Large portfolios are commonly rated for risk as if aggregate losses are bound to fall within a few standard deviations of the perceived mean. This can be wildly inappropriate, especially for senior debt tranches. Ignorance, greed, and bad regulatory incentives have driven gross misuse verging on fraud. We’ll examine some healthier alternatives.

Readers who have been fretting the dearth of moneymaking advice in this book need look no further. This chapter will demonstrate how to rake in fortunes through creative statistical fraud. Granted, others have beaten us to much booty. Still, foolish regulations, negligent rating practices, and irresponsible fiduciaries invite more ill-gotten gains.

Here is a basic recipe:

- Take a big bunch of different credit risks, just like a normal bank does, only with far less supervision.
- Repackage the claims in tranches ranked by seniority, and give them impressive names like Collateralized Debt Obligations (CDOs).
- Let the lowest tranches shoulder the mean observed risks and the middle tranches the risks of three to four standard deviation outliers.
- Assume the risks are normally distributed to prove that the highest tranches are nearly risk free.
- Hire rating agencies to certify the calculations, and rely on regulators and major institutional investors to accept the certifications on faith.
- Offer the certified tranches for sale as tradable securities.
• Once a credit boom gathers steam, offer the highest tranches to anyone seeking something for nothing.

Like any sting, this recipe works best if the buyers don’t realize they’re being had. For awe and admiration, hire finance PhDs and MBAs with impressive qualifications, great networking skills, and little real understanding of what they’re hawking. When in doubt, speak of Gaussian copulas. Faced with doubt, emphasize that securitization itself inspires trust by providing liquidity.

This chapter will expose the statistical underpinnings to these facades. It will also show how to analyze portfolio risks better. Hopefully it will encourage higher standards for portfolio risk analysis.

Exchangeability

The easiest way to understand dependence across risks is to conceive each individual risk as very simple and statistically exchangeable with other risks. Simple we already know: default risk. So let’s take $n$ credit assets and tag them in some order. The portfolio outcome is a string of zeros and ones indicating which pay and which default. Exchangeable means that if the tags get mixed up we can’t detect any difference in the probabilities; all that matters is how many assets default or don’t.

Given 100 default risks, there are $2^{100}$ distinct portfolio outcomes. If those risks are exchangeable, 100 probabilities completely define the rest. That is a colossal savings over the general case. We need something like that to make the analysis tractable.

However, even tracking 100 probabilities is a lot of trouble. So we look for summary statistics like mean and standard deviation. Unfortunately, means and standard deviations don’t do a very good job of indicating tail risks, even in perfectly exchangeable portfolios.

Osband (2002) introduced a contraption that tries to make this intuitively clear. He called it Devlin’s Triangle after a fictional character and hoped risk analysts would adopt it. They haven’t, and since I hate to see an orphan, I will claim it for my own.

To build an Osband triangle, lay a bed of $m + 1$ nonnegative numbers on the bottom, with gaps in between. For the layer above, add neighboring numbers and fasten the sum over the gap. Keep adding layers until there’s
a single number. Then divide every element by the top number. Here is an
example for $m = 3$.

\[
\begin{array}{cccc}
1 \\
\frac{1}{2} & \frac{1}{2} \\
\frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\
\frac{1}{4} & 0 & \frac{1}{4} & 0 \\
\end{array}
\]

Every triangle constructed in this fashion turns out to represent a feasible set of exchangeable default risks. The $k^{th}$ entry on the $n+1^{th}$ row denotes the probability that in any subset of $n$ assets, the first $k$ assets default and the rest do not.

To calculate the probability that $k$ out of $n$ assets default in any order, multiply these entries by the corresponding entries in a much better-known triangle. It is called Pascal’s triangle in the West and Yang Hui’s triangle in China, although Persian and Indian mathematicians invented it centuries before. That triangle also starts from one at the top but adds neighboring elements down rather than up. The $k^{th}$ entry on the $n+1^{th}$ row indicates the number of distinct combinations of $k$ hits and $n-k$ misses. Unlike in an Osband triangle, the numbers are always the same and look like this:

\[
\begin{array}{cccc}
1 \\
1 & 1 \\
1 & 2 & 1 \\
1 & 3 & 3 & 1 \\
1 & 4 & 6 & 4 & 1 \\
\end{array}
\]

In the example above, multiplying the two triangles tells us level by level that—

1. The probability of the whole is 1.
2. Every asset is equally likely to default or service.
3. Pairs of assets have a 25% chance of both defaulting and a 25% of both servicing.
4. Three assets together have a 25% chance of no default and a 75% chance of two defaults.
As first glance it might seem impossible for exchangeable risks to take this form. The following example refutes this. Three companies are competing for a government contract to set a new networking standard. Securing a contract is life-and-death for the companies. The government considers four alternative awards: to the first, to the second, to the third, and to all three jointly. If each award is equally likely it generates the risk structure above.

This example starts out looking like a standard coin toss but isn’t. Usually it’s impossible to infer from any given row what an Osband triangle looks like below. The main exceptions can’t be extended at all. In particular, there’s no way to extend the row $[\frac{1}{4} \ 0 \ \frac{1}{4} \ 0]$ or any other row with an element more than twice the sum of its neighbors.

For more generic behavior, we might exclude Osband triangles that can’t be extended indefinitely, even if they hit bedrock only after millions of layers. De Finetti (1931a) showed that only one class of default risks is infinitely exchangeable. This “mixed binomial” class denotes, in effect, that assets are conditionally independent given the default rates.

In general, most exchangeable random variables can be viewed as conditionally independent. This is known as de Finetti’s Theorem. It is an extremely important result in probability theory. It carves a world of interdependent randomness into things that move completely in sync and things that move completely on their own.

Imagine that a physicist is measuring molecular motion inside a balloon when a child runs away with the balloon. It would be helpful to distinguish the random movements inside the balloon from the child’s motion that shifts the balloon. It would be even more helpful to separate child from balloon. That creates a controlled experiment. Unfortunately, most financial analyses are so loosely controlled we have to settle for distinguishing things in our minds.

The mixing distribution corresponds to the child’s motion. It determines which $\theta$ applies. Given $\theta$, we assume the remaining risks are independent and identically distributed.

**Creation of Correlation**

It’s a lot easier to think about mixing distributions for conditionally independent risks than to estimate separate probabilities for every joint outcome.
However, many risk analysts lack patience even for that. They prefer instead to think about summary statistics of the joint distribution, particularly the mean and the correlation.

As we’ve seen, the mean is hard to measure. The correlation is even harder to measure. But let’s imagine we nail them both. Indeed, let’s imagine that we can measure any moment we please, and eliminate any default risk triangle yielding a different value. What is left?

The answer is remarkably neat. The first \( n \) moments of the portfolio uniquely identify the first \( n + 1 \) rows and vice-versa. (The first row can be viewed as matching the 0th moment, which sums the probabilities to one.) In particular, the mean and correlation indicate the probabilities of zero, one, or two defaults in any pair of credits. That leaves a lot of wiggle room for rows further down.

Let’s work through an example. Suppose we own 1,000 exchangeable default risks. We estimate a mean of 0.5% and a correlation of 2%. This implies a 0.9751% risk of one default in a pair and a 0.01245% risk of two defaults in a pair.

Now let us infer the risk of experiencing 50 or more defaults. That is at least ten times the mean and marks a 4.4 standard deviation outlier. Moreover, if we cheat and claim zero correlation, the standard deviation of defaults will collapse to 2.2 from 10.2, making this a 20 standard deviation event. Very few risk analysts—let alone their employers, clients, or regulators—knowingly worry about 20 standard deviation events. So it’s tempting to round the number to zero and move on.

However, all our certainty technically implies is a 50 bps mean and 99.75 bps standard deviation on the mixing distribution. One mixing distribution meeting those criteria puts a 3.8% probability on a 550+ bps risk and all remaining probability on a 30 bps risk. In that case, there’s over a 3% chance of 50 or more defaults. Another feasible mixing distribution attaches 25.1% probability to a 199 bps risk and the rest to zero risk, for a probability of 50+ defaults of barely one in 500 million.

In general, all the correlation \( \rho \) tells us is the ratio of two variances:

\[
\rho = \frac{\text{var}(\text{commonRisk})}{\text{var}(\text{singleCredit})}. \tag{9.1}
\]

Hence, from a mixing distribution perspective, correlation isn’t causing anything. It’s simply a measure induced by the variance of the common risk.
And since the variance can’t describe the full tail of risks, correlation can’t either. Neither can the skewness or kurtosis, which specify another one or two rows down the Osband triangle.

Consequently, when practitioners infer tail risks from estimates of moments, they fold in explicit or implicit assumptions about the forms of the mixing distribution. There’s nothing wrong with that per se. The problem is rather that the most popular assumptions chronically understate the tail risks.

The biggest single error is to take recent history as representative of all likely regimes. If recent history is turbulent, people are reluctant to take the risk. Once bitten, twice shy. Conversely, if recent history is calm, people presume it will persist and tend to take on too much risk. However, the next few pages assume the correlation is correctly estimated.

Binomial and Normal Approximations

The simplest mixing distribution doesn’t mix. It puts all its weight on a single default risk $\theta$. That yields a standard binomial distribution for the portfolio. With $n$ equally-weighted assets, the standard deviation of portfolio returns will be approximately $\sqrt{\theta/n}$. As $n$ increases, risk gradually fades away.

With positive correlation $\rho$, the standard deviation of portfolio returns converges to $\sqrt{\rho\theta}$ rather than zero as $n$ gets large. That can make a huge difference in regulatory capital. It’s as if $n$ were capped at $\frac{1}{\rho}$. Quadrupling $\rho$ doubles the minimum standard deviation. So even though default correlations are generally low and might not appear significantly greater than zero, a lot rides on the value selected.

Adjusting $n$ to make a binomial distribution generate the target standard deviation is the essence of Moody’s Binomial Expansion Technique, or BET. First proposed by Cifuentes and O’Connor (1996), it improves on standard binomial models by allowing a wider dispersion. The adjusted $n$, called the diversity score, can be interpreted as the effective number of independent assets.

The diversity score typically ranges from five to a few dozen. With exchangeable assets, it can never exceed $\frac{1}{\rho}$. This makes risk estimates quite “lumpy” and ill-suited to calibration of high safety thresholds. As Fender and Kiff (2004) note, BET is particularly prone to understate the risk on senior tranches of portfolios with low diversity scores.
Gaussian (normal) approximations are far more popular. By allowing continuous returns and making all risks scale with standard deviation, they avoid the lumpiness of BET and make thresholds easier to calculate. Adding to the attraction, a multivariate normal distribution with a single correlation parameter $\rho$ has a neat interpretation. For every unit of default risk, a common component contributes $\rho$ and an independent component contributes $\sqrt{1-\rho^2}$. That’s a convenient way to think about risk.

However, the convenience comes at a steep price. It can’t handle a crisis, when the common factor surges in intensity. This became glaringly evident when the mortgage bubble burst. Gaussian models denied that senior CDO tranches were seriously threatened, even as the market imploded.

The implausibility of normal distributions shouldn’t surprise. They can’t fit high odds of zero losses without also assigning high odds to negative losses. They presume symmetry around the mean. They tie all risk judgments to standard deviations.

Nevertheless, the normal distribution has abnormal appeal. It’s the most widely used distribution in finance, often with good reason. Force of habit encourages using it even with bad reason. Few risk analysts grasp the central limitations of the Central Limit Theorem: how much it depends on independent components, and how slowly it converges in the tails.

Fat-Tailed Approximations

Single-parameter approximations needn’t be thin-tailed approximations. The exponential distribution is the best-known example. Moody’s Correlated Binomial (CorBin) distribution, developed by Witt (2004), provides another.

CorBin assumes that the conditional correlation when all other observed assets have defaulted equals the ordinary correlation. This provides a recursive formula for calculating the right edge of the Osband triangle, where everything defaults. While the probability of complete default shrinks as the portfolio expands, it stays bounded away from zero. The other entries get calculated through successive subtractions.

CorBin can match the first two moments of any BET distribution. Yet it is much fatter-tailed. To choose between CorBin and BET scientifically, one needs to consider higher moments or try to estimate the outer tail risk directly.
I have found no evidence that Moody’s made such tests. Still, it shelved CorBin quickly. During the credit bubble, other rating agencies avoided fat-tailed distributions, too. While some of their quants warned, the business side overruled. Acknowledging fat tails would have made it harder to award lucrative top ratings.

Apart from money there was little excuse. The rating agencies knew from experience that portfolio risks tended to be fat-tailed. If they found CorBin too complex, they could have opted for the tractable CreditRisk + methodology, published by Credit Suisse First Boston (Wilde 1997).

CreditRisk + acknowledges the instability of default risk and the merit of mixed binomial models. It proceeds to approximate a low-mean binomial risk as Poisson and a low-mean beta mixing distribution as gamma. The resulting Poisson-gamma mixture is known as negative binomial (NegBin) and is very easy to work with.

NegBin tails decay at rates that converge to constant fractions. In contrast, binomial and normal tails decay at accelerating rates. Hence, NegBin distributions tend to be fatter-tailed than binomial or normal.

To make this more concrete, let’s suppose we have a portfolio of 100 exchangeable credits with a mean default risk of 2% and correlation 5%.

Figure 9.1
Default Risks for Correlated Debt Portfolios
Each of the four models generates probability distribution for defaults. I chart them all in Figure 9.1 in log terms. Only three lines are visible because NegBin and CorBin coincide on the scale shown. The jaggedness in BET reflects lumpiness that my linear interpolations couldn’t smooth out.

Charts for other plausible examples look broadly similar. Apart from twists near the origin, the logarithm of probability decays quadratically for Gaussian and near-linearly for NegBin and CorBin. BET tails are fatter than Gaussian tails but thinner than NegBin or CorBin tails.

How much fatter or thinner? That depends on the part of the tail we’re looking at, as well as the default risk and correlation it’s attached to. In the range we’re interested in, the differences are big even at the 95% confidence level and widen as confidence gets more demanding. Table 9.1 presents a table for the previous example.

Although CorBin tightened confidence standards, it was introduced during a credit boom, when assessments of mean risk were dropping. The correlation confused a journalist for Bloomberg (Smith 2008), who made CorBin a symbol of rating agency laxness. Spurred by the article, a Congressional hearing demanded that Moody’s executives defend themselves on why they had used the CorBin model at all. Unfortunately, confusion pervades public discussion about the reform of financial risk analysis.

### False Confidence

A negative binomial distribution is the best simple way to model risky debt portfolios. It is based on a gamma distribution for default risk, which

<table>
<thead>
<tr>
<th>Confidence</th>
<th>Gaussian</th>
<th>BET</th>
<th>CorBin</th>
<th>NegBin</th>
</tr>
</thead>
<tbody>
<tr>
<td>95%</td>
<td>7</td>
<td>6</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>99%</td>
<td>9</td>
<td>11</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>99.9%</td>
<td>12</td>
<td>17</td>
<td>26</td>
<td>27</td>
</tr>
<tr>
<td>99.99%</td>
<td>14</td>
<td>21</td>
<td>36</td>
<td>38</td>
</tr>
</tbody>
</table>
provides both a natural way to model uncertainty in default risk and a long-term equilibrium for actual risk. A correlated binomial distribution offers similar predictions, although it is far less tractable. The main BET and Gaussian alternatives are implausible and unduly rigid.

Shifting to NegBin or CorBin would expose much false confidence in our credit ratings. Imagine a portfolio derivative that pays 1 if \( k \) or fewer defaults occur and 0 otherwise. Let’s call it a \( k \)-CDO and use the historical average default rates reported in Chapter 8 to assign it a credit grade.

In the example above, a 10-CDO would get rated as low triple-B by Gaussian, double-B by BET, and single-B by NegBin. A 12-CDO would get rated as double-A by Gaussian, a low triple-B by BET, and a low double-B by NegBin. A 14-CDO would get rated as triple-A by Gaussian, triple-B by BET, and double-B by NegBin.

These are huge differences. If rating agencies had applied stricter methodologies, a lot fewer claims on debt portfolios would have been rated investment grade, much less triple-A. It would have been much harder to gain leverage by arbitraging the loopholes in Basel II. With less leverage, housing prices wouldn’t have bubbled up so much or come crashing down so hard. Never in world history have toxic statistics stirred more false hopes or rude awakenings.

NegBin models also have dramatic implications for bank regulation. Recall from Chapter 6 that Basel II tied bank capital to estimated buffers against excess losses. Since ordinary reserves are supposed to cover expected losses, and since 99.9% confidence is a fashionable definition of regulatory safety, the buffer calculation can be stylized as the “99.9% threshold less the mean.”

For a Gaussian distribution, the buffer is always 3.09 standard deviations and will scale approximately with the square root of the mean risk. For NegBin, the buffer will be higher but scale less rapidly with the mean. We saw that in Chapter 5 with beta and gamma distributions: low means are associated with high uncertainty and a long tail. As NegBins are gamma-Poisson mixtures, the same reasoning extends to them.

Figure 9.2 compares the size of Gaussian and NegBin buffers at 99.9% confidence. As mean default rises from 2 bps to 512 bps, the Gaussian buffer increases from less than a quarter of the recommended NegBin buffer to more than half. While the chart below sticks with 100 assets and 5% correlation, experiments with other values generate similar pictures.

Whereas a Gaussian or binomial approach advises nearly ten times the buffer for a 500 bps risk as for a 5 bps risk, NegBin advises less than five
times the buffer. Since most of Basel II presumes binomial or Gaussian dependence, one quick patch would halve the gradient between strong and weak credits. Granted, no quick patch can fix Basel II. Still, I want to emphasize that thoughtful uncertainty can generate regulatory advice that is at least as clear as the feigned certainty underlying Basel II.

Copulas

Over the past decade it has become fashionable to use copulas to model dependence. Copulas describe the joint distribution once each individual asset is transformed to a uniform random draw of real numbers between 0 and 1. That transformation is sufficiently unnatural for discrete default risk that it is hard to convey good intuition for copula application to CDOs. Ironically, that appears to be one of the main attractions of copulas.

Think of it this way. By distinguishing clearly between conditionally independent risks and common drivers, de Finetti’s Theorem begs investigation into the form of the mixing distribution and the uncertainty of estimation. Copulas obscure that with complex formulation. Most critics feel too intimidated to poke around.
Li (2000) introduced the Gaussian copula to finance. It quickly became an industry standard thanks to its one-parameter depiction of common and independent components in default intensity. However, as we have seen, a fixed Gaussian distribution can’t handle crisis.

This is now widely recognized. In an article that captured the public eye, Salmon (2009) described Li’s model as “the formula that killed Wall Street.” Patterson’s (2010) best seller casts the blame more broadly on “a new breed of math whizzes.”

Brigo, Pallavicini, and Torresetti (2010) defend the new breed and its tool kit. They cite numerous efforts to refine Li’s model and make copulas more flexible. They present a tractable framework for pricing credit derivatives in line with market beliefs. They also note that models should not be held responsible for their misuse.

Both sets of arguments have merits. On the one hand, quants unwittingly fed the exaggeration of credit derivatives safety. On the other hand, better modeling of credit derivatives would doubtless have warned sooner of crisis and thereby tempered some extremes. It will doubtless temper extremes going forward.

For a historical parallel, consider the product known as portfolio insurance. It is supposed to neutralize the overall market exposure or “beta” of a given basket of securities. Usually these securities are equities, and the market exposure is summarized in an index like the S&P 500.

The classic Black-Scholes (1973) model of options pricing presumes that any derivative’s risk can be instantaneously measured and hedged out using a specified fraction or “delta” of the underlying asset. By equating delta with beta, it suggests a potentially automatic way to provide portfolio insurance. By the mid-1980s, few major trading houses were not using Black-Scholes models. Portfolio insurance became one of their most lucrative products.

Portfolio insurance worked well until it was most needed, during the 1987 crash. Prices plunged too fast for the hedging programs to keep up. Their attempt to keep up aggravated the crash, because the lower prices fell the more the programs sold in order to neutralize their net delta exposure.

After the crash, there were heated calls to ban program trading in options or to restrict derivatives trading more broadly. With time the heat faded. Soon program trading and derivatives had regrown even bigger than before, with even more sophisticated models.

However, two notions were discredited. One was that market prices followed steady Gaussian distributions as assumed by Black-Scholes. The other was that portfolio insurance could costlessly lock in returns.
As the mortgage wreckage clears, the credit derivatives market will likely regrow even bigger than before. Copulas or related techniques for analyzing portfolio dependence will become even more vital to modeling. Hopefully we’ll pay more attention to common driving factors.

However, credit markets will need to abandon two illusions. One is that credit tails are thin enough to be analyzed using Gaussian or binomial distributions. The other is that quant chefs can easily convert risky debt portfolios into an abundance of nearly risk-free derivatives.

Securitization transforms the tips of sow’s ears into silk purses. Insecuritization tries to smuggle in the rest of the sow. To better distinguish the two, we’ll need better thinking, higher rating agency standards, and wiser regulation. For more insight on current deficiencies, see SEC (2008) and Witt (2010b).

Let me caution that there’s no panacea. In practice, underestimation of the variance or correlation often dwarfs the choice of model. That was clearly the case in 2005/6. The low defaults and calm markets experienced in the previous few years cut nearly all models’ estimates of tail risks. The main exceptions explicitly focused on identifying market bubbles—e.g., Zhou and Sornette (2006).

“Portfolio risk is tricky,” said Prometheus. “The aggregate tail risk often behaves far differently from the average of the individual tail risks. Almost anything can happen, given enough assets.”

“Yes, and knowing the correlations won’t suffice to pin the tail risks down,” said Pandora. “It’s shocking how many skilled risk analysts get this wrong.”

“If they’re getting it wrong, they’ve got the wrong skills. Too much normality drilled into their brains. They should study Osband’s risk triangles.”

“Perhaps the diversity of portfolio tails is too daunting. Risk analysts have to give practical guidance. They can’t just shrug their shoulders. Normality is a convenient crutch.”

“It’s the wrong crutch,” said Prometheus. “Normality in highly rated debt portfolios is highly abnormal. Better to choose a distribution with much slower-decaying tails. Negative binomial, correlated binomial, gamma, lognormal—take your pick. They all decay approximately exponentially where it matters. But not normal, and not standard binomial, either.”
“I would urge more focus on common risk drivers. People lose sight of the iceberg for the ice cubes.”

“Or their absence. When analysts looked back 60 years and couldn’t find a U.S.-wide mortgage crisis, it was easy to assume there wouldn’t be one.”

“Strange, the price explosions were right under their noses,” said Pandora. “They should have looked for the massive leverage supporting it,”

“Good point. But our approaches aren’t contradictory. Fluctuations in common risk drivers help explain slow-decaying tails. In using normal and standard binomial distributions, regulators implicitly presume a fixed common risk.”

“I see. With slower decaying tail risks, it becomes harder to portray senior claims on weak portfolios as strong. If uncertainty gets graded as well, obscure multilayered securitizations will lose most of their appeal. Securitizers will have less incentive to create them. Credit graders will have less incentive to turn a blind eye to them.”