Math 455 Class 31 April 11

New homework assigned.

One assignment: interpretation of regression coeff. in a model.

$$\log(Y) \sim \beta_0 + \beta_1 X_1 + \cdots + \beta_p X_p + \varepsilon$$

We discussed this in class. A brief explanation in the text book is on p. 134 in Ch. 9.

Also: hints in the problem set.

Important when reading p. 134: the preferred answer is in terms of "% increase", rather than "change by a factor of $e^{\hat{\beta}_1}$" when possible.

There will be a problem of this type on midterm 2.

Q: When do we change our model to a model of $\log Y$ instead of a model of Y?

A: Some quantities tend to change by a factor instead of a fixed amount.

(Multiplicative, not additive)

This is common in economics.

You'll know this is true for domain-specific reasons.

(E.g. you've worked on income data before and know it has this property.)

Q: What if we don't have any prior expectation?

A: If there is no prior expectation and a model for Y does not fit well, there is a standard way to search for a transformation: Box-Cox. (9.1)

There is also a problem in HW Ch. 8 where you search for a transformation.

This Box-Cox method is basically "maximum likelihood". We consider a family of transformations

$$g_\lambda(y) = \begin{cases} \dfrac{y^\lambda - 1}{\lambda} & \lambda > 0 \\ \log(y) & \lambda = 0 \end{cases}.$$

This funny function $\frac{y^\lambda - 1}{\lambda}$ is chosen

for continuity: $\lim_{\lambda \to 0} \frac{y^\lambda - 1}{\lambda} = \log y$.

This works only for positive values of Y:

if your data has negative values of Y, you can't use it.

What is produced is a graph of the likelihood.

You look at the graph and choose the *interpretable* value of λ closest to the maximum.

(bottom of p. 134)

[$\lambda = 0$ corresponds to $\log(Y)$]

Then you transform $Y \to Y^\lambda$.

(No need to do $\frac{Y^\lambda - 1}{\lambda}$ because this is linear in Y^λ and won't change anything fundamental.)

What does "interpretable" mean?

Changing $Y \to \log Y$ is "interpretable";
see the homework.

Changing $Y \to \sqrt{Y}$ might be interpretable.

It's hard to see how to interpret the change

$$Y \to Y^{3.719}$$

Note: this method is powerfully influenced by outliers, if you're finding $\lambda = 9.637$, this might be why.

FYI We looked at code in R which produces a graph of log-likelihood.
boxcox fcn in MASS package.

Typical test question: shows command and graph, asks what to do.

Answer: transform $Y \mapsto Y^{\lambda}$ (or $\log Y$, if $\lambda = 0$)
where λ is interpretable value close to maximum.