The Latticework:

Statistics & Probabilities

What I noted since the really big ideas carry 95% of the freight, it wasn't at all hard for me to pick up all the big ideas from all the big disciplines and make them a standard part of my mental routines. Once you have the ideas, of course, they are no good if you don't practice – if you don't practice you lose it. So, I went through life constantly practicing this model of the multidisciplinary approach. Well, I can't tell you what that's done for me. It's made life more fun, it's made me more constructive, it's made me more helpful to others, it's made me enormously rich, you name it, that attitude really helps…

…It doesn't help you just to know them enough just so you can give them back on an exam and get an A. You have to learn these things in such a way that they're in a mental latticework in your head and you automatically use them for the rest of your life.

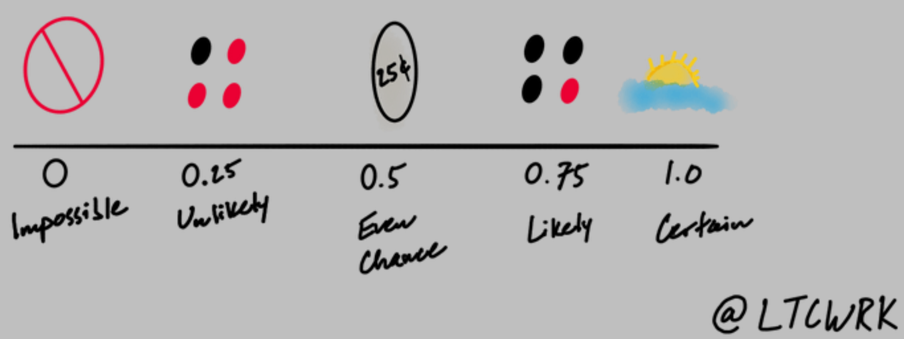
– Charlie Munger,

2007 USC Gould School of Law Commencement Speech

Statistics & Probabilities

Statistics and probabilities are an offshoot of mathematics. These ideas are incredibly important to understand to effectively operate in the world and make better decisions. Thinking in this manner is quite difficult and often counterintuitive, but a basic understanding of statistics and probabilities, coupled with concepts we'll tackle later on like compound interest, the time value of money, and opportunity costs, provide the foundation for making effective decisions. This process will help us see reality for what it is, rather than how we wish it to be.

Many of these ideas aren't complex, but a number of biases make it hard for us to adhere to statistical and probabilistic thinking. People can be overconfident, misunderstand the role of luck and skill, overweight losses compared to gains, overvalue the things they currently own, and much more. However, by keeping some of these big ideas in mind, we can mitigate these biases and make more effective decisions that align with the probabilities we face.



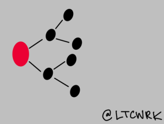
These ideas in isolation, if not understood to the point that they're applicable in your life, are not fulfilling their purpose. The point is to enhance the way you think and live in the real world and while it can be intellectually stimulating when these concepts sound beautiful on paper, it misses the point. We are seeking ideas that concretely improve our lives and the process of becoming a more thorough statistical and probabilistic thinker is an uncomfortable, yet lucrative journey that accomplishes this.

The most important questions of life are indeed, for the most part, really only problems of probability.

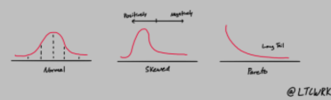
– Pierre Simon Laplace

The Big Ideas of Statistics & Probabilities:

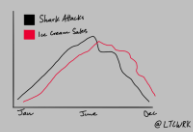
1. Bayes' Theorem



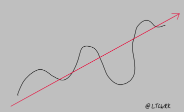
2. Probability Distributions



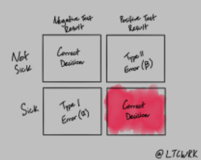
3. Correlation vs. Causation



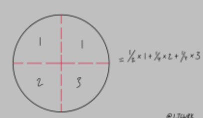
4. Regression to the Mean



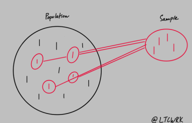
5. Hypothesis Testing



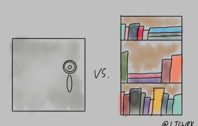
6. Expected Value



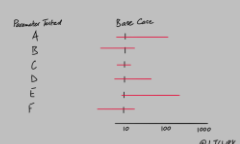
7. Sample Size



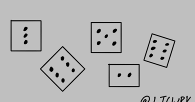
8. Permutations & Combinations



9. Sensitivity Analysis



10. Randomness

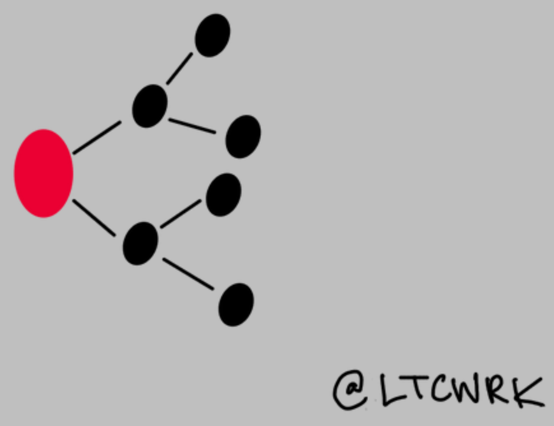


Bayes' Theorem

The Bayesian method is an approach to thinking where you first consider all known information and then actively update your forecast as newer information becomes available.

This method is especially productive given the fundamentally probabilistic and non-deterministic world we live in. We must use prior odds and new information in combination to arrive at our best decisions.

This is not necessarily our intuitive decision-making engine and this process helps us manage some natural human biases by encouraging us to recalculate the odds of success after a subsequent event has occurred. Past performance tells you something (but not everything!) about future results. We must be willing to make the effort to understand the probabilities we face and act accordingly.



The theorem itself isn't so hard: the probability that a proposition is true, given some new data, is proportional to the probability it was true before that data came in, times the likelihood of the new data if the proposition were true.

– *Sean Carroll, Bayes' Theorem*

Probability Distributions

A probability distribution provides the chance of occurrence of different possible outcomes in a given context. It is concerned with the analysis and approximate estimation of any specific or random phenomena occurring.

Because we live in an unknowable, complex world that is dominated by probabilistic outcomes, this is important to understand. Although we cannot predict the future with great certainty, we are wise to ascribe odds to more and less probable events. We do this every day unconsciously – from driving our cars to choosing where to travel – but the trick is to make it a conscious habit and to track your guesses and their outcomes (this is where Journaling could be game-changing).

Understanding the probability of success, or failure, in any given situation is crucial to making good decisions and improving our decision-making process over time. Thinking probabilistically is about managing risk and understanding potential rewards.

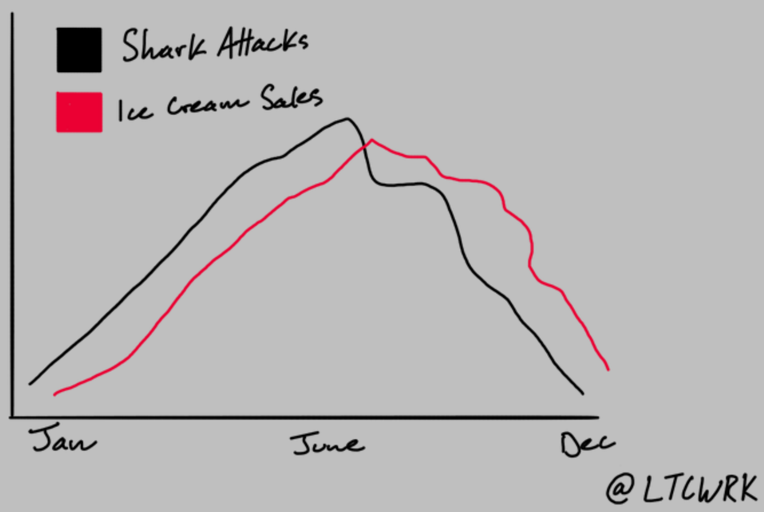
How can investors deal with the limitations on their ability to know the future? The answer lies in the fact that not being able to know the future doesn't mean we can't deal with it. It's one thing to know what's going to happen and something every different to have a feeling for the range of possible outcomes and the likelihood of each one happening. Saying we can't do the former doesn't mean we can't do the latter. The information we're able to estimate – the list of events that might happen and how likely each one is – can be used to construct a probability distribution. Key point number one in this memo is that the future should be viewed not as a fixed outcome that's destined to happen and capable of being predicted, but as a range of possibilities and, hopefully on the basis of insight into their respective likelihoods, as a probability distribution.…Bruce Newberg says, "There's a big difference between probability and outcome." Unlikely things happen – and likely things fail to happen – all the time. Probabilities are likelihoods and very far from certainties.

– Howard Marks, Dare to Be Great II

Correlation vs. Causation

When two things seem linked, it is easy and tempting to assume that one causes another. It may be true, but make it a conscious decision rather than simply assuming it is so.

As we discussed in mental models, our models about the world will never perfectly represent reality, but, if at least grounded in reality, they can be good enough. However, we get into trouble when we fool ourselves into thinking that the world is a linear or predictable series of events. Like the tip of the iceberg above the water, events are the most visible aspect of a larger complex but not always the most important. We are less likely to be surprised if we can see how events accumulate into dynamic patterns of behavior. This is difficult work, but necessary (Systems Thinking)



Cars with flames painted on the hood might get more speeding tickets. Are the flames making the car go fast? No. Certain things just go together. And when they do, they are correlated. It is the darling of all human errors to assume, without proper testing, that one is the cause of the other.

– Barbara Kingsolver, Flight Behavior

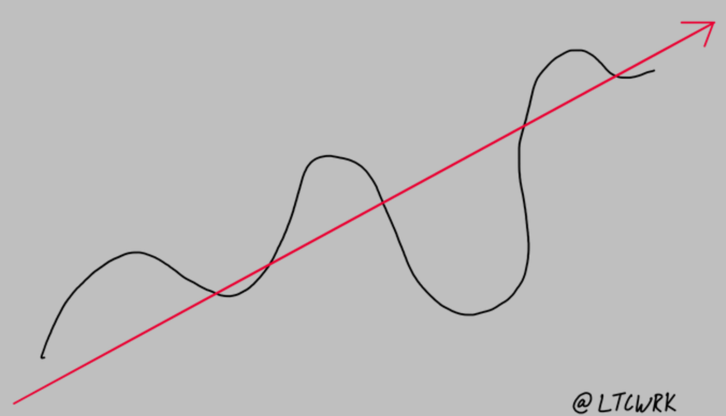
Regression to the Mean

Regression to the Mean is the phenomenon that if a variable is extreme on its first measurement, it will tend to be closer to the average in future measurement.

For example, did the remedy cure the sick patient or was it simply the body regressing to the mean? Could very well be the remedy, but we should at least consider regression to the mean.

This fools people all the time. The addicted gambler who thinks things are bound to turn his way. They might, eventually. But, as Warren Buffett said, the market can stay irrational longer than you can stay solvent. We might not know when or how quickly things will regress back to the mean, simply that they eventually should.

This is the danger with averages. They often mask the underlying distribution. Howard Marks is a big fan of the line: "You must never forget the six-foot-tall man who drowned crossing the river that was five feet deep on average."



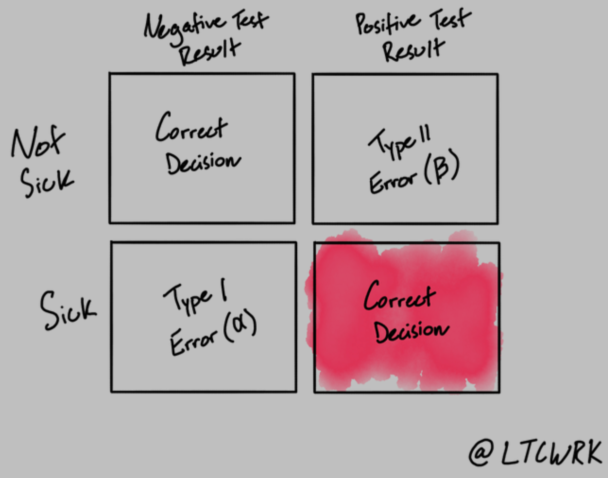
Regression to the mean is the most powerful law in financial physics: Periods of above average performance are inevitably followed by below-average returns, and bad times inevitably set the stage for surprisingly good performance.

– Jason Zweig

Hypothesis Testing

Hypothesis testing uses statistics to determine the probability that a given hypothesis is true. The usual process of hypothesis testing consists of four steps.

1. Formulate the null hypothesis H0 (commonly, that the observations are the result of pure chance) and the alternative hypothesis Ha (commonly, that the observations show a real effect combined with a component of chance variation).
2. Identify a test statistic that can be used to assess the truth of the null hypothesis.
3. Compute the P-value, which is the probability that a test statistic is at least as significant as the one observed would be obtained assuming that the null hypothesis were true. The smaller the P-value, the stronger the evidence against the null hypothesis.
4. Compare the P-value to an acceptable significance value alpha (sometimes called an alpha value). If p <= alpha, that the observed effect is statistically significant, the null hypothesis is ruled out, and the alternative hypothesis is valid.



The idea is if you put ill-determined numbers and equations (garbage) in then you can only get ill-determined results (garbage) out. By implication the converse is tacitly assumed, if what goes in is accurate then what comes out must be accurate. I shall show both of these assumptions can be false.

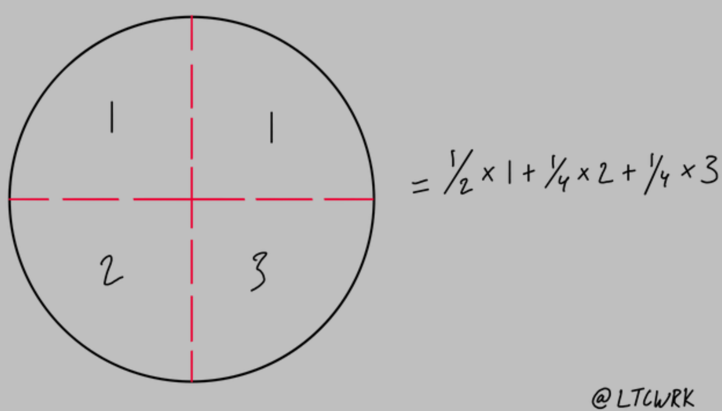
– Richard Hamming, Simulation

Expected Value

Expected value is the long run average value of an experiment or series of events.

The value is determined by multiplying each possible outcome by its probability of occurring and summing the results.

The expected value (and the variance) is important to know when making a decision. It can inform how confident you should be in your assessment and how big of a bet to place.



Often the best way to choose between alternative courses of action is by figuring out which has the highest "expected value": the total value arrived at by multiplying each possible outcome by its probability of occurring and summing the results. As I learned from my first textbook at Wharton fifty years ago (Decisions Under Uncertainty by C. Jackson Grayson, Jr.), if one act has a higher expected value than another and "…if the decision maker is willing to regard the consequences of each act-event in purely monetary terms, then this would be the logical act to choose. Keeping in mind, however, that only one event and its consequences will occur (not the weighted average consequence)," agents may not be able to choose on the basis of expected value or the weighted average of all possible outcomes. If a given action has potential bad consequences that are absolutely unacceptable, the expected value of all its consequences – both good and bad – can be irrelevant.

– Howard Marks, Dare to Be Great II

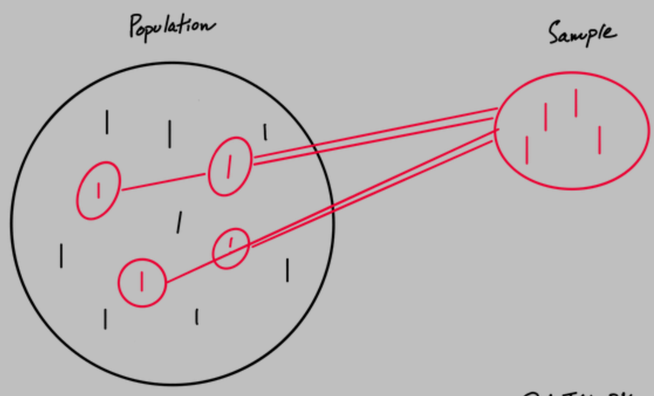
Sample Size

Sample Size is the number of observations to include in a statistical sample. This is an important consideration as too few or irrelevant samples would lead to incorrect conclusions. That is why we said earlier that a large, relevant sample size is a statistician's best friend. The bigger the sample size, the more reliable the conclusions (remember The Three Buckets?)

One of the fundamental underlying assumptions of probability theory is that as more instances of an event occur, the actual results will converge on the expected value and trend closer as more trials are performed. This is known as the law of large numbers.

For example, if I know that the average man is 5 feet 10 inches tall, I am far more likely to get an average of 5' 10" by selecting 500 men at random than 5 men at random.

The opposite of this model is the law of small numbers, which states that small samples can and should be looked at with great skepticism as much larger fluctuations, randomness, and noise should be expected. Anecdotes and broad generalizations are common examples of the law of small numbers coming into play and influencing our perspective.



A few billion years is vastly more proof than a thousand days of survival, and the oldest system around is clearly Mother Nature.

– Nassim Taleb, The Black Swan

Permutations & Combinations

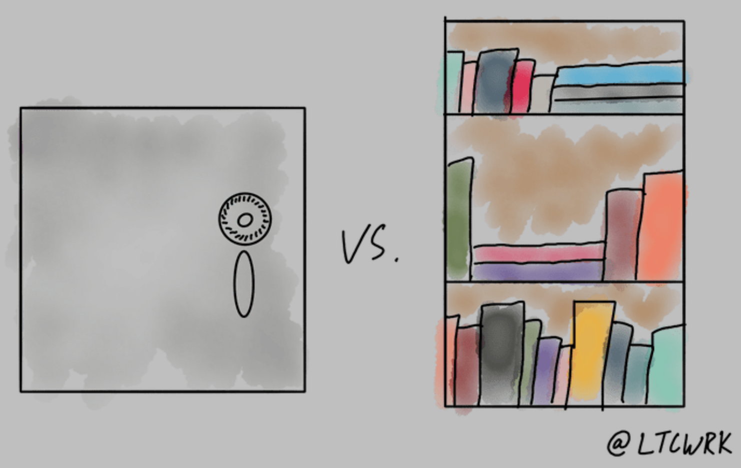
In everyday English, we use the word "combination" without thinking if the order of things is important. However, in mathematics, there are distinct words for groupings in which the order does and does not matter.

When the order does matter it is a Permutation.

When the order doesn't matter, it is a Combination.

In other words, a permutation is an ordered combination.

The mathematics of permutations and combinations leads us to understand the practical probabilities of the world around us, how things can be ordered, how we should think about things, and help us improve our intuition around probabilities and make good decisions based on expected value.



No one in the world is going to beat you at being you. You're never going to be as good at being me as I am. I'm never going to be as good at being you as you are. Certainly, listen and absorb, but don't try to emulate. It's a fool's errand. Instead, each person is uniquely qualified at something. They have some specific knowledge, capability, and desire nobody else in the world does, purely from the combinatorics of human DNA and development. The combinatorics of human DNA and experience are staggering. You will never meet any two humans who are substitutable for each other.

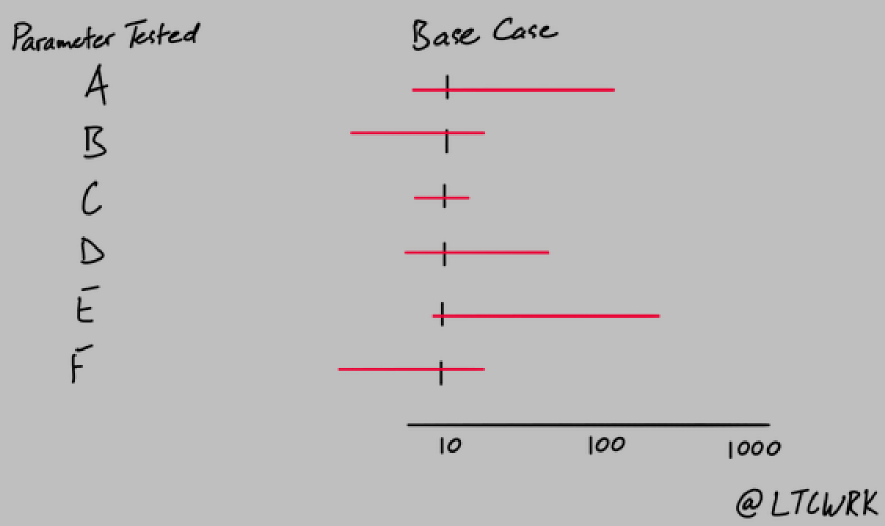
– Naval Ravikant, The Almanack of Naval Ravikant

Sensitivity Analysis

Sensitivity Analysis is the process of analyzing how sensitive a model is to its various inputs, helping determine the impact of these inputs.

This process helps you quickly uncover the key inputs which impact the results, showing you where you should focus your time and efforts.

For example, to better understand the reliability of systems such as bridges, tunnels, railways, and more, Monte Carlo simulations can be run an infinite number of times to better understand how the failure of the various components would impact the reliability of the overall system. The simulations tests how sensitive each component is under a wide variety of circumstances.



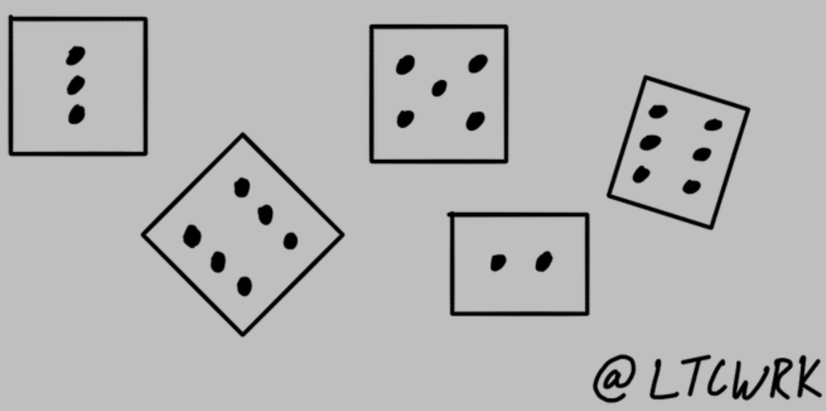
Limit risk with: Deep analysis, bargain purchase, sensitivity analysis.

– Seth Klarman

Randomness

Randomness is a measure of uncertainty of an outcome and, although the human brain has trouble comprehending it, much of the world is composed of random, non-sequential, non-ordered events.

As Taleb says, we are "fooled by randomness" when we attribute causality to things that are actually outside of our control (Correlation vs. Causation). If we don't course-correct for this fooled by randomness effect – our faulty sense of pattern-seeking and our love of narrative – we will tend to see things as being more predictable than they are and act accordingly, making ourselves vulnerable to randomness and black swans. This should humble us and make clear why considering redundancy and a margin of safety into thinking and decision-making is so valuable.



A large group of physicists, certainly, created a healthy flow of ideas. But Kelly believed the most valuable ideas arose when the large group of physicists bumped against other departments and disciplines, too. "It's the interaction between fundamental science and applied science, and the interface between many disciplines, that creates new ideas," explains Herwig Kogelnik, the laser scientist. This may indeed have been Kelly's greatest insight.

– Jon Gertner, The Idea Factory

As the African proverb goes, “If you want to go fast, go alone. If you want to go far, go together.”

Thank you for choosing to go together.