

The following content is provided under a Creative Commons license. Your support will help MIT OpenCourseWare continue to offer high quality educational resources for free. To make a donation or view additional materials from hundreds of MIT courses, visit MIT OpenCourseWare at [ocw.mit.edu](https://ocw.mit.edu).

**KENNETH  
ABBOTT:**

As I said, my name is Ken Abbott. I'm the operating officer for Firm Risk Management at Morgan Stanley, which means I'm the everything else guy. I'm like the normal stuff with a bar over it. The complement of normal-- I get all the odd stuff. I consider myself the Harvey Keitel character. You know, the fixer? And so I get a lot of interesting stuff to do. I've covered commodities, I've covered fixed income, I've covered equities, I've covered credit derivatives, I've covered mortgages.

Now I'm also the Chief Risk Officer for the buy side of Morgan Stanley. The investment management business and the private equity holdings that we have. And I look after lot of that stuff and I sit on probably 40 different committees because it's become very, very, very bureaucratic. But that's the way it goes.

What I want to talk about today is some of the core approaches we use to measure a risk in a market risk setting. This is part of a larger course I teach at a couple places. I'm a triple alum at NYU-- no I'm a double alum and now I'm on their faculty [INAUDIBLE]. I have a masters in economics from their arts and sciences program. I have a masters in statistics from Stern when Stern used to have a stat program. And now I teach at [INAUDIBLE]. I also teach at Claremont and I teach at [INAUDIBLE], part of that program. So I've been through this material many times.

So what I want to do is lay the foundation for this notion that we call risk, this idea of var. [INAUDIBLE] put this back on. Got it. I'll make it work. I'll talk about it from a mathematical standpoint and from a statistical standpoint, but also give you some of the intuition behind what it is that we're trying to do when we measure this thing.

First, a couple words about risk management. What is the risk do? 25 years ago,

maybe three firms had risk management groups. I was part of the first risk management group at Bankers Trust in 1986. No one else had a risk management group as far as I know. Market risk management really came to be in the late '80s. Credit risk management had obviously been around in large financial institutions the whole time.

So our job is to make sure that management knows what's on the books. So step one is, what is the risk profile of the firm? How do I make sure that management is informed about this? So it requires two things. One, I have to know what the risk profile is because I have to know it in order to be able to communicate it. But the second thing, equally important, particularly important for you guys and girls, is that you need to be able to express relatively complex concepts in simple words and pretty pictures. All right?

Chances are if you go to work for big firm, your boss won't be a quant. My boss happens to have a degree from Carnegie Mellon. He can count to 11 with his shoes on. His boss is a lawyer. His boss is the chairman. Commonly, the most senior people are very, very intelligent, very, very articulate, very, very learned. But not necessarily quants. Many of them have had a year or two of calculus, maybe even linear algebra.

You can't show them-- look, when you and I chat and we talk about regression analysis, I could say  $x^T x^{-1} x^T y$ . And those of you that have taken a refresher course think, ah, that's beta hat. And we can just stop it there. I can just put this form up there and you may recognize it. I would have to spend 45 minutes explaining this to people on the top floor because this is not what they're studying.

So we can talk the code amongst ourselves, but when we go outside our little group-- getting bigger-- we have to make sure that we can express ourselves clearly. That's done in clear, effective prose, and in graphs. And I'll show you some of that stuff as we go on. So step one, make sure management knows what the risk profile is. Step two, protect the firm against unacceptably large concentrations.

This is the subjective part. I can know the risk, but how big is big? How much is too much? How much is too concentrated? If I have \$1 million of sensitivity per basis point, that's a 1/100th of 1% move in a rate. Is that big? Is that small? How do I know how much? How much of a particular stock issue should I own? How much of a bond issue? How much futures open interest? How big a limit should I have on this type of risk?

That's where intuition and experience come into play. So that's the second part of our job is to protect against unacceptably large losses. So the third, no surprises, you can liken the trading business-- it's taking calculated risks. Sometimes you're going to lose. Many times you're going to lose. In fact, if you win 51% of the time, life is pretty good. So what you want to do is make sure you have the right information so you can estimate, if things get bad, how bad will they get? And to use that, we leverage a lot of relatively simple notions that we see in statistics.

And so I should use a coloring mask here, not a spotlight. We do a couple things. Just like the way when they talk about the press in your course about journalism, we can shine a light anywhere we want, and we do all the time. You know what? I'm going to think about this particular kind of risk. I'm going to point out that this is really important. You need to pay attention to it. And then I could shade it. I can make it blue, I can make a red, I can make it green. I'd say this is good, this is bad, this is too big, this is too small, this is perfectly fine. So that's just a little bit of quick background on what we do.

So I'm going to go through as much of this as I can. I'm going to fly through the first part and I want to hit these because these are the ways that we actually estimate risk. Variance, covariance [? as ?] a quadratic form. Monte Carlo simulation, the way I'll show you is based on a quadratic form. And historical simulation is Monte Carlo simulation without the Monte Carlo part. It's using historical data. And I'll go through that fairly quickly. Questions, comments? No? Excellent.

Stop me-- look, if any one of you doesn't understand something I say, probably many of you don't understand it. I don't know you guys, so I don't know what you

know and what you don't know. So if there's a term that comes up, you're not sure, just say, Ken, I don't have a PhD. I work for a living. I make fun of academics. I know you work for a living too. All right. There's a guy I tease at Claremont [INAUDIBLE] in this class, I say, who is this pointy headed academic [INAUDIBLE]. Only kidding.

All right, so I'm going to talk about one asset value at risk. First I'm going to introduce the notion of value at risk. I'm going to talk about one asset. I'm going to talk about price based instruments. We're going to go into yield space, so we'll talk about the conversions we have to do there. One thing I'll do after this class is over, since I know I'm going to fly through some of the material-- and since this is MIT, I'm sure you're used to just flying through material. And there's a lot of this, the proof of which is left to the reader as an exercise. I'm sure you get a fair amount of that.

I will give you papers. If you have questions, my email is on the first page. I welcome your questions. I tell my students that every year. I'm OK with you sending me an email asking me for a reference, a citation, something. I'm perfectly fine with that. Don't worry, oh, he's too busy. I'm fine. If you've got a question, something is not clear, I've got access to thousands of papers. And I've screened them. I've read thousands of papers, I say this is a good one, that's a waste of time.

But I can give you background material on regulation, on bond pricing, on derivative algorithms. Let me know. I'm happy to provide that at any point in time. You get that free with your tuition. A couple of key metrics. I don't want to spend too much time on this. Interest rate exposure, how sensitive am I to changes in interest rates, equity exposure, commodity exposure, credit spread exposure. We'll talk about linearity, we won't talk too much about regularity of cash flow. We won't really get into that here. And we need to know correlation across different asset classes. And I'll show you what that means.

At the heart of this notion of value at risk is this idea of a statistical order statistic. Who here has heard of order statistics? All right, I'm going to give you 30 seconds. The best simple description of an order statistic.

**PROFESSOR:** The maximum or the minimum of a set of observations.

**KENNETH ABBOTT:** All right? When we talk about value at risk, I want to know the worst 1% of the outcomes. And what's cool about order statistics is they're well established in the literature. Pretty well understood. And so people are familiar with it. Once we put our toe into the academic water and we start talking about this notion, there's a vast body of literature that says this is how this thing is. This is how it pays. This is what the distribution looks like. And so we can estimate these things.

And so what we're looking at in value at risk, if my distribution of returns, how much I make. In particular, if I look historically, I have a position. How much would this position have earned me over the last  $n$  days,  $n$  weeks,  $n$  months. If I look at a frequency distribution of that, I'm likely-- don't have to-- I'm likely to get something that's symmetric. I'm likely to get something that's unimodal. It may or may not have fat tails. We'll talk about that a little later.

If my return distribution were beautifully symmetric and beautifully normal and independent, then the risk-- I could measure this 1% order statistic. What's the 1% likely worst case outcome tomorrow? I might do that by integrating the normal function from negative infinity-- for all intents and purposes five or six standard deviations. Anyway, from negative infinity to negative 2.33 standard deviations. Why? Because the area under the curve, that's 0.01.

Now this is a one sided confidence interval as opposed to a two sided confidence interval. And this is one of these things that as an undergrad you learn two sided, and then the first time someone shows you one sided you're like, wait a minute. What is this? Then you say, oh, I get it. You're just looking at the area. I could build a gazillion two sided confidence intervals. One sided, it's got to stop at one place.

All right so this set of outcomes-- and this is standardized-- this is in standard deviation space-- negative infinity to 2.33. If I want 95%, or 5% likely loss, so I could say, tomorrow there's a 5% chance my loss is going to be  $x$  or greater, I would go to 1.645 standard deviations. Because the integral from negative infinity to 1.645 standard deviations is about 0.05. It's not just a good idea, it's the law. Does that

make sense?

And again, I'm going to say assuming the normal. That's like the old economist joke, assume a can opener when he's on a desert island. You guys don't know that one. I got lots of economics jokes. I'll tell them later on maybe-- or after class. If I'm assuming normal distribution, and that's what I'm going to do, what I want to do is I'm going to set this thing up in a normal distribution framework. Now doing this approach and assuming normal distributions, I liken it to using Latin. Nobody really uses it anymore but everything we do is based upon it.

So that's our starting point. And it's really easy to teach it this way and then we relax the assumptions like so many things in life. I teach you the strict case then we relax the assumptions to get to the way it's done now. So this makes sense? All right. So let's get there. This is way oversimplified-- but let's say I have something like this. Who has taken intermediate statistics? We have the notion of stationarity that we talk about all the time. The mean and variance constant is one simplistic way of thinking about this. Do you have a better way for me to put that to them? Because you know what their background would be.

**PROFESSOR:** No.

**KENNETH ABBOTT:** All right. Just, mean and variance are constant. When I look at the time series itself, the time series mean and the time series variance are not constant. And there also could be other time series stuff going on. There could be seasonality, there could be autocorrelation. This looks something like a random walk but it's not stationary. It's hard for me to draw inference by looking at that alone. So we want to try to predict what's going to happen in the future, it's kind of hard.

And the game, here, that we're playing, is we want to know how much money do I need to hold to support that position? Now, who here has taken an accounting course? All right, word to the wise-- there's two things I tell students and quant finance programs. First of all, I know you have to take a time series course-- I'm sure-- this is MIT. If you don't get a time series course, get your money back because you've got to take time series. Accounting is important. Accounting is

important because so much of what we do, the way we think about things is predicated on the dollars. And you need to know how the dollars are recorded.

Quick aside. Balance sheet. I'll give you a 30 second accounting lecture. Assets, what we own. Everything we own-- we have stuff, it's assets. We came to that stuff one of two ways. We either pay for it out of our pocket, or we borrowed money. There's no third way. So everything we own, we either paid for out of our pocket or borrowed money. The amount we paid for out of our pocket is the equity. The ratio of this to this is called leverage among other things.

All right? If I'm this company. I have this much stuff and I bought it with this much debt, and this much equity. Again, that's a gross oversimplification. When this gets down to zero, it's game over. Belly up. All right? Does that make sense? Now you've taken a semester of accounting. No, only kidding. But it's actually important to have a grip on how that works. Because what we need to make sure of is that if we're going to take this position and hold it, we need to make sure that with some level of certainty-- every time we lose money this gets reduced. When this goes down to zero, I go bankrupt.

So that's what we're trying to do. We need to protect this, and we do it by knowing how much of this could move against us. Everybody with me? Anybody not with me? It's OK to have questions, it really is. Excellent. All right, so if I do a frequency distribution of this time series, I just say, show me the frequency with which this thing shows. I get this thing, it's kind of trimodal. It's all over the place. It doesn't tell me anything. If I look at the levels-- the frequency distribution, the relative frequency distribution of the levels themselves, I don't get a whole lot of intuition.

If I go into return space, which is either looking at the log differences from day to day, or the percentage changes from day to day, or perhaps the absolute changes from day to day-- it varies from market to market. Oh, look, now we're in familiar territory. So what I'm doing here-- and this is why I started out with a normal distribution because this thing is unimodal. It's more or less symmetric. Right? Now is it a perfect measure? No, because it's probably got fat tails. So it's a little bit like

looking for the glasses you lost up on 67th Street down on 59th street because there's more light there. But it's a starting point.

So what I'm saying to you is once I difference it-- no, I won't talk about [INAUDIBLE]. Once I difference the timeshares, once I take the timeshares and look at the percentage changes, and I look at the frequency distribution of those changes, I get this which is far more minimal. And I can draw inference from that. I can say, ah, now if this thing is normal, then I know that  $x\%$  of my observations will take place over here. Now I can start drawing inferences.

And a thing to keep in mind here, one thing we do constantly in statistics is we do parameter estimates. And remember, every time you estimate something you estimate it with error. I think that maybe the single most important thing I learned when I got my statistics degree. Everything you estimate you estimate with error. People do means, they say, oh, it's  $x$ . No, that's the average and that's an unbiased estimator, but guess what, there's a huge amount of noise. And there's a certain probability that you're wrong by  $x\%$ .

So every time we come up with a number, when somebody tells me the risk is 10, that means it's probably not 10,000, it's probably not zero. Just keep that in mind. Just sort of throw that in on the side for nothing. All right, so when I take the returns of this same time series, I get something that's unimodal, symmetric, may or may not have fat tails. That has important implications for whether or not my normal distribution underestimates the amount of risk I'm taking.

Everybody with me on that more or less? Questions? Now would be the time. Good enough? He's lived this. All right. So once I have my time series of returns, which I just plotted there, I can gauge their dispersion with this measure called variance. And you guys probably know this. Variance the expected value of  $x_i$  minus  $\bar{x}$ -- I love these thick chalks-- squared. And it's the sum of  $x_i$  minus  $\bar{x}$  squared over  $n$  minus 1.

It's a measure of dispersion. Variance has [INAUDIBLE]. Now, I should say that this is  $\hat{\sigma}^2$ . Right? Estimate-- parameter estimate. Parameter. Parameter



estimate. This is measured with error. Anybody here know what the distribution of this is? Anyone? \$5. Close.  $m$  chi squared. Worth \$2. Talk to me after class. It's a chi squared distribution. What does that mean? That means that we know it can't be 0 or less than 0. If you figure out a way to get variances less than zero, let's talk. And it's got a long right tail, but that's because this is squared. [INAUDIBLE] one point can move it up.

Anyway, once I have my returns, I have a measure of the dispersion of these returns called variance. I take the square root of the variance, which is the standard deviation, or the volatility. When I'm doing it with a data set, I usually refer to it as the standard deviation. When I'm referring to the standard deviation of the distribution, I usually call it the standard error. Is that a law or is that just common parlance?

**PROFESSOR:** Both The standard error is typically for something that's random, like an estimate. Whereas the standard deviation is more like for sample--

**KENNETH ABBOTT:** Empirical. See, it's important because when you first learn this, they don't tell you that. And they flip them back and forth. And then when you take the intermediate courses, they say, no, don't you standard deviation when you mean standard error. And you'll get points off on your exam for that, right? All right, so, the standard deviation is the square root of the variance, also called the volatility. In a normal distribution, 1% of the observations is outside of 2.33 standard deviations. For 95%, it's out past 1.64, 1.645 standard deviations.

Now you're saying, wait a minute, where did my 1.96 go that I learned as an undergrad. Two sided. So if I go from the mean to 1.96 standard deviations on either side, that encompasses 95% of the total area of the integral from negative infinity to positive infinity. Everybody with me on that? Does that make sense? The two sided versus one sided. That's confused me. When I was your age, it confused me a lot. But I got there.

All right so this is how we do it. Excel functions are var and-- you don't need to know that. All right, so in this case, I estimating the variance of this particular time series. I

took the standard deviation by taking the square root of the variance. It's in percentages. When you do this, I tell you, it's like physics, your units will screw you up every time. What am I measuring? What are my units? I still make units mistakes. I want you to know that. And I'm in this business 30 years. I still make units mistakes. Just like physics.

I'm in percentage change space, so I want to talk in terms of percentage changes. The standard deviation is 1.8% of that time series I showed you. So 2.33 standard deviations times the standard deviation is about 4.2%. What that says, given this data set-- one time series-- I'm saying, I expect to lose, on any given day, if I have that position, 99% of the time I'm going to lose 4.2% of it or less. Very important. Think about that. Is that clear? That's how I get there.

I'm making a statement about the probability of loss. I'm saying there's a 1% probability, for that particular time series-- which is-- all right? If this is my historical data set and it's my only historical data set, and I own this, tomorrow I may be 4.2% lighter than I was today because the market could move against me. And I'm 99% sure, if the future's like the past, that my loss tomorrow is going to be 4.2% or less. That's [? var. ?] Simplest case, assuming normal distribution, single asset, not fixed income. Yes, no? Questions, comments?

**AUDIENCE:** Yes, [INAUDIBLE] positive and [INAUDIBLE].

**KENNETH ABBOTT:** Yes, yes. Assuming my distribution is symmetric. Now that's the right assumption to point out. Because in the real world, it may not be symmetric. And when we go into historical simulation, we use empirical distributions where we don't care if it's symmetric because we're only looking at the downside. And whether I'm long or short, I might care about the downside or the pretty upside. Because it'll be short, and I care about how much is going to move up. Make sense? That's the right question to ask. Yes?

**AUDIENCE:** [INAUDIBLE] if you're doing it for upside as well?

**KENNETH** Yes.

**ABBOTT:**

**AUDIENCE:** Could it just be the same thing?

**KENNETH**  
**ABBOTT:** Yes. In fact, in this case, in what we're doing here of variance covariance or closed form var, it's for long or short. But getting your signs right, I'm telling you, it's like physics. I still make that mistake. Yes?

**AUDIENCE:** [INAUDIBLE] symmetric. Do you guys still use this process to say, OK--

**KENNETH**  
**ABBOTT:** I use it all the time as a heuristic. All right? Because let's say I've got-- and that's a very good question-- let's say I've got five years worth of data and I don't have time to do an empirical estimate. It could be lopsided. If you tell me a two standard deviation move is  $x$ , that means something to me. Now, there's a problem with that. And the problem is that people extrapolate that. Sometimes people talk to me and, oh, it's an eight standard deviation move. Eight standard deviation moves don't happen. I don't think we've seen an eight standard deviation move in the Cenozoic era. It just doesn't happen.

Three standard deviation-- you will see a three standard deviation move once every 10,000 observations. Now, I learned this the hard way by just, see how many times do I have to do this? And then I looked it up in the table, oh, I was right. When we oversimplify, and start to talk about everything in terms of that normal distribution, we really just lose our grip on reality. But I use it as a heuristic all the time. I'll do it even now, and I know better. But I'll go, what's two standard deviations? What's three standard deviations?

Because by and large-- and I still do this, I get my data and I line it up and I do frequency distributions. Hold on, I do this all the time with my data. Is it symmetric? Is it fat tailed? Is it unimodal? So that's a very good question. Any other questions?

**AUDIENCE:** [INAUDIBLE] have we talked about the [? standard t ?] distribution?

**PROFESSOR:** We introduced it in the last lecture. And the problems set this week does relate to that.

**KENNETH ABBOTT:** All right, perfect lead in. So the statement I made, it's 1% of the time I'd expect to lose more than 4.2 pesos on 100 peso position. That's my inferential statement. In fact, over the same time period I lost 4.2% 1.5% of the time instead of 1% of the time. What that tells me, what that suggests to me, is my data set has fat tails. What that means is the likelihood of a loss-- a simple way of thinking about it [INAUDIBLE] care whether what that means in a metaphysical sense, a way to interpret it. The likelihood of a loss is greater than would be implied by the normal distribution. All right?

So when you hear people say fat tails, generally, that's what they're talking about. There are different ways you could interpret that statement, but when somebody is talking about a financial time series, it has fat tails. Roughly 3/4 of your financial time series will have fat tails. They will also have time series properties, they won't be true random walks. True random walks says that I don't know whether it's going to go up or down based on the data I have. The time series has no memory.

When we start introducing time series properties, which many financial time series have, then there's seasonality, there's mean reversion, there's all kinds of other stuff, other ways that we have to think about modeling the data. Make sense?

**AUDIENCE:** [INAUDIBLE] higher standard deviation than [INAUDIBLE].

**KENNETH ABBOTT:** Say it once again.

**AUDIENCE:** Better yield, does it mean that we have a higher standard deviation than [INAUDIBLE]?

**KENNETH ABBOTT:** No. The standard deviation is the standard deviation. No matter what I do, this is standard deviation, that's it. Don't have a higher standard deviation. But the likelihood of-- the put it this way-- the likelihood of a move of 2.33 standard deviations is more than 1%. That's the way I think of it. Make sense?

**AUDIENCE:** Is there any way for you to [INAUDIBLE] to--

**KENNETH**

What?

**ABBOTT:**

**AUDIENCE:**

Sorry, is there any way to put into that graph what a fatter tail looks like?

**KENNETH**

Oh, well, be patient. If we have time. In fact, we do that all the time. And one of our techniques doesn't care. It goes to the empirical distribution. So it captures the fat tails completely. In fact, the homework assignment which I usually precede this lecture by has people graphing all kinds of distributions to see what these things look like. We won't have time for that. But if you have questions, send them to me. I'll send you some stuff to read about this.

**ABBOTT:**

All right, so now you know one asset var, now you're qualified to go work for a big bank. All right? Get your data, calculate returns. Now I usually put in step 2b, graph your data and look at it. All right? Because everybody's data has dirt in it. Don't trust anyone else. If you're going to get fired, get fired for being incompetent, don't get fired for using someone else's bad data. Don't trust anyone. My mother gives me data, Mom, I'm graphing it. Because I think you let some poop slip into my data.

Mother Theresa could come to me with a thumb drive [INAUDIBLE] S&P 500. Sorry, Mother Teresa. I'm graphing it before I use it. All right? So I don't want to say that this is usually in here. We do extensive error testing. Because there could be bad data, there could be missing data. And missing data is a whole other lecture that I give. You might be shocked at [INAUDIBLE].

So for one asset var, get my data, create my return series. Percentage changes, log changes. Sometimes that's what the difference is. Take the variance, take the square root of the variance, multiply by 2.33. Done and dusted. Go home, take your shoes off, relax. OK. Percentage changes versus log changes. For all intents and purposes, it doesn't really matter and I will often use one or the other.

The way I think about this-- all right, there'll be a little bit of bias at the ends. But for the overwhelming bulk of the observations whether you use percentage changes or log changes doesn't matter. Generally, even though I know the data is closer to log

normally distributed than normally distributed, I'll use percentage changes just because it's easier. Why would we use log normal distribution? Well, when we're doing simulation, the log normal distribution has this very nifty property of keeping your yields from going negative. But, even that-- I can call that into question because there are instances of yields going negative. It's happened. Doesn't happen a lot, but it happens.

All right. So I talked about bad data, talked about one sided versus two sided. I'll talk about longs and shorts a little bit later when we we're talking multi asset. I'm going to cover a fixed income piece. We use this thing called a PV01 because what I measure in fixed income markets isn't a price. [? Asian ?] measure a yield. I have to get from a change of yield to a change of price. Hm, sounds like a Jacobian, right? With kind of a poor man's Jacobian. It's a measure that captures the fact that my price yield relationship-- price, yield-- is non-linear.

For any small approximation I look at the tangent. And I use my PV01 which has a similar notion to duration, but PV01 is a little more practical. The slope of that tells me how much my price will change for a given change of yield. See, there it is. You knew you were going to use the calculus, right? You're always using the calculus. You can't escape it. But the price yield line is non-linear.

But for all intents and purposes, what I'm doing is I'm shifting the price yield relationship-- I'm shifting my yield change into price change by multiplying my yield change by my PV01 which is my price sensitivity to 1/100th percent move in yields. Think about that for a second. We don't have time to-- I would love to spend an hour on this, and about trading strategies, and about bull steepeners and bear steepeners in barbell trades, but we don't have time for that.

Suffice to say if I'm measuring yields the thing is going to trade as a 789 or a 622 or a 401 yield. How do I get that in the change in price? Because I can't tell my boss, hey, I had a good day. I bought it at 402 and sold it at 401. No, how much money did you make? Yield to coffee break yield to lunch time, yield to go home at the end of day. How do I get from change in yield to change in price? Usually PV01. I could

use duration. Bond traders who think in terms of yield to coffee break, yield to lunch time, yield to go home at the end of the day typically think in terms of PV01. Do you agree with that statement?

**AUDIENCE:** [INAUDIBLE]

**KENNETH** How often on the fixed income [? desk ?] did you use duration measures?

**ABBOTT:**

**AUDIENCE:** Well, actually, [INAUDIBLE].

**KENNETH** Because of the investor horizon? OK, the insurance companies. Very important point I want to reach here as a quick aside. You're going to hear this notion of PV01, which is called PVBP or DV01. That's the price sensitivity to a one basis point move. One basis point is 1/100th of a percent in yield. Duration is the half life, essentially, of my cash flow. What's the weighted expected time to owe my cash flows?

If my duration is 7.9 years, my PV01 is probably about \$790 per million. In terms of significant digits, they're roughly the same but they have different meanings and the units are different. Duration is measured in yield, PV01 is measured in dollars. In bond space I typically think in PV01. If I'm selling to long term investors they have particular demands because they've got cash flow payments they have to hedge. So they may think of it in terms of duration. For our purposes, we're talking DV01 or PV01 or PVBP, those three terms more or less equal. Make sense? Yes?

**AUDIENCE:** [INAUDIBLE] in terms of [INAUDIBLE] versus [INAUDIBLE]?

**KENNETH** We could. In some instances, in some areas and options we might look at an overall 1% move. But we have to look at what trades in the market. What trades in the market is the yield. When we quote the yield, I'm going to quote it going from 702 to 701. I'm not going to have the calculator handy to say, a 702 move to a 701. What's 702 minus 701 divided by 702? Make sense? It's the path of least resistance.

What's the difference between a bond and a bond trader? A bond matures. A little fixed income humor for you. Apparently very little. I don't want to spend too much

time on this because we just don't have the time. I provide an example here. If you guys want examples, contact me. I'll send you the spreadsheets I use for other classes if you just want to play around with it.

When I talk about PV01, when I talk about yields, I usually have some kind of risk free rate. Although this whole notion of the risk free rate, which is-- so much of modern finance is predicated on this assumption that there is a risk free rate, which used to be considered the US treasury. It used to be considered risk free. Well, there's a credit spread out there for US Treasury. I don't mean to throw a monkey wrench into the works. But there's no such thing. I'm not going to question 75 years of academic finance. But it's troublesome.

Just like when I was taking economics 30 years ago, inflation just mucked with everything. All of the models fell apart. There were appendices to every chapter on how you have to change this model to address inflation. And then inflation went way and everything was better. But this may not go away. I've got two components here. If the yield is 6%, I might have a 450 treasury rate and 150 basis point credit spread. The credit spread reflects the probability of default.

And I don't want to get into measures of risk neutrality here. But if I'm an issuer and I have a chance of default, I have to pay my investors more. Usually when we measure sensitivity we talk about that credit spread sensitivity and the risk free sensitivity. We say, well, how could they possibly be different? And I don't want to get into detail here, but the notion is, when credit spreads start getting high, it implies a higher probability of default.

You have to think about credit spreads sensitivity a little differently. Because when you get to 1,000 basis points, 1,500 basis points credit spread, it's a high probability of default. And your credit models will think different. Your credit models will say, ah, that means I'm not going to get my next three payments. There's an expected, there's a probability of default, there's a loss given default, and there's recovery. A bunch of other stochastic measures come into play. I don't want to spend any more time on it because it's just going to confuse you now. Suffice to say we have these



yields and yields are composed of risk free rates and credit spreads.

And I apologize for rushing through that, but we don't have time to do it. Typically you have more than one asset. So in this framework where I take 2.33 standard deviations times my dollar investment, or my renminbi investment or my sterling investment. That example was with one asset. If I want to expand this, I can expand this using this notion of covariance and correlation.

You guys covered correlation and covariance at some point in your careers? Yes, no? All right? Both of them measure the way one asset moves vis a vis another asset. Correlation is scaled between negative 1 and positive 1. So I think of correlation as an index of linearity. Covariance is not scaled. I'll give you an example of the difference between covariance and correlation.

What if I have 50 years of data on crop yields and that same 50 years of data on tons of fertilizer used? I would expect a positive correlation between tons of fertilizer used and crop yields. So the correlation would exist between negative 1 and positive 1. The covariance could be any number, and that covariance will change depending on whether I measure my fertilizer in tons, or in pounds, or in ounces, or in kilos. The correlation will always be exactly the same. The linear relationship is captured by the correlation.

But the units-- in covariance, the units count. If I have covariance-- here it is. Covariance matrices are symmetric. They have the variance along the diagonal. And the covariance is on the off diagonal. Which is to say that the variance is the covariance of an item with itself. The correlation matrix, also symmetric, is the same thing scaled with correlations, where the diagonal is 1.0.

If I have covariance-- because correlation is covariance-- covariance divided by the product of the standard deviations. Gets me-- sorry-- correlation hat. This is like the apostrophe in French. You forget it all the time. But the one time you really need it, you won't do it and you'll be in trouble. If you have the covariances, you can get to the correlations. If you have the correlations, you can't get to the covariances unless you know the variances. That's a classic mid-term question. I give that almost-- not

every year, maybe every other year. Don't have time to spend much more time on it.

Suffice to say this measure of covariance says when  $x$  is a certain distance from its mean, how far is  $y$  from its mean and in what direction? Yes? Now this is just a little empirical stuff because I'm not as clever as you guys. And I don't trust anyone. I read it in the textbook, I don't trust anyone.  $a, b$ , here's  $a + b$ . Variance of  $a + b$  is variance of  $a$  plus variance of  $b$  plus 2 times covariance  $a, b$ . It's not just a good idea, it's the law.

I saw it in a thousand statistics textbooks, I tested it anyway. Because if I want to get fired, I'm going to get fired for making my own mistake, not making someone else's mistake. I do this all the time. And I just prove it empirically here. The proof of which will be left to the reader as an exercise. I hated when books said that.

**PROFESSOR:** I actually kind of think that's a proven point, that you really should never trust output from computer programs or packages--

**KENNETH** Or your mother, or Mother Teresa.

**ABBOTT:**

**PROFESSOR:** It's good to check them. Check all the calculations.

**KENNETH** Mother Teresa will slip you some bad data if she can. I'm telling you, she will. She's  
**ABBOTT:** tricky that way. Don't trust anyone. I've caught mistakes in software, all right? I had a programmer-- it's one of my favorite stories-- we're doing one of our first Monte Carlo simulations, and we're factoring a matrix. If we have time, we'll get-- so I factor a covariance matrix into  $E^T \lambda E$ . It's our friend the quadratic form. We're going to see this again. And this is a diagonal matrix of eigenvalues. And I take the square root of that. So I can say this is  $E^T \lambda^{1/2} E$ .

And so my programmer had gotten this, and I said, do me a favor. I said, take this, and transpose and multiply by itself. So take the square root and multiply it by the other square root, and show me that you get this. Just show me. He said I got it. I

said you got it? He said out to 16 decimals. I said stop. On my block, the square root of 2 times the square root of 2 equals 2.0. All right? 2.0000000 what do you mean out to 16 decimal places? What planet are you on?

And I scratched the surface, and I dug, and I asked a bunch of questions. And it turned out in this code he was passing a float to a [? fixed. ?] All right? Don't trust anyone's software. Check it yourself. Someday when I'm dead and you guys are in my position, you'll be thanking me for that. Put a stone on my grave or something.

All right so covariance. Covariance tells me some measure of when x moves, how far does y move? [? Or ?] for any other asset? Could I have a piece of your cookie? I hardly had lunch. You want me to have a piece of this, right? It's just looking very good there. Thank you. It's foraging. I'm convinced 10 million years ago, my ape ancestors were the first one at the dead antelope on the planes. All right. So we're talking about correlation covariance. Covariance is not unit free. I can use either, but I have to make sure I get my units right. Units screw me up every time. They still screw me up. That was a good cookie.

All right. So more facts. Variance of  $x$  times  $y$   $x$  squared variance  $a$   $y$  squared variance  $b$  plus  $2xy$  covariance  $ab$ . You guys seen this before? I assume you have. Now I can get pretty silly with this if I want.  $xayb$  you get the picture, right? But what you should be thinking, this is a covariance matrix,  $\sigma$  squared,  $\sigma$  squared,  $\sigma$  squared. It's the sum of the variances plus 2 times the sum of the covariances. So if I have one unit of every asset, I've got  $n$  assets, all have to do to get the portfolio variance is sum up the whole covariance matrix.

Now, you never get only one unit, but just saying. But you notice that this is kind of a regular pattern that we see here. And so what I can do is I can use a combination of my correlation matrix and a little bit of linear algebra [? ledger ?] domain, to do some very convenient calculations. And here I just give an example of a covariance matrix and a correlation matrix. Note the correlation matrices between negative 1 and positive 1. All right. Let me cut to the chase here.

I'll draw it here because I really want to get into some of the other stuff. What this

means, if I have a covariance structure,  $\sigma$ . And I have a vector of positions,  $x$  dollars in dollar yen,  $y$  dollars in gold,  $z$  dollars in oil. And let's say I've got a position vector,  $x_1, x_2, x_3, x_n$ . If I have all my positions recorded as a vector. This is asset one, asset two, and this is in dollars. And I have the covariance structure, the variance of this portfolio that has these assets and this covariance structure-- this is where the magic happens-- is  $x^T \sigma x = \sigma^2$  portfolio.

Now you really could go work for a bank. This is how portfolio variance, using the variance covariance method, is done. In fact, when we were doing it this way 20 years ago, spreadsheets only have 256 columns. So we tried to simplify everything into 256-- or sometimes you had to sum it up using two different spreadsheets. We didn't have multitab spreadsheets. That was a dream, multitab spreadsheets. This was Lotus 1-2-3 we're talking about here, OK? You guys don't even know what Lotus 1-2-3 is. It's like an abacus but on the screen. Yes?

**AUDIENCE:** What's  $x$  again in this?

**KENNETH ABBOTT:** Position vector. Let's say I tell you that you've got dollar yen, gold, and oil. You've got \$100 of dollar yen, \$50 of oil, and \$25 of gold. It would be 100, 50, 25. Now, I should say \$100 of dollar yen, your position vector would actually show up as negative 100, 50, 25. Why is that? Because if I'm measuring my dollar yen-- and this is just a little aside-- typically, I measure dollar yen in yen per dollar. So dollar yen might be 95. If I own yen and I'm a dollar investor and I own yen, and yen go from 95 per dollar to 100 per dollar, do I make or lose money? I lose money. Negative 100.

Just store that. You won't be tested on that, but we think about that all the time. Same thing with yields. Typically, when I record my PV01-- and I'll record some version, something like my PV01 in that vector, my interest rate sensitivity, I'm going to record it as a negative. Because when yields go up and I own the bond, I lose money. Signs, very important.

And, again, we've covered-- usually I do this in a two hour lecture. And we've

covered it in less than an hour, so pretty good. All right. I spent a lot more time on the fixed income.

[STUDENT COUGHING]

Are you taking something for that? That does not sound healthy. I don't mean to embarrass you. But I just want to make sure that you're taking care of yourself because grad students don't-- I was a grad student, I didn't take care of myself very well. I worry. All right. Big picture, variance covariance. Collect data, calculate returns, test the data, matrix construction, get my position vector, multiply my matrices. All right? Quick and dirty, that's how we do it. That's the simplified approach to measuring this order statistic called value at risk using this particular technique.

Questions, comments? Anyone? Anything you think I need to elucidate on that? And this is, in fact, how we did this up until the late '90s. Firms used variance covariance. I heard a statistic in Europe in 1996 that 80% of the European banks were using this technique to do their value at risk. It was no more complicated than this.

I use a little flow diagram. Get your data returns, graph your data to make sure you don't screw it up. Get your covariance matrix, multiply your matrices out.  $x$  transpose  $\sigma$   $x$ . Using the position vectors and then you can do your analysis. Normally I would spend some more time on that bottom row and different things you can do with it, but that will have to suffice for now.

A couple of points I want to make before we move on about the assumptions. Actually, I'll fly through this here so we can get into Monte Carlo simulation. Where am I going to get my data? Where do I get my data? I often get a lot of my data Bloomberg, I get it from public sources, I get it from the internet. Especially when you get it from-- look, if it says so on the internet, it must be true. Right? Didn't Abe Lincoln say, don't believe everything you read on the internet? That was a quote, I saw that some place.

You get data from people, you check it. There's some sources that are very reliable.

If you're looking for yield data or foreign exchange data, the Federal Reserve has it. And they have it back 20 years, daily data. It's the H.15 and the H.10. It's there, it's free, it's easy to download, just be aware of it. Exchange--

**PROFESSOR:** [INAUDIBLE] study posted on the website that goes through computations for regression analysis and asset pricing models and the data that's used there is from the Federal Reserve for yields.

**KENNETH** It's H15 It's for yields, it's probably from the H.15.

**ABBOTT:**

[INTERPOSING VOICES]

**PROFESSOR:** Those files, you can see how to actually get that data for yourselves.

**KENNETH** Now, another great source of data is Bloomberg. Now the good thing about  
**ABBOTT:** Blumberg data is everybody uses it, so it's clean. Relatively clean. I still find errors in it from time to time. But what happens is when you find an error in your Bloomberg data, you get on the phone to Bloomberg right away and say I found an error in your data. They say, oh, what date? June 14, you know, 2012. And they'll say, OK, we'll fix it.

All right? So everybody does that, and the data set is pretty clean. I found consistently that Bloomberg data is the cleanest in my experience. How much data do we use in doing this? I could use one year of data, I can use two weeks of data. Now, times series, we usually want 100 observations. That's always been my rule of thumb. I can use one year of data. There are regulators that require you to use at least a year of data. You could use two years of data. In fact, some firms use one year of data. There's one firm that uses five years of data.

And there, we could say, well, am I going to weight it. Am I going to weight my more recent data heavily? I could do that with exponential smoothing, which we won't have time to talk about. It's a technique I can use to lend more credence to the more recent data. Now, I'm a relatively simple guy. I tend to use equally weighted data because I believe in Occam's razor, which is, the simplest explanation is

usually the best. I think we get too clever by half when we try to parameterize. How much more does last week's data have an impact than from two weeks ago, three weeks ago.

I'm not saying that it doesn't, what I am saying is, I'm not smart enough to know exactly how much it does. And assuming that everything's equally weighted throughout time is just as strong an assumption. But it's a very simple assumption, and I love simple. Yes?

**AUDIENCE:** [INAUDIBLE] calculate covariance matrix?

**KENNETH ABBOTT:** Yes. All right, quickly. Actually I think I have some slides on that. Let me just finish this and I'll get to that. Gaps in data. Missing data is a problem. How do I fill in missing data? I can do a linear interpolation, I can use the prior day's data. I can do a Brownian bridge, which is I just do a Monte Carlo between them. I can do a regression based, I can use regression to project changes from one onto changes in another. That's usually a whole other lecture I gave on how to do missing data. Now you've got that lecture for free. That's all you need to know. It's not only a lecture, it's a very hard homework assignment.

But how frequently do I update my data? Some people update their covariance structures daily. I think that's an overkill. We update our data set weekly. That's what we do. And I think that's overkill, but tell that to my regulators. And we use daily data, weekly data, monthly data. We typically use daily data. Some firms may do it differently. All right. Here's your exponential smoothing.

Remember, I usually measure covariance sum of  $x_i - \bar{x}$  times  $y_i - \bar{y}$  divided by  $n - 1$ . What if I stuck an  $\omega$  in there? And I use this calculation instead, where the denominator is the sum of all the  $\omega$ s-- you should be thinking finite series. You have to realize, I was a decent math student, I wasn't a great math student. And what I found when I was studying this, I was like, wow, all that stuff that I learned, it actually-- finite series, who knew? Who knew that I'd actually use it?

So I take this, and let's say I'm working backwards in time. So today's observations is  $t=0$ . Yesterday's observation is  $t_1, t_2, t_3$ . So today's observation would get-- and let's assume for the time being that this  $\omega$  is on the order 0.95. It could be anything. So today would be  $0.95^0$  divided by the sum of all the  $\omega$ s. Tomorrow it will be  $0.95^1$  divided by the sum of the  $\omega$ s. The next would be  $0.95^2$  squared divided by the sum of the  $\omega$ s.  $0.95^3$  cubed and get smaller and smaller.

For example, if you use 0.94, 99% of your weight will be in the last 76 days. 76 observations, I shouldn't say 76 days. 76 observations. So there's this notion that the impact declines exponentially. Does that make sense? People use this pretty commonly, but what scares me about it-- somebody stuck these fancy transitions in between these slides. Anyway, is that here's my standard deviation [INAUDIBLE] the rolling six [INAUDIBLE] window. And here's my standard deviation using different weights.

The point I want to make here, and it's an important point, my assumption about my weighting coefficient has a material impact on the size of my measure volatility. Now when I see this, and this is just me. There's no finance or statistics theory behind this, any time the choice-- any time an assumption has this material an impact, bells and whistles go off and sirens. All right, and red lights flash. Be very, very careful.

Now, lies, damn lies, and statistics. You tell me the outcome you want, and I'll tell you what statistics to use. That's where this could be abused. Oh, you want to show high volatility? Well let's use this. You want to show low volatility, let's use this? See, I choose to just take the simplest approach. And that's me. That's not a terribly scientific opinion, but that's what I think. Daily versus weekly, percentage changes log changes. Units.

Just like dollar yen, interest rates. Am I long or am I short? If I'm long gold, I show it as a positive number. And if I'm short gold, in my position vector, I show it as a negative number. If I'm long yen, and yen is measured in yen per dollar, then I show it as a negative number. If I'm long yen, but my covariance matrix measures yen as



dollars per yen-- 0.000094, whatever-- then I show it as a positive number.

It's just like physics only worse because it'll cost you real-- no, I guess physics would be worse because if you get the units wrong, you blow up, right? This will just cost you money. I've made this mistake. I've made the units mistake. All right, we talked about fixed income. So that's what I want to cover from the bare bones setup for var. Now I'm going to skip the historical simulation and go right to the Monte Carlo because I want to show you another way we can use covariance structures.

[POWERPOINT SOUND EFFECT]

That's going to happen two or three more times. Somebody did this, somebody made my presentation cute some years ago. And I just-- I apologize. All right, see, there's a lot to meat in this presentation that we don't have time to get to. Another approach to doing value at risk is rather than use this parametric approach, is to simulate the outcomes. Simulate the outcomes 100 times, 1,000 times, 10,000 times, a million times, and say, these are all the possible outcomes based on my simulation assumptions. And let's say I simulate 10,000 times, and I have 10,000 possible outcomes for tomorrow. And I wanted to measure my value at risk at the 1% significance level. All I would do is take my 10,000 outcomes and I would sort them and take my hundredth worst. Put it in your pocket, go home. That's it.

This is a different way of getting to that order statistic. Lends a lot more flexibility. So I can go and I can tweak the way I do that simulation, I can relax my assumptions of normality. I don't have to use normal distribution, I could use a t distribution, I could do lots, I could tweak my distribution, I could customize it. I could put mean reversion in there, I could do all kinds of stuff.

So another way we do value at risk is we simulate possible outcomes. We rank the outcomes, and we just count them. If I've got the 10,000 observations and I want my 5% order statistic, well I just take my 500th. Make sense? It's that simple. Well, I don't want to make it seem like it's that simple because it actually gets a little messy in here. But when we do Monte Carlo simulation, we're simulating what we think is

going to happen all subject to our assumptions.

And we run through this Monte Carlo simulation. Simulation of method using sequences of random numbers. Coined during the Manhattan Project, similar to games of chance. You need to describe your system in terms of probability density functions. What type of distribution? Is this normal? Is it  $t$ ? Is it chi squared? Is it  $F$ ? All right? That's the way we do it. So quickly, how do I do that?

I have to have random numbers. Now they're truly random numbers. Somewhere at MIT you could buy-- I used to say tape, but people don't use tape. They'll give you a website where you can get the atomic decay. That's random. All right? Anything else is pseudo random. What you see when you go into MATLAB, you have a random number generator, it's an algorithm. It probably takes some number and takes the square root of that number and then goes 54 decimal places to the right and takes the 55 decimal places to the right, multiplies those two numbers together and then takes the fifth root, and then goes 16 decimal places to the right to get that-- it's some algorithm.

True story, before I came to appreciate that these were all highly algorithmically driven, I was in my 20's, I was taking a computer class, I saw two computers, they were both running random number of generators and they were generating the same random numbers. And I thought I was at the event horizon. I thought that light was bending and the world was coming to an end, all right? Because this this stuff can't happen, all right? It was happening right in front of me. It was a pseudo random number generator. I didn't know, I was 24.

Anyway. quasi random numbers, it's sort of a way of imposing some order on your random numbers. You random numbers, one particular set of draws may not have enough draws in a particular area to give you the numbers you want. I can impose some conditions upon that. I don't want to get into a discussion of random numbers. How do I get from random uniform-- most remember generous give you random uniform number between 0 and 1. What you'll typically do is you'll take that random uniform number, you'll map it over to the cumulative density function, and map it

down.

So this gets you from random uniform space into standard deviation space. We used to worry about how we did this, now your software does it for you. I've gotten comfortable enough, truth be told. I usually trust my random number generators in Excel, in MATLAB. So I kind of violate my own rules, I don't check. But I think most of your standard random number of generators are decent enough now. And you can go straight to normal, you don't have to do random uniform and back into random normal. You can get it distributed in any way you want.

What I do when I do a Monte Carlo simulation-- and this is going to be rushed because we've only got like 20 minutes. If I take a covariance matrix-- you're going to have to trust me on this because again, I'm covering like eight hours of lecture in an hour and a half. You guys go to MIT so I have no doubt you're going to be all over this.

Let's take this out of here for a second. I can factor my covariance structure. I can factor my covariance structure like this. And this is the transpose of this. I didn't realize that the first time we did this commercially I saw this instead of this and I thought we had sent bad data to the customer. I got physically sick. And then I remembered  $AB^T = B^T A$ . These things keep happening. My high school math keeps coming back to me. But I had forgotten this and I got physically sick because I thought we'd sent bad data because I was looking at this when it's just the transpose of this.

Anyway, I can factor this into this where this is the a matrix of eigenvectors. This is a diagonal matrix of the eigenvalues. All right? This is the vaunted Gaussian copula. This is it. Most people view it as a black box. If you've had any more than introductory statistics, this should be a glass box to you. That's why I wanted to go through this even though I'd love to spend another hour and a half and do about 50 examples. Because this is how I learned this, I didn't learn it from looking at this equation and saying, oh, I get it. I learned it from actually doing it about 1,000 times in a spreadsheet, and sunk in like water into a store.

So I factor this matrix, and then I take this, which is the square root matrix, which is my transpose of my eigenvector matrix and diagonal matrix contain the square root of my eigenvalues. Now, could this ever be negative and take me into imaginary root land? Well, if my variances are positive or zero, then that will be a problem. So here we get into this-- remember you guys studied positive semidefinite, positive definite. Once again, it's another one of these high school math things. Like, here it is. I had to know this.

Suddenly I care whether it's positive semidefinite. Covariance structures have to be positive semidefinite. If you don't have a complete data set, let's say you've got 100 observations, 100 observations, 100 observations, 25 observations, 100 observations, you may have a negative eigenvalue. If you just measure the covariance with the amount of data that you have. My intuition-- and I doubt this is the [INAUDIBLE]-- is that you're measuring with error and you have fewer observations you measure with more error.

So it's possible if some of your covariance measures have 25 observations and some of them have 100 observations that there's more error in some than in others. And so there's the theoretical possibility for negative variance. True story, we didn't do this in the '90s. I took this problem to the chairman of the statistics department at NYU said, I'm getting negative eigenvalues. And he didn't know. He had no idea, he's a smart guy. You have to fill in your missing data. You have to fill in your missing data.

If you've got 1,000 observations, 1,000 observations, 1,000 observations, 200 observations, and you want to make sure you won't have a negative eigenvalue, you've got to fill in those observations. Which is why missing data is a whole other thing we talk about. Again, I could spend a lot of time on that. And I learned that the hard way. But anyway, so I take this square root matrix, if I premultiply that square root matrix by row after row of normals, I will get out an array that has the same covariance structure as that with which I started.

Another story here, I've been using the same eigenvalue-- I believe in full

attribution, I'm not a clever guy. I have not an original thought in my head. And whenever I use someone else's stuff, I give them credit for it. And the guy who wrote the code that did the eigenvalue [? decomposition-- ?] this is something that was translated from Fortran IV. It wasn't even [INAUDIBLE], there's a dichotomy in the world. There are people that have written Fortran, and people that haven't. I'm guessing that there are two people in this room that have ever written a line of Fortran. Anyone here? Just saying. Yeah, with cards or without cards?

**PROFESSOR:** [INAUDIBLE].

**KENNETH ABBOTT:** I didn't use cards. See, you're an old-timer because you used cards. The punch line is, I've been using this guy's code. And I could show you the code. It's like the Lone Ranger, I didn't even get a chance to thank him. Because he didn't put his name on the code. On the internet now, if you do something clever on the quant newsgroups, you're going to post your name all over it. I've been wanting to thank this guy for like 20 years and I haven't been able to. Anyway, [INAUDIBLE] code that's been translated.

Let me show you what this means. Here's some source data. Here's some percentage changes. Just like we talked about. Here is the empirical correlation of those percentage changes. So the correlation of my government 10 year to my AAA 10 year is 0.83. To my AA, 8.4. All right, you see this. And I have this covariance matrix which is the-- the correlation matrix is a scaled version of the covariance matrix.

And I do a little bit of statistical ledger domain. Eigenvalues and eigenvectors. Take the square root of that. And again, I'd love to spend a lot more time on this, but we just don't-- suffice to say, I call this a transformation matrix, that's my term. This matrix here is this. If we had another hour and a half I'd take the step by step to get you there. The proof of which is left to the reader as an exercise. I'll leave this spreadsheet for you, I'll send it to you.

I have this matrix. This matrix is like a prism. I'm going to pass white light through it, I'm going to get a beautiful rainbow. Let me show you what I mean. So remember

that matrix, this matrix I'm calling  $t$ . Remember my matrix is 10 by 10. One, two, three, four, five, six, seven, eight, nine, ten. 10 columns of data. 10 by 10 correlation matrix. Let's check. Now I've got row vectors of sorry-- uncorrelated random normals.

So what I'm doing then is I'm premultiplying that transformation matrix row by row by each row of uncorrelated random normals. And what I get is correlated random normals. So what I'm telling you here is this array happens to be 10 wide and 1,000 long. And I'm telling you that I started with my historical data-- let me see how much data have there. A couple hundred observations of historical data.

And what I've done is once I have that covariance structure, I can create a data set here which has the same statistical properties as this. Not quite the same. It can have the same means and the same variances. This is what Monte Carlo simulation is about. I wish we had another hour because I'd like to spend time and-- this is one of these things, and again, when I first saw this, I was like, oh my god. I felt like I got the keys to the kingdom.

And I did it manually, did it all on a spreadsheet. Didn't believe anyone else's code, did it all on a spreadsheet. But what that means quickly, let me just go back over here for a second. I happen to have about 800 observations here. Historical observations. What I did was I happened to generate 1,000 samples here. But I could generate 10,000 or 100,000, or a million or 10 million or a billion just by doing more random normals. I could generate-- in effect, what I'm generating here is synthetic time series that have properties similar to my underlying data.

That's what Monte Carlo simulation is about. The means and the variances and the covariances of this data set are just like that. Now, again, true story, when somebody first showed me this I did not believe them. So I developed a bunch of little tests. And I said, let me just look at the correlation of my Monte Carlo data versus my original correlation matrix. So 0.83, 0.84, 0.85, 0.85, 0.67, 0.81. You look at the corresponding ones of the random numbers I just generated, 0.81, 0.82, 0.84, 0.84, 0.64, 0.52. 0.54 versus 0.52. 0.18 Versus 0.12. 0.51 versus 0.47.

Somebody want to tell me why they're not spot on? Sampling error.

The more data I use the closer it will get to that. If I do 1 million, I'd better get right on top of that. Does that make sense? So what I'm telling you here is that I can generate synthetic time series. Now, why would I generate so many? Well because, remember, I care what's going on out in that tail. If I only have 100 observations and I'm looking empirically at my tail, I've only got one observation out in the 1% tail. And that doesn't tell me a whole lot about what's going on. If I can simulate that distribution exactly, I can say, you know what, I want a billion observations in that tail. Now we can look at that tail.

If I have 1 billion observations, let's say I'm looking at some kind of normal distribution. I'm circling it out here, I'm seeing-- I can really dig in and see what the properties of this thing are. In fact, this can really only take two distributions, and really, it's only one. But that's another story. So what I do in Monte Carlo simulations, I'm and simulated these outcomes so we can get a lot more meat in this tail to understand what's happening out there. Does it drop off quickly? Does it not drop off quickly?

That's kind of what it's about. So we're about out of time. We just covered like four weeks of material, all right? But you guys are from MIT. I have complete confidence in you. I say that to the people who work for me. I have complete confidence in your ability to get that done by tomorrow morning. Questions or comments? I know you're sipping from the fire hose here. I fully appreciate that.

So those are examples. When I do this with historical simulation I won't generate these Monte Carlo trials, I'll just use historical data. And my fat tails are built into it. But what I've shown you today is what we developed a one asset var model, then we developed a multi-asset variance covariance model. And then I showed you quickly, and in far less time than I would like to have shown you is how I can use another statistical technique, which is called the Gaussian copula, to generate has data sets that will have the same properties as my source historical data. All right?

There you have it.

[APPLAUSE]

Oh you don't have to-- please, please, please. And I'll tell you, for me, one of the coolest things was actually being able to apply so much of the math I learned in high school and in college and never thought I'd apply again. One of my best moments was actually finding a use for trigonometry. If you're not an engineer, where are you going to use it?

Where do you use it? Seasonals. You do seasonal estimation. And what you do is you do fast Fourier transform. Because I can describe any seasonal pattern with a linear combination of sine and cosine functions. And it actually works. I have my students do it as an exercise every year. I say, go get New York city temperature data. And show me some linear combination of sine and cosine functions that will show me the seasonal pattern of temperature data. And when I first realized I could use trigonometry, yes! It wasn't a waste of time. I still pull the coordinates, I still haven't found a use for that one. But it's there. I know it's there. All right? Go home.