Bank Example, polynomial regression

Data from Table 3.5

A sample of branch offices of a commercial bank is selected to study the relationship between the size of the minimum deposit necessary to receive a free gift (x) and the number of new accounts opened at that office (y)

:

```
deposit.size = c(75,75,100,100,125,125,150,175,175,200,200)
new.accounts = c(28,42,112,136,160,150,152,156,124,124,104)
bank.data = data.frame(deposit.size,new.accounts)


attach(bank.data)

sat.mod = lm(new.accounts ~ poly(deposit.size,5))
summary(sat.mod)

##
## Call:
## lm(formula = new.accounts ~ poly(deposit.size, 5))
##
## Residuals:
##          1          2          3          4          5          6
7
## -7.000e+00  7.000e+00 -1.200e+01  1.200e+01  5.000e+00 -5.000e+00  3.464e-
14
##          8          9         10         11
##  1.600e+01 -1.600e+01  1.000e+01 -1.000e+01
##
## Coefficients:
##                        Estimate Std. Error t value Pr(>|t|)
## (Intercept)             117.091      4.569  25.629 1.69e-06 ***
## poly(deposit.size, 5)1   71.703     15.153   4.732 0.005186 **
## poly(deposit.size, 5)2 -112.598     15.153  -7.431 0.000696 ***
## poly(deposit.size, 5)3   29.356     15.153   1.937 0.110442
## poly(deposit.size, 5)4   -6.934     15.153  -0.458 0.666434
## poly(deposit.size, 5)5   -2.337     15.153  -0.154 0.883474
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 15.15 on 5 degrees of freedom
## Multiple R-squared:  0.9423, Adjusted R-squared:  0.8845
## F-statistic: 16.32 on 5 and 5 DF,  p-value: 0.004085
```

```
three.mod = lm(new.accounts ~ poly(deposit.size,3))
summary(three.mod)

##
## Call:
## lm(formula = new.accounts ~ poly(deposit.size, 3))
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -13.163  -8.693  -4.289   7.617  18.837
##
## Coefficients:
##                         Estimate Std. Error t value Pr(>|t|)
## (Intercept)               117.09       3.95  29.641 1.28e-08 ***
## poly(deposit.size, 3)1     71.70      13.10   5.473 0.000933 ***
## poly(deposit.size, 3)2   -112.60      13.10  -8.594 5.75e-05 ***
## poly(deposit.size, 3)3     29.36      13.10   2.241 0.060018 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 13.1 on 7 degrees of freedom
## Multiple R-squared:  0.9396, Adjusted R-squared:  0.9137
## F-statistic: 36.28 on 3 and 7 DF,  p-value: 0.0001233

quad.mod = lm(new.accounts ~ poly(deposit.size,2))
summary(quad.mod)

##
## Call:
## lm(formula = new.accounts ~ poly(deposit.size, 2))
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -25.306  -7.535  -1.242   7.666  24.927
##
## Coefficients:
##                         Estimate Std. Error t value Pr(>|t|)
## (Intercept)              117.091      4.842  24.181 9.12e-09 ***
## poly(deposit.size, 2)1    71.703     16.060   4.465 0.002098 **
## poly(deposit.size, 2)2  -112.598     16.060  -7.011 0.000111 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 16.06 on 8 degrees of freedom
## Multiple R-squared:  0.8962, Adjusted R-squared:  0.8703
## F-statistic: 34.55 on 2 and 8 DF,  p-value: 0.000116
```

```
newdat = data.frame(deposit.size = seq(min(deposit.size)-50,
max(deposit.size)+50, length.out = 100))
newdat$sat.pred = predict(sat.mod, newdata = newdat)
newdat$three.pred = predict(three.mod, newdata = newdat)
newdat$quad.pred = predict(quad.mod, newdata = newdat)

plot(bank.data$deposit.size,bank.data$new.accounts,
     xlim = c(min(deposit.size)-50, max(deposit.size)+50),ylim = c(0,180),
     xlab = "deposit size",ylab = "new accounts")
points(newdat$deposit.size,newdat$sat.pred,type="l",col="blue")
points(newdat$deposit.size,newdat$three.pred,type="l",col="red")
points(newdat$deposit.size,newdat$quad.pred,type="l",col="green")
```
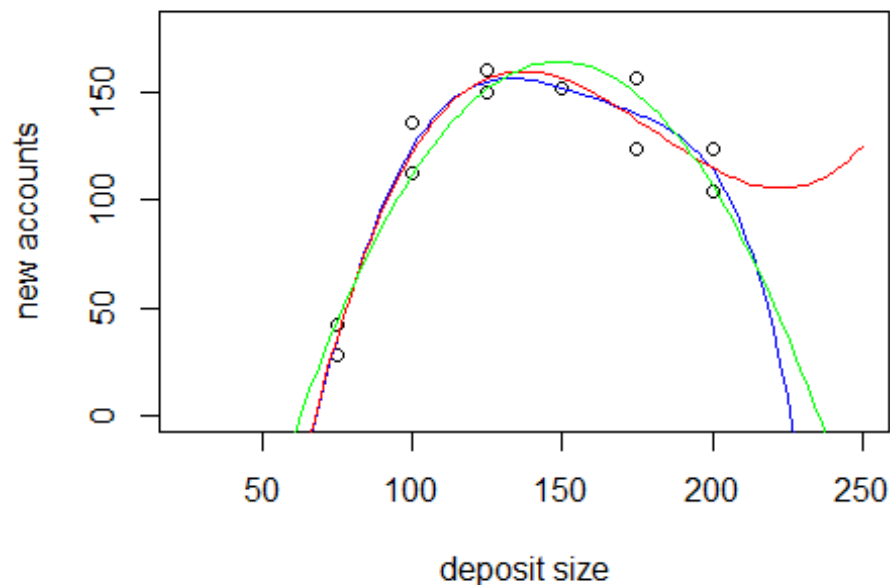
```
anova(quad.mod,sat.mod)

## Analysis of Variance Table
##
## Model 1: new.accounts ~ poly(deposit.size, 2)
## Model 2: new.accounts ~ poly(deposit.size, 5)
##   Res.Df    RSS Df Sum of Sq      F Pr(>F)
## 1      8 2063.3
## 2      5 1148.0  3    915.32 1.3289 0.3635

accept.mod = lm(new.accounts ~ deposit.size + I(deposit.size^2) +
I(deposit.size^3))
summary(accept.mod)

##
## Call:
## lm(formula = new.accounts ~ deposit.size + I(deposit.size^2) +
##     I(deposit.size^3))
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -13.163  -8.693  -4.289   7.617  18.837
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)      -6.855e+02  1.670e+02  -4.105  0.00455 **
## deposit.size      1.553e+01  4.064e+00   3.822  0.00652 **
## I(deposit.size^2) -9.160e-02  3.115e-02  -2.940  0.02170 *
## I(deposit.size^3)  1.697e-04  7.575e-05   2.241  0.06002 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 13.1 on 7 degrees of freedom
## Multiple R-squared:  0.9396, Adjusted R-squared:  0.9137
## F-statistic: 36.28 on 3 and 7 DF,  p-value: 0.0001233
```