Alice-Data-Analysis

Daniel Gotthardt

2024-08-15

variance function regression

```r
library(dglm)
# Fit variance function regression with dglm
vfr <- dglm(data=mydata,
            # formula specifies the mean model
            formula = Y.obs ~ X.mean + I(X.mean^2) + Z,
            # dformula specifies the variance (dispersion) model
            # the model does not have to be the same!
            dformula = Y.obs ~ X.mean + I(X.mean^2) + Z,
            # Alice models the mean model as gaussian
            family = gaussian(link = "identity"),
            # dlink specifies the link-function for the dispersion
            dlink = "log",
            # reml should be used for all but very large sample sizes
            method="reml")
# Summary shows all important results, while print only shows the mean model!
summary(vfr)
```

```
##
## Call: dglm(formula = Y.obs ~ X.mean + I(X.mean^2) + Z, dformula = Y.obs ~
##     X.mean + I(X.mean^2) + Z, family = gaussian(link = "identity"),
##     dlink = "log", data = mydata, method = "reml")
##
## Mean Coefficients:
##                  Estimate Std. Error     t value     Pr(>|t|)
## (Intercept)  0.3870976321 0.03679239 10.52113422 1.284997e-21
## X.mean       0.0444272880 0.11144231  0.39865729 6.904915e-01
## I(X.mean^2)  0.4666474103 0.08104667  5.75776158 2.530410e-08
## Z           -0.0009699684 0.01199996 -0.08083099 9.356421e-01
## (Dispersion Parameters for gaussian family estimated as below )
##
##     Scaled Null Deviance: 1531.514 on 249 degrees of freedom
## Scaled Residual Deviance: 245.4988 on 246 degrees of freedom
##
## Dispersion Coefficients:
##                Estimate Std. Error     z value     Pr(>|z|)
## (Intercept) -3.14280361  0.4360171 -7.20798244 5.678695e-13
## X.mean      -0.56624079  1.7626882 -0.32123707 7.480307e-01
## I(X.mean^2) -5.02882333  1.8092341 -2.77953152 5.443737e-03
## Z           -0.02846807  0.2990737 -0.09518749 9.241659e-01
## (Dispersion parameter for Gamma family taken to be 2 )
##
##     Scaled Null Deviance: 408.9292 on 249 degrees of freedom
## Scaled Residual Deviance: 287.0631 on 246 degrees of freedom
##
## Minus Twice the Log-Likelihood: -502.9904
## Number of Alternating Iterations: 8
```

quantile regression

```r
library(quantreg)
# Define quantiles to fitted
quantiles <- c(0.1,0.25,0.5,0.75,0.9)
# Fit quantile regression
quant <- rq(formula = Y.obs ~ X.mean + I(X.mean^2) + Z,
            tau = quantiles,
            data = mydata)
# Print provides a nice overview of all results but does not compute SE
print(quant)
```

```
## Call:
## rq(formula = Y.obs ~ X.mean + I(X.mean^2) + Z, tau = quantiles,
##     data = mydata)
##
## Coefficients:
##                 tau= 0.10   tau= 0.25   tau= 0.50   tau= 0.75   tau= 0.90
## (Intercept) -5.029103e-02  0.23547860 0.363057656  0.58952106  0.72029299
## X.mean       8.762219e-01  0.22854610 0.069089619 -0.23084801 -0.53899212
## I(X.mean^2)  5.735323e-02  0.44545626 0.467691987  0.55473140  0.72907021
## Z            9.461185e-05 -0.01255031 0.007839453 -0.02726065  0.01416658
##
## Degrees of freedom: 250 total; 246 residual
```

```r
# Use summary with se = "rank" or "boot" to not require iid errors
summary(quant, se = "rank")
```

```
##
## Call: rq(formula = Y.obs ~ X.mean + I(X.mean^2) + Z, tau = quantiles,
##     data = mydata)
##
## tau: [1] 0.1
##
## Coefficients:
##             coefficients lower bd upper bd
## (Intercept) -0.05029     -0.12203  0.11364
## X.mean       0.87622      0.34086  1.13543
## I(X.mean^2)  0.05735     -0.10134  0.44799
## Z            0.00009     -0.02774  0.03356
##
## Call: rq(formula = Y.obs ~ X.mean + I(X.mean^2) + Z, tau = quantiles,
##     data = mydata)
##
## tau: [1] 0.25
##
## Coefficients:
##             coefficients lower bd upper bd
## (Intercept)  0.23548      0.11673  0.35468
## X.mean       0.22855     -0.01706  0.57558
## I(X.mean^2)  0.44546      0.13521  0.58939
## Z           -0.01255     -0.03780  0.02044
```

```
##
## Call: rq(formula = Y.obs ~ X.mean + I(X.mean^2) + Z, tau = quantiles,
##     data = mydata)
##
## tau: [1] 0.5
##
## Coefficients:
##             coefficients lower bd upper bd
## (Intercept)  0.36306      0.33732  0.46078
## X.mean       0.06909     -0.14378  0.17343
## I(X.mean^2)  0.46769      0.38149  0.66459
## Z            0.00784     -0.03582  0.03007
##
## Call: rq(formula = Y.obs ~ X.mean + I(X.mean^2) + Z, tau = quantiles,
##     data = mydata)
##
## tau: [1] 0.75
##
## Coefficients:
##             coefficients lower bd upper bd
## (Intercept)  0.58952      0.51474  0.73543
## X.mean      -0.23085     -0.82130 -0.03068
## I(X.mean^2)  0.55473      0.40447  1.00486
## Z           -0.02726     -0.04893  0.02390
##
## Call: rq(formula = Y.obs ~ X.mean + I(X.mean^2) + Z, tau = quantiles,
##     data = mydata)
##
## tau: [1] 0.9
##
## Coefficients:
##             coefficients lower bd upper bd
## (Intercept)  0.72029      0.68592  0.76228
## X.mean      -0.53899     -0.73780 -0.36091
## I(X.mean^2)  0.72907      0.57864  1.00007
## Z            0.01417     -0.08383  0.04030
```

```r
# Anova.rq with joint = FALSE allows to compare the coeficients between all taus
anova(quant, joint = FALSE, iid = FALSE)
```

```
## Quantile Regression Analysis of Deviance Table
##
## Model: Y.obs ~ X.mean + I(X.mean^2) + Z
## Tests of Equality of Distinct Slopes: tau in {  0.1 0.25 0.5 0.75 0.9  }
##
##             Df Resid Df F value    Pr(>F)
## X.mean       4     1246  7.0513 1.299e-05 ***
## I(X.mean^2)  4     1246  2.8885   0.02139 *
## Z            4     1246  1.8426   0.11830
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
# Fit quantiles seperately
out50 <- rq(Y.obs ~ X.mean+I(X.mean^2)+Z,
            data=mydata,tau=c(0.5))
out90 <- rq(Y.obs ~ X.mean+I(X.mean^2)+Z,
            data=mydata,tau=c(0.9))
out75 <- rq(Y.obs ~ X.mean+I(X.mean^2)+Z,
            data=mydata,tau=c(0.75))
out25 <- rq(Y.obs ~ X.mean+I(X.mean^2)+Z,
            data=mydata,tau=c(0.25))
out10 <- rq(Y.obs ~ X.mean+I(X.mean^2)+Z,
            data=mydata,tau=c(0.1))
# Anova can be used to compare specific quantiles for seperate fits
anova(out90, out10, joint = FALSE, iid = FALSE)
```

```
## Quantile Regression Analysis of Deviance Table
##
## Model: Y.obs ~ X.mean + I(X.mean^2) + Z
## Tests of Equality of Distinct Slopes: tau in {  0.9 0.1  }
##
##             Df Resid Df F value    Pr(>F)
## X.mean       1      499 26.7465 3.364e-07 ***
## I(X.mean^2)  1      499 10.1400  0.001541 **
## Z            1      499  0.2095  0.647322
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```