Stockholms universitet

Value-at-Risk: Prediktion med GARCH(1,1), RiskMetrics och Empiriska kvantiler

Lukas Martin

Kandidatuppsats i matematisk statistik
Bachelor Thesis in Mathematical Statistics

Kandidatuppsats 2016:5
Matematisk statistik
Juni 2016

www.math.su.se

Matematisk statistik
Matematiska institutionen
Stockholms universitet
106 91 Stockholm

Stockholms universitet

Matematisk statistik
Stockholms universitet
Kandidatuppsats **2016:5**
http://www.math.su.se/matstat

Value-at-Risk: Prediktion med GARCH(1,1), RiskMetrics och Empiriska kvantiler.

Lukas Martin*

Juni 2016

Sammanfattning

Value-at-Risk, VaR, mäter den potentiella förlusten i en investering och används frekvent inom riskhantering. Prediktion av VaR ligger till grund för att uppskatta och hantera risk. Vi beskriver i det här arbetet grundläggande finansiell tidsserieanalys där vi tar fram olika volatilitetsmodeller som vi sedan tillämpar på VaR-prediktion.

Syftet med arbetet är undersöka olika volatilitetsmodeller och att med Kupiec's test utvärdera deras förmåga att prediktera VaR.

Vi undersöker logaritmerade avkastningar för aktieindexet OMXS 30 som vi kommer fram till inte är autokorrelerade men konditionellt heteroskedastiska. Sedan anpassar vi volatilitetsmodellerna, GARCH(1,1) med normal- resp. t-fördelade innovationer,
IGARCH(1,1) med antaganden enligt RiskMetrics och Empiriska kvantiler. De fyra modellerna används till att prediktera 1-dags VaR 250 dagar framåt med rullande observationer som underlag för modellanpassning.

Resultaten vid backtesting med Kupiec's test ger oss ingen god uppfattning om hur bra de olika modellerna är på att prediktera VaR. Samtliga modeller får både bra och dåliga resultat i testet beroende på hur många rullande observationer vi har som skattningsunderlag.

Vi utför sedan VaR-prediktioner på 9 överlappande perioder om 250 observationer och får backtesting-resultat som tyder på att GARCH-modellerna är bra och att modellen Empiriska kvantiler är dålig. Resultaten motsäger de testen vi gjorde vid en enstaka period.

Slutligen görs backtesting på simulerad data. Utifrån hur våra simuleringar utformats och tolkats blir vår slutsats att vi inte med säkerhet kan avgöra om volatilitetsmodellerna predikterar VaR väl baserat på Kupiec's test med 250 prediktioner.

*Postadress: Matematisk statistik, Stockholms universitet, 106 91, Sverige.
E-post: j.lukas.martin@gmail.com. Handledare: Joanna Tyrcha, Filip Lindskog och Mathias Lindholm.

Abstract

Value-at-Risk measures the potential loss of an investment and is frequently used in Risk Management. Prediction of VaR is used to assess and manage risk. In this paper we describe basic financial timeseries analysis by analyzing different volatility models and using them for VaR-prediction.

The aim of this paper is to look at different volatility models and with Kupiec's test analyze their ability to predict Value-at-Risk.

We look at log-returns for the Swedish stock index OMXS 30 and draw the conclusion that they are not autocorrelated but conditionally heteroskedastic. We then fit the models, GARCH(1,1) with normal- and t-distributed innovations, IGARCH(1,1) with the assumptions by RiskMetrics and Empirical Quantiles. The four models are used to predict 1-day VaR for 250 days with rolling observations used for model-fitting.

The results obtained by backtesting with Kupiec's test give no conclusions on how well the models predict VaR. All models get good as well as bad results in the test depending on how many rolling observations we have used.

We do VaR-predictions on 9 overlapping periods of 250 observations and obtain good backtesting results for the GARCH-models and bad results for the Emperical Quantiles-model. The results for the overlapping periods contradict the results obtained testing at the one single period.

Finally backtesting is preformed on simulated data. Based on how our simulations were made and interpreted our conclusion is that we can't with certainty decide if our volatility models predict VaR well based on Kupiec's test with 250 predictions.

Tack

Jag vill tacka mina handledare Joanna Tyrcha, Filip Lindskog och Mathias Lindholm för god feedback och handledning. Jag vill även tacka andra studenter på institutionen för diskussioner kring kandidatarbetet.

Innehåll

1 Inledning

Value-at-Risk, förkortat VaR, är en statistiskt beräkningsmetod för att uppskatta finansiell risk hos en investering. VaR går ut på att med viss säkerhet uppskatta den potentiella förlusten.

Riskmåttet består av tre komponenter, potentiell förlust, sannolikhetsgrad och tid. Value-at-Risk uttrycker exempelvis att sannolikheten att vi förlorar mer än 15% av vår investering den kommande veckan är 5%. Det innebär att vi med 95% säkerhet kan säga att vi inte kommer förlora mer än 15% av investeringen, 15% av investeringen är således vårt Value-at-Risk.

Hantering av risk är fundamentalt inom finansbranchen där risktagande är nödvändigt för avkastning medan för stora risktaganden kan leda till att aktörer på finansmarknaden går i konkurs och i extrema fall världsomfattande finanskriser.

Riskmått som Value-at-Risk används därför av finansiella aktörer som vill kontrollera sina risker och är även av intresse för myndigheter vill kontrollera att aktörerna har god riskhantering.

1.1 Bakgrund och Syfte

Value-at-Risk kan uppskattas på flera olika sätt. En god prediktion av VaR kräver en bra modell för att beskriva variationen, så kallat volatiliteten, i en investering. Det har utvecklats en rad olika mer eller mindre avancerade volatilitetsmodeller. Vi kommer i det här arbetet främst fokusera på några av de mest etablerade modellerna baserat på tidsserieanalys.

Syftet med detta kandidatarbete är att beskriva och tillämpa olika volatilitetsmodeller för att uppskatta risken på marknaden. Vi använder volatilitetsmodellerna för att prediktera Value-at-Risk och utvärderar sedan hur väl prediktionerna stämmer överens med verkligheten. De modeller vi kommer tillämpa är normal respektive t-fördelad GARCH(1,1), RiskMetrics och Empiriska kvantiler vilka är väl etablerade akademiskt såväl som professionellt.

2 Teori

I den här delen av rapporten beskriver vi den bakomliggande teorin som sedan tillämpas i analysen i Kapitel 4. Vi går kort igenom grunden för finansiell tidsserieanalys och volatilitetsmodeller. I Kapitel 2.3 introducerar vi begreppet Value-at-Risk och presenterar de modeller vi kommer att använda för prediktion i analysen. Slutligen i Kapitel 2.4 beskriver vi Backtesting som metod att utvärdera Value-at-Risk prediktionerna.

2.1 Avkastningar

När vi studerar finansiell data tittar vi på avkastningar för finansiella tillgångar som aktier eller som i vårt fall aktieindex. Det finns flera sätt att definiera avkastningar på. Vi kommer i vår analys betrakta dagliga logaritmerade avkastningar på grund av att de har fördelaktiga statistiska egenskaper. Vi definierar avkastningarna i enlighet med [Tsa10, s.3-5] och använder samma beteckningar.

För dag t definierar vi den dagliga relativa avkastningen som

$$R_t = \frac{P_t}{P_{t-1}} - 1 = \frac{P_t - P_{t-1}}{P_{t-1}}. \tag{1}$$

Där P_t är stängningspriset dag t. Vidare definierar vi den dagliga logaritmerade avkastningen för dag t som

$$r_t = ln(1 + R_t) = ln(\frac{P_t}{P_{t-1}}). \tag{2}$$

En av de främsta fördelarna med logaritmerade avkastningar, r_t, mot relativa avkastningar, R_t, är att vi kan anta att r_t är normalfördelad. Vi har att $P_t > 0$ samt $P_{t-1} > 0$ i ekvationerna (1) och (2) vilket ger att $R_t > -1$ och således att $1 + R_t > 0$. Eftersom normalfördelningen är definierad för alla reella tal håller inte ett anatagande om normalfördelning. För logaritmerade avkastningar gäller dock att $r_t > ln(0) = -\infty$ och vi kan anta normalfördelning. Detsamma gäller för att anta t-fördelning för data, båda fördelningarna är vanligt förekommande vid modellering av finansdata.

2.2 Finansiell tidsserieanalys

Framöver kommer vi att betrakta våra logaritmerade avkastningar, r_t, som en samling stokastiska variabler över tid vilket ger oss tidsserien $\{r_t\}_{t=1}^{T}$ där T är antalet dagar vi observerar, se [Tsa10, s.29].

Med linjär tidsserieanalys kan vi beskriva linjära samband mellan r_t och seriens tidigare värden. Framöver kallar vi dessa samband för autokorrelationer.

Vi kommer främst vara intresserade av volatilitetsprocessen som beskriver variansen för returserien över tid, betingat på tidigare observerade värden. Vidare undersöker vi

olika modeller för volatiliteten som sedan används i vårt syfte att uppskatta risken i avkastningsserien.

2.2.1 Volatilitetsmodeller

Vi härleder volatilitetsprocessen ur avkastningsserien och beskriver här den grundläggande strukturen för volatilitetsmodeller.

Vi definierar medelvärde, μ_t, respektive varians, σ_t^2, för avkastningarna, r_t betingat på information tillgängligt vid tid t, F_{t-1}, som

$$\mu_t = E(r_t|F_{t-1}), \quad \sigma_t^2 = Var(r_t|F_{t-1}) = E[(r_t - \mu_t)^2|F_{t-1}],$$

där F_{t-1} vanligtvis består av alla tidigare linjära funktioner av avkastningarna vid tid $t-1$, se [Tsa10, s.22].

Vidare kan vi dela upp avkastningsserien i två delar, en medelvärdesekvation, μ_t och innovationer, a_t, som betraktas vid modellering av volatiliteten.

$$r_t = \mu_t + a_t$$

av vilket fås att

$$\sigma_t^2 = Var(r_t|F_{t-1}) = E[(r_t - \mu_t)^2|F_{t-1}] = E[(a_t)^2|F_{t-1}] = Var(a_t|F_{t-1}).$$

Volatiliteten beskrivs alltså av a_t som inte är autokorrelerad men beroende i den mening att a_t^2 är autokorrelerad. Vi säger att serien är konditionellt heteroskedastisk vilket uttrycker sig som perioder av ökad volatilitet, kallat volatilitetskluster. Denna effekt kallas även Autoregressive Conditional Heteroskedasticity (ARCH) och modelleras med ARCH-modeller som vi kommer beskriva senare i detta kapitel.

Tillvägagånssättet för att konstruera en volatilitetsmodell beskrivs av [Tsa10, s.113] med följande steg.

1. Uttryck en medelvärdesekvation, om autokorrelation förekommer för r_t kan den anpassas med exempelvis en AR-modell[1]. Om autokorrelation inte förekommer reduceras medelvärdesekvationen till stickprovsmedelvärdet för r_t.
2. Ansätt residualerna från medelvärdesekvationen till a_t och undersök om det förekommer autokorrelation för a_t^2, så kallad ARCH-effekt. Om ARCH-effekt förekommer anpassas en modell för innovationerna, a_t som utgör själva volatilitetsmodellen.

Vi kommer i den här rapporten fokusera på olika modeller för modellering av ARCH-effekten och deras förmåga att prediktera riskmåttet Value-at-Risk som beskrivs i Kapitel 2.3.

[1] AR modellen beskrivs i Kapitel 2.2.4

2.2.2 Autokorrelationsfunktionen (ACF)

Betrakta en tidsserie r_t där korrelationen mellan r_t och dess tidigare värden r_{t-l} beskrivs av lag$-l$ autokorrelationen,

$$\rho_l = \frac{Cov(r_t, r_{t-l})}{\sqrt{Var(r_t)Var(r_{t-l})}}.$$

Det gäller att $\rho_0 = 1$ och $\rho_l \in [-1, 1]$, se [Tsa10, s.31]

2.2.3 Ljung-Box test

Ljung-Box testet används för att testa om det föreligger någon autokorrelation hos en tidsserie. $H_0 : \rho_1 = ... = \rho_m = 0$ testas mot $H_a : \rho_i \neq 0$ för något $i \in \{1, ..., m\}$ med,

$$Q(m) = T(T+2)\sum_{l=1}^{m} \frac{\hat{\rho}_l^2}{T-l} \sim \chi_\alpha^2(m).$$

H_0 förkastas då $Q(m) > \chi_\alpha^2(m)$ där m beskriver hur många lag vi tar hänsyn till i testet såväl som antalet frihetsgrader. Val av m kan påverka utfallet; i standardfall har $m \approx ln(T)$ visat sig vara lämpligt där T betecknar antalet observationer i serien, se [Tsa10, s.33].

2.2.4 AR

AR-modellen är en linjär tidsseriemodell som kan tillämpas på exempelvis avkastningsserien, r_t, om den är autokorrelerad. AR-modellen ligger till grund för mer avancerade tidsseriemodeller och vi har även användning av AR-modellen för att modellera medelvärdesekvationen i en volatilitetsmodell. Vi beskriver här AR-modellen utifrån [Tsa10, s.37].

Om det för serien r_t finns signifikant lag-1 autokorrelation, dvs. ρ_1 är signifikant skild från noll kan r_t modelleras med hjälp av r_{t-1} enligt

$$r_t = \Phi_0 + \Phi_1 r_{t-1} + a_t,$$

där serien a_t antas vara vitt brus[2] med väntevärdet noll och varians σ_a^2.

Ovanstående modell kallas AR(1) där ettan står för antalet lag autokorrelationer vi betraktar. Modellen har samma form som en enkel linjär regressionsmodell med r_t som responsvariabel och r_{t-1} som förklaringsvariabel och har liknande egenskaper[3]. Det är

[2] "White Noise", är en sekvens av iid. variabler med begränsat medelvärde och varians.
[3] AR-modelens egenskaper diskuteras i [Tsa10, s.38]

vanligt att utvidga modellen till att ta med autokorrelationer av högre lag. Generellt skrivs för lag-p, även kallat med order p, AR(p) som

$$r_t = \Phi_0 + \Phi_1 r_{t-1} + \cdots + \Phi_p r_{t-p} + a_t,$$

som kan liknas vid multipel linjär regression.

2.2.5 ARCH

ARCH-modellen är baserad på antagandet om konditionell heteroskedsticitet. För en volatilitetsmodell $r_t = \mu_t + a_t$ antar vi att innovationerna a_t inte är autokorrelerade men beroende i den mening att autokorrelation förekommer för innovationernas kvadrerade värden a_t^2.

I [Tsa10, s.116] skrivs ARCH-modellen som

$$a_t = \sigma_t \epsilon_t, \quad \sigma_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2 + \cdots + \alpha_m a_{t-}^2,$$

där ϵ_t är en sekvens av oberoende likafördelade (iid.) slumpvariabler med väntevärdet 0 och varians 1.

I den här rapporten kommer vi fokusera mer på olika GARCH-modeller vilka är vidareutvecklade ur ARCH-modellen. Vidare beskrivning av ARCH-modellens egenskaper och hur dess parametrar skattas finns i [Tsa10].

2.2.6 GARCH

GARCH-modellen utgår ifrån ARCH-modellen men beskrivs även av tidigare modellerade värden dvs. σ_{t-j}^2. Detta leder i flera fall till en enklare modell med färre parametrar i jämförelse med ARCH-modellen.

Vi definierar GARCH-modellen enligt [Tsa10, s.132] för order (m, s) som

$$a_t = \sigma_t \epsilon_t, \qquad \sigma_t^2 = \alpha_0 + \sum_{i=1}^{m} \alpha_i a_{t-i}^2 + \sum_{j=1}^{s} \beta_j \sigma_{t-j}^2, \tag{3}$$

där ϵ_t är en sekvens av oberoende likafördelade(iid.) slumpvariabler med väntevärde 0 och varians 1, exempelvis standardnormalfördelad eller t-fördelad.

Speciellt skriver vi GARCH(1,1) som

$$\sigma_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2 + \beta_1 \sigma_{t-1}^2, \tag{4}$$

där $m = s = 1$ sätts in i Ekv. (3).

I Value-at-Risk modelleringen som beskrivs i Kap. 2.3 använder vi prediktioner för framtida volatilitet. Betrakta GARCH(1,1)-modellen vid tid $t = h$ dvs. σ_h^2 utifrån vilken

vi vill prediktera variansen för i morgon, σ^2_{h+1}. Vi utgår från Ekv. (4) och skriver 1-dagsprediktionen som

$$\sigma^2_h(1) = \sigma^2_{h+1} = \alpha_0 + \alpha_1 a^2_h + \beta_1 \sigma^2_h. \tag{5}$$

Parameterskattningarna $(\hat{\alpha}_0, \hat{\alpha}_1, \hat{\beta}_1)$ beräknas med olika maximum-likelihood metoder beroende på vilken fördelning vi antar för ϵ_t. Metoderna för paramterskattning i GARCH-modellen och IGARCH-modellen beskrivs i Appendix A.1.

2.2.7 IGARCH

Den integrerade GARCH-modellen, kallad IGARCH är en variant på GARCH där man antar att $\sum_i \alpha_i + \sum_j \beta_j = 1$.

Vi kan skriva IGARCH(1,1)-modellen genom att sätta $\alpha_1 = 1 - \beta_1$ i Ekv. (4) vilket ger

$$a_t = \sigma_t \epsilon_t, \qquad \sigma^2_t = \alpha_0 + \beta_1 \sigma^2_{t-1} + (1 - \beta_1) a^2_{t-1},$$

där ϵ_t är definierat som i GARCH-modellen. Med samma resonemang som för GARCH skriver vi, vid tid h, 1-dagsprediktionen för volatiliteten som

$$\sigma^2_h(1) = \sigma^2_{h+1} = \alpha_0 + (1 - \beta_1) a^2_h + \beta_1 \sigma^2_h.$$

En fördel med IGARCH-modellen är att den ger ett enklare uttryck för flerdagarsperioder dvs. $\sigma^2_h(l) = \sigma^2_h(1) + (l-1)\alpha_0$ för $l \geq 1$. Speciell under antagandet att α_0=0 gäller att $\sigma^2_h(l) = \sigma^2_h(1)$, se [Tsa10, s.141]. Vi kommer att diskutera användningen av IGARCH under detta antagande i Kap. 2.3.3 där vi beskriver RiskMetrics, en väl etablerad modell för Value-at-Risk-modellerig.

2.3 Value-at-Risk

Value-at-risk, vanligen förkortat till VaR, är ett mått på hur stor förlusten för en tillgång kan vara för en viss period givet en viss sannolikhet, p. VaR kan tolkas som att vi med en viss säkerhet kan säga att den potentiella förlusten över en viss period är mindre än VaR, se [Tsa10, s.327].

Vi uttrycker en generell matematisk definition av VaR som

$$VaR_p = min\{m : P(L \leq m) \geq 1 - p\},$$

där L är förlustfunktionen för tillgången. Vi kommer att anta olika fördelningar för förlustfunktionen och betrakta dess kvantiler. Vi kan skriva VaR_p som den $(1-p) : te$ kvantilen för L.

$$VaR_p = F^{-1}_L(1 - p) \tag{6}$$

där F_L är fördelningsfunktionen för L, se [Lin+12, s.166].

Vi kommer framöver betrakta VaR som en procentuell förlust snarare än ett absolut belopp. Utgår vi från Ekv.(6) kan vi skriva

$$VaR_p(r_t) = F^{-1}_{-r_t}(1-p) = -F^{-1}_{r_t}(p), \tag{7}$$

där vi låter L vara lika med $-r_t$.

2.3.1 Empiriska kvantiler

Med empiriska kvantiler kan vi med ett enkelt och icke-parametriskt förfarande räkna ut VaR utan att behöva anta någon fördelning för log-avkastningarna. Vi antar däremot att fördelningen i stickprovsperioden håller över prediktionsperioden.

Vi skriver ordningsstatistikan [4] som

$$r_{(1)} \leq r_{(2)} \leq ... \leq r_{(n)},$$

där $r_{(1)}$ är det minsta av alla r_t och $r_{(n)}$ är det största. Vidare antar vi att avkastningarna r_t är *iid.* med fördelningsfunktionen $F(x)$, då blir ordningsstatistikan normalfördelad enligt

$$r_{(l)} \sim N\Big(x_p, \frac{p(1-p)}{n[f(x_p)]^2}\Big), \quad l = np,$$

där x_p är p-te kvantilen för $F(x)$ och $f(x)$ är täthetsfunktionen för r_t.

Av detta fås att vi kan använda $r_{(l)}$ som skattning för x_p. I fallet då l inte är ett heltal definierar vi de närmsta heltalen $l_1 < l < l_2$ och $p_i = \frac{l_i}{n}$ och skattar x_p enligt

$$\hat{x}_p = \frac{p_2 - p}{p_2 - p_1} r_{(l_1)} + \frac{p - p_1}{p_2 - p_1} r_{(l_2)}.$$

I enlighet med vårt uttryck för VaR på procentuell form i Ekv.(7) kan vi skriva

$$VaR_p(r_t) = -F^{-1}_{r_t}(p) = -\hat{x}_p.$$

2.3.2 Ekonometrisk VaR

Det ekonometriska tillvägagångssättet för beräkning av VaR använder tidsseriemodeller för en avkastningsserie, r_t och dess volatilitet σ_t^2. Vi utgår från predikterad avkastning och volatilitet samt den betingade fördelningen för r_{t+1} vid tid t.

För $r_t = a_t + \mu_t$ och $a_t = \sigma_t \epsilon_t$ skriver vi 1-dags VaR som

$$VaR_p = \hat{\mu}_{t+1} + F^{-1}_{\epsilon_t}(1-p)\sigma_t(1),$$

[4] "Order statistic" på engelska, beskrivs i [Tsa10, s.338].

där är 1-dagspediktionen $\sigma_t(1)$ fås genom exempelvis Ekv.(5). Vi får således olika estimat för VaR beroende på vilka fördelningsantaganden vi gör för ϵ_t.

Under antagandet att ϵ_t är standardiserat t-fördelad med v frihetsgrader gäller

$$VaR = \hat{\mu}_{t+1} + t_v^*(1-p)\sigma_t(1),$$

där $t_v^*(1-p)$ är kvantilerna för den standardiserade t-fördelningen. Låt $t_v(p)$ vara den p-te kvantilen för t-fördelningen, då gäller sambandet

$$t_v^*(1-p) = \frac{t_v(1-p)}{\sqrt{v/(v-2)}},$$

se [Tsa10, s.334]. I vår analys kommer vi betrakta det speciella fallet då $E(r_t) = \mu_t$ inte är autokorrelerad och skriver 1-dagsprediktionen för logavkastningen som $\hat{r}_t(1) = \hat{\mu}$. Notera att om r_t är autokorrelerad bör prediktionen för $\mu_t(1)$ vara baserad på en tidsseriemodell, exempelvis en AR-modell.

2.3.3 RiskMetrics

RiskMetrics är en metod för beräkning av VaR som utvecklades av J.P. Morgan under 90-talet. Med metoden görs antagandet att log-avkastningarna är normalfördelade med väntevärde noll samt att volatiliteten anpassas baserat på IGARCH(1,1)-modellen vi beskriver i Kap. 2.2.7 förutom att RiskMetrics även antar att $\alpha_0 = 0$. Vi modellerar då volatiliteten enligt

$$\mu_t = 0, \quad \sigma_t^2 = \beta_1 r_{t-1}^2 + (1-\beta_1)\sigma_{t-1}^2, \quad \beta_1 \in [0,1],$$

där $r_{t-1} = a_{t-1}$ eftersom $\mu_t = 0$, se[Tsa10, s.329].

Med modellen ovan blir 1-dagsprediktionen

$$\sigma_t^2(1) = \beta_1 r_t^2 + (1-\beta_1)\sigma_t^2$$

vilket under antagandet om normalfördelning och $\mu_t = 0$ ger VaR prediktionen

$$VaR_p = \Phi^{-1}(1-p)\sigma_t(1),$$

för tid $t+1$ där $\Phi^{-1}(1-p)$ är den (1-p)-te standardnormalkvantilen.

2.4 Backtest för VaR

För att undersöka hur korrekta prediktionerna för VaR har varit utför man olika så kallade backtest. Testen går huvudsakligen ut på att i efterhand jämföra antalet tillfällen då VaR har överskridits, dvs. då den verkliga förlusten blev större än predikterad VaR_p, med antalet överskridanden som borde inträffat under perioden enligt sannolikhet p.

Något mer avancerade backtest undersöker hurvida dessa tillfällen är oberoende av varandra och tar hänsyn till flera sannolikheter, p, i samma test, se [Cam05].

Låt tillfällena då VaR överskrids beskrivas av träff-funktionen[5]

$$I_{t+1}(p) = \begin{cases} 1, & \text{om } r_{t+1} \leq -VaR_p \\ 0, & \text{om } r_{t+1} > -VaR_p \end{cases} \tag{8}$$

av vilken vi kan definiera träff-sekvensen $[I_{t+1}(p)]_{t=1}^{t=T}$, en sekvens av ettor och nollor. Vi betraktar träff-sekvensen och testar om den besitter egenskaperna Unconditional Coverage och Independence.

2.4.1 Kupiec's test för Unconditional Coverage

Under Unconditional Coverage gäller att sannolikheten att VaR_p överskrids är p. Uttryckt med träff-funktionen, Ekv.(8), gäller att $P(I_{t+1}(p) = 1) = p$. Vidare gäller att om $P(I_{t+1}(p) = 1) < p$ så överskattas risken med Var_p; Motsvarande underskattas risken om $P(I_{t+1}(p) = 1) > p$.

Kupiec's test använder POF[6] för att testa
$H_0 : P(I_{t+1}(p) = 1) = p$ mot
$H_a : P(I_{t+1}(p) = 1) \neq p$.

$$POF = 2log\Big(\Big(\frac{1-\hat{p}}{1-p}\Big)^{T-I(p)}\Big(\frac{\hat{p}}{p}\Big)^{I(p)}\Big) \sim \chi^2(1),$$

där T är antalet observationer, $I(p) = \sum_{t=1}^{T} I_t(p)$ och $\hat{p} = I(p)/T$.

2.4.2 Christoffersen's test för Independence

Under Independence gäller att information om tidigare överskridanden inte innehåller någon information om kommande överskridanden. Dvs. att två element i träff-funktionen, $(I_{t+j}(p), I_{t+k}(p))$ är oberoende av varandra. Christoffersen(1998) presenterar ett Markov test som testar om sannolikheten för VaR överskridande beror på gårdagens överskridande, dvs. om $(I_{t+1}(p), I_t(p))$ är oberoende eller ej, se [Cam05, s.8].

Betrakta träff-funktionen, Ekv.(8) som en Markov-kedja med stokastisk matris

$$\Pi_1 = \begin{bmatrix} 1-\pi_{01} & \pi_{01} \\ 1-\pi_{11} & \pi_{11} \end{bmatrix},$$

där $\pi_{ij} = P(I_t = j | I_{t-1} = i)$ vilket ger likelihood-funktionen

$$L(\Pi_1; I_1, ..., I_T) = (1-\pi_{01})^{n_{00}}\pi_{01}^{n_{01}}(1-\pi_{11})^{n_{10}}\pi_{11}^{n_{11}},$$

[5] 'Hit function' på engelska, beskrivs i [Cam05, s.3].
[6] Proportion of Failure, se [Cam05, s.5].

där n_{ij} är antalet par observationer med värde i följt av j, se [Chr98, s.845].

Låt motsvarande Markov-kedja under nollhypotesen om oberoende beskrivas av den stokastiska matrisen

$$\Pi_2 = \begin{bmatrix} 1-\pi_2 & \pi_2 \\ 1-\pi_2 & \pi_2 \end{bmatrix},$$

med likeliood-funktionen

$$L(\Pi_2; I_1, ..., I_T) = (1-\pi_2)^{n_{00}+n_{10}} \pi_2^{n_{01}+n_{11}}.$$

Vi kan nu med hjälp av maximum likelihood-skattningarna[7] testa
H_0 : Oberoende, $LR_{ober} = 0$ mot
H_a : Ej oberoende, $LR_{ober} \neq 0$ med

$$LR_{ober} = -2log[L(\hat{\Pi}_2; I_1, ..., I_T)/L(\hat{\Pi}_1; I_1, ..., I_T)] \sim \chi^2(1)$$

2.4.3 Gemensamt test för Coverage och Independence

I [Chr98] presenterar Christoffersen även ett test som tar hänsyn till både Unconditional Coverage och Independence. Här testas nollhypotesen i Unconditional Coverage-testet mot alternativhypotesen i Independence-testet vilket resulterar i likelihood-kvot statistikan

$$LR_{cc} = -2log[L(p; I_1, ..., I_T)/L(\hat{\Pi}_1; I_1, ..., I_T)] \sim \chi^2(2),$$

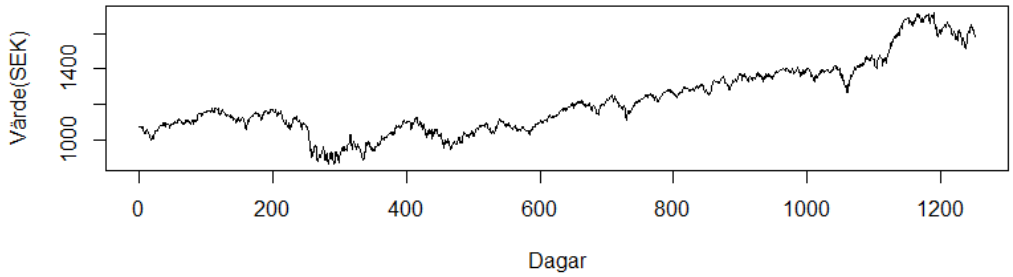
där $L(p; I_1, ...I_T) = (1-p)^{T-I(p)} p^{I(p)}$ från Kupiec's test.

[7] ML-skattningarna härleds i Appendix A.2.

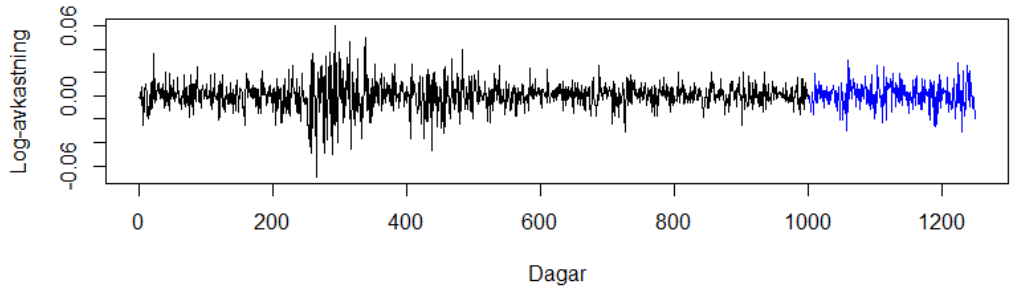
3 Metod

3.1 Data

Vi kommer i våra tillämpningar analysera data från aktieindexet Nasdaq OMX Stockholm 30, hämtat från nasdaqomxnordic.com. Vi betraktar de dagliga stängningspriserna under perioden 2010-08/06 till 2015-07-31 vilken består av 1251 börsdagar. Dessa priser transformeras[8] till 1250 stycken dagliga logaritmerade avkastningar. Priserna och avkastningarna kan ses i Figur 1 och 2.



Figur 1: Värde för OMXS30 under perioden 2010-08-06 till 2015-07-31.



Figur 2: Logaritmerad avkastning för OMXS30 under perioden 2010-08-09 till 2015-07-31. Den blå kurvan representerar de faktiska avkastningarna under perioden för utvärdering.

I den deskriptiva analysen betraktar vi log-avkastningarna för dag 1 till 1000, och gör

[8]Priserna transformeras enligt formeln för log-avkastningar i Kap. 2.1.

antaganden om korrelation samt diverse fördelningar utifrån dessa. När vi modellerar VaR kommer vi göra skattningar för de sista 250 dagarna för att sedan utvärdera dessa skattningar mot de faktiska avkastningarna från dag 1001 till 1250.

3.2 Programvara

För analysen i arbetet använder vi R som är ett populärt programspråk för statistiska beräkningar. R är tillgängligt för alla och kan hämtas gratis på `https://www.r-project.org/about.html`. Vi har även använt oss av RStudio, ett populärt program för programering i R, som finns tillgängligt på `https://www.rstudio.com/products/rstudio/`.

Det är öppet för alla att författa och dela paket med funktioner för R. Vi använder paketen ”fGarch” [WC15] och ”rugarch” [Gha15] som är populära för tidsserieanalys.

3.3 Modeller

Vi kommer att modellera VaR med följande modeller:

- Empiriska kvantiler.
- GARCH(1,1) med normal-fördelade innovationer.
- GARCH (1,1) med t-fördelade innovationer.
- IGARCH(1,1) med de specifika antagandena under RiskMetrics.

De olika modellerna beskrivs i Kap. 2.3. Likheten i modellerna ligger i att de gör en prediktion av VaR för dag $t+1$ baserat på information om avkastningarna t.o.m. dag t.

I vår analys antar vi ett rullande skattningsunderlag på 1000 respektive 500 dagar. Det innebär att vi använder avkastningar 1-1000 för att prediktera VaR för dag 1001, avkastningar 2-1001 för att prediktera VaR för dag 1002 osv för 1000 rullande observationer. Vi använder avkastningar 501-1000 för att prediktera VaR dag 1001 avkastningar 502-1001 för att prediktera VaR dag 1002 osv i fallet av 500 rullande observationer. Sedan utvärderar vi VaR-prediktionerna för de olika modellerna och de olika rullande skattningsunderlagen med backtesting.

Vårt tillvägagångssätt resulterar i 250 prediktioner för VaR vilket på ett ungefär motsvarar ett börsår. Ramverket för backtesting av VaR, se [Ban96], specificerar att backtesting ska implementeras på just 250 dagar.

4 Analys

I denna del presenteras hur analysen genomförts. Vi utgår från det teoretiska ramverket framlagt i Kapitel 2 som vi tillämpar och anpassar efter data. Vi inleder med en deskriptiv analys för förståelse av data och slutsatser som ligger till grund för hur vi modellerar volatiliteten. Sedan anpassar vi våra modeller och predikterar VaR.

4.1 Deskriptiv analys

Inledningsvis är vi intresserade av hur våra log-avkastningar, r_t, är fördelade, främst för att kunna motivera de antaganden vi gör vid GARCH-modelleringen. Eftersom vi ska göra prediktioner för 250 dagar är det rimligt att vi analyserar data under föreliggande dagar. Vi undersöker därför egenskaperna hos avkastningarna för dag 1 till 1000.

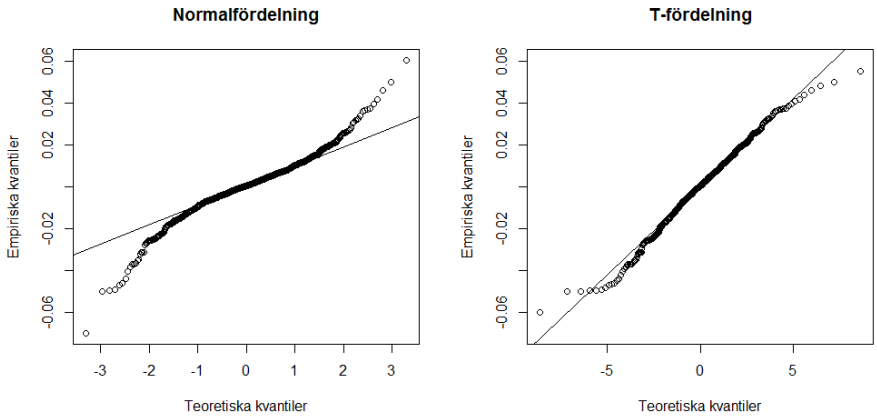
Vi betraktar först stickprovs-medelvärdet, skickprovs-skewness och stickprovs-kurtosis. Vi förväntar oss ett medelvärde nära noll men något positivt eftersom sista priset i perioden är högre än första så att den totala avkastningen för hela perioden är positiv. Skewness beskriver symmetrin i data, exempelvis innebär ett negativt värde på skewness att extrema negativa avkastningar är mer vanligt förekommande än extrema positiva. Kurtosis beskriver hur vanligt förekommande extrema värden är, kurtosis antar värdet 3 för standardnormalfördelningen. I Tabell 1 ser vi grundläggande statistiska värden för log-avkastningarna.

Medelvärde	Standardavvikelse	Skewness	Kurtosis	Excess Kurtosis
0.000267453	0.01223118	-0.3269246	6.514914	3.514914

Tabell 1: Grundläggande statistiska värden för log-avkastningarna, r_t.

Medelvärdet är positivt som förväntat men samtidigt väldigt litet, vi testar framöver om det är signifikant skilt från noll. Vanligtvis antar man att skewness på intervallet $[-0.5, 0.5]$ indikerar symmetrisk data, vi ser även nedan att en modell som antar symmetrisk fördelning fungerar. Tydligast i Tabell 1 är att kurtosis är större än fallet för normalfördelat data; vi behöver rimligtvis anta en fördelning med längre svansar än normalfördelningen.

Vi är främst intresserade av att huruvida vi kan anta att avkastningarna är normal eller t-fördelade vilket avgör hur vi skattar våra parametrar i GARCH-modellen. [Tsa10] skriver att t-fördelningen har visat sig beskriva log-avkastningar bättre än normalfördelningen samtidigt som normalfördelningen ändå antas frekvent i praktiken. Vidare är, enligt [PS07], en t-fördelning med fyra frihetsgrader empiriskt den fördelning som beskriver log-avkastningar för aktieindex bäst generellt sett. För att undersöka fördelningen för avkastningarna jämför vi deras empiriska kvantiler med de teoretiska kvantilerna för standard normal-fördelningen respektive t-fördelningen med fyra frihetsgrader. Detta gör vi i Figur 3 med en QQ-graf för varje fördelning.



Figur 3: QQ-grafer för OMXS30 mot normalfördelning(vänster) och mot t-fördelning med fyra frihetsgrader(höger) med respektive fördelnings teoretiska kvantiler som referenslinje.

Utifrån QQ-graferna drar vi slutsatsen att ingen av fördelningarna beskriver avkastningarna korrekt eftersom de empiriska kvantilerna avviker från de teoretiska, speciellt med avseende på extrema observationer. Däremot är avvikelsen betydligt mindre för t-fördelningen. Vi antar därför att t-fördelningen beskriver vår data bättre än normalfördelningen.

För modellering av volatilitet antar vi att det finns seriell korrelation för kvadrerade avkastningar, r_t^2, som uttrycker sig som volatilitetskluster, dvs. perioder med sammanhängande hög volatilitet. Vidare antar vi att det inte finns någon seriell korrelation för log-avkastning. Om så vore fallet eller om medelvärdet för r_t är signifikant skilt fån noll måste vi korrigera för detta genom att uttrycka en medelvärdesekvation.

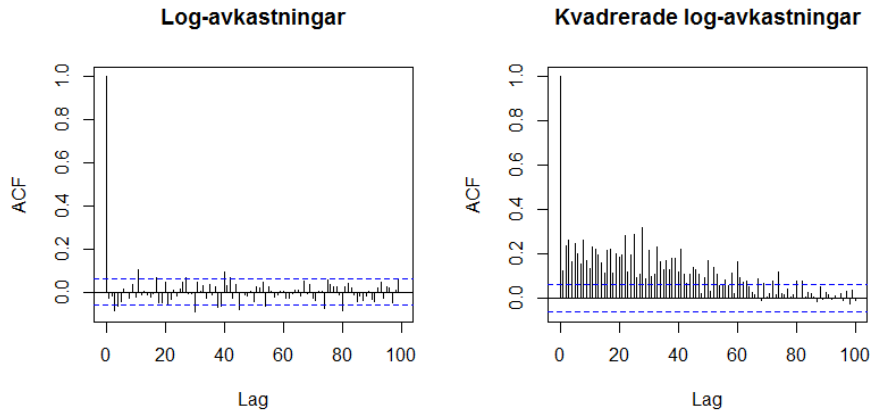
I Tabell 1 framgår att medelvärdet, μ, är större än noll men litet. Vi testar
$H_0 : \mu = 0$ mot
$H_a : \mu \neq 0$
med ett t-test vilket ger

$$t = \frac{\bar{r}_t}{s/\sqrt{n}} = 0.6914789 < t_{0.05}(n-1) \approx 1.96,$$

där s är stickprov-standardavvikelsen och n är antalet observationer. Vi förkastar således inte H_0 och antar att medelvärdet är noll.

Vårt antagande om seriell korrelation bekräftas av att det förekommer volatilitetskluster för r_t vilket är synligt som sammanhängande extrema avkastningar i Figur 2. Vi har visat att μ inte är signifikant skilt från noll och vill även försäkra oss om att det inte finns någon seriell korrelation för r_t. Vidare vill vi bekräfta att volatiliteten är seriellt korrelerad.

Vi gör detta genom att ta fram Autokorrelationfunctionen (ACF) för avkastningarna, r_t, och för r_t^2. ACF uttrycks som $\hat{\rho}_l$ för laggar $l = \{1, ...100\}$ och visas i Figur 4. Om mer än 5% av skattningarna, $\hat{\rho}_l$, ligger utanför det 95%-iga konfidensintervallet tyder det på att seriell korrelation förekommer. Förekomst av volatilitetskluster borde uttrycka tydlig korrelation för de första laggarna för att sedan avta för högre lagg.



Figur 4: ACF för log-avkastningar(vänster) och kvadrerad log-avkastningar(höger) över 100 laggar.

Det framgår att ingen tydlig korrelation förekommer för avkastningarna eftersom de korrelationer vi ser förekommer till synes slumpartat över en period av 100 laggar. Däremot kan vi se att att korrelation förekommer för volatiliteten, långsamt avtagande över 100 laggar.

För att matematiskt säkerställa att detta stämmer gör vi Ljung-Boxtest för log-avkastningarna samt volatiliteten dvs. de kvadrerade log-avkastningarna. Vi testar nollhypotesen att serien inte är korrelerad mot alternativhypotesen att den är det och presenterar resultaten i Tabell 2.

Typ	Q(m)	P-värde
Log-avkastningar	Q(7)=15.0327	0.03558
Volatilitet	Q(7)=296.3906	$< 2.2 \cdot 10^{-16}$

Tabell 2: Ljung-boxtest för $m = 7 \approx log(1000)$ för log-avkastningar samt kvadrerade log-avkastningar med motsvarande p-värde.

Vi kan förkasta nollhypotesen om ingen korrelation för de kvadrerade avkastningarna på alla signifikansnivåer. För avkastningsserien kan vi förkasta nollhypotesen på 5%-ig signifikansnivå men inte på 1%-ig nivå.

För att göra ett antagande om seriell korrelation bör vi ta hänsyn till både ACF-grafen och Ljung-Box testet. I Figur 4 framgår inget tydligt mönster för autokorrelationen, vilket innebär att modelleringen för korrelationen blir komplicerad och svår att motivera teoretiskt. Vi gör antagandet att seriell korrelation för avkastningarna inte håller över kommande period för prediktion. Ställt i relation till att medelvärdet, μ, inte är signifikant skilt från noll gör vi valet att inte specificera en medelvärdesekvation vid GARCH-modelleringen.

Sammanfattningsvis gör vi utifrån den deskriptiva analysen följande antaganden,

- Seriell korrelation förekommer för de kvadrerade log-avkastningarna vilket utgör grunden för modellering av volatiliteten med GARCH.
- Väntevärdet är inte signifikant skilt från noll och det finns ingen seriell korrelation för log-avkastningarna. Vi behöver således inte specificera någon medelvärdesekvation, μ_t, vid GARCH-modelleringen.
- Avkastningarnas fördelning beskrivs bättre av t-fördelningen än normalfördelningen, vilket rimligtvis innebär att en GARCH-modell med t-fördelade innovationer beskriver volatiliteten bättre.

Vi utgår från dessa antaganden och slutsatser vid modelleringen av volatiliteten och VaR. Vi tar även fram modeller som bygger på andra antaganden för jämförelse. Vi modellerar GARCH(1,1) med både normal- och t-fördelade innovationer. IGARCH-modellen anpassar vi utifrån de standardiserade antagandena som görs under ramverket RiskMetrics och VaR-prediktion med empiriska kvantiler gör inga antaganden om fördelning eller autokorrelation.

4.2 GARCH(1,1)

I den här delen använder vi GARCH(1,1)-modellen för att beskriva volatiliteten hos avkastningarna och prediktera 1-dags VaR. Vi kommer använda rullande observationsunderlag för att prediktera volatiliteten en dag framåt och använda dessa prediktioner för att perdiktera 1-dags VaR för sannolikheter 5% och 1%.

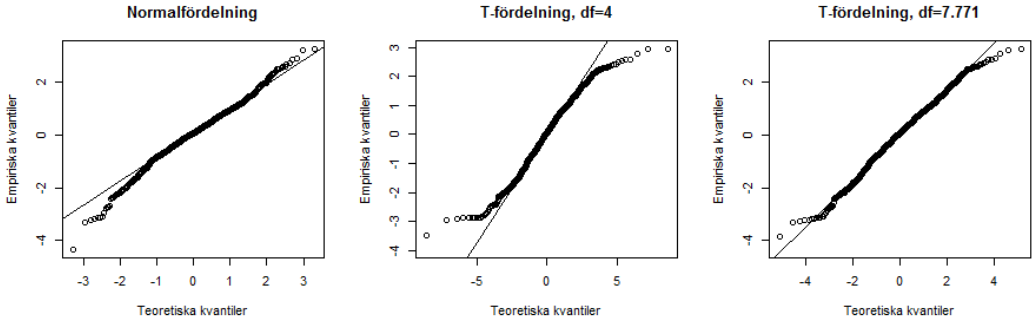
Utifrån de antaganden vi gjort baserat på den deskriptiva analysen specificerar vi modellen. Eftersom vi antar att $\mu_t = 0$ blir $r_t = a_t$ och vi skriver GARCH(1,1) modellen som

$$r_t = \sigma_t \epsilon_t, \qquad \sigma_t^2 = \alpha_0 + \alpha_1 r_{t-1}^2 + \beta_1 \sigma_{t-1}^2,$$

från vilken vi specificerar två modeller, för $\epsilon_t \sim i.i.d.\ N(0,1)$ och för $\epsilon_t \sim i.i.d.\ t(0,1,df)$.

Innan prediktion över rullande fönster vill vi kontrollera om GARCH(1,1)-modellen fångar upp autokorrelationen så att antagande om oberoende för ϵ_t stämmer. Vi anpassar därför modellen för våra första 1000 observationer och undersöker ACF för de standardiserade residualerna, ϵ_t, och ϵ_t^2 i Figur 6. Vi skattar ϵ_t med $\frac{a_t}{\sigma_t} = \frac{r_t}{\sigma_t}$ från den anpassade GARCH-modellen.

För modellering med t-fördelade ϵ_t kan vi välja att specificera antalet frihetsgrader i förhand eller estimera det[9]. Vi anpassar modellen baserat på de olika fördelningsantagandena för ϵ_t och undersöker deras kvantiler mot respektive teoretiska kvantiler med QQ-grafer i Figur 5. Baserat på antagandet om fyra frihetsgrader för avkastningsserien testar vi detta antagande även för ϵ_t.

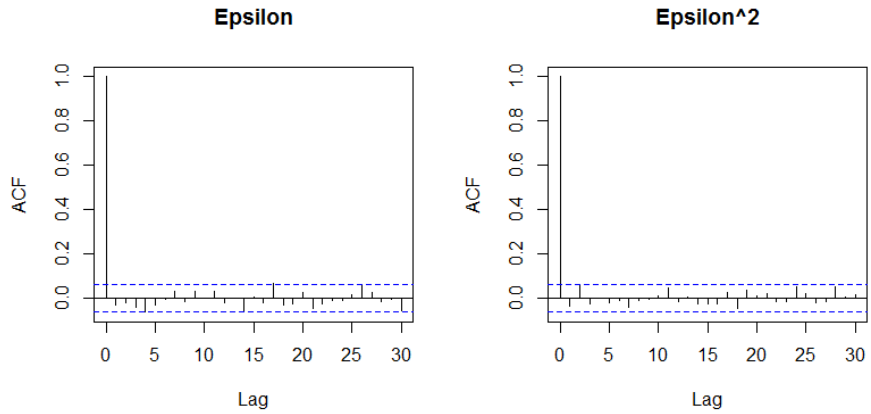


Figur 5: QQ-grafer för de standardiserade residualerna för respektive fördelningsantagande. Residualernas kvantiler jämförs med motsvarande teoretiska kvantiler. Normalfördelade (vänster), t-fördelade med df=4 (mitten) och t-fördelade med estimerade frihetsgrader df=7.771 (höger).

Vi ser att antagandet att ϵ_t är normalfördelat ger standardiserade residualer som följer den antagna normalfördelningen. Det samma gäller för t-fördelningen med estimerat antal frihetsgrader medan antagandet om fyra frihetsgrader inte håller lika väl. Framöver modellerar vi GARCH(1,1) med normalfördelade ϵ_t samt t-fördelade med estimerat antal frihetsgrader.

För att kontrollera att GARCH(1,1) fångar upp heteroskedasticiteten i avkastningarna tar vi fram ACF för de standardiserade residualerna och undersöker om de kan antas vara oberoende likafördelade, dvs. ingen korrelation förekommer för ϵ_t eller ϵ_t^2. Vi väljer godtyckligt att testa för antagandet $\epsilon_t \sim N(0,1)$. ACF visas i Figur 6 nedan.

[9]Estimering av frihetsgrader härleds i Appendix A.1.3.



Figur 6: ACF för ϵ_t(vänster) och ϵ_t^2(höger) över 30 laggar.

Vi ser inga signifikanta korrelationer då ACF ligger inom intervallet över alla 30 laggar. Vi kan således dra slutsatsen att vi lyckats modellera heteroskedastisiteten och får oberoende standardiserade residualer som resultat.

I Tabell 3 undersöker vi för GARCH(1,1) med normalfördelade respektive t-fördelade ϵ_t parameterskattningarna[10] samt AIC och BIC[11].

	T-fördelning		Normal-fördelning	
	Estimat	**P-värde**	**Estimat**	**P-värde**
α_0	$1.610{\cdot}10^{-06}$	0.0475	$1.376{\cdot}10^{-06}$	0.0294
α_1	$6.550{\cdot}10^{-02}$	$1.84{\cdot}10^{-05}$	$5.580{\cdot}10^{-02}$	$2.5{\cdot}10^{-07}$
β_1	$9.227{\cdot}10^{-01}$	$< 2{\cdot}10^{-16}$	$9.328{\cdot}10^{-01}$	$< 2{\cdot}10^{-16}$
AIC	-6.267989	-	-6.247497	-
BIC	-6.248358	-	-6.232774	-

Tabell 3: Parameterskattningar, AIC och BIC för GARCH(1,1) modellen med t- respektive normal-fördelade innovationer. μ_t är antaget som 0.

Vi noterar att AIC och BIC är marginellt lägre för t-fördelade än för normalfördelade ϵ_t vilket tyder på att modellen under t-fördelning anpassar data något bättre. Värt att notera är att α_0 har betydligt högre p-värde än de andra parametrarna och är samtidigt väldigt nära noll. Vi behåller α_0 som parameter i GARCH(1,1)-modellen men antar att den är noll i IGARCH-modellen för RiskMetrics. IGARCH-modellen antar även att $\alpha_1 + \beta_1 = 1$ vilket vi kommer ganska nära med skattningarna under normal-fördelade ϵ_t, $\hat{\alpha}_1 + \hat{\beta}_1 = 0.9886$.

[10] Parameterskattningara gör med Maximum Likelihood-metoden vilken härleds i Appendix A.1.
[11] AIC och BIC förklaras i Appendix A.3.

Vi skattar 1-dags VaR_p med 1000 respektive 500 rullande observationer. Vi gör detta för GARCH(1,1) med normal- och t-fördelade innovationer samt för sannolikheterna p=0.01 och p=0.05.

Vi utgår från dag 1000 och predikterar VaR för dag 1001, sedan utgår vi från dag 1001 när vi predikterar för dag 1002 osv. Detta resulterar i 250 stycken prediktioner för VaR baserat på de senaste 1000 respektive 500 rullande observationerna. Resultaten över 250 dagar visas i jämförelse med faktiska avkastningar i Figur 7. Eftersom vi vill se hur VaR prediktionerna följer negativa avkastningar plottar vi negativa VaR i figuren.

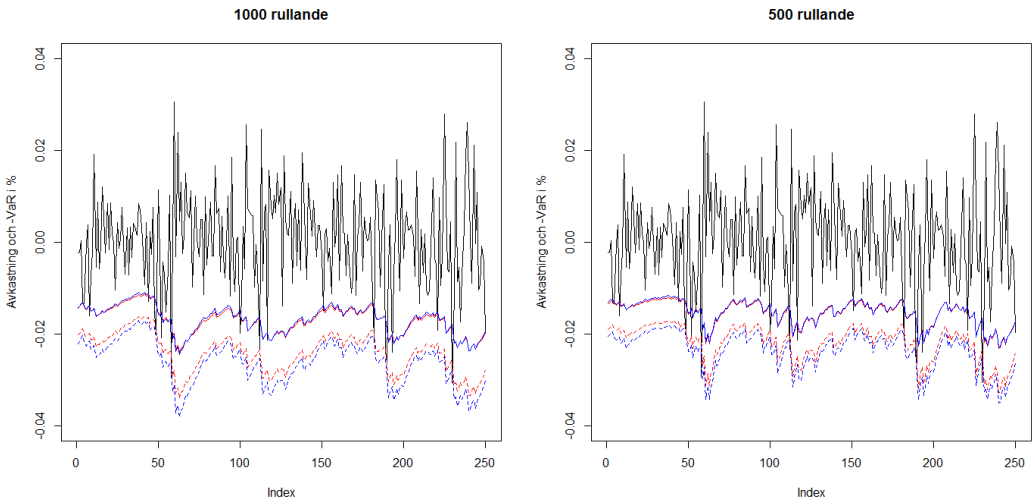
Under antagandet att $\mu_t = 0$ skattar vi VaR_p enligt följande för normalfördelade ϵ_t:

$$VaR_p = \Phi^{-1}(1-p)\sigma_t(1),$$

där $\Phi^{-1}(1-p)$ är den $1-p$-te kvantilen för normalfördelningen. För t-fördelade ϵ_t skriver vi

$$VaR_p = t_v^*(1-p)\sigma_t(1),$$

där $t_v^*(1-p)$ är den $1-p$-te kvantilen för den standardiserade t-fördelningen med v frihetsgrader. Antalet frihetsgrader skattas om för varje prediktion precis som de andra parametrarna i modellen.



Figur 7: 5%-ig(hela linjer) resp. 1%-ig(streckade linjer) VaR-prediktioner för normalfördelade(röd) och t-fördelade(blå) innovationer. För 1000(vänster) och 500(höger) rullande observationer.

Vi ser att VaR generellt sett skattas högre för normalfördelade innovationer än t-fördelade på 5%-nivån men tvärt om på 1%-nivån. Vi noterar skillnaden för VaR under de två fördelningsantagandena är större för 1000 observationer än för 500. Vi utvärderar modellernas förmåga att prediktera VaR med backtest i Kapitel 5.1.

4.3 IGARCH(1,1) för RiskMetrics

IGARCH(1,1) är reducerad GARCH(1,1)-modellen som under ramverket för RiskMetrics reduceras ytterligare. De antaganden vi gör är $\mu_t = 0$,$\alpha_0 = 0$, $\alpha_1+\beta_1 = 1$ och $\epsilon_t \sim N(0,1)$ vilket ger modellen

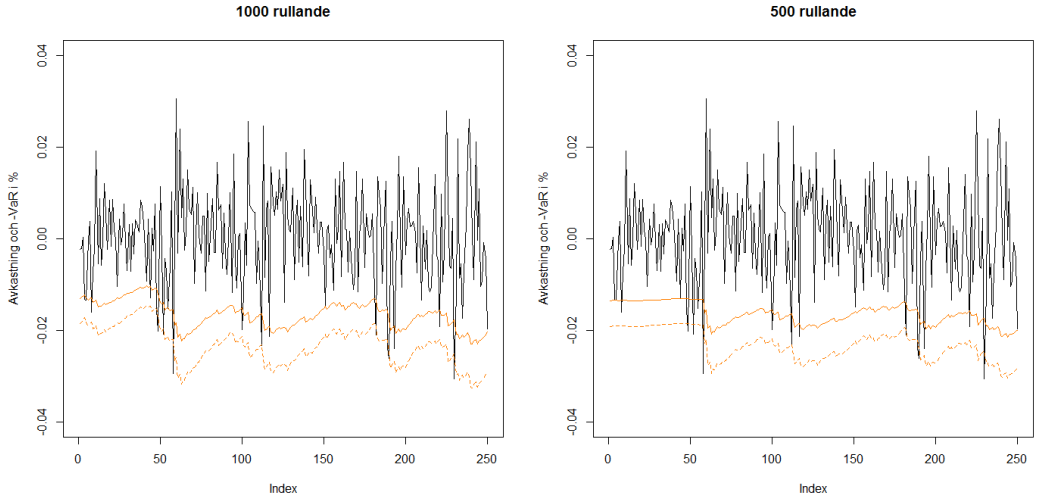
$$r_t = \sigma_t \epsilon_t, \quad \sigma_t^2 = \beta_1 r_{t-1}^2 + (1-\beta_1)\sigma_{t-1}^2, \quad \beta_1 \in [0,1].$$

Lämpligheten för dessa antaganden för vår data diskuteras i föregående del, Kapitel 4.2. Vi predikterar 1-dags VaR enligt

$$VaR_p = \Phi^{-1}(1-p)\sigma_t(1),$$

där $\Phi^{-1}(1-p)$ är den $1-p$-te kvantilen för normalfördelningen.

Som tidigare predikterar vi VaR_p över 250 dagar baserat på 1000 respektive 500 rullade observationer och jämför de negativa värdena på VaR mot de faktiska avkastningarna för de 250 dagarna. Vi tar fram VaR_p för sannolikheterna $p = 0.05$ och $p = 0.01$; resultaten presenteras i Figur 8.



Figur 8: 5%-ig(hela linjer) resp. 1%-ig(streckade linjer) VaR-prediktioner med RiskMetrics. För 1000(vänster) och 500(höger) rullande observationer.

Modellernas förmåga att prediktera VaR utvärderas med backtest i Kapitel 5.1.

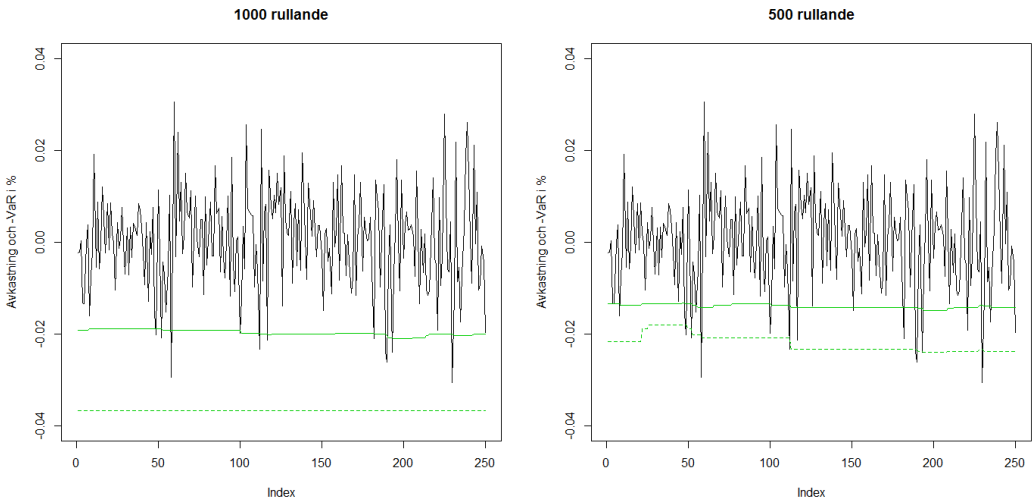
4.4 Empiriska kvantiler

VaR estimering med empiriska kvantiler skiljer sig markant från de varianter av GARCH-modellen vi använder ovan. Vi antar enkelt att fördelningen för r_{t+1} är densamma som

tidigare observerade r_t för vilka vi tar fram empiriska kvantiler, x_p. VaR kan då uttryckas enligt

$$VaR_p(r_t) = -F_{r_t}^{-1}(p) = -\hat{x}_p.$$

Precis som tidigare predikterar vi negativa VaR för 250 dagar baserat på 1000 respektive 500 rullande observationer jämfört med faktiska avkastningar i Figur 9.



Figur 9: 5%-ig(hela linjer) resp. 1%-ig(streckade linjer) VaR-prediktioner med Empiriska Kvantiler. För 1000(vänster) och 500(höger) rullande observationer.

Vi ser tydlig skillnad från GARCH-modellernas VaR-prediktioner i hur väl prediktionerna följer volatiliteten. Detta är inte oväntat eftersom vi inte har modellerat för heteroskedasticitet här, dvs. vi tar inte hänsyn till volatilitets-kluster när vi modellerar VaR med empiriska kvantiler. Prediktionerna ser ut att i vissa fall överskatta VaR och i vissa fall underskatta, beroende på antalet rullande observationer vi betraktar och vilken sannolikhet, p, vi testar för. Vidare utvärdering av prediktionsförmågan gör i nästa del, Kapitel 5.1.

5 Resultat och Slutsatser

5.1 Backtesting

I den här delen utvärderar vi de olika modellerna för prediktion av VaR i Kapitel 4.2 till 4.4. Vi kommer att använda Kupiec's test för Unconditional Coverage och Christoffersen's test för Coverage och Independence.[12] Kort beskrivet testar Kupiec's test huruvida VaR_p-prediktionerna överskrids av negativa avkastningar i $(p \cdot 100)\%$ av fallen. Christoffersen's test för Independence testar huruvida överskridande av VaR dag t är beroende av överskridande dag $t-1$. Christoffersen's test för Coverage och Independence är ett sammanslaget test för de två testen.

Med Kupiec's test testar vi
$H_0 : VaR_p-$prediktionerna överskrids av negativa avkastningar i $(p \cdot 100)\%$ av fallen.
Med Christoffersen's test testar vi
$H_0 : VaR_p-$prediktionerna överskrids av negativa avkastningar i $(p \cdot 100)\%$ av fallen och överskridanden av VaR dag t är oberoende av överskridande dag $t-1$.

P-värdena för de två testen presenteras i Tabell 4. De fyra olika modellerna baserade på 500 respektive 1000 rullande observationer testas för förmåga att prediktera 1-dags VaR_p för $p = 5\%$ samt $p = 1\%$. Vi presenterar även andelen överskridelser uttryckt i procent för jämförelse med den p-% vi eftersträvar. Exempelvis för $VaR_{5\%}$ undersöker vi om andelen överskridelser också är 5% eller om skillnaden är signifikant i den meningen att vi kan förkasta H_0.

Vi betraktar p-värdena för testen, där höga p-värden indikerar en bra modell. P-värden lägre än 0.05 innebär att vi förskastar respektive nollhypotes som beskrivs ovan. Vi får en del odefinierade p-värden vilket sker vid för få observationer av vissa slag. Specifikt vid kupiec's test innebär det att vi inte har några överskridelser alls; vid Christoffersen's test kan det även innebära att överskridelser två dagar i rad inte inträffar. Om överskridande aldrig inträffar indikerar det att risken överskattas av modellen. Om två överskridanden i rad inte inträffar indikerar det oberoende mellan överkridanden.

[12] Kupiec's och Christoffersen's tester beskrivs i Kapitel 2.4.

Modell	rullande obs.	VaR_p p	Kupiec's överskrid.(%)	Kupiec's p-värde	Christoffersen's p-värde
GARCH-N	500	5%	8%	0.044	0.066
		1%	3.2%	0.005	NaN
	1000	5%	6.4%	0.329	0.406
		1%	2%	0.162	NaN
GARCH-t	500	5%	8%	0.044	0.030
		1%	3.6%	0.001	0.003
	1000	5%	7.2%	0.133	0.254
		1%	0.8%	0.742	NaN
Risk-Metrics	500	5%	6.4%	0.329	0.406
		1%	3.2%	0.005	0.010
	1000	5%	7.6%	0.079	0.098
		1%	3.6%	0.001	0.0005
Empiriska kvantiler	500	5%	7.6%	0.079	0.183
		1%	3.2%	0.005	0.010
	1000	5%	4.4%	0.657	0.717
		1%	0%	NaN	NaN

Tabell 4: Resultat vid backtesting av 250 VaR-prediktioner. Vi testar 4 olika modeller baserat på 500 respektive 1000 rullande observationer och predikterar $VaR_{5\%}$ och $VaR_{1\%}$. Andelen VaR-prediktioner som överskrids av negativa avkastningar uttrycks i procent. P-värden anges för Kupiec's test samt Christoffersen's sammanslagna test, där högt p-värde är önskvärt.

Resultatet är någorlunda otydligt, vi gör totalt 32 test för olika kombinationer av modell, antal observationer och sannolikhet, p. Ingen modell presterar betydligt bättre än de andra. Det är däremot tydligt att våra prediktioner generellt sett presterar bättre med 1000 rullande observationer än 500 rullande observationer vilket gäller för alla modeller utom RiskMetrics.

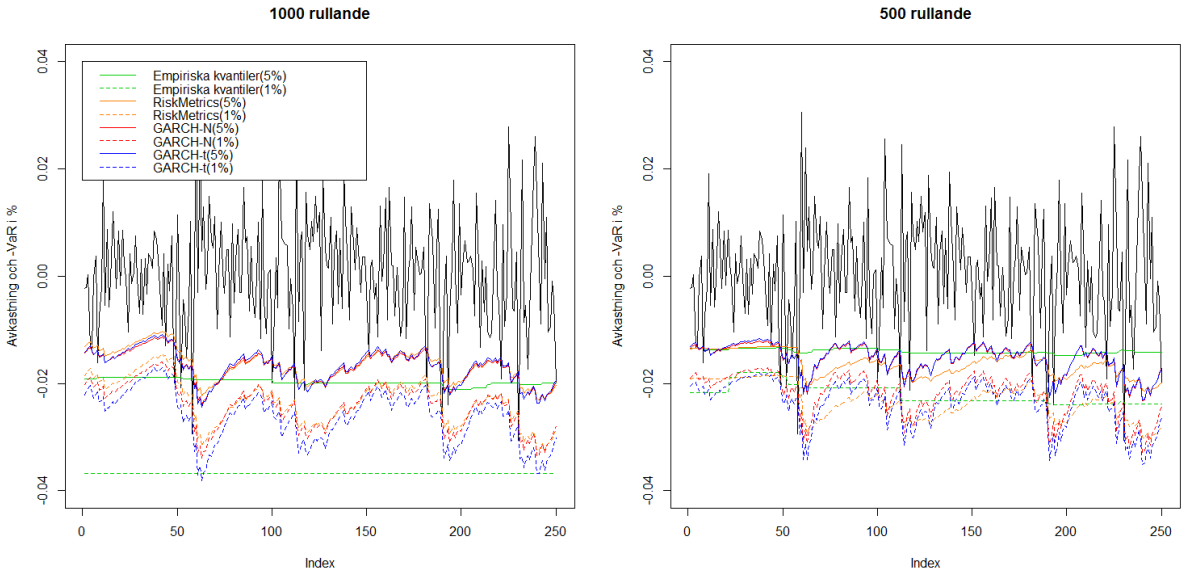
Värt att notera är även att de flesta modeller underskattar risken dvs. andelen övserskridanden är högre än motsvarande teoretiska. Undantagsvis överskattas risken för modellen empiriska kvantiler med 1000 observationer samt GARCH-t med 1000 observationer för $VaR_{1\%}$. Det är även i dessa fall prediktionerna presterar som bäst med höga p-värden.

Något oväntat får vi ganska bra resultat för modellen Empiriska kvantiler. Rent teoretiskt är det vår sämsta modell, där volatilitetskluster inte tas hänsyn till. Därav borde överskridanden också ske i kluster vilket vi kan se i Figur 9, speciellt för 500 observationer. Detta lyckas inte Christoffersen's test fånga upp, testets lämplighet diskuteras vidare i Kapitel 6.

För att dra slutsatser om huruvida en modell lyckas prediktera VaR väl bör vi därför inte enbart utgå från resultaten i backtesten utan även ta hänsyn till hur prediktionerna

följer de negativa avkastningarna. Vilken modell som bäst modellerar Value-at-Risk bör undersökas genom att betrakta resultaten vid backtesting i relation till den teoretiska bakgrunden och de resultat vi sett under analysens gång.

I Figur 10 kan vi se VaR-prediktionerna för alla modeller i samma graf för 1000 respektive 500 rullande observationer.



Figur 10: VaR-prediktioner för alla modeller.

I backtestet får modellen Empiriska kvantiler med 1000 observationer högst p-värde för $VaR_{5\%}$ med andelen överskridelser 4.4%. Vi skulle om vi enbart tittar på resultaten från backtestet kunna påstå att Empiriska kvantiler är bäst för att modellera $VaR_{5\%}$. Däremot ser vi i Figur 10 att VaR prediktionerna för Empiriska kvantiler ständigt ligger på samma nivå. Prediktionerna har en bra teckningsgrad men överskattar risken i perioder av lägre volatilitet.

För 500 rullande observationer beter sig prediktionerna med Empiriska metoder på ett motsatt sätt. Prediktionerna ser ut att skatta risken rätt under perioder av lägre volatilitet men hänger inte med i perioder av högre volatilitet vilket även ger utslag i backtestet där vi kan förkasta att andelen överskridelser ligger på 5%. Vi drar slutsatsen att modellen Empiriska kvantiler är sämre på att prediktera VaR än de andra modeller vi undersöker.

Eftersom modellen empiriska kvantiler inte modellerar för konditionell heteroskedasticitet kan vi anta att skillnaden i modellen för 500 respektive 1000 rullande observationer förklaras av variansen i tidigare data. Det kan ses i Figur 2 att avkastningarna är mer volatila under perioden dag 1-1000 än dag 500-1000.

Denna skillnad i periodernas volatilitet förklarar rimligtvis även varför GARCH-modellerna presterar sämre med 500 rullande observationer, vi drar alltså inte slutsatsen att 500 observationer är för få för att anpassa en GARCH-modell utan att de resultaten i backtesten förklaras av volatiliteten i just vår data. Vi noterar att andelen överskridanden tenderar att vara högre än motsvarande teoretiska. Teoretiskt innebär det att fördelningsantagandena vi gör inte tillräckligt beskriver de lägre kvantilerna i data, dvs. att extrema negativa avkastningar är mer vanligt förekommande än vi har modellerat för.

Även för 1000 rullande observationer ser det ut som att andelen överskridanden är högre än motsvarande teoretiska. Enligt [Tsa10] är det vanligt förekommande att finansdata har tjockare svansar än t-fördelningen. I den deskriptiva analysen antar vi att t-fördelningen med fyra frihetsgrader beskriver avkastningarna bättre än normalfördelningen och den volatilitetsmodell som tar störst hänsyn till detta är GARCH-modellen med t-fördelade innovationer. Eventuellt måste volatilitetsmodelleringen ta ytterligare hänsyn till extrema avkastningar.

5.2 Kvartalsvis Backtesting

En nackdel med att utvärdera volatilitetsmodellerna utifrån backtesten enligt ovan är att vi enbart betraktar ett test per modell och gör detta för en specifik period. Det vill säga att vi endast gör prediktioner och test för perioden dag 1001-1250. Resultaten riskerar att ge en felaktig bild av modellernas lämplighet att prediktera VaR på grund av vad som händer just under perioden vi testar för.

För att undkomma denna problematik utför vi backtest kvartalsvis med 500 rullande observationer. Kvartalsvis innebär att vi börjar med att backtesta 250 prediktioner för dag 501-750; sedan hoppar vi fram 62 dagar (ett kvartal) och backtestar 250 prediktioner från dag 563-812 och så vidare. Förfarandet resulterar i 9 stycken överlappande perioder om 250 dagar för vilka vi utför prediktioner med 500 rullande observationer och sedan backtestar.

Vi utför Kupiec's test för de fyra volatilitetsmodellerna och presenterar i Tabell 5. för 9 test den genomsnittliga andelen överskridanden och hur många av de 9 testen som resulterar i p-värde över 0.05.

Modell	VaR_p p	Kupiec's medel.överskrid.	p-värde>0.05
GARCH-N	5%	6.222%	9/9
	1%	1.689%	8/9
GARCH-t	5%	6.222%	9/9
	1%	1.156%	9/9
RiskMetrics	5%	6.267%	9/9
	1%	2%	8/9
Empiriska kvantiler	5%	2.978%	4/9
	1%	0.844%	3/9

Tabell 5: 9 stycken backtest utförs kvartalsvist för respektive modell enligt ovan beskrivning på 250 VaR_p-prediktioner för $p = 5\%$ resp. 1%. Genomsnittlig procentuella överskridanden för de 9 testen presenteras i tabellen. Antal test som resulterar i p-värde över 5% presenteras också för resp. modell.

Resultaten för de kvartalsvisa backtesten i Tabell 5 ger en tydligare bild än backtesten för en enstaka period. Främst ser vi tydlig skillnad mellan de ekonometriska modellerna och Empiriska kvantiler. Medelvärdet av andelen överskridanden avslöjar att modellen tenderar att överskatta risken och skillnaden är signifikant i fler än hälften av testen.

Eftersom Empiriska kvantiler modellen har en genomsnittlig överskridelseandel som är lägre än p kan vi dra slutsatsen att i de flesta av de 9 testen är skattingsunderlaget mer volatilt än data i periden vi predikterar VaR för.

De ekonomteriska modellerna klarar testet samtliga gånger för $VaR_{5\%}$, och nästan alla gånger för $VaR_{1\%}$ där GARCH-t modellen presterar bäst.

Vi observerar också att de ekonometriska modellerna har högre genomsnittlig andel överskridelser än motsvarande teoretiska. Rimligtvis kan detta innebära att trotts att vi predikterar VaR över en mindre volatil period än skattningsunderlaget så tenderar de ekonometriska modellerna att underskatta volatiliteten.

En nackdel med vårt kvartalsvisa förfarande är att perioderna överlappar varandra och således blir resultaten i backtesten beroende av varandra. Vidare har vi kvar problematiken att vi godtyckligt valt en period att testa för, vilket i kombination med att perioderna ej är oberoende gör att slutsatserna om hur väl modellerna predikterar VaR blir osäkra och möjligtvis felaktiga.

5.3 Backtesting på simulerad data

För att undkomma problematiken med att backtesten är gjorda över en godtyckligt vald period kan vi simulera data på vilken vi predikterar VaR och gör backtest. Många

upprepade simuleringar och test enligt en bootstrapmetod[13] resulterar i att vi kan uppskatta precisionen i backtesten. Vi kommer här utgå från den generella beskrivningen av bootstrap-metoder i [HS14, s.65-70] samt en beskrivning av bootstrapping med avseende att uppskatta precisionen i VaR-backtesting i [Dow02]. Vi väljer här att begränsa undersökningen till GARCH-modellen med t-fördelade innovationer.

Vi utför simuleringar, modellanpassning, VaR-prediktion och Backtest enligt en för ändamålet utformad bootstrap-metod. Vi beskriver vår metod med följande steg.

1. Simulera 250 dagar avkastningsdata baserat på GARCH-t modellen med skattade parametrar baserat på vår data. Standardiserade residualer, ϵ_t simuleras enligt standardiserad t-fördelning och sätts in i vår skattade GARCH-t modell.
2. Prediktera VaR över de 250 simulerade datapunkterna med 500 rullande observationer på samma sätt som gjorts tidigare.
3. Jämför predikterad VaR_p med simulerad avkastningsdata och utför Kupiec's test för $p = 5\%$ och $p = 1\%$.
4. Upprepa steg 1-3, 1000 gånger.

Metoden som beskrivs ovan producerar 1000 resultat från backtest för vilka vi kan undersöka osäkerheten genom att ta fram ett enkelt konfidensintervall baserat på resultatens kvantiler.

Det innebär att vi för t.ex. Kupiec's test för $VaR_{5\%}$ noterar andelen överskridanden för varje test på simulerad data. Vi har då för 1000 test 1000 uppskattningar av andelen överskridelser som, om modellen predikterat $VaR_{5\%}$ väl, ligger i närheten av 5%. Medelvärdet och konfidensintervallet av dessa uppskattningar ger oss en uppfattning om förväntad andel överskridelser och hur säkra vi kan vara på denna förväntning.

Vi kan även på samma sätt ta fram p-värdena för de 1000 testen på simulerad data och undersöka deras medelvärde och konfidensintervall. Detta ger oss en uppfattning om hur VaR-prediktionerna kan förväntas prestera i backtesten. Om hela konfidensintervallet för p-värdena ligger under 0.05 kan vi med 95% säkerhet säga att VaR-modellen kommer underkännas i testet. Analogt om hela konfidensintervallet ligger över 0.05 kan vi med 95% säkerhet säga att VaR-modellen kommer att klara Kupiec's test.

Medelvärden samt konfidensintervall för resultaten av de 1000 backtesten för simulerad data presenteras i Tabell 6. Vi betraktar andelen vid de 1000 testen andelen överskridanden samt motsvarande p-värden för Kupiec's test.

[13] Bootstrapping är en så kallad Monte Carlo-metod som används för att beskriva fördelningen hos en parameter baserat på observationer från många upprepade simuleringar.

VaR_p	Mått	$q_{0.025}$	Medelvärde	$q_{0.975}$
$VaR_{5\%}$	överskr.	3.6%	6.4812%	9.6%
	p-värde	0.00289	0.3502	0.8853
$VaR_{1\%}$	överskr.	0.4%	1.5304%	3.2%
	p-värde	0.0003	0.3964	0.7580

Tabell 6: Medelvärden samt kvantiler för resultaten av de 1000 backtesten för simulerad data. Kvantilerna utgör konfidensintervallen för våra andelar överskridanden och motsvarande p-värden. Resultaten tas fram för backtesting av $VaR_{5\%}$ samt $VaR_{1\%}$.

Vi får med vårt bootstrap-förfarande medelvärden för andelen överskridanden som tyder på att modellen underskattar risken i data. Däremot ligger en stor osäkerhet i att dra den slutsatsen eftersom konfidensintervallen inkluderar överskridelse-andelar både över och under motsvarande teoretiska under nollhypotesen.

Om vi studerar konfidensintervallen för p-värdena ser vi de innehåller både värden under 0.05 och värden över 0.05. Baserat på detta kan vi alltså inte med 95% säkerhet dra slutsatsen att Kupiec's test kommer att acceptera modellen, men inte heller kan vi dra slutsatsen att modellen kommer förkastas.

5.4 Slutsatser

Sammanfattningsvis har vi utvecklat resonemangen kring resultaten i backtestingen i tre steg. Först utfördes test vid en enstaka period utifrån vilka vi ville avgöra om volatilitetsmodellerna predikterar VaR väl. För modellen GARCH-t ser vi i Tabell 4. att andelen överskridelser är 8% vilket är signifikant skilt från 5%. Från det enstaka testet skulle vi kunna dra en felaktig slutsats att GARCH-t modellen underskattar risken vid prediktion av $VaR_{5\%}$.

Vi gjorde sedan kvartalsvisa backtest för 9 överlappande perioder där GARCH-t modellens andel överskridanden inte var signifikant skild från noll i något test. Det motsäger resultatet vid det enstaka testet vilket gör det osäkert att avgöra huruvida modellen lyckas prediktera VaR.

Osäkerheten bekräftas sedan av de spridda resultaten vid backtesting på simulerad data. Baserat på hur våra simuleringar utformats och tolkats blir vår sista slutsats blir att vi inte med säkerhet kan avgöra huruvida GARCH-t modellen predikterar VaR väl baserat på Kupiec's test med 250 prediktioner.

6 Diskussion

Slutsatsen vi gör i sammanfattningen innebär att vi förkastar backtesting med 250 prediktioner, den metod som följer ramverket i [Ban96], som otillräcklig för att avgöra huruvida en modell kan prediktera VaR väl. Detta är ett ganska starkt antagande som rimligtvis behöver stärkas av en mer genomgående analys av bootstrap-metoden vi har använt. Det kan finnas flera möjliga felaktigheter eller otillräckligheter i hur simulering utförts som bör kommenteras.

Vi har valt att simulera data från fördelning, så kallad parametrisk bootstrap. Det går även att använda ickeparametrisk bootstrap, där vi plockar från referensdata med återläggning. Valet har gjorts godtyckligt utan att diskutera hur det påverkar resultaten och hur konfidensintervallet ska utformas.

Det fins flera sätt att utforma ett bootstrap-konfidensintervall, vi har valt det enklaste där vi endast använder kvantilerna för bootstrap-resultaten. I både [HS14] och [Dow02] beskrivs hur bias i resultaten kan resultera i ett dåligt kalibrerat konfidensintervall, något som vi inte tar hänsyn till i det här arbetet.

Backtesten på simulerad data i det här kandidatarbetet har inte varit det huvudsakliga syftet utan har kommit som följd av att försöka tolka resultaten från backtesten på observerad data. Resultaten och slutsatserna därur bör inte antas med största säkerhet men ligger till grund för vidare analys.

I den deskriptiva analysen avgörs vilka volatilitetsmodeller vi väljer, vilka parametrar vi antar vara signifikant skilda från noll och huruvida vi ska specificera en medelvärdeskvation eller inte. Syftet med arbetet är att undersöka etablerade modellers förmåga att prediktera VaR, snarare än att anpassa den ultimata modellen. Vi begränsar oss därför till att göra analysen på en specifik tidsperiod utan att ta hänsyn till hur exempelvis Ljung-Box testets resultat kan variera över tid.

Christoffersen's gemensamma test för Coverage och Independence utförs i arbetet för att undersöka om det finns ett beroende mellan överskridelser. Testet undersöker egentligen bara om det finns beroende mellan överskridande idag och överskridande igår över en tidsperiod. Beroende kan förekomma på annat sätt, exempelvis kan det finnas ett beroende mellan överskridanden med en veckas mellanrum vilket i så fall inte skulle upptäckas av testet vi utfört. Det finns mer avancerade test för att undersöka beroende, bla. har Christoffersen presenterat ett test där man undersöker om tiden mellan två överskridelser är oberoende av tiden sen den senaste överskridelsen.

Referenser

[Ban96] Basle Committee on Banking Supervision. *Supervisory framework for the use of 'backtesting' in conjunction with the internal models approach to market risk capital requirements*. 1996. URL: http://www.bis.org/publ/bcbs22.htm.

[Chr98] Peter F. Christoffersen. "Evaluating Interval Forecasts". I: *International Economic Review* 39.4 (1998), s. 841–862.

[Dow02] Kevin Dowd. *A bootstrap back-test*. 2002. URL: http://www.risk.net/data/Pay_per_view/risk/technical/2002/1002_dowd.pdf.

[Cam05] Sean D. Campbell. "A review of backtesting and backtesting procedures". I: *Finance and Economics Discussion Series Divisions of Research & Statistics and Monetary Affairs Federal Reserve Board, Washington, D.C.* (2005). DOI: http://www.federalreserve.gov/pubs/feds/2005/200521/200521pap.pdf.

[PS07] E. Platen och R. Sidorowicz. *Empirical Evidence on Student-t Log-Returns of Diversified World Stock Indicies*. 2007. URL: http://www.qfrc.uts.edu.au/research/research_papers/rp194.pdf.

[Tsa10] Ruey S. Tsay. *Analysis of Financial Time Series, Third Edition*. Wiley Series in Probability and Statistics. John Wiley & Sons, Inc., 2010. ISBN: 9780470414354.

[Lin+12] F. Lindskog m. fl. *Risk and Portfolio Analysis, Principles and Methods*. Springer Series in Operations Research and Financial Engineering. Springer New York Heidelberg Dordrecht London, 2012. ISBN: 978-1-4614-4102-1.

[HS14] L. Held och D. Sabnés Bové. *Applied Statistical Inference, Likelihood and Bayes*. Springer New York Heidelberg Dordrecht London, 2014. ISBN: 978-3-642-37886-7.

[Gha15] Alexios Ghalanos. *(rugarch) Univariate GARCH Models*. 2015. URL: https://cran.r-project.org/web/packages/fGarch/fGarch.pdf.

[WC15] Diethelm Wuertz och Yohan Chalabi. *(fGarch) Rmetrics - Autoregressive Conditional Heteroskedastic Modelling*. 2015. URL: https://cran.r-project.org/web/packages/fGarch/fGarch.pdf.

A Appendix

Vissa matematiska beskrivningar av test, ML-skattningar och dyl. som används i Metoddelen.

A.1 Maximum likelihood-skattningar för GARCH(1,1)

Maximum likelihood skattarna tas fram genom derivering av log likelihoodfunktionen som definieras baserat på den fördelning som antas för innovationerna i GARCH-modellen. För GARCH(1,1) modellen

$$a_t = \sigma_t \epsilon_t, \quad \sigma_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2 + \beta_1 \sigma_{t-1}^2,$$

definierar vi ML-skattarna som

$$\hat{\theta} = (\hat{\alpha}_0, \hat{\alpha}_1, \hat{\beta}_1) = argmax_\theta \; L(\theta),$$

där $L(\theta)$ är likelihoodfunktionen. Vi följer härledningarna för likelihoodfunktionerna i [Tsa10, s.120-121] som görs för ARCH-modellen. Skillnaden i hur vi skattar parametrarna i ARCH från GARCH är hur vi uttrycker volatiliteten, σ_t och vilka parametrar vi skattar.

A.1.1 Normal-fördelade innovationer

För GARCH(1,1)-modellering med normal-fördelade innovationer gör vi antagandet att a_t är normalfördelade med väntevärde noll och betingad varians så att $a_t \sim N(0, \sigma_t^2 | F_{t-1})$ där F_{t-1} betecknar informationen tillgänglig till och med observation t. Vi skriver täthetsfunktionen för innovationerna som

$$f_t(a_t) = \frac{1}{\sqrt{2\pi\sigma_t^2}} exp(-\frac{1}{2}\frac{a_t^2}{\sigma_t^2}).$$

Vidare skriver vi likelihoodfunktionen och härleder log-likelihoodfunktionen enligt

$$L(\theta) = L(\sigma_t^2(\theta), a_1, ...a_n) = \prod_{t=1}^{n} f_t(a_t) = (2\pi)^{-n/2} \cdot \prod_{t=1}^{n} \left[\sigma_t^{-1} exp(-\frac{1}{2}\frac{a_t^2}{\sigma_t^2})\right]$$

$$\Rightarrow l(\theta) = -\frac{n}{2} log(2\pi) - \frac{1}{2}\sum_{t=1}^{n} log(\sigma_t^2) - \frac{1}{2}\sum_{t=1}^{n} \frac{a_t^2}{\sigma_t^2}.$$

ML-skattningarna fås genom att sätta derivatan av log-likelihoodfunktionen till noll och lösa ekvationen.

$$\frac{l(\theta)}{\partial\sigma_t^2} = \frac{1}{2}\sum_{t=1}^{n}\Big[\frac{a_t^2}{(\sigma_t^2)^2} - \frac{1}{\sigma_t^2}\Big]\frac{\partial\sigma_t^2}{\partial\theta} = 0,$$

där

$$\frac{\partial\sigma_t^2}{\partial\theta} = (1, a_{t-1}^2, \omega_{t-1}^2) + \beta_1\frac{\partial\sigma_{t-1}^2}{\partial\theta}.$$

A.1.2 IGARCH(1,1) för RiskMetrics

Vi antar för RiskMetrics att innovationera är normalfördelade, därav gäller samma log-likelihoodfunktion som härleds för GARCH(1,1)-modellen med normal-fördelade innovationer. IGARCH(1,1) för RiskMetrics skrivs som

$$a_t = \sigma_t\epsilon_t, \quad \sigma_t^2 = (1-\beta_1)a_{t-1}^2 + \beta_1\sigma_{t-1}^2.$$

Skillnaden blir hur vi definierar θ, eftersom vi i RiskMetrics antar att $\alpha_0 = 0$ och $\alpha_1+\beta_1 = 1$ behöver vi bara skatta β_1.[14] Vi ansätter alltså $\theta = \beta_1$ som löses av ekvationen

$$\frac{l(\theta)}{\partial\sigma_t^2} = \frac{1}{2}\sum_{t=1}^{n}\Big[\frac{a_t^2}{(\sigma_t^2)^2} - \frac{1}{\sigma_t^2}\Big]\frac{\partial\sigma_t^2}{\partial\theta} = 0,$$

där

$$\frac{\partial\sigma_t^2}{\partial\theta} = \sigma_{t-1}^2 - a_{t-1}^2 + \beta_1\frac{\partial\sigma_{t-1}^2}{\partial\beta_1}.$$

A.1.3 T-fördelade innovationer

För t-fördelade innovationer betraktar vi två fall, då antalet frihetsgrader är specificerat a priori samt då antalet frihetsgrader skattas tillsammans med de andra parametrarna.

För ett givet antal frihetsgrader, v, skriver vi täthetsfunktionen för innovationerna som

$$f_t(a_t) = \frac{\Gamma(\frac{v+1}{2})}{\Gamma(\frac{v}{2})\sqrt{\pi(v-2)\sigma_t^2}}\Big[1 + \frac{a^2}{(v-2)\sigma_t^2}\Big]^{-\frac{v+1}{2}}.$$

Vi skriver likelihoodfunktionen och vidare härleder log-likelihoodfunktionen som

$$L(\theta) = L(\sigma_t^2(\theta), a_1, ..., a_n) = \prod_{t=1}^{n} f_t(a_t) \Rightarrow$$

[14] Vi kan även välja att skatta α_1 istället för β_1, insatt i modellen resulterar i det sama skattning för σ_t^2 oavsett.

$$l(\theta) = n \cdot log\left[\frac{\Gamma(\frac{v+1}{2})}{\Gamma(\frac{v}{2})\sqrt{\pi(v-2)}}\right] - \frac{1}{2}\sum_{t=1}^{n} log\ \sigma_t^2 - \frac{v+1}{2}\sum_{t=1}^{n} log\Big[1 + \frac{a_t^2}{(v-2)\sigma_t^2}\Big].$$

För ML-skattning av $\theta = (\alpha_0, \alpha_1, \beta_1)$ görs liksom beskrivet för normal-fördelade innovationer. Derivatan av log-likelikhoodfunktionen sätts till noll vilket ger ett ekvationssystem som löses för θ.

Log-Likelihoodfunktionen ovan gäller för ett, a priori, specificerat antal frihetsgrader. Då vi vill skatta antalet frihetsgrader görs det gemensamt med de andra parametrarna. ML-skattningarna kan då inte lösas explicit som i tidigare beskrivna fall, men kan beräknas med numeriska metoder.

A.2 Maximum likelihood-skattningar för Christoffersens Independence test

För härledning av ML-skattarna för Christoffersen's test för Independence skriver vi upp en kontingenstabell, Tabell 7, som beskriver resultaten för sekvensen $[I_t, I_{t-1}]_{t=1}^{T}$.

	$I_{t-1} = 0$	$I_{t-1} = 1$	Σ
$I_t = 0$	n_{00}	n_{01}	n_{0+}
$I_t = 1$	n_{10}	n_{11}	n_{1+}
Σ	n_{+0}	n_{+1}	n_{++}

Tabell 7: Kontingenstabell över summerade utfall för träfffunktionen tid t i kombination med utfall för träffunktionen i tid $t-1$. Antalet gånger $I_t = i$ givet att $I_{t-1} = j$ betecknas av n_{ij} där $t = \{1, ..., T\}$.

Vi följer [Chr98] och betraktar träff-funktionen som en Markov-kedja med stokastisk matris

$$\Pi_1 = \begin{bmatrix} 1 - \pi_{01} & \pi_{01} \\ 1 - \pi_{11} & \pi_{11} \end{bmatrix},$$

där $\pi_{ij} = P(I_t = j | I_{t-1} = i)$ vilket ger likelihood-funktionen och härledningsvis log-likelihoodfunktionen

$$L(\Pi_1) = (1 - \pi_{01})^{n_{00}} \pi_{01}^{n_{01}} (1 - \pi_{11})^{n_{10}} \pi_{11}^{n_{11}} \Rightarrow$$

$$l(\Pi_1) = n_{00} \log(1 - \pi_{01}) + n_{01} \log \pi_{10} + n_{10} \log(1 - \pi_{11}) + n_{11} \log \pi_{11}.$$

Vi sätter $\frac{\partial l(\hat{\Pi}_1)}{\partial \Pi_1} = 0$ vilket resulterar i fyra ekvationer vilkas lösningar ges av

$$\hat{\Pi}_1 = \begin{bmatrix} \frac{n_{00}}{n_{00}+n_{01}} & \frac{n_{01}}{n_{00}+n_{01}} \\ \frac{n_{10}}{n_{10}+n_{11}} & \frac{n_{11}}{n_{10}+n_{11}} \end{bmatrix}.$$

Låt motsvarande Markov-kedja under nollhypotesen om oberoende beskrivas av den stokastiska matrisen

$$\Pi_2 = \begin{bmatrix} 1-\pi_2 & \pi_2 \\ 1-\pi_2 & \pi_2 \end{bmatrix},$$

med likeliood-funktionen och log-likelihoodfunktionen

$$L(\Pi_2; I_1, ..., I_T) = (1-\pi_2)^{n_{00}+n_{10}} \pi_2^{n_{01}+n_{11}} \Rightarrow$$

$$l(\Pi_2) = (n_{00}+n_{10}) \log(1-\pi_2) + (n_{01}+n_{11}) \log \pi_2.$$

ML-skattarna få genom att lösa $\frac{\partial \hat{\Pi}_2}{\pi_2} = 0$ vilken löses av

$$\hat{\pi}_2 = \frac{n_{01}+n_{11}}{n_{01}+n_{11}+n_{00}+n_{10}},$$

vilket insatt i den stokastiska matrisn ger $\hat{\Pi}_2$.

A.3 AIC och BIC

För att avgöra hur väl en modell anpassar data, speciellt för att jämföra modeller brukar testen AIC och BIC användas där lägre värde indikerar bättre anpassning. Generellt skrivs testen enligt

$$AIC = -\frac{2}{T} \cdot \log(L) + \frac{2}{T} \cdot k$$

och

$$BIC = -\frac{2}{T} \cdot \log(L) + \frac{2}{T} \cdot \log(T),$$

där L är likelihoodfunktionen, k är antalet skattade parametrar i modellen och T är antalet observationer, se [Tsa10].