

# Quantile regression for trend analysis

**CURRENT STATUS:** POSTED

James Elsner  
Florida State University

**DOI:**

10.1038/nprot.2008.203

**SUBJECT AREAS**

*Computational biology and bioinformatics*

**KEYWORDS**

*quantiles, quantile regression, R project for statistical computing, trend analysis, tropical cyclones*

## Introduction

Quantile regression extends ordinary least-squares regression to quantiles of the response variable.

Ordinary regression is a model for the conditional mean, where the mean is conditional on the value of the explanatory variable. Likewise, quantile regression is a model for the conditional quantiles. For trend analysis the explanatory variable is time. Quantiles are points taken at regular intervals from the cumulative distribution function of a random variable. The quantiles mark a set of ordered data into equal-sized data subsets.

The software is downloaded from the internet and installed on a computer. A data set from the internet is imported into a software session. An exploratory plot of the data is created to visualize the trends. A quantile regression model is fit to the data to quantify the trends and determine their statistical significance.

## Equipment

1. Computer running Macintosh, Linux/Unix, or Windows
2. Internet connection
3. R environment for statistical computing - "<http://www.r-project.org/>":<http://www.r-project.org> <sup>1</sup>

## Procedure

1. Download and install R. **Tip:** Only the base directory is needed.
2. Click on the icon to start R. With Linux/Unix, type the letter R from a command window.
3. Read the data into R by typing on the command line: `_StormMax=read.csv - "http://garnet.fsu.edu/jelsner/extspace/extremedatasince1899.csv":http://garnet.fsu.edu/jelsner/ex`

**Caution:** The quotes must be bidirectional. With a copy/paste directly from this document, the quotes will most likely be unidirectional causing an error message.

4. Subset the cyclones by basin (cyclones away from the U.S. coastline) and by year after 1977 (satellite era) by typing:

```
StormMaxBasin=subset(StormMax,Region=="Basin");
```

```
StormMaxBasin=subset(StormMaxBasin,Yr>1977)
```

5. Make the columns of the data set available by name by typing:

```
attach(StormMaxBasin)
```

6. Create an exploratory plot of the annual lifetime maximum wind speed (intensity) as a function of year by typing: *x=boxplot(Wmax~as.factor(Yr),plot=F)*

```
boxplot(Wmax~as.factor(Yr),ylim=c(35,175),xlab="Year",ylab="Intensity (kt)")
```

```
xx=1:29
```

```
abline(lm(x$stats[5,]~xx),col="red")
```

```
abline(lm(x$stats[4,]~xx),col="blue")
```

```
abline(lm(x$stats[3,]~xx),col="green")
```

7. Install and load the *quantreg* package developed by Roger Koenker<sup>2</sup>. Then print the reference citation. *install.packages("quantreg")*

```
library(quantreg)
```

```
citation("quantreg")
```

8. Summarize the results of a quantile regression at the upper quantiles 0.75, 0.9, and 0.95. *summary(rq(Wmax~Yr,tau=c(0.75,0.9,0.95)),se="iid")*

**Tip:** The standard errors (*se=* argument) can be estimated using other methods, type: *?summary.rq*

9. Plot the model results. *model=rq(Wmax~Yr,tau=seq(0.2,0.8,0.1))*

```
plot(summary(model,alpha=.05,se="iid"),parm=2,pch=19,cex=1.2,mar=c(5,5,4,2)+0.1,ylab="Trend (kt/yr)",xlab="Quantile")
```

## Timing

20 minutes

## Troubleshooting

Make sure there is permission on the computer to install software.

Make sure the commands are typed exactly as shown with bidirectional double quotes.

If copy/paste is used, make sure to change the quotations to bidirectional.

## Anticipated Results

The exploratory plot should verify the lack of trend in the median lifetime maximum intensity. It

should also show a tendency for the strongest cyclones (higher quantiles) to get stronger during the past 30 years. The statistical significance of the trends is assessed with a quantile regression model and the results are plotted.

## References

1. R Development Core Team, R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, "<http://www.R-project.org>":<http://www.R-project.org> (2006).
2. Koenker, R. *quantreg: Quantile Regression*. R package version 4.10. "<http://www.r-project.org>":<http://www.r-project.org> (2007).

## Acknowledgements

Thanks go to all involved with the R project for statistical computing. Special thanks go to Thomas Jagger for his statistical help. The work is supported by the U.S. National Science Foundation, Risk Prediction Initiative of the Bermuda Institute for Ocean Studies, and the Florida Catastrophic Storm Risk Management Center of Florida State University.

## The increasing intensity of the strongest tropical cyclones

by James B. Elsner, James P. Kossin & Thomas H. Jagger

Nature (04 July, 2008)