The R Package regtools and the Mystery of P-Values

Norm Matloff
University of California at Davis

Central Iowa R User Group

April 28, 2016

These slides at `http://heather.cs.ucdavis.edu/Iowa.pdf`

Two Rather Distant Topics

Norm Matloff
University of California at Davis

Two Rather Distant Topics

I'll talk about two very different topics:

Two Rather Distant Topics

I'll talk about two very different topics:

- Introduction to my R package **regtools**, especially in terms of regression diagnostics.

Two Rather Distant Topics

I'll talk about two very different topics:

- Introduction to my R package **regtools**, especially in terms of regression diagnostics.
- Comments on the dramatic recent ASA announcement regarding p-values.

regtools

regtools

R package for regression and classification.

Norm Matloff
University of California at Davis

regtools

R package for regression and classification.

- On my GitHub site, `https://github.com/matloff`

regtools

R package for regression and classification.

- On my GitHub site, `https://github.com/matloff`
- “Various tools for linear, nonlinear and nonparametric regression.”

regtools

R package for regression and classification.

- On my GitHub site, `https://github.com/matloff`
- "Various tools for linear, nonlinear and nonparametric regression."
- Meant to accompany my forthcoming book, *From Linear Models to Machine Learning: Regression and Classification, with R Examples*.

regtools

R package for regression and classification.

- On my GitHub site, `https://github.com/matloff`
- "Various tools for linear, nonlinear and nonparametric regression."
- Meant to accompany my forthcoming book, *From Linear Models to Machine Learning: Regression and Classification, with R Examples*.
- In many senses, both the package and the book *take a very nontraditional point of view.*

The Book

The Book

The Book

50% rough draft at
`http://heather.cs.ucdavis.edu/draftregclass.pdf`.

The Book

50% rough draft at `http://heather.cs.ucdavis.edu/draftregclass.pdf`. Takes unconventional views, such as:

Norm Matloff University of California at Davis

Central Iowa R User Group

The Book

50% rough draft at `http://heather.cs.ucdavis.edu/draftregclass.pdf`.
Takes unconventional views, such as:

- Inference procedures based on normality assumptions for ϵ greatly de-emphasized.

The Book

50% rough draft at
`http://heather.cs.ucdavis.edu/draftregclass.pdf`.
Takes unconventional views, such as:

- Inference procedures based on normality assumptions for ϵ greatly de-emphasized.
- Use of transformations discouraged.

The Book

50% rough draft at `http://heather.cs.ucdavis.edu/draftregclass.pdf`. Takes unconventional views, such as:

- Inference procedures based on normality assumptions for ϵ greatly de-emphasized.
- Use of transformations discouraged.
- A full chapter on measuring and interpreting factor effects.

The Book

50% rough draft at `http://heather.cs.ucdavis.edu/draftregclass.pdf`. Takes unconventional views, such as:

- Inference procedures based on normality assumptions for ϵ greatly de-emphasized.
- Use of transformations discouraged.
- A full chapter on measuring and interpreting factor effects.
- Debunks incorrect views of unbalanced classification problems.

The Book

50% rough draft at
`http://heather.cs.ucdavis.edu/draftregclass.pdf`.
Takes unconventional views, such as:

- Inference procedures based on normality assumptions for ϵ greatly de-emphasized.
- Use of transformations discouraged.
- A full chapter on measuring and interpreting factor effects.
- Debunks incorrect views of unbalanced classification problems.
- Interweaves nonparametric methods with linear and nonlinear parametric regression models.

Norm Matloff
University of California at Davis

Central Iowa R User Group

The regtools Package

Norm Matloff
University of California at Davis

Central Iowa R User Group

The regtools Package

- Written for the book, but usable by all.

The regtools Package

- Written for the book, but usable by all.
- As with the book, interweaves nonparametric methods with linear and nonlinear parametric regression models.

The regtools Package

- Written for the book, but usable by all.
- As with the book, interweaves nonparametric methods with linear and nonlinear parametric regression models.
- Includes some unusual functions, both in the sense of new ways of doing old things, and ways of doing new things.

The regtools Package

- Written for the book, but usable by all.
- As with the book, interweaves nonparametric methods with linear and nonlinear parametric regression models.
- Includes some unusual functions, both in the sense of new ways of doing old things, and ways of doing new things.
- Work in progress, adding more functions over time.

Norm Matloff
University of California at Davis

Central Iowa R User Group

Example: Multiclass Classification

Example: Multiclass Classification

- One-vs.-All or All-vs.-All?

Norm Matloff
University of California at Davis

Central Iowa R User Group

Example: Multiclass Classification

- One-vs.-All or All-vs.-All? Have code for both, argues (somewhat) in favor of OVA.

Norm Matloff
University of California at Davis

Example: Multiclass Classification

- One-vs.-All or All-vs.-All? Have code for both, argues (somewhat) in favor of OVA.
- Example: UCI vertebrae data. Logit model, 3 classes.

Example: Multiclass Classification

- One-vs.-All or All-vs.-All? Have code for both, argues (somewhat) in favor of OVA.
- Example: UCI vertebrae data. Logit model, 3 classes.

```
> trnidxs <- sample(1:310,225)
> predidxs <- setdiff(1:310,trnidxs)
> ovout <- ovalogtrn(3,vert[trnidxs,])
> predy <- ovalogpred(ovout,vert[predidxs,1:6])
> mean(predy == vert[predidxs,7])
[1] 0.8823529
> avout <- avalogtrn(3,vert[trnidxs,])
> predy <- avalogpred(3,avout,vert[predidxs,1:6])
> mean(predy == vert[predidxs,7])
[1] 0.8588235
```

Example: Multiclass Classification

- One-vs.-All or All-vs.-All? Have code for both, argues (somewhat) in favor of OVA.
- Example: UCI vertebrae data. Logit model, 3 classes.

```
> trnidxs <- sample(1:310,225)
> predidxs <- setdiff(1:310,trnidxs)
> ovout <- ovalogtrn(3,vert[trnidxs,])
> predy <- ovalogpred(ovout,vert[predidxs,1:6])
> mean(predy == vert[predidxs,7])
[1] 0.8823529
> avout <- avalogtrn(3,vert[trnidxs,])
> predy <- avalogpred(3,avout,vert[predidxs,1:6])
> mean(predy == vert[predidxs,7])
[1] 0.8588235
```

Similar success rates, but AVA much more computation.

Example: Letter Recognition

Example: Letter Recognition

- Again, famous UCI data set.

Example: Letter Recognition

- Again, famous UCI data set.
- But, classes are artificially balanced, not reflecting that, e.g., 'z' is much rarer in English than 'a'.

Example: Letter Recognition

- Again, famous UCI data set.
- But, classes are artificially balanced, not reflecting that, e.g., 'z' is much rarer in English than 'a'. The book derives adjustment procedures, implemented in **ova*()**.

Example: Letter Recognition

- Again, famous UCI data set.
- But, classes are artificially balanced, not reflecting that, e.g., 'z' is much rarer in English than 'a'. The book derives adjustment procedures, implemented in **ova*()**.

```
> data(ltrfreqs)
> ltrfreqs <- ltrfreqs[order(ltrfreqs[,1]),]
> truepriors <- ltrfreqs[,2]/100  # not Bayesian!
# success rate with sample priors 0.75
> trnout1 <- ovaknntrn(lrtrn[,1],xdata,26,50,
     truepriors)
> ypred <- ovaknnpred(trnout1,lrtest1[,-1])
> mean(ypred == lrtest1[,1])
[1] 0.8787988  # nice!
```

Norm Matloff
University of California at Davis

Model Fit Assessment

Norm Matloff
University of California at Davis

Central Iowa R User Group

Model Fit Assessment

- Classical approach: Plot residuals.

Norm Matloff University of California at Davis

Central Iowa R User Group

Model Fit Assessment

- Classical approach: Plot residuals.
- NM book/**regtools** approach: Use the nonparametric to help assess the parametric.

Example: Currency Data (Fong & Ouliaris, 1995)

Example: Currency Data (Fong & Ouliaris, 1995)

- Predict *yen* from Can. $, *mark*, *franc*, *pound*.

Example: Currency Data (Fong & Ouliaris, 1995)

- Predict *yen* from Can. $, *mark*, *franc*, *pound*.
- Linear model:

Example: Currency Data (Fong & Ouliaris, 1995)

- Predict *yen* from Can. $, *mark*, *franc*, *pound*.
- Linear model:

```
> fout <- lm(Yen ~ ., data=cur1)
> summary(fout)
...
             Estimate Std. Error t value Pr(>|t|)
(Intercept)   102.855     14.663   7.015 5.12e-12
Can           -45.941     11.979  -3.835 0.000136
Mark          147.328      3.325  44.313  < 2e-16
Franc         -21.790      1.463 -14.893  < 2e-16
Pound         -48.771     14.553  -3.351 0.000844
...
Mult. R-squared:  0.8923,  Adj. R-squared:  0.8918
```

Currency Data, contd.

Currency Data, contd.

- Nonparametric fit: k-NN, k det. by cross-val:

```
> xdata ← preprocessx(curr1[,−5],150,xval=TRUE)
> kminout ← kmin(curr1$Yen,xdata,predwrong,nk=30
> kminout$kmin
[1] 5
> kout ← knnest(curr1[,5],xdata,5)
> cor(kout$regest,curr1[,5])^2
[1] 0.9920137
```

Currency Data, contd.

- Nonparametric fit: k-NN, k det. by cross-val:

```
> xdata ← preprocessx(curr1[,−5],150,xval=TRUE)
> kminout ← kmin(curr1$Yen,xdata,predwrong,nk=30
> kminout$kmin
[1] 5
> kout ← knnest(curr1[,5],xdata,5)
> cor(kout$regest,curr1[,5])^2
[1] 0.9920137
```

We're "leaving money on the table"!

Currency

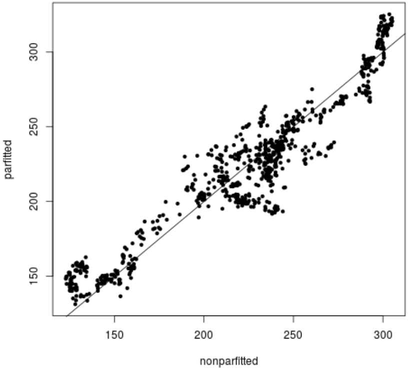
Currency

Plot parametric vs. nonparametric fits:

Currency

Plot parametric vs. nonparametric fits:

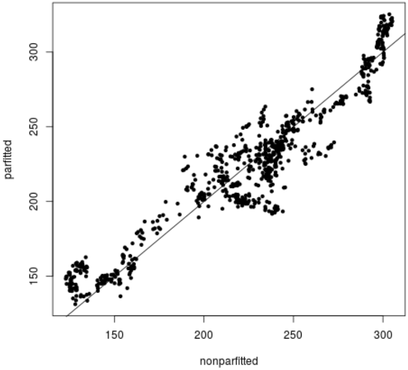
```
> parvsnonparplot(fout1,kout)
```



Currency

Plot parametric vs. nonparametric fits:

```
> parvsnonparplot(fout1,kout)
```

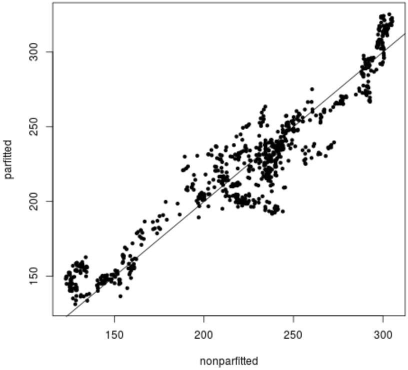


Interesting features, especially "hook" at the left end and "tail" near the middle.

Currency

Plot parametric vs. nonparametric fits:

```
> parvsnonparplot(fout1,kout)
```



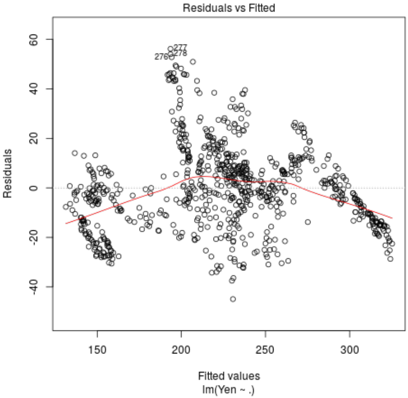
Interesting features, especially "hook" at the left end and "tail" near the middle. Bring in the domain experts!

Currency

Currency

R builit-in plot:

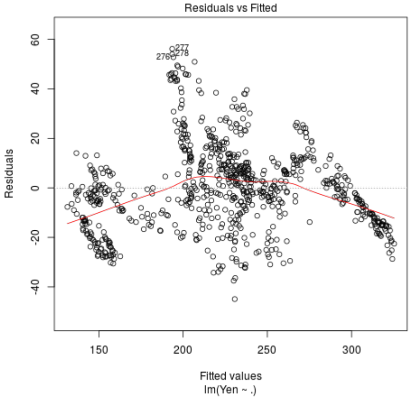
```
> plot(lmout)
```



Currency

R builit-in plot:

```
> plot(lmout)
```



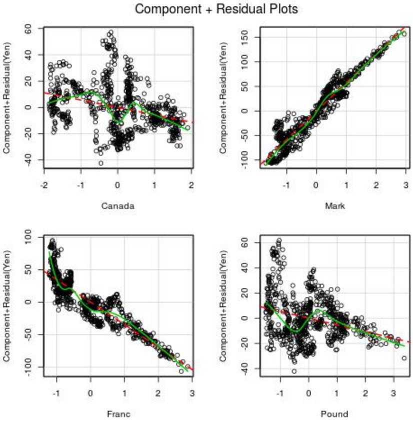
Hook, tail visible here too, but arguably less clearly.

Currency

Currency

Draw partial residual plots, using the **car** package.

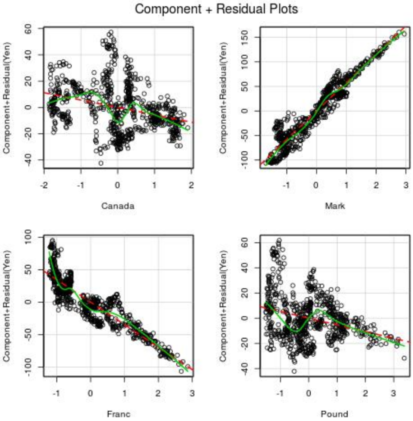
```
> crPlots(fout1)
```



Currency

Draw partial residual plots, using the **car** package.

```
> crPlots(fout1)
```



Mark plot looks rather "clean." And yet...

Norm Matloff
University of California at Davis

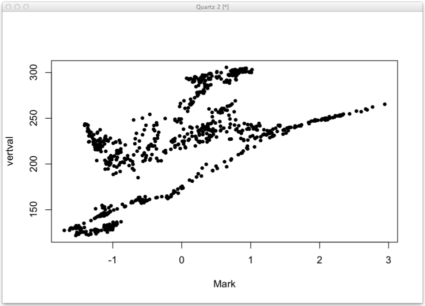
Central Iowa R User Group

Currency

Currency

Drawing a similar plot from **regtools**, which uses smoothing:

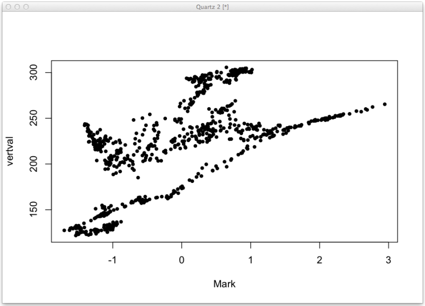
```
> nonparvsxplot(kout)
next plot
next plot
next plot
next plot
```



Currency

Drawing a similar plot from **regtools**, which uses smoothing:

```
> nonparvsxplot(kout)
next plot
next plot
next plot
next plot
```

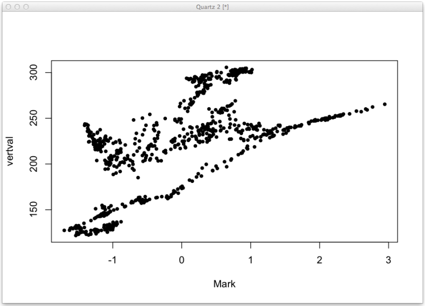


Not so clean at all!

Currency

Drawing a similar plot from **regtools**, which uses smoothing:

```
> nonparvsxplot(kout)
next plot
next plot
next plot
next plot
```



Not so clean at all! Again, need domain experts.

P-Values

P-Values

- Many of us have been saying for years that p-values and hypothesis tests

P-Values

- Many of us have been saying for years that p-values and hypothesis tests
  - are addressing an irrelevant question, and
  - can be highly misleading.
- The problems are well known. Even people like Einstein and Feynman complained.

P-Values

- Many of us have been saying for years that p-values and hypothesis tests
  - are addressing an irrelevant question, and
  - can be highly misleading.
- The problems are well known. Even people like Einstein and Feynman complained.
- Yet they are ENTRENCHED in statistical education and practice.

P-Values

- Many of us have been saying for years that p-values and hypothesis tests
  - are addressing an irrelevant question, and
  - can be highly misleading.
- The problems are well known. Even people like Einstein and Feynman complained.
- Yet they are ENTRENCHED in statistical education and practice.
- Fortunately, ASA decided to address the issue recently.

P-Values

- Many of us have been saying for years that p-values and hypothesis tests
  - are addressing an irrelevant question, and
  - can be highly misleading.
- The problems are well known. Even people like Einstein and Feynman complained.
- Yet they are ENTRENCHED in statistical education and practice.
- Fortunately, ASA decided to address the issue recently.
- The report is “a camel designed by a committee,” thus not as strong as it should be.

P-Values

- Many of us have been saying for years that p-values and hypothesis tests
  - are addressing an irrelevant question, and
  - can be highly misleading.
- The problems are well known. Even people like Einstein and Feynman complained.
- Yet they are ENTRENCHED in statistical education and practice.
- Fortunately, ASA decided to address the issue recently.
- The report is "a camel designed by a committee," thus not as strong as it should be. But at the very least, one can say that the report is very negative about typical usage today of p-values.

Norm Matloff University of California at Davis

Example: MovieLens data.

Example: MovieLens data.

```
> head(uu)
  userid age gender      occup   zip  avg_rat
1      1  24      0 technician 85711 3.610294
2      2  53      0      other 94043 3.709677
> q <- lm(avgrat ~ age + gender, data=uu)
> summary(q)
...
Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) 3.4725821  0.0482655  71.947  < 2e-16 **
age         0.0033891  0.0011860   2.858  0.00436 **
gender      0.0002862  0.0318670   0.009  0.99284
...
Multiple R-squared:  0.008615,  Adjusted R-squared:
0.006505
```

Example: MovieLens data.

```
> head(uu)
  userid age gender        occup   zip  avg_rat
1      1  24      0 technician 85711 3.610294
2      2  53      0      other 94043 3.709677
> q <- lm(avgrat ~ age + gender, data=uu)
> summary(q)
...
Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) 3.4725821  0.0482655  71.947  < 2e-16 **
age         0.0033891  0.0011860   2.858  0.00436 **
gender      0.0002862  0.0318670   0.009  0.99284
...
Multiple R-squared:  0.008615,  Adjusted R-squared:
0.006505
```

Age effect is “highly significant” — yet **highly unimportant**.

Problems with Change

Problems with Change

- People like **simple, crisp answers** ("A is signficantly related to B"), not messy hedging ("Well, there seems to be a mild effect but...").

Problems with Change

- People like **simple, crisp answers** ("A is signficantly related to B"), not messy hedging ("Well, there seems to be a mild effect but...").
- Stat instructors like telling "favorite bedtime stories" to their kids, and **don't want to change**

Norm Matloff University of California at Davis

Central Iowa R User Group

Problems with Change

- People like **simple, crisp answers** ("A is signficantly related to B"), not messy hedging ("Well, there seems to be a mild effect but..").
- Stat instructors like telling "favorite bedtime stories" to their kids, and **don't want to change**
- Will the ASA statement have any effect????