Carols for a Cornell Stats-Mass

Inspired by the Cornell Glee Club and the Statistics Holiday Party 2019.

Sigma and Mu

(All I Want for Christmas)

I don’t want a lot from data
There are just two things I need
I don’t need no fancy models
For some skedasticity
There's so many quantities
Estimates and summaries
But I need just two
All I want are sigma and mu

I don’t want a lot from data
There are just two things I need
I don't care about robustness
or trying to trim my means
I don't need no rank-based testi
Nor to use some divergence
Huberized won't make it better
Nor a biweight estimator
There's so many quantities
Estimates and summaries
But I need just two
All I want are sigma and mu

I won't look at a regression
I won't even use anova
I'm just going report numbers
That I really get to know
I won't make a list and send it
Of pvalues or statistics
No false discovery rating
Is ever going to click
I just want a straight insight
That I can use to quantify
What more can I do,
All I want are sigma and mu.

mu baby

All the models shining
So brightly everywhere
And the sound of CPU's fills the air.
And everyone optimizing
Some model they're surmising
Oh won't you just bring me the numbers I need
Won't you please just give them to me?

I don't need a lot from data
This is all I'm asking for
I just want to see my numbers
Put in my final report.
I just want a straight insight
That I can use to quantify
What more can I do,
All I want are sigma and mu.

Clean Data

(White Christmas)

With thanks to Jacob Bien

I'm dreaming of some clean numbers
Just like the ones I saw in class
There was no missing data
And y was just X beta
Plus error: nothing more to ask!

I'm dreaming of balanced designs
All factors randomly assigned!
When collinearity
And exogeneity
Could not be farther from my mind.

I'm dreaming of a few features
No selection to be done
And a nice effect size
Where good power lies
But small enough to still be fun

I'm dreaming of some clean data
With every paper that I write!
May your console be merry and bright
And may all of your results come out right!

Tidy Little R-Space

(Merry Little Christmas)

Thanks to Indrayudh Ghosal for teaching me just enough about the tidyverse, just in time.

Make yourself a tidy little R space
Tibble data frames
Use dplyr to make sure you've consistent names.

Pipe your data through so many functions
Hiding arguments
Ggplot results so that it all makes sense

Shiny apps made using forcats
Put together with Stringr
Readr in to Tidyr out
And you'll only want to purr!

Someday we will all publish with markdown
If journals allow
With built-in code our pdf's will be endowed!
So make yourself a tidy little R space now!

Vanilla R

(We Three Kings)

We three kings of vanilla R
Have packages but used none so far
Neither CRAN nor github can
Review them or give them a star.

CHORUS:

R so shiny R so bright
But will I know my answer's right
To check the coding that I'm loading
Keeps me working through the night!

Not that base-R is without quirks
Another surprise always out there lurks
Dropped dimensions raise my tensions
Bugging up half my works.

CHORUS

All your tests sequential must be
From some now-lost philosophy
A type-II square we cannot bear
It must be done manually

CHORUS

Sometimes I want functionality
To fit models like mgcv
Or fda or refund, hey
Even something like nlme

CHORUS

But RandomForrest, gbf or ranger?
Will xgboost lead me to danger?
caret, rpart, gbm, neuralnet
e1071 or MLR?

CHORUS

What if code is requested of me
to verify reproducibility?
Submit to CRAN
And you, too, can
Be yelled at by Brian Ripley!

CHORUS

Thousands of packages on CRAN alone
Documentation that makes my head groan
What's the one
To get the job done?
I think I'll just roll my own!

CHORUS

Likelihoods!

Giles Hooker
December 2020

(Jingle Bells)

CHORUS:

Likelihoods, likelihoods, optimized today!
Let your model tell you what
your data has to say, hey!
Likelihoods, likelihoods, averaged o'er your prior
a posterior inference for
anything you desire!

Write your model down
Probabilistic-ly
Now an estimate,
Is very easy!
Just find the top of l,
Or if you like Bayes,
MCMC, naturally,
In oh so many ways! Hey!

CHORUS

Tests from ratios
Intervals from profiles
You can be credible
If you wait a while!
Or if you need it fast
Just differentiate
An approximate Laplace
Will never keep you late!

CHORUS