R作业5程序及结果(仅供参考)

April 13, 2016

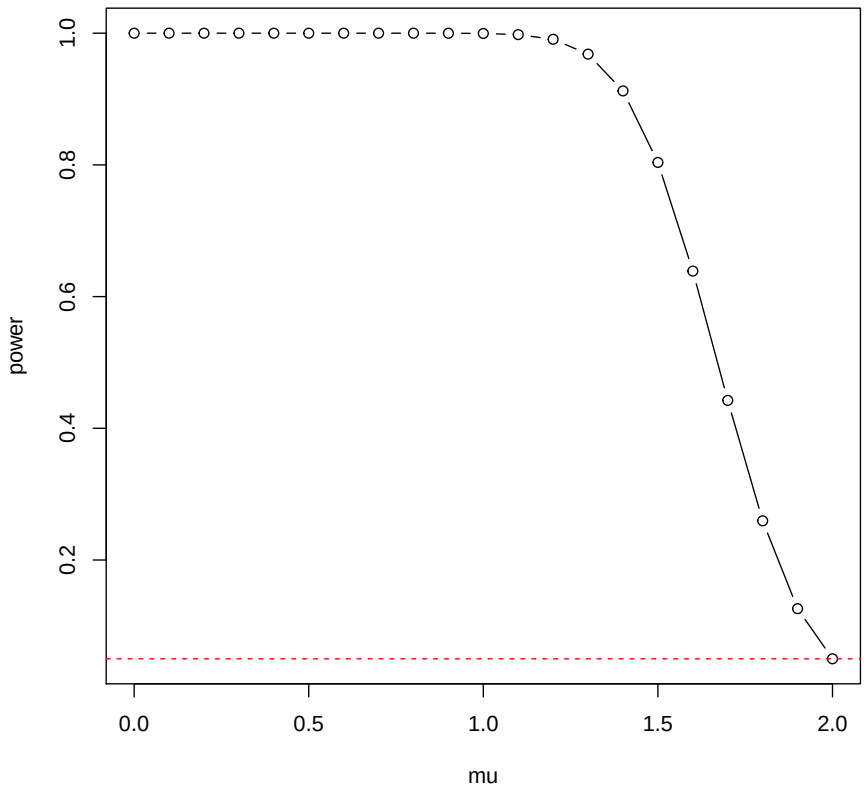
1. 单个正态总体方差已知时对均值的检验的功效函数
数据

```
mu0=2
sigma=2
n=100
x=rnorm(n,mu0,sigma)
alpha=0.05
```

(1)

$$H_0 : \mu = 2, \quad H_1 : \mu < 2.$$

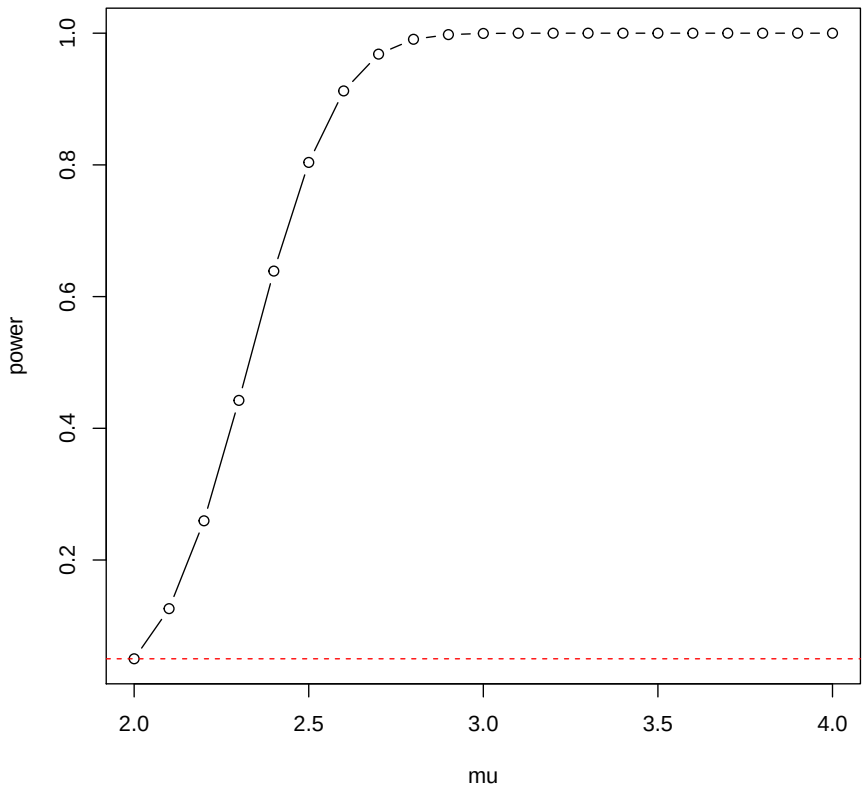
```
mu1=seq(0,2,0.1)
delt=(mu1-mu0)/(sigma/sqrt(n))
power=pnorm(qnorm(alpha)-delt)
plot(mu1,power,type="b",xlab="mu")
abline(h=alpha,col=2,lty=2)
```



(2)

$$\mathrm{H}_0 : \mu = 2, \quad \mathrm{H}_1 : \mu > 2.$$

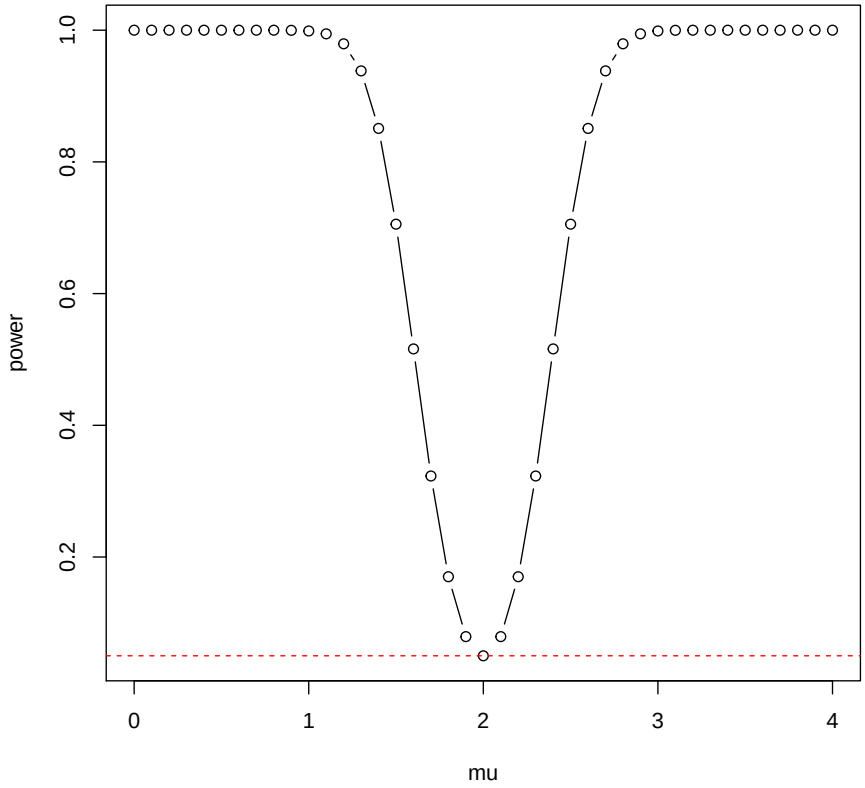
```
mu1=seq(2,4,0.1)
delt=(mu1-mu0)/(sigma/sqrt(n))
power=1-pnorm(qnorm(1-alpha)-delt)
plot(mu1,power,type="b",xlab="mu")
abline(h=alpha,col=2,lty=2)
```



(3)

$$\mathrm{H}_0 : \mu = 2, \ \ \mathrm{H}_1 : \mu \neq 2.$$

```
mu1=seq(0,4,0.1)
delt=(mu1-mu0)/(sigma/sqrt(n))
power=1-pnorm(qnorm(1-alpha/2)-delt)+
  pnorm(qnorm(alpha/2)-delt)
plot(mu1,power,type="b",xlab="mu")
abline(h=alpha,col=2,lty=2)
```



2. 单个正态总体方差未知时对均值的检验的功效函数
数据

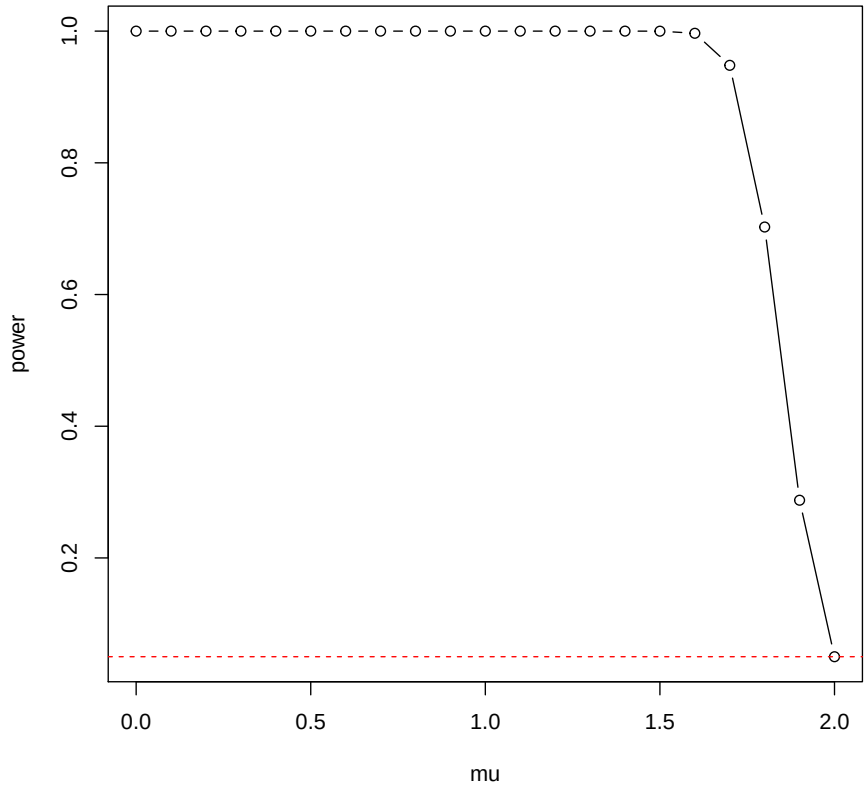
```
mu0=2
sigma=2
n=100
set.seed(123)
x=rnorm(n,mu0,sigma)
alpha=0.05
```

(1)

$$\mathrm{H}_0 : \mu = 2, \ \ \mathrm{H}_1 : \mu < 2.$$

```
mu1=seq(0,2,0.1)
delt=(mu1-mu0)/(sd(x)/sqrt(n))
```

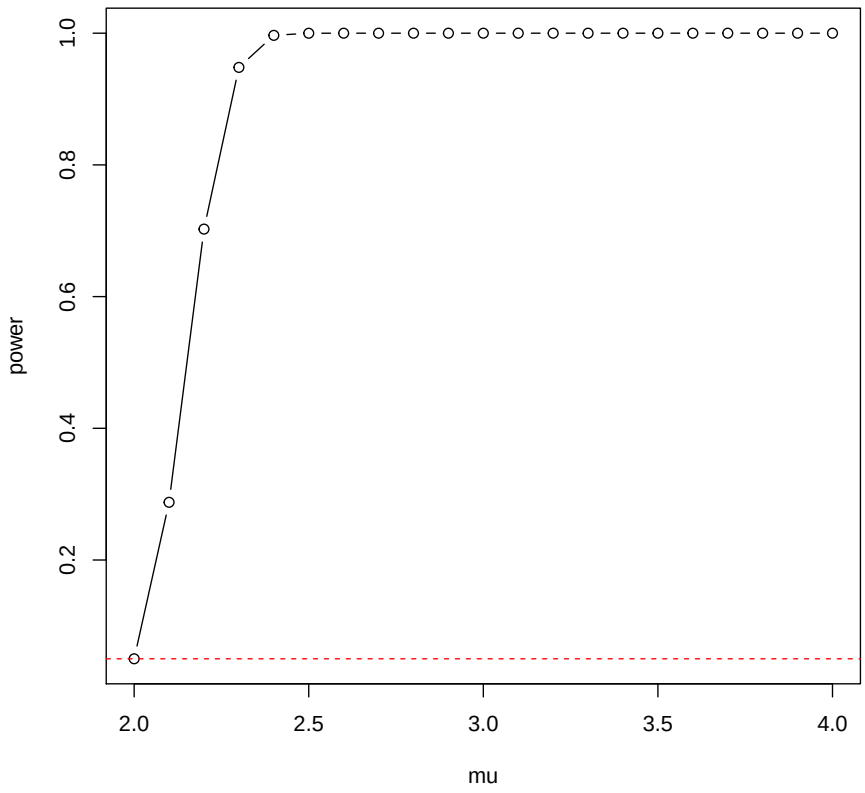
```r
power=pt(qt(alpha,n-1)-delt,n-1,delt) #delt为t分布的非中心参数
plot(mu1,power,type="b",xlab="mu")
abline(h=alpha,col=2,lty=2)
```



(2)

$$\mathrm{H}_0 : \mu = 2, \quad \mathrm{H}_1 : \mu > 2.$$

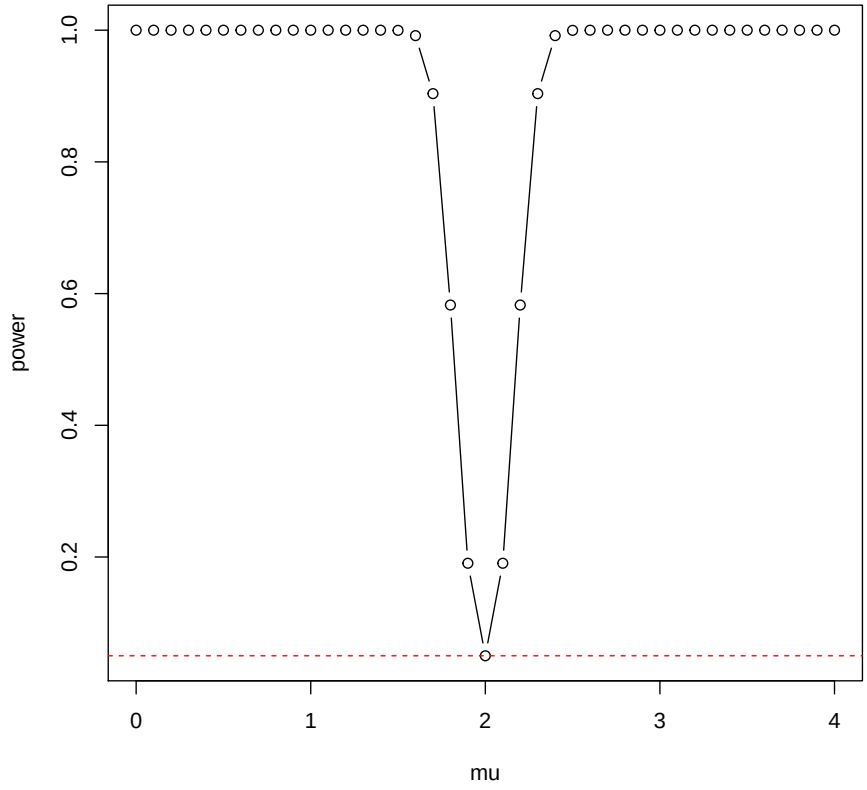
```r
mu1=seq(2,4,0.1)
delt=(mu1-mu0)/(sd(x)/sqrt(n))
power=1-pt(qt(1-alpha,n-1)-delt,n-1,delt)
plot(mu1,power,type="b",xlab="mu")
abline(h=alpha,col=2,lty=2)
```



(3)

$$\mathrm{H}_0 : \mu = 2, \quad \mathrm{H}_1 : \mu \neq 2.$$

```
mu1=seq(0,4,0.1)
delt=(mu1-mu0)/(sd(x)/sqrt(n))
power=1-pt(qt(1-alpha/2,n-1)-delt,n-1,delt)+
  pt(qt(alpha/2,n-1)-delt,n-1,delt)
plot(mu1,power,type="b",xlab="mu")
abline(h=alpha,col=2,lty=2)
```



3. 一元线性回归模型的统计推断

加载数据

```
load("RABE5.Rdata")
P054
```

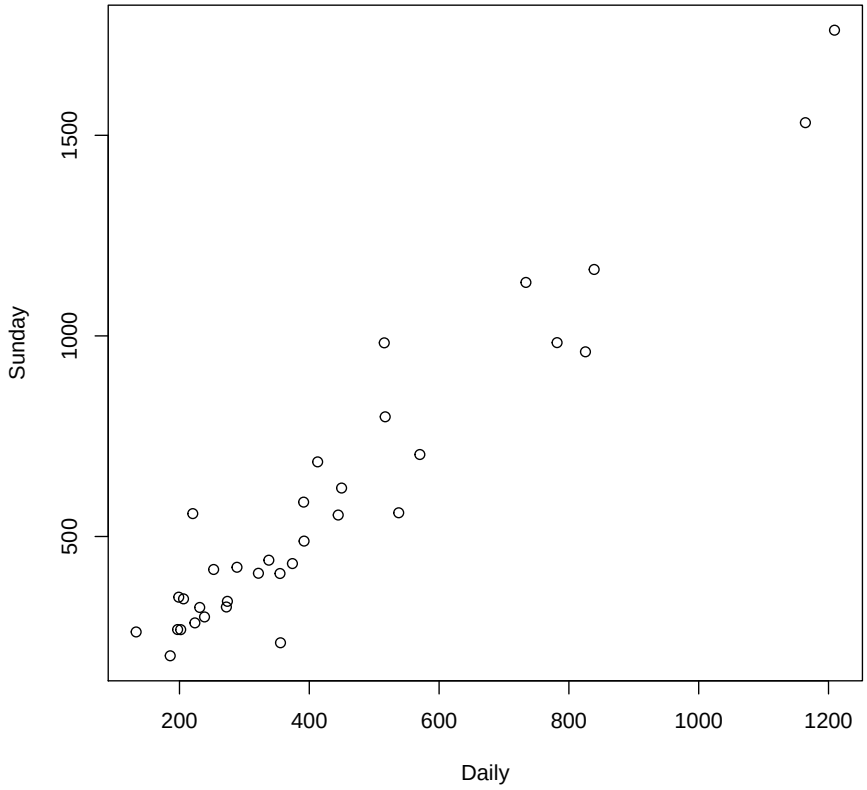
```
##                                  Daily   Sunday
## Baltimore Sun                  391.952  488.506
## Boston Globe                   516.981  798.298
## Boston Herald                  355.628  235.084
## Charlotte Observer             238.555  299.451
## Chicago Sun Times              537.780  559.093
## Chicago Tribune                733.775 1133.249
## Cincinnati Enquirer            198.832  348.744
## Denver Post                    252.624  417.779
## Des Moines Register            206.204  344.522
## Hartford Courant               231.177  323.084
```

```
## Houston Chronicle                449.755  620.752
## Kansas City Star                 288.571  423.305
## Los Angeles Daily  News          185.736  202.614
## Los Angeles Times               1164.388 1531.527
## Miami Herald                     444.581  553.479
## Minneapolis Star Tribune         412.871  685.975
## New Orleans Times-Picayune       272.280  324.241
## New York Daily News              781.796  983.240
## New York Times                  1209.225 1762.015
## Newsday                          825.512  960.308
## Omaha World Herald               223.748  284.611
## Orange County Register           354.843  407.760
## Philadelphia Inquirer            515.523  982.663
## Pittsburgh Press                 220.465  557.000
## Portland Oregonian               337.672  440.923
## Providence Journal-Bulletin      197.120  268.060
## Rochester Democrat & Chronicle   133.239  262.048
## Rocky Mountain News              374.009  432.502
## Sacramento Bee                   273.844  338.355
## San Francisco Chronicle          570.364  704.322
## St. Louis Post-Dispatch          391.286  585.681
## St. Paul Pioneer Press           201.860  267.781
## Tampa Tribune                    321.626  408.343
## Washington Post                  838.902 1165.567

attach(P054)
```

(1) 散点图

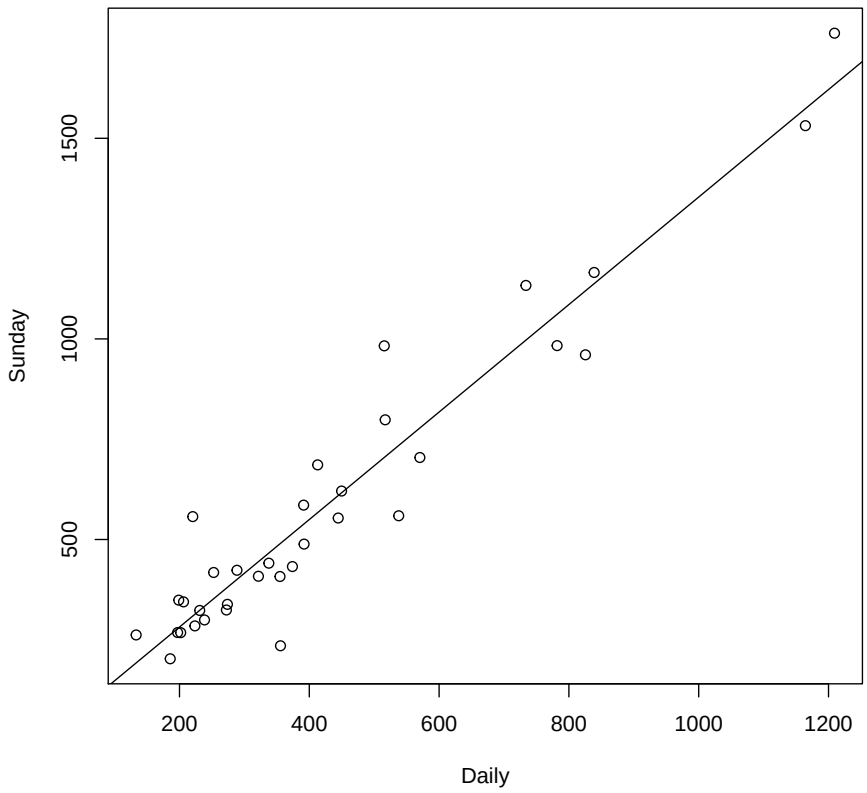
```
plot(Daily,Sunday)
```



该图显示二者有较好的线性关系

(2) 作回归，添加回归直线，作预测

```
plot(Daily,Sunday)
lm.sol=lm(Sunday~Daily)
abline(lm.sol)
```




```
predict(lm.sol)
```

```
       1        2        3        4        5        6        7
538.9395 706.4427 490.2757 333.4313 734.3074 996.8848 280.2138
       8        9       10       11       12       13       14
352.2797 290.0902 323.5469 616.3790 400.4385 262.6689 1573.7834
      15       16       17       18       19       20       21
609.4474 566.9650 378.6132 1061.2193 1633.8522 1119.7862 313.5941
      22       23       24       25       26       27       28
489.2240 704.4894 309.1958 466.2198 277.9202 192.3379 514.9010
      29       30       31       32       33       34
380.7085 777.9607 538.0473 284.2705 444.7227 1137.7250
```

(3) 求置信区间

```
beta.int=function(fm,alpha=0.05){
  A=summary(fm)$coefficients
  df=fm$df.residual
  left=A[,1]-A[,2]*qt(1-alpha/2,df)
  right=A[,1]+A[,2]*qt(1-alpha/2,df)
  rowname=dimnames(A)[[1]]
  colname=c("Estimate","Left","Right")
  matrix(c(A[,1],left,right),ncol=3,dimnames=list(rowname,colname))
}

beta.int(lm.sol)

##               Estimate       Left     Right
## (Intercept) 13.835630 -59.094743 86.766003
## Daily        1.339715   1.195594  1.483836
```

(4) 假设检验

```
summary(lm.sol)

##
## Call:
## lm(formula = Sunday ~ Daily)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -255.19  -55.57  -20.89   62.73  278.17
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 13.83563   35.80401   0.386    0.702
## Daily        1.33971    0.07075  18.935   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 109.4 on 32 degrees of freedom
## Multiple R-squared:  0.9181,Adjusted R-squared:  0.9155
## F-statistic: 358.5 on 1 and 32 DF,  p-value: < 2.2e-16
```

结果显示对假设$\mathrm{H}_0:\beta_1=0,\ \mathrm{H}_1:\beta_1\neq 0$，检验统计量的观测值为18.935，远远大于$t_{1-\alpha/2}(32)=2.037$，检验的p值小于$10^{-16}$，极度显著。说明周日发行量与平日发行量之间有显著的线性关系。

(5) 解释

由summary的结果可知，$R^2=0.9181$，故周日发行量的变化中能由平日发行量解释的比例为0.9181。

(6) 周日发行量均值的置信区间

```
new=data.frame(Daily=500)
predict(lm.sol,new,interval="confidence",level=0.95)

      fit      lwr     upr
1 683.693 644.1951 723.191
```

(7) Daily=500 时，周日发行量的预测区间

```
new=data.frame(Daily=500)
predict(lm.sol,new,interval="prediction",level=0.95)

      fit      lwr      upr
1 683.693 457.3367 910.0493
```

该预测区间与(6)中的置信区间相比，要大很多。

(8) Daily=2000 时，周日发行量的预测区间

```
new=data.frame(Daily=2000)
predict(lm.sol,new,interval="prediction",level=0.95)

       fit      lwr      upr
1 2693.265 2373.463 3013.068
```

该预测区间与(7)中的预测区间相比，要大很多，不如(7)中的精确。