

Quantile
Regression

Samuel
Croker
Stephen
Dawkins

Introduction
Motivation

Medicare
Data Analysis

Data Source
EDA

Quantile Regression
in `ggplot2`
`quantreg` Package

Suggested
Reading

Quantile Regression

Samuel Croker
Stephen Dawkins

June 3, 2013

Medicare Charges and Payments

Quantile
Regression

Samuel
Crocker
Stephen
Dawkins

Introduction
Motivation

Medicare
Data Analysis

Data Source
EDA

Quantile Regression
in ggplot2
quantreg Package

Suggested
Reading

Top 100 Diagnosis Related Groups (DRG) by hospital for
FY 2011.

```
http://www.cms.gov/  
Research-Statistics-Data-and-Systems/  
Statistics-Trends-and-Reports/  
Medicare-Provider-Charge-Data/index.html
```

Densities

Quantile
Regression

Samuel
Croker
Stephen
Dawkins

Introduction

Motivation

Medicare
Data Analysis

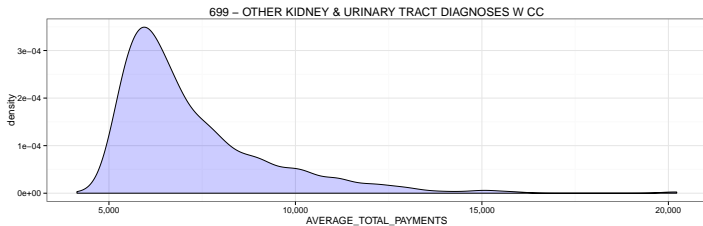
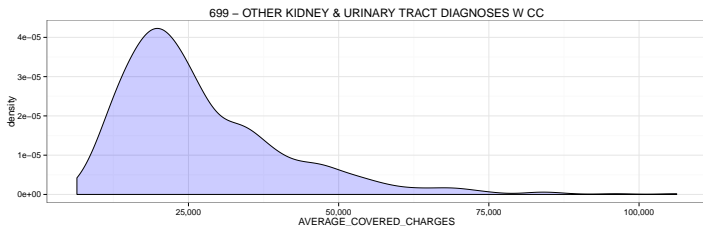
Data Source

EDA

Quantile Regression
in ggplot2

quantreg Package

Suggested
Reading



ggplot2 Density Code

Quantile
Regression

Samuel
Crocker
Stephen
Dawkins

Introduction
Motivation

Medicare
Data Analysis
Data Source
EDA

Quantile Regression
in ggplot2
quantreg Package

Suggested
Reading

```
library(ggplot2)

ggplot(data=medplot) +
  geom_density(aes(x=AVERAGE_COVERED_CHARGES,y=..density..)
    ,fill='blue',alpha=0.2) +
  scale_x_continuous(labels=comma) +
  theme_bw() +
  ggtitle(paste(medplot[1,1]))
```

Regression Quantiles for EDA

Quantile Regression

Samuel
Croker
Stephen
Dawkins

Introduction Motivation

Medicare Data Analysis

Data Source
EDA

Quantile Regression
in `ggplot2`
`quantreg` Package

Suggested Reading

`ggplot2` has a quantile regression layer that facilitates exploratory data analysis.

```
geom_quantile(quantiles=c(.25, 0.5, 0.75, 0.90))  
stat_quantile(quantiles=c(.25, 0.5, 0.75, 0.90))
```

Quantile Regression Plot

Quantile Regression

Samuel
Crocker
Stephen
Dawkins

Introduction

Motivation

Medicare Data Analysis

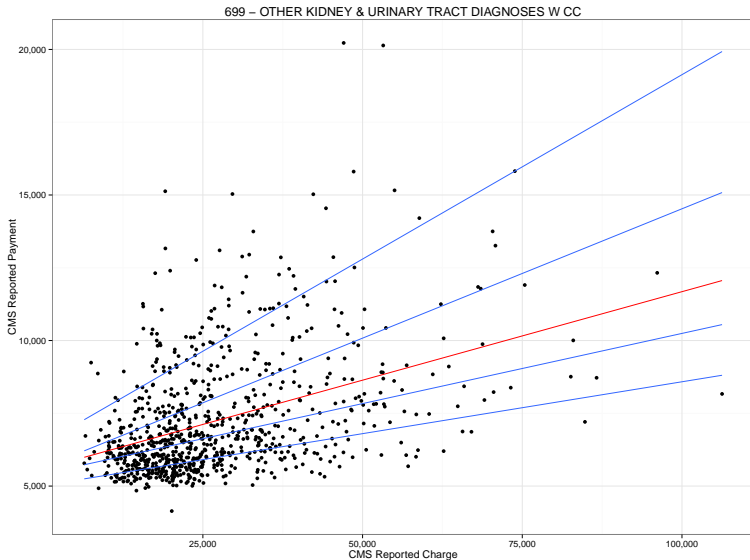
Data Source

EDA

Quantile Regression
in `ggplot2`

`quantreg` Package

Suggested Reading



ggplot2 Quantile Regression Code

Quantile Regression

Samuel
Croker
Stephen
Dawkins

Introduction

Motivation

Medicare Data Analysis

Data Source

EDA

Quantile Regression
in ggplot2

quantreg Package

Suggested Reading

```
library(ggplot2)
ggplot(data=medplot, aes(x=AVERAGE_COVERED_CHARGES
                          ,y=AVERAGE_TOTAL_PAYMENTS)) +
  geom_point() +
  geom_smooth(method=lm, color='red', se=F, size=1.2) +
  geom_quantile(quantiles=c(.25,0.5,0.75,0.90), size=1.2) +
  scale_x_continuous("CMS Reported Charge", labels=comma) +
  scale_y_continuous("CMS Reported Payment", labels=comma) +
  theme_bw() +
  ggtitle(paste(medplot[1,1]))
```

The quantreg Package

Quantile
Regression

Samuel
Croker
Stephen
Dawkins

Introduction

Motivation

Medicare
Data Analysis

Data Source

EDA

Quantile Regression
in ggplot2

quantreg Package

Suggested
Reading

```
library(quantreg)
#Author: Roger Koenker <rkoenker@illinois.edu>
#Maintainer: Roger Koenker <rkoenker@illinois.edu>

rq.fit <- rq(data=medplot
             , AVERAGE_TOTAL_PAYMENTS ~ AVERAGE_COVERED_CHARGES
             , tau=seq(.01, .99, .01)
             , method="br")
```


Quantile Regression Partial Plots

Quantile Regression

Samuel
Crocker
Stephen
Dawkins

Introduction

Motivation

Medicare Data Analysis

Data Source

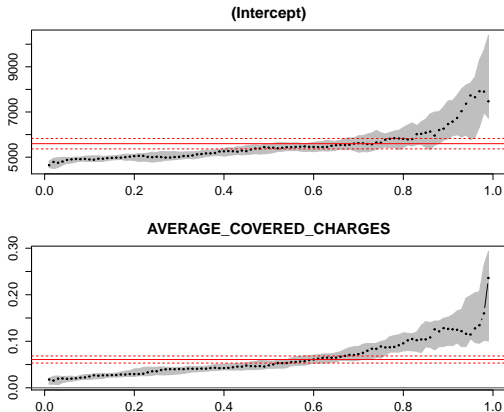
EDA

Quantile Regression
in `ggplot2`

`quantreg` Package

Suggested Reading

```
plot.summary.rqs(summary(rq.fit))
```



Parameter Estimates

Quantile
Regression

Samuel
Croker
Stephen
Dawkins

Introduction

Motivation

Medicare
Data Analysis

Data Source

EDA

Quantile Regression
in ggplot2

quantreg Package

Suggested
Reading

```
> summary(rq.fit)
```

```
tau: [1] 0.9
```

```
Coefficients:
```

	coefficients	lower bd	upper bd
(Intercept)	6465.40776	5812.89728	7491.64496
AVERAGE_COVERED_CHARGES	0.12668	0.09160	0.15189

```
Call: rq(formula = AVERAGE_TOTAL_PAYMENTS ~ AVERAGE_COVERED_CHARGES,  
tau = seq(0.01, 0.99, 0.01), data = medplot, method = "br")
```

```
> summary.rq(rq.fit, se='boot')
```

```
Call: rq(formula = AVERAGE_TOTAL_PAYMENTS ~ AVERAGE_COVERED_CHARGES,  
tau = 0.9, data = medplot, method = "br")
```

```
tau: [1] 0.9
```

```
Coefficients:
```

	Value	Std. Error	t value	Pr(> t)
(Intercept)	6465.40776	393.60603	16.42609	0.00000
AVERAGE_COVERED_CHARGES	0.12668	0.01547	8.18823	0.00000

Suggested Reading

Quantile
Regression

Samuel
Croker
Stephen
Dawkins

Introduction
Motivation

Medicare
Data Analysis

Data Source
EDA

Quantile Regression
in ggplot2
quantreg Package

Suggested
Reading

Quantile Regression, Koenker & Hallock <http://www.econ.uiuc.edu/~roger/research/rq/QRJEP.pdf>

QR in R, Koenker <http://www.econ.uiuc.edu/~roger/research/rq/vig.pdf>

Quantile Regression book, Koenker <http://www.amazon.com/Quantile-Regression-Econometric-Society-Monographs-dp/0521608279>

QR for Ecologists, Cade <http://www.fort.usgs.gov/products/publications/21137/21137.pdf>

Mostly Harmless Econometrics, Angrist & Pischke
<http://cces.lnu.edu.cn/upload/1245335295453.pdf>