Shenzhen Lecture 1 (May 17)

Conditional Expectation.

(Introduce concept).

1. Review of distributions.

1) Univariate

① cumulative distribution function.

$$F(x) = \Pr(X \leq x)$$

non-decreasing. $F(-\infty) = 0$, $F(\infty) = 1$

② pdf: $f(x) = \frac{\partial}{\partial x} F(x)$.

③ ~~[illegible]~~ expectation: $E(x) = \int x f(x) dx$.

4. Bivariate distribution

$$F(x, y) = \Pr(X \leq x, Y \leq y)$$

$$f(x,y) = \frac{\partial F(x,y)}{\partial x \partial y}$$

Conditional expectation.

$$E(y|x) = \int y f(y|x) dy.$$

$$= \int y \frac{f(y,x)}{f(x)} dy$$

Law of iterated expectation

$$E(E(y|x)) = E(y)$$

the average of conditional means — unconditional mean.

$$\int\left(\int y \frac{f(x,y)}{f(x)}\,dy\right) f(x)\,dx$$

$$= \int\int y f(x,y)\,dy\,dx$$

$$\int (E(y|x)) f(x)\,dx$$

example: average score

= female avg score × Pr(female)

+ male avg score × Pr(male)

properties:

1. $E(E(y|x_1, x_2)|x_1) = E(y|x_1)$

$E(E(y|x_1)|x_1, x_2) = E(y|x_1)$

2. $E(g(x)y|x) = g(x)E(y|x)$

$E(g(x)y) = E(g(x)E(\ldots$

Conditional expectation function

$E(y \mid x) = m(x)$

$y = m(x) + e$

where $e = y - m(x)$

By definition $E(e \mid x)$

$= E[(y - m(x)) \mid x] = 0.$

Properties:

① $E(e \mid x) = 0$

② $E(e) = 0$

③ $E(h(x)e) = 0$ (by LIE)

e is uncorrelated with any function of x.

e is mean indep of x. Weaker ~~than~~ "fully indep".

Regression variance

$\sigma^2 = E(e^2)$ is the variance of CEF error.

Interpretation: the variance unexplained by

Lecture 2 (1st)

Property:

$$Var(y) \geq Var(y - E(y|X)) \geq Var[y - E(y|X_1, X_2)].$$

Conditional mean as the best predictor of y in MSE

$$E(y - g(x))^2 = E(y - E(y|X) + E(y|X) - g(x))^2$$

$$= ① + E\{[y - E(y|X)][E(y|X) - g(x)]\} \quad ③$$

only indep on the selection of x.

independent of g(x)

$$③ = E\{(E(y|X) - g(x))\, E[(y - E(y|X))|X]\}$$

again, ② is irrelevant to g(x).

$\cancel{E(y - E(y|X))[E(y|X) - g(x)]}$

$\cancel{\overset{LIE}{=} E(E[(E(y|X) - g(x))(y - E(y|X))|X])}$

$\cancel{= E(E(y|X) - g(x))}$

The last term is minimized at g

Conditional variance

Def $\sigma^2(X) = E(e^2 | X]$

$= E((y - E(y|X))^2 | X]$.

Def. homoskedastici

$\sigma^2(X) = \sigma^2$ is indep of X.

Regression derivative

$\frac{\partial}{\partial X_1} m(X_1, \dots, X_k)$

partial derivative. Hold other variables constant.

if X_1 is binary,

derivative is $m(1, X_2 \; X_k) - m(0, X_2 \dots X_k)$

The condition mean doesn't hold all else constant.

(doesn't hold for these variables not included.

Linear CEF

is a special case. $m(X) = X\beta$

In this model, the CEF derivative is the coefficient.

The "marginal effect.".

The linear CEF can add non-linear effect.

$$m(X_1, X_2) = X_1\beta_1 + X_2\beta_2 + X_1^2\beta_3 + X_2^2\beta_4 + X_1X_2\beta_5 + e.$$

$$\frac{\partial m}{\partial X_1} = \beta_1 + X_2\underline{\underline{\beta_5}} + 2\beta_3 X_1$$

interaction effect.

Linear CEF with dummy variables

If all regressors take a finite set of values, CEF can be written as a linear function of regressors (exact). Not a special case.

If ~~x~~ takes two values, ~~no matter what,~~

~~the~~ usually assign $\{0, 1\}$

E[M ...] ... sex ...

then $E(y|x) = \mu_0 + (\mu_1 - \mu_0)\cdot sex$.

two dummies married unmarried.

μ_{00} μ_{01}

μ_{10} μ_{11}.

$$y = \mu_0 + \beta_1 mar + \beta_2 sex + \beta_3 mar \cdot sex.$$

categorical variables. $\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & & \end{pmatrix}$ 1, 2, 3 is not meaningful

---

upto p28.