Math 3070 § 1. **Cement Blocksd: Weibull PP Plot** Name: Example
Treibergs October 10, 2012

This demonstration illustrates using the PP-plot to check if data is reasonably distributed as Weibull. We use the data 6.9.8 provided in Navidi, *Statistics for Engineers and Scientists, 2nd. ed.,* Mc Graw Hill, 2008. The measurements are of the compressive strength in MPa of cement blocks that were cured for six days.

To generate a probability plot for Weibull variables, we follow the prescription in Devore, *Probability and Statistics for Engineers and Scientists, 5th ed.,* Duxbury, 2000, in section 4.6. Observe that the cumulative distribution function for Weibull random variable X with parameters $\alpha, \beta > 0$ is

$$\begin{aligned} P(X \le x) = F(x) &= 1 - \exp\left(-\left(\frac{x}{\beta}\right)^{\alpha}\right) \\ &= 1 - \exp\left(-\exp\left(\alpha(\log x - \log\beta)\right)\right) \end{aligned}$$

After $u = \alpha(\log x - \log\beta)$, a linear change of variables, $\log x$ has the cumulative distribution

$$F(u) = 1 - e^{-e^{u}},$$

which is called the *extreme value distribution* because of other applications.

Thus we may plot quantiles of logarithm of the observed variable vs. the quantiles of the extreme value distribution to get a PP-plot for X. The observed quantiles are the sorted variables $\log(x_i)$, for $i = 1, 2, \ldots, n$. To get the theoretical quantiles, one solves $p_i = F(\eta(p_i))$ for the midpoints of equal intervals of $[0, 1]$ so for $i = 1, 2, \ldots, n$,

$$p_i = \frac{i - \frac{1}{2}}{n}$$

and so

$$q_i = \eta(p_i) = \log\bigl(-\log(1 - p_i)\bigr).$$

The Weibull PP-plot consists of the points $(q_i, \log(x_i))$.

After we're convinced that Weibull is a reasonable distribution for our data, how do we determine the parameters α and β? We use the method of moments, which is just a fancy way of saying we pick the Weibull parameters in such a way that their mean and variance is the sample mean and sample variance of the data. For Weibull variables, the population mean and variance are given by

$$\mu = \beta\Gamma\left(1 + \frac{1}{\alpha}\right),$$

$$\sigma^2 = \beta^2\left\{\Gamma\left(2 + \frac{1}{\alpha}\right) - \left[\Gamma\left(1 + \frac{1}{\alpha}\right)\right]^2\right\}$$

By adding the square of the mean, we isolate α.

$$\frac{\sigma^2 + \mu^2}{\mu^2} = \frac{\Gamma\left(1 + \frac{2}{\alpha}\right)}{\left[\Gamma\left(1 + \frac{1}{\alpha}\right)\right]^2} = g(\alpha)$$

Using the sample statistics, we solve for α in

$$r = \frac{s^2 + \bar{x}^2}{\bar{x}^2} = g(\alpha).$$

Once α is known, we solve the equation for the mean for β:

$$\beta = \frac{\bar{x}}{\Gamma\left(1 + \frac{1}{\alpha}\right)}.$$

One way to solve $r = g(\alpha)$ for α is to use the root finder `uniroot()` in **R**©. By plotting, we see that $g(\alpha)$ is a strictly decreasing function. We observe that $g(1) > r > g(100)$ so that we supply the interval for the root finder

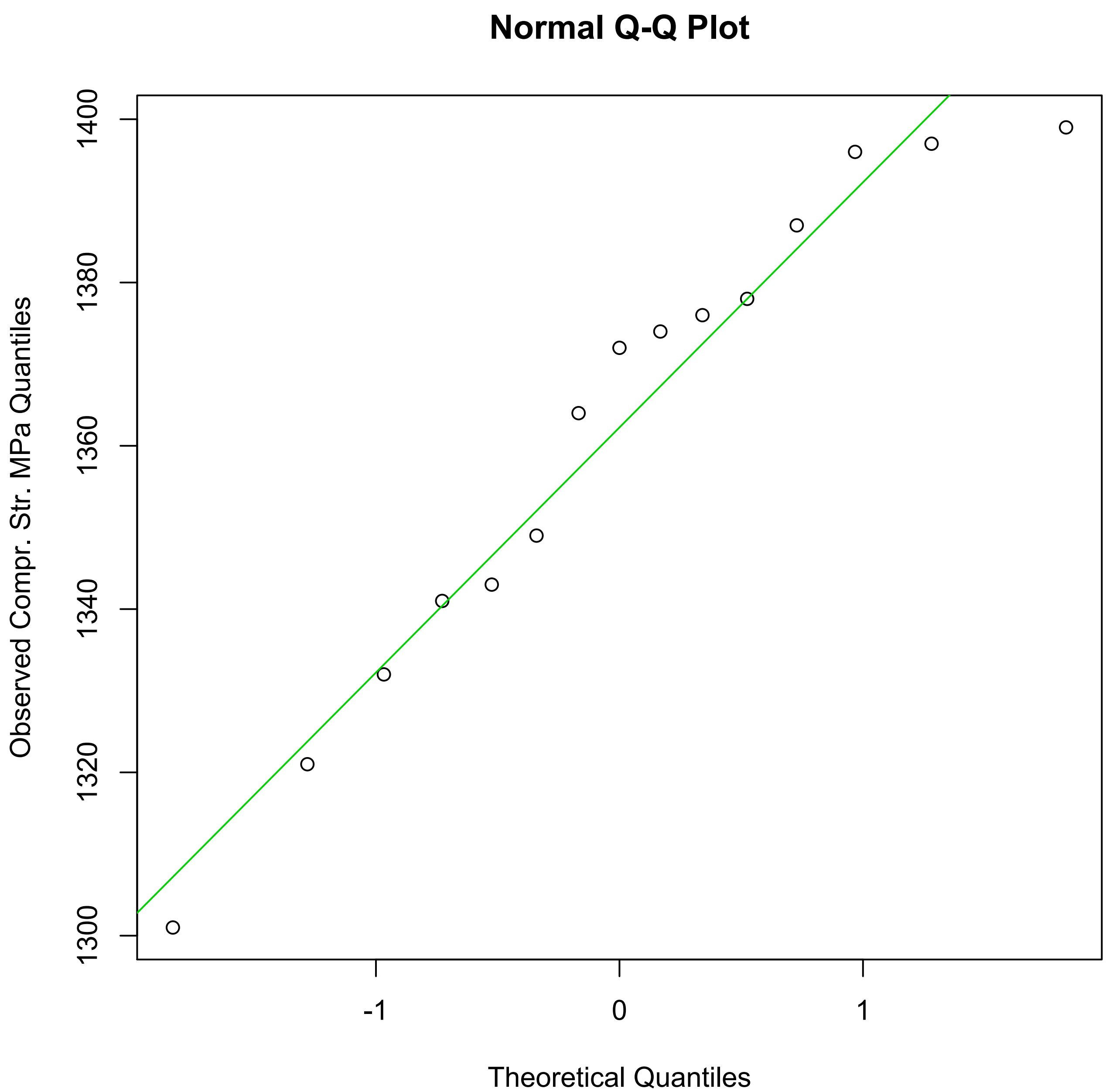
```
uniroot(g(x)-r,c(1,100))
```

which spits out the root α and its accuracy.

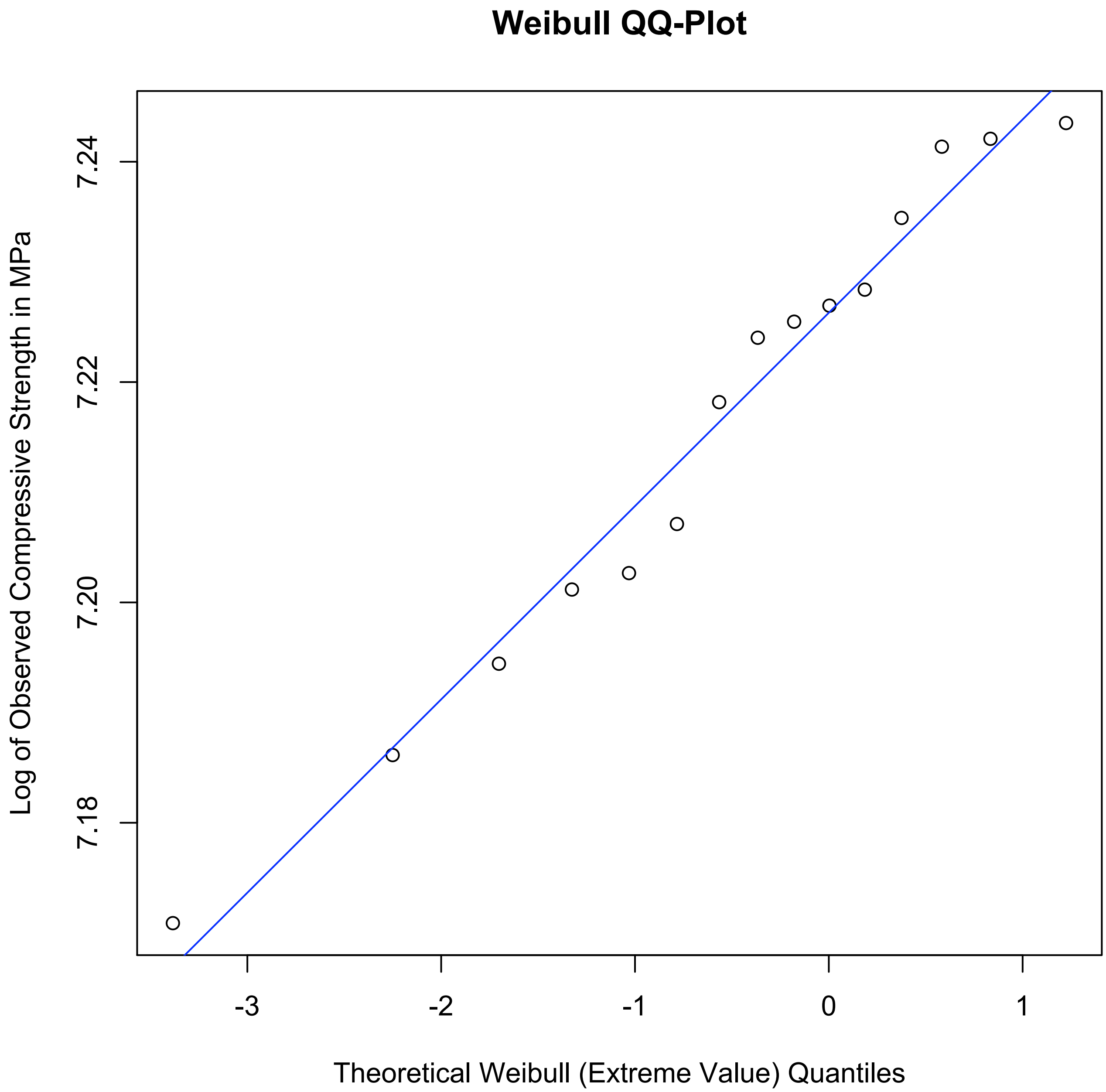
R Session:

```
R version 2.10.1 (2009-12-14)
Copyright (C) 2009 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.

[R.app GUI 1.31 (5538) powerpc-apple-darwin8.11.1]



> ########## READ THE COMPRESSIVE STRENGTH DATA (NAVIDI 6.9.8) ###############
> tii <- scan()
1: 1387 1301 1376 1397 1399 1378 1343 1349 1321 1364 1332 1396 1372
14: 1341 1374
16:
Read 15 items
> tii
 [1] 1387 1301 1376 1397 1399 1378 1343 1349 1321 1364 1332 1396 1372
[14] 1341 1374

> n <- length(tii);n
[1] 15

> # Sort
> sii <- sort(tii); sii
 [1] 1301 1321 1332 1341 1343 1349 1364 1372 1374 1376 1378 1387 1396
[14] 1397 1399
```

```
> # Take logs
> lsii <- log(sii); lsii
 [1] 7.170888 7.186144 7.194437 7.201171 7.202661 7.207119 7.218177
 [8] 7.224025 7.225481 7.226936 7.228388 7.234898 7.241366 7.242082
[15] 7.243513

> ################# CHECK IF DATA IS REASONABLY NORMALLY DISTRIBUTED ##########
> qqnorm(tii,ylab="Observed Compr. Str. MPa Quantiles")
> qqline(tii,col=3)
> # M3075Cement2.pdf
> # The plot is somewhat "r" shaped, so the normality may not hold. Although
> # n=15 is small and the normal QQ-Plot does not strongly indicate non-normal.
> # Here we suppose that there is strong engineering reason to suspect Weibull.
```

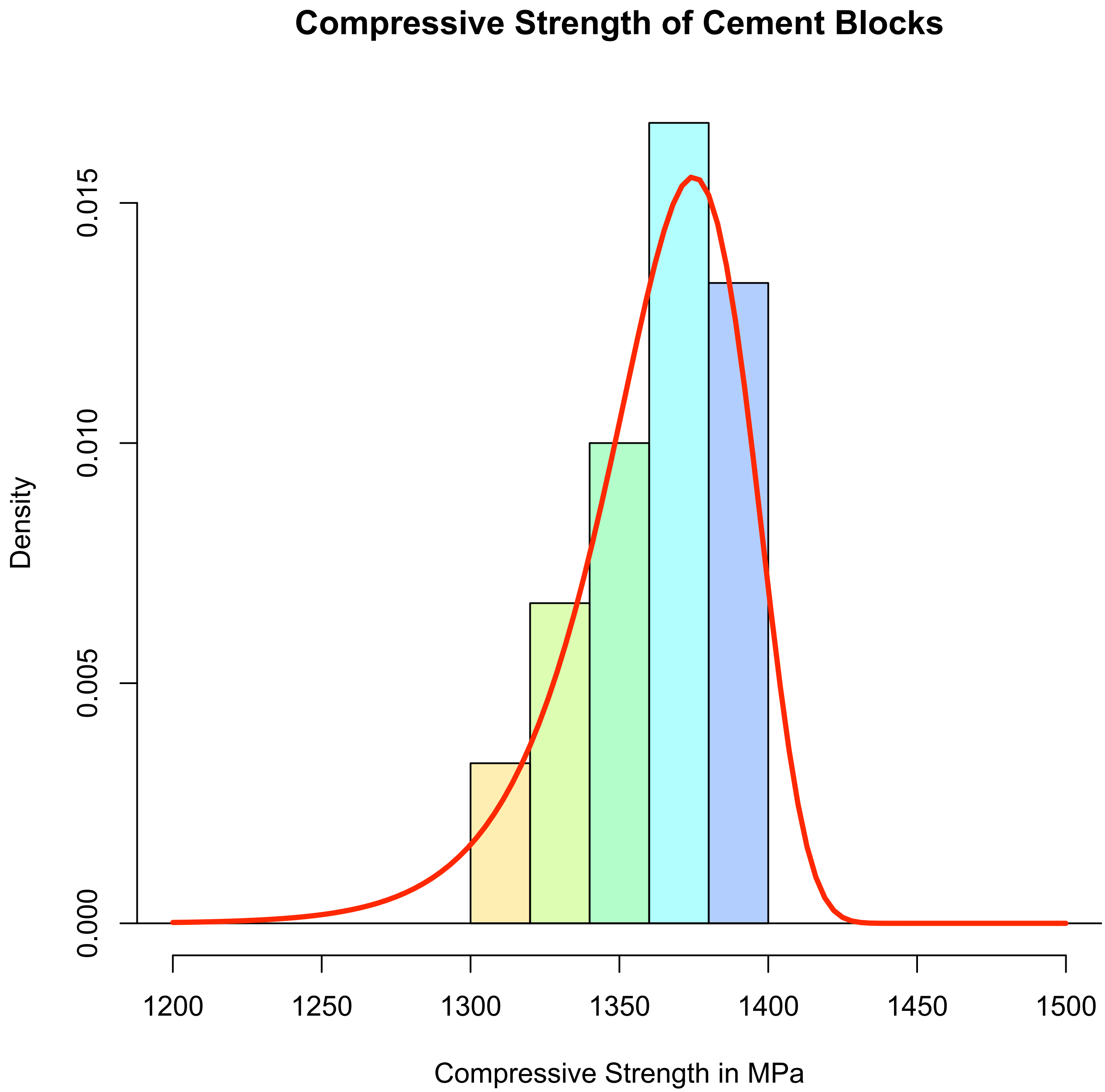

Normal Q-Q Plot
Observed Compr. Str. MPa Quantiles
1300
1320
1340
1360
1380
1400
-1
0
1
Theoretical Quantiles

```
> ##################### CONSTRUCT THE THEORETICAL QUANTILES ##################
> # "ppoints" gives midpoints of  n  equal intervals of [0,1]
> q <- ppoints(n);q
 [1] 0.03333333 0.10000000 0.16666667 0.23333333 0.30000000 0.36666667
 [7] 0.43333333 0.50000000 0.56666667 0.63333333 0.70000000 0.76666667
[13] 0.83333333 0.90000000 0.96666667

> # For Weibull, the quantiles (Extreme value Quantiles)
> qq <- log(-log(1-q)); qq
 [1] -3.384294493 -2.250367327 -1.701983355 -1.325375512 -1.030930433
 [6] -0.783600688 -0.565661963 -0.366512921 -0.178830030  0.003296669
[11]  0.185626759  0.375203292  0.583198081  0.834032445  1.224127541

> qq <- log(-log(1-pp))
>
> ############# TEST IF DATA IS WEIBULL USING WEIBULL QQ-PLOT ################
>
> qqplot(qq,lsii,main="Weibull QQ-Plot",ylab="Log of Observed Compressive
  Strength in MPa",xlab="Theoretical Weibull (Extreme Value) Quantiles")
> abline(lm(lsii ~ qq),col=4)
> # M3075Cemen3.pdf
> # The points line up better. Thus Weibull is a reasonable distribution.
```


Weibull QQ-Plot
Log of Observed Compressive Strength in MPa
7.18
7.20
7.22
7.24
-3
-2
-1
0
1
Theoretical Weibull (Extreme Value) Quantiles

```
>
> ############### FIND THE PARAMETERS OF THE WEIBULL DISTRIBUTION ############
>
> # We shall use the method of moments. We compute the sample mean  m  and
> # sample variance  v  and find  Weibull  alpha and beta that have
> # mu = m  and  sigma^2 = v.
>
> m<- mean(tii)
> v <- var(tii)
> ratio <- (v+m^2)/m^2;ratio
[1] 1.000475
> # define a function of alpha
> f <- function(x){gamma(1+2/x)/gamma(1+1/x)^2}
> #  f  is a decreasing function: see it by plotting
> curve(f(x),1,10)
> #
> # To solve  f(alpha)=ratio, we need to find an interval that brackets
> f(1)
[1] 2
> f(10)
[1] 1.014475
> f(100)
[1] 1.000162
> # Thus F(1) > ratio > f(100).
> # "uniroot" solves a one variable equation f(x)-ratio=0
>
> uniroot(function(x){gamma(1+2/x)/gamma(1+1/x)^2-ratio},c(1,100))
$root
[1] 58.10109

$f.root
[1] -1.074045e-10

$iter
[1] 8

$estim.prec
[1] 6.103516e-05

> alpha<- 58.10109
> beta <- m/gamma(1+1/alpha);beta
[1] 1375.266
>
> # Check that these alpha and beta give the mean and variance
> beta*gamma(1+1/alpha)
[1] 1362
> # same as  m
>
> beta^2*(gamma(1+2/alpha)-gamma(1+1/alpha)^2)
[1] 881.9998
> # almost the same as  v
```

```
> ################ PLOT THE DENSITY FUNCTION ON TOP OF THE HISTOGRAM #######
>
> hist(tii,freq=F,xlim=c(1200,1500),col=rainbow(8,alpha=.3)[2:6], main=
  "Compressive Strength of Cement Blocks", xlab="Compressive Strength in MPa")
> abline(h=0)
> curve(dweibull(x,scale=beta,shape=alpha),add=T,col=2,lwd=3)
> # M3075Cemen1.pdf
>
```



Compressive Strength of Cement Blocks
Density
0.000
0.005
0.010
0.015
1200
1250
1300
1350
1400
1450
1500
Compressive Strength in MPa