2.3 Multiple Regression

Wednesday, September 8, 2021 10:01 AM

- Cont. Issues
- Economic analysis
- Game theory
- Econometrics: Causal inference, Panel + survey data

Credibility Revolution

1. Context. Why? – Hail
2. What was it? – Graham
3. Ill effects
   a. Not enough focus on external validity – Kori
   b. Too much focus on small questions – Jake
4. Rebuttal to #3 – Maverick

Regression

	Sandwiches	Sleep
Logan	8	7
Hail	2	8
Maverick	2	6
Gus	2	8
Kori	1	5
Nicole	1	6
Graham	5	7

Average Sandwiches:
- Females: 1.33
- Males: 4.25
- Total: 4

$min \sum_i (S_i - \hat{S})^2$

$2\sum(S_i - \hat{S}) = 0$

$\sum S_i = \sum \hat{S} = 7\hat{S} = n\hat{S}$

$\hat{S} = \frac{\sum S_i}{n}$

Split the prediction equation for males and females
Best predictor value is the average number for that group

Regression: Compute $E(y|x, x_1, x_2, x_3, \dots, x_k)$

$\min_{\hat{y}} \sum_{i \in g} (y_i - \hat{y}_g)^2 \rightarrow \hat{y}_g = \bar{y}_g = E(y|g)$

$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 + \dots + \hat{\beta}_k x_k$

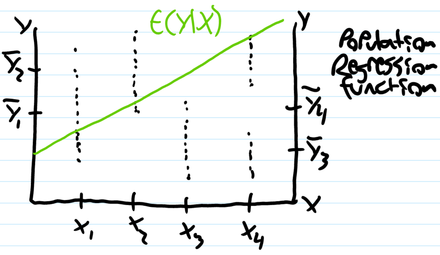
Small # groups, Categorical vars

ANOVA

$E(y|X) = X\beta$ ← linear function of x

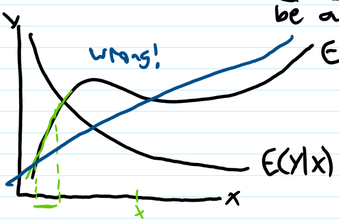
↑ vector of x

Y	X1	X2	X3	X4	X0	B
	1	40	3		1	B1
	0	90	8		1	B2
	0	75	12		1	B3
	1	66	8		1	B4
	1	12	7		1	B5
	1	11	9		1	B6
	0	103	6		1	B7
	1	41	5		1	B8
	0	52	4		1	B9



$\bar{y}$ = Averages

IF I drew it right, there would be a line for the regression



$\ln Y = \beta_0 + \beta_1 \ln X$

$Y = \beta_0 + \beta_1 X_1 + \beta_2 X^2 + \beta_3 X^3$

$Y = \beta_0 + \beta_1 / X$

β is the unknown

$\min \sum (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_{1i} - \dots - \hat{\beta}_k x_{ki})^2 \rightarrow \min \sum e_i^2$

↳ $\min \sum (y_i - \hat{y}_i)^2$

$\hat{\beta} = (X^T X)^{-1} X^T y$

$\frac{dSS_i}{d\hat{\beta}_1} = -2\sum(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_{1i} - \dots) x_{1i} = 0$

Residual

$\sum (x_{ki} e_i) = 0 \quad \forall k$

Picks the line so there's no correlation between the explanatory + unpredicted part

$y_i = \hat{y} + \hat{e}_i$

$E(\beta) = \sum \sum (X\beta + \epsilon) = (X^T X^{-1}) X^T X \beta + (X^T X)^{-1} X^T \epsilon =$

$\beta + (X^T X)^{-1} X^T \epsilon$

$= 0$

$Sandwich_i = \beta_0 + \beta_1 sex_i + \beta_2 sleep_i + \beta_3 Income + \epsilon_i$

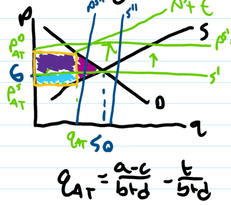
$= \beta_0 + \beta_1 sex_i + \beta_2 sleep_i + \beta_3 Income_i + \beta_4 Ad_i + \epsilon_i$

For causation, need sleep, sex, income, and uncorrelated with ϵ_i

ϵ_i: combined effect of other causes

$P^S = c + dq^S$ → supply
$P^D = a - bq^D$ → demand

for iced coffee → $P^S = P^D$, $q^S = q^D = q$



after tax... $t + P^S = P^D$

$t + P^S = a - bq$

$t + c + dq = a - bq$

$(b+d)q = a - c - t$

$q = (a - c - t)/(b + d)$

$q_T = \frac{a-c}{b+d} - \frac{t}{b+d}$

Consumer surplus
Producer surplus
Deadweight loss
Government revenue