CARM

Table of Contents

# 2. Probability theory overview

## Distributions

**The Bernoulli distribution**

with :

**The binomial distribution** with

For later use we set for every

Note that –

The **uniform distribution** with :

The **normal or Gaussian distribution** with :

The or **Student distribution** for :

where is the Euler function defined for with by

**The exponential distribution** with :

**The lognormal distribution** with :

**The Pareto distribution** with :

## Mean variance

Definition. Let . (If well-defined) The **mean** of is given by

The **variance** of is defined by

### Mean and variance examples



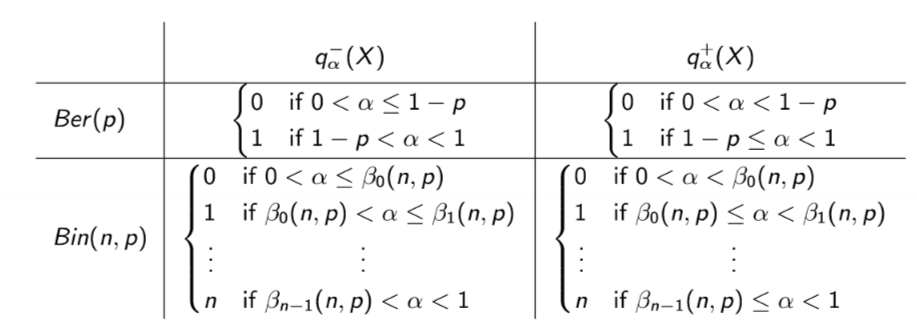
Mean variance examples

## Quantiles

Proposition. For all and the following are equivalent: (1) For all with we have .

Interpretation. The lower and upper -quantiles are distinct only when the distribution function of is flat at the level . Corollary. Let and . If is strictly increasing and continuous in a neighborhood of with then

### Examples of quantiles



## Correlation

Definition. Let and assume neither nor is null. (If well-defined) The correlation between and is given by

**Correlation provides an indication about linear dependence**. Two RV can be uncorrelated and yet display strong nonlinear dependence!

# 3. Capital adequacy

= set of all future *market scenarios*  
 = set of all observable *market events*  
P = historical probability

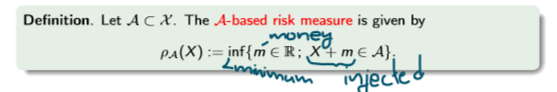
A = aggregate value of assets  
L = aggregate value of liabiliites  
X = RV, net capital, where X = A - L

For every terminal scenario , we have:

## Two main take-homes:

* **How to determine whether a company is adequately capitalized**
* **What to do if not**

**Variance** is unsifficient risk measure for CARM, because it doesn’t tell what to do in case of inadequacy.



A-based risk measure

The higher the figure :  
- the more costly to reach acceptability - the riskier position X

### Acceptance set

allows us to define **capital adequacy tests**

# Examples of risk measures

## Value at Risk (VaR)

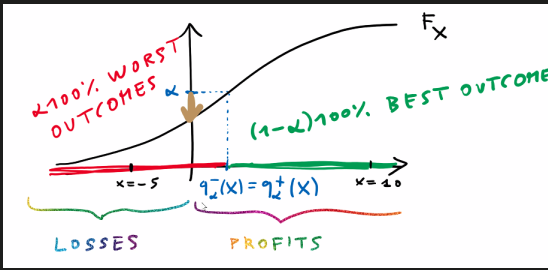
Definition. Let . The Value at Risk (VaR) at level is defined by

where upper last quantile

### Interpretation

The quantity is the **worst realization of X that may occur in the cases**. Alpha is close to zero IRL, between 0.005-0.05.

**Easy one** : VaR is the **worst** outcome of the best cases. That is precisely the the q.



Plot title.

It tells us, how much capital to inject into X in order to **ensure that default prob does not exceed alpha**.

### Pitfalls

* in discrete cases, the level might not be the worst cases
* because the way it is defined, it might not be the maximum loss- it can be also a profit.

### Operational interpretation

represent the **smallest amount of capital** that has to be injected into X to \*\*ensure that default prob. does not exceed level $

## Expected shortfall

### Interpretation

ES takes all VaR quantiles left of a given quantile and then normalizes is by .

### Operational interpretation

Recurring description: Acceptance set is such set, which has negative ES (because ES is itself negative, - times - = +)

**ES can be interpreted as the smallest amount of capital that has to be injected into X to ensure NO DEFAULTS occur on average in the tail beyond the upper -quantile of X.** company is AQ under ES at level if it is solvent on average over the tail beyond the upper -quantile.

*Alternative name* - average value at risk

### Question - How to choose ?

With calibration. There is no a-priori theory, usually is set to app. 0.025. In regulation, it is prescribed to 0.01.

## Worst-case risk measures

Definition. Let . The worst-case risk measure based on is defined by

### Interpretation

WC can be interpretated as the best realization of *-X* over the event A, or up to a sign as the worst realization of X over the event A.

### Operational interpretation

Smallest amount of capital injected into X in order to ensure that NO DEFAULTS occur in the scenarios belonging to A. company is adequately capitalized based on A if it solvent in each of the test scenarios prescribed by A.

## Loss Value at Risk

Definition. Let be increasing. The Loss Value at Risk (LVaR) at level is defined by

\*sup = supremum, i.e. highest value

### Interpretation

The parameter x is loss level. For each x we apply VaR to the transformed cap. position X-x, we let the prob. level (x) depend on the loss size. Monotonicity ensure highers loss = lower prob. LVaR is the most conservative. It is generalization of Var. If =, we get VaR.

### Operational interpretation

Smallest amount of cap. that has to be injected into X that for every loss threshold x, prob. of incurring a default beyond size x does not exceed the level (x).

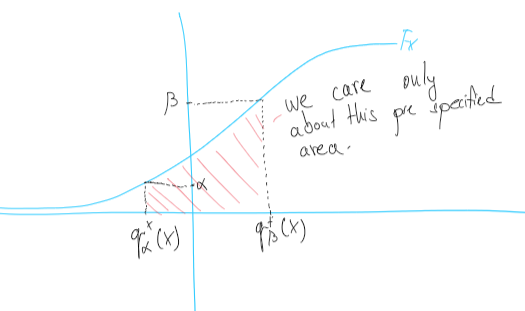
VaR controls only for prob. of default, LVaR controls for prob. of losing some specific value.

## Range Value at Risk

Definition. Let . The Range Value at Risk (RVaR) at levels and is defined by

### Interpretation

can be interpreted, up to a sign, as a kind of average realization of among all the realizations that are contained between and

 **We take the cut range, because at extreme ends, we have no data**

# 4. and 5. Properties of Risk Measures

**Main take home**: keep in mind what is a apt risk measure for CA

Proposition. Let . For all we have - is always finite. - is always finite. - - if and only if . - consider only the negative amounts and check, whether it’s finite or not

## Impact of capital injections

**KEY PROPERTY** satisfies the following cash-additivity property:

*Interpretation*: Cap. injections decrease cap. req. in a linear way cap. test are some defnites as **target capital**

where is capital at time 0 . Indeed, by cash additivity, we have

### Monotonicty

If company A has more capital than company B, which is adeq. cap., than A is also adeq. cap. Important for holdings, consolidation, etc.

### Stability under scaling

We say that is stable under scaling whenever

Stab. under scaling means acceptability is **independent of the size** od the cap. positions.

!! LVaR is NOT homogenous, scaling doesnt work for it.

### Stability under aggregation

If two positions are independtly adeq. cap., than their sum is also adeq. cap. Again important for consolidation, subsidiaries etc.

The sign, meaning sometimes two positions aggregated yield an “overcapitalized” positions, meaning: - **incentivizes** aggregation (always true) - **implied** diversification (if certain assumption are full-filled - they have to be i.i.d)

#### Three main implications:

* sum of cap. req. of indi. subposition is conservative proxy for the cap. req. of the agg. position (challenging to estimate)
* CA can be enforec in a decentralized way
* no incentive to break up balance sheet

Following risk measures are (not) subadditive:

* VaR: NO
* ES: YES
* WC: YES
* LVaR: NO
* RVaR: NO

### Coherence

**We say that is COHERENT iff it’s BOTH stable under scaling AND aggregation**.

Following risk measures are (not) coherent

* VaR: NO
* ES: YES
* WC: YES
* LVaR: NO
* RVaR: NO

!!! Careful:

If we have a fully-leveraged positions (zero-cost LS portfolio e.g.), we accumulate **loss peaks in the tail**, but because of subaddivity, the position even improves CA.

# 6. Tail risks

## Sensitivity to tail risk

### Definition

### Interpretation

If is sensitivity to tail risk, a large accum. of loss peaks will always result in a positive cap. req. There risk measures are sensitive to tail risk:

* VaR: NO
* ES: YES
* WC: NO
* LVaR: NO
* RVaR: NO

ONLY *ES* is sensitivity to tail risks!!

## Acceptability and distributions

We say that is law invariant whenever

### Interpretation

Law invariance says that acceptability depends only on the **probability distribution** of a capital position. Meaning we can use statistical estimation to compute capital requirements.

# 7. Surplus and default profiles

* shareholders’ surplus: random var, : what is extra from the acc. set, shareholders’ payoff
* shareholders’ default option, : incorporates the limited liability.If L-A is negative, meaning there is less cash than liabilities to satisfy them, creditors get 0 (max condition) and can suck it

#### For every capital positiion we have X=Sx-Dx

## Protection of liability holders

Definition. We say that is surplus invariant if

### Interpretation

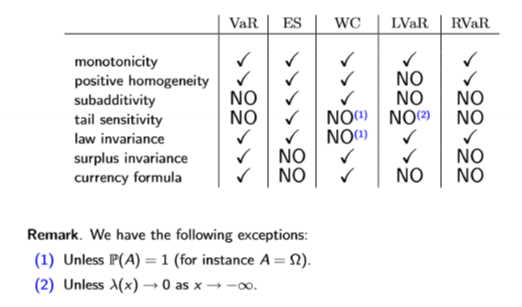
Surplus invariance = company whose default option is smaller than of an acceptable company, it’s also acceptable.

## Surplus invariance

Proposition. For every acceptance set the following are equivalent: - For every capital position we have

is surplus invariant. Interpretation. The outcome of a surplus-invariant capital adequacy test is independent of (the size of) the surplus and is only driven by the default profile.

# Summary (SUPER IMPORTANT)



# Review of Statistics

## Statistical inference

The basic goal of statistical inference is to provide **assesments about the PD that is assumed to have “generated” the sample**. Basically, we infer some characteristics from the sample distribution, which should have generated the sample and assume those characteristics are to remain constant and keep generating sample subject to the underlying rules and mechanics.

Standard assumption is the sample is i.i.d., even though it might not be necessarily true.

## Statistics (backtesting through hypothesis testing)

Definition. Let be a set. Any function is called statistic. We say that is a - (point) estimator if consists of points in . - interval/region estimator if consists of subsets of . - test statistic if contains two elements. - rank statistic if is finite and contains more than two elements.

### Interpretation:

the set (statistic) consists of **model parameters** (mean, variance) or **decision parameters** (“accept, reject”).

A statistic gives us a **rule to link the data with a specific model parameter or decision variable**

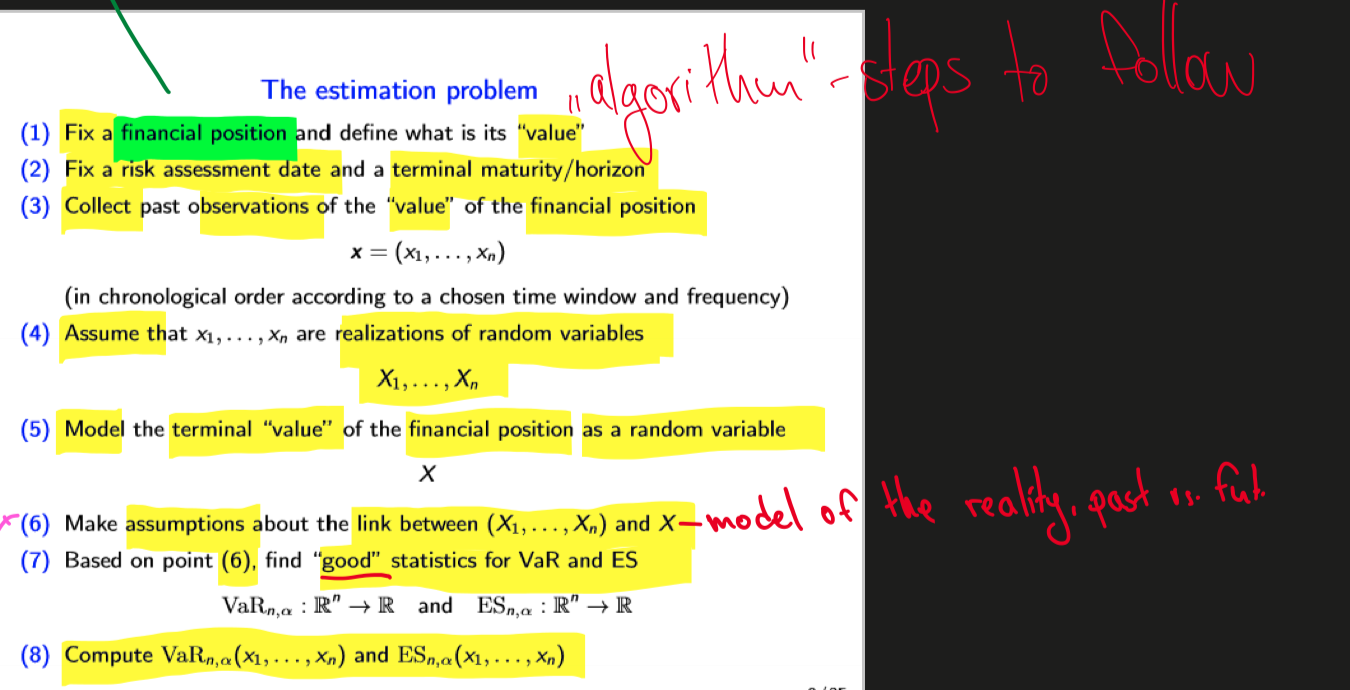
Example. Let . The statistic given by

is called the sample mean. Example. Let . The statistic given by for )

is called the sample variance.

# Statistical Estimation of VaR and ES

## Algoritm for estimation



Plot title.

### The estimation problem

1. Make assumptions about the link

* **unconditional approach**: the random variables X1 …. Xn and X are assumed to be **independent and identically distributed**
* **conditional approach**: the random vars are assumed to be *dependent* and to display different distributions

1. Choice of the sample size is crucial:

* **small sample size**: the RM statistics may be very volatile and hence encentivize too frequent portfolio rebalancing
* **large sample size**: the RM statistics may be too slow to react to losses or incorporate different volatility clusters

## From CA to PM

we do not have data about capital position of a financial institution.

We consider

### Interpretation:

Quantity is the capital to be injected to ensure we have a profit V1>V0 at an acceptable level of confidence

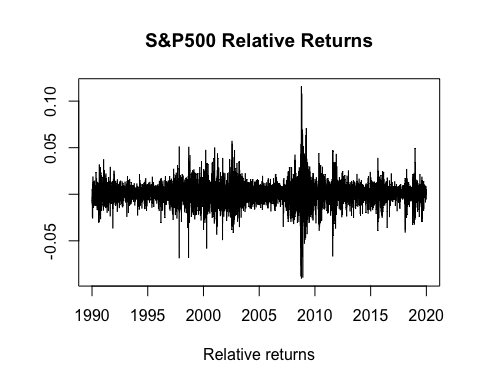
# Estimation setting

library ( zoo )

##   
## Attaching package: 'zoo'

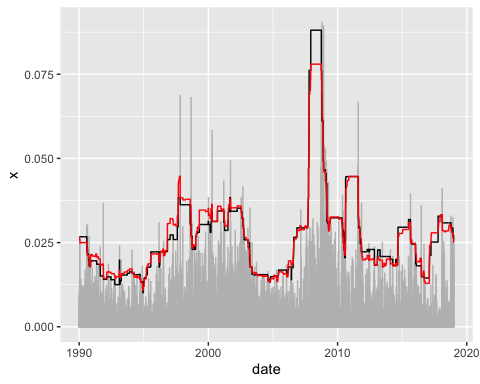
## The following objects are masked from 'package:base':  
##   
## as.Date, as.Date.numeric

library ( pdfetch )  
library (ggplot2)  
data = pdfetch\_YAHOO("^GSPC", fields="close", from="1990-01-01", to= "2019-12-31")  
price = coredata ( data ) #very similar to xlsread, coredata takes only the  
x = diff ( price )/ price [1:( length ( price ) -1)]  
  
  
time = index ( data )  
t = time [2: length ( time )]  
plot (t ,x , type ="l" , xlab ="Relative returns" , ylab = "", main="S&P500 Relative Returns")



hVaR = function (x , alpha ){  
n = length (x)  
k = floor (n\* alpha )  
r = - sort (x )[ k +1]  
return ( r)  
}  
compute\_hVaR = function (x ,n , alpha ){  
N = length (x)- n +1  
r = rep (0 , N)  
for (i in 1: N ){  
r[i ] = hVaR (x[ i :( i+n -1)] , alpha )  
}  
return ( r)  
}  
# % x is the time series of our relative returns  
hVarOutput = compute\_hVaR (x ,250 ,0.01)

hES = function (x , alpha ){  
n = length (x)  
k = floor (n\* alpha )  
r = - sum ( sort ( x )[1: k ]) / (n\* alpha )- sort ( x )[ k +1] \*( alpha -k/n )/ alpha  
return ( r)  
}  
compute\_hES = function (x ,n , alpha ){  
N = length (x)- n +1  
r = rep (0 , N)  
for (i in 1: N ){  
r[i ] = hES (x[i :( i+n -1)] , alpha )  
}  
return ( r)  
}  
#% x is the time series of ou  
hESOutput = compute\_hES(x ,250 ,0.025)  
  
date = t[1:length(hESOutput)]  
  
x = diff ( price )/ price [1:( length ( price ) -1)]  
b = length(x)  
for (i in 1 : b) {  
 if (x[i] < 0) {  
 x[i] = abs(x[i])}  
 else {  
 x[i] = 0  
 }  
 }  
x = x[1:length(hESOutput)]   
   
df = data.frame(date,hVarOutput,hESOutput,x)  
ggplot(df, aes(date)) +   
 geom\_col(aes(y = x), color = "gray") +# basic graphical object  
 geom\_line(aes(y=hVarOutput), colour="black") + # first layer  
 geom\_line(aes(y=hESOutput), colour="red") #second layer



# Estimatoin of VaR and ES - The Weighted Historical Approach

## The weighted sample distribution function

Definition. Let and consider a vector satisfying: The weighted sample distribution function associated with and is the distribution function given by

### Interpretation

The sample distribution is a step-wise function constructed by \*\*weighing past observations through the (probability) weights . It is a probability

For every we define the index

Weighted historical estimators. The weighted historical VaR is given by

The weighted historical ES is given by

### Interpretation:

The weighted historical VaR and ES coincide with the VaR and ES of the weighted sample distribution function

### Setting probability weights

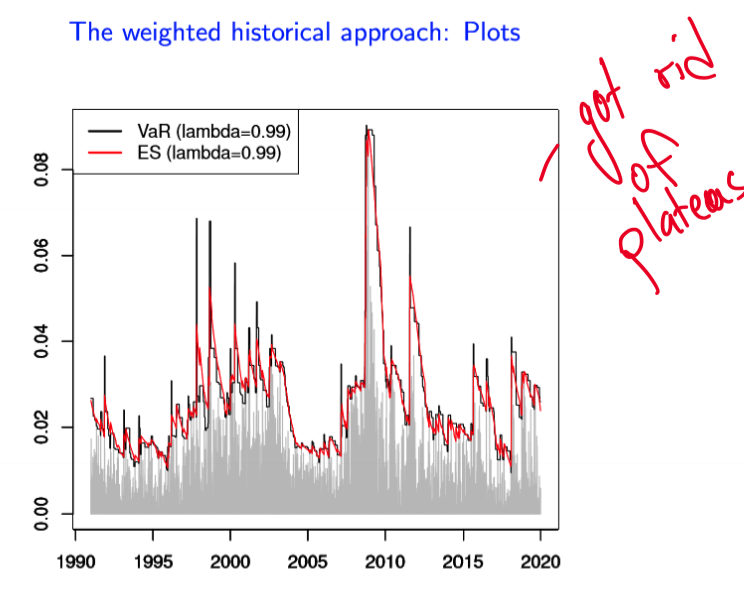
pweights = function ( lambda ){  
i = 1:250  
pi = lambda ^(250 - i )\* (1 - lambda )/ (1 - lambda ^250)  
return(pi)  
}

### Compute historical weighted Var

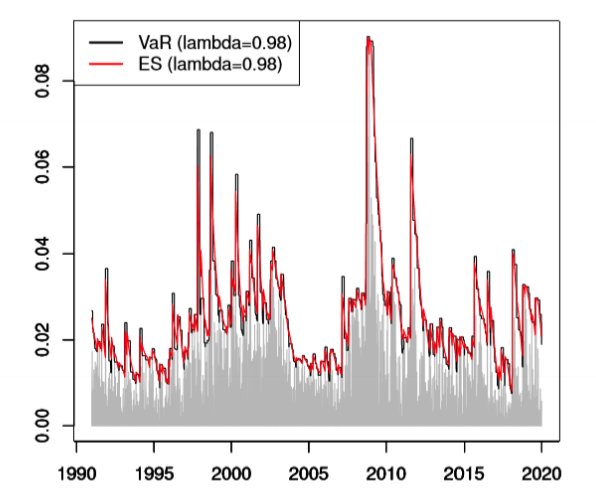
wVaR = function (x , lambda , alpha ){  
pi = pweights ( lambda )  
k = min ( which ( cumsum ( pi [ order (x )]) > alpha )) -1  
r = - sort (x )[ k +1]  
return ( r)  
}  
hold4 = compute\_wVaR = function (x ,n , lambda , alpha ){  
N = length (x)- n +1  
r = rep (0 , N)  
for (i in 1: N ){  
r[i ] = wVaR (x[ i :( i+n -1)] , lambda , alpha )  
}  
return ( r)  
}  
#% x is the time series of our relative returns  
hold1 = compute\_wVaR(x ,250 ,0.99 ,0.01)  
hold2 = compute\_wVaR(x ,250 ,0.98 ,0.01)

### Historical weighted ES

wES = function (x , lambda , alpha ){  
pi = pweights ( lambda )  
q = pi [ order (x )]  
k = min ( which ( cumsum (q) > alpha )) -1  
r1 = (k ==0) \* sort (x )[ k +1]  
r2 = (k >0) \*( sum ( sort (x )[1: k ]\*q [1: k ])+ sort (x )[ k +1] \* ( alpha - sum (q [1: k ])))  
r = -( r1 + r2 )/ alpha  
return ( r)  
}  
holder3 = compute\_wES = function (x ,n , lambda , alpha ){  
N = length (x)- n +1  
r = rep (0 , N)  
for (i in 1: N ){  
r[i ] = wES (x[i :( i+n -1)] , lambda , alpha )  
}  
return ( r)  
}  
#% x is the time series of our relative returns  
holder = compute\_wES(x ,250 ,0.99 ,0.025)  
holder2 = compute\_wES(x ,250 ,0.98 ,0.025)



Plot title.



Plot title.

# Estimation of VaR and ES: The Risk Metrics Approach

## VaR and ES under normality

Let be the distribution function of a standard normal random variable, i.e.

Proposition. Let Then, for every we have: - -

## Estimating mean and variance

Definition. For a given , the EWMA variance is defined by

or recursively by

The EWMA standard deviation is given by

Interpretation. The acronym EWMA stands for exponential weighted moving average. This is akin to the variance in an (I)GARCH model (see Lecture 10).

We define variance by (as above)

Lambda is usually .99 or .98. Interpretation: the older the observation is (lower i), the less weight it gives into the EWMA (given by ) and thus receives less importance

### Setting EWMA weights

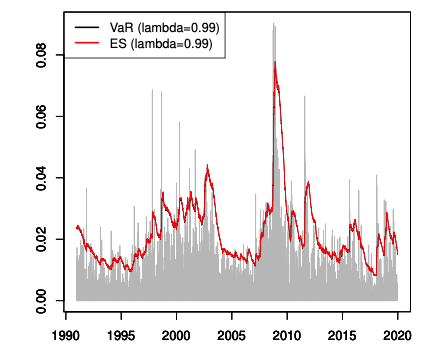
EWMAweights = function(lambda){ i = 1:250  
pi = (1-lambda)\*lambda^(250-i)   
return(pi)  
}

### Computing VaR in Risk Metrics Approach

rmVaR = function(x,lambda,alpha){  
n = length(x)  
m = sum(x)/n  
sd = sqrt(sum(EWMAweights(lambda)\*(x-m)^2))   
q = qnorm(alpha ,0,1)  
r = -m-sd\*q  
return(r)  
}

### Computing ES in Risk Metrics Approach

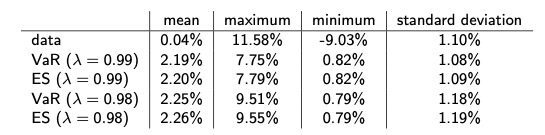
compute\_rmVaR = function(x,n,lambda,alpha){ N = length(x)-n+1  
r = rep(0,N)  
for (i in 1:N){  
r[i] = rmVaR(x[i:(i+n-1)],lambda ,alpha)  
}  
return(r)  
}  
# x is the time series of our relative returns compute\_rmVaR(x,250,0.99,0.01) compute\_rmVaR(x,250,0.97,0.01)



Plot title.



Plot title.



Plot title.

# Estimation of VaR and ES: Simulation Approach

## Generating distribution using quantiles

Proposition. Let . For every define the random variable

Then, we have

This process is called **“quantile inversion”**

Idea: By taking (upper) quantile of a DF F to a uniform random variable, we obtain a random variable with the same distribution F.

That shows that if we want to simulate realizations of a random variable with distribution F, it suffices to apply the upper quantile of F to simulated realizations of a uniform random variable.

## The simulation approach

The idea is that if we have a set of reutrns Z, even if we konw their distribution, it’s no trivial to know distribution of the whole set X. That is why we combine them with *f*, the DF from before.

### Algorithm

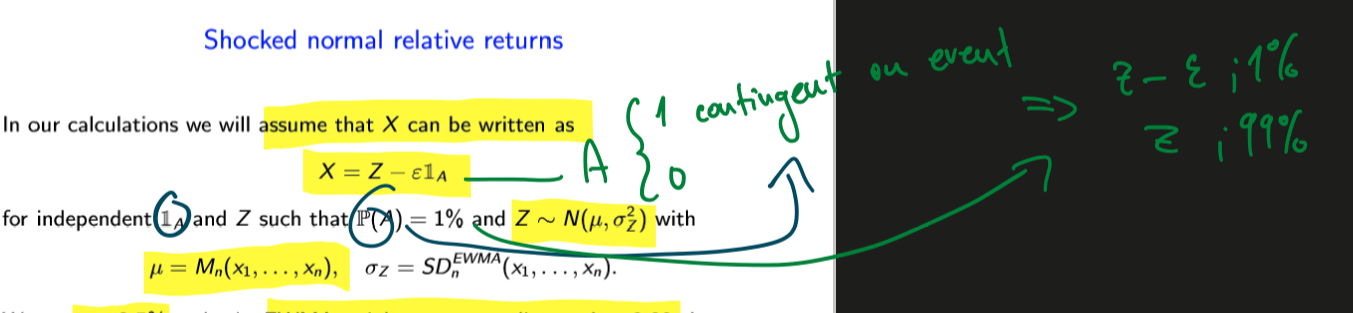
The simulation approach is based on the following steps: - Let (number of iterations) and (number of simulations per iteration). - For every let be realizations of independent uniform random variables on . - For every and set

- For every set

Remark. Note that we have to ensure independence also across the different iterations.

Then we perform VaR and ES for each iteration *h* and average across all iterations.

#### Shocked normal relative returns



Plot title.

We set and take EWMA weights corresponding to .

Interpretation: We assume that X follows a normal distr. subject to external shocks with probability 1%. The shock lowers relative returns by *2.5%*. Note that X is not normal.

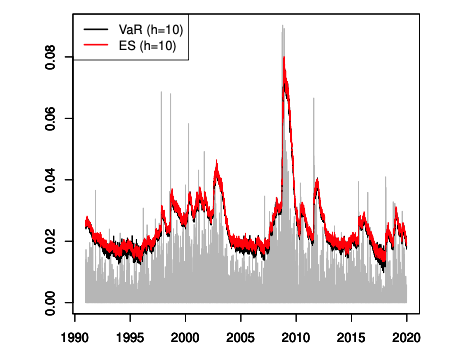
### Simulation VaR

simVaR = function(x,h,s,lambda,alpha){  
m = sum(x)/length(x)  
sd = sqrt(sum(EWMAweights(lambda)\*(x-m)^2))  
hr = rep(0,h)  
u = runif(2\*s\*h)  
sim\_z = qnorm(u[1:(s\*h)],m,sd)  
sim\_ind = qbinom(u[(s\*h+1):(2\*s\*h)],1,0.01)  
  
for (k in 1:h){  
z = sim\_z[((k-1)\*s+1):(k\*s)]  
ind = sim\_ind[((k-1)\*s+1):(k\*s)]  
x = z-0.025\*ind  
hr[k] = hVaR(x,alpha)  
}  
r = sum(hr)/h  
  
return(r)  
}  
compute\_simVaR = function(x,n,h,s,lambda,alpha){ N = length(x)-n+1  
r = rep(0,N)  
for (i in 1:N){  
r[i] = simVaR(x[i:(i+n-1)],h,s,lambda ,alpha)  
}  
return(r)  
} # x is the time series of our relative returns compute\_simVaR(x,250,1,500,0.99,0.01)  
 simulatedVaR = compute\_simVaR(x,250,10,500,0.99,0.01)

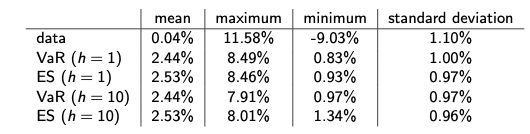
simES = function(x,h,s,lambda,alpha){  
m = sum(x)/length(x)  
sd = sqrt(sum(EWMAweights(lambda)\*(x-m)^2))  
hr = rep(0,h)  
u = runif(2\*s\*h)  
sim\_z = qnorm(u[1:(s\*h)],m,sd)  
sim\_ind = qbinom(u[(s\*h+1):(2\*s\*h)],1,0.01)  
for (k in 1:h){  
z = sim\_z[((k-1)\*s+1):(k\*s)]  
ind = sim\_ind[((k-1)\*s+1):(k\*s)]  
x = z-0.025\*ind  
hr[k] = hES(x,alpha)  
}  
r = sum(hr)/h  
return(r)  
}  
compute\_simES = function(x,n,h,s,lambda,alpha){ N = length(x)-n+1  
r = rep(0,N)  
for (i in 1:N){  
r[i] = simES(x[i:(i+n-1)],h,s,lambda ,alpha)  
}  
return(r)  
} #% x is the time series of our relative returns   
simES1 = compute\_simES(x,250,1,500,0.99,0.025)   
simES2 = compute\_simES(x,250,10,500,0.99,0.025)



Plot title.



Plot title.



Plot title.

# Estimation VaR and ES - Conditional Approach

So far, we assumed i.i.d. Empirically we know wolatility clusters, so no good.

## Filtrations

Definition. An ( -dimensional) filtration is any sequence of -fields

such that We always assume that .

**Interpretation**: filtration models the *flow of information* - at each point in time we get new information and we want to discard some outcomes as impossible.

Definition. Let and consider an -dimensional filtration . We say that is -adapted whenever is -measurable for every .

## The general time series model

Setup. Let and take a filtration Set and assume that are part of an -adapted process such that

where and are is the innovation term. We stipulate the following assumptions: - the elements of are independent - the elements of are identically distributed as - and - is strictly positive (almost surely) Remark: The time series , with standardized i.i.d. elements and zero unconditional mean, is also known in time series analysis as **“white noise”** - that contains no information



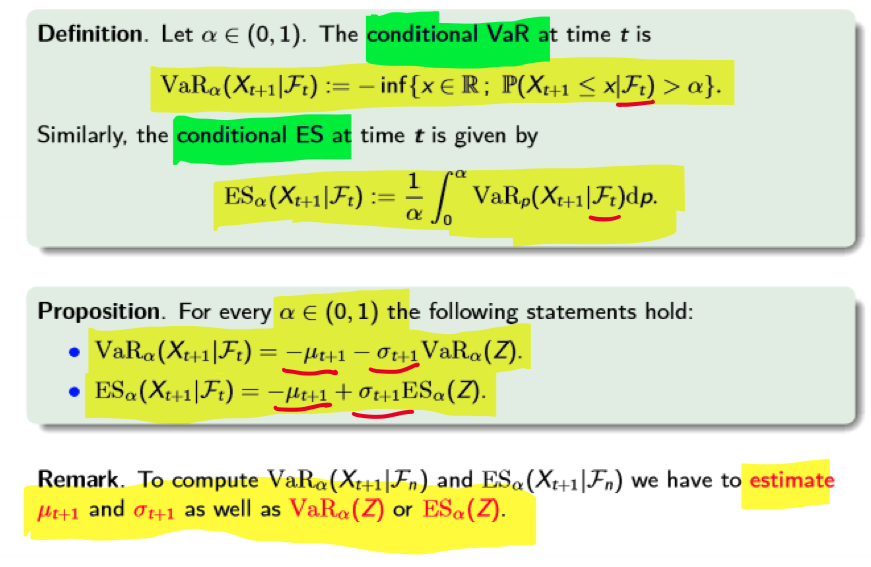
Plot title.

## Maximum-Likelihood estimation

Definition. Let and take a filtration . Consider an -adapted process with joint density function , with an unknown parameter vector. The likelihood function and log-likelihood function are defined respectively as

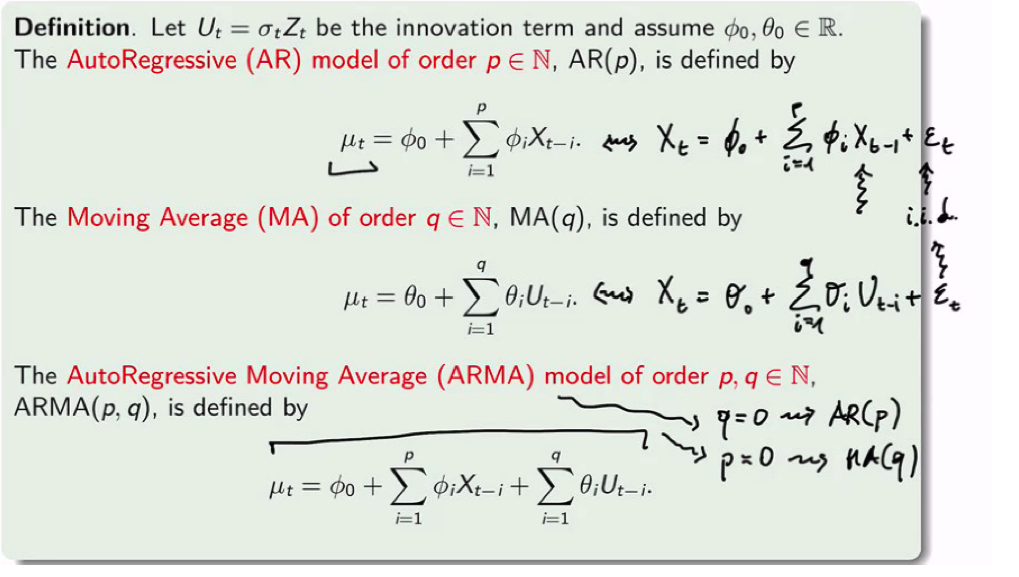
The maximum likelihood estimator (MLE) of is the maximizer of .

## Conditional VaR and ES



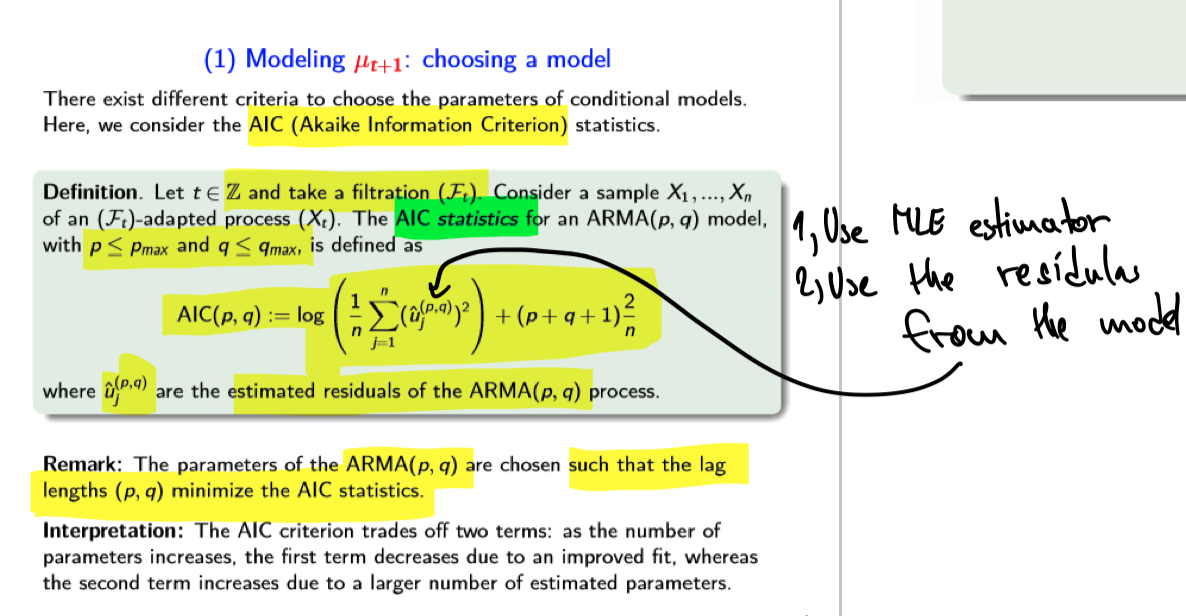
Plot title.

## Modeling AR, MA and ARMA models



Plot title.

### Choosing the model



Plot title.

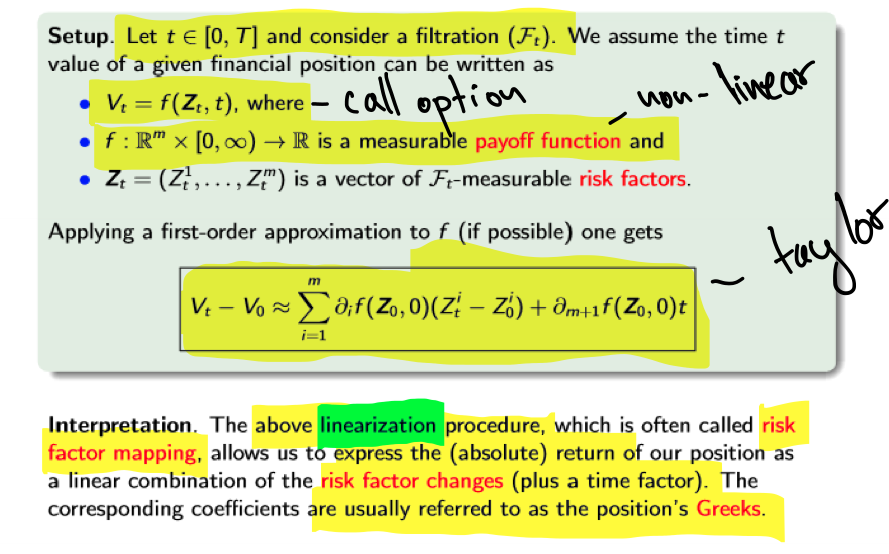
## Modelling conditional heteroskedastic models

It conditional and unconditional variance is the same, it’s called *conditional homoskedasticity*.

the elements of are serially uncorrelated but dependent, i.e. if we emove the 2nd “i” in i.i.d., their conditional variance is not constant anymore

This property is called conditional heteroskedasticity: the dependence across nakes the flow of information relevant to determine their variance, which varies ser time as new information becomes available.

## Risk Factor Mapping



Plot title.

Linearization is regression

### Delta VaR and ES

Definition. Let and set The Delta VaR is given by

Similarly, the Delta ES is given by

Interpretaion: Delta approach is useful when risk factors can be estimated - we have a lot of data. Then we believe payoff of our position can be safely approximated in a linear way.

#### Delta VaR and ES under normal risk factors

Proposition. For let and be an invertible positive definite (symmetric) matrix. Then, for every the following statements are equivalent: - for every -

Proposition. Let and be an invertible positive definite (symmetric) matrix. Assume that . Then, we have:

### Estimating covariance

Risk Metrics estimators. The Risk Metrics VaR estimator is given by The Risk Metrics ES estimator is given by

Interpretation: above esimators coincide with the Delta VaR and ES computed under the assumption that the risk factor changes are jointly normal

### Time conversoin rules

* if we want 1- day risk measure, we use daily data
* if we want 5-day RM, we use weekly

**Warning 1:** if chose horizon is too far, we have sparse data set. By doing so - We hide loss peaks between consecutive data points - include data points not include

**Warning2:** If we have a multi-day risk measure using a rolling window, we have to roll the window according to the chosen horizon

Let and . In this case is safer to estimate - is trickier to estimate We ask whether we can express in terms of as in

In this case, we could rely on the estimation of and derive the estimates of by exploiting the above link.

Easy interpretation: if we want longer horizons, because of the time thing, we esimate a shorter, i.e. safer VaR and then transform it to longer horizon with the distribution. In practice, it is done by the square-root rule

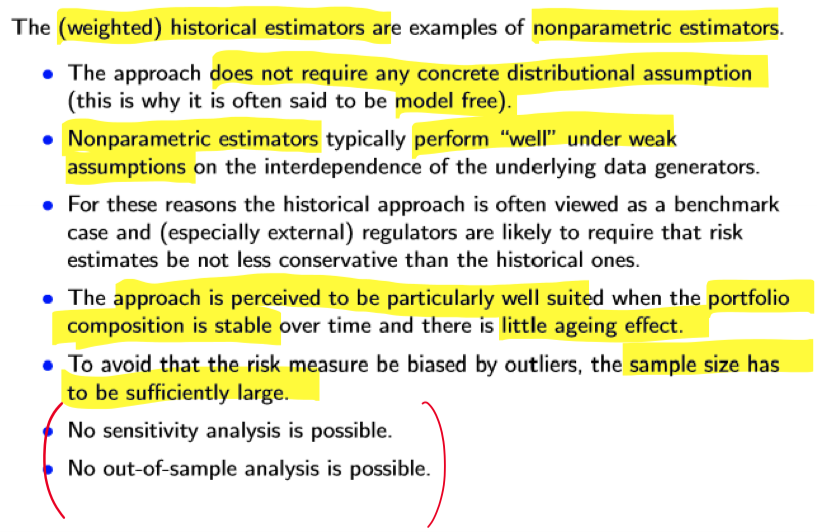
#### Square root rule

Done in practice but impossible to prove mathematically.

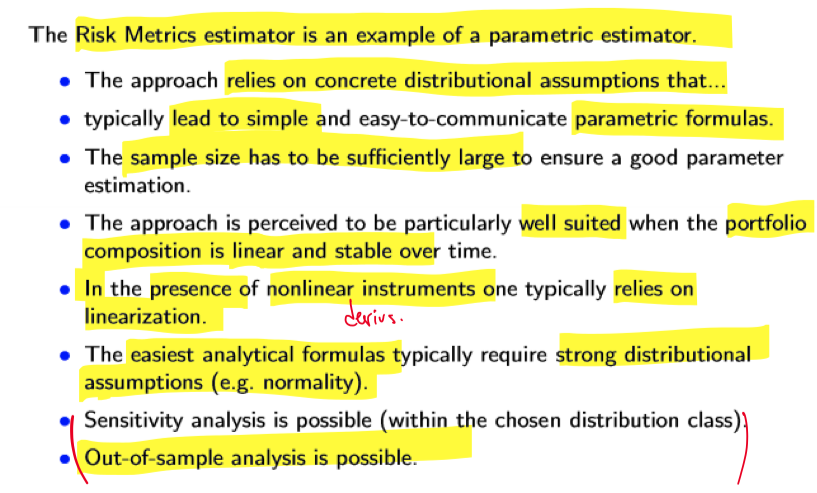
The square-root rule. Let and . The square-root rule for VaR postulates that

The square-root rule for ES postulates that

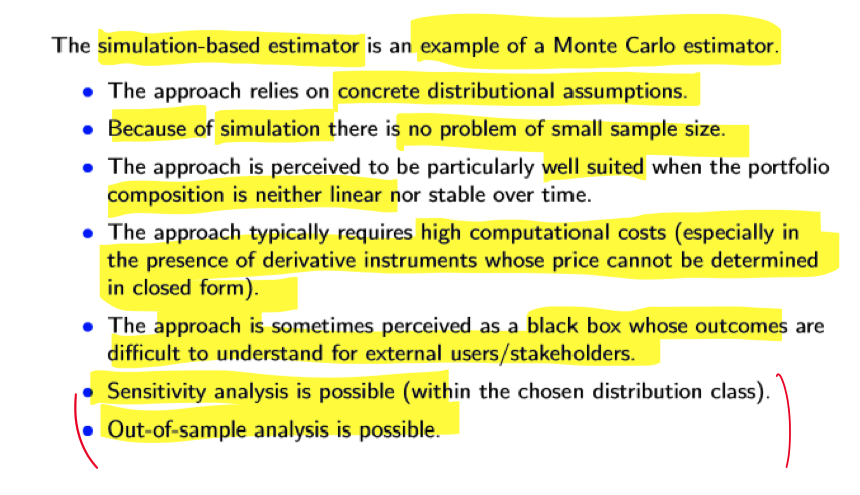
## Overview



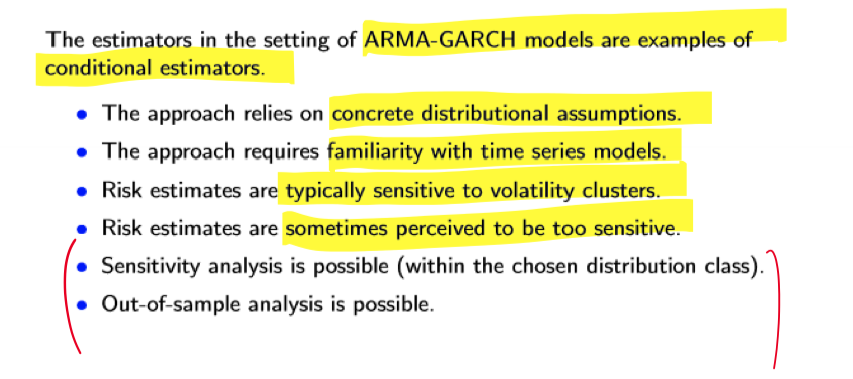
**Nonparametric**



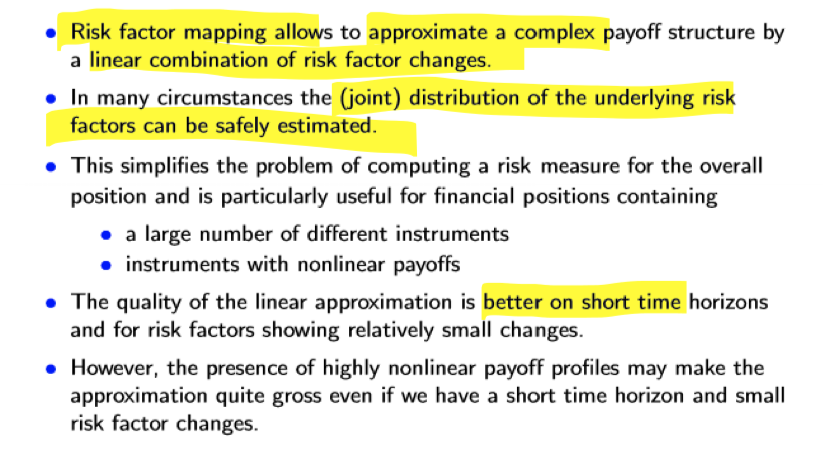
**Parametric**



**Monte Carlo**



**Conditional estimators**



**Risk factor mapping**

# Backtesting under VaR and ES

The term **backtesting** is a *process of assesing if ex-ante estimation are aligned with ex-post observations*.

## Model selection

Setup. Let be a law-invariant risk measure. To perform model selection using one can do the following: - Collect past observations . - Use to come up with competing models for , say

- Compute the risk estimates . - Assume that materializes into . - We want to use to select the best model out of

By doing backtest for all models, we look which performs the best. More specifically, we are selecting our model \*\*based on the performance of the risk measure .

**Question:** How to decide which model is better?

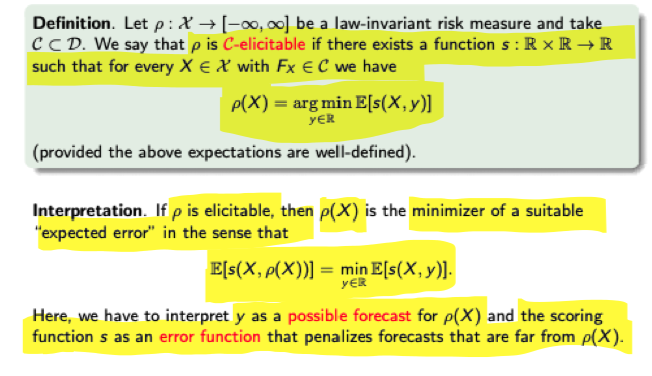
## Model validation

Setup. Let be a law-invariant risk measure. To perform model validation using one can do the following: - Collect past observations - Use to come up with a model, say , for . - Compute the risk estimate . - Assume that materializes into . - We want to use to validate the choice of or not.

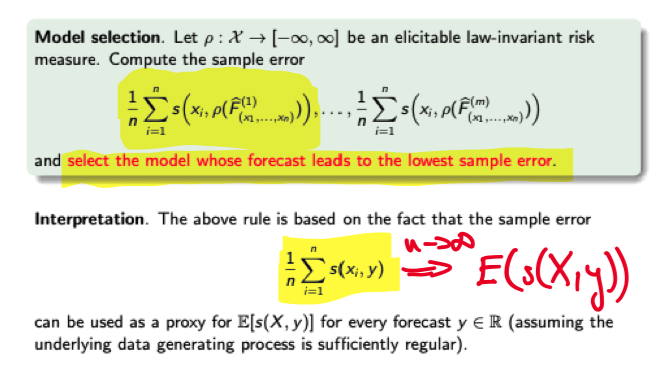
We compare ex-post observations with ex-ante model forecasts to assess whether our model assumption should be kept or not. This is **model valiatoin**.

**Question:** How to decide whether the model is validaated or not?

### Elicitability



Plot title.



Plot title.

#### Elicitability under VaR

Lemma. Let For every we have

Moreover, the infimum is attained at any -quantile of .

\*\* Intepreation\*\*: VaR is the minimizer and ES is the results. That also leads to the fact that ES is not elicitable.

compute\_scoreVaR = function(x,risk,n,alpha){ N = length(x)-n+1  
s = rep(0,N)  
for (i in 1:N){  
s[i] = mean(risk[i]-(risk[i]+x[i:(i+n-1)])\*(-x[i:(i+n-1)]-risk[i]>=0)/alpha) }  
m = mean(s)  
return(m)  
}  
#% x is the time series of our relative returns  
#VaR is the time series of our VaR estimates compute\_scoreVaR(x,VaR,250,0.01)

## Identifiability

Definition. Let be a law-invariant risk measure and take . We say that is -identifiable if there exists a function such that for every with we have: - if and only if - is strictly increasing (in (provided the above expectations are well-defined).

**Interpretation**: If is identifiable, the is the unique zero of a suitable “expected distance” in the sense that:

We have to interpret *y* as possible forecast for and the function as a *penalty function* that penalizes forcasts that are from from . assigns a negative average value to underestimated forecasts and a psoitive average value to overestimated ones.

Identifiability and model validation/selection Model validation and selection. Let be an identifiable law-invariant risk measure. Compute the sample penalty

and validate the model if the sample penalty is sufficiently close to zero. Similarly, compute the sample penalties

and select the model whose forecast leads to the smallest sample penalty. Interpretation. The above rules are based on the fact that the sample distance

can be used as a proxy for for every forecast (assuming the underlying data generating process is sufficiently regular).

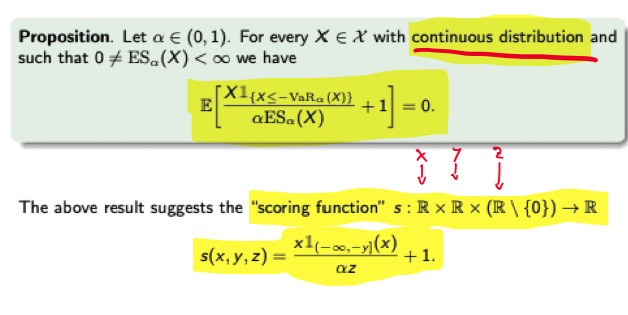
VaR is identifiable, ES is not.

compute\_penaltyVaR = function(x,risk,n,alpha){ N = length(x)-n+1  
p = rep(0,N)  
for (i in 1:N){  
p[i] = abs(mean(1-(1/alpha)\*(-x[i:(i+n-1)]-risk[i]>0)))   
m = mean(p)  
}  
return(m)  
}  
# x is the time series of our relative returns  
# VaR is the time series of our VaR estimates compute\_penaltyVaR(x,VaR,250,0.01)

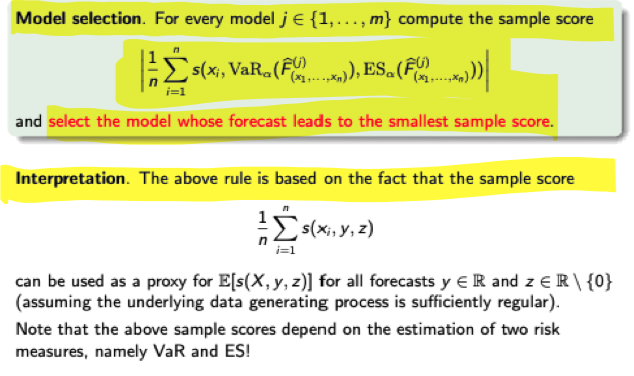
## Backtesting under ES

ES is neither identifiable nor elicitable - does **not** mean we cannot perform backtests.

We can use a “scoring function” that is yero on average in correspondence of ES.



Plot title.



Plot title.

compute\_scoreES = function(x,risk1,risk2,n,alpha){ N = length(x)-n+1  
s = rep(0,N)  
for (i in 1:N){  
s[i] = abs(mean(x[i:(i+n-1)]\*(x[i:(i+n-1)]+risk1[i]<=0)/(alpha\*risk2[i])+1)) }  
m = mean(s)  
return(m)  
}  
#x is the time series of our relative returns  
#VaR is the time series of our VaR estimates  
# ES is the time series of our ES estimates compute\_scoreES(x,VaR,ES,250,0.025)

# Backtesting Under VaR : The Basel Framework

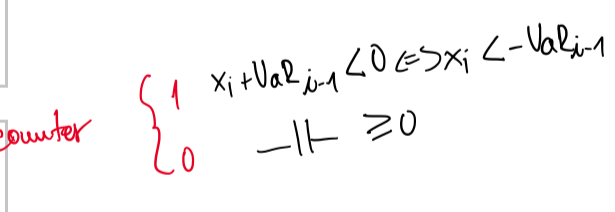
## The Traffic Light Rule (TLR)

TLR is based on VaR and says a good model for **underlying data generating process** (data generation bscl.) should not give rise to too many losses exceeding the corresponding VaR level.

**Testing of model is testing of the underlying distribution itself**.

We want to calculate the **coverage ratio**, which test how often the above conditon (arrised default) is violated.

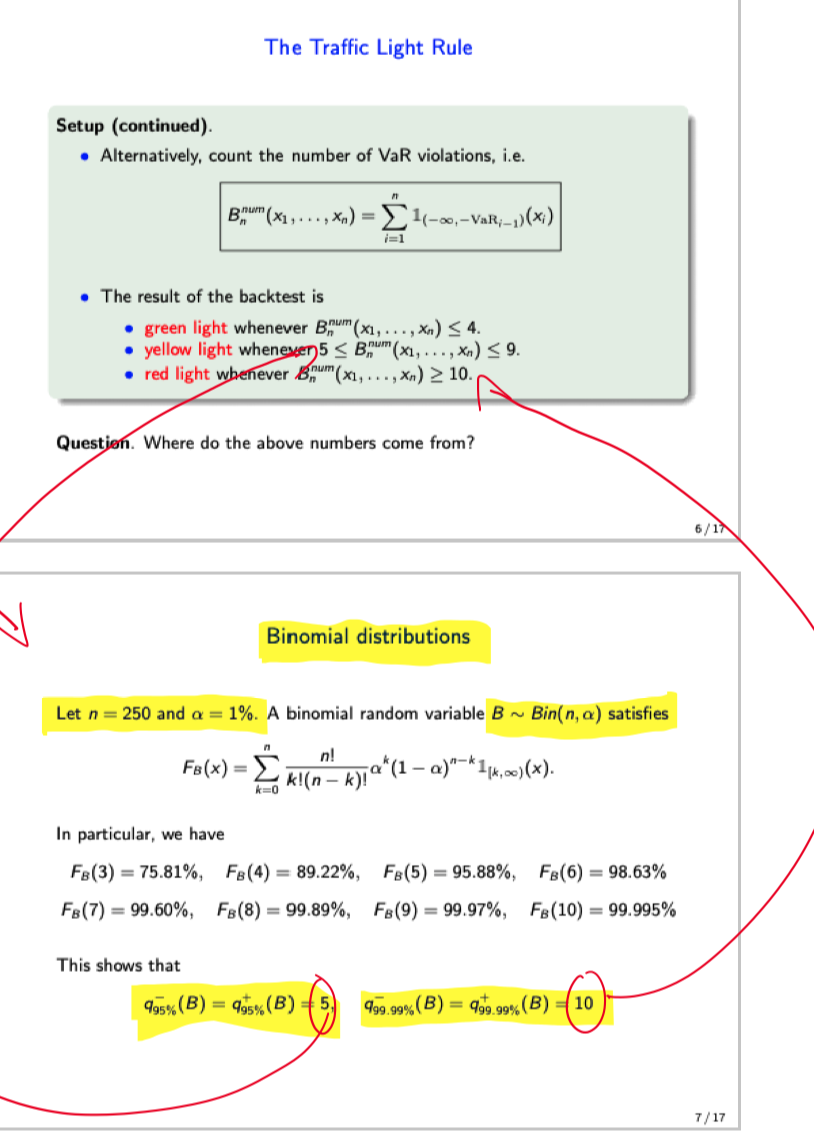
Determine the frequency of VaR violations, also called coyerage ratio, by



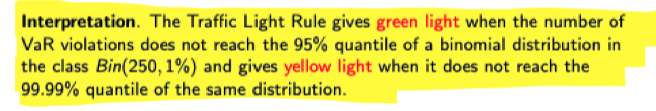
Plot title.

* green light if freq. ≤ 1.6%
* yellow light if 1.6% ≤ freq. ≤ 3.6%
* red light if freq > 3.6%

These percentages, or rather absolute number in relation to 252 (basel framework) come from Biniomal distributions.



Plot title.



Plot title.