Problem 1: Chapter 6, Problem 8

Skinfold body measurements are used to approximate the body density of individuals. The data on n = 92 men, aged 20-25, where x = skinfold measurement and Y = body density are given available in the file skinfolddata.txt posted on the course website.

R Code to Fit Simple Linear Regression Model

```
x<-skinfolddata$Skinfold
# relabel Skinfold variate as x
y<-skinfolddata$BodyDensity
# relabel Body Density variate as y
# run regression y = alpha+beta*x
RegModel<-lm(y~x)
# parameter estimates and p-value for test of no
# relationship
summary(RegModel)
```

R Output

Call: lm(formula = y ~ x)

Residuals:

Min	1Q	Median	3Q	Max
-0.0251400	-0.0040412	-0.0001752	0.0041324	0.0192336

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	1.161139	0.005429	213.90	<2e-16 ***
x	-0.062066	0.003353	-18.51	<2e-16 ***

--- Signif. codes: 0 ‘*’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1**

Residual standard error: 0.007877 on 90 degrees of freedom

Multiple R-squared: 0.7919, Adjusted R-squared: 0.7896

F-statistic: 342.6 on 1 and 90 DF, p-value: < 2.2e-16

Find:

(a) the least squares estimate of β

(b) the maximum likelihood estimate of α

(c) the equation of the fitted least squares line

(d) the estimate of σ

(e) the value of the test statistic for testing H_0: $\beta = 0$

(f) the p-value for testing H_0: $\beta = 0$

(g) conclusion with justification regarding the hypothesis H_0: $\beta = 0$

R Output

Call: lm(formula = y ~ x)

Residuals:

Min	1Q	Median	3Q	Max
-0.0251400	-0.0040412	-0.0001752	0.0041324	0.0192336

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	1.161139	0.005429	213.90	<2e-16 ***
x	-0.062066	0.003353	-18.51	<2e-16 ***

--- Signif. codes: 0 '*' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1**

Residual standard error: 0.007877 on 90 degrees of freedom

Multiple R-squared: 0.7919, Adjusted R-squared: 0.7896

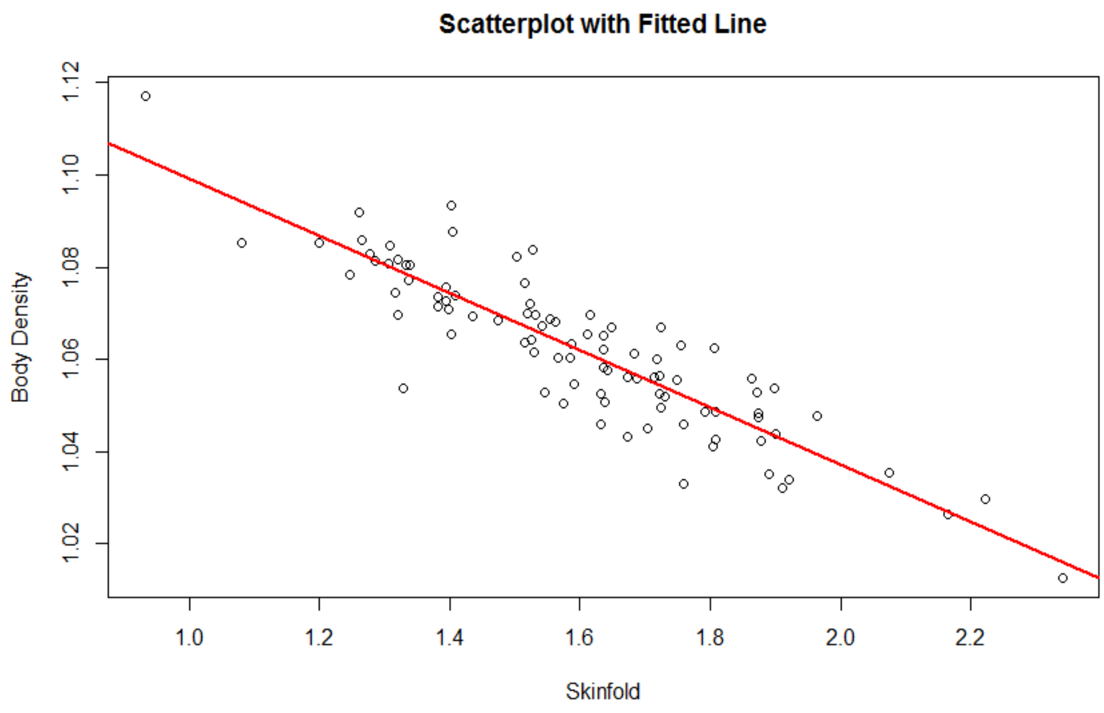
F-statistic: 342.6 on 1 and 90 DF, p-value: < 2.2e-16

95% Confidence interval for β

```
n<-length(x)    # n=sample size
betahat<-RegModel$coefficients[2]
# estimate of slope
se<-summary(RegModel)$sigma
# 95% Confidence interval for slope
a<-qt(0.975,n-2)      # value from t table for 95%
confidence interval
Sxx<-sum(x^2)-sum(x)^2/n      # value of Sxx
c(betahat-a*se/sqrt(Sxx),betahat+a*se/sqrt(Sxx))
          x           x
-0.06872823 -0.05540425
```

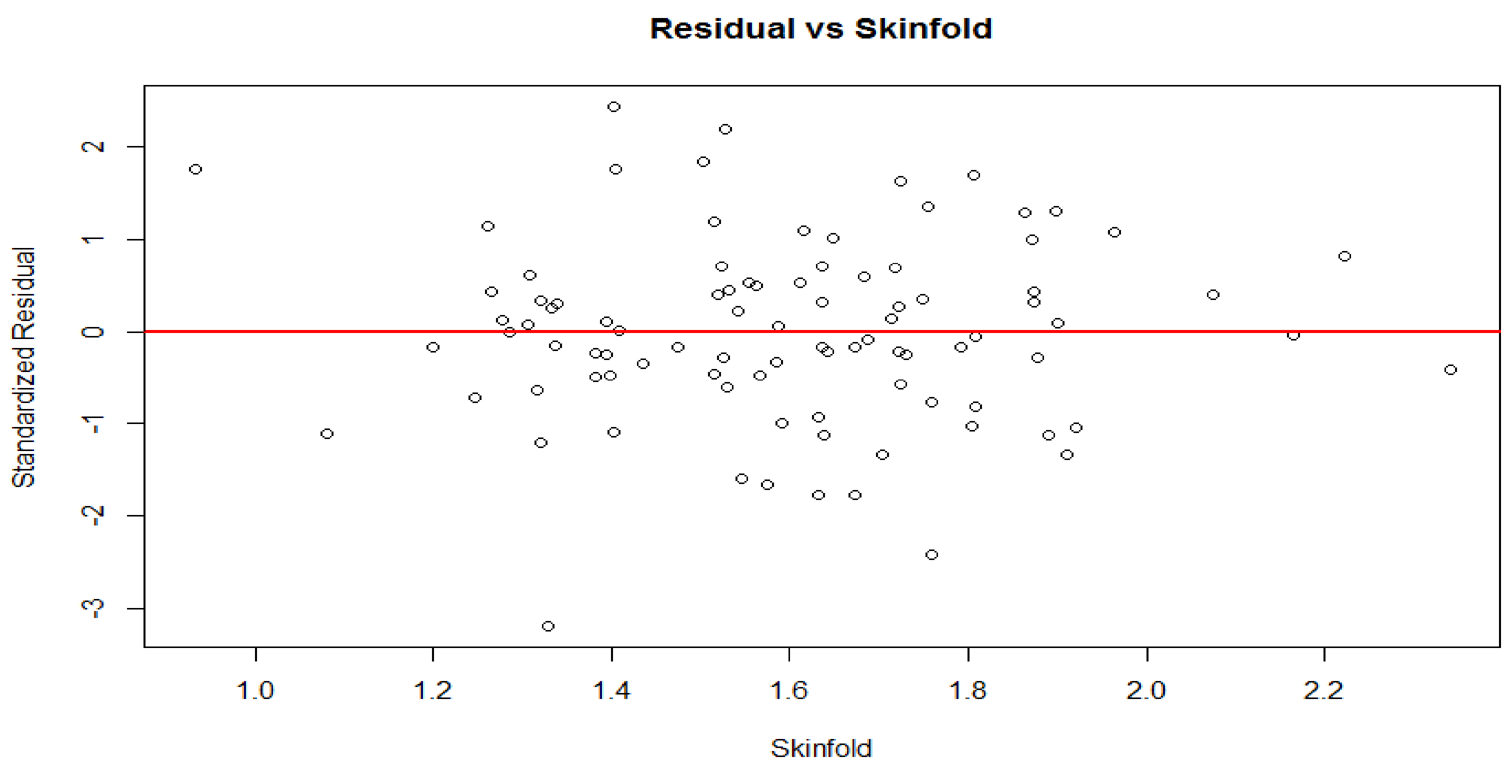
Plots to Check Model

```
plot(x,y,xlab="Skinfold",ylab="Body Density")
title(main="Scatterplot with Fitted Line")
abline(a=alphahat,b=betahat,col="red",lwd=2)
```



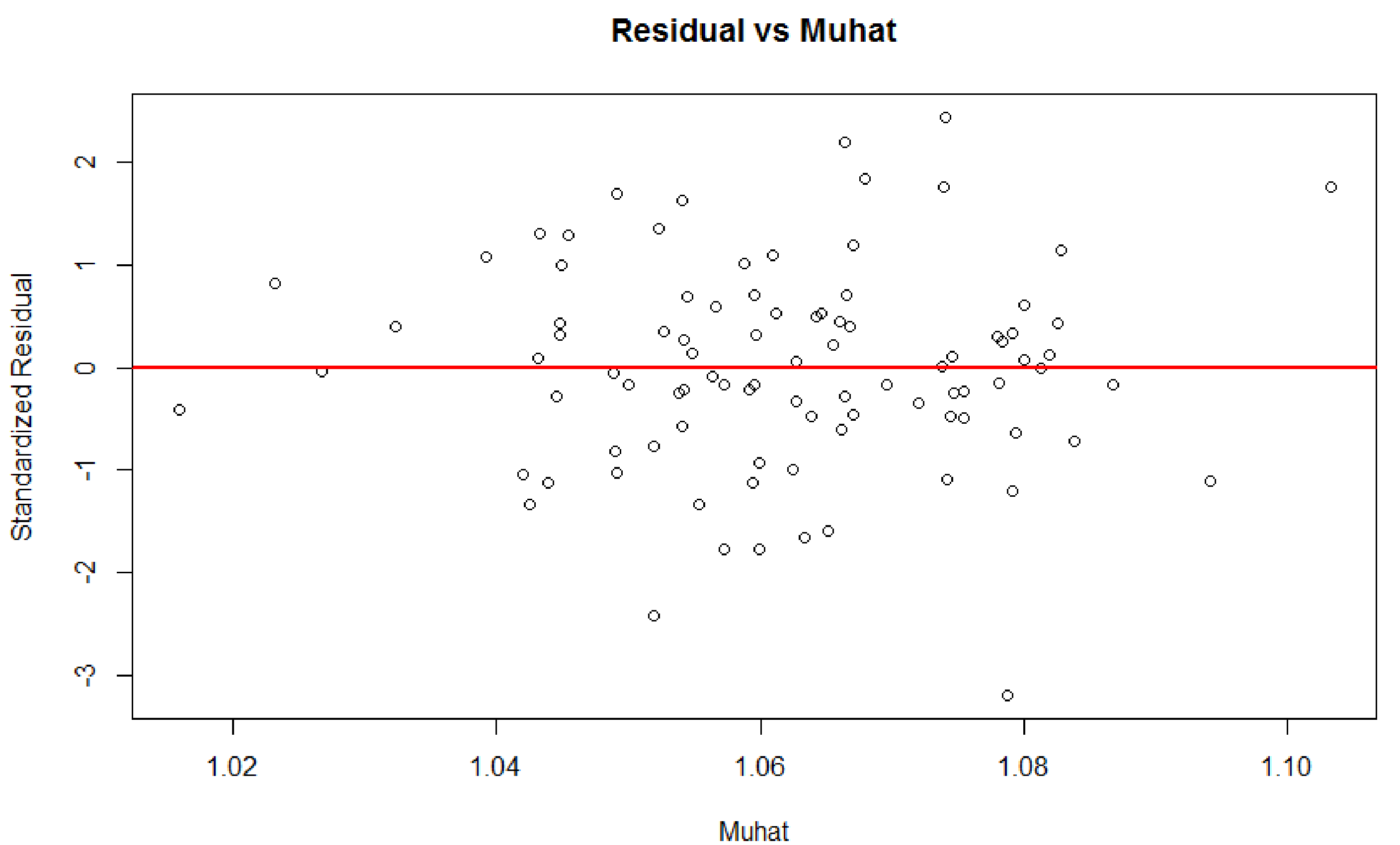
Plots to Check Model

```
rstar <- (RegModel$residuals)/se   # the standardized residuals
plot(x,rstar,xlab="Skinfold",ylab="Standardized Residual")
title(main="Residual vs Skinfold")
abline(a=0,b=0,col="red",lwd=2)
```



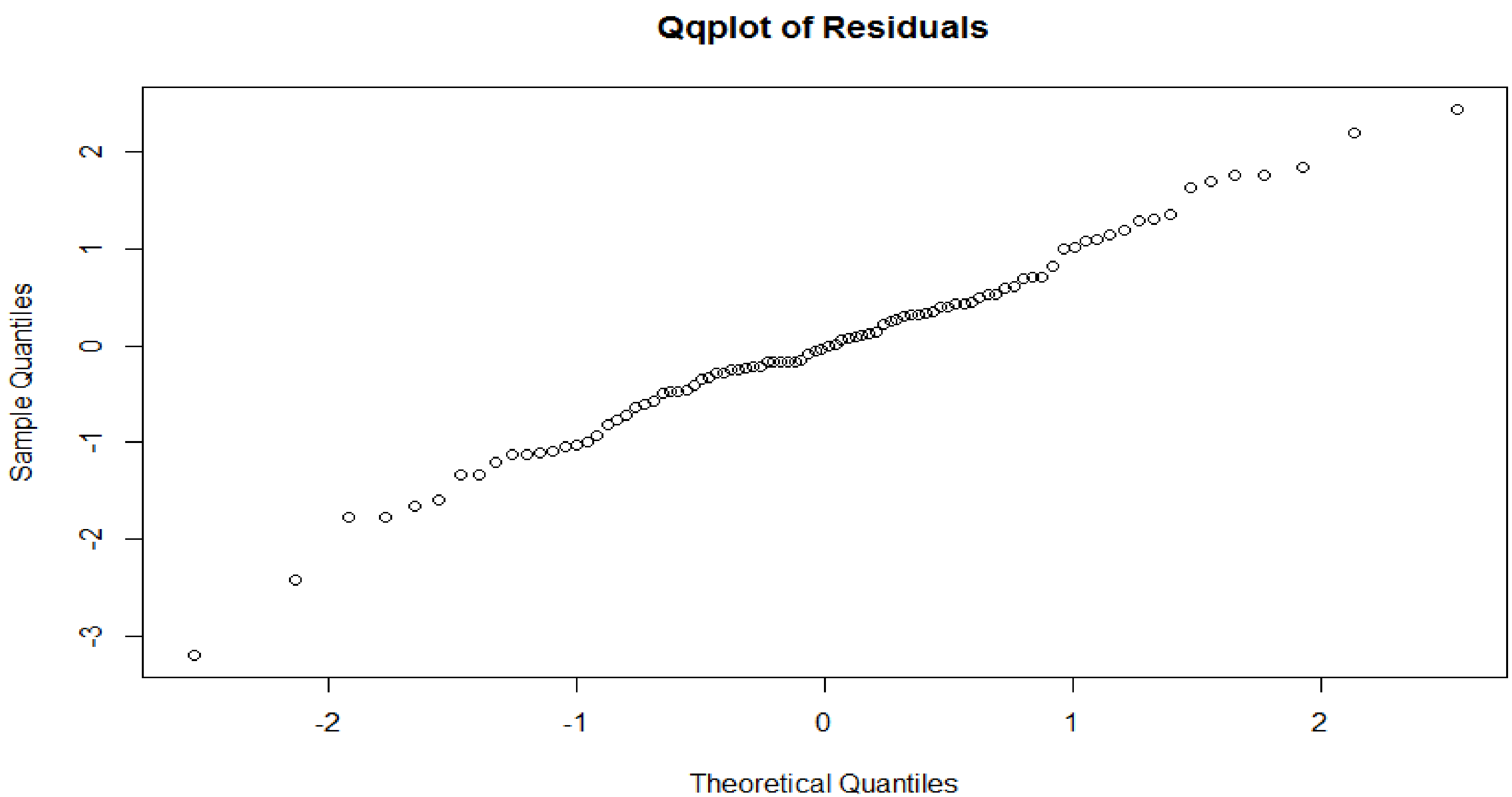
Plots to Check Model

```
plot(muhat,rstar,xlab="Muhat",ylab="Standardized Residual")
title(main="Residual vs Muhat")
abline(a=0,b=0,col="red",lwd=2)
```



Plots to Check Model

```
qqnorm(rstar,main="")
title(main="Qqplot of Residuals")
```



Problem 2

Suppose $y_1, y_2, \ldots, y_n$ is an observed random sample from the distribution with probability density function

$$f(y;\theta) = \frac{\theta}{y^{\theta+1}} \quad y \geq 1, \;\; \theta > 0$$

(a) Find the likelihood ratio test statistic for testing H_0: $\theta = \theta_0$.

(b) If $n = 30$ and $\prod_{i=1}^{30} y_i = 72$ then use the likelihood ratio test to test H_0: $\theta = 6$.