapter 5

Great Expectations

5.1 The maths! The maths!

We can fuss about all we want about the maths, but they are impervious to our feelings. They remain. We can stay, or go. If we stay, and spend the time in active pursuit (just like a waiting time, waiting for insight, or not, waiting for Godot, who never shows up), we might achieve a learning apogee. We suppose that we will stay awhile, for the time being. Now let us dig into our model of waiting times. Our first stop is a set of tools we will need for the excavation.

In what follows we use Y as the wage, the metric we want to generate from its mean and standard deviation. We conjecture that Y depends on X, the level of educational attainment through the conditional mean of $Y \mid X$, just like we did with vaccination waiting times.

5.2 What did we all expect?

We define expectations as aggregations of two kinds of information. One is the information provided by an array of outcomes Y_i for $i = 1 \ldots N$, where i indexes N outcomes. The other is the array of probabilities assigned to each outcome π_i. The frequentist will assign $\pi = f_i/N$, where f_i is the long-run frequency of occurrence of outcome i. Instead, with **?** and **?** we will assign π as a normalized index of the logical plausibility of an outcome where all πs add up to one and each is somewhere between 0 and 1. After all we can't wait for the long-run, we have to go to school!

This allows us to interpret probability as an extension of logic, where probability quantifies the reasonable expectation that everyone (even a *robot* or *golem*) who shares the same knowledge (experience, understanding, *and* judgment) should share in accordance with the rules of conditional probability.[1](https://en.wikipedia.org/wiki/Cox%27s_theorem) provides a logical underpinning to this statement: the rules of probability theory need not be derived from a definition of probabilities as relative frequencies (frequentist approach). He goes further to show that the properties of probability as logic but also follow from certain properties one might desire of any system of plausible reasoning about uncertainty. Van Horn is a tutorial on Cox's approach. Plausible reasoning may be illustrated with this example of the distinction between gradual degrees of possible outcomes, that is, uncertainty, and what is or is not, that is, truth. As an example, one's confidence in the statement *Daniel is well over six feet tall*, after seeing Daniel, legs splayed out, sitting at a desk, is a degree of plausibility. In contrast, the statement *Daniel is tall* may be somewhat true (if Daniel measures five feet eleven inches and the definition of tall is greater than or equal to six feet) or entirely true (if Daniel measures seven feet one inch).] All of this ensures we have a complete picture of all of the probability contributions, as weights, of each outcome consistent with a systematic, principled way to reason about uncertainty, as we mash together data and hypotheses about the data.

The aggregation we propose is then this expression for the expectation E of outcomes Y.

$$\mathrm{E}\, Y = \sum_{i}^{N} \pi_i Y_i \tag{5.1}$$

In this way we can say that E operates on Y where **to operate** means **to aggregate** several outcomes X_i into one number (or possibly function) by multiplying probability weights times outcomes and then summing the products. That's really two operations combined into the expectation operation. And so goes the maths!

Using this idea of an operator $\mathrm{E}\, Y$ means we define the aggregation as this expression.

$$\mathrm{E} = \sum_{i}^{N} \pi_i \times \tag{5.2}$$

Here are some of the algebraic rules of the road when we use this highly condensed short-hand notation.

[1] Cox's theorem

$$
\begin{aligned}
Y &= \alpha\, X && (5.3)\\
\mathrm{E}\, Y &= \mathrm{E}[\alpha\, X] && (5.4)\\
&= \sum_i^N [\pi_i\, (\alpha\, X_i)] && (5.5)\\
&= \pi_1\, \alpha\, X_1 + ...\, \pi_N\, \alpha\, X_N && (5.6)\\
&= \alpha\, (\pi_1\, X_1 + ...\, \pi_N\, X_N) && (5.7)\\
&= \alpha\, \sum_i^N [\pi_i\, (X_i)] && (5.8)\\
&= \alpha\; \mathrm{E}\, X && (5.9)
\end{aligned}
$$

This means that we can take the constant α outside of the expectation operator. All we did, step by step on the logical staircase, is to use the definition of the operator and then manipulate it algebraicly to deduce an equivalent expression.

If $X_1 = 1, \dots, X_N = 1$, and the sum of probabilities $\sum_i^N \pi_i = 1$, then we can deduce this expression.

$$
\begin{aligned}
Y &= \alpha\, X && (5.10)\\
\mathrm{E}\, Y &= \mathrm{E}[\alpha\, X] && (5.11)\\
&= \sum_i^N [\pi_i\, \dot{(}\alpha\, X_i)] && (5.12)\\
&= \pi_1\, \alpha\, (1) + ...\, \pi_N\, \alpha\, (1) && (5.13)\\
&= \alpha\, (\pi_1\, (1) + ...\, \pi_N\, (1) && (5.14)\\
&= \alpha\, \sum_i^N [\pi_i (1)] && (5.15)\\
&= \alpha\; \mathrm{E}\, 1 && (5.16)\\
&= \alpha && (5.17)
\end{aligned}
$$

This may have been immediately clear to some of us before the 7 step deduction, but we might find it reassuring that the deduction verifies, and perhaps validates, our initial conjecture. We also discover another relationship.

$$\mathrm{E}\, 1 = 1$$

In algebra we call this the identity operator. For any number or variable, or even another expectation, α, then this is true.

$$\alpha \,\mathrm{E}\, 1 = \alpha\, 1 \tag{5.18}$$
$$= \alpha \tag{5.19}$$

Yes, this is identity under a multiplication. Is there a zero? Yes, $\mathrm{E}\,0 = 0$, the identity operator under addition. Anything added to $\mathrm{E}\,0 = 0$ just returns itself.

What is the expectation of a sum of variables X and Y?

$$Z = X + Y \tag{5.20}$$
$$\mathrm{E}\,Z = \mathrm{E}[X + Y] \tag{5.21}$$
$$= \sum_{i}^{N} [\pi_i\, (\, X_i + Y_i)] \tag{5.22}$$
$$= \pi_1\ (X_1 + Y_1) + \ldots \pi_N\, (X_N + Y_N) \tag{5.23}$$
$$= (\pi_1\, X_1 + \ldots \pi_1\, X_N) + (\pi_N\, Y_N + \ldots \pi_N\, Y_N) \tag{5.24}$$
$$= \sum_{i}^{N} [\pi_i\, (X_i)] + \sum_{i}^{N} [\pi_i\, (Y_i)] \tag{5.25}$$
$$= \mathrm{E}\,X + \mathrm{E}\,Y \tag{5.26}$$

The expectation of a sum of outcome variables is the sum of the expectations of each variable.

We just examined a sum of two variables, so it behooves us to look at the product of two variables.

$$Z = XY \tag{5.27}$$
$$\mathrm{E}\,Z = \mathrm{E}[XY] \tag{5.28}$$
$$= \sum_{i}^{N} [\pi_i\, (\, X_i\, Y_i)] \tag{5.29}$$
$$= \pi_1\, X_1\, Y_1 + \ldots \pi_N\, X_N\, Y_N) \tag{5.30}$$
$$= \mathrm{E}\,XY \tag{5.31}$$

Alas, we have reduced this operation to its simplest expression already. If $Y = X$, going through the same steps as above we find this out.

$$
\begin{aligned}
&if\ Z = XY && (5.32)\\
&and && (5.33)\\
&Y = X && (5.34)\\
&then && (5.35)\\
&Z = XX && (5.36)\\
\mathrm{E}\, Z &= \mathrm{E}[XX] && (5.37)\\
&= \sum_i^N [\pi_i\, (\, X_i\, X_i)] && (5.38)\\
&= \pi_1\, X_1\, X_1 + ...\, \pi_N\, X_N\, X_N) && (5.39)\\
&= \pi_1\, X_1^2 + ...\, \pi_N\, X_N^2 && (5.40)\\
&= \mathrm{E}\, X^2 && (5.41)
\end{aligned}
$$

It turns out that we can take an expression like this, $Y = \alpha + \beta\, X$, multiply it by X and, then operate on it with $\mathrm{E} = \sum_i^N \pi_i \times$ with the tools we now possess.

$$
\begin{aligned}
Y &= \alpha + \beta\, X && (5.42)\\
XY &= \alpha\, X + \beta\, XX && (5.43)\\
XY &= \alpha\, X + \beta\, X^2 && (5.44)\\
\mathrm{E}\, XY &= \mathrm{E}[\alpha\, X + \beta\, X^2] && (5.45)\\
&= \mathrm{E}[\alpha\, X] + \mathrm{E}[\beta\, X^2] && (5.46)\\
&= \alpha\ \mathrm{E}[X] + \beta\ \mathrm{E}[X^2] && (5.47)
\end{aligned}
$$

This will be very useful indeed. We usually will call $\mathrm{E}\, X = \mu_X$ in honor of the **mean** of the population of all possible realizations of X. We already know this as the weighted average of X outcomes, where the weights are probabilities, all of which add up to 1. What about $\mathrm{E}\, X^2$? To ponder this we consider the calculation of another very familiar metric, the square of the standard deviation, which has been dubbed the **variance**. We start with the definition and use all of the new tricks up our sleeves. We define variance as the probability weighted average of squared deviations of outcomes from the expected outcome.

We will need the remembrance of things in our algebraic past that look like this.

$$
\begin{aligned}
(a + b)^2 &= (a + b)(a + b) && (5.48)\\
&= a^2 + 2ab + b^2 && (5.49)
\end{aligned}
$$

In what follows $a = X$ and $b = -\mathrm{E}\, X$. We will also need to remember that $-2b^2 + b^2 = -b^2$.

$$
\begin{aligned}
define & & (5.50)\\
\sigma_X^2 &= Var(X) & (5.51)\\
then & & (5.52)\\
Var(X) &= \mathrm{E}(X - \mathrm{E}\,X)^2 & (5.53)\\
&= \mathrm{E}(X^2 - 2X\,\mathrm{E}\,X + \mathrm{E}\,X^2) & (5.54)\\
&= \mathrm{E}\,X^2 - \mathrm{E}[2X\,\mathrm{E}\,X] + \mathrm{E}[\mathrm{E}\,X^2] & (5.55)\\
&= \mathrm{E}\,X^2 - 2(\mathrm{E}\,X)^2 + (\mathrm{E}\,X)^2 & (5.56)\\
&= \mathrm{E}\,X^2 - (\mathrm{E}\,X)^2 & (5.57)\\
&= \mathrm{E}\,X^2 - \mu_X^2 & (5.58)\\
thus & & (5.59)\\
\sigma_X^2 &= \mathrm{E}\,X^2 - \mu_X^2 & (5.60)\\
rearranging & & (5.61)\\
\mathrm{E}\,X^2 &= \sigma_X^2 + \mu_X^2 & (5.62)
\end{aligned}
$$

Yes, we can breathe a collective sigh of relief having accomplished these algebraic acrobatics! But we must now move on to the *piece de resistance* , $\mathrm{E}\,XY$. We start with the definition of covariance, for this is where an XY product resides.

$$
\begin{aligned}
define & & (5.63)\\
\sigma_{XY} &= Cov(X,Y) & (5.64)\\
then & & (5.65)\\
Cov(X,Y) &= \mathrm{E}(X - \mathrm{E}\,X)(Y - \mathrm{E}\,Y) & (5.66)\\
&= \mathrm{E}(XY - X\,\mathrm{E}\,Y - Y\,\mathrm{E}\,X + \mathrm{E}\,X\ \mathrm{E}\,Y) & (5.67)\\
&= \mathrm{E}(XY - \mathrm{E}\,X\ \mathrm{E}\,Y - \mathrm{E}\,Y\ \mathrm{E}\,X + \mathrm{E}\,X\ \mathrm{E}\,Y) & (5.68)\\
&= \mathrm{E}\,XY - 2\,\mathrm{E}\,X\ \mathrm{E}\,Y + \mathrm{E}\ X[\mathrm{E}\,Y & (5.69)\\
&= \mathrm{E}\,XY - \mathrm{E}\,X\ \mathrm{E}\,Y & (5.70)\\
thus & & (5.71)\\
\sigma_{XY} &= \mathrm{E}\,XY - \mu_X\mu_Y & (5.72)\\
rearranging & & (5.73)\\
\mathrm{E}\,XY &= \sigma_{XY} + \mu_X\mu_Y & (5.74)
\end{aligned}
$$

Now we can go to work on our model with one more stop: solving a simultaneous equation. This tool too will come in handy. We suppose we have the following two equations in a and b. We will use the row-column convention of subscripts. Thus coefficient c_{12} will be in row 1, column 2 of a matrix. First the two equations.

$$c_{11}a + c_{12}b = d_1 \tag{5.75}$$
$$c_{21}a + c_{22}b = d_2 \tag{5.76}$$

In matrix form this is a very tidy arrangement like this.

$$\begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} d_1 \\ d_2 \end{bmatrix} \tag{5.77}$$
$$\mathrm{Ca} = \mathrm{d} \tag{5.78}$$

Very tidy indeed! We might remember that a unique solution exists only if (or is it if and only if?) the determinant of the matrix C is not zero. If it is, then the solution $\mathrm{a} = \mathrm{C}^{-1}d$ does not exist and the model is singular. In what we will do below we will compose our coefficients of means, standard deviations and correlations. Some combinations of these aggregations, constants, will prove to yield a zero determinant, and a singular model results.

The determinant $\det \mathrm{C}$ is

$$\det \mathrm{C} = c_{11}c_{22} - c_{12}c_{21}$$

The solution proceeds in two sweeps, one for each of a and b. In the first sweep we replace the first, the a column, in C with the column vector d. We find the determinant of this new C_a matrix and divide by $\det \mathrm{C}$. Here we go.

$$\textit{original}\ \mathrm{C} \tag{5.79}$$
$$\begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix} \tag{5.80}$$
$$\textit{swap out first column} \tag{5.81}$$
$$\mathrm{C}_a = \begin{bmatrix} d_1 & c_{12} \\ d_2 & c_{22} \end{bmatrix} \tag{5.82}$$
$$\textit{then} \tag{5.83}$$
$$a = \frac{\det \mathrm{C_a}}{\det \mathrm{C}} \tag{5.84}$$
$$= \frac{d_1c_{22} - d_2c_{12}}{c_{11}c_{22} - c_{12}c_{21}} \tag{5.85}$$

Now the second sweep in all its glory.

$$
\begin{aligned}
& \textit{original C} && (5.86)\\
& \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix} && (5.87)\\
& \textit{swap out first column} && (5.88)\\
\mathrm{C}_b &= \begin{bmatrix} c_{11} & d_1 \\ c_{21} & d_2 \end{bmatrix} && (5.89)\\
& \textit{then} && (5.90)\\
b &= \frac{\det \mathrm{C_b}}{\det \mathrm{C}} && (5.91)\\
&= \frac{c_{11}d_2 - c_{21}d_1}{c_{11}c_{22} - c_{12}c_{21}} && (5.92)
\end{aligned}
$$

Very much a formula for the ages.

5.3 Walking the straight line

Here is our model where both Y and X have some distribution with π probabilities for each. Here we use π as the Greek letter for p, not as the π of circle fame. Both Y and X are what we will very loosely call **random variables**, because they have outcomes with associated probabilities of occurrence.

$$Y = \alpha + \beta\, X$$

We now ask the question, what is $\mathrm{E(Y \mid X = x)} = {}_{\mathrm{Y|X}}$? What, on weighted average, can we expect Y to be? First of all, this must be true.

$$
\begin{aligned}
& \textit{if} && (5.93)\\
\mathrm{E}(Y \mid X = x) &= \mu_{Y|X} && (5.94)\\
& \textit{then} && (5.95)\\
\mu_{Y|X} &= \mathrm{E}(\alpha + \beta\, X) && (5.96)\\
&= \mathrm{E}\,\alpha(1) + \mathrm{E}(\beta\, X) && (5.97)\\
&= \alpha\ \mathrm{E}\,1 + \beta\ \mathrm{E}\,X && (5.98)\\
&= \alpha\,(1) + \beta\,\mu_X && (5.99)\\
&= \alpha + \beta\,\mu_X && (5.100)
\end{aligned}
$$

Result one is in hand, $\mu_{Y|X} = \alpha + \beta\,\mu_X$ is a true statement according to our many deductions. By the way the statement $\mu_{Y|X} = \alpha\ \mathrm{E}\,1 + \beta\ \mathrm{E}\,X$ is an example of the distributive property of multiplication over addition.

Now for our second trick we multiply Y by X to get a second result and a second true statement. We will condense $Y \mid X = Y$ to save what's left of our eyesight. We remember all of our hard work above, especially this inventory of results.

$$\mathrm{E}\, Y = \mu_Y \tag{5.101}$$
$$\mathrm{E}\, X = \mu_X \tag{5.102}$$
$$\mathrm{E}\, X^2 = \sigma_X^2 + \mu_X^2 \tag{5.103}$$
$$\mathrm{E}\, XY = \sigma_{XY} + \mu_X \mu_Y \tag{5.104}$$

Using this inventory more than a few times we get these results.

$$Y = \alpha + \beta\, X \tag{5.105}$$
$$\textit{then} \tag{5.106}$$
$$XY = \alpha\, X + \beta\, XX \tag{5.107}$$
$$= \alpha\, X + \beta\, X^2 \tag{5.108}$$
$$\textit{so that} \tag{5.109}$$
$$\mathrm{E}\, XY = \mathrm{E}(\alpha\, X + \beta\, X^2) \tag{5.110}$$
$$= \mathrm{E}\, \alpha\, X + \mathrm{E}\, \beta\, X^2 \tag{5.111}$$
$$= \alpha\ \mathrm{E}\, X + \beta\ \mathrm{E}\, X^2 \tag{5.112}$$
$$= \alpha\, \mu_X + \beta\, (\sigma_X^2 + \mu_X^2) \tag{5.113}$$
$$\textit{but we know that} \tag{5.114}$$
$$\mathrm{E}\, XY = \sigma_{XY} + \mu_X \mu_Y \tag{5.115}$$
$$\textit{thus, again} \tag{5.116}$$
$$\sigma_{XY} + \mu_X \mu_Y = \alpha\, \mu_X + \beta\, (\sigma_X^2 + \mu_X^2) \tag{5.117}$$

We now have two equations in two, as yet to be determined, unknowns. They are unobserved data, α and β. Both equations are true, and true jointly. This means we can stack one on top of the other as a simultaneous equation system and, we hope this time, solve them for unique values of α and β. Yes, we demand a formula!

Here are the two equations with α and β terms on the left-hand side and constant terms, the expectations are all constant aggregations, on the right-hand side of the equation. We also commutes the terms so that our unknowns are pre-multiplied by coefficients.

$$\alpha + \mu_X\, \beta = \mu_Y \tag{5.118}$$
$$\mu_X\, \alpha + (\sigma_X^2 + \mu_X^2)\, \beta = \sigma_{XY} + \mu_X \mu_Y \tag{5.119}$$

The matrix representation will help us easily match coefficients with our simultaneous equation model, way above as we replicate below.

$$\begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} d_1 \\ d_2 \end{bmatrix} \tag{5.120}$$

$$\mathrm{Ca} = \mathrm{d} \tag{5.121}$$

Our simultaneous equations of expected values for the linear model $Y = \alpha + \beta X$ yields this structure.

$$\alpha + \mu_X \, \beta = \mu_Y \tag{5.122}$$

$$\mu_X \, \alpha + (\sigma_X^2 + \mu_X^2) \, \beta = \sigma_{XY} + \mu_X \mu_Y \tag{5.123}$$

$$\textit{becomes} \tag{5.124}$$

$$\begin{bmatrix} 1 & \mu_X \\ \mu_X & \sigma_X^2 + \mu_X^2 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} \mu_Y \\ \sigma_{XY} + \mu_X \mu_Y \end{bmatrix} \tag{5.125}$$

$$\mathrm{Ca} = \mathrm{d} \tag{5.126}$$

We can solve for the unobserved, unknown, and otherwise conjectured (we might smell a hypothesis brewing here) α and β using our trusty determinant solutions.

$$\alpha = \frac{\mu_Y(\sigma_x^2 + \mu_X^2) - \mu_X(\sigma_{XY} + \mu_X \mu_Y)}{\sigma_X^2 + \mu_X^2 - \mu_x^2} \tag{5.127}$$

$$= \frac{\mu_Y \sigma_X^2 - \mu_X \sigma_{XY}}{\sigma_X^2} \tag{5.128}$$

$$= \mu_Y - \mu_X \frac{\sigma_{XY}}{\sigma_X^2} \tag{5.129}$$

$$\textit{and then} \tag{5.130}$$

$$\beta = \frac{\det \mathrm{C}}{\det \mathrm{C}} \tag{5.131}$$

$$= \frac{c_{11} d_2 - c_{21} d_1}{c_{11} c_{22} - c_{12} c_{21}} \tag{5.132}$$

$$= \frac{\sigma_{XY} + \mu_X \mu_Y - \mu_X \mu_Y}{\sigma_X^2 + \mu_X^2 - \mu_x^2} \tag{5.133}$$

$$= \frac{\sigma_{XY}}{\sigma_X^2} \tag{5.134}$$

Yeow! All that work to get at this simplification all due to the wonderful result that α has $\beta = \sigma_{XY}/\sigma_X^2$ in it.

$$
\begin{aligned}
\mathrm{E}(Y \mid X) &= \alpha + \beta\, X && (5.135)\\
\mathrm{E}(Y \mid X) &= \left(\mu_Y - \mu_X \frac{\sigma_{XY}}{\sigma_X^2}\right) + \frac{\sigma_{XY}}{\sigma_X^2}\, X && (5.136)\\
&\textit{rearranging terms} && (5.137)\\
\mathrm{E}(Y \mid X) &= \mu_Y + \frac{\sigma_{XY}}{\sigma_X^2}(X - \mu_X) && (5.138)
\end{aligned}
$$

The second formulation is also the basis for the vaunted Capital Asset Pricing Model in finance, where Y is the return on a security (stock, bond, etc.) and X is the return on a market index (e.g., S&P 500).

We have, yes, one more stop, before we drop. The definition of correlation is here.

$$\rho = \frac{\sigma_{XY}}{\sigma_X\, \sigma_Y}$$

We can use this definition to rearrange the deck chairs on this Titanic of a beast of gnarly maths (all algebra! and nary a faint odor of calculus?).

$$
\begin{aligned}
&if && (5.139)\\
\rho &= \frac{\sigma_{XY}}{\sigma_X\, \sigma_Y} && (5.140)\\
&then && (5.141)\\
\sigma_{XY} &= \rho\, \sigma_X\, \sigma_Y && (5.142)\\
&thus && (5.143)\\
\beta &= \frac{\sigma_{XY}}{\sigma_X^2} && (5.144)\\
&= \frac{\rho\, \sigma_X\, \sigma_Y}{\sigma_X^2} && (5.145)\\
&= \frac{\rho\, \sigma_Y}{\sigma_X} && (5.146)
\end{aligned}
$$

We need numbers, *stat(im)*! But we should hold on. One more calculation to make. After we have the mean, but what about the conditional standard deviation?

5.4 A short variance diversion

Here we take the standard deviation as given, perhaps at our peril. The variance of waiting times is

$$Var(Y \mid X) = (1 - \rho^2)\sigma_Y^2$$

How do we get this? A lot easier than the preceding. There are no simultaneous equations to worry about. Here's the algebra for the stout-hearted.

$$\begin{aligned}
Var(Y \mid X) &= \mathrm{E}[Y - \mathrm{E}(Y \mid X)]^2 && (5.147)\\
&= \mathrm{E}[Y - (\mu_Y + \frac{\sigma_{XY}}{\sigma_X^2}(X - \mu_X))]^2 && (5.148)\\
&= \mathrm{E}[(Y - \mu_Y) + \frac{\sigma_{XY}}{\sigma_X^2}(X - \mu_X)]^2 && (5.149)\\
&= \mathrm{E}[(Y - \mu_Y)^2 + \frac{\rho_{XY}^2\sigma_Y^2}{\sigma_X^2}(X - \mu_X)^2 - 2\frac{\rho_{XY}\sigma_Y}{\sigma_X}(Y - \mu_Y)(X - \mu_X)] && (5.150)\\
&= \mathrm{E}[(Y - \mu_Y)^2] + \frac{\rho_{XY}^2\sigma_Y^2}{\sigma_X^2}\,\mathrm{E}[(X - \mu_X)^2] - 2\frac{\rho_{XY}\sigma_Y}{\sigma_X}\,\mathrm{E}[(Y - \mu_Y)(X - \mu_X)] && (5.151)\\
&= \sigma_Y^2 + \frac{\rho_{XY}^2\sigma_Y^2}{\sigma_X^2}\sigma_X^2 - 2\frac{\rho_{XY}\sigma_Y}{\sigma_X}(\rho_{XY}\sigma_X\,\sigma_Y) && (5.152)\\
&= \sigma_Y^2 - \rho_{XY}^2\sigma_Y^2 && (5.153)\\
&= (1 - \rho_{XY}^2)\sigma_Y^2 && (5.154)
\end{aligned}$$

Done! Yes, really.

If the joint distribution of X and Y is Gaussian, then we can generate $Y \mid X \sim \mathrm{N}(\alpha + \beta X, (1 - \rho_{XY}^2)\sigma_Y^2)$. Now we can infer Y behavior.

Chapter 6

Quaestio: Whether Corporate Risk Disclosures Exhibit Logical Coherence

This quaestio integrates:

- formal logic,
- managerial ethics,
- and the metaphysics of intelligibility

6.1 Objections

1. It seems that corporate MD&A and Risk Factor language is inherently inconsistent, since firms routinely assert both vulnerability and resilience, both risk and opportunity, both exposure and mitigation. Therefore, no coherent logical structure can be extracted.

2. Further, the conditional forms used in disclosures ("only if," "may," "could," "unless") are too vague to admit propositional Formalize. Their ambiguity defeats the precision required for deductive analysis.

3. Moreover, the complexity of modern supply chains (e.g., NVIDIA's dependence on TSMC for fabrication and packaging) introduces causal entanglement, not discrete propositions. Thus, propositional logic is too coarse to capture the real structure of managerial claims.

6.2 On the contrary

The *Rules* section of this chapter teaches that even complex statements can be decomposed into binary propositions and tested for consistency using the algebra of binary.

And the Apostle says: *"Let your yes be yes and your no be no."*
Therefore, intelligibility is not only possible but required.

6.3 I respond that

Corporate disclosures do exhibit a latent logical structure, even when the prose is hedged, qualified, or strategically ambiguous.

Three insights guide the analysis:

6.3.1 1. Emergence of structure beneath ambiguity

Risk language often appears fuzzy, but its *functional role* in MD&A is to articulate conditional dependencies: - If X occurs, Y may follow.
- Only if A is achieved, B will result.
- Unless C is mitigated, D could arise.

These are precisely the forms your binary conditional $1 + x + xy$ is designed to encode.

6.3.2 2. Ethical intelligibility

Managers owe stakeholders clarity about: - what conditions threaten liquidity,
- what dependencies constrain operations,
- what assumptions underlie forecasts.

Logical consistency is not merely a technical property; it is a moral one.
Incoherent disclosures obscure responsibility.

6.3.3 3. Pedagogical emergence

By beginning with synthetic passages, we learn to: - identify propositions,
- translate English into logical form,
- encode implications in binary,
- test for consistency.

Then, by moving to actual MD&A-style passages, we encounter the complexity of actual managerial communication.

6.4 Application: Synthetic Passages (Option A Content)

Below are three synthetic but realistic NVIDIA-style passages, each followed by a propositional structure suitable for binary analysis.

6.4.1 Passage 1 — Supply-Chain Concentration

> "We depend on a limited number of suppliers for wafer fabrication and advanced packaging. If these suppliers cannot meet our demand, our ability to deliver products will be adversely affected. Our suppliers will meet fabrication demand. Our suppliers may not meet packaging demand."

Let: - F: fabrication demand is met
- P: packaging demand is met
- D: we deliver products on schedule

Formalize: - $\neg F \oplus \neg P \to \neg D$
- F
- $\neg P$

we test whether D must be false.

6.4.2 Passage 2 — Geopolitical Exposure

> "A disruption in Taiwan could materially affect our supply chain. Taiwan will not experience a major disruption. Our supply chain will be materially affected."

Let: - T: Taiwan disruption
- S: supply chain materially affected

Formalize: - $T \to S$
- $\neg T$
- S

we test whether the set is consistent.

6.4.3 Passage 3 — Customer Concentration

> "A small number of customers account for a substantial portion of our revenue. If one of these customers reduces orders, our financial results may be adversely affected. No major customer will reduce orders. Our financial results will be adversely affected."

Let: - C: major customer reduces orders
- R: financial results adversely affected

Formalize: - $C \to R$
- $\neg C$
- R

We test whether the set is consistent.

6.5 Replies to the Objections

6.5.1 Reply to Objection 1

While managerial language is often hedged, its underlying structure is conditional.
Once decomposed, many passages are perfectly consistent—or reveal subtle contradictions.

6.5.2 Reply to Objection 2

Ambiguity does not preclude formalization. The binary algebra handles "only if," "unless," and "may" through standard encodings in a principled framework.

6.5.3 Reply to Objection 3

Even complex supply chains (e.g., NVIDIA $\to$ TSMC $\to$ Foxconn $\to$ Supermicro) can be represented as binary dependencies. Propositional logic does not capture all causal nuance, but it captures the core logical commitments managers make.

6.6 Buildinig a logicks technology

Below we'll treat each passage's propositions as variables in $\{0, 1\}$, with:

- Negation:

$$\neg x = 1 + x$$

- Conjunction:

$$x \wedge y = xy$$

- Disjunction (exclusive "or"):

$$x \oplus y = x + y + xy$$

- Implication $x \to y$: encoded as the truth-value polynomial

$$x \to y \equiv 1 + x + xy$$

  (This equals 1 in all truth-table rows except $x = 1, y = 0$, where it equals 0.)

We clarify our approach with rules learnt from basic arithmetic. Addition is defined as addition of integers, but restricted to modulo 2, with only two possible values $\{0, 1\}$.[1] These values are the terms and the propositions themselves, not to mention the possible outcomes of sequences of propositions we would call an argument. If we remember our basic maths and modulo arithmetic as a clock, the hands on this clock can sit at 0 or 1. If the hand is at 1 and we add 1 to this setting we end up at, yes, 0. A regular OR in many languages would often mean and/or, an inclusive OR, an either one or the other or both. The XOR excludes any possibility in common between two propositions. Propositions are mutually exclusive of one another. Here is a table to help us.

$+(XOR)$	0	1
0	0	1
1	1	0

Addition with boolean values $\{0, 1\}$ is the same as that of the logical exclusive OR operations, XOR. Since each element equals its opposite, subtraction is thus the same operation as addition, so that $0 - 0 = 0$, $0 - 1 = 1$ (remembering the clock with just two settings!), $1 - 0 = 1$, and $1 - 1 = 0$, anyway, all equivalent to the + rules. In words, it is **one or the other, not both**.

The multiplication is multiplication modulo 2 (see the table below), and on boolean variables corresponds to the logical AND operation. Here is the **both** proposition with two terms.

$\times(AND)$	0	1
0	0	0
1	0	1

[1] **?** develops a sweeping history of the progress, decline, and revitalization of algebraic logic. **?** augment Boole's algebraic logic with a Rule of 0 and 1 founded on Horn sentences. This is significant because of the hermeneutical anti-compositional principle of building up an understanding of reasoned arguments from its propositional elements and in turn from terms. We take into account the argument as a whole and the reciprocal interaction of the whole argument with its components.

The NOT logic simply takes a 0 and flips it to 1, and if 1 flips it to zero.

$\neg(NOT)$	$1+x$
0	1
1	0

With these three operations we can evaluate any logical statement. And we must remember that NOT is not the subtraction operation we are very familiar with from non-binary algebra. Some folks will write $\overline{x}$ instead of $\neg$.

Here is a table of several useful expressions based on AND, XOR, and NOT. For any two propositions x and y whose values can only take on a 0 or 1, so that when $x = 0$, we mean x is false, otherwise (that is, $x=1) true.[2]

		AND	XOR	NOT	$y \mid x$	SUB
x	y	xy	$x+y$	$1+x$	$1+x+xy$	$x-y=x+y$
0	0	0	0	1	1	0
0	1	0	1	1	1	1
1	0	0	1	0	0	1
1	1	1	0	0	1	0
x	y	$x \wedge y$	$x \oplus y$	$\neg x$, $\overline{x}$	$x \to y$	
x	y	$-$	$(x \vee y) \wedge \neg(x \wedge y)$	$\neg x \vee y$	$\neg(x \oplus (x \wedge y))$	

The last row describes the algebraic operations in term of propositional logic as applied to Boolean values of $\{0, 1\}$. I can easily get lost in that row. I tend to prefer 9th grade algebraic expressions.

The last column uses the negative to SUBtract y from x. We might be tempted to say that $x - y = x + (1 + y) = 1 + x + y = 1 + (x + y)$ so that the subtraction operator is "NOT the ADD" operator. Nice try! We must always go back to the most literal, most primitive operations, namely the addition and multiplication rules which govern these more derived expressions.

That next to the last column will deserves some important attention since it forms the foundation of conditional (reasonably expected) probability. But first some very helpful derived relationships will make our work going forward a lot smoother.

x	x	$xx = x^2$	$xxx = x^3$	$x + x = 2x$	$x + x + x = 3x$	$x - x = x + x$
0	0	0	0	0	0	0
0	0	0	0	0	0	0
1	1	1	1	0	1	0
1	1	1	1	0	1	0
	$=$	x	x	0	x	0

[2]Boole has the negation of x as 1-x. We follow **?** 1+x, given that all subtractions end up being addition in mod 2 arithmetic.

What happens here is that there are no monomial terms of any higher degree than 1; no quadratic or cubic terms at all. Two propositions literally cancel each other, as the Germans might say an *Aufhebung* event. But three return the single proposition. All of this is the result of the modulo 2 arithmetic to which we constrain ourselves.

We now study $y|x = 1 + x + xy$ in three moves. The first move is to realize that, at least for binary data, Aristotle discovered four logical forms, two in the affirmative, two negative; two universal, two contingent. Medieval logicians called the two affirmative Forms A and I for the first two vowels in *AffIrmo*, Latin for "I affirm," and the two negative forms E and O from the Latin *nEgO*, "I deny." Together they form the **Square of Opposition**. Here S is the subject and P the predicate. Any subject S signifies what it is we are talking about, say, rain. Any predicate P signifies what the subject is about, say, falling to the ground.

Equations and identities do not have a subject or a predicate and themselves might be the subject or predicate. But all propositions do have a subject and a predicate, just like in 3rd grade when we learned to write and speak in complete sentences, that is, in propositions, which contain a subject (usually a noun) and a predicate (usually a verb). We assume that when we apply the forms to concrete examples of propositions, the content of the form, that is, the S and the P, exist. For Thomas Aquinas signs are physical manifestations that allow us to understand something beyond their immediate appearance, like a footprint manifesting someone's presence or smoke manifesting fire. This something with an immediate appearance we will assume without further bother, that it somehow exists. Perhaps we append the particle *any* to S and P to get any rain and any falling (of rain).

We build a table of logical forms in this first move. In the table, *decisions* is the subject S, what we are talking about, and *are rational* is the predicate P, what the subject is about.

Form	*Proposition*	*Sentence*	*Algebra*
A	All S is P	"All decisions are rational."	$a(1+b)=0$
E	No S is P	"Not all decisions are rational."	$ab=0$
I	Some S is P	"Some decisions are rational."	$ab \neq 0$
O	Some S is not P	"Some decisions are not rational."	$a(1+b) \neq 0$

The second move is to parse the Form A proposition "All decisions (a) are rational (b)." Logically a and not b is false (0) in Form A, that is, algebraicly, $a(1+b) = 0$. "Decisions" and "not-rational" is false, that is, inconsistent according to Form A. The obverse must be true, that "No-decisions are not-rational," as if we might be able to interpret this double negative. One more swipe at interpretation is called for. To say that "all of anything is something else" is, effectively to identify "anything" with "something else." If this statement is true, as we are positing here in the form, then it cannot be true that both "anything $= a$" and "not-something else $= (1 + b)$" can coexist. Thus the logical Form A seems to admit a very basic principle we would hold alongside all other principles (Aristotle called this *metaphysics*), namely, the *Principle of Non-Contradiction.* Yes, we cannot, so far it seems, to be in two places at the same time.

In a third move, we use the NOT operation on Form A. We recall we are using "+" as exclusive OR, XOR with this algebraic rearrangement.

statement	*reason*	
1	$a(1 + b) = 0$	Form A definition
2	$1 + a(1 + b) = 1$	Symmetric property
3	$1 + a + ab = 1$	Distribution of multiplication over addition property
4	$a \to b = b \mid a$	a gives some information to b

Simply negating (using 1+) the Form A definition reveals a conditional relationship between a and b. Saying that a AND $not - b$ is possible when we negate the Form A requirement, means a does share something with b. So-called "implication" means that a shares a's information with b. This is the primary meaning we ascribe to "a conditions b", $b \mid a$. We now have all the ingredients to use this framework to discover a principled path to the expectations approach to probability, that is, counting the ways in which data are consistent with hypotheses. But before we enter that particular room, we need to argue a bit more, that is, discover valid and invalid logical arguments.

Each sentence in the passage becomes a polynomial equation of the form "expression = 1" (true) or "variable = 0/1".

6.6.0.1 Passage 1: Supply-chain concentration

> "We depend on a limited number of suppliers for wafer fabrication and advanced packaging. If these suppliers cannot meet our demand, our ability to deliver products will be adversely affected. Our suppliers will meet fabrication demand. Our suppliers may not meet packaging demand."

Let:

- F: fabrication demand is met.
- P: packaging demand is met.
- D: we deliver products on schedule.

We formalized:

1. $\neg F \oplus \neg P \to \neg D$

2. F

3. $\neg P$

Step 1: Encode each statement.

- (2) F is true:

$$F = 1$$

- (3) $\neg P$ is true:

$$\neg P = 1 + P = 1 \quad \Rightarrow \quad P = 0$$

- (1) $\neg F \oplus \neg P \to \neg D$:
  First compute $\neg F$ and $\neg P$:

$$\neg F = 1 + F, \quad \neg P = 1 + P$$

Disjunction:

$$\neg F \oplus \neg P = (1 + F) \oplus (1 + P)$$
$$= (1 + F) + (1 + P) + (1 + F)(1 + P)$$

Simplify:

$$(1 + F) + (1 + P) = F + P$$
$$(1 + F)(1 + P) = 1 + F + P + FP$$

So

$$\neg F \oplus \neg P = F + P + 1 + F + P + FP = 1 + FP$$

So the antecedent is $1 + FP$. The consequent is $\neg D = 1 + D$.
The implication polynomial is:

$$(1 + FP) \to (1 + D) = 1 + (1 + FP) + (1 + FP)(1 + D)$$

Compute:

$$1 + (1 + FP) = FP$$

$$(1 + FP)(1 + D) = 1 + D + FP + FPD$$

So:

$$1+(1+FP)+(1+FP)(1+D) = FP+1+D+FP+FPD = 1+D+FPD$$

Therefore (1) says:

$$1 + D + FPD = 1 \quad \Rightarrow \quad D + FPD = 0$$

Step 2: Substitute known values.

We already have $F = 1, P = 0$. Then:

$$FPD = (1)(0)D = 0$$

So the equation $D + FPD = 0$ becomes:

$$D + 0 = 0 \quad \Rightarrow \quad D = 0$$

Conclusion

A satisfying assignment is:

$$F = 1, \quad P = 0, \quad D = 0$$

All three statements can be true simultaneously. The set is logically consistent, and the algebra forces the natural managerial reading: if packaging capacity fails, on-time delivery fails.

6.6.0.2 Passage 2: Geopolitical exposure

> "A disruption in Taiwan could materially affect our supply chain. Taiwan will not experience a major disruption. Our supply chain will be materially affected."

Let:

- T: Taiwan disruption.
- S: supply chain materially affected.

Formalize:

1. $T \to S$
2. $\neg T$
3. S

Step 1: Encode each statement.

- (1) $T \to S$:

$$T \to S = 1 + T + TS = 1$$

  So:

$$1 + T + TS = 1 \quad \Rightarrow \quad T + TS = 0$$

- (2) $\neg T$:

$$1 + T = 1 \quad \Rightarrow \quad T = 0$$

- (3) S:

$$S = 1$$

Step 2: Substitute into (1).

With $T = 0$:

$$T + TS = 0 + 0 \cdot S = 0$$

So equation (1) becomes $0 = 0$, automatically satisfied for any S. We also impose $S = 1$ from (3).

Conclusion

A satisfying assignment is:

$$T = 0, \quad S = 1$$

All three statements are jointly consistent. The algebra highlights a conceptual point for we: supply-chain stress may have causes other than the Taiwan scenario flagged in (1).

6.6.0.3 Passage 3: Customer concentration

> “A small number of customers account for a substantial portion of our revenue. If one of these customers reduces orders, our financial results may be adversely affected. No major customer will reduce orders. Our financial results will be adversely affected.”

Let:

- C: a major customer reduces orders.
- R: financial results adversely affected.

Formalize:

1. $C \to R$
2. $\neg C$
3. R

Step 1: Encode each statement.

- (1) $C \to R$:

$$C \to R = 1 + C + CR = 1$$

  So:

$$1 + C + CR = 1 \quad \Rightarrow \quad C + CR = 0$$

- (2) $\neg C$:

$$1 + C = 1 \quad \Rightarrow \quad C = 0$$

- (3) R:

$$R = 1$$

Step 2: Substitute into (1).

With $C = 0$:

$$C + CR = 0 + 0 \cdot R = 0$$