

There is a variant of the normal PP-Plot described in Devore's Problem 4.96. It is called the "Half-Normal" PP Plot, and is used for data that is expected to have mean zero and when n -size is small. The idea is to flip the negative points of the distribution into the first quadrant and "double up" the number of points.

Given observations $\{x_1, x_2, \dots, x_n\}$, we order their absolute values $\{w_1, w_2, \dots, w_n\}$. The corresponding quantiles are

$$q_i = \Phi^{-1} \left(\frac{1}{2} \left[\frac{i - .5}{n} + 1 \right] \right)$$

Then the Half Normal PP Plot are the points (q_i, w_i) in the first quadrant. If the data was normal, then these points would line up along a line through the origin.

The typical dataset with mean zero is the set of residuals from a statistical model. For example, the sample mean is the best estimate of the center, and by subtracting the sample mean from the data gives the set of residuals which have mean zero. In fact, the standardized residuals

$$z_i = \frac{x_i - \bar{X}}{s}$$

also have mean zero, and would line up with the 45° line if the data were normal.

R Session:

```
R version 2.11.1 (2010-05-31)
Copyright (C) 2010 The R Foundation for Statistical Computing
ISBN 3-900051-07-0
```

```
R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.
```

```
Natural language support but running in an English locale
R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.
```

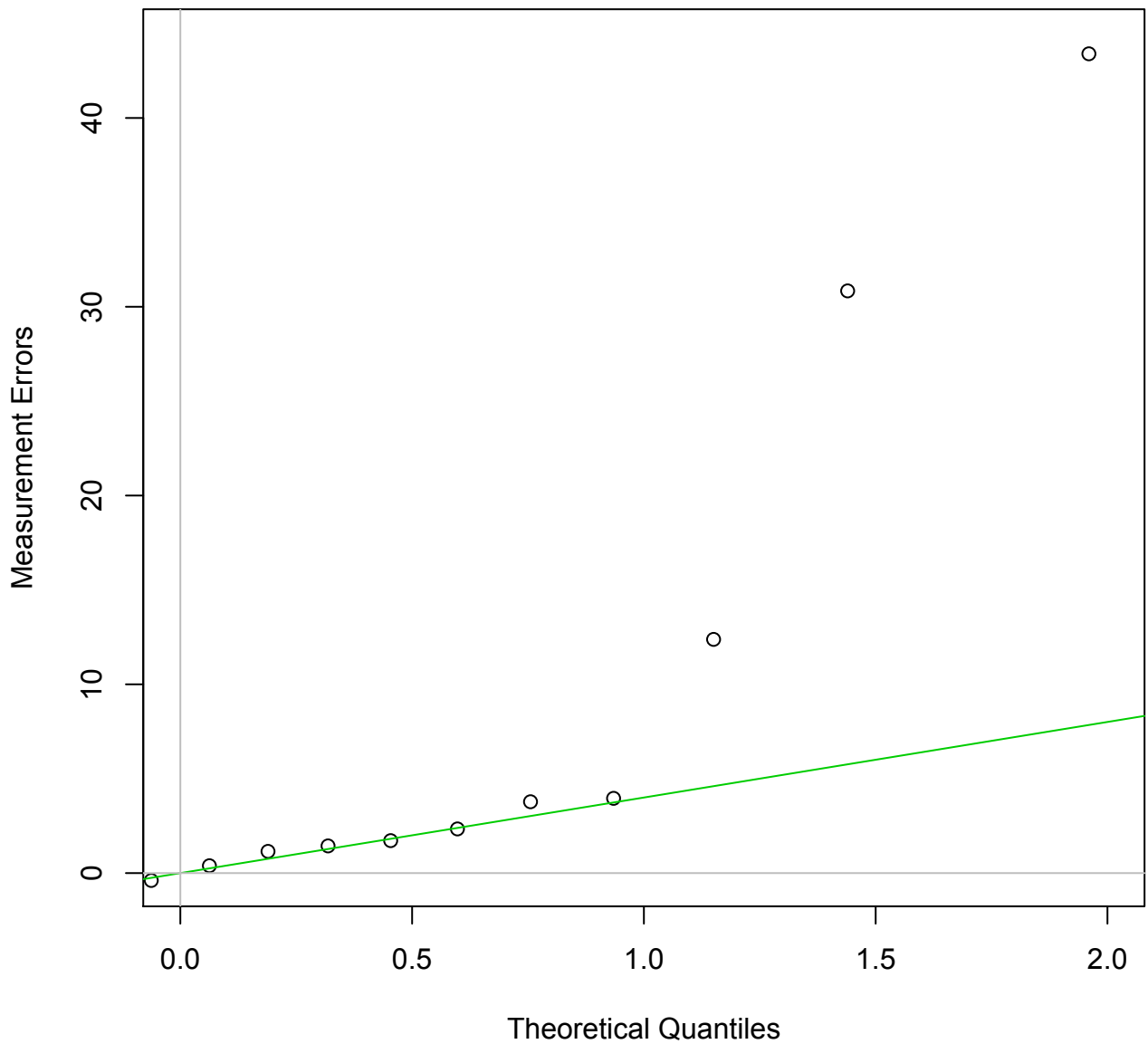
```
Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.
```

```
[R.app GUI 1.34 (5589) i386-apple-darwin9.8.0]
[Workspace restored from /home/1004/ma/treibergs/.RData]
```

```
> ##### SCAN THE DATA FROM DEVORE 4.96 #####
> # These are "measurement errors"
> err <- scan()
1: -3.78 -1.72 1.44 -.39 12.38 -43.40 1.15 -3.96 -2.34 30.84
11:
Read 10 items
> err
 [1] -3.78 -1.72 1.44 -0.39 12.38 -43.40 1.15 -3.96 -2.34 30.84
> summary(err)
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
-43.400 -3.420  -1.055  -0.978  1.368  30.840
```

```
> ##### DOUBLE THE DATA, PLOT HALF OF CANNED qqplot() #####
> twoerr <- c(abs(err),-abs(err))
> qqnorm(twoerr,xlim=c(0,2),ylim=c(0,44),main="Half-Normal QQ Plot",ylab="Measurement Errors")
> qqline(twoerr,col=3)
> abline(h=0,col="gray");abline(v=0,col="gray")
> # Plot is not very normal: it is heavy tailed.
```

Half-Normal QQ Plot



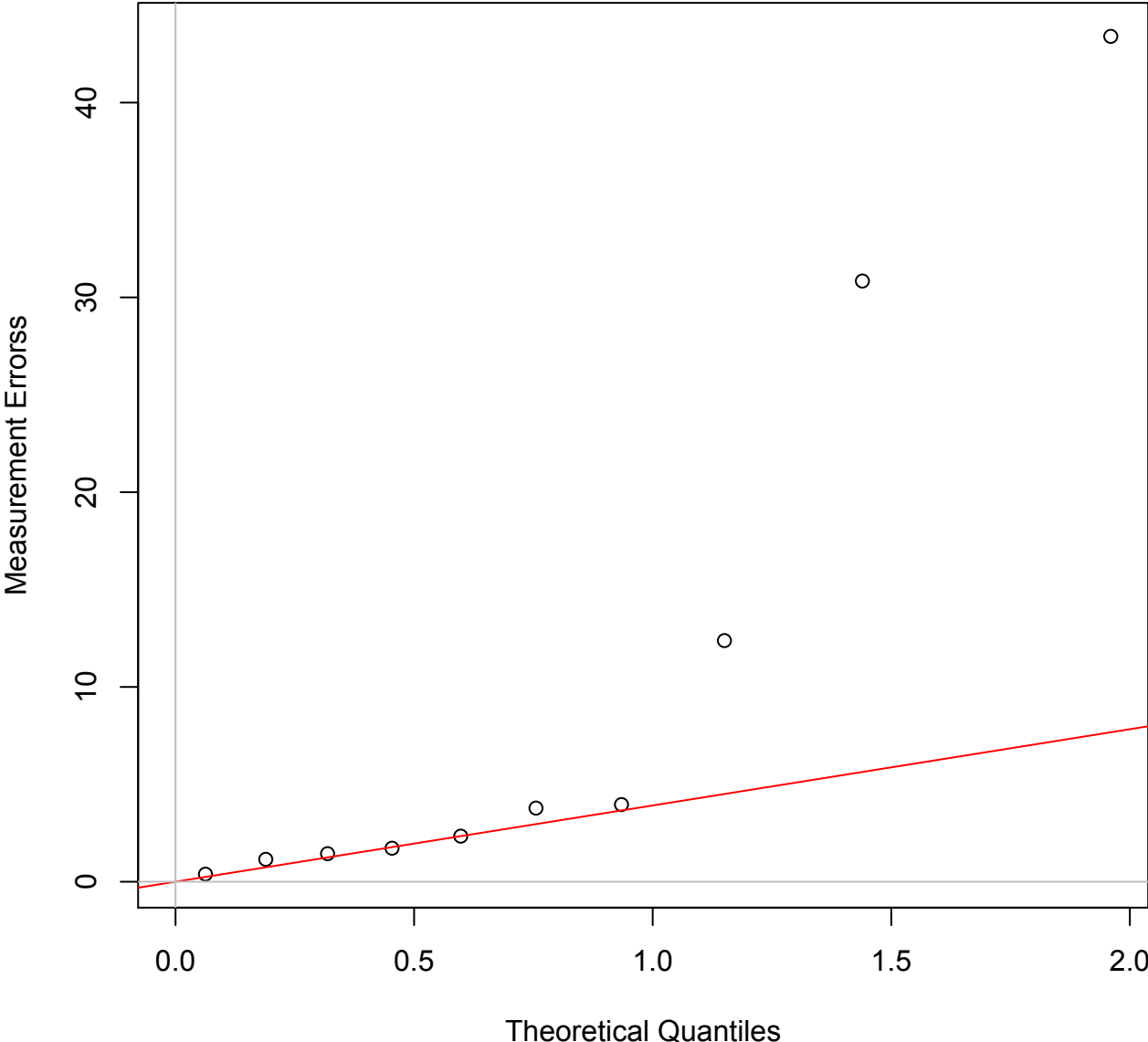
```

> ##### HALF-NORMAL PLOT "BY HAND" #####
> # Half-Normal plot by hand

> w <- sort(abs(err))
> w
[1] 0.39 1.15 1.44 1.72 2.34 3.78 3.96 12.38 30.84 43.40
> n <- length(err)
> n
[1] 10
> pts <- ((1:10-.5)/n+1)/2;pts
[1] 0.525 0.575 0.625 0.675 0.725 0.775 0.825 0.875 0.925 0.975
> # this gives even spacing of percentages between .5 and 1.
> # This function is canned in R: ppoints()
> # It produces (1:m - a)/(m + (1-a)-a) where default a= 1/2 for large sets.
> ppoints(1:(2*n))
[1] 0.025 0.075 0.125 0.175 0.225 0.275 0.325 0.375 0.425 0.475
[11] 0.525 0.575 0.625 0.675 0.725 0.775 0.825 0.875 0.925 0.975
> pts <- ppoints(2*n)[(n+1):(2*n)]; pts
[1] 0.525 0.575 0.625 0.675 0.725 0.775 0.825 0.875 0.925 0.975
> qs <- qnorm(pts); qs
[1] 0.06270678 0.18911843 0.31863936 0.45376219 0.59776013
[6] 0.75541503 0.93458929 1.15034938 1.43953147 1.95996398
> plot(qs,w,xlim=c(0,max(qs)),ylab="Measurement Errors",xlab="Theoretical Quantiles")
> # To add the line, recall that "abline(a,b)" draws line y = a + bx
> # I put in the line that is through the origin and the quartile. (midway from 1 to n)
> # I use the "n %% 2"th point. (integer divide)
> abline(0,w[n%%2]/qs[n%%2],col=2)
> abline(h=0,col="gray");abline(v=0,col="gray")
> title("Half-Normal PP Plot")
> # M3074HalfNormal2.pdf

```

Half-Normal PP Plot



```

> ##### ENTER normtemp DATA FROM UsingR #####
> library(UsingR)
Loading required package: MASS
> # Data set "normtemp" is simulated from values contained in Mackowiak, P. A.,
> # Wasserman, S. S., and Levine, M. M. (1992), "A Critical Appraisal of
> # 98.6 Degrees F, the Upper Limit of the Normal Body Temperature, and Other
> # Legacies of Carl Reinhold August Wunderlich," Journal of the American Medical
> # Association, 268, 1578-1580.
> # gender = 1 for male. Temp is Fahrenheit
> head(normtemp)
  temperature gender hr
1          96.3      1 70
2          96.7      1 71
3          96.9      1 74
4          97.0      1 80
5          97.1      1 73
6          97.1      1 75
> # Pick off males temperature.
> attach(normtemp)
> z <- subset(temperature,gender==1,data=normtemp)
> length(z)
[1] 65
> summary(z)
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  96.3   97.6   98.1   98.1   98.6   99.5
> # Standardize the data. Plot Standardized Observations vs normal quantiles
> znor <- (z-mean(z))/sd(z)
> ##### PLOT USUAL QQ-PLOT #####
> qqnorm(znor,ylab="Normalized Male Body Temperatures")
> abline(h=0,col="gray");abline(v=0,col="gray");abline(0,1,col=4)
> #M3074HalfNormal3.pdf
>
> # Data is quite normal.
> shapiro.test(znor)

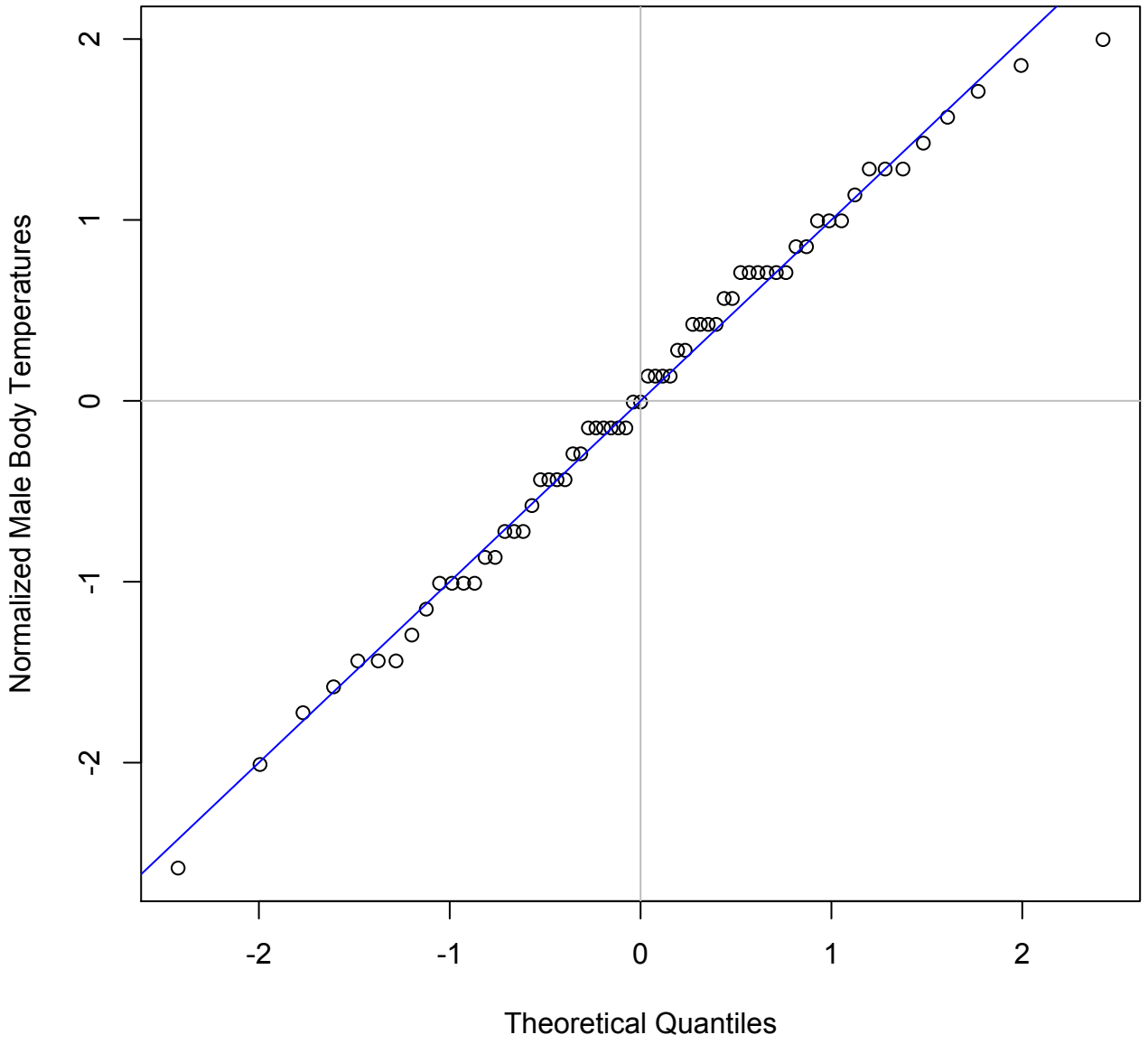
Shapiro-Wilk normality test

data:  znor
W = 0.9894, p-value = 0.8545

> # Yup. Can't reject H0: znor is normal.

```

Normal Q-Q Plot



```
> ##### HALF NORMAL PLOT OF znor #####
>
> zns <- sort(abs(znor))
> n <- length(zns);n
[1] 65
> pots <- ppoints(2*n)[(n+1):(2*n)]
> qots <- qnorm(pots)
> plot(qots, zns, xlim = c(0,max(qots)), ylab = "Standardized Male Temperatures",
+ xlab = "Theoretical Quantiles", main = "Half-Normal PP Plot")
> abline(h=0,col=8); abline(v=0,col=8)
> abline(h=0,col=8); abline(v=0,col=8)
> abline(0,1,col=rainbow(8)[7])
> # M3074HalfNormal4.pdf
```

Half-Normal PP Plot

