Building a Bayesian Burrito:

John McDonnell, MD

Description

San Diego has long been widely celebrated for its prime location, breezy ocean views, and fine dining experiences. One thing that only the locals fully appreciate is how great its Mexican food offerings are– no surprise given that the Mexican border is only 15 miles away.

As a burrito aficionado (I gained a remarkable 15 pounds during my first year in medical school due to a diet of [almost] exclusively Chipotle burritos) and former denizen of San Diego, when I saw that a like-minded searcher from the San Diego area had collected years of data on the subject of local burrito options in San Diego I grew excited– and a little bit hungry…

Data Source

This project uses data from Scott Cole's Github site. A few years ago, Scott and some San Diego friends started rating local burritos on a 10 point scale. In his words:

1. "Volume - "size matters," "bigger is better," or whatever your favorite innuendo is fits because there's nothing more disconcerting than ordering a burrito and not being full.
2. Tortilla quality
3. Temperature - the Goldilocks zone
4. Meat quality
5. Non-meat filling quality
6. Meat : filling - The ratio between meat and non-meat. Perhaps the golden ratio: 1.6180339887…
7. Uniformity - Bites full of sour cream and cheese with no meat are disappointing.
8. Salsa quality - and variety!
9. Flavor synergy - "That magical aspect a great burrito has, making everything come together like it is a gift from the skies" - A wise Dutchman
10. Wrap integrity - you ordered a burrito, not a burrito bowl."

Based on the above 10 point scale, Scott and his burrito-loving friends also rendered a final burrito rating (on a 5 point, noninteger scale).

Study Population

Our study population is the burritos of the greater San Diego area. Mr. Cole and colleagues rated 423 burritos from May 2011 to August 2019. In addition to the 10 point ratings we saw before, he also collected data on burrito ingredients.

Variables

A description of the data and the variables of interest is seen below.

```
. describe

Contains data from burrito_clean.dta
  obs:           423
 vars:            60                          23 Jul 2020 09:27
-----------------------------------------------------------------------------
-----------------------------------------------------------------------------
-----------
              storage   display    value
variable name   type    format     label      variable label
-----------------------------------------------------------------------------
-----------------------------------------------------------------------------
-----------
burrito         str30   %30s                  burrito name
date            str10   %10s                  date of rating
neighborhood    str18   %18s                  neighborhood in San Diego
yelp            float   %9.0g                 yelp rating
google          float   %9.0g                 google rating
cost            float   %9.0g                 cost of burrito ($)
hunger          float   %9.0g                 hunger level of reviewer
massg           int     %8.0g                 mass of burrito (g)
densitygml      float   %9.0g                 density of burrito (g/mL)
length          float   %9.0g                 length of burrito
circum          float   %9.0g                 circumference of burrito
volume          float   %9.0g                 volume of burrito
tortilla        float   %9.0g                 rating of quality of tortilla
temp            float   %9.0g                 rating of temperature of
burrito
meat            float   %9.0g                 rating of meat in burrito
fillings        float   %9.0g                 rating of non-meat fillings in
the burrito
meatfilling     float   %9.0g                 ratio between meat and non-meat
uniformity      float   %9.0g                 rating of uniformity of burrito
salsa           float   %9.0g                 rating of salsa used in burrito
synergy         float   %9.0g                 rating of the burrito's synergy
wrap            float   %9.0g                 integrity of burrito's wrap job
overall         float   %9.0g                 overall rating of the burrito
reviewer        str12   %12s                  reviewer name
date_num        float   %d                    date of rating
l_location      str51   %51s                  restaurant name
chips_num       float   %9.0g                 indicator for chips
recommend       float   %9.0g                 would reviewer recommend
burrito
beef_ind        float   %9.0g                 indicator for beef
pico_ind        float   %9.0g                 indicator for pico
guac_ind        float   %9.0g                 indicator for guac
cheese_ind      float   %9.0g                 indicator for cheese
fries_ind       float   %9.0g                 indicator for fries
sourcream_ind   float   %9.0g                 indicator for sourcream
pork_ind        float   %9.0g                 indicator for pork
chicken_ind     float   %9.0g                 indicator for chicken
shrimp_ind      float   %9.0g                 indicator for shrimp
fish_ind        float   %9.0g                 indicator for fish
rice_ind        float   %9.0g                 indicator for rice
```

```
beans_ind       float   %9.0g                indicator for beans
lettuce_ind     float   %9.0g                indicator for lettuce
tomato_ind      float   %9.0g                indicator for tomato
bellpeper_ind   float   %9.0g                indicator for bellpeper
carrots_ind     float   %9.0g                indicator for carrots
cabbage_ind     float   %9.0g                indicator for cabbage
sauce_ind       float   %9.0g                indicator for sauce
cilantro_ind    float   %9.0g                indicator for cilantro
onion_ind       float   %9.0g                indicator for onion
taquito_ind     float   %9.0g                indicator for taquito
pineapple_ind   float   %9.0g                indicator for pineapple
ham_ind         float   %9.0g                indicator for ham
chilerelleno_~d float   %9.0g                indicator for chilerelleno
nopales_ind     float   %9.0g                indicator for nopales
lobster_ind     float   %9.0g                indicator for lobster
egg_ind         float   %9.0g                indicator for egg
mushroom_ind    float   %9.0g                indicator for mushroom
bacon_ind       float   %9.0g                indicator for bacon
sushi_ind       float   %9.0g                indicator for sushi
avocado_ind     float   %9.0g                indicator for avocado
corn_ind        float   %9.0g                indicator for corn
zucchini_ind    float   %9.0g                indicator for zucchini
-----------------------------------------------------------------------------
-----------------------------------------------------------------------------
-----------
Sorted by:
     Note: Dataset has changed since last saved.
```

I've attached labels to each variable to specify its meaning in a clear way.

We will go into specifics in a minute, but for now let's examine the string/character variables. There are five of them in total: *burrito*, *date*, *neighborhood*, *reviewer*, and *l_location.* These variables are not going to be very useful to us in the regression analysis, but we can get a "flavor" of them (pun most certainly intended) by looking at a small subset.

```
. list burrito date neighborhood reviewer l_location in 1/5

     +-----------------------------------------------------------------------
-+
     |    burrito       date   neighbor~d   reviewer             l_location
|
     |-----------------------------------------------------------------------
-|
  1. |  california   1/18/2016      miramar      Scott     donato's taco shop
|
  2. |  california   1/24/2016   san marcos      Scott   oscar's mexican food
|
  3. |    carnitas   1/24/2016                   Emily   oscar's mexican food
|
  4. | carne asada   1/24/2016                 Ricardo   oscar's mexican food
|
  5. |  california   1/27/2016     carlsbad      Scott           pollos maria
|
```

```
    +------------------------------------------------------------------------
-+
```

For our analysis, we are definitely going to need to look at the continuous variables. Many of these correspond to the reviewers' 1-5 rating scale for individual elements of burrito quality, while others pertain to burrito measurements. As we will see later, there are a lot of missing data with the objective burrito measurements but most of the subjective ratings are complete. Additionally, we see here variables corresponding to the burritos' Yelp and Google reviews; however, many of these are missing.

```
. summarize yelp-overall

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
        yelp |         87    3.887356    .4753957        2.5        4.5
      google |         87    4.167816    .3736975        2.9          5
        cost |        416    7.065216    1.503645       2.99         25
      hunger |        420    3.496095    .8114664         .5          5
       massg |         22    546.1818    144.4456        350        925
-------------+---------------------------------------------------------
   densitygml |         22    .6752774    .0804682        .56   .8656716
      length |        284     20.0469    2.084957         15         26
      circum |        282    22.13174    1.777526         17         29
      volume |        282    .7864894    .1522597         .4       1.54
    tortilla |        423    3.519385    .7933014          1          5
-------------+---------------------------------------------------------
        temp |        403    3.780397    .9800436          1          5
        meat |        409    3.622249    .8283836          1          5
    fillings |        420    3.542024    .8012532          1          5
 meatfilling |        414    3.589082    .9962922         .5          5
  uniformity |        421    3.434086    1.069349          0          5
-------------+---------------------------------------------------------
       salsa |        398    3.372613    .9224345          0          5
     synergy |        421    3.587767     .886277          1          5
        wrap |        420     3.98119    1.115803          0          5
     overall |        421    3.620887     .755718          1          5
```

Finally, we also have a number of indicator variables, generally corresponding to whether a given ingredient was included in the burrito. As we might expect, when we look at the summarized proportions (I included them this way to spare the reader a long list of tabulated output), we can see that certain ingredients were common (guacamole, beef) and others were uncommon (lobster).

```
. codebook chips_num-zucchini_ind, compact

Variable      Obs Unique      Mean  Min  Max  Label
-----------------------------------------------------------------------------
-----------------------------------------------------------------------------
-----------
chips_num     423      2  .0520095    0    1  indicator for chips
recommend     233      2  .6995708    0    1  would reviewer recommend
burrito
```

```
beef_ind        423     2  .4255319    0    1  indicator for beef
pico_ind        423     2  .3758865    0    1  indicator for pico
guac_ind        423     2  .3664303    0    1  indicator for guac
cheese_ind      423     2  .3782506    0    1  indicator for cheese
fries_ind       423     2  .3026005    0    1  indicator for fries
sourcream_~d    423     2  .2174941    0    1  indicator for sourcream
pork_ind        423     2  .1205674    0    1  indicator for pork
chicken_ind     423     2  .0496454    0    1  indicator for chicken
shrimp_ind      423     2  .0496454    0    1  indicator for shrimp
fish_ind        423     2  .0141844    0    1  indicator for fish
rice_ind        423     2  .0851064    0    1  indicator for rice
beans_ind       423     2  .0827423    0    1  indicator for beans
lettuce_ind     423     2  .0260047    0    1  indicator for lettuce
tomato_ind      423     2  .0165485    0    1  indicator for tomato
bellpeper_~d    423     2  .0165485    0    1  indicator for bellpeper
carrots_ind     423     2  .0023641    0    1  indicator for carrots
cabbage_ind     423     2  .0189125    0    1  indicator for cabbage
sauce_ind       423     2  .0898345    0    1  indicator for sauce
cilantro_ind    423     2   .035461    0    1  indicator for cilantro
onion_ind       423     2  .0401891    0    1  indicator for onion
taquito_ind     423     2  .0094563    0    1  indicator for taquito
pineapple_~d    423     2  .0165485    0    1  indicator for pineapple
ham_ind         423     2  .0047281    0    1  indicator for ham
chilerelle~d    423     2  .0094563    0    1  indicator for chilerelleno
nopales_ind     423     2  .0094563    0    1  indicator for nopales
lobster_ind     423     2  .0023641    0    1  indicator for lobster
egg_ind         423     2  .0118203    0    1  indicator for egg
mushroom_ind    423     2  .0070922    0    1  indicator for mushroom
bacon_ind       423     2  .0070922    0    1  indicator for bacon
sushi_ind       423     2  .0047281    0    1  indicator for sushi
avocado_ind     423     2  .0307329    0    1  indicator for avocado
corn_ind        423     2  .0070922    0    1  indicator for corn
zucchini_ind    423     2  .0023641    0    1  indicator for zucchini
-----------------------------------------------------------------------------
-----------------------------------------------------------------------------
-----------
```

Response Variable

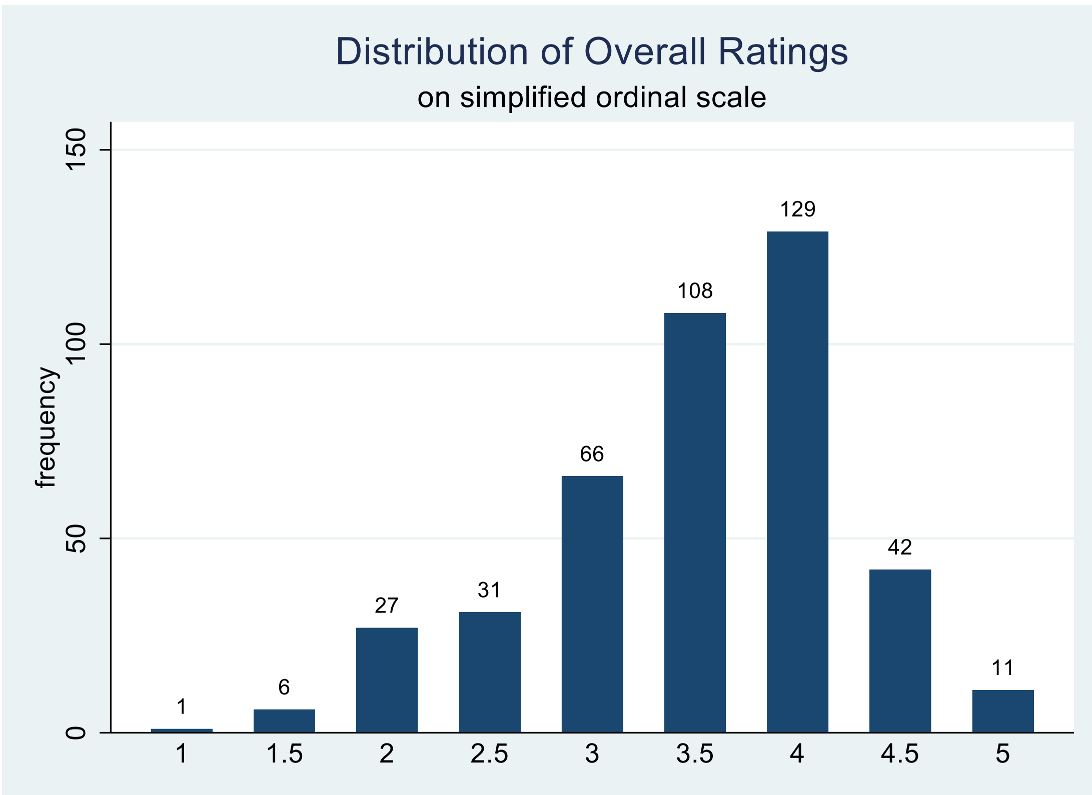
Our response variable in this analysis is *top_burrito*. This variable is a binary corresponding to 1 if a burrito was ranked in the top 2 categories for overall rating (*overall_ord*).

The distribution of the "overall" burrito rating is given below, first in table form (note that "." indicates missing in Stata):

```
. tab overall_ord, missing

overall_ord |      Freq.     Percent        Cum.
------------+-----------------------------------
          1 |          1        0.24        0.24
        1.5 |          6        1.42        1.65
          2 |         27        6.38        8.04
        2.5 |         31        7.33       15.37
          3 |         66       15.60       30.97
```

```
        3.5 |        108       25.53       56.50
          4 |        129       30.50       87.00
        4.5 |         42        9.93       96.93
          5 |         11        2.60       99.53
          . |          2        0.47      100.00
------------+-----------------------------------
      Total |        423      100.00
```

and then in graphical form (focusing on the bars corresponding to ratings of 4, 4.5, and 5):



To be honest, I would have vastly preferred to treat overall rating as an ordinal outcome variable, but ran into some catastrophic issues with convergence when I fit the Bayesian ordinal regression model. These problems disappeared when I treated the outcome as a binary and ran Bayesian logistic regression instead (of course, at the expense of some loss of information in the outcome). In any case, our outcome variable, *top_burrito*, will simply be binary for all the variables that were rated as 4 or higher.

```
. gen top_burrito = 0

. replace top_burrito = 1 if overall_ord >=4 & !missing(overall_ord)
(182 real changes made)

.
. tab overall_ord top_burrito

overall_or |      top_burrito
```

```
          d |          0          1 |     Total
------------+----------------------+----------
          1 |          1          0 |         1
        1.5 |          6          0 |         6
          2 |         27          0 |        27
        2.5 |         31          0 |        31
          3 |         66          0 |        66
        3.5 |        108          0 |       108
          4 |          0        129 |       129
        4.5 |          0         42 |        42
          5 |          0         11 |        11
------------+----------------------+----------
      Total |        239        182 |       421
```

Predictors

While a number of predictor variables will be considered here, we are ultimately going to have to choose just a few of them in the final model. We will spend a lot more time discussing our variable selection process a little bit later, but here is a "sneak peak" of the predictor variables we will end up using.

```
. list synergy meat meatfilling fillings in 1/5

     +---------------------------------------+
     | synergy   meat   meatfi~g   fillings |
     |---------------------------------------|
  1. |       4      3          4        3.5 |
  2. |     2.5    2.5          2        2.5 |
  3. |       3    2.5        4.5          3 |
  4. |       4    3.5          4          3 |
  5. |     4.5      4        4.5        3.5 |
     +---------------------------------------+
```

Objective of Study

What makes a burrito good? I will use Bayesian logistic regression methods to answer this question. In this dataset, with N = 421 observations and a pool of over 60 candidate predictor variables, I will follow the overall strategy:

1. determine reasonable predictor variables using stepwise logistic regression and lasso methods
2. fit a Bayesian logistic regression model
3. examine fit characteristics of the Bayesian model and diagnose any problems with autocorrelation and convergence
4. make some predictions
5. compare to a standard/frequentist logistic regression model and summarize results

Background: What is known on this subject?

Perhaps unsurprisingly, there is a dearth of good information on what makes a good burrito. The most interesting previous research on the subject comes from Five Thirty Eight. Nate Silver and his co-investigator Anna Barry-Jester looked for the best burrito in America with their project In Search of America's Best Burrito. Their process involved data mining Yelp to create an overall rating called "Value Over Replacement Burrito." forming a "Burrito Selection Committee" to determine the most promising 64 burrito joints (from 60k+ ratings) in the United States, and then taste-testing each of these. The ultimate winner was La Taqueria's (San Francisco, CA) carnitas burrito whose "bombardment of liquid and flavor … are enough to stop any woman in her tracks, even one who'd been eating burritos daily for two months straight." Although I suspect I may have an appetite sufficient to replicate this task, I don't think it will fit into my school, work, and home life at this point in time.

In his own analysis of the burrito data, Scott Cole performed a principal components analysis predicting rating as an outcome variable, and found fillings, meat, meat : filling, and uniformity to be the most important predictors. The details of his analysis are summarized here.

Contribution to the literature

To my knowledge, this will be the first study evaluating predictors of burrito quality using Bayesian methods.

Exploratory Analysis

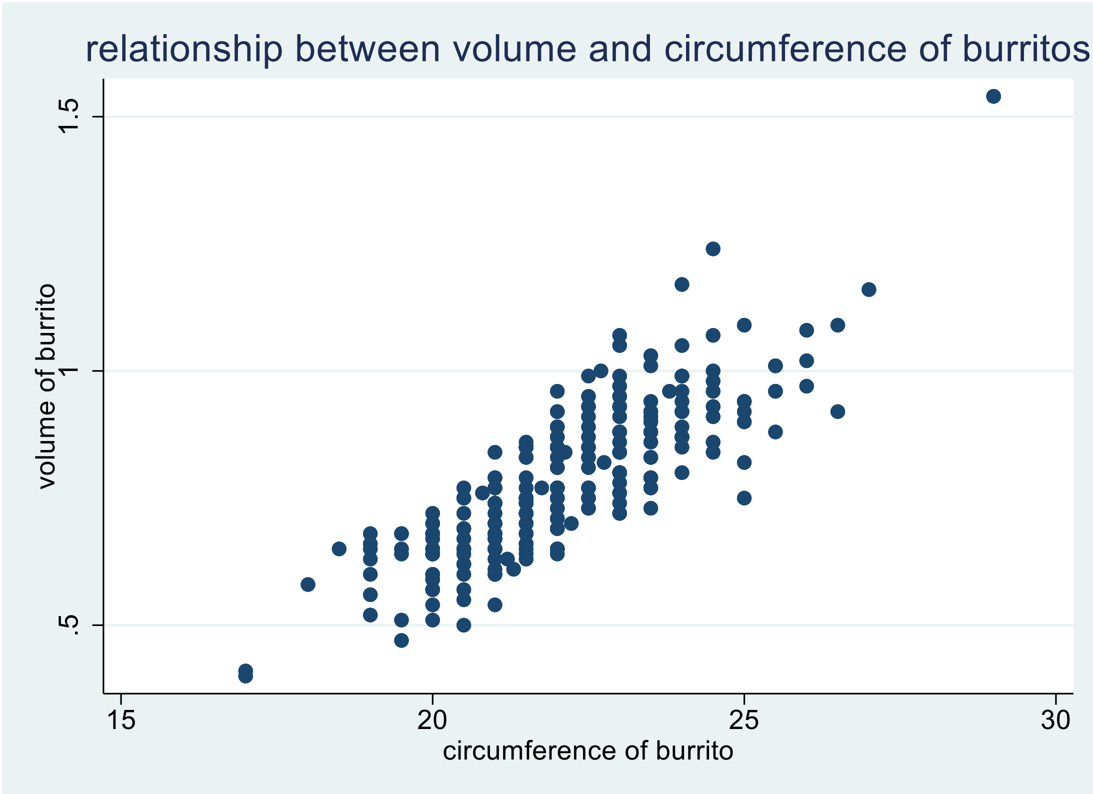
Having spoken a bit about the background and objectives to this analysis, let's look more into the dataset itself, and particularly how our outcome of interest relates to the potential predictor variables.

A first step is to examine some scatterplots of the relationship between continuous variables. This approach can give us an overall sense of the distribution and correlations of our variables, and perhaps clue us into relationships that may be important for our model.

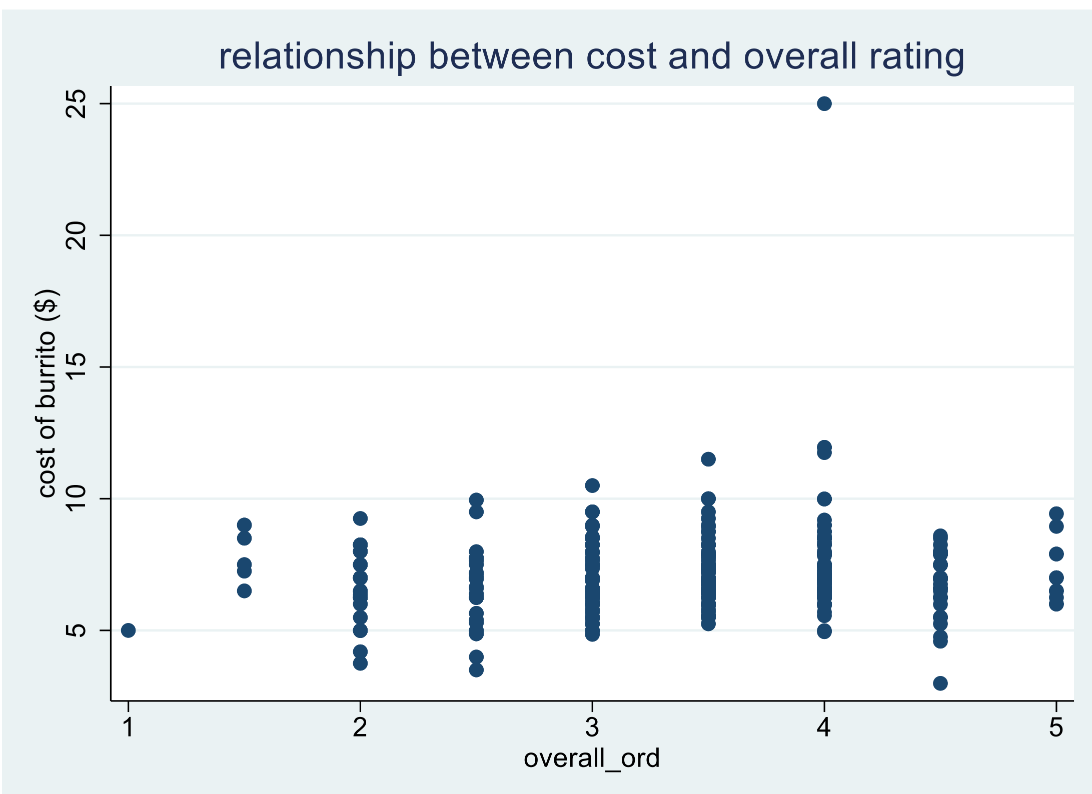
For the purposes of my exploratory data analysis, I ran the following command (output omitted to spare the reader's sanity):

```
graph matrix yelp google cost hunger massg densitygml length circum volume
tortilla temp meat fillings meatfilling uniformity salsa synergy wrap
overall_ord
```

After running this command, the first thing I noted was that the strongest linear relationships in these data were just where I would expect them– in the measures of the physical dimensions of the burritos. For example:

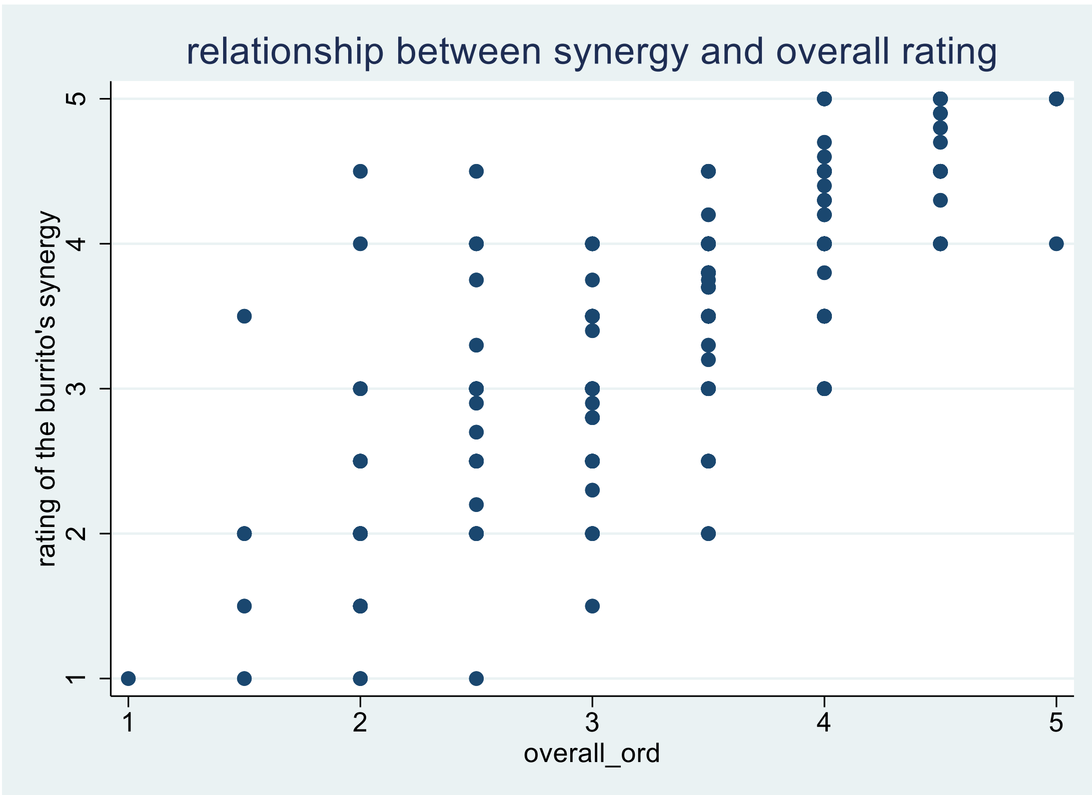


The relationship between the overall rating and the cost seems less strong. See:

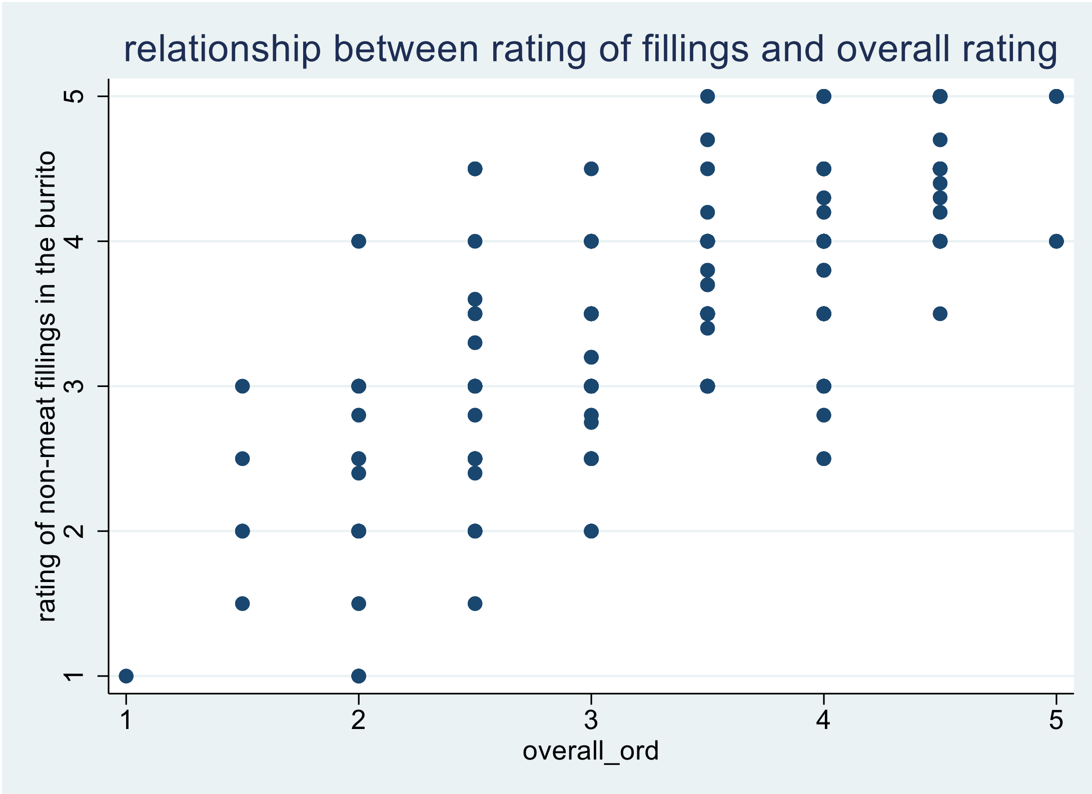


(Incidentally, I tracked down the outlier in this plot. It turns out that this was not a miscode - scroll to the "specials" section for more on this.)

There turned out to be a monotonic relationship between the burrito's synergy and its overall rating. This makes intuitive sense - a reviewer who thought a burrito had great synergy would be likely to rate it higher.



Another noteworthy relationship is seen here, this time between the fillings and the overall rating of the burrito.



Unfortunately, we have a problem that I alluded to before– namely, many of the continuous variables we might be interested in have a lot of missing data. One reasonable strategy would be to impute, but for the purposes of focusing on the main goals of this project, I will just leave these variables out instead.

```
. inspect massg-volume

massg:  mass of burrito (g)                   Number of Observations
----------------------------             ----------------------------------------
-
                                                 Total      Integers
Nonintegers
|  #                        Negative                -             -
-
|  #   #                    Zero                    -             -
-
|  #   #                    Positive               22            22
-
|  #   #   #                             -----------   -----------   ----------
-
|  #   #   #                Total                  22            22
-
|  #   #   #   .   #        Missing               401
+-----------------------                 -----------
350                    925                        423
  (18 unique values)

densitygml:  density of burrito (g/mL)           Number of Observations
```

```
--------------------------------------- ---------------------------------------
-
                                                     Total      Integers
Nonintegers
|  #       #                 Negative             -             -
-
|  #   #   #                 Zero                 -             -
-
|  #   #   #   #             Positive            22             -
22
|  #   #   #   #                         -----------   -----------   ----------
-
|  #   #   #   #             Total               22             -
22
|  #   #   #   #   #         Missing            401
+----------------------                  -----------
.56            .8656716                         423
  (21 unique values)

length:  length of burrito                  Number of Observations
--------------------------               ---------------------------------------
-
                                                     Total      Integers
Nonintegers
|          #                 Negative             -             -
-
|          #                 Zero                 -             -
-
|      #   #                 Positive           284           162
122
|      #   #   #                         -----------   -----------   ----------
-
|      #   #   #             Total              284           162
122
|  #   #   #   #   .         Missing            139
+----------------------                  -----------
15                   26                         423
  (29 unique values)

circum:  circumference of burrito           Number of Observations
---------------------------------        ---------------------------------------
-
                                                     Total      Integers
Nonintegers
|          #                 Negative             -             -
-
|      #   #                 Zero                 -             -
-
|      #   #                 Positive           282           162
120
|      #   #                             -----------   -----------   ----------
-
|      #   #                 Total              282           162
120
|  .   #   #   #   .         Missing            141
+----------------------                  -----------
17                   29                         423
```

```
  (30 unique values)

volume:  volume of burrito                        Number of Observations
--------------------------             ---------------------------------------
-
                                                   Total      Integers
Nonintegers
|      #                     Negative                  -             -
-
|      #                     Zero                      -             -
-
|      #                     Positive                282             2
280
|      #                                 -----------   -----------   ----------
-
|      #   #                 Total                   282             2
280
|  #   #   #   .   .         Missing                 141
+----------------------                  -----------
.4                  1.54                             423
  (64 unique values)

. drop massg-volume
```

The variables corresponding to Yelp and Google reviews also have a lot of missing values. Additionally, these might not be good predictors for another reason– they would be expected to be highly correlated with the outcome rating. Therefore, we will leave them out.

```
. inspect yelp-google

yelp:  yelp rating                                Number of Observations
------------------                     ---------------------------------------
-
                                                   Total      Integers
Nonintegers
|            #               Negative                  -             -
-
|            #               Zero                      -             -
-
|            #               Positive                 87            48
39
|            #                           -----------   -----------   ----------
-
|        #   #   #           Total                    87            48
39
|  .   #   #   #   #         Missing                 336
+----------------------                  -----------
2.5                  4.5                             423
   (6 unique values)

google:  google rating                            Number of Observations
----------------------                 ---------------------------------------
-
                                                   Total      Integers
Nonintegers
```

```
|             #            Negative          -            -
-
|         #   #            Zero              -            -
-
|         #   #            Positive         87           10
77
|         #   #                    -----------  -----------  ----------
-
|         #   #            Total            87           10
77
|  .   .   #   #   #       Missing         336
+----------------------            -----------
2.9                   5                    423
  (18 unique values)

. drop yelp-google
```

Let's explore the categorical variables now. Perhaps the best way to get a sense of the relationships at play here is by using tables. For example, we can see that many of the "top burritos" contained beef, but not the majority.

```
. tab top_burrito beef, row

+-----------------+
| Key             |
|-----------------|
|   frequency     |
| row percentage  |
+-----------------+

top_burrit |  indicator for beef
         o |         0          1 |     Total
-----------+----------------------+----------
         0 |       128        113 |       241
           |     53.11      46.89 |    100.00
-----------+----------------------+----------
         1 |       115         67 |       182
           |     63.19      36.81 |    100.00
-----------+----------------------+----------
     Total |       243        180 |       423
           |     57.45      42.55 |    100.00
```

In fact, a large number of these fillings were rare, with counts <= 5. For this reason, we won't seriously consider them as predictors. I'll also drop any indicator variables where the "d" cell (in standard epidemiological format, corresponding in this case to counts where top_burrito == 1 and topping == 1) is <=5. The other indicators, however, are fair game for possible inclusion.

```
.
. drop zucchini corn sushi bacon mushroom egg lobster nopales chile ham
taquito carrot /* low counts overall */

. drop avocado pineapple onion cilantro cabbage bellpeper tomato lettuce
beans fish shrimp chicken /* low counts in "d" */
```

```
.
. foreach v of varlist beef-sauce{
  2.     tab top_burrito `v', row
  3. }

+----------------+
| Key            |
|----------------|
|   frequency    |
| row percentage |
+----------------+

top_burrit |  indicator for beef
         o |         0          1 |     Total
-----------+----------------------+----------
         0 |       128        113 |       241
           |     53.11      46.89 |    100.00
-----------+----------------------+----------
         1 |       115         67 |       182
           |     63.19      36.81 |    100.00
-----------+----------------------+----------
     Total |       243        180 |       423
           |     57.45      42.55 |    100.00

+----------------+
| Key            |
|----------------|
|   frequency    |
| row percentage |
+----------------+

top_burrit |  indicator for pico
         o |         0          1 |     Total
-----------+----------------------+----------
         0 |       137        104 |       241
           |     56.85      43.15 |    100.00
-----------+----------------------+----------
         1 |       127         55 |       182
           |     69.78      30.22 |    100.00
-----------+----------------------+----------
     Total |       264        159 |       423
           |     62.41      37.59 |    100.00

+----------------+
| Key            |
|----------------|
|   frequency    |
| row percentage |
+----------------+

top_burrit |  indicator for guac
         o |         0          1 |     Total
-----------+----------------------+----------
         0 |       149         92 |       241
           |     61.83      38.17 |    100.00
-----------+----------------------+----------
```

```
           1 |       119         63 |       182 
             |     65.38      34.62 |    100.00 
-------------+----------------------+----------
       Total |       268        155 |       423 
             |     63.36      36.64 |    100.00 

+-----------------+
| Key             |
|-----------------|
|   frequency     |
| row percentage  |
+-----------------+

top_burrit | indicator for cheese
         o |         0          1 |     Total
-----------+----------------------+----------
         0 |       145         96 |       241 
           |     60.17      39.83 |    100.00 
-----------+----------------------+----------
         1 |       118         64 |       182 
           |     64.84      35.16 |    100.00 
-----------+----------------------+----------
     Total |       263        160 |       423 
           |     62.17      37.83 |    100.00 

+-----------------+
| Key             |
|-----------------|
|   frequency     |
| row percentage  |
+-----------------+

top_burrit |  indicator for fries
         o |         0          1 |     Total
-----------+----------------------+----------
         0 |       166         75 |       241 
           |     68.88      31.12 |    100.00 
-----------+----------------------+----------
         1 |       129         53 |       182 
           |     70.88      29.12 |    100.00 
-----------+----------------------+----------
     Total |       295        128 |       423 
           |     69.74      30.26 |    100.00 

+-----------------+
| Key             |
|-----------------|
|   frequency     |
| row percentage  |
+-----------------+

           |     indicator for
top_burrit |       sourcream
         o |         0          1 |     Total
-----------+----------------------+----------
         0 |       188         53 |       241 
           |     78.01      21.99 |    100.00 
```

```
-----------+----------------------+----------
         1 |       143         39 |       182 
           |     78.57      21.43 |    100.00 
-----------+----------------------+----------
     Total |       331         92 |       423 
           |     78.25      21.75 |    100.00 

+-----------------+
| Key             |
|-----------------|
|    frequency    |
| row percentage  |
+-----------------+

top_burrit |  indicator for pork
         o |         0          1 |     Total
-----------+----------------------+----------
         0 |       208         33 |       241 
           |     86.31      13.69 |    100.00 
-----------+----------------------+----------
         1 |       164         18 |       182 
           |     90.11       9.89 |    100.00 
-----------+----------------------+----------
     Total |       372         51 |       423 
           |     87.94      12.06 |    100.00 

+-----------------+
| Key             |
|-----------------|
|    frequency    |
| row percentage  |
+-----------------+

top_burrit |  indicator for rice
         o |         0          1 |     Total
-----------+----------------------+----------
         0 |       213         28 |       241 
           |     88.38      11.62 |    100.00 
-----------+----------------------+----------
         1 |       174          8 |       182 
           |     95.60       4.40 |    100.00 
-----------+----------------------+----------
     Total |       387         36 |       423 
           |     91.49       8.51 |    100.00 

+-----------------+
| Key             |
|-----------------|
|    frequency    |
| row percentage  |
+-----------------+

top_burrit |  indicator for sauce
         o |         0          1 |     Total
-----------+----------------------+----------
         0 |       216         25 |       241 
           |     89.63      10.37 |    100.00 
```

```
-----------+---------------------+----------
         1 |       169         13 |       182
           |     92.86       7.14 |    100.00
-----------+---------------------+----------
     Total |       385         38 |       423
           |     91.02       8.98 |    100.00
```

For our remaining indicator variables, there are none with an obviously strong relationship to the outcome.

Model Selection

After excluding the variables mentioned above, we are left with the following:

```
. describe burrito-sauce

              storage   display    value
variable name   type    format     label      variable label
-----------------------------------------------------------------------------
-----------------------------------------------------------------------------
-----------
burrito         str30   %30s                  burrito name
date            str10   %10s                  date of rating
neighborhood    str18   %18s                  neighborhood in San Diego
cost            float   %9.0g                 cost of burrito ($)
hunger          float   %9.0g                 hunger level of reviewer
tortilla        float   %9.0g                 rating of quality of tortilla
temp            float   %9.0g                 rating of temperature of
burrito
meat            float   %9.0g                 rating of meat in burrito
fillings        float   %9.0g                 rating of non-meat fillings in
the burrito
meatfilling     float   %9.0g                 ratio between meat and non-meat
uniformity      float   %9.0g                 rating of uniformity of burrito
salsa           float   %9.0g                 rating of salsa used in burrito
synergy         float   %9.0g                 rating of the burrito's synergy
wrap            float   %9.0g                 integrity of burrito's wrap job
overall         float   %9.0g                 overall rating of the burrito
reviewer        str12   %12s                  reviewer name
date_num        float   %d                    date of rating
l_location      str51   %51s                  restaurant name
chips_num       float   %9.0g                 indicator for chips
recommend       float   %9.0g                 would reviewer recommend
burrito
beef_ind        float   %9.0g                 indicator for beef
pico_ind        float   %9.0g                 indicator for pico
guac_ind        float   %9.0g                 indicator for guac
cheese_ind      float   %9.0g                 indicator for cheese
fries_ind       float   %9.0g                 indicator for fries
sourcream_ind   float   %9.0g                 indicator for sourcream
pork_ind        float   %9.0g                 indicator for pork
rice_ind        float   %9.0g                 indicator for rice
sauce_ind       float   %9.0g                 indicator for sauce
```

I will employ two strategies for model selection. The first is stepwise forward regression. Note that Stata accomplishes this using changes in p values, not an information criterion like AIC or BIC.

```
. set seed 453

.
. stepwise, pe(.2): logit top_burrito c.cost c.hunger c.(tortilla-wrap)
i.(chips beef-sauce)
note: 0b.chips_num dropped because of estimability
note: 0b.beef_ind dropped because of estimability
note: 0b.pico_ind dropped because of estimability
note: 0b.guac_ind dropped because of estimability
note: 0b.cheese_ind dropped because of estimability
note: 0b.fries_ind dropped because of estimability
note: 0b.sourcream_ind dropped because of estimability
note: 0b.pork_ind dropped because of estimability
note: 0b.rice_ind dropped because of estimability
note: 0b.sauce_ind dropped because of estimability
                      begin with empty model
p = 0.0000 <  0.2000  adding  synergy
p = 0.0000 <  0.2000  adding  meat
p = 0.0000 <  0.2000  adding  meatfilling
p = 0.0002 <  0.2000  adding  temp
p = 0.0007 <  0.2000  adding  fillings
p = 0.0035 <  0.2000  adding  tortilla
p = 0.0532 <  0.2000  adding  salsa
p = 0.1975 <  0.2000  adding  1.rice_ind
p = 0.1991 <  0.2000  adding  cost

Logistic regression                             Number of obs     =
358
                                                LR chi2(9)        =
278.23
                                                Prob > chi2       =
0.0000
Log likelihood =  -103.6348                     Pseudo R2         =
0.5731

-----------------------------------------------------------------------------
-
 top_burrito |      Coef.   Std. Err.      z    P>|z|     [95% Conf.
Interval]
-------------+---------------------------------------------------------------
-
     synergy |   1.716266   .4052795     4.23   0.000     .9219331
2.5106
        meat |   1.699645   .3629123     4.68   0.000     .9883498
2.41094
 meatfilling |   1.069974   .2472165     4.33   0.000     .5854387
1.55451
        temp |   .7140077   .1992907     3.58   0.000     .3234051
1.10461
    fillings |   1.039525   .3543918     2.93   0.003     .3449296
1.73412
```

```
    tortilla |   .8013605    .279028     2.87   0.004     .2544756
1.348245
       salsa |   .3800812   .2372236     1.60   0.109    -.0848686
.8450309
  1.rice_ind |  -.9454823   .6962917    -1.36   0.175    -2.310189
.4192243
        cost |   .2000829   .1558271     1.28   0.199    -.1053326
.5054985
       _cons |  -29.16476   3.500084    -8.33   0.000     -36.0248   -
22.30473
----------------------------------------------------------------------------
-

.
```

Our stepwise selection technique suggests we include *synergy*, *meat*, *meatfilling*, *temp*, *fillings*, *tortilla*, *salsa*, *cost* and *rice_ind* in the model.

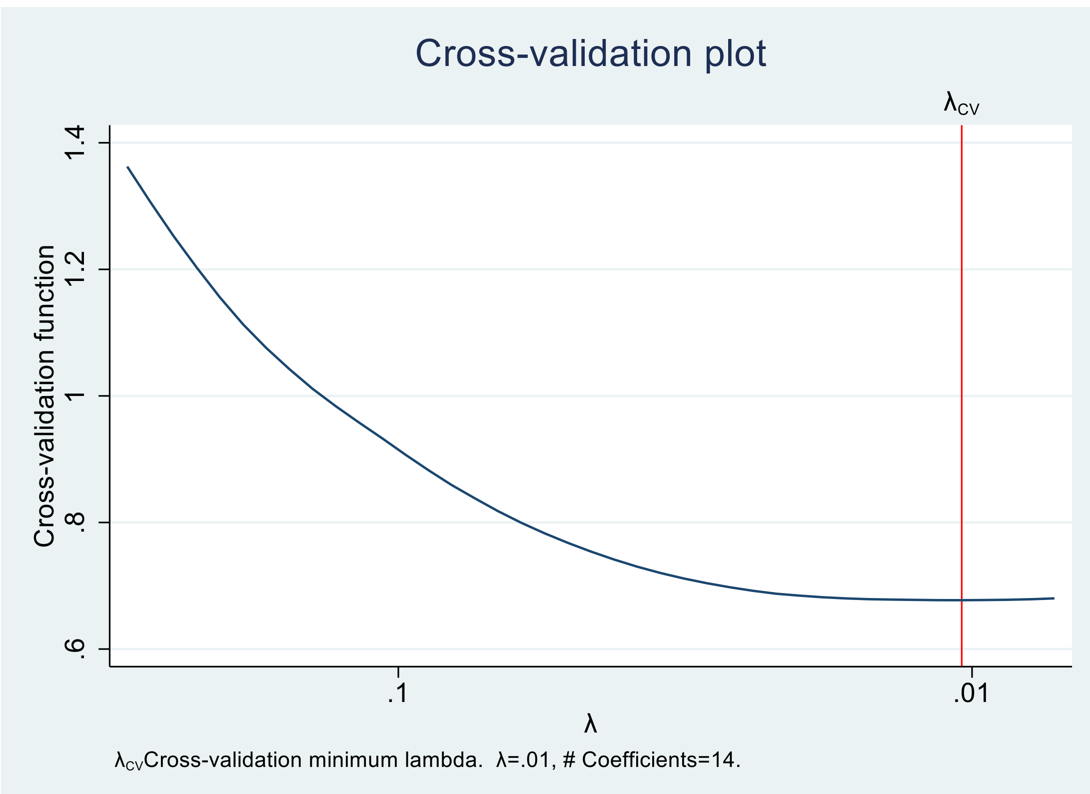
The other strategy for variable selection I want to try is Lasso penalized regression.

```
. set seed 453

. lasso logit top_burrito c.cost c.hunger c.(tortilla-wrap) i.(chips beef-
sauce)

10-fold cross-validation with 100 lambdas ...
Grid value 1:     lambda = .2968057   no. of nonzero coef. =       0
Folds: 1...5....10   CVF = 1.362455
Grid value 2:     lambda = .2704383   no. of nonzero coef. =       1
Folds: 1...5....10   CVF = 1.306353
Grid value 3:     lambda = .2464133   no. of nonzero coef. =       1
Folds: 1...5....10   CVF = 1.252595
Grid value 4:     lambda = .2245226   no. of nonzero coef. =       2
Folds: 1...5....10   CVF = 1.202766
Grid value 5:     lambda = .2045767   no. of nonzero coef. =       2
Folds: 1...5....10   CVF = 1.155658
Grid value 6:     lambda = .1864026   no. of nonzero coef. =       2
Folds: 1...5....10   CVF = 1.113017
Grid value 7:     lambda = .1698431   no. of nonzero coef. =       3
Folds: 1...5....10   CVF = 1.07567
Grid value 8:     lambda = .1547548   no. of nonzero coef. =       3
Folds: 1...5....10   CVF = 1.042219
Grid value 9:     lambda = .1410068   no. of nonzero coef. =       3
Folds: 1...5....10   CVF = 1.011164
Grid value 10:    lambda = .1284801   no. of nonzero coef. =       3
Folds: 1...5....10   CVF = .9836577
Grid value 11:    lambda = .1170663   no. of nonzero coef. =       4
Folds: 1...5....10   CVF = .9579422
Grid value 12:    lambda = .1066664   no. of nonzero coef. =       4
Folds: 1...5....10   CVF = .9330114
Grid value 13:    lambda = .0971905   no. of nonzero coef. =       5
Folds: 1...5....10   CVF = .9072486
Grid value 14:    lambda = .0885564   no. of nonzero coef. =       6
Folds: 1...5....10   CVF = .8825074
Grid value 15:    lambda = .0806893   no. of nonzero coef. =       6
Folds: 1...5....10   CVF = .8591103
```

```
Grid value 16:    lambda =  .073521   no. of nonzero coef. =       6
Folds: 1...5....10   CVF = .8381328
Grid value 17:    lambda = .0669896   no. of nonzero coef. =       6
Folds: 1...5....10   CVF = .8178472
Grid value 18:    lambda = .0610385   no. of nonzero coef. =       6
Folds: 1...5....10   CVF = .7995145
Grid value 19:    lambda =  .055616   no. of nonzero coef. =       7
Folds: 1...5....10   CVF = .7829686
Grid value 20:    lambda = .0506752   no. of nonzero coef. =       7
Folds: 1...5....10   CVF = .7679245
Grid value 21:    lambda = .0461734   no. of nonzero coef. =       7
Folds: 1...5....10   CVF = .7541
Grid value 22:    lambda = .0420714   no. of nonzero coef. =       7
Folds: 1...5....10   CVF = .7414924
Grid value 23:    lambda = .0383339   no. of nonzero coef. =       7
Folds: 1...5....10   CVF = .7301799
Grid value 24:    lambda = .0349285   no. of nonzero coef. =       7
Folds: 1...5....10   CVF = .720204
Grid value 25:    lambda = .0318255   no. of nonzero coef. =       7
Folds: 1...5....10   CVF = .7115774
Grid value 26:    lambda = .0289982   no. of nonzero coef. =       8
Folds: 1...5....10   CVF = .7040825
Grid value 27:    lambda = .0264221   no. of nonzero coef. =       8
Folds: 1...5....10   CVF = .697619
Grid value 28:    lambda = .0240748   no. of nonzero coef. =       9
Folds: 1...5....10   CVF = .6920241
Grid value 29:    lambda = .0219361   no. of nonzero coef. =       9
Folds: 1...5....10   CVF = .6874287
Grid value 30:    lambda = .0199873   no. of nonzero coef. =      10
Folds: 1...5....10   CVF = .6844194
Grid value 31:    lambda = .0182117   no. of nonzero coef. =      12
Folds: 1...5....10   CVF = .6818702
Grid value 32:    lambda = .0165938   no. of nonzero coef. =      12
Folds: 1...5....10   CVF = .6799953
Grid value 33:    lambda = .0151197   no. of nonzero coef. =      12
Folds: 1...5....10   CVF = .6787508
Grid value 34:    lambda = .0137765   no. of nonzero coef. =      13
Folds: 1...5....10   CVF = .6781779
Grid value 35:    lambda = .0125526   no. of nonzero coef. =      13
Folds: 1...5....10   CVF = .6776702
Grid value 36:    lambda = .0114375   no. of nonzero coef. =      13
Folds: 1...5....10   CVF = .6772458
Grid value 37:    lambda = .0104214   no. of nonzero coef. =      14
Folds: 1...5....10   CVF = .6771735
Grid value 38:    lambda = .0094956   no. of nonzero coef. =      14
Folds: 1...5....10   CVF = .6774453
Grid value 39:    lambda =  .008652   no. of nonzero coef. =      14
Folds: 1...5....10   CVF = .6779057
Grid value 40:    lambda = .0078834   no. of nonzero coef. =      14
Folds: 1...5....10   CVF = .6787364
Grid value 41:    lambda = .0071831   no. of nonzero coef. =      15
Folds: 1...5....10   CVF = .6801017
... cross-validation complete ... minimum found

Lasso logit model                           No. of obs        =        358
                                            No. of covariates =         31
Selection: Cross-validation                 No. of CV folds   =         10
```

```
--------------------------------------------------------------------------
         |                                No. of      Out-of-      CV mean
         |                               nonzero       sample
      ID |     Description      lambda     coef.   dev. ratio     deviance
---------+----------------------------------------------------------------
       1 |    first lambda    .2968057         0       0.0046     1.362455
      36 |   lambda before    .0114375        13       0.5006     .6772458
    * 37 | selected lambda    .0104214        14       0.5007     .6771735
      38 |    lambda after    .0094956        14       0.5005     .6774453
      41 |     last lambda    .0071831        15       0.4985     .6801017
--------------------------------------------------------------------------
* lambda selected by cross-validation.
```

Our selected lambda is indicated by an asterisk, and can be seen visually in the cross-validation plot. Here, the Lasso has selected a lambda with 14 coefficients, some of which we are going to want to prune down.




```
. lassocoef, display(coef, standardized)

---------------------------
                |    active
----------------+----------
           cost |  .1421825
       tortilla |    .43437
           temp |  .4847077
           meat |  1.072969
```

```
        fillings |    .593496
     meatfilling |   .7481762
      uniformity |   .0191915
           salsa |   .2297641
         synergy |    1.15989
      0.pico_ind |   .1321554
      0.guac_ind | -.0710488
 0.sourcream_ind | -.0071801
      0.rice_ind |   .0933653
     0.sauce_ind |   .0757354
           _cons | -.9800552
---------------------------
Legend:
  b - base level
  e - empty cell
  o - omitted
```

Focusing on the Lasso standardized coefficients that are >=0.5 in absolute value (an admittedly arbitrary decision, but I want to focus on the more important variables), we would include *meat*, *fillings*, *meatfillings*, and *synergy*. The Lasso gives us several of the same variables suggested by forward suggestion, but (using our >=0.5 cutoff) is overall more parsimonious. For this reason, we'll stick with the Lasso.

Statistical Methods

Based on the above, our planned Bayesian logistic model will be as follows:

top_burrito ~ synergy + meat + meatfillings + fillings

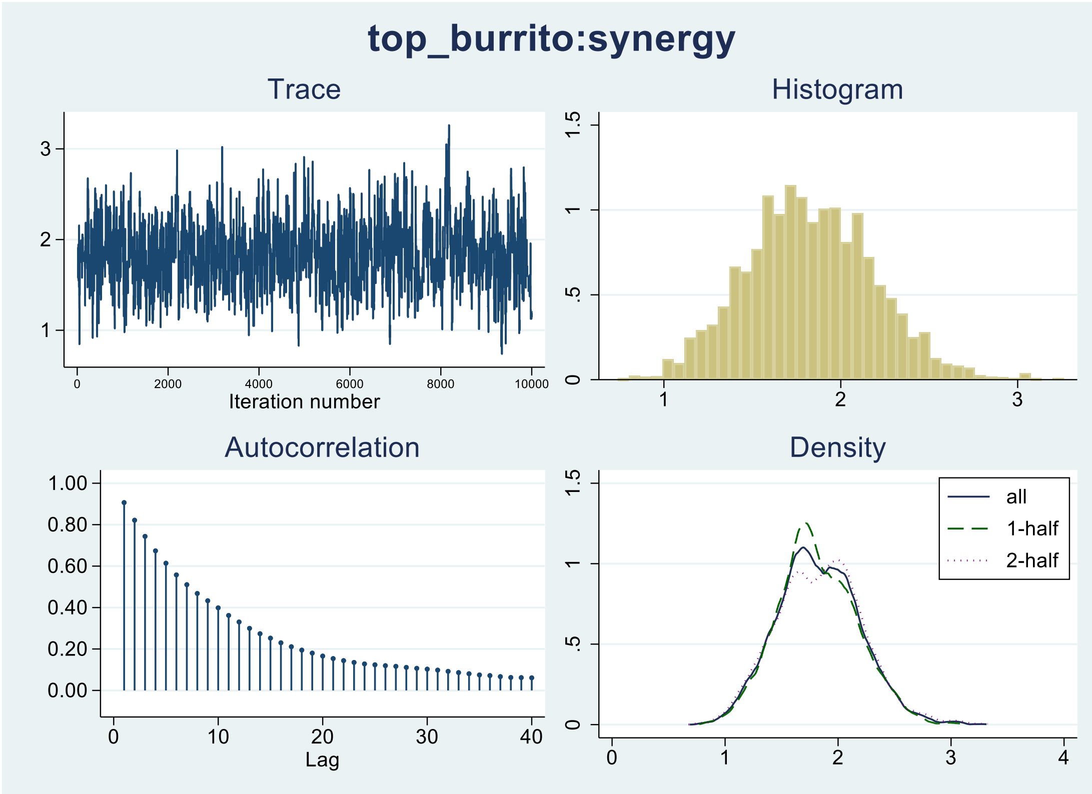
We fit the model below and obtain the following output:

```
. set seed 453

. bayes, burnin(5000):logit top_burrito c.(synergy meat meatfilling fillings)

Burn-in ...
Simulation ...

Model summary
-----------------------------------------------------------------------------
-
Likelihood:
  top_burrito ~ logit(xb_top_burrito)

Prior:
  {top_burrito:synergy meat meatfilling fillings _cons} ~ normal(0,10000)
(1)
-----------------------------------------------------------------------------
-
(1) Parameters are elements of the linear form xb_top_burrito.

Bayesian logistic regression                       MCMC iterations  =
15,000
```

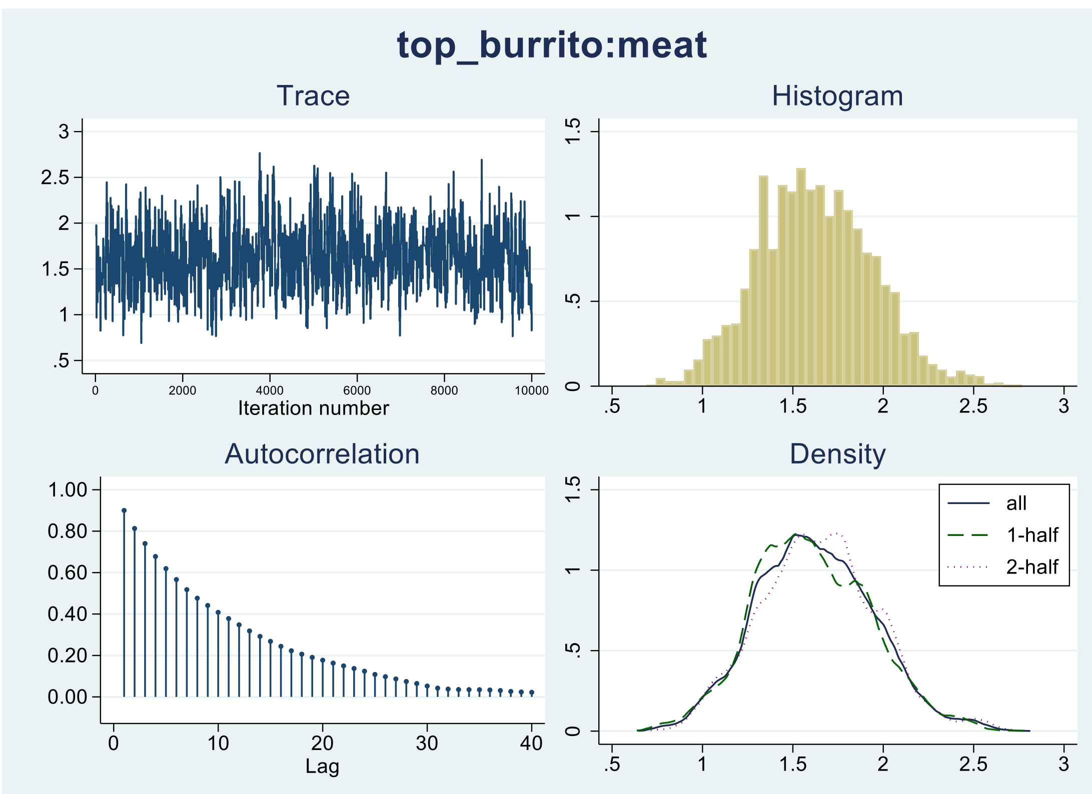
```
Random-walk Metropolis-Hastings sampling         Burn-in          =
5,000
                                                 MCMC sample size =
10,000
                                                 Number of obs    =
404
                                                 Acceptance rate  =
.256
                                                 Efficiency:  min =
.03947
                                                              avg =
.04696
Log marginal-likelihood = -160.34572                          max =
.05357

--------------------------------------------------------------------------------
-
             |                                                Equal-tailed
 top_burrito |      Mean   Std. Dev.     MCSE     Median  [95% Cred.
Interval]
-------------+------------------------------------------------------------------
-
     synergy |  1.826937   .3628998   .017686   1.810796    1.15451
2.556837
        meat |  1.624753   .3205945   .014894   1.612789   1.019055
2.276121
 meatfilling |  1.051048   .2255301   .009766   1.053912   .5946789
1.497281
    fillings |  1.203905   .3318068   .016702   1.198985   .5724701
1.848112
       _cons | -21.52936   2.332865   .100791      -21.5  -26.42333  -
17.18702
--------------------------------------------------------------------------------
-
Note: Default priors are used for model parameters.
```

Note that we used independent normal priors in the formulation of this model– these are weakly informative at best. Additionally, I requested a longer burn-in period in the call than Stata’s default in order to give the model a longer adaptation period.

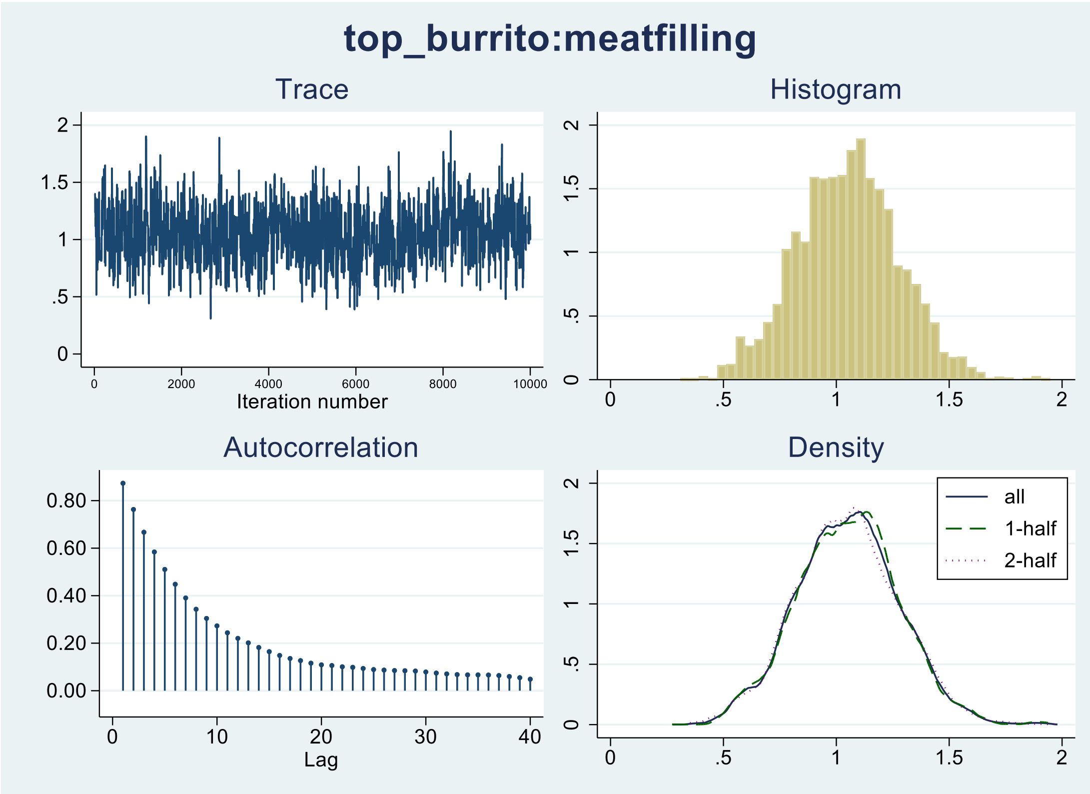
Before we get into the interpretation of the model, we should perform convergence diagnostics.

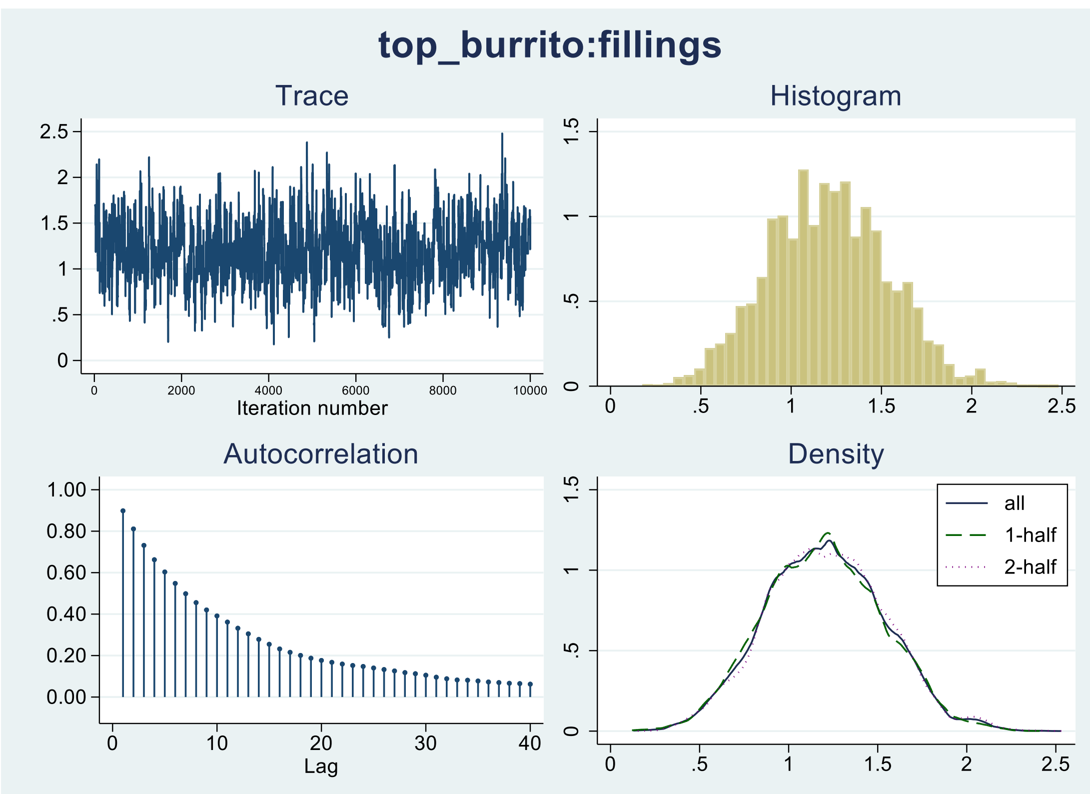


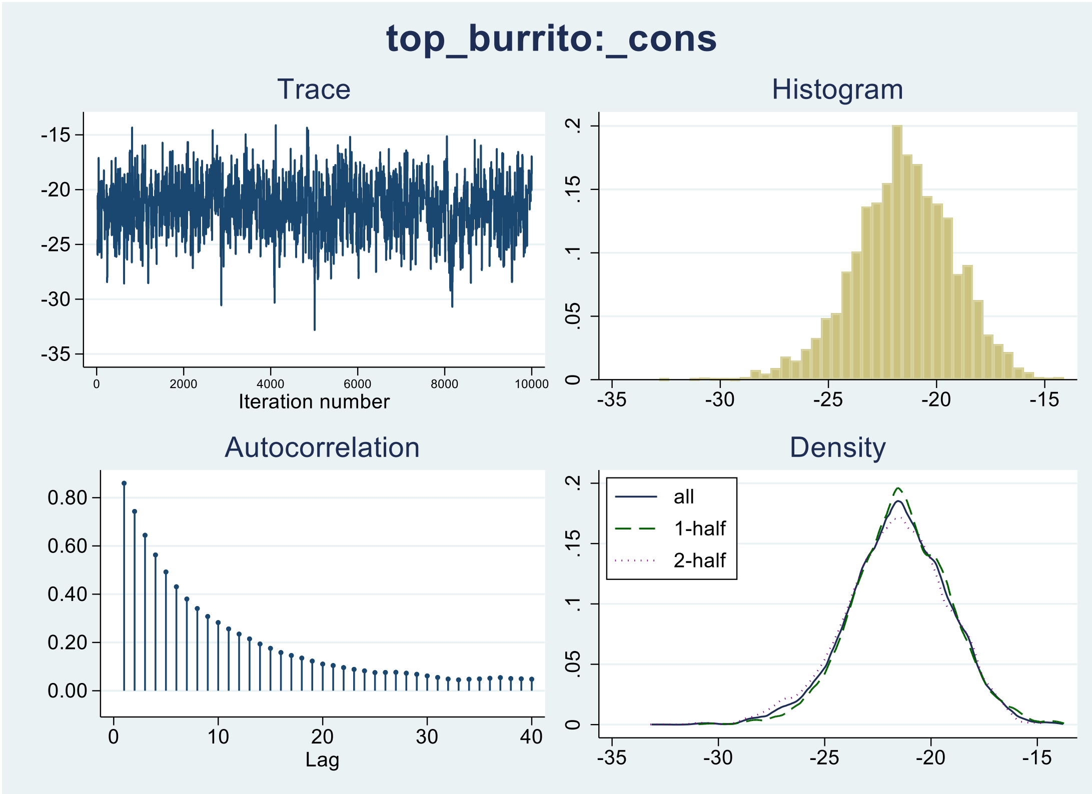
top_burrito:synergy
Trace
Histogram
Autocorrelation
Density
Iteration number
Lag
all
1-half
2-half



top_burrito:meat
Trace
Histogram
Autocorrelation
Density
Iteration number
Lag
all
1-half
2-half


top_burrito:meatfilling
Trace
Iteration number
Histogram
Autocorrelation
Lag
Density
all
1-half
2-half


top_burrito:fillings
Trace
Iteration number
Histogram
Autocorrelation
Lag
Density
all
1-half
2-half



Overall, these convergence plots look fine. The trace is constant for each and it looks like the algorithim mixed through the posterior distribution adequately. Our autocorrelations become negligible over time, another thing we expect to see in a converged model. Additionally, the histograms are in good agreement with the normal distribution.

Referencing the model output, our acceptance rate of 0.256 is in the desired range.

Our effective samples sizes for each covariate are given below.

```
. bayesstats ess

Efficiency summaries    MCMC sample size =    10,000
                        Efficiency:  min =    .03947
                                     avg =    .04696
                                     max =    .05357

----------------------------------------------------
top_burrito  |        ESS   Corr. time    Efficiency
-------------+--------------------------------------
     synergy |     421.03        23.75        0.0421
        meat |     463.32        21.58        0.0463
 meatfilling |     533.27        18.75        0.0533
    fillings |     394.68        25.34        0.0395
       _cons |     535.72        18.67        0.0536
----------------------------------------------------
```

We would ideally like to see ESS measures closer to the MCMC size, and efficiencies closer to 10%. One solution to this problem would be to run the model with higher sample sizes, or increase the number of chains.

Note: I took this step (rerunning with 2 chains and MCMC size of 30,0000) but still had issues with efficiency. I'm not including these results here for ease of exposition.

Results

Having examined the convergence diagnostics for our model, we turn to interpretation. I'll run our model again, but request odds ratios in the command.

```
. bayes, or

Model summary
--------------------------------------------------------------------------------
-
Likelihood:
  top_burrito ~ logit(xb_top_burrito)

Prior:
  {top_burrito:synergy meat meatfilling fillings _cons} ~ normal(0,10000)
(1)
--------------------------------------------------------------------------------
-
(1) Parameters are elements of the linear form xb_top_burrito.

Bayesian logistic regression                      MCMC iterations  =
15,000
Random-walk Metropolis-Hastings sampling          Burn-in          =
5,000
                                                  MCMC sample size =
10,000
                                                  Number of obs    =
404
                                                  Acceptance rate  =
.256
                                                  Efficiency:  min =
.03997
                                                               avg =
.06036
Log marginal-likelihood = -160.34572                           max =
.1211

--------------------------------------------------------------------------------
-
             |                                                Equal-tailed
 top_burrito |Odds Ratio   Std. Dev.     MCSE     Median  [95% Cred.
Interval]
-------------+------------------------------------------------------------------
-
     synergy |  6.649253   2.604627   .130276   6.115314    3.17247
12.89497
```

```
        meat |  5.350071   1.808814   .083706   5.016781   2.770574
9.738832
 meatfilling |  2.934869    .678627   .029336   2.868853   1.812449
4.46952
    fillings |  3.523644   1.217598   .060496    3.31675    1.77264
6.347822
       _cons |  5.05e-09   2.81e-08   8.1e-10   4.60e-10   3.35e-12   3.43e-
08
------------------------------------------------------------------------------
-
Note: _cons estimates baseline odds.
Note: Default priors are used for model parameters.
```

The baseline odds are given by the constant, which is close to zero. For an example of interpretation, the estimated odds of a "top burrito rating" (>= 4) are 6.6 times as large for an increase in synergy rating of one point, where the other covariates are held constant. The 95% credible associated with this odds ratio is (3.17, 12.89).

How do we obtain predictions for our model? Because of some peculiarities of Stata, the easiest way to do so is to code the model in a slightly different (but fairly equivalent) form using a different command which requests the Random-walk Metropolis-Hastings method specifically.

```
. set seed 453

. bayesmh top_burrito c.(synergy meat meatfilling  fillings),
likelihood(logit) prior({synergy} {meat} {meatfilling} {fillings} {_cons},
normal(0,10000)) burnin(5000
> ) saving(bayes, replace)

Burn-in ...
Simulation ...

Model summary
------------------------------------------------------------------------------
-
Likelihood:
  top_burrito ~ logit(xb_top_burrito)

Prior:
  {top_burrito:synergy meat meatfilling fillings _cons} ~ normal(0,10000)
(1)
------------------------------------------------------------------------------
-
(1) Parameters are elements of the linear form xb_top_burrito.

Bayesian logistic regression                     MCMC iterations  =
15,000
Random-walk Metropolis-Hastings sampling         Burn-in          =
5,000
                                                 MCMC sample size =
10,000
                                                 Number of obs    =
404
```

```
                                                Acceptance rate  =
.256
                                                Efficiency:  min =
.03947
                                                             avg =
.04696
Log marginal-likelihood = -160.34572                         max =
.05357

------------------------------------------------------------------------------
-
             |                                                Equal-tailed
 top_burrito |      Mean   Std. Dev.     MCSE     Median  [95% Cred.
Interval]
-------------+----------------------------------------------------------------
-
     synergy |  1.826937   .3628998   .017686   1.810796    1.15451
2.556837
        meat |  1.624753   .3205945   .014894   1.612789   1.019055
2.276121
 meatfilling |  1.051048   .2255301   .009766   1.053912   .5946789
1.497281
    fillings |  1.203905   .3318068   .016702   1.198985   .5724701
1.848112
       _cons | -21.52936   2.332865   .100791      -21.5  -26.42333  -
17.18702
------------------------------------------------------------------------------
-

file bayes.dta saved
```

Doing so, we can take advantage of Stata's suite of **bayesspredict** commands.

```
. bayespredict {_ysim1}, saving(predfile, replace)

Computing predictions ...

file predfile.dta saved
file predfile.ster saved


. use predfile.dta, clear


. list _chain _index _ysim1_1 _mu1_1 _frequency in 1/10

     +----------------------------------------------------+
     | _chain   _index   _ysim1_1      _mu1_1   _frequ~y |
     |----------------------------------------------------|
  1. |      1        1          0   .22726647          1 |
  2. |      1        2          0   .22726647          1 |
  3. |      1        3          0   .22726647          1 |
  4. |      1        4          0   .22726647          1 |
  5. |      1        5          0   .22726647          1 |
     |----------------------------------------------------|
  6. |      1        6          1   .22726647          1 |
```

```
  7. |       1         7          0    .22726647          1 |
  8. |       1         8          0    .22726647          1 |
  9. |       1         9          0    .22726647          1 |
 10. |       1        10          1    .22726647          1 |
     +----------------------------------------------------+
```

Focusing only on the first burrito, we see a list of the first 10 simulated predictions. This dataset of predictions is actually quite large; there are a total of 10,000 similar predicted values for each of the 423 burritos in the original data set. Essentially, the **bayesspredict** command uses samples the posterior predictive distribution for our outcome of *top_burrito* 10,000 times for each burrito.

Note that we can apply Stata's normal commands to this simulated data set. For example, let's summarize all 10,000 results for the first two burritos.

```
. bayesstats summary {_ysim[1] _ysim[2] _mu[1] _mu[2]} using predfile

Posterior summary statistics                      MCMC sample size =
10,000

------------------------------------------------------------------------------
-
             |                                                Equal-tailed
             |      Mean   Std. Dev.     MCSE     Median  [95% Cred.
Interval]
-------------+----------------------------------------------------------------
-
    _ysim1_1 |     .2881    .4529005    .00472          0          0
1
    _ysim1_2 |     .0007    .0264496   .000264          0          0
0
      _mu1_1 |  .2874842    .0643705   .002389   .2863605   .1678432
.4182073
      _mu1_2 |  .0005911    .0005618   .000022   .0004209    .000066
.0020245
------------------------------------------------------------------------------
-
```

There is much more that Stata can do for postestimation after Bayesian analysis. For example, we could perform out of sample predictions. However, in the interest of time I'll leave this here.

Comparison to Previous Analyses

Our main point of comparison is Scott Cole's principal components analysis, referenced previously in the "Background" section. This analysis found that the ratings for fillings, meat, meat/filling ratio, and uniformity were the most important (note: he excluded synergy rating as a predictor due to potential difficulty disentangling it from the outcome).

Our Bayesian model is largely in agreement with this analysis, with pertinent differences being our inclusion of synergy and lack of inclusion of uniformity, as it did not seem to be important in our variable selection process.

Limitations/Statistical Issues

The main limitation of this analysis is use of noninformative prior probabilities. Bayesian models using noninformative priors like ours may be less subjective, but this comes at the price of loss of customizability and practical application of probability knowledge. Arguably, we lose the chief advantages of the Bayesian approach in the first place.

```
. set seed 453

. bayes, burnin(5000): logit top_burrito c.(synergy meat meatfilling
fillings)

Burn-in ...
Simulation ...

Model summary
------------------------------------------------------------------------------
-
Likelihood:
  top_burrito ~ logit(xb_top_burrito)

Prior:
  {top_burrito:synergy meat meatfilling fillings _cons} ~ normal(0,10000)
(1)
------------------------------------------------------------------------------
-
(1) Parameters are elements of the linear form xb_top_burrito.

Bayesian logistic regression                       MCMC iterations  =
15,000
Random-walk Metropolis-Hastings sampling           Burn-in          =
5,000
                                                   MCMC sample size =
10,000
                                                   Number of obs    =
404
                                                   Acceptance rate  =
.256
                                                   Efficiency:  min =
.03947
                                                                avg =
.04696
Log marginal-likelihood = -160.34572                            max =
.05357

------------------------------------------------------------------------------
-
             |                                                Equal-tailed
 top_burrito |      Mean   Std. Dev.     MCSE     Median  [95% Cred.
Interval]
-------------+----------------------------------------------------------------
-
     synergy |  1.826937   .3628998   .017686   1.810796    1.15451
2.556837
```

```
        meat |  1.624753   .3205945   .014894   1.612789   1.019055
2.276121
 meatfilling |  1.051048   .2255301   .009766   1.053912   .5946789
1.497281
    fillings |  1.203905   .3318068   .016702   1.198985   .5724701
1.848112
       _cons | -21.52936   2.332865   .100791      -21.5  -26.42333  -
17.18702
------------------------------------------------------------------------------
-
Note: Default priors are used for model parameters.

. logit top_burrito c.(synergy meat meatfilling fillings)

Iteration 0:   log likelihood = -275.85367
Iteration 1:   log likelihood = -141.97749
Iteration 2:   log likelihood = -130.84006
Iteration 3:   log likelihood = -130.56487
Iteration 4:   log likelihood = -130.56421
Iteration 5:   log likelihood = -130.56421

Logistic regression                             Number of obs     =
404
                                                LR chi2(4)        =
290.58
                                                Prob > chi2       =
0.0000
Log likelihood = -130.56421                     Pseudo R2         =
0.5267

------------------------------------------------------------------------------
-
 top_burrito |      Coef.   Std. Err.      z    P>|z|     [95% Conf.
Interval]
-------------+----------------------------------------------------------------
-
     synergy |   1.765252    .356725     4.95   0.000     1.066084
2.46442
        meat |   1.566101   .3074037     5.09   0.000     .9636007
2.168601
 meatfilling |   1.024583   .2191053     4.68   0.000     .5951446
1.454022
    fillings |   1.168341   .3259696     3.58   0.000     .5294522
1.80723
       _cons |  -20.84379   2.233457    -9.33   0.000    -25.22128   -
16.46629
------------------------------------------------------------------------------
-
```

Looking at our output, now back on the log odds scale, we note that our estimated coefficients and confidence/credible intervals are quite similar between the two approaches. This would not be the case if we had different priors, perhaps from a pilot study or subject matter expertise. This would have likely helped us with the efficiency problems as well.

Another limitation is the logistic design of the outcome variable. It would have been better to keep the outcome as a multilevel, ordinal rating (I had convergence problems when I tried this). An additional limitation, in my opinion, is the lack of inclusion of any indicator variables corresponding to ingredients. It would have been nice to discover, for example, that the presence of guacamole or bacon or something else helped elevate a burrito's rating independent of other factors.

Future Directions

Research, as properly conceived, is dynamic– it is not a static event and cannot be contained in a single study or analysis. Having seen some of the things that factor into a quality burrito, and in the process learning a whole lot about Bayesian methods, I am not content to stop here. I suspect that it will take years of further study and the ingestion of hundreds of burritos for me to really get to the state of zen-like knowledge and culinary satisfaction I am hoping to achieve. However, I'm a scientist, and I'm willing to take the time, and the pounds, to see this labor of burrito love to completion.