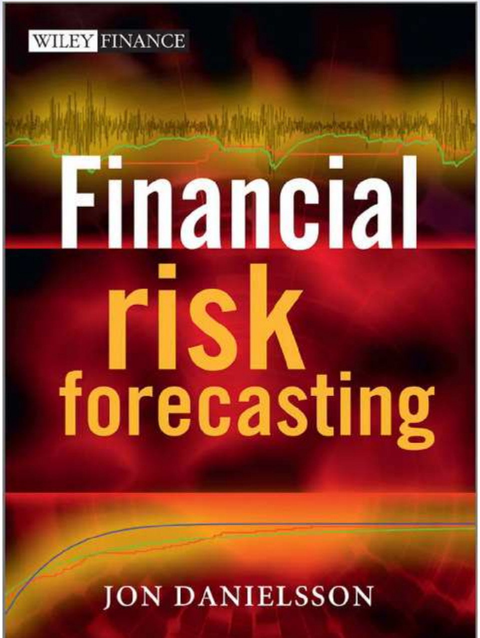
# Financial Risk Forecasting
# Chapter 4
# Risk Measures

Jon Danielsson ©2023
London School of Economics

To accompany
*Financial Risk Forecasting*
www.financialriskforecasting.com
Published by Wiley 2011

Version 8.0, August 2023


WILEY FINANCE
Financial risk forecasting
JON DANIELSSON

# Risk Measures

# The Focus of This Chapter

- Defining and measuring risk
  - Volatility
  - Value–at–Risk (VaR)
  - Expected Shortfall (ES)
- Theoretical issues
- Holding periods
- Scaling and the square root of time

# Notation

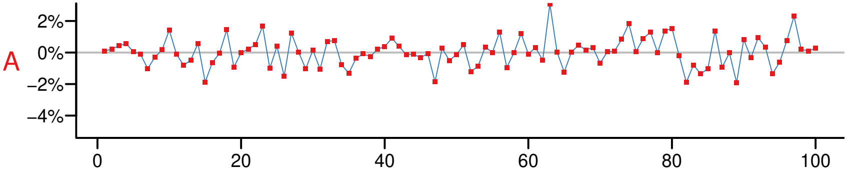
| | |
|---|---|
| $\rho$ | Probability |
| $Q$ | Profit and loss |
| $q$ | Observed profit and loss |
| $w$ | Vector of portfolio weights |
| $X$ and $Y$ | Refer to two different assets |
| $\varphi(.)$ | Risk measure |
| $\vartheta$ | Portfolio value |

# General Definition

- No universal definition of what constitutes risk
- On a very *general level*, financial risk could be defined as *"the chance of losing a part or all of an investment"*
- Large number of such statements could equally be made, many of which would be contradictory

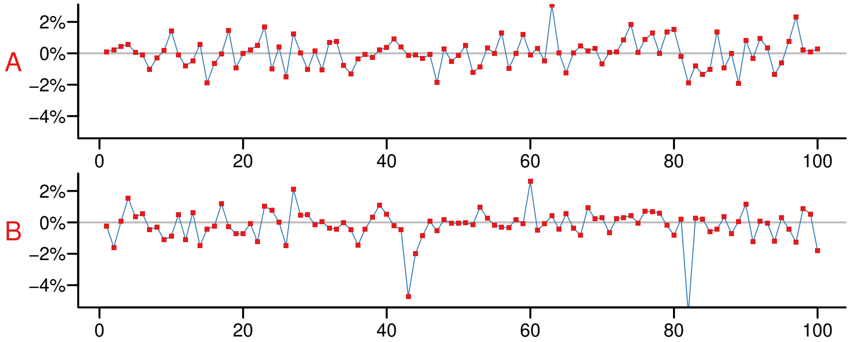
# Which Asset Do You Prefer?

## All three assets have volatility one and mean zero



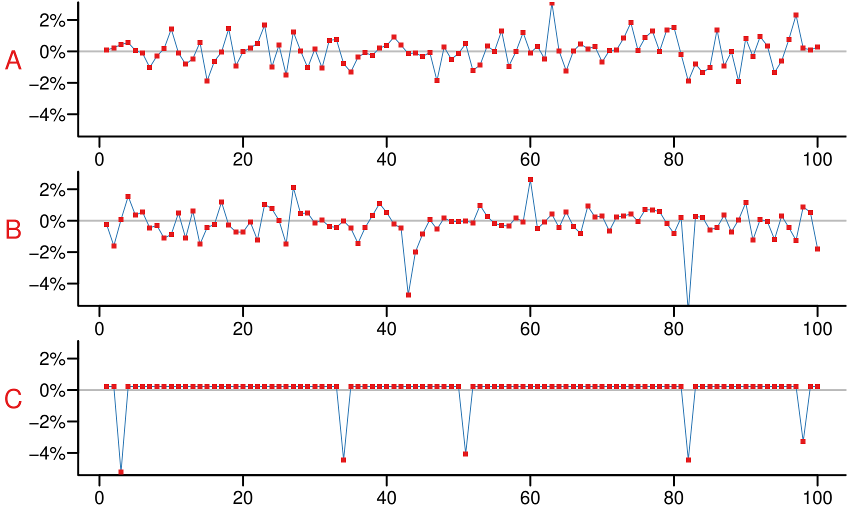
# Which Asset Do You Prefer?

## All three assets have volatility one and mean zero



# Which Asset Do You Prefer?

## All three assets have volatility one and mean zero



# Which Asset Do You Prefer? (cont.)

- *Standard mean-variance analysis* indicates that all three assets are equally risky and preferable
- Since we have the same mean

$$\mathrm{E}(A) = \mathrm{E}(B) = \mathrm{E}(C) = 0$$

- And the same volatility

$$\sigma_A = \sigma_B = \sigma_C = 1$$

- If one uses mean-variance analysis, one is indifferent between all three

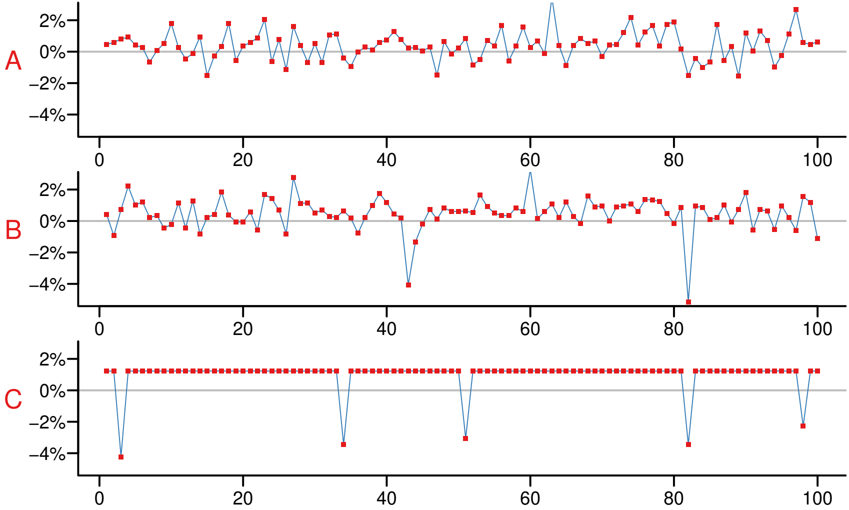
# Which Asset Is Preferred by MV?

- If, however, one asks anybody which asset they would prefer, they probably would have a personal preference for one
- A trader investing someone's else money, is likely to prefer $C$. Why?
- The most popular speculative financial asset in the world is inverted $C$ – lottery tickets

# So What Happened?

- The model – mean variance – comes with a set of assumptions that are beneficial for creating a practical investment model
- But at the same time inconsistent with people's risk preference
- Oftentimes that is not important
- But sometimes it is
- Recall "all models are wrong, some models are useful"

# Mean $A = {}^1\!/_3$, mean $B = {}^2\!/_3$, mean $C = 1$



# Now the Preferences Change

- Because all three asset still have the same volatility
- But the mean of $C$ is highest, mean variance tells us to pick $C$
- However, most people would pick one of these using some criteria private to them

# Three Investment Choices

- Suppose Yiying, Alvaro, and Mary all have the same amount of money to invest
- All have access to the same investment technology (All have taken this course)
- All are contemplating putting $1 million into Amazon

**Yiying** is a day trader, aiming to buy and sell within a week

**Alvaro** is a fund manager, and his bonuses depend on quarterly performance

**Mary** is 22 years old, planning to retire in 40 years and expects to die in 70. She is saving for her pension that needs to be available when she's 90 years old and far from able to manage her own money

# Their Choices

- While faced with same technology, their preferences are different
- And consequently they will evaluate the three investment choices $A$, $B$, $C$ differently
- Mary will unequivocally pick $C$
- It's quite possible Yiying does as well because she would not be a day trader if she didn't like risk
- Alvaro would pick $A$ or $B$, probably the latter

# Which Asset Is “Better”?

- There is no obvious way to discriminate between the assets
- One can try to model the *underlying distribution* of market prices and returns of assets, but it is generally *unknown*.
  - Can identify by maximum likelihood methods
  - Or test the distribution against other other distributions by using methods such as the Kolmogorov-Smirnov test
- Practically, it is impossible to accurately identify the distribution of financial returns

# Risk Is a Latent Variable

- Financial risk cannot be measured directly
- Risk has to be *inferred* from the behaviour of observed market prices
    - For example, at the end of a trading day, the return of the day is known while the risk is unknown

# Risk Measure and Risk Measurement

**Risk measure:** a mathematical concept of risk

**Risk measurement:** a number that captures risk, obtained by applying data to a risk measure

# Volatility

- Volatility is the *standard deviation of returns*
- Main measure of risk in most financial analysis
- It is a *sufficient* measure of risk when returns are *normally distributed*
    - For this reason, in mean-variance analysis, the efficient frontier shows the best investment decision
    - If returns are not normally distributed, solutions on the efficient frontier may be inefficient

# Volatility (cont.)

- The assumption of *normality of return is violated* for most, if not all financial returns
  - See Chapter 1 on the non-normality of returns
- For most applications in financial risk, volatility is likely to systematically *underestimate risk*

# Value–at–Risk (VaR)

# History

- Until 1994, the only risk measure was volatility
- Then the JP Morgan bank proposed a risk measure called *Value-at-Risk* (VaR) and a method to measure it, called Riskmetrics, what we now call EWMA
- Why would JP Morgan do that – to be able to reduce its level of capital
- It used to be called the $4^{15}$ report because it was created because the chairman of the bank wanted a single measurement of the bank's risk in time for the treasury meeting at $4^{15}$

# Value-at-Risk

**Definition: Value-at-Risk.** The loss on a trading portfolio such that there is a probability $\rho$ of losses equaling or exceeding VaR in a given trading period and a $(1-\rho)$ probability of losses being lower than the VaR.

- The most common risk measure after volatility
- It is *distribution independent* in theory, but not in practice

# Quantiles and Profit and Loss ($Q$)

- VaR is a *quantile* on the distribution of profit and loss ($Q$)
- In the case of holding *one unit* of an asset, we have

$$Q_t = P_t - P_{t-1}$$

- More generally, if the portfolio value is $\vartheta$:

$$Q_t = \vartheta Y_t = \vartheta \frac{P_t - P_{t-1}}{P_{t-1}}$$

- That is, the $Q$ is the portfolio value ($\vartheta$) multiplied by the returns

# VaR and The Profit and Loss Density

- The density of $Q$ is denoted by $f_q(.)$, then VaR is given by:

$$\mathbb{P}[Q \leq -\text{VaR}(\rho)] = \rho$$

or,

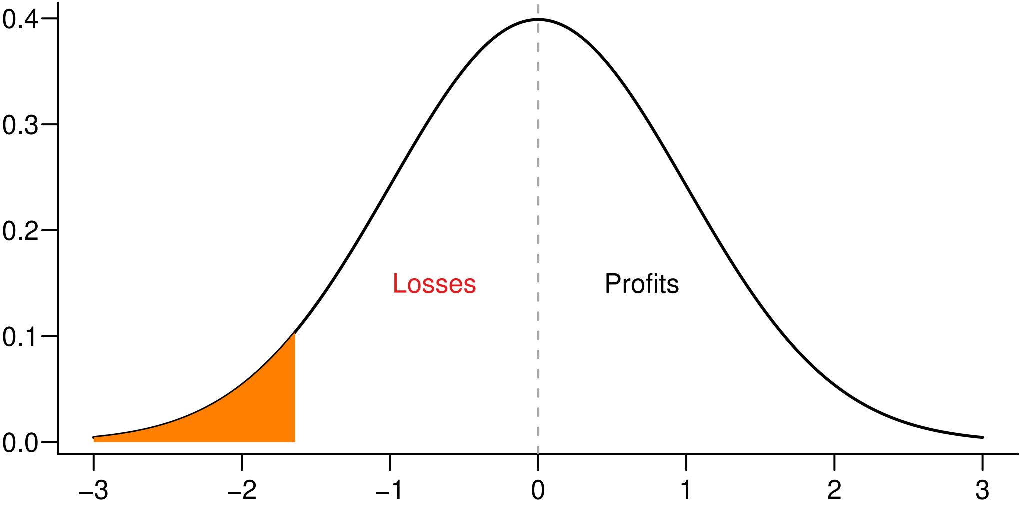
$$\rho = \int_{-\infty}^{-\text{VaR}(\rho)} f_q(x)dx$$

- We usually write it as $\text{VaR}(\rho)$ or $\text{VaR}^{100\times\rho\%}$
  - For example, $\text{VaR}(0.05)$ or $\text{VaR}^{5\%}$

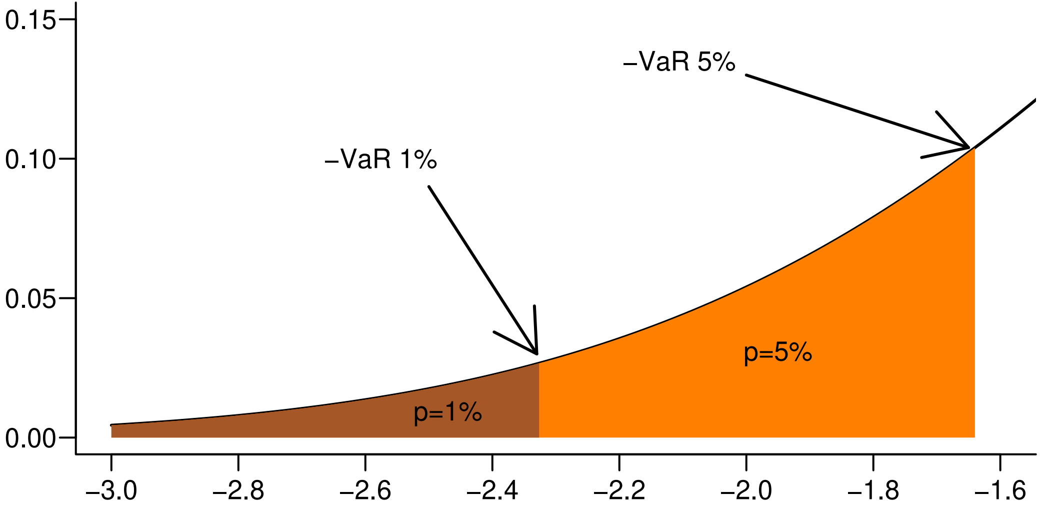
# Is VaR a Negative or Positive Number?

- VaR can be stated as a negative or positive number
- Equivalently, probabilities can be stated as close to one or close to zero – for example, VaR(0.95) or VaR(0.05)
- We take the more common approach of referring to VaR as a *positive number* using *low-probability terminology* (for example, 5%)
- However, almost nobody is able to be consistent, (not me certainly) and I will use 5% and 95% VaR interchangeably

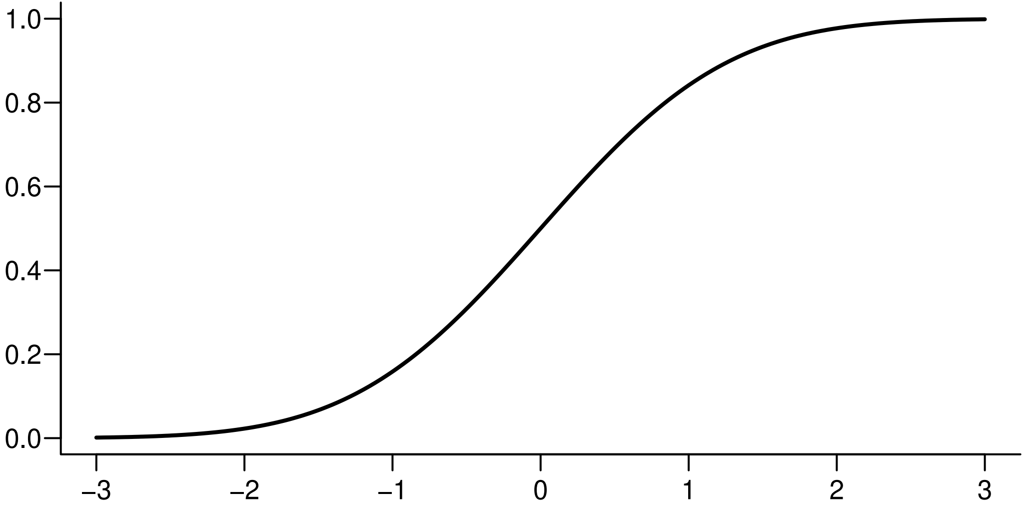
# VaR Graphically: Density of profit and loss



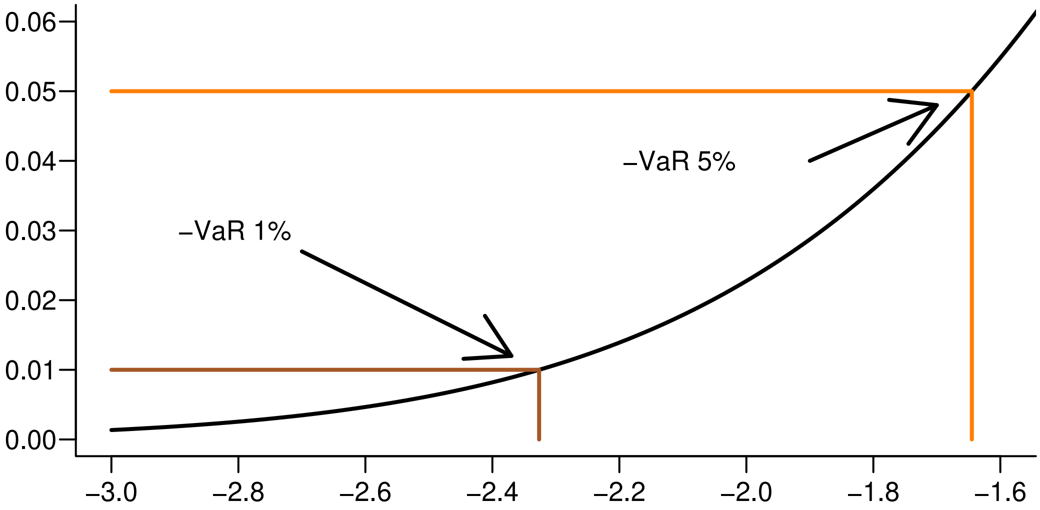
# VaR Graphically: Density zoomed in



# VaR Graphically: Cumulative distribution of profit and loss



# VaR Graphically: Cumulative distribution zoomed in



# Example

- The commodities' trading book is worth £1 billion and daily $\text{VaR}^{1\%}$=£10 million
- This means we expect to loose £10 million or more once every 100 days, or about once every 5 months

# The Three Steps in VaR Calculations

1. The *probability* of losses exceeding VaR, $\rho$
2. The *holding period*, the time period over which losses may occur
3. The *probability distribution* of the P/L of the portfolio

# Which Probability Should We Use?

- VaR levels of 1% to 5% are very common in practice
- Regulators (Basel II) demand 1%
- But less extreme numbers, such as 10%, are often used in risk management on the trading floor
- More extreme lower numbers, such as 0.1%, may be used for applications like economic capital, survival analysis, or long-run risk analysis for pension funds

# Holding Period

- The *holding period* is the time period over which losses may occur
  - It is usually one day
  - Can be minutes or hours
  - Or several days, but it does not make sense to use more than two weeks, and even that is on the high side
- Holding periods can vary depending on different circumstances
- Many proprietary trading desks focus on intraday VaR
- For institutional investors and nonfinancial corporations, it is more realistic to use longer holding periods

# Probability Distribution of Profit and Loss (P/L)

- The identification of the probability distribution is difficult
- The standard practice is to estimate the distribution by using past observations and a statistical model
- We will use EWMA and GARCH later

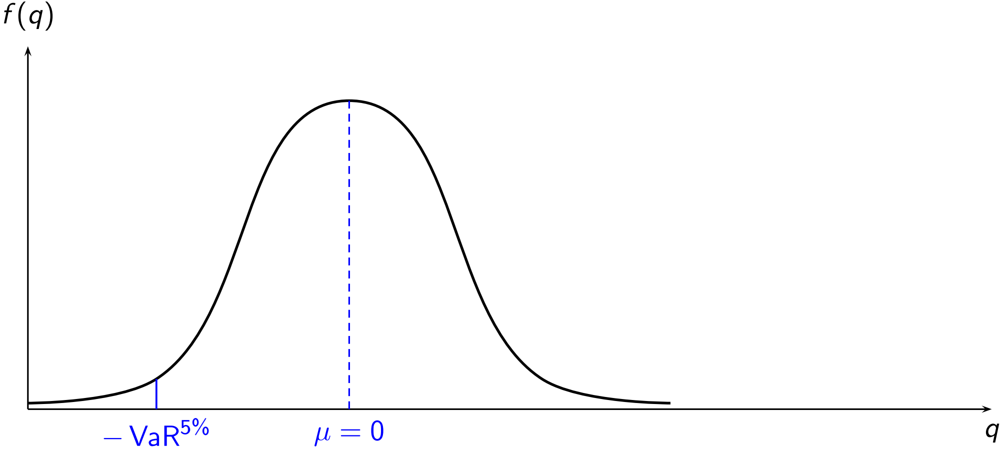
# VaR and Normality

- VaR *does not* imply normality of returns and we can use any distribution in calculating VaR
- Even if I am often surprised by people who assume it does
- However, the most common distribution assumption for returns in the calculation of VaR is *conditional* normality
  - In this case, volatility provides the same information as VaR. Why?

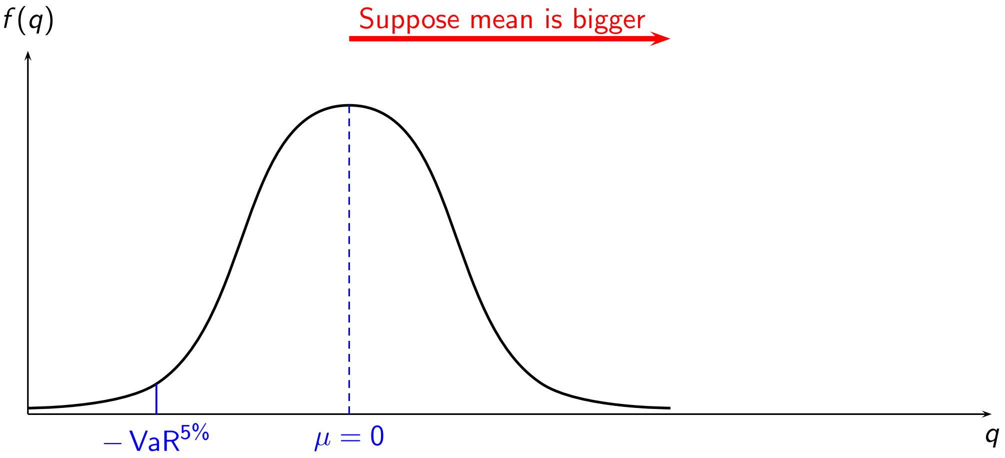
# Sign of VaR (Part 1)

- If the mean of the density of P/L is sufficiently large, the probability $\rho$ quantile (the VaR), might easily end up on the other side of zero
- This means that the *relevant losses have become profits*
- In such situations, we either specify a more extreme $\rho$ or use a different measure of risk

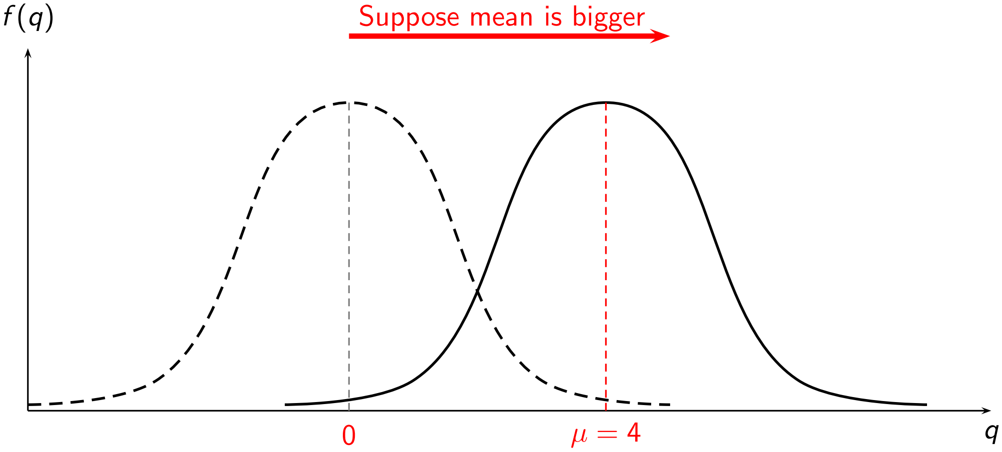
# Sign of VaR (Part 2)



# Sign of VaR (Part 2)

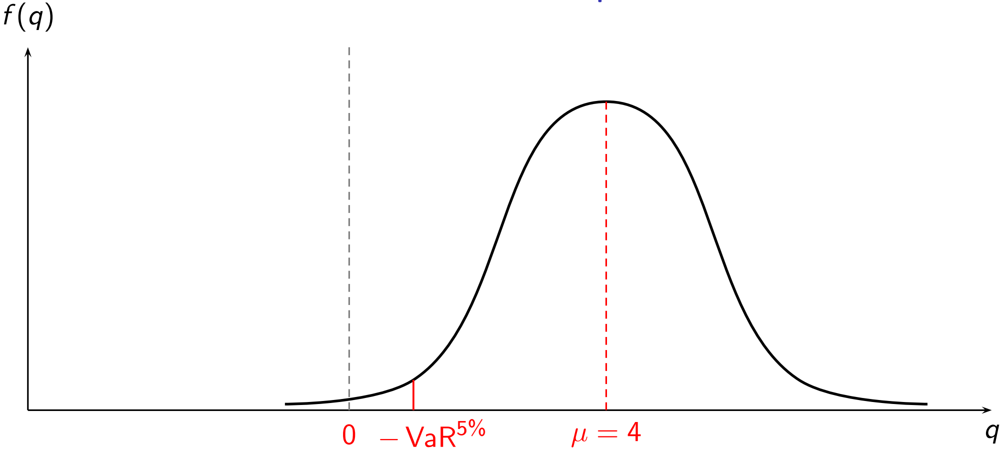


# Sign of VaR (Part 2)



# Sign of VaR (Part 2)

## Note that -VaR is now positive



# Issues in Applying VaR

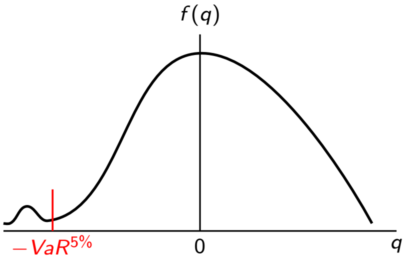
# Main Issues in the Implementation

1. VaR is only a quantile on the profit/loss distribution
2. VaR is not a coherent risk measure
3. VaR is easy to manipulate

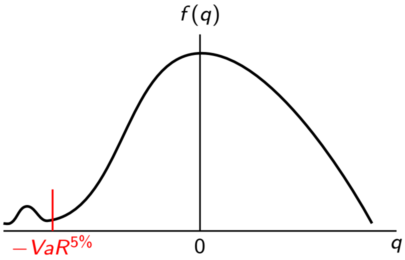
# VaR Is Only a Quantile

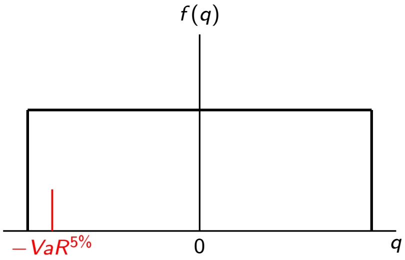
- VaR gives the *"best of worst case scenarios"* and, as such, it inevitably underestimates the potential losses associated with a probability level
- That is, VaR($\rho$) is incapable of capturing the risk of extreme movements that have a probability of less than $\rho$
- If VaR=\$1,000, are potential losses \$1001 or \$10000000?
- The shape of the tail before and after VaR need not have any bearing on the actual VaR number

# VaR in Unusual Cases

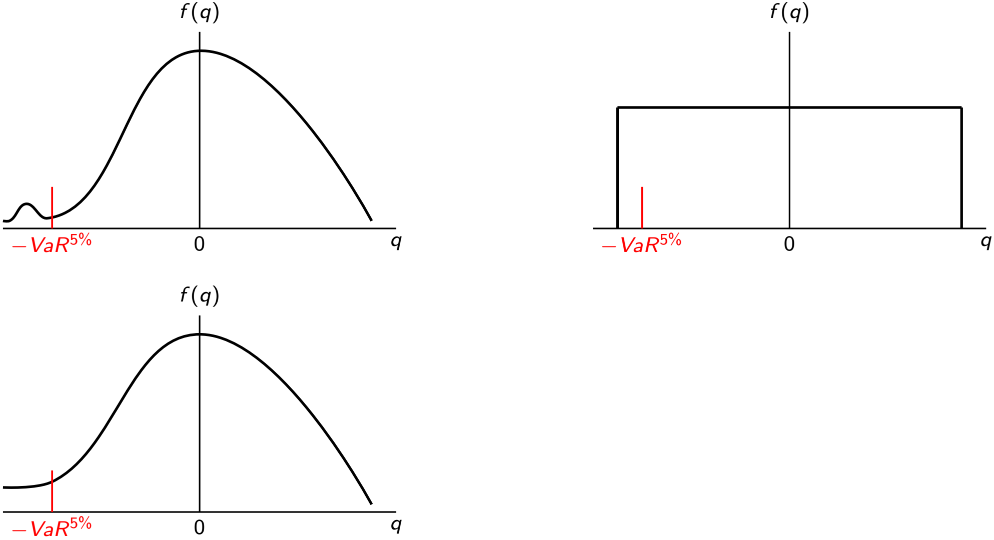


# VaR in Unusual Cases

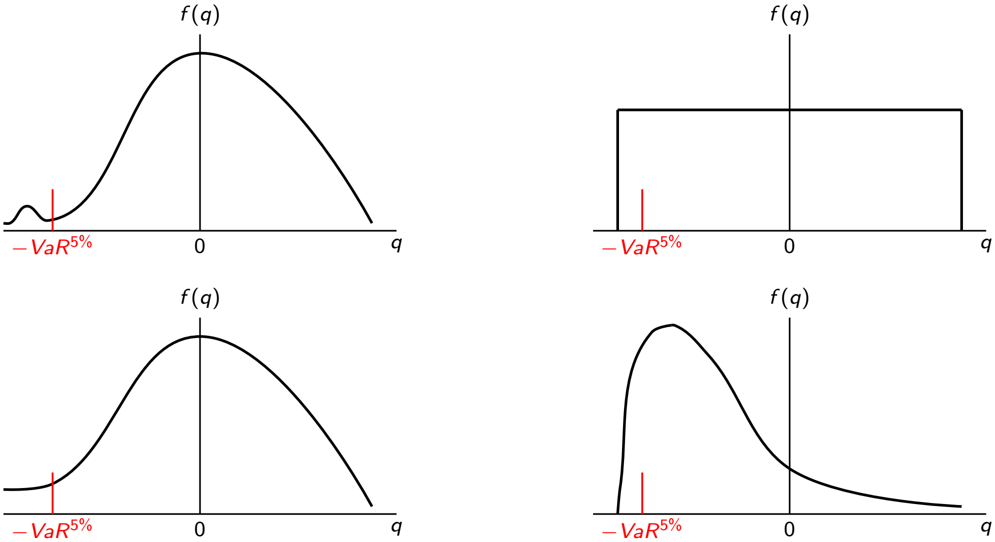


# VaR in Unusual Cases



# VaR in Unusual Cases



# Ideal Properties of Risk Measures

# Ideal Properties of a Risk Measure

- The ideal properties of any financial risk measure were proposed by Artzner et al. (1999)
- To them, coherence is ideal
- Other authors have added more “ideal conditions”
- While others have dismissed the importance of some of these

# Coherence

- Suppose we have two assets, $X$ and $Y$
- Denote some arbitrary risk measure by $\varphi(\cdot)$; It could be volatility, VaR, or something else
- $\varphi(\cdot)$ is then some function that maps some observations of an asset, like $X$, onto a risk measurement
- Further define some arbitrary constant $c$
- We say that $\varphi(\cdot)$ is a coherent risk measure if it satisfies the following four axioms:
  1. Monotonicity
  2. Translation invariance
  3. Positive homogeneity
  4. Subadditivity

# Monotonicity

- If

$$X \leq Y$$

- And

$$\varphi(X) \geq \varphi(Y)$$

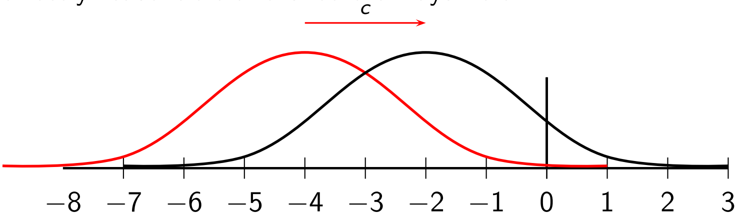
- Then risk measure $\varphi$ satisfies monotonicity
- What this means is that if outcomes for asset $X$ are always more negative than outcomes for $Y$
  - Suppose $X$ and $Y$ are daily returns on AMZN and GOOG, and the returns on GOOG are always higher than for AMZN, then GOOG risk is lower than that of AMZN
- Then the risk of $Y$ should never exceed the risk of $X$
- This is perfectly reasonable and should always hold

# Translation Invariance

- If

$$\varphi(X+c) = \varphi(X)-c$$

- Then risk measure $\varphi$ satisfies translation invariance
- In other words, if we add a positive constant to the returns of AMZN, then the risk will go down by that constant
- This is perfectly reasonable and should always hold



# Positive Homogeneity

- If $c > 0$ and

$$\varphi(cX) = c\varphi(X)$$

- Then risk measure $\varphi$ satisfies positive homogeneity
- Positive homogeneity means risk is *directly proportional* to the value of the portfolio
- For example, suppose a portfolio is worth \$1,000 with risk \$10, then doubling the portfolio size to \$2,000 will double the risk to \$20

# Positive Homogeneity (Cont.)

- As relative shareholdings increase, the risk may increase more rapidly than the portfolio size
- In this case, *positive homogeneity is violated*:

$$\varphi(cX) > c\varphi(X)$$

- This is because when we are trying to sell, the price of the stock falls, therefore the eventual selling price is lower than the initial market price
  - See chapter 10, Endogenous risk

# Subadditivity

- If
$$\varphi(X+Y) \leq \varphi(X) + \varphi(Y)$$
- Then risk measure $\varphi$ satisfies *subadditivity*
- Subadditivity means a portfolio of assets is measured as less risky than the sum of the risks of individual assets
- That is, *diversification reduces risk*

# Volatility Is Subadditive

- Recall how portfolio variance is calculated when we have two assets
- $X$ and $Y$, with volatilities $\sigma_X$ and $\sigma_Y$, respectively, correlation coefficient $\kappa$ and portfolio weights $w_X$ and $w_Y$
- The portfolio variance is:

$$\sigma^2_{\text{portfolio}} = w_X^2\sigma_X^2 + w_Y^2\sigma_Y^2 + 2w_X w_Y \kappa \sigma_X \sigma_Y$$

# Volatility Is Subadditive (Cont.)

- Rewriting, we get

$$\sigma^2_{\text{portfolio}} = (w_X\sigma_X + w_Y\sigma_Y)^2 - 2w_X w_Y(1-\kappa)\sigma_X\sigma_Y$$

  where the last term is positive
  - $W_X, W_Y \geq 0$, $-1 \leq \kappa \leq 1$
- Volatility is therefore *subadditive* because:

$$\sigma_{\text{portfolio}} \leq w_X\sigma_X + w_Y\sigma_Y$$

# VaR Can Violate Subadditivity

- Asset $X$ has probability of 4.9% of a return of -100, and 95.1% of a return of 0
- Hence we have

$$\text{VaR}^{5\%}(X) = 0$$
$$\text{VaR}^{1\%}(X) = 100$$

# VaR Can Violate Subadditivity (Cont.)

- Consider another asset $Y$, *independent* of $X$ and with the *same distribution* as $X$
- Suppose we hold an equally weighted portfolio of assets X and Y, the 5% VaR of the portfolio is

$$\text{VaR}^{5\%}_{\text{portfolio}} = \text{VaR}^{5\%}(0.5X + 0.5Y) = 50$$

| Outcome | $X$ | $Y$ | $\frac{1}{2}X + \frac{1}{2}Y$ | Probability | Cumulative |
|---|---|---|---|---|---|
| 1 | -100 | -100 | -100 | 0.2% | 0.2% |
| 2 | -100 | 0 | -50 | 4.7% | 4.9% |
| 3 | 0 | -100 | -50 | 4.7% | 9.6% |
| 4 | 0 | 0 | 0 | 90.4% | 100% |

# VaR Can Violate Subadditivity (Cont.)

- In this case, $\mathrm{VaR}^{5\%}$ *violates subadditivity*

$$\mathrm{VaR}^{5\%}_{\text{portfolio}} > 0.5\,\mathrm{VaR}^{5\%}(X) + 0.5\,\mathrm{VaR}^{5\%}(Y) = 0$$

- This is because the probability of a loss (4.9%) of a single asset is slightly smaller than the VaR probability (5%)
- While the portfolio probability is higher than 5%

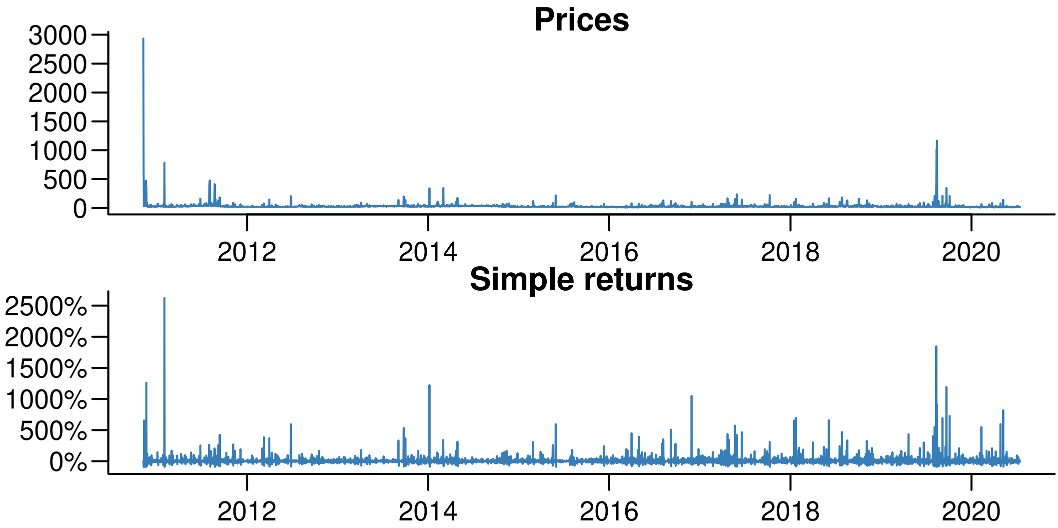
# Does VaR Really Violate Subadditivity?

- VaR is subadditive in the special case of normally distributed returns
- Subadditivity for the VaR is violated when the tails are *super fat*
    - For example, a Student-t where the degrees of freedom are less than one
    - Imagine you go to a buffet restaurant where you suspect one of the dishes might give you food poisoning
    - Then the optimal strategy is only to eat one dish, not to diversify
- Most assets do not have super-fat tails, this includes most equities, exchange rates and commodities

# Does VaR Really Violate Subadditivity? (Cont.)

- VaR of assets that are subject to occasional very large negative returns tends to suffer subadditivity violations, for example
  - Exchange rates in countries that peg their currency but are *subject to occasional devaluations*
  - Electricity prices subject to very extreme price swings (see next Slide)
  - *Junk bonds* where most of the time the bonds deliver a steady positive return
  - Short deep out of the money options

# Houston Power Prices



# Manipulation

# Manipulating VaR

- VaR is easily manipulated, perhaps to reduce the VaR measurement lower without risk management, internal control, or the authorities realising
- There are many ways to do this, for example
    1. Cherry pick assets that make a VaR measure low
    2. Particular derivative trading strategies

# Cherry Pick Assets

- Suppose a trader manages an equity portfolio worth \$100 million
- Where the VaR = \$10 million
- And then her boss says "Can you reduce the VaR without ~~getting caught~~ violating any rules"?
- Easy, ask the computer to search for stocks that give the same expected return, but lower VaR
- You can easily do that with the CRSP stocks

# London Whale

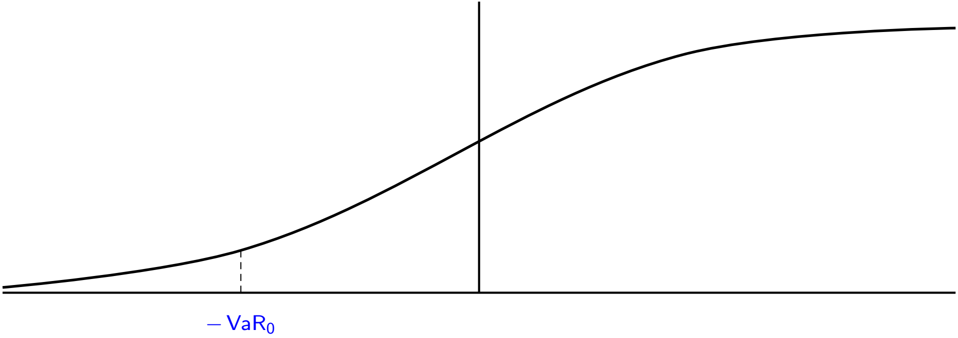
- JP Morgan bank in 2013
- VaR for the "chief investment office" division exceeded \$95 million
- Total target VaR for the entire bank was \$125 million
- Person in charge of the VaR sent an email from a Yahoo account to his colleagues with the subject "Optimising regulatory capital"
- JP Morgan's lost \$5.8 billion
- JP Morgan's quarterly securities filing: "This portfolio has proven to be riskier, more volatile and less effective as an economic hedge than the firm previously believed"

# When All You Need Is a Number

- A friend of mine, Rupert Goodwin, sold risk systems
- One day he went to a bank that had just been audited by the local financial authority
- The regulator came in and asked the risk manager if he used a risk model
- When the risk manager said yes, the regulator ticked off a box and left

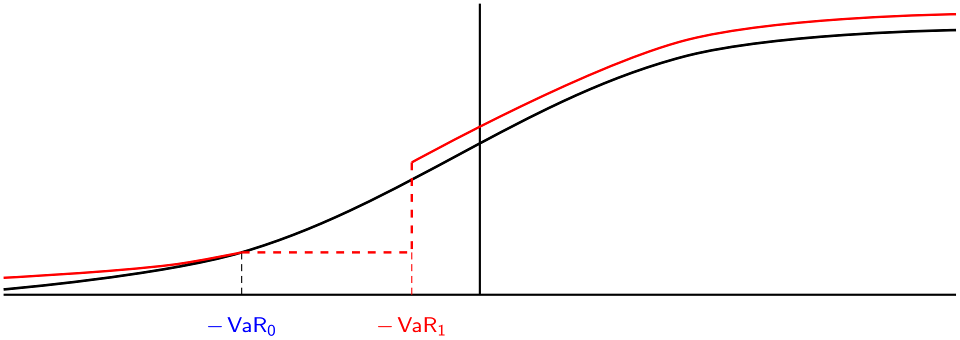
# VaR Manipulation With Derivatives (Part 1)

- Suppose the VaR before any manipulation is $\text{VaR}_0$ and that a bank prefers the VaR to be $\text{VaR}_1$
- Where, $0 < \text{VaR}_1 < \text{VaR}_0$



# VaR Manipulation With Derivatives (Part 2)

- Suppose the VaR before any manipulation is $\text{VaR}_0$ and that a bank prefers the VaR to be $\text{VaR}_1$
- Where, $0 < \text{VaR}_1 < \text{VaR}_0$



# VaR Manipulation With Derivatives (Part 3)

- This can be achieved by
  1. Buying put with a strike above $\text{VaR}_1$
  2. Writing a put option with a strike price below $\text{VaR}_0$
- This will result a lower expected profit
  - The fee received from writing the option is lower than the fee payed from buying the option
- And an increase in downside risk
  - Because it the potential for large losses (makes the tail fatter)
- While this may be an obvious manipulation
- It can be very hard to identify in the real world

# Expected Shortfall (ES)

# Expected Shortfall (ES)

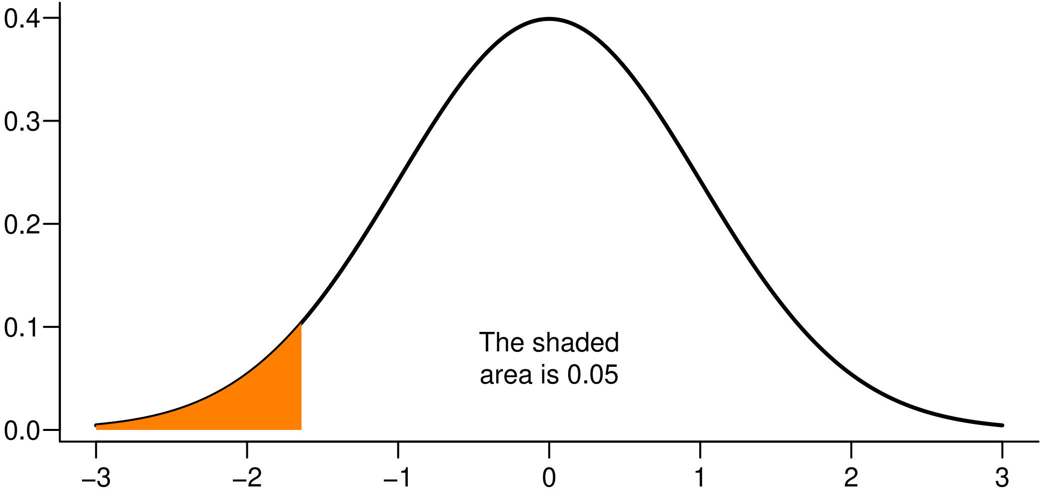
- A large number of other risk measures have been proposed
- The only one to get traction is ES
- It is known by several names, including
  1. ES
  2. Expected tail loss
  3. Tail VaR

**Definition: Expected shortfall.** Expected loss conditional on VaR being violated (that is, expected P/L, $Q$, when it is lower than negative VaR)

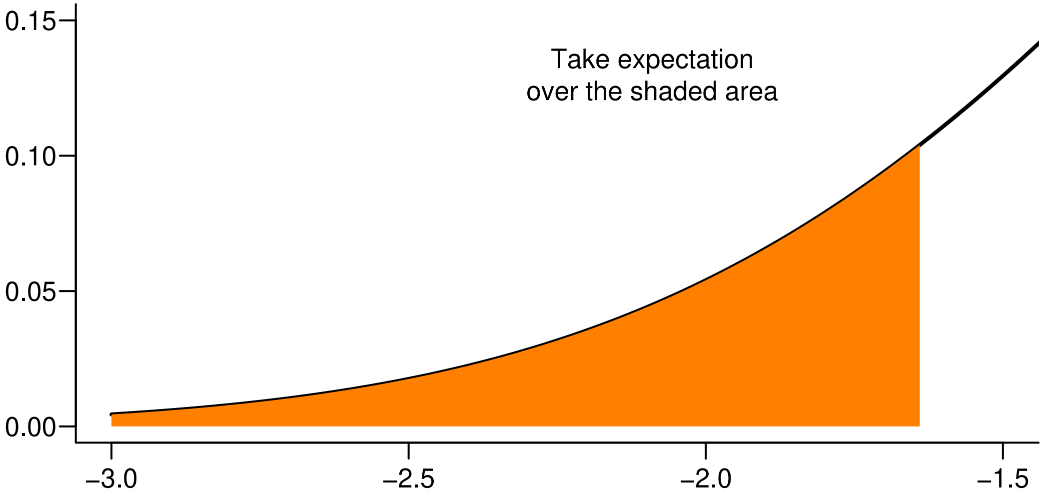
$$\mathrm{ES} = -\mathrm{E}[Q|Q \le -\mathrm{VaR}(\rho)]$$

- ES is an alternative risk measures to VaR which overcomes the problem of subadditivity violation
- It is aware of the shape of the tail distribution while VaR is not

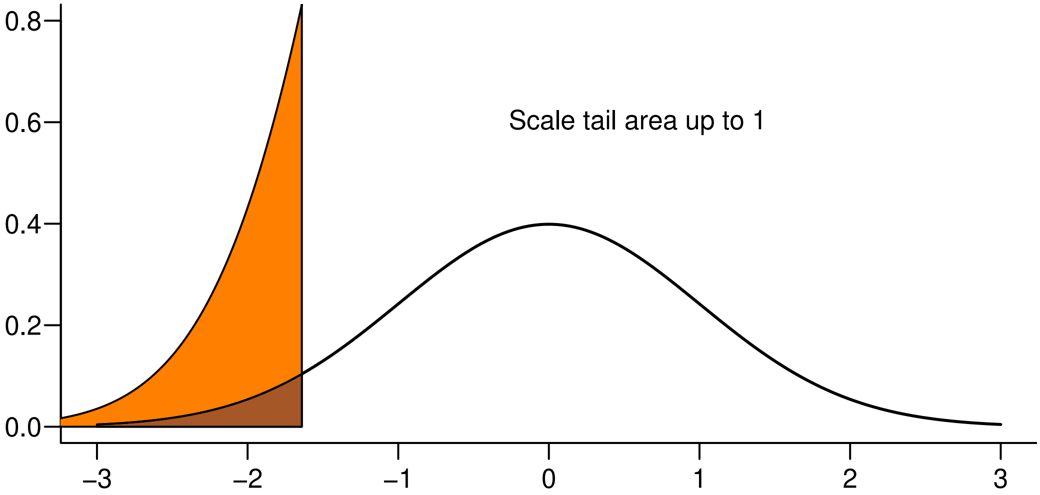
# ES and VaR for Profit/Loss Outcomes: Density of P/L and VaR



# ES and VaR for Profit/Loss Outcomes: Left tail of the density



# ES and VaR for Profit/Loss Outcomes: Blow up the tail



# Expected Shortfall

- The expectation over the VaR (tail) density, $f_{\text{VaR}}(x)$

$$\text{ES} = \int_{-\infty}^{-\text{VaR}(\rho)} x f_{\text{VaR}}(x) dx$$

- Where the VaR density, $f_{VaR}(.)$ is given by:

$$1 = \int_{-\infty}^{-\text{VaR}(\rho)} f_{\text{VaR}}(x) dx = \frac{1}{\rho} \int_{-\infty}^{-\text{VaR}(\rho)} f_q(x) dx$$

- ES is then

$$ES = \frac{1}{\rho} \int_{-\infty}^{-\text{VaR}(\rho)} x f_q(x) dx$$

# Under the Standard Normal Distribution, ...

- If the portfolio value is 1, then:

$$\text{ES} = -\frac{\phi(\Phi^{-1}(p))}{p}$$

where $\phi$ and $\Phi$ are the normal density and distribution, respectively

- VaR and ES for different levels of confidence for a portfolio with a face value of \$1 and normally distributed P/L with mean 0 and volatility 1

| $p$ | 0.5 | 0.1 | 0.05 | 0.025 | 0.01 | 0.001 |
|---|---|---|---|---|---|---|
| VaR | 0 | 1.282 | 1.645 | 1.960 | 2.326 | 3.090 |
| ES | 0.798 | 1.755 | 2.063 | 2.338 | 2.665 | 3.367 |

# Pros and Cons

*Pros*

- ES in *coherent* and VaR is not
- It is *harder to manipulate* ES than VaR

*Cons*

- See next slides

# Measurement

- To calculate ES, we first have to know VaR and then integrate over the tail from VaR to minus infinity
- That means in practice that we need more calculations for ES than VaR
- And that the estimation error multiplies

# Backtesting

- We discuss backtesting in Chapter 8
- Backtesting is a technique for evaluating the quality of a risk forecast model
- As it turns out, it is much harder to backtest ES because it requires estimates of the tail expectation to compare with the ES forecast
- While VaR can be compared to actual market outcomes

# Holding Periods, Scaling, and the Square Root of Time

# Length of Holding Periods

- In practice, *the most common holding period is daily*
- Shorter holding periods are common for risk management on the trading floor
  - Where risk managers use hourly, 20-minute, and even 10 minute, holding periods
  - This is technically difficult because intraday data has complicated diurnal patterns

# Longer Holding Periods

- Holding periods exceeding one day are also demanding
  - The effective date the sample becomes much smaller
  - One could use scaling laws
- Most VaR forecasts require at least a few hundred observations to estimate risk accurately
- For a 10-day holding period, will need at least 3,000 trading days, or about 12 years
- In most cases, data from 12 years ago are fairly useless

# Scaling Laws

- If data comes from a particular stochastic process it may be possible to use VaR estimates at high frequency (for example daily) and *scale them up* to lower frequencies (eg biweekly)
- This would be possible because we know the stochastic process and how it aggregates
- That is not usually the case

# Variances, IID and Square Root of Time Scaling

- Suppose we observe an IID random variable $\{X_t\}$ with variance $\sigma^2$ over time
  - The variance of the sum of two consecutive $X$'s is then:

$$\mathrm{Var}(X_t + X_{t+1}) = \mathrm{Var}(X_t) + \mathrm{Var}(X_{t+1}) = 2\sigma^2$$

  - So the standard deviation (volatility) scales by

$$\sigma_{\mathrm{2days}} = \sqrt{2}\sigma$$

  - And generally over $T$ days

$$\sigma_{T\mathrm{days}} = \sqrt{T}\sigma$$

  - The scaling law for variances is time
  - It is the same for the mean
  - While the scaling law for standard deviations (volatility) is square root of time
  - This holds regardless of the underlying distribution (provided the variance is defined)

# Square Root of Time Scaling

- The square root of time rule applies to volatility when data is IID
- It does not apply when data is not IID, like GARCH
- This rule applies to volatility regardless of the underlying distribution provided that the returns are IID
- For VaR, the square root of time rule only applies when returns are IID *normally* distributed
  - Note we need an additional assumption, normality for the VaR
- It is possible to derive the scaling law for IID fat-tailed data