Multiple Linear Equation Models

Ethan Ligon

March 13, 2022

Beginning

The beauty of econometrics and modern ways of thinking about estimation and inference are easy to lose sight of when we get caught up in technical details. But we're developing tools that really help us to understand the world we live in, and in particular to learn about economic behavior. This is a profound endeavor!

Naïveté

You've spent the first half of the semester developing a toolkit of linear methods in econometrics. For today, I want you to set that toolkit aside, and to think about some very basic issues, starting from a very naïve perspective. You've probably seen *all* of this material before, but I want to be sure we're all aware of the forest as we walk along examining trees.

Notation for Random Variables

Setting	Scalar	Vector	Matrix
Statistics	X	$\boldsymbol{x}$	$\boldsymbol{X}$
Econometrics	x	$\boldsymbol{x}$	$\boldsymbol{X}$
PDF	X or x	$\boldsymbol{x}$	$\boldsymbol{X}$
Jupyter	`x`	`x`	`X`
Handwriting			

Defining Random Variables in python

See `random_variables0.ipynb` on `datahub`.

The Fundamental Linear Regression Model

Start with

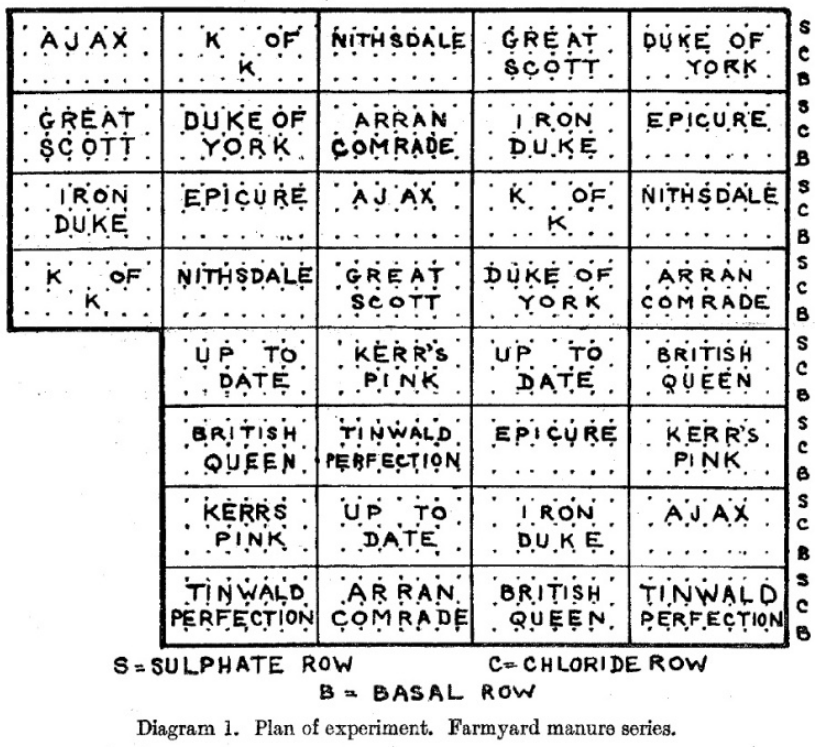
$$y = \boldsymbol{X}\beta + u.$$

- Allowing $(y, \boldsymbol{X}, u)$ to all be random.
- $\boldsymbol{X}$ has full column rank.

Compare Classical Approach

E.g., R.A. Fisher; Fisher Box 1980

$$y = \boldsymbol{X}\beta + u.$$



AJAX	K OF K	NITHSDALE	GREAT SCOTT	DUKE OF YORK	S
GREAT SCOTT	DUKE OF YORK	ARRAN COMRADE	IRON DUKE	EPICURE	C
IRON DUKE	EPICURE	AJAX	K OF K	NITHSDALE	B
K OF K	NITHSDALE	GREAT SCOTT	DUKE OF YORK	ARRAN COMRADE	S
	UP TO DATE	KERR'S PINK	UP TO DATE	BRITISH QUEEN	C
	BRITISH QUEEN	TINWALD PERFECTION	EPICURE	KERR'S PINK	B
	KERRS PINK	UP TO DATE	IRON DUKE	AJAX	S
	TINWALD PERFECTION	ARRAN COMRADE	BRITISH QUEEN	TINWALD PERFECTION	C
					B

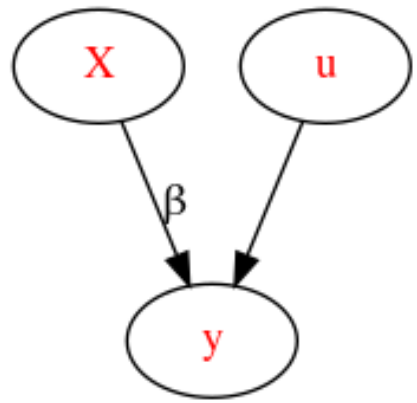
S = SULPHATE ROW C = CHLORIDE ROW
B = BASAL ROW

Diagram 1. Plan of experiment. Farmyard manure series.

Figure: "Triplicate Chessboard": Diagram I from Ronald A Fisher and Winifred A Mackenzie. 1923. Studies in crop variation. II. the manurial response of different potato varieties. *The Journal of Agricultural Science* 13 (3): 311–320

Classical Interpretation

The "dependent" variable y is determined by some random variables X with observations realized, and some random unobserved u. Critically, u and X are orthogonal; i.e., $\mathbb{E}(X^\top u) = \mathbf{0}$.



- The orthogonality of X and u is *not testable*, since u isn't observed.
- In general the causal diagram above imposes needed structure for interpreting regression.
- With this structure, β is "effect of variation in X on y."

Classical Regression in python

See `classical_regression.ipynb` on `datahub`.

Compare Bayes

$$y = \boldsymbol{X}\beta + u.$$

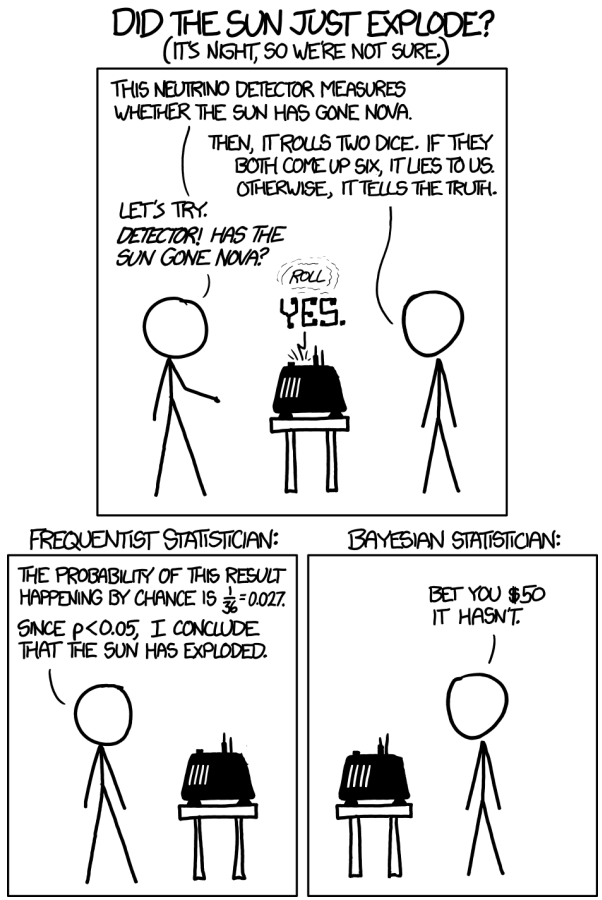


Figure: `https://xkcd.com/1132/`

Bayesian Interpretation

The disturbance u is independently distributed $Q(u|\sigma)$, and there's a "prior" $\Pr(\beta = \beta, \sigma = \sigma) = \Pr(\beta, \sigma)$ over these unknown vectors. The variables (β, σ) are ordinarily assumed to be distributed independently of $(\boldsymbol{X}, y)$.