دانلود جزوه اقتصاد سنجی

پروفسور احمد جعفری صمیمی

استاد دانشگاه مازندران

انتشار جزوه : سال 91

انواع داده‌ها:
۱. مقطعی ← داده‌های گروهی از کشورها ... در یک زمان معین
۲. سری زمانی ← داده‌هایی که در طول زمان مربوط به یک شخص، استان، یا کشور جمع‌آوری می‌شود
۳. پنل (تابلویی) ← ترکیبی از دو حالت قبلی ← داده‌های مقطعی در طی زمان

مثال: درآمد اتفاقی الف و ب در سالهای ۸۹ و ۹۰

مثال ۱: درآمد آقای الف و ب و ج در سال ۸۹ — تورم در مرز دو هزار و ...
درآمد کشور x در سال ۸۸

آقای	درآمد
الف	۱۰۰
ب	۹۰
ج	۸۰

$\bar{x} = \frac{100+90+80}{3} = 90$

مثال ۲: درآمد کشور ایران در سال ۷۰ و ۷۱

زمان	x
۸۹	۱۰۰
۹۰	۸۰

مثال حالت سوم:

درآمد	۸۹	۹۰	۹۱
الف	۱۰۰	۱۵۰	۴۹
ب	۳۰	۲۰	۵
ج	۰	۰	-

۳×۳

	درآمد x_i	$x_i - \bar{x}$
الف	۱۰۰	+۱۰
ب	۹۰	۰
ج	۸۰	-۱۰
		۰

$\bar{x} = 90$

$\Sigma(x_i - \bar{x}) = +10 + 0 + (-10) = 0$

نکته:
مجموع انحراف داده‌ها از میانگین برابر صفر است.

$\Sigma(x_i - \bar{x})^2 = 100 + 0 + 100 = 200$

$s^2 = \frac{\Sigma(x_i - \bar{x})^2}{\underbrace{n-1}_{\text{درجه آزادی}}} = \frac{200}{2} = 100$

$\sqrt{s^2} = s$
انحراف معیار
خطای معیار

PAPCO
۱

Subject:
Year: Month: Date: ( )

داده‌ها متقابل در برای دو گروه

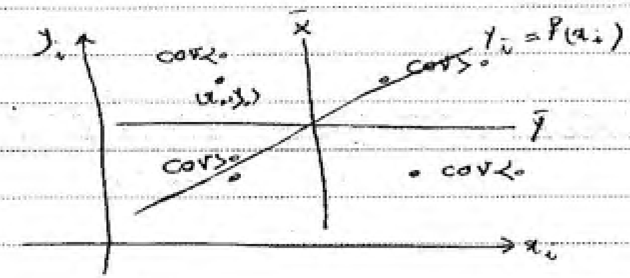
	نمره آمار x_i	نمره معارف y_i	$x_i-\bar{x}$	$y_i-\bar{y}$	$(x_i-\bar{x})(y_i-\bar{y})$
الف	۱۰۰	۹۵	۱۰	۸,۳	۸۳
ب	۹۰	۹۰	۰	۳,۳	۰
ج	۸۰	۷۵	−۱۰	−۱۱,۶	۱۱۶
	$\bar{x}=۹۰$	$\bar{y}=\frac{۲۶۰}{۳}$			۱۹۹

$\bar{y}=۸۶,۶$

برای بدست آوردن رابطه بین x و y باید از کوواریانس استفاده کرد.

$$cov(x,y)=E(x_i y_i)-E(x_i)E(y_i)=E(x_i-\bar{x})\,E(y_i-\bar{y})$$
$$=\frac{۱۹۹}{۳}$$

تفسیر کوواریانس:



۱. cov>0: رابطه مستقیم دو داده‌ها است.

۲. cov<0: رابطه معکوس دو داده‌ها است.

۳. cov=0: دو تابع مستقل اند

$$E(x_i y_i)=E(x_i)\,E(y_i)$$

شرط استقلال وقتی دو تابع مستقل اند

به زبان احتمالات ⟵ $P(A\cap B)=P(A)\cdot P(B)$

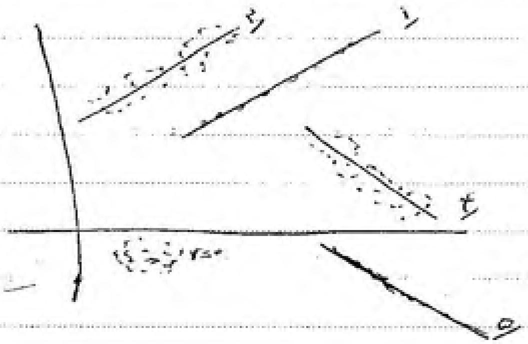
نکته: اشکالات اساسی کوواریانس در این است که دامنه احتمالات آن از منفی تا مثبت است. و میزان ضعف یا قدرت همبستگی بین دو متغیر را نشان نمی‌دهد برای رفع این مشکل

کوواریانس را بر انحراف معیار دو گروه تقسیم می‌کنیم تا ضریب همبستگی را دامنه آن بین

PaPCO

... را بدست آوردیم

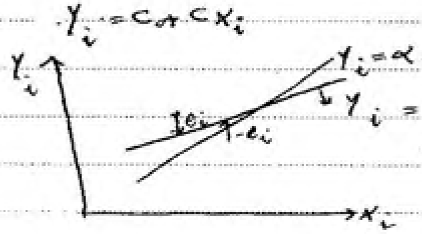
ضریب همبستگی خطی $= r = \dfrac{cov(x,y)}{S_x S_y}$



$r=+1$ ← رابطه خطی کامل مستقیم
$0<r<+1$ ← رابطه مستقیم ناقص
$r=0$ ← ~~رابطه~~ استقلال دو متغیر
$-1<r<0$ ← رابطه غیر مستقیم ناقص
$r=-1$ ← رابطه خطی کامل غیر مستقیم

رگرسیون: یعنی بدست آوردن یک خط بین دو گروه داده.
در واقع تخمین ضرایب خطی α و β $\qquad Y = \alpha+\beta X_i + U_i$
پسماند، انحراف داده ها از خط رگرسیون می باشد

$Y_i = c_0 + c X_i$



$Y_i = \alpha + \beta X_i + U_i$ ← خط جامعه
$Y_i = a + b X_i + e_i$ ← خط نمونه

X_i متغیر مستقل
Y_i متغیر وابسته

از جامعه نمونه گیری می کنیم زیرا برآورد در جامعه هزینه بر است و وقت گیر.
در واقع ما خط جامعه را نمی بینیم ولی موجود است، باید بهترین خطی که با جامعه انطباق پیدا می کند را پیدا کنیم.
روش OLS: بهترین برآورد کننده خطی
فروض روش OLS:
الف) $E e_i = 0$ یا $\sum e_i = 0$
ب) $\sum x_i e_i = 0$ ← $\sum x_i$ تصادفی نبوده درحالیکه e_i ها تصادفی اند.
یعنی استقلال بین x_i و e_i

ج) $\sigma^2_{e_i} = \sigma^2$ واریانس همسانی

د) $\sum (e_i e_j) = 0$ عدم همبستگی بین eها

ه) e_i توزیع نرمال دارند.

روش اول برای برآورد a,b : $Y_i = a + bX_i + e_i$

فرض اول
فرض دوم

معادلات نرمال
$$\begin{cases} \sum Y_i = \sum a + \sum bX_i + \sum e_i \\ \sum X_iY_i = \sum aX_i + \sum bX_i^2 + \sum X_i e_i \end{cases}$$

$$\begin{cases} \sum Y_i = na + b\sum X_i + 0 \\ \sum X_iY_i = a\sum X_i + b\sum X_i^2 \end{cases}$$

$\to \{a = \bar{Y} - b\bar{X}\}$

$\to \left\{ b = \frac{n\sum X_iY_i - \sum X \sum Y_i}{n\sum X_i^2 - (\sum X_i)^2} \right\}$ به اثبات مهم است

$a = \frac{\sum Y_i}{n} - b\frac{\sum X_i}{n}$

$b = \frac{cov(X,Y)}{var(X)}$

$$b = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sum (X_i - \bar{X})^2} = \frac{\sum X_iY_i - n\bar{X}\bar{Y}}{\sum X_i^2 - n\bar{X}^2}$$

روش دوم برآورد a,b

* eها پسماند یا خطای برازش
چون $\sum e_i = 0$ بنابراین $\sum e_i^2$ را مینیمم می‌کنیم

روش حداقل مربعات $\min \sum (Y_i - a - bX_i)^2$

مشتق نسبت به a: $-2\sum (Y_i - a - bX_i) = 0 \to \sum Y_i - \sum a - b\sum X_i = 0$

$\to \sum Y_i = na + b\sum X_i \to a = \bar{Y} - b\bar{X}$

مشتق نسبت به b: $-2\sum (Y_i - a - bX_i)(X_i) = 0 \to \sum X_iY_i - \sum aX_i - b\sum X_i^2$

$\to \sum X_iY_i = \sum aX_i + b\sum X_i^2 \to$ ادامه دارد

PAPCO

مثال: بدست آوردن خط رگرسیون

Y_i (مصرف)	X_i (درآمد)	X_iY_i	X_i^2	Y_i^2
۲۲	۸	۱۷۶	۶۴	۴۸۴
۲۳	۱۰	۲۳۰	۱۰۰	
۱۸	۷	۱۲۶	۴۹	۵۲۹
۹	۲	۱۸	۴	
۱۴	۴	۵۶	۱۶	۳۲۴
۲۰	۶	۱۲۰	۳۶	۸۱
۲۱	۷	۱۴۷	۷	
۱۸	۶	۱۴۸	۶	
۱۶	۴	۶۴	۱۶	
۱۹	۶	۱۱۴	۳۶	
۱۸۰ مصرف	۶۰ درآمد	۱۱۵۹	۴۰۶	

$\bar{X} = 6 \qquad \bar{Y} = 18$

$$b = \frac{\sum X_iY_i - n\bar{X}\bar{Y}}{\sum X_i^2 - n\bar{X}^2} = \frac{1159 - 10 \times 6 \times 18}{406 - 10(6)(6)} = 1.77$$

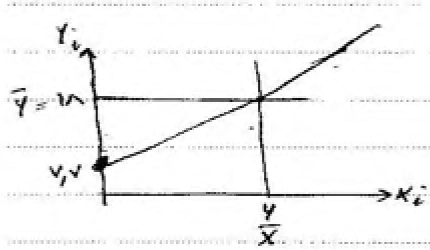
$$a = 18 - 1.77 \times 6 = 7.69 \qquad \hat{Y} = 7.69 + 1.77\hat{X}$$

$$Y_i = 7.69 + 1.77X_i + e_i$$

به ازای تغییر X_iها به میزان ۱٫۷۷ به Y_i اضافه می‌شود.

اگر X_i صفر شود میزان Y_i برابر ۷٫۶۹ است.

اگر X_i یک واحد تغییر کند به میزان ۱٫۷۷ واحد Y تغییر کند.



$$b = \frac{\Delta Y_i}{\Delta X_i}$$

کاربرد اصلی مدلها پیش بینی

۳

به a و b ها می گویند پارامتر
به α و β می گویند آماره. صفحه ۴۷ به بعد

Subject:
Year. Month. Date. ( )

قضیه گاوس-مارکو:

این قضیه بیان می دارد که روش ols در بین بهترین روش تخمین زننده های خطی می باشد با خصوصیات:

۱- a و b خطی هستند نسبت به Y_i

$$b = \frac{\sum Y_i (x_i - \bar{x})}{\sum (x_i - \bar{x})^2} = \sum Y_i w_i$$

متغیر x مستقل است یعنی مقادیر ثابت دارد و تصادفی نیست.

$$a = \frac{\sum Y_i}{n} - b\bar{x} = \frac{\sum Y_i}{n} - (\sum Y_i w_i)\bar{x} = \sum Y_i \underbrace{\left(\frac{1}{n} - w_i\bar{x}\right)}_{c_i} = \sum Y_i c_i$$

۲- a و b نااریب (ناتور) هستند

$$E(a) = \alpha$$
$$E(b) = \beta$$

آماره نمونه — پارامتر جامعه

$$b = \sum w_i Y_i = \sum w_i (\alpha + \beta x_i + U_i) = \sum \alpha w_i + \sum \beta x_i w_i + \sum U_i w_i$$

$$b = \alpha \underbrace{\sum w_i}_{0} + \beta \underbrace{\sum x_i w_i}_{1} + \sum U_i w_i \longrightarrow E(b) = \beta + E(\underbrace{\sum U_i w_i}_{0})$$
$$E(b) = \beta$$

$$\sum w_i x_i = \frac{\sum (x_i - \bar{x}) x_i}{\sum (x_i - \bar{x})^2} = 1 \longrightarrow$$ صورت و مخرج هر دو جمله واریانس هستند

$$\sum w_i = \frac{\sum (x_i - \bar{x})}{\sum (x_i - \bar{x})^2} = 0$$

تمرین: ثابت کنید $E(a) = \alpha$

$$a = \bar{Y} - b\bar{x} \longrightarrow \bar{Y} = \alpha + \beta\bar{x}$$
$$\Rightarrow a = \alpha + \beta\bar{x} - b\bar{x} \longrightarrow E(a) = \alpha + \cancel{\beta}\bar{x} - \bar{x}\,\underbrace{\cancel{E(b)}}_{\beta}$$
$$\longrightarrow E(a) = \alpha$$

PAPCO

Subject:
Year: Month: Date: ( )

(۵) ~~[illegible]~~ (بهترین تخمین خطی) a و b کمترین واریانس را در میان تخمین زننده‌های خطی دارا باشند.

$$var(b) = E(b-E(b))^2 = E(b-\beta)^2 = E(\sum u_i w_i)^2 = E(\sum u_i^2 w_i^2 + 2\sum_i \sum_j u_i w_i u_j w_j)$$

$$(u_1 w_1 + u_2 w_2 + \dots + u_n w_n)^2$$

$$b = \beta + \sum u_i w_i$$

$$b - \beta = \sum u_i w_i$$

فرض سوم کلاسیک u_i ها مستقل اند

$$var(b) = E(\sum u_i^2 w_i^2) = \sigma^2 \sum w_i^2 = \sigma^2 \frac{1}{var x_i} = \frac{\sigma^2}{var(x_i)}$$

$$E(\sum u_i^2)\ \sum w_i^2 \qquad \sum w_i^2 = \frac{\sum (x_i-\bar{x})^2}{(\sum (x_i-\bar{x})^2)^2} = \frac{1}{\sum (x_i-\bar{x})^2}$$

$$var(a) = E(a - E(a))^2$$

نکته $var(x-y) = var(x) + var(y)$

$$\therefore a = \bar{y} - b\bar{x} \longrightarrow \sigma_a^2 = \sigma_{\bar{y}}^2 + \bar{x}^2 \sigma_b^2 = \frac{\sigma^2}{n} + \bar{x}^2 \frac{\sigma^2}{\sum (x_i-\bar{x})^2} =$$

$$\sigma_{\bar{y}}^2 = \frac{\sigma^2}{n}$$

$$\sigma_a^2 = \sigma^2 \left(\frac{1}{n} + \frac{\bar{x}^2}{\sum (x_i-\bar{x})^2}\right)$$

علت اینکه OLS بهترین تخمین زننده می‌باشد طبق قضیه گاوس مارکوف

واریانسی که از طریق OLS بدست می‌آید به حداقل ممکن رسیده زیرا اطلاعات [illegible] است پیدا می‌کند.

$$Y_i = a + bx_i + e_i$$

$$\bar{Y} = \bar{a} + b\bar{x} + \bar{e}$$

$$\sum \left[Y_i - \bar{Y} = bx_i - b\bar{x} + e_i - \bar{e}\right]^2 \longrightarrow \sum (Y_i - \bar{Y})^2 = b^2 \sum (x_i - \bar{x})^2 + \sum e_i^2 + 2\sum e_i (x_i - \bar{x}) b$$

$$b(x_i - \bar{x}) + e_i$$

$$\sum e_i^2 = \sum (Y_i - \bar{Y})^2 - b^2 \sum (x_i - \bar{x})^2$$

می‌توان نشان داد $\frac{\sum e_i^2}{n-2} = \sigma_u^2 = \sigma^2$ که $n-2$ درجه آزادی نام دارد و علت کم شدن ۲ واحد

PAPCO

۲

از تعداد نمونه‌ها این است که تخمین a و b داریم.

مفهوم درجه آزادی چیست؟

* ۳ عدد را طوری انتخاب کنید که میانگین ۱۰ شود.
اگر قرار باشد میانگین ۳ عدد به فرض ۱۰ باشد ۲ عدد هرچه می‌خواهد باشد ولی عدد سوم باید طوری باشد که جمع اعداد تقسیم بر ۳، ۱۰ شود.

ضریب همبستگی $r = b\frac{S_x}{S_y} = \frac{cov(x,y)}{\sqrt{S_x^2 \cdot S_y^2}} = \frac{cov}{S_x \cdot S_y} \cdot \frac{S_x}{S_x} = \frac{cov}{S_x^2} \cdot \frac{S_x}{S_y} = b\frac{S_x}{S_y}$

$$b = \frac{cov}{S_x^2}$$

$$r^2 = b^2 \frac{S_x^2}{S_y^2} = b^2 \frac{\frac{\sum(x_i-\bar{x})^2}{n}}{\frac{\sum(y_i-\bar{y})^2}{n}} = b^2 \frac{\sum(x_i-\bar{x})^2}{\sum(y_i-\bar{y})^2} = \frac{\text{میزان تغییرات رگرسیون}}{\text{میزان کل تغییرات}}$$

$0 \le r^2 \le 1$

$$r^2 = \frac{\sum(y_i-\bar{y})^2 - \sum e_i^2}{\sum(y_i-\bar{y})^2} = 1 - \frac{\sum e_i^2}{\sum(y_i-\bar{y})^2}$$

آزمون F:

$$F = \frac{b^2 \sum(x_i-\bar{x})^2}{\frac{\sum e_i^2}{n-2}}$$

توزیع نرمال: ۱) ۶۸٪ افراد جامعه بین $\bar{X} \pm \sigma$
۲) ۹۵٪ $\bar{X} \pm 2\sigma$
۳) ۹۹٪ $\bar{X} \pm 3\sigma$

متغیری که دارای این ویژگی‌ها باشد دارای توزیع نرمال است.

مثال ۵: فرض کنید قوطی رب توسط شرکت تبرک دارای وزنی به اندازه $\underset{\bar{x}}{950}$ گرم و انحراف معیار ۱۵ واحد یک فاصله اطمینان ۹۵٪ برای میانگین بسازید.

σ این بین جمع قوطی ها ۹۵۰ گرم وزن ممکن است ۱۵ گرم کم و زیاد داشته باشد.

آیا قوطی ربی که تصادفی انتخاب شده و وزن ۹۸۵ گرم را دارد در سطح اطمینان ۹۵ درصد قرار گرفته است؟

چون ۹۵٪ است در فاصله $\bar{x} \pm 2\sigma$ قرار دارد.

$\mu: \bar{x} \pm Z_{\alpha/2}\,\sigma$

$\mu: 950 \pm 2 \times 15 \longrightarrow 920 < \bar{x} < 980$ فاصله اطمینان

تقریب مناسب نیست زیرا از $Z = \frac{x_0 - \bar{x}}{\sigma} = \frac{985-950}{15} = \frac{35}{15} = 2.3$

$Z = 2.3$ بزرگتر است و از فاصله اطمینان خارج است.

پس آزمون رد می‌شود.

توزیع t: وقتی σ و μ جامعه نامعلوم هستند از این توزیع استفاده می‌کنیم.
وقتی $n < 30$ باشد از توزیع t استفاده می‌کنیم.

توزیع χ^2: وقتی مورد استفاده قرار می‌گیرد که متغیرهای توزیع نرمال را به توان ۲ برسانیم و جمع بزنیم.

درجه آزادی n می‌باشد. $\chi^2_n = N_1^2 + N_2^2 + N_3^2 + \dots + N_n^2$

برای واریانس‌ها استفاده می‌شود.

توزیع F و t برای فاصله اطمینان و آزمون فرض میانگین مورد استفاده قرار می‌گیرند.
در حالیکه توزیع χ^2 برای آزمون فرض و فاصله اطمینان واریانس جامعه مورد استفاده قرار می‌گیرد.
برای مقایسه واریانس ۲ جامعه از توزیع χ^2 استفاده می‌شود.

$\mu = \bar{x} \pm t_{\frac{\alpha}{2}}\, s_{\bar{x}}$, $\frac{\bar{x}-\mu}{s_{\bar{x}}}$ | $\mu \xrightarrow[\text{n نمونه}]{\text{انتخاب}} \bar{X}$ و $s_{\bar{x}}$

s_x انحراف معیار نمونه
σ_x انحراف معیار جامعه

$\sigma^2 \to \chi^2_{n-1}$ برای مقایسه یک واریانس

توزیع F: یک دنباله و مثبت بوده ولی دارای ۲ درجه آزادی است.

برای مقایسه دو واریانس $F = \frac{\chi^2 / d.f}{\chi^2 / d.f}$

آزمون کلی رگرسیون می باشد برای بررسی اینکه آیا تخمین مدل رگرسیون در حالت کلی درست است باید از توزیع F استفاده می شود همچنین این توزیع با t^2 رابطه مستقیم دارد. توزیع F در آنالیز واریانس نیز مورد استفاده قرار می گیرد که دوباره با t^2 رابطه مستقیم پیدا می کند.

توزیع t برای معنی دار بودن تک تک آماره ها مورد استفاده قرار می گیرد.

آنالیز واریانس:

$H_0: \beta_1 = \beta_2 = \dots = \beta_K = 0$
$H_1:$ حداقل یکی از β ها مخالف صفر است

منبع تغییرات		درجه آزادی	میانگین مجذورات
بوسیله رگرسیون	$b^2 \sum (x_i - \bar{x})^2$	k 1	$F = \frac{b^2 \sum (x_i - \bar{x})^2 / 1}{\sum e_i^2 / (n-2)}$
باقیمانده	$\sum e_i^2 = \sum (y_i - \bar{y})^2 - b^2 \sum (x_i - \bar{x})^2$	$n-2$	
کل	$\sum (y_i - \bar{y})^2$	$n-1$ $n-k-1$	

Subject: ۳
Year. Month. Date. ( )

اگر F محاسباتی از F جدول بزرگتر باشد رگرسیون رد شده (H0 رد شده) و آزمون معنی دار است

اگر H0 پذیرفته شود یعنی آزمون رگرسیون معنی دار نمی باشد.

$$\begin{cases} \beta_1 = 1 \\ \beta_1 \neq 1 \end{cases} \qquad t = \frac{b-\beta}{s_b} \qquad \beta = b \pm t_{\alpha/2}\, s_b$$

$$\begin{cases} \alpha = 0 \\ \alpha \neq 0 \end{cases} \qquad t = \frac{a-\alpha}{s_a} \qquad \alpha = a \pm t_{\alpha/2}\, s_a$$

مثال: برای مثال قبل s_a و s_b و σ^2 را محاسبه کنید. مثال صفحه ۳

$$s_{xy}^2 = \frac{\sum e_i^2}{n-2} = \frac{\sum (y_i-\bar{y})^2 - b^2 \sum (x_i-\bar{x})^2}{n-2} = \frac{\sum y_i^2 - n\bar{y}^2 - b^2(\sum x_i^2 - n\bar{x}^2)}{n-2} = \frac{20.39}{8}$$

$$s_{xy}^2 = 2.5 = \sigma^2$$ واریانس جمعیت

$$s_b^2 = \frac{\sigma^2}{\mathrm{var}\,x} = \frac{\sigma^2}{\sum (x_i-\bar{x})^2} = \frac{2.5}{46} = 0.055$$

$$s_a^2 = \sigma^2 \left[\frac{1}{n} + \frac{\bar{x}^2}{\sum (x_i-\bar{x})^2}\right] = 2.5 \left[\frac{1}{10} + \frac{36}{46}\right] = 2.25$$

$$r^2 = b^2 \frac{\sum (x_i-\bar{x})^2}{\sum (y_i-\bar{y})^2} = \frac{135}{154} = 0.87$$ یعنی ۸۷٪ رگرسیون معنی دار است.

$$F = \frac{135}{\frac{20.39}{8}} = \frac{b^2 \sum (x_i-\bar{x})^2}{\frac{\sum e_i^2}{n-2}} = 53$$

H0 رد می شود یعنی رگرسیون معنی دار است. جدول $F_{1,8,0.05} = 5.32$

PAPCO
۶/

Subject:
Year. Month. Date. ( )

$$\begin{cases} H_0: \beta=1 \\ H_1: \beta\neq 1 \end{cases} \quad b=۱,۷۷ \quad t=\frac{۱,۷-۱}{S_b}=\frac{b-\beta}{S_b}=\frac{۰/۷}{۰/۲۳۵}=۳$$

$t_{\%۵,n}$ جدول $=۲$

آزمون فرض رد می‌شود یعنی $\beta\neq 1$ است.

فاصله اطمینان $\beta: b\pm t_{\alpha/2}\cdot S_b=۱,۷\pm ۲\cdot ۰/۲۳ \rightarrow ۱,۲۴<\beta<۲,۱۷$

$$\begin{cases} H_0: \alpha=0 \\ H_1: \alpha\neq 0 \end{cases} \quad \alpha: a\pm t_{\alpha/2}\cdot S_a=۷,۷\pm \underbrace{۲,۷\cdot ۱,۵}_{۳}$$

$$۴,۷<\alpha<۱۰,۷$$

$$t=\frac{a-\alpha}{S_a}=\frac{۷,۷}{۲,۲۵}=۳,۴۲ > t_{جدول}=۲$$

H_0 رد می‌شود یعنی $\alpha\neq 0$

ساختن فاصله اطمینان برای میانگین شرطی y به شرط x:

$$y=۷,۷+۱,۷۷x$$

$$y/x_0 = y/\wedge = ۷,۷+۱,۷(\wedge)=۲۱ \qquad a+bx_0$$

$$y/x_0 = a+bx_0$$

$$\bar{y}_i=a+b\bar{x}_i$$

$$\hat{y}=a+bx_0 \rightarrow \hat{y}=(\bar{y}-b\bar{x})+bx_0=\bar{y}+b(x_0-\bar{x})$$

$$S^2_{y|x_0}=\frac{\sigma^2}{n}+(x_0-\bar{x})^2 S^2_b=\frac{\sigma^2}{n}+\frac{(x_0-\bar{x})^2\sigma^2}{\sum(x_i-\bar{x})^2}=\sigma^2\left[\frac{1}{n}+\frac{(x_0-\bar{x})^2}{\sum(x_i-\bar{x})^2}\right]$$

$$\bar{y}=\frac{\sum y_i}{n} \longrightarrow S_{\bar{y}}=\frac{\sigma^2}{n}$$

$$S^2_{y|x_0}=۲,۵\left[\frac{1}{۱۰}+\frac{(\wedge-۷)^2}{۴۲}\right]=۰/۴\wedge$$

$$\hat{y}|x_0=a+bx_0\pm ۲(۰/۴\wedge)=۲۱\pm ۲(۰/۴\wedge)$$

جلسه ۱۵ و ۱۰ و ۱۲ فصل سوم

Subject:
Year. Month. Date.

نکته: کمترین مقدار واریانس شرطی $Y|x$ وقتی است که

عوامل مؤثر بر حدود اطمینان:

۱. حجم نمونه: هرچه حجم نمونه بیشتر باشد فاصله محدودتر و دقت بیشتر می‌شود.
واریانس کمتر می‌شود.

۲. هرچه α کمتر شود فاصله محدودتر و دقت بیشتر باشد.

۳. هرچه x_0 به $\bar{x}$ نزدیکتر باشد فاصله محدود و دقت بیشتر می‌شود.

۴. هرچه واریانس داده‌ها (پراکندگی داده‌های x) بیشتر باشد فاصله محدودتر و دقت بیشتر است.

نکته: چنانچه یک متغیر توضیحی به رگرسیون اضافه گردد مثلاً در مدل زیر

$$\underbrace{Y}_{\text{مصرف}} = \alpha + \beta_1 \underbrace{X_1}_{\text{درآمد}} + \beta_2 X_2 + \beta_3 X_3 + [U_i] \; X_4$$

باعث می‌شود مقداری از پسماندها کم شده و به رگرسیون اضافه شود این حالت باعث افزایش R^2 می‌شود.

$$\bar{R}^2 \quad R^2 = 1 - \frac{\sum e_i^2 / n-k}{\sum (Y_i - \bar{Y})^2 / n-1} = \frac{b^2 \sum (x_i - \bar{x})^2}{\sum (Y - \bar{Y})^2}$$

آزمون R^2 و $\bar{R}^2$ صرفاً برای اضافه کردن یک متغیر جدید است.

PAPCO

ماتریس:

$$Y_i = \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + \cdots + \beta_K X_{Ki} + U_i$$

$$Y_1 = \beta_1 + \beta_2 X_{21} + \beta_3 X_{31} + \cdots + \beta_K X_{K1} + U_1$$
$$Y_2 = \beta_1 + \beta_2 X_{22} + \beta_3 X_{32} + \cdots + \beta_K X_{K2} + U_2$$
$$\vdots$$
$$Y_n = \beta_1 + \beta_2 X_{2n} + \beta_3 X_{3n} + \cdots + \beta_K X_{Kn} + U_n$$

$$\underset{Y}{\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix}_{n\times 1}} = \begin{bmatrix} 1 & X_{21} & X_{31} & \cdots & X_{K1} \\ 1 & X_{22} & X_{32} & \cdots & X_{K2} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & X_{2n} & X_{3n} & \cdots & X_{Kn} \end{bmatrix}_{n\times K} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_K \end{bmatrix}_{K\times 1} + \begin{bmatrix} U_1 \\ U_2 \\ \vdots \\ U_n \end{bmatrix}_{n\times 1}$$

$$Y_{n\times 1} = X_{n\times K}\, \beta_{K\times 1} + U_{n\times 1}$$

فروض کلاسیک:

$X'U = 0$

X ها از U ها مستقل هستند.

$$X'Y = X'X\beta + \cancel{X'U} \longrightarrow \underline{X'Y = (X'X)\beta}$$

در تمام محاسبات اقتصادسنجی از این فرمول استفاده گردد زیرا نیازی به معکوس کردن ندارد.

$$\beta = (X'X)^{-1} X'Y$$

$$\begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & & x_n \end{bmatrix} \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix} = \begin{bmatrix} n & \sum x_i \\ \sum x_i & \sum x_i^2 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_n \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} \sum y_i \\ \sum x_i y_i \end{bmatrix}$$

$$\begin{bmatrix} n & \sum x_i \\ \sum x_i & \sum x_i^2 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} \sum y_i \\ \sum x_i y_i \end{bmatrix} \longrightarrow \begin{cases} na + b\sum x_i = \sum y_i \\ a\sum x_i + b\sum x_i^2 = \sum x_i y_i \end{cases}$$ معادلات نرمال

• روش کرامر برای بدست آوردن a و b

$$\begin{bmatrix} ۱۰ & ۶۰ \\ ۶۰ & ۴۰۲ \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} ۱۸۰ \\ ۱۱۵۹ \end{bmatrix}$$

$$a = \frac{\begin{vmatrix} ۱۸۰ & ۶۰ \\ ۱۱۵۹ & ۴۰۲ \end{vmatrix}}{\begin{vmatrix} ۱۰ & ۶۰ \\ ۶۰ & ۴۰۲ \end{vmatrix}} \qquad b = \frac{\begin{vmatrix} ۱۰ & ۱۸۰ \\ ۶۰ & ۱۱۵۹ \end{vmatrix}}{\begin{vmatrix} ۱۰ & ۶۰ \\ ۶۰ & ۴۰۲ \end{vmatrix}}$$

$$\begin{bmatrix} n & \sum x_{i1} & \sum x_{i2} \\ \sum x_{i1} & \sum x_{i1}^2 & \sum x_{i1}x_{i2} \\ \sum x_{i2} & \sum x_{i1}x_{i2} & \sum x_{i2}^2 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} \sum y \\ \sum x_{i1}y \\ \sum x_{i2}y \end{bmatrix}$$

اثبات واریانس $E(b)$ و $Var(b)$

$$Y = X\beta + U$$

$$b = (X'X)^{-1}X'Y \longrightarrow b = (X'X)^{-1}X'(X\beta + U) = \underbrace{(X'X)^{-1}(X'X)}_{I}\beta + (X'X)^{-1}X'U$$

$$E(b) = E(\beta) + (X'X)^{-1}X'E(U) \longrightarrow \boxed{E(b) = \beta}$$

$$b - \beta = (X'X)^{-1}X'U$$

$$Var(b) = E(b-\beta)(b-\beta)' = E\left[(X'X)^{-1}X'UU'X(X'X)^{-1}\right] =$$

$$\underbrace{(X'X)^{-1}(X'X)}_{I}(X'X)^{-1}E(UU') = (X'X)^{-1}\sigma^2$$

Subject: ۹۰/۱۲/۱۸
Year Month Date

* بهترین برآورد نقطه‌ای $\bar{x}$

* بهترین برآورد فاصله‌ای $\bar{x} \pm \overset{t_{\alpha/2}}{z_{\alpha/2}} \cdot \sigma$ (هرجا z را ندادند ۲ می‌گیریم.)

$$\bar{x} = \frac{\sum x_i}{n} \qquad \sigma_x^2 = \frac{\sum (x_i - \bar{x})^2}{n-1}$$

$$\operatorname{cov} = E(xy) - E(x) \cdot E(y)$$

$$\operatorname{cov} = E(x_i - \bar{x})(y_i - \bar{y})$$

$$y_i = b_1 + b_2 x_{2i} + b_3 x_{3i} + e_i$$

$$\begin{cases} \sum y_i = \sum b_1 + \sum b_2 x_{2i} + \sum b_3 x_{3i} + \cancel{\sum e_i} \\ \sum y_i x_2 = \sum b_1 x_2 + \sum b_2 x_2^2 + \sum b_3 x_2 x_3 + \cancel{\sum e_i x_2} \\ \sum y_i x_3 = \sum b_1 x_3 + \sum b_2 x_2 x_3 + \sum b_3 x_3^2 + \cancel{\sum e_i x_3} \end{cases}$$ معادلات نرمال

$$\underbrace{\begin{bmatrix} \sum y_i \\ \sum y_i x_2 \\ \sum y_i x_3 \end{bmatrix}}_{\text{معلوم}} = \underbrace{\begin{bmatrix} n & \sum x_2 & \sum x_3 \\ \sum x_2 & \sum x_2^2 & \sum x_2 x_3 \\ \sum x_3 & \sum x_2 x_3 & \sum x_3^2 \end{bmatrix}}_{\text{معلوم}} \underbrace{\begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}}_{\text{مجهول}}$$

اگر دستگاه معادلات نرمال را بنویسیم با کمک ماتریس حل می‌کنیم از این روش استفاده می‌کنیم.

PAPCO

Year.   Month.   Date.   ( )

$e_i = y_i - \hat{y}$

عددهای بدست آمده را تفسیر کنید

b_1 عرض از مبدا

y_i	x_{1i}	x_{2i}	x_{1i}^2	x_{2i}^2	$x_{1i}x_{2i}$	$y_i x_{1i}$	$y_i x_{2i}$
۲۲	۸	۶	۶۴	۳۶	۴۸	۱۷۶	۱۳۲
۲۳	۱۰	۷	۱۰۰	۴۹	۷۰	۲۳۰	۱۶۱
۱۸	۷	۵	۴۹	۲۵	۳۵	۱۲۶	۹۰
۹	۲	۲	۴	۴	۴	۱۸	۱۸
۱۴	۴	۳	۱۶	۹	۱۲	۵۶	۴۲
۲۰	۶	۴	۳۶	۱۶	۲۴	۱۲۰	۸۰
۲۱	۷	۴	۴۹	۱۶	۲۸	۱۴۷	۸۴
۱۸	۶	۳	۳۶	۹	۱۸	۱۰۸	۵۴
۱۶	۴	۳	۱۶	۹	۱۲	۶۴	۴۸
$x_i =$ ۱۹	۶	۳	۳۶	۹	۱۸	۱۱۴	۵۷
۱۸۰	۶۰	۴۰	۴۰۶	۱۸۲	۲۶۹	۱۱۵۹	۷۶

$$\begin{bmatrix} ۱۸۰ \\ ۱۱۵۹ \\ ۷۶ \end{bmatrix} = \underbrace{\begin{bmatrix} ۱۰ & ۹۰ & ۴۰ \\ ۶۰ & ۴۰۶ & ۲۶۹ \\ ۴۰ & ۲۶۹ & ۱۸۲ \end{bmatrix}}_{A} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

روش کرامر

$$b_1 = \frac{\begin{vmatrix} ۱۸۰ & ۶۰ & ۴۰ \\ ۱۱۵۹ & ۴۰۶ & ۲۶۹ \\ ۷۶ & ۲۶۹ & ۱۸۲ \end{vmatrix}}{\text{دترمینان } A} = \frac{۸۲۵۴۰}{۱۷۱۰}$$

$b_1 = ۷٫۹$

$b_2 = ۲٫۳۶$

$b_3 = -۱٫۲$

$$b_2 = \frac{\begin{vmatrix} ۱۰ & ۱۸۰ & ۴۰ \\ ۶۰ & ۱۱۵۹ & ۲۶۹ \\ ۴۰ & ۷۶ & ۱۸۲ \end{vmatrix}}{\text{دترمینان } A} = \frac{\quad}{۱۷۱۰}$$

$$b_3 = \frac{\begin{vmatrix} ۱۰ & ۶۰ & ۱۸۰ \\ ۶۰ & ۴۰۶ & ۱۱۵۹ \\ ۴۰ & ۲۶۹ & ۷۶ \end{vmatrix}}{\text{دترمینان } A} = \frac{\quad}{۱۷۱۰}$$

$\text{دترمینان } A = ۱۰(۱۵۳۱) - ۶۰(۱۶۰) + ۴۰(-۱۰۰) = ۱۷۱۰$

PAPCO

۹/

Subject:
Year Month Date ( )

$\overset{\text{مصرف}}{y} = ۷٫۹ + ۲٫۳۶\overset{\text{درآمد}}{x_2} - ۱٫۰۲\overset{\text{تعداد اعضا}}{x_3} + e_i$

جمله با تئوری اقتصاد سازگار است.

اگر شخص درآمد نداشته باشد و خانواده‌ای نداشته باشد برای زنده ماندن مصرف او ۷٫۹ است.

با داشتن انحراف معیارهای b_1 و b_2 و b_3 و آزمون فرض می‌توان معنی‌دار بودن آنها را بدست آورد.

اگر یک واحد به درآمد اضافه شود ۲٫۳۶ واحد به مصرف اضافه می‌شود.
اگر یک نفر به اعضا اضافه شود ۱٫۰۲ واحد از مصرف کم می‌شود که این با تئوری اقتصاد مغایر است.

PAPCO

$$Y = \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \cdots + \beta_K X_K + U_i$$

$$n \to Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \cdots + \beta_K X_{Ki} + U_i$$ (X_1 ها همه یک هستند)

$$\begin{cases} Y_1 = \beta_1 + \beta_2 X_{21} + \beta_3 X_{31} + \cdots + \beta_K X_{K1} + U_1 \\ Y_2 = \beta_1 + \beta_2 X_{22} + \beta_3 X_{32} + \cdots + \beta_K X_{K2} + U_2 \\ \vdots \\ Y_n = \beta_1 + \beta_2 X_{2n} + \beta_3 X_{3n} + \cdots + \beta_K X_{Kn} + U_n \end{cases}$$

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix}_{n\times 1} = \begin{bmatrix} 1 & X_{21} & X_{31} & \cdots & X_{K1} \\ 1 & X_{22} & X_{32} & \cdots & X_{K2} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & X_{2n} & X_{3n} & \cdots & X_{Kn} \end{bmatrix}_{n\times K} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{bmatrix}_{K\times 1} + \begin{bmatrix} U_1 \\ U_2 \\ \vdots \\ U_n \end{bmatrix}_{n\times 1}$$

$$Y = X_{n\times K}\, \beta_{K\times 1} + U_{n\times 1}$$

فرض اول کلاسیک:

$$E(U) = E\begin{bmatrix} U_1 \\ \vdots \\ U_n \end{bmatrix} = \begin{bmatrix} E(U_1) \\ E(U_2) \\ \vdots \\ E(U_n) \end{bmatrix} = 0 \Rightarrow \begin{bmatrix} E(U_1) = 0 \\ E(U_2) = 0 \\ \vdots \\ E(U_n) = 0 \end{bmatrix}$$

فرض دوم کلاسیک:

$$Y = X\beta + U$$

$$\sum X_i e_i = 0 \xrightarrow{\text{برداری } U \text{ و } \beta} X'U = 0$$

PAPCO

۱۵

X' ماتریس ترانهاده X است:

$$\begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ x_{21} & x_{22} & x_{23} & \cdots & x_{2n} \\ x_{31} & x_{32} & x_{33} & \cdots & x_{3n} \\ \vdots & \vdots & & & \\ x_{k1} & x_{k2} & x_{k3} & \cdots & x_{kn} \end{bmatrix}$$

$$X'Y = X'X\beta \qquad \Rightarrow X'Y = X'X\beta \Rightarrow \beta = (X'X)^{-1}X'Y$$

روش بدست آوردن β

$$(X'X)^{-1}X'Y = \underbrace{(X'X)^{-1}(X'X)}_{I}\beta = \beta$$

$$(X'X)^{-1}X'Y = \beta \qquad \text{مهم}$$

مثال:

Y	x
5	3
6	6
7	9
8	5

$$(X'X)\beta = X'Y$$

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 3 & 6 & 9 & 5 \end{bmatrix} \begin{bmatrix} 1 & 3 \\ 1 & 6 \\ 1 & 9 \\ 1 & 5 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 3 & 6 & 9 & 5 \end{bmatrix} \begin{bmatrix} 5 \\ 6 \\ 7 \\ 8 \end{bmatrix}$$

$$\begin{bmatrix} 4 & 23 \\ 23 & 151 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} = \begin{bmatrix} 26 \\ 154 \end{bmatrix} \qquad \begin{cases} 4\beta_1 + 23\beta_2 = 26 \\ 23\beta_1 + 151\beta_2 = 154 \end{cases}$$

$$\beta_1 = \frac{\begin{vmatrix} 26 & 23 \\ 154 & 151 \end{vmatrix}}{\begin{vmatrix} 4 & 23 \\ 23 & 151 \end{vmatrix}} = \frac{384}{75} = 5/12$$

$$\beta_2 = \frac{\begin{vmatrix} 4 & 26 \\ 23 & 154 \end{vmatrix}}{\begin{vmatrix} 4 & 23 \\ 23 & 151 \end{vmatrix}} = \frac{18}{75} = 0/24$$

P4PCO

مثال 8

y	x_2	x_3
5	3	2
6	6	5
7	9	3
8	5	4

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 3 & 6 & 9 & 5 \\ 2 & 5 & 3 & 4 \end{bmatrix}\begin{bmatrix} 1 & 3 & 2 \\ 1 & 6 & 5 \\ 1 & 9 & 3 \\ 1 & 5 & 4 \end{bmatrix}\begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 3 & 6 & 9 & 5 \\ 2 & 5 & 3 & 4 \end{bmatrix}\begin{bmatrix} 5 \\ 6 \\ 7 \\ 8 \end{bmatrix}$$

$$\begin{bmatrix} 4 & 23 & 14 \\ 23 & 151 & 83 \\ 14 & 83 & 54 \end{bmatrix}\begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} = \begin{bmatrix} 26 \\ 154 \\ 128 \end{bmatrix}$$

$$\beta_1 = \frac{\begin{vmatrix} 26 & 23 & 14 \\ 154 & 151 & 83 \\ 128 & 83 & 54 \end{vmatrix}}{\begin{vmatrix} 4 & 23 & 14 \\ 23 & 151 & 83 \\ 14 & 83 & 54 \end{vmatrix}} = \frac{-13870}{350} = -39.6$$

$$\beta_2 = \frac{\begin{vmatrix} 4 & 26 & 14 \\ 23 & 154 & 83 \\ 14 & 128 & 54 \end{vmatrix}}{\begin{vmatrix} 4 & 23 & 14 \\ 23 & 151 & 83 \\ 14 & 83 & 54 \end{vmatrix}} = \frac{-680}{350} = -1.94$$

$$\beta_3 = \frac{\begin{vmatrix} 4 & 23 & 26 \\ 23 & 151 & 154 \\ 14 & 83 & 128 \end{vmatrix}}{\begin{vmatrix} 4 & 23 & 14 \\ 23 & 151 & 83 \\ 14 & 83 & 54 \end{vmatrix}} = \frac{5740}{350} = 16.37$$

PAPCO_

11

فروض کلاسیک ماتریسی:

۱) $E(U)=0$

۲) $X'U=0$

۳) $E(UU')=0 \Rightarrow E\begin{bmatrix}u_1\\u_2\\ \vdots\\u_n\end{bmatrix}_{n\times1}\begin{bmatrix}u_1 & \cdots & u_n\end{bmatrix}_{1\times n} = E\begin{bmatrix}u_1^2 & u_1u_2 & \cdots & u_1u_n\\ u_2u_1 & u_2^2 & \cdots & \\ \vdots & & & \\ & & & u_n^2\end{bmatrix}$

$$=\begin{bmatrix}E(u_1^2) & E(u_1u_2) & \cdots & \\ E(u_2u_1) & E(u_2^2) & \cdots & \\ & & & \\ & & & E(u_n^2)\end{bmatrix}=\begin{bmatrix}\sigma^2 & 0 & 0 & \cdots\\ 0 & \sigma^2 & 0 & \cdots\\ 0 & 0 & \sigma^2 & \cdots\\ 0 & \cdots & \cdots & \sigma^2\end{bmatrix}=\sigma^2\begin{bmatrix}1 & 0 & \cdots\\ 0 & 1 & \cdots\\ \vdots & \vdots & \ddots\end{bmatrix}$$

$$=\sigma^2 I=\sigma^2$$

فرض اضافی در روش ماتریسی این است که ماتریس X دارای k بردار ستونی مستقل باشد

زیرا در بدست آوردن $\beta=(X'X)^{-1}X'Y$ دترمینان ماتریس $X'X$ صفر میشود

یعنی حتماً باید $|X'X|\neq 0$

در غیر اینصورت با هم خطی متغیرهای X روبرو هستیم یعنی $X_2=f(X_3)$

y	x_1	x_2
3	3	5
1	1	4
8	5	6
3	2	4
5	4	6

$$\begin{bmatrix} \sum y_i \\ \sum y_i x_1 \\ \sum y_i x_2 \end{bmatrix} = \begin{bmatrix} n & \sum x_1 & \sum x_2 \\ \sum x_1 & \sum x_1^2 & \sum x_1 x_2 \\ \sum x_2 & \sum x_1 x_2 & \sum x_2^2 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix}$$

$$\begin{bmatrix} 20 \\ 76 \\ 109 \end{bmatrix} = \begin{bmatrix} 5 & 15 & 25 \\ 15 & 55 & 81 \\ 25 & 81 & 129 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix}$$

$$\beta_0 = \frac{\begin{vmatrix} 20 & 15 & 25 \\ 76 & 55 & 81 \\ 109 & 81 & 129 \end{vmatrix}}{\begin{vmatrix} 5 & 15 & 25 \\ 15 & 55 & 81 \\ 25 & 81 & 129 \end{vmatrix}} = 4$$

$$\beta_1 = \frac{\begin{vmatrix} 5 & 20 & 25 \\ 15 & 76 & 81 \\ 25 & 109 & 129 \end{vmatrix}}{\begin{vmatrix} 5 & 15 & 25 \\ 15 & 55 & 81 \\ 25 & 81 & 125 \end{vmatrix}} = 2.5$$

$$\beta_2 = \frac{\begin{vmatrix} 5 & 15 & 20 \\ 15 & 55 & 76 \\ 25 & 81 & 109 \end{vmatrix}}{\begin{vmatrix} 5 & 15 & 25 \\ 15 & 55 & 81 \\ 25 & 81 & 125 \end{vmatrix}} = -1.5$$

$$\hat{y} = 4 + 2.5 x_1 - 1.5 x_2$$

خاصیت اول: باید امید برآوردکننده ها با پارامترهای جامعه برابر باشد $E(b)=\beta$

جامعه $Y = X\beta + U$

تخمین نمونه $Y = Xb + e$

$$b = (X'X)^{-1}X'Y = (X'X)^{-1}X'(X\beta+U) = \underbrace{(X'X)^{-1}(X'X)}_{I}\beta + (X'X)^{-1}X'U$$

$$b = \beta + (X'X)^{-1}X'U \longrightarrow E(b) = E(\beta) + E(\underbrace{(X'X)^{-1}X'U}_{0}) \longrightarrow E(b)=\beta$$

$$\hookrightarrow b - \beta = (X'X)^{-1}X'U$$

خاصیت دوم: $\operatorname{var} b = ?$

$$\operatorname{var} b = E(b-\beta)(b-\beta)' = E[(X'X)^{-1}X'UU'X(X'X)^{-1}] =$$

$$\underbrace{E(UU')}_{\text{فرض دوم کلاسیک}}\ (X'X)^{-1}\underbrace{(X'X)(X'X)^{-1}}_{I} = (X'X)^{-1}\sigma^2 \ ?$$

$$\operatorname{var} b = \frac{\sigma^2}{\sum (x_i-\bar{x})^2}$$

دو متغیره

ت همچنین زننده های خاه در بین برآوردکننده های خطی بهترین هستند به این معنی که $E(b)=\beta$

و واریانس آن کمترین می باشد.

PAPCO

دو متغیره $\sigma^2_{y\cdot x} = \dfrac{\sum e_i^2}{n-2}$

چند متغیره $\sigma^2_y = \dfrac{ee'}{n-k}$

$$y = xb + e$$
$$e = y - xb$$
$$e'e = (y-xb)'(y-xb) = y'y - 2b'x'y + b'x'xb \xrightarrow{b=(x'x)^{-1}x'y}$$

$$e'e = y'y - 2b'x'y + b'\underbrace{x'x(x'x)^{-1}}_{I}x'y = y'y - 2b'x'y + b'x'y$$

$$e'e = y'y - b'x'y$$

رگرسیون شده ، کل تغییرات ، بیان نشده با رگرسیون شده

$$y'y = e'e + b'x'y$$

کل تغییرات ، رگرسیون نشده ، رگرسیون شده

$$R^2 = \frac{\text{رگرسیون شده}}{\text{کل}} = \frac{b'x'y}{y'y} = 1 - \frac{e'e}{y'y}$$

$$F = \frac{\frac{b'x'y}{k-1}}{\frac{e'e}{n-k}}$$

اثبات $y = xb + e$

$$y' = b'x' + e' \qquad y'y = e'e + b'x'y$$
$$y'y = b'x'xb + e'e + 0 + 0$$

تا اینجا درس

مثال صفحه ۱۲ را بدون در نظر گرفتن عرض از مبدأ حل کنید.

$$\hat{y} = 2.5x_2 - 1.5x_3$$

$$b = \frac{\sum (x_i-\bar{x})(y_i-\bar{y})}{\sum (x_i-\bar{x})^2} = \frac{\sum x_i y_i}{\sum x_i^2}$$

اعداد کوچک یعنی تفاوت اعداد بزرگ را از میانگین بدست آوردیم

PAPCO

۱۳

$$\begin{bmatrix} \sum x_{2i}^2 & \sum x_{2i}x_{3i} \\ \sum x_{2i}x_{3i} & \sum x_{3i}^2 \end{bmatrix} \begin{bmatrix} \beta_2 \\ \beta_3 \end{bmatrix} = \begin{bmatrix} \sum x_{2i}y_i \\ \sum x_{3i}y_i \end{bmatrix}$$

وقتی مدلی بدون عرض از مبدأ می‌خواهیم یعنی $y-\bar{y}$ را بدست آوردیم:

$$Y = \beta_1 + \beta_2 X_2 + \beta_3 X_3$$

$$\bar{Y} = \beta_1 + \beta_2 \bar{X}_2 + \beta_3 \bar{X}_3$$

$$Y - \bar{Y} = \beta_1 + \beta_2 X_2 + \beta_3 X_3 - \beta_1 - \beta_2 \bar{X}_2 - \beta_3 \bar{X}_3 \Rightarrow y = \beta_2 x_2 + \beta_3 x_3$$

$$\sum x_i^2 = \sum (X_2 - \bar{X}_2)^2 = \sum X_2^2 - n\bar{X}^2$$

$$\sum x_2 y = \sum (X_2 - \bar{X})(Y - \bar{Y}) = \sum X_2 Y - n\bar{X}_2\bar{Y}$$

$$\begin{bmatrix} ۵۵٫۴۵ & ۸۱٫۷۵ \\ ۸۱٫۷۵ & ۱۲۹٫۱۲۵ \end{bmatrix} \begin{bmatrix} \beta_2 \\ \beta_3 \end{bmatrix} = \begin{bmatrix} ۷۴٫۱۰ \\ ۱۹٫۱۰ \end{bmatrix} \Rightarrow \begin{bmatrix} ۱۰ & ۶ \\ ۶ & ۴ \end{bmatrix} \begin{bmatrix} \beta_2 \\ \beta_3 \end{bmatrix} = \begin{bmatrix} ۱۶ \\ ۹ \end{bmatrix}$$

$$\beta_2 = \frac{\begin{vmatrix} ۱۶ & ۶ \\ ۹ & ۴ \end{vmatrix}}{\begin{vmatrix} ۱۰ & ۶ \\ ۶ & ۴ \end{vmatrix}} = \frac{۶۴-۵۴}{۴۰-۳۶} = \frac{۱۰}{۴} = ۲٫۵$$

$$\beta_3 = \frac{\begin{vmatrix} ۱۰ & ۱۶ \\ ۶ & ۹ \end{vmatrix}}{\begin{vmatrix} ۱۰ & ۶ \\ ۶ & ۴ \end{vmatrix}} = \frac{۹۰-۹۶}{۴۰-۳۶} = \frac{-۶}{۴} = -۱٫۵$$

X_2	$(X_2 - \bar{X}_2)$	$\sum x_2^2$	X_3	x_3	x_3^2	$x_2 x_3$	Y	y
۳	۰	۰	۵	۰	۰	۰	۳	
۱	-۲	۴	۴	-۱	۱	۲	۱	
۵	۲	۴	۶	۱	۱	۲	۸	
۲	-۱	۱	۴	-۱	۱	۱	۳	
۴	+۱	۱	۶	+۱	۱	۱	۵	
$\bar{X}_2 = ۳$		۱۰	$\bar{X}_3 = ۵$		۴ $\sum x_3^2$	۶	$\bar{Y} = ۴$	

PAPCO

فصل ۵ سئوال ۲۵ و ۱۰

$$R^2 = \frac{b'x'y}{y'y} = \frac{\overbrace{\begin{bmatrix} 4.5 & -1.5 \end{bmatrix}}^{b'} \overbrace{\begin{bmatrix} 12 \\ 9 \end{bmatrix}}^{x'y}}{28} = \frac{26.5}{28} \longrightarrow \begin{array}{l} y'y = e'e + b'x'y \\ 28 = 1.5 + 26.5 \end{array}$$

$$F = \frac{\frac{26.5}{3-1}}{\frac{1.5}{5-3}} = \frac{26.5}{1.5}$$

$$\sigma_y^2 = \frac{e'e}{n-k} = \frac{1.5}{2}$$

$$\operatorname{var} b = \sigma^2 (x'x)^{-1} = 0.75 \times \frac{1}{4} \begin{bmatrix} 4 & -6 \\ -6 & 10 \end{bmatrix} = \begin{bmatrix} \frac{4}{0.75} & \frac{-6}{0.16} \\ \frac{-6}{0.16} & \frac{10}{0.16} \end{bmatrix}$$

$var(b_2)$; $cov(b_2, b_3)$; $var(b_3)$

PAPCO

۱۴

مدل ساده یک مدل خطی است.

خطی بودن این روش به این معنا است که مدل نسبت به پارامترهای α و β خطی می باشد نه نسبت به متغیرها.

$$\underbrace{Y}_{\text{متغیر وابسته}} = \underbrace{\alpha}_{\text{پارامتر عرض از مبدا}} + \underbrace{\beta}_{\text{پارامتر شیب}} \underbrace{X}_{\text{متغیر مستقل}}$$

این مدل هم نسبت به پارامترها و هم نسبت به متغیرها خطی است.

$$Y = \alpha + \beta_1 X + \beta_2 X^2$$

فقط نسبت به پارامترها خطی است.

$$Y = \alpha^2 + \beta X$$

کاملاً غیرخطی

خطی نسبت به پارامترها می گوییم نه نسبت به متغیرها

مثال:

Y	X	X'
2	2	4
4	1	1
5	0	0
6	3	9
7	2	4

$$\begin{bmatrix} n & \Sigma x \\ \Sigma x & \Sigma x^2 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} \Sigma y \\ \Sigma xy \end{bmatrix}$$

$$\begin{bmatrix} 5 & 8 \\ 8 & 18 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} 25 \\ 42 \end{bmatrix}$$

$$\alpha = \frac{\begin{vmatrix} 25 & 8 \\ 42 & 18 \end{vmatrix}}{5 \times 18 - 64} = \frac{114}{26} = 4.38$$

$$\beta = \frac{\begin{vmatrix} 5 & 25 \\ 8 & 42 \end{vmatrix}}{90 - 64 = 26} = \frac{10}{26} = 0.38$$

$$Y = \alpha + \beta X + \delta X^2$$

برای بدست آوردن δ و β و α در معادله درجه ۲ فرض می کنیم $X'_2 = X^2$ باشد و مدل $Y = \alpha + \beta X + \delta X'$ را تخمین بزنید.

Subject.
Year. Month. Date.

$$\begin{bmatrix} n & \Sigma x & \Sigma x^2 \\ \Sigma x & \Sigma x^2 & \Sigma x\cdot x^2 \\ \Sigma x^2 & \Sigma x\cdot x^2 & \Sigma x^4 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \\ \delta \end{bmatrix} = \begin{bmatrix} \Sigma y \\ \Sigma y x \\ \Sigma y x^2 \end{bmatrix}$$

برازش منحنی با استفاده از $y = \alpha + \beta x + \delta \frac{1}{x}$ را تخمین بزنید.

$x = 0 \longrightarrow y = 2$

اگر $x = 0$ باشد نمی‌توان مقدار را تخمین زد.

$$\begin{bmatrix} 5 & 8 & 18 \\ 8 & 18 & 44 \\ 18 & 44 & 114 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \\ \delta \end{bmatrix} = \begin{bmatrix} 25 \\ 42 \\ 98 \end{bmatrix}$$

$$\alpha = \frac{\begin{vmatrix} 25 & 8 & 18 \\ 42 & 18 & 44 \\ 98 & 44 & 114 \end{vmatrix}}{\begin{vmatrix} 5 & 8 & 18 \\ 8 & 18 & 44 \\ 18 & 44 & 114 \end{vmatrix}} = \frac{604}{124} = 4{,}87$$

$$\beta = \frac{\begin{vmatrix} 5 & 25 & 18 \\ 8 & 42 & 44 \\ 18 & 98 & 114 \end{vmatrix}}{124} = \frac{-116}{124} = -0{,}93$$

$$\delta = \frac{\begin{vmatrix} 5 & 8 & 25 \\ 8 & 18 & 42 \\ 18 & 44 & 98 \end{vmatrix}}{124} = \frac{56}{124} = 0{,}45$$

$$\underline{y = 4{,}87 - 0{,}93x + 0{,}45x^2}$$

PAPCO

۱۵

Subject.
Year. Month. Date.

1) $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2$

2) $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_1^2$ فرض : $X_3 = X_1^2$

3) $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 \frac{1}{X_2}$ فرض : $X_3 = \frac{1}{X_2}$

4) $\underbrace{\log Y}_{Y'} = \beta_0 + \beta_1 \underbrace{\log X_1}_{X'_1} + \beta_2 \underbrace{\log X_2}_{X'_2}$

با جانشین کردن $Y' = \log Y$ و $X'_1 = \log X_1$ و $X'_2 = \log X_2$

متغیرهای غیرخطی را به خطی تبدیل می کنیم.

5) $Y = \alpha + \beta \log X_1$

Y	$X'_1 = \log X_1$
2	$\log 2 = 0{,}3$
4	0
5	0,3
6	0,47
7	0,3

$$\begin{bmatrix} 5 & 1{,}37 \\ 1{,}37 & 0{,}49 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} 25 \\ 7{,}02 \end{bmatrix}$$

$$\alpha = \frac{2{,}6326}{0/5731} = 4{,}59$$

$$Y = 4{,}59 + 1{,}48 \log X_1$$

$$\beta = \frac{0/85}{0/5731} = 1{,}48$$

معادلات ذاتاً خطی :

$$Y = \alpha X_1^{\beta} X_2^{\delta} \xrightarrow{\text{تبدیل به خطی می کنیم}} \log Y = \log \alpha + \beta \log X_1 + \delta \log X_2$$

چون $\log \alpha$ داریم غیرخطی است پس فرض می کنیم $\log \alpha = \alpha'$ و داریم :

α و α' به آنتی لوگ همدیگر هستند

$\log Y = \alpha' + \beta \log X_1 + \delta \log X_2$ $\alpha = \text{antilog}\, \alpha'$

PAPCO

مثال ۳: این معادله را تخمین بزنید: $y = \alpha x_1^{\beta}$

y	x_1	y'	x'
۳	۲	۰/۴۷	۰/۳
۴	۱	۰/۶	۰
۵	۲	۰/۶۹	۰/۳
۶	۳	۰/۷۷	۰/۴۷
۷	۲	۰/۸۴	۰/۳

$$\log y = \underbrace{\log \alpha}_{\alpha'} + \beta \log x_1 \longrightarrow \underbrace{\log y}_{y'} = \alpha' + \beta \underbrace{\log x}_{x'}$$

$$y' = \alpha' + \beta x'$$

$$\begin{bmatrix} 5 & 1.37 \\ 1.37 & 0.49 \end{bmatrix} \begin{bmatrix} \alpha' \\ \beta \end{bmatrix} = \begin{bmatrix} 3.37 \\ 1.84 \end{bmatrix}$$

$$\alpha' = \frac{-0.8695}{0.5731} = -1.51 \qquad \alpha' = \log \alpha \longrightarrow -1.51 = \log \alpha \longrightarrow \alpha = 10^{-1.51}$$

$$\beta = \frac{4.5831}{0.5731} = 7.9 \qquad y = 10^{-1.51} x_1^{7.9}$$

ذاتاً خطی

$$y = \alpha x_1^{\beta} x_2^{\delta} u$$

$$\log y = \log \alpha + \beta \log x_1 + \delta \log x_2 + \log u$$

$$\log y = \alpha' + \beta \log x_1 + \delta \log x_2 + \log u$$

ذاتاً غیرخطی

$$y = \alpha x_1^{\beta} x_2^{\delta} + u$$

نمی‌توان این مدل را تبدیل به خطی کرد

مثال: تبدیل به خطی کنید.

$$y = e^{\beta_0 + \beta_1 x_1 + u} \longrightarrow \ln y = \beta_0 + \beta_1 x_1 + u$$

$$y = \frac{1}{\beta_0 + \beta_1 x_1 + \beta_2 x_2 + u} \longrightarrow \frac{1}{y} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + u$$

PAPCO

۱۲

سؤال ۷ ص ۱۲۹

$x_i y_i = \beta_0 + \beta_1 x_i$

$y_i = \beta_0 \frac{1}{x_i} + \beta_1 \longrightarrow y = \beta_0 x' + \beta_1$

$x' = \frac{1}{x_i}$

نقض فرض x_i ها مستقل خطی است.

$Var(b) = \sigma^2 (x'x)^{-1}$

$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2$

وابستگی خطی بین x_1 و x_2 هست.

$$\begin{bmatrix} n & \sum x_1 & \sum x_2 \\ \sum x_1 & \sum x_1^2 & \sum x_1 x_2 \\ \sum x_2 & \sum x_1 x_2 & \sum x_2^2 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix} = \begin{bmatrix} \sum y \\ \sum x_1 y \\ \sum x_2 y \end{bmatrix}$$

فرض کنید که $x_2 = cx_1$

$$x'x = \begin{bmatrix} n & \sum x_1 & \sum cx_1 \\ \sum x_1 & \sum x_1^2 & \sum cx_1^2 \\ \sum cx_1 & \sum cx_1^2 & \sum c^2 x_1^2 \end{bmatrix} \xrightarrow{R_3 - cR_2} \begin{bmatrix} n & \sum x_1 & \sum cx_1 \\ \sum x_1 & \sum x_1^2 & \sum cx_1^2 \\ 0 & 0 & 0 \end{bmatrix}$$

چون دترمینان این ماتریس صفر است پس نمی‌توان از این روابط معادلات نرمال مجهولات را
چون سطر دوم و سوم یکی است.
بدست آورد.

$x'x\beta = x'y \longrightarrow \beta = (x'x)^{-1} x'y$

$x_2 = cx_1 \longrightarrow |x'x| = 0 \longrightarrow$ نمی‌توان مجهولات را بدست آورد.

$\longrightarrow (x'x)^{-1}$ قابل محاسبه نمی‌باشد.

مثال: $x_2 = 2x_1$

$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2$

$Y = \beta_0 + \beta_1 x_1 + \beta_2 (2x_1)$

$Y = \beta_0 + \underbrace{(\beta_1 + 2\beta_2)}_{\gamma} x_1$

PAPCO

Y	X_1	X_2
۳	۲	۴
۴	۱	۲
۵	۳	۶

$$\begin{bmatrix} 3 & 6 & 12 \\ 6 & 14 & 28 \\ 12 & 28 & 56 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix} = \begin{bmatrix} 12 \\ 25 \\ 50 \end{bmatrix}$$

دترمینان این ماتریس صفر است چون
سطر سوم ۲ برابر سطر دوم است در نتیجه نمی‌توان
از این روابط مجهولات را بدست آورد پس نتیجه می‌گیریم متغیرها مشکل هم خطی دارند.

$X_2 = 2X_1 \qquad Y = \beta_0 + \beta_1 X_1 + 2\beta_2 (X_1) = \beta_0 + \underbrace{(\beta_1 + 2\beta_2)}_{\delta} X_1$

$Y = \beta_0 + \delta X_1$

$$\begin{bmatrix} 3 & 6 \\ 6 & 14 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \delta \end{bmatrix} = \begin{bmatrix} 12 \\ 25 \end{bmatrix} \qquad Y = 3 + 0.5 X_1 \qquad \beta_1 + 2\beta_2 = 0.5$$

$$\beta_0 = \frac{12 \times 14 - 25 \times 6}{3 \times 14 - 36} = \frac{18}{6} = 3 \qquad \delta = \frac{3 \times 25 - 12 \times 6}{3 \times 14 - 36} = \frac{3}{6} = \frac{1}{2}$$

مشکلات هم خطی

$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 \quad ; \quad x_3 = 2x_1 + 3x_2$

$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 (2x_1 + 3x_2) = \beta_0 + (\beta_1 + 2\beta_3) x_1 + (\beta_2 + 3\beta_3) x_2$

$Y = \beta_0 + \underbrace{(\beta_1 + 2\beta_3)}_{\delta} x_1 + \underbrace{(\beta_2 + 3\beta_3)}_{\gamma} x_2 = \beta_0 + \delta x_1 + \gamma x_2$

نکته: گاهی اوقات دترمینان $X'X$ صفر نیست ولی خیلی به صفر نزدیک است در نتیجه مشکل
که پیش می‌آید این است که در رابطه $\sigma_b = \sigma^2 (X'X)^{-1}$ بزرگ تخمین می‌خورند.

چنانچه بین متغیرهای X هم خطی کامل وجود داشته باشد ضرایب و واریانس غیرقابل محاسبه

Subject.
Year. Month. Date. ( )

می باشد و در جهت رفع این مشکل باید یکی از متغیرها حذف شود ولی چنانچه

ولی چنانچه هم خطی ناقص باشد و شدت هم خطی زیاد در این حالت دترمینان ماتریس $X'X$

نزدیک به صفر بوده و اعداد معکوس ماتریس $(X'X)^{-1}$ اعداد نسبتاً بزرگی می باشند دترمینان کوچک

و معکوس می

واریانس β ها که $\sigma_b^2 = \sigma^2 (X'X)^{-1}$

در آزمون فرض $t = \frac{\beta}{Var(\beta)}$

* فرض کنید یک مدل تخمین می خورد آزمون F معنی دار و R^2 عدد بالایی را نشان می دهد ولی پارامترها

به صورت تک به تک معنی دار نمی باشد این مشکل هم خطی بین X های می باشد.

* رابطه خطی با R^2 مشخص می شود.

Δ حل دوم:
در این حالت که X_1 و X_2 رگرسیون می شوند اگر R^2 معادله $X_1 = \hat{\alpha}_1 + \hat{\alpha}_2 X_2$ زیاد و بزرگ بود مشخص

می شود که بین X_1 و X_2 یک رابطه هم خطی وجود دارد.

PAPCO

Subject:
Year: Month: Date:

مثال:

X_1	X_2	X_3
1	0	1
1	0	1
1	0	1
1	0	1
0	1	0
0	1	0
0	1	0
0	1	0

$X_3 = X_1 + X_2$

R ضریب همبستگی $R_{x_2x_3} = -\frac{1}{3}$

R^2 ضریب تعیین $R_{x_1x_2} = -0.109$

$R^2 = \frac{cov(xy)}{\sigma_x \sigma_y}$

راههای از بین بردن هم خطی:

۱) ترکیب داده ها: فرضاً اگر مواقع جمعیت کشورها با تولید ناخالص ملیشان رابطه هم خطی وجود دارد.

اگر بخواهیم یک تابع مصرف به شکل $C = \beta_0 + \beta_1 GDP + \beta_2 N$ تخمین بزنیم باید دو متغیر

را با هم ترکیب کرد: $C = \alpha_0 + \alpha_1$ تولید سرانه $\qquad$ تولید سرانه $= \frac{GDP}{N}$

۲) روش دوم: حذف متغیرها مثلاً N را از معادله حذف کنیم:
$C = \alpha_0 + \alpha_1 GDP$

۳) روش سوم: تخمین β_2 از روش دو مرحله ای
فرض می کنید β_2 را از تابع دیگری مثل بیرون اندازه $(1-\beta_2)N = $ قابل محاسبه

باشد سپس در معادله $C - \beta_2 N = \beta_0 + \beta_1 GDP$ جایگذاری می کنیم و داریم:

$C' = \beta_0 + \beta_1 GDP$

۴) روش چهارم: روی قطر ماتریس $X'X$ یک مقدار عددی فرضی (۰٫۱) اضافه نمایید

۳	۲
۴	۲٫۵

افزودن روی قطر ماتریس را تغییر می دهیم. $7.5 - 8 = -0.5$
$14 - 8 = 6$
در صورت هم خطی ناقص

۱۸

Subject.
Year. Month. Date.

(۵) روش پنجم: برای حل مشکل هم‌خطی می‌توان اطلاعات آماری را بازیاد کرد و یا آنها را تغییر داد. مثلاً اگر داده‌ها زیاد بود اطلاعات را فصلی و یا ماهانه وارد کنید.

مسأله ۳: صفحه ۲۰۳

$$C = 4153 + 0.93\,Y_t - 0.16\,Y_{t-1}$$
$$\quad (308) \qquad (0.11) \qquad (0.12)$$

اعداد داخل پرانتز انحراف معیار ضرایب هستند.

$$R_{Y_t, Y_{t-1}} = 0.996$$

۰٫۱۶- مورد تائید نمی‌باشد زیرا عدد منفی است و با تئوری اقتصادی سازگاری ندارد.
زیرا اگر درآمد زیاد شود مصرف هم زیاد می‌شود.

برای رفع مشکل هم‌خطی باید ۰٫۱۶- حذف گردد.

$$t = \frac{\alpha}{S_\alpha} = \frac{4153}{308} = 13$$

قدر مطلق این اعداد باید از ۲ بزرگتر باشد.

$$t = \frac{\beta}{S_\beta} = \frac{0.93}{0.11} = 8$$

پس ۰٫۱۶- را حذف می‌کنیم.

$$t = \frac{\gamma}{S_\gamma} = \frac{-0.16}{0.12} \simeq -1.3$$

4PCO

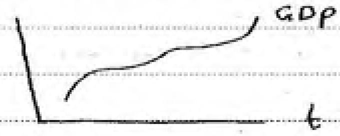
Subject: ۹۱/۲/
Year. Month. Date.

«همبستگی»

۱. علت چیست؟
۲. اثرات آن چیست؟
۳. چگونه تشخیص دهیم؟

هرگاه $E(U_i U_j) \neq 0$ خودهمبستگی — در این حالت روش OLS دیگر بهترین نیست.

علل وقوع خودهمبستگی:



۱. سری زمانی بودن متغیرها (طبیعت متغیرها)

۲. حذف یک متغیر (خطای تصریح)

معادله اصلی: $Y_t = \beta_0 + \beta_1 X_{1t} + \underline{\beta_2 X_{2t} + U_t}$

فرض می‌کنیم که می‌خواهیم مدل زیر را برآورد کنیم: $\Rightarrow Y_t = \beta_0 + \beta_1 X_{1t} + V_t$

$V_t = \beta_2 X_{2t} + U_t$

در این صورت اگر X_{2t} خودهمبسته باشد V_t ها دچار خودهمبستگی گردیده و در مجموع معادله (مدل) دچار خودهمبستگی می‌باشد.

۳. فرم نادرست معادله

$$C_t = \alpha + \beta_1 Y_t + \underline{\beta_2 Y_t^2 + U_t}$$

$$\Rightarrow C_t = \alpha + \beta_1 Y_t + V_t$$

در این صورت نیز احتمال ایجاد خودهمبستگی وجود دارد.

PAPCO

۱۹

۴- وقفه زمانی متغیر در مدل (مثل تارعنکبوتی)

مدل تارعنکبوتی
$$\begin{cases} Q_{s_t} = \beta_0 + \beta_1 P_{t-1} + V_t \\ Q_d = \alpha_0 + \alpha_1 P_t + V_t \end{cases}$$
عرضه دچار وقفه زمانی است.

تابع مصرف کینز — $C_t = \alpha_0 + \alpha_1 Y_t + \alpha_2 C_{t-1} + V_t$

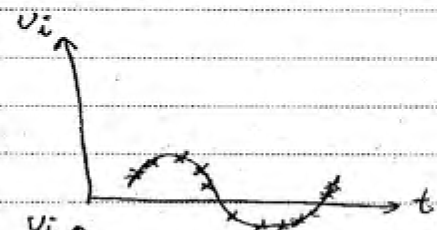
۵- دستکاری اطلاعات آماری : تبدیل آمارهای فصلی به آمارهای ماهانه
تبدیل آمارهای سالانه به آمارهای فصلی یا ماهانه

می‌توانند موجب خود همبستگی شوند.

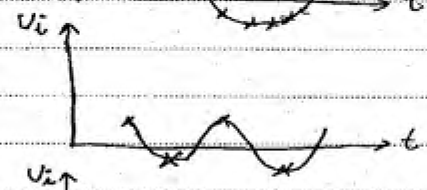
در مدل تحلیل استفاده نمی‌باشد زیرا در این حالت ها ماتریس بالا برده و دیگر نمی‌توان از آزمون‌های DW استفاده کرد.

تشخیص : اینکه یک مدل خود همبسته هست یا خیر؟

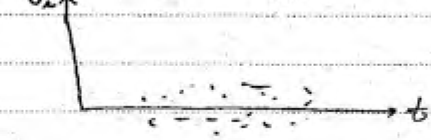
۱- روش هندسی یا



خود همبستگی



خود همبستگی



در این حالت خود همبستگی وجود ندارد
غیر همبسته است.

$$Y_i = \alpha + \beta X_i + U_i$$
$$\hat{Y}_i = a + b X_i \longrightarrow U_i = Y_i - \hat{Y}_i$$

آزمون دوربین واتسون : D.W

$$Y_i = \alpha + \beta X_i + U_i$$

فرض می‌کنیم یک همبستگی بین اجزاء اخلال U_i وجود داشته باشد.

U_i متوالی‌اند $\quad U_i = \rho U_{i-1} + \varepsilon_i$

$|\rho| < 1$ ضریب همبستگی

$$\varepsilon_i \begin{cases} E(\varepsilon_i) = 0 \\ E(\varepsilon_i \varepsilon_j) = 0 \\ E(\varepsilon_i^2) = \sigma^2 \end{cases}$$

$$\rho = \frac{E(U_i U_{i-1})}{\sqrt{E(U_i^2) E(U_{i-1}^2)}} = \frac{\sum U_i U_{i-1}}{\sqrt{\sum U_i^2 \sum U_{i-1}^2}} \simeq \frac{\sum e_i e_{i-1}}{\sqrt{\sum e_i^2 \sum e_{i-1}^2}}$$

$$d = \frac{\sum (e_i - e_{i-1})^2}{\sum e_i^2} = \frac{\sum e_i^2 - 2\sum e_i e_{i-1} + \sum e_{i-1}^2}{\sum e_i^2} \longrightarrow \sum e_i^2 \simeq \sum e_{i-1}^2$$

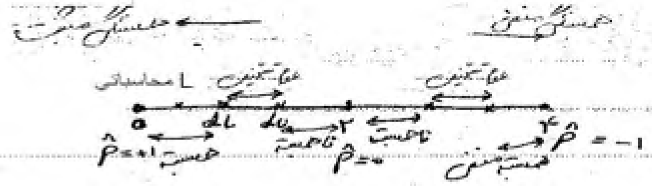
$$d \simeq \frac{2\sum e_i^2 - 2\sum e_i e_{i-1}}{\sum e_i^2} = 2 - 2\frac{\sum e_i e_{i-1}}{\sum e_i^2}$$

$\sum e_i^2 = \sum e_{i-1}^2$ با فرض $\quad \rho \simeq \dfrac{\sum e_i e_{i-1}}{\sqrt{\sum e_i^2 \sum e_{i-1}^2}} \Rightarrow \hat{\rho} = \dfrac{\sum e_i e_{i-1}}{\sum e_i^2}$

$d = 2 - 2\hat{\rho} \longrightarrow d = 2(1 - \hat{\rho})$ خیلی مهم

$$\begin{cases} \hat{\rho} = 0 \longrightarrow d = 2 \longrightarrow \text{عدم خود همبستگی} \\ \hat{\rho} = +1 \longrightarrow d = 0 \longrightarrow \text{همبستگی کامل مثبت} \\ \hat{\rho} = -1 \longrightarrow d = 4 \longrightarrow \text{همبستگی کامل منفی} \end{cases}$$

۲۰



$d < d_L$ ← همبستگی منفی

$d_L < d < d_U$ ← عدم تشخیص ? مثال‌های صفحه ۴۲

$d > d_U$ ← عدم همبستگی

چگونگی از بین بردن خود همبستگی:

$$Y_i = \alpha + \beta X_i + U_i$$

$$U_i = \rho U_{i-1} + \varepsilon_i \longrightarrow \varepsilon_i = U_i - \rho U_{i-1}$$

$$Y_{i-1} = \alpha + \beta X_{i-1} + U_{i-1}$$

$$\times\rho \quad \rho Y_{i-1} = \alpha\rho + \beta\rho X_{i-1} + \rho U_{i-1}$$

$$Y_i - \rho Y_{i-1} = \alpha - \alpha\rho + \beta X_i - \beta\rho X_{i-1} + U_i - \rho U_{i-1}$$

$$Y_i - \rho Y_{i-1} = \alpha(1-\rho) + \beta(X_i - \rho X_{i-1}) + (U_i - \rho U_{i-1})$$

این مدل تمام خصوصیات نرمال را دارا می‌باشد. $Y^* = \alpha^* + \beta X_i^* + \varepsilon_i$

نکات مهم در مورد این روش:

۱. یک مشاهده را به خاطر وقفه از دست می‌دهیم.

۲. دو روش وجود دارد که $\hat{\rho}$ را بدست آوریم: روش اول (کوکران-اورکات)

ρ وجود ندارد پس در نتیجه باید ρ را پیدا کرد.

PAPCO

۲ تفاضل معادله اولیه $Y_i = \alpha + \beta X_i + U_i$ را با OLS برآورد می‌کنیم سپس با استفاده از پسماند مقدار ρ را برآورد می‌کنیم.

روش دوم : «دوربین»

$$Y_i - \rho Y_{i-1} = \alpha(1-\rho) + \beta(X_i - \rho X_{i-1}) + \varepsilon_i$$

$$Y_i = \underline{\alpha(1-\rho)} + \beta\, X_i + \beta\rho X_{i-1} + \rho Y_{i-1} + \varepsilon_i$$

$$Y_i = f(Y_{i-1}, X_i, X_{i-1})$$

Subject: ۹۱/۲/۲۴
Year: Month: Date: ( )

معادلات همزمان ۳

از یک معادله بیشتر است و باید از طریق سیستم دستگاه حل کنیم.

Q	P
۱۰	۲۰
۸۰	۲۵
۳۰	۱۵

$$0/5\,(Q_t^d = \alpha_0 + \alpha_1 P_t + U_t)$$

$$0/5\,(Q_t^s = \beta_0 + \beta_1 P_t + V_t)$$ معادله بالا معادله همزمان است

Q و P هر دو درون زا هستند.

$$Q = 0/5\,Q_t^d + 0/5\,Q_t^s = 0/5\,(\alpha_0+\beta_0) + 0/5\,(\alpha_1+\beta_1)P_t + 0/5\,(U_t+V_t)$$

معادلات عرضه و تقاضا غیر قابل تشخیص هستند چون عرضه و تقاضا هر دو شبیه به هم شکل است

$$C_t = \bar{C} + \alpha Y_t + U_t$$

دو معادلات ساختاری متغیرهای درون زا و برون زا بدون نظم خاصی چیده می‌شوند.

$$\begin{cases} C_t = \alpha_0 + \alpha_1 Y_t + U_t \\ Y = C + I \end{cases}$$

و طبق فروض کلاسیک باید مستقل باشد

معادلات ساختاری

Y_t توسط C_t مشخص می‌شود و در نتیجه Y مستقل نیست

بنابراین C_t و Y معادلات همزمان هستند.

تنها متغیر برون زا $I=\bar{I}$ می‌باشد که مستقل است $\longrightarrow$ $I = \bar{I}$

$$C_t = \alpha_0 + \alpha_1 (C_t + I) + U_t \longrightarrow C_t - \alpha_1 C_t = \alpha_0 + \alpha_1 I + U_t$$

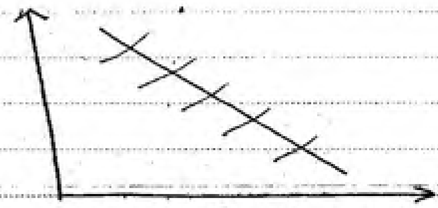
$$\longrightarrow (1-\alpha_1) C_t = \alpha_0 + \alpha_1 I + U_t$$

$$\xrightarrow{\text{فرم حل شده}} \left\{ C_t = \frac{\alpha_0}{1-\alpha_1} + \frac{\alpha_1}{1-\alpha_1} I + \frac{1}{1-\alpha_1} U_t \right\}$$

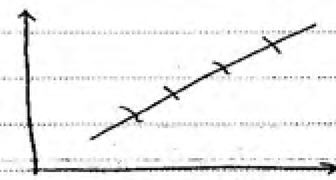
$$Y = \frac{c_0 + \bar{I}}{1-c} = \frac{\alpha_0 + \bar{I}}{1-\alpha_1} = \frac{\alpha_0}{1-\alpha_1} + \frac{\bar{I}}{1-\alpha_1} \pm \frac{1}{1-\alpha_1} u_t$$

حل خودم $\quad Y = C_t + I = \frac{\alpha_0}{1-\alpha_1} + \underbrace{\frac{\alpha_1}{1-\alpha_1}\bar{I} + \bar{I}}_{\frac{\bar{I}}{1-\alpha_1}} + \frac{1}{1-\alpha_1} u_t$

فرم حل شده: متغیرهای درون زا بر حسب متغیرهای برون زا چیده شده باشند و در طرف راست پارامترها و متغیرهای برون زا باشند.



تقاضا قابل شناسایی است.



عرضه قابل شناسایی است.

۳

۲ –

$$Q = 0.5(\alpha_0 + \beta_0) + 0.5(\alpha_1 + \beta_1)P_t + 0.5\,\alpha_2 Y_t$$

$$\begin{cases} 0.5(Q_t^d = \alpha_0 + \alpha_1 P_t + \alpha_2 Y_t) \\ 0.5(Q_t^s = \beta_0 + \beta_1 P_t) \end{cases}$$

تابع فوق متفاوت از تابع تقاضا می‌تواند باشد و می‌توان تابع عرضه را شناسایی کرد. مانند شکل ۲

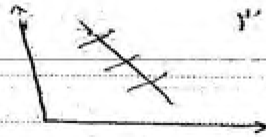
برای تشخیص معادلات هم زمان یک میانگین از معادلات می‌گیریم اگر معادله جدید شبیه معادلات عرضه یا تقاضا بود در این حالت آن معادله شبیه قابل تشخیص نبوده و معادله‌ای که شبیه معادله میانگین نباشد قابل تشخیص است.

PAPCO

۲۲

$$0.5\ (Q_t^d = \alpha_0 - \alpha_1 P_t)$$

$$0.5\ (Q_t^s = \beta_0 + \beta_1 P_t + \beta_2 W)$$ ۳

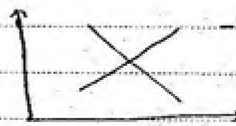


$$Q = 0.5(\alpha_0 + \beta_0) + 0.5(\alpha_1 + \beta_1)P_t + 0.5\beta_2 W$$

تابع تقاضا قابل تشخیص وعرضه غیرقابل تشخیص

$$0.5(Q_t^d = \alpha_0 + \alpha_1 P_t + \alpha_2 Y)$$ ۴-

$$0.5(Q_t^s = \beta_0 + \beta_1 P_t + \beta_2 W)$$



$$Q = 0.5(\alpha_0 + \beta_0) + 0.5(\alpha_1 + \beta_1)P_t + 0.5\beta_2 W + 0.5\alpha_2 Y$$

درحالت چهارم معادله میانگین شبیه هیچ کدام نبوده بنابراین Q^d و Q^s هردوقابل تشخیص است.

معادلات با وقفه زمانی :

$$C_t = \alpha_0 + \alpha_1 Y_t + \alpha_2 Y_{t-1} + \alpha_3 Y_{t-2} + \alpha_4 Y_{t-3}$$

مثال : تابع مصرف مودیگلیانی $\sum C_t = \sum Y_t$

یعنی فرد هرچقدر درآمد دارد همه آن را مصرف میکند وهیچ چیزی برای ارث باقی نمی‌گذارد.

Subject
Date

$$C_t = f(Y_t, Y_{t-1}, Y_{t-2}, \ldots)$$

مصرف وابسته است به درآمد امسال و سالهای گذشته.

هرچه زمان دورتر و دورتر می رود αها کوچکتر می شود.

مشکلات الگوهای با وقفه زمانی :

۱- همبستگی نسبتاً شدید بین Yها (هم خطی بین Yها)

۲- ایجاد مشکل خود همبستگی

۳- از دست دادن درجه آزادی

C	Y_t	Y_{t-1}	Y_{t-2}	Y_{t-3}	—	Y_{t-n}
۱۰۰	۸۰	—	—	—		—
	۹۰	۱۰	—	—		—
	۱۰۰	۱۰	۲۰	—		—

راه حل مشکل مدل وقفه زمانی :

① مدل کویک

$$C_t = \alpha_1 Y_t + \alpha_2 Y_{t-1} + \alpha_3 Y_{t-3}$$

یک تصاعد هندسی تشکیل می دهیم

$$C_t = \alpha_1 Y_t + \alpha_1 \lambda Y_{t-1} + \alpha_1 \lambda^2 Y_{t-2} + \alpha_1 \lambda^3 Y_{t-3}$$

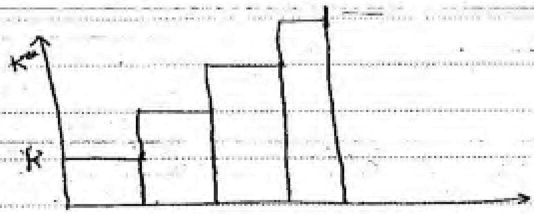
PAPCO

۴۳

* فرض انتظارات تطبیقی $\qquad K_t - K_{t-1} = \lambda (K^* - K_t)$

تطبیق دهیم خودمان را با

انتظارات را تمامان برابر با ضریب λ



* انتظارات عقلایی بیان می‌کند روندها را با یک بار انجام دهند.

$$C_t = \frac{\alpha_1}{1-\lambda} y_t$$

کویک فرض می‌کند هر سال با یک روندی خاص ضرایب کوچک و کوچکتر می‌شود.

$$C_t = \beta_0 y_t + \beta_1 y_{t-1} + \beta_2 y_{t-2} + \cdots + \beta_k y_{t-k}$$

تبدیل کویک

$$\begin{cases} \beta_1 = \lambda \beta_0 \\ \beta_2 = \lambda \beta_1 = \lambda(\lambda \beta_0) = \lambda^2 \beta_0 \\ \beta_3 = \lambda \beta_2 = \lambda(\lambda^2 \beta_0) = \lambda^3 \beta_0 \end{cases}$$

$$C_t = \beta_0 y_t + \underbrace{\beta_0 \lambda y_{t-1} + \beta_0 \lambda^2 y_{t-2} + \beta_0 \lambda^3 y_{t-3} + \cdots + \lambda^k \beta_0 y}_{}$$

$$C_t = \beta_0 y_t + \lambda C_{t-1}$$

$$\lambda C_{t-1} = \lambda \beta_0 y_{t-1} + \lambda \beta_1 y_{t-2} + \cdots$$

مزایای معادله $C_t = \beta_0 Y_t + \lambda C_{t-1}$

۱- خود همبستگی نیست.

۲- هم خطی ندارد.

۳- درجه آزادی یک دارد.

می توان معادله مصرف کوتاه مدت و بلندمدت را بدست آورد

$$C_t = C_{t-1} \xrightarrow{\text{فرض بلندمدت}} C_t - \lambda C_{t-1} = \beta_0 Y$$

تابع مصرف بلندمدت

$$\longrightarrow (1-\lambda) C_t = \beta_0 Y \longrightarrow C_t = \left(\frac{\beta_0}{1-\lambda}\right) Y$$

$\frac{\beta_0}{1-\lambda}$ میل نهایی به مصرف در بلندمدت β_0 میل نهایی به مصرف در کوتاه مدت

متغیرهای مجازی :

خیلی از متغیرها در اقتصاد متغیر کیفی است مثل: زن، مرد، سرد، گرم، سفید و سیاه، فصول سال، ماه های سال برای استفاده از متغیرهای کیفی از متغیر مجازی یا دامی Dummy variable استفاده می شود. استفاده از این متغیرها این کمک را می کند که به جای تخمین چند معادله یک معادله بیشتر تخمین نخورد.

PAPCO

۲۴

$$\begin{cases} \text{زن} & y = \alpha_1 + \alpha_2 X \\ \text{مرد} & y = \beta_0 + \beta_1 X \end{cases}$$

حال می‌توانیم به جای استفاده از دو معادله از یک معادله استفاده کنیم که هر دو گروه را شامل می‌شود و فرض کنید که تنها اختلاف مربوط به عرض از مبدأ α و β باشد و شیب برابر باشد.

زن $y = \alpha_1 + \beta x$
مرد $y = \alpha_2 + \beta x$

متغیر مجازی

$$y = \alpha_1 + (\alpha_2 - \alpha_1) D + \beta x$$

زن if $D = 0 \longrightarrow y = \alpha_1 + \beta x$
مرد if $D = 1 \longrightarrow y = \alpha_2 + \beta x$

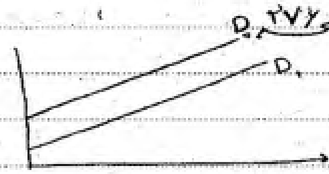
اگر ضریب D بعد از برآورد تابع معنی دار بود تابع تخمین مربوط به مردان می‌باشد در غیر اینصورت مربوط به زنان.

مثال ۲: صفحه ۲۷۶

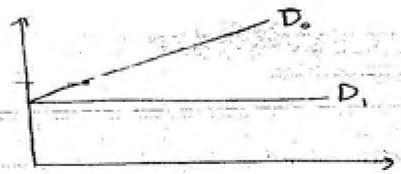
$C = \alpha_1 + (\alpha_2 - \alpha_1) D + \beta Y$

$D = 1$ برای سالهای جنگ

$D = 0$ برای سالهای غیرجنگ

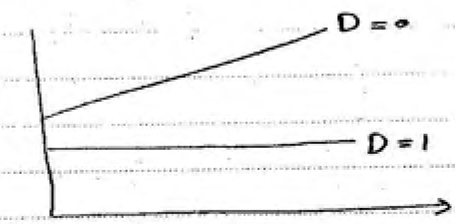


شیب تغییر نکرده و فقط عرض از مبدأ تغییر کرده



شیب تغییر کرده عرض از مبدأ تغییر نکرده

$$y = \alpha + \beta_1 x + (\beta_2 - \beta_1) D x \qquad \begin{array}{l} D=0 \\ D=1 \end{array} \begin{cases} y = \alpha + \beta_1 x \\ y = \alpha + \beta_2 x \end{cases}$$



۲ متغیر ه

$$y = \alpha_1 + (\alpha_2 - \alpha_1) D_1 + \beta_1 x + (\beta_2 - \beta_1) D x$$

هم عرض از مبدأ تغییر کرده و هم شیب

مثال ۴ ص ۲۷۸

$$\begin{cases} y = \alpha_1 + \beta x & \text{زمستان} \\ y = \alpha_2 + \beta x & \text{بهار} \\ y = \alpha_3 + \beta x & \text{تابستان} \\ y = \alpha_4 + \beta x & \text{پاییز} \end{cases}$$

?

$$y = \alpha_1 + (\alpha_2 - \alpha_1) D_1 + (\alpha_3 - \alpha_2) D_2 + (\alpha_4 - \alpha_3) D_3 + \beta x$$

تابستان $\begin{cases} D_1 = 0 \\ D_2 = 1 \\ D_3 = 0 \end{cases}$ $y = \alpha_3 + \beta x$

پاییز $\begin{cases} D_1 = 0 \\ D_2 = 0 \\ D_3 = 1 \end{cases}$ $y = \alpha_4 + \beta x$

زمستان $\begin{cases} D_1 = 0 \\ D_2 = 0 \\ D_3 = 0 \end{cases}$ $y = \alpha_1 + \beta x$

بهار $\begin{cases} D_1 = 1 \\ D_2 = 0 \\ D_3 = 0 \end{cases}$ $y = \alpha_2 + \beta x$

PAPCO

۲۵

۷- همیشه تعداد متغیرهای مجازی یکی کمتر از تقسیم بندی متغیر می باشد زیرا اگر به تعداد تقسیمات متغیر دامی تعریف کنیم دارای صورت هم خطی کامل بین عرض از مبدأ و متغیر مجازی رخ می دهد.

مثال:

c	y	مرد D_1	زن D_2
۱۰۰	۸۰	۱	۰
۷۰	۴۰	۱	۰
۹۰	۸۰	۱	۰
۷۰	۶۰	۰	۱
۳۰	۴۰	۰	۱
۴۰	۵۰	۰	۱

۱	۱۰۰	۱	۰
۱	۷۰	۱	۰
۱	۹۰	۱	۰
۱	۷۰	۰	۱
۱	۳۰	۰	۱
۱	۴۰	۰	۱

جمع می شود

ستون اصلی که همان عرض از مبدأ است.

معادلات هم زمان:

با توجه به مسأله ۱ صفحه ۲۰۳

$$\begin{cases} c = \underset{\pi_{11}}{\frac{a}{1-b}} + \underset{\pi_{12}}{\frac{b}{1-b}}\,\bar{I} \\ y = \underset{\pi_{21}}{\frac{a}{1-b}} + \underset{\pi_{22}}{\frac{1}{1-b}}\,\bar{I} \end{cases}$$

۱. شناسایی معادله

۲. تخمین

دیگر یک رابطه بین π ها و a و b درست آورده‌ایم:

PAPCO

۱- غیر قابل تشخیص

justified — دقیقاً مشخص

۲- قابل تشخیص

over — بیش از حد مشخص شود

روش 2sls

فرم ساختاری را حل می‌کنیم و تخمین می‌زنیم و کاری به ضرایب نداریم و جدا استفاده می‌کنیم

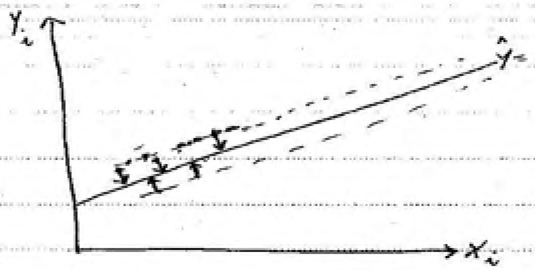
و در روش ۲ مرحله‌ای باید ۲ بار تخمین بزنیم.

همسانی واریانس: فروض کلاسیک

$V(U_i) = \sigma_{U}^2 = \sigma^2$

نشان می‌دهد پراکندگی داده‌ها چقدر است؟

$$\sigma^2 = \frac{\sum (X_i - \bar{X})^2}{n-1}$$



$\sum U_i = 0$

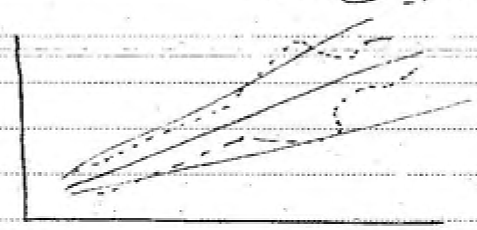
$E(U_i U_j) = 0$

$E(U_i U_i) = E(U_i^2) = \sigma^2$

فرض واریانس همسانی

پراکندگی داده‌ها یکسان می‌باشد.

$\sigma_{U_i}^2 = \sigma_i^2$



ناهمسانی واریانس

در شکل روبرو

$C_{\text{مصرف}} = \alpha + \beta Y_{\text{درآمد}}$

وقتی همسانی واریانس به ناهمسانی واریانس تبدیل می‌شود دیگر روش کمال بهترین نیست یا بلو.

σ^2 از پیش فاکتور می‌گیریم و بقیه را در V_i قرار می‌دهیم.

$V(U_i) = \sigma_{U_i}^2 = \sigma_i^2 = \sigma^2 V_i$

PAPCO

۱ ۲۷

* اثرات واریانس ناهمسانی :

۱. نااریب بودن آماره‌ها از بین می‌رود.

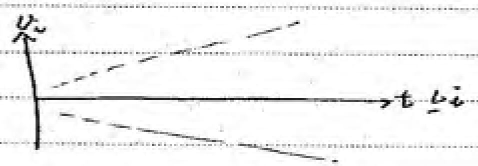
۲. واریانس دیگر کمترین نیست.

* پیش بینی‌ها دچار خطا می‌کند

* فواصل اطمینانی که برای پارامترهای α و β حول خط برازش انجام می‌گیرد با افزایش X_i ها فاصله‌ها بزرگتر می‌شوند. و بزرگ شدن فاصله اطمینان، نااطمینانی ایجاد می‌کند.

راههای تشخیص واریانس ناهمسانی : ص ۲۴۳

روش هندسی : ابتدا مدل را تخمین می‌زنیم و سپس گراف U_i^2 ها را در برابر زمان یا X_i رسم نماییم.



$$Y_i = \alpha + \beta X_i + U_i$$

ص ۲۴۴

روش اول : تقسیم داده‌ها به دو گروه یعنی حجم نمونه را به دو قسمت تقسیم می‌کنیم و سپس واریانس جمله پسماندها را برای ۲ گروه محاسبه می‌کنیم اگر این واریانس‌ها اختلاف شدیدی داشتند نتیجه می‌گیریم واریانس ناهمسانی دارند.

PAPCO

آزمون گولدفلد - کوانت:

داده‌ها را مرتب می‌کنیم و c آماری که در وسط قرار دارند را حذف می‌کنیم و برای هر دو گروه

u_i^2 ها و RSS را حساب می‌کنیم و آزمون F تشکیل می‌دهیم و

این F حتماً از یک بزرگتر است. پس با F جدول مقایسه $F_{df_1, df_2} = \frac{RSS_2}{RSS_1}$

می‌کنیم اگر این F از F جدول بزرگتر باشد ناهمسانی واریانس داریم.

آزمون پارک:

ابتدا تابع اصلی $Y = \alpha + \beta X_i + U_i$ را تخمین بزنید. u_i^2 ها را تشکیل دهید.

پس تابع $\log u_i^2 = \lambda + d \log X_i + Z_i$

را تخمین بزنید.

در صورتی که d معنی دار بود آنگاه ناهمسانی واریانس وجود دارد.

در صورتی که ضریب d اگر معنی دار بود ناهمسانی واریانس می‌باشد.

$$\begin{cases} H_0 : d = 0 \\ H_1 : d \neq 0 \end{cases}$$

در صورتی که d معنی دار است که H_0 رد می‌شود.

اثبات:

$Y = \alpha + \beta X_i + U_i$

$var(U_i) = \sigma_i^2 = \sigma^2 X_i^d e^{Z_i}$

$\log \sigma_i^2 = \log \sigma^2 + d \log X_i + Z_i \longrightarrow \log \sigma_i^2 = \lambda + d \log X_i + Z_i$

PAPCO

۲ ۲۸/

Subject
Date

$$Y_i = \alpha + \beta X_i + U_i$$

$$V(Y_i) = V(\alpha + \beta X_i + U_i)$$

$$V(Y_i) = V(\alpha) + \beta^2 var(X_i) + var(U_i)$$

$$var(U_i) = \sigma^2 U_i = \sigma^2 X_i^2 \quad + X_i^2 \sigma^2$$

$$\frac{Y_i}{X_i} = \frac{\alpha}{X_i} + \beta \frac{X_i}{X_i} + \frac{U_i}{X_i} \qquad \frac{1}{\cancel{X_i^2}} var(U_i) \; (\cancel{X_i^2}\sigma^2) = \sigma^2$$

روش حل ناهمسانی واریانس این است که کل متغیرها Y_i و X_i و U_i را بر $Z_i^{1/2}$ تقسیم کنیم

که Z_i در درون واریانس U_i به شکل $\sigma^2 Z_i = \sigma^2_{U_i}$ وجود دارد

$$Y_i = \alpha + \beta X_i + U_i$$

حالت اول : $\sigma^2_{U_i} = X_i^2 \sigma^2 \xrightarrow{\text{حل}} \frac{Y_i}{X_i} = \frac{\alpha}{X_i} + \beta + \frac{U_i}{X_i}$

حالت دوم : $\sigma^2_{U_i} = X_i \sigma^2 \longrightarrow \frac{Y_i}{X_i^{1/2}} = \frac{\alpha}{X_i^{1/2}} + \beta \frac{X_i}{X_i^{1/2}} + \frac{U_i}{X_i^{1/2}}$

حالت سوم : $\sigma^2_{U_i} = X_i^{1/2} \sigma^2$

واریانس U_i به شکل $\sigma^2 Z_i = \sigma^2_{U_i}$ وجود دارد.

مثال :

$$V(U_i) = \frac{\sigma^2}{X_i}$$

$$Y_i = \alpha + \beta X_i + U_i \Rightarrow Y_i X_i^{1/2} = \alpha X_i^{1/2} + \beta X_i^{3/2} + U_i X_i^{1/2}$$

$$V(U_i X_i^{1/2}) = (X_i^{1/2})^2 \, var(U_i) = \cancel{X_i} \cdot \frac{\sigma^2}{\cancel{X_i}} = \sigma^2$$

PAPCO

آزمون گلسجر:

باید قدر مطلق U ها را بدست آوریم و روی
X ها مثل فرمها تخمین زده و بیشتر t ها که X
معنی دار است . C حالت واریانس است .

$$|U_i| = \alpha + \beta X_i$$

$$|U_i| = \alpha + \beta X_i^{-1}$$

$$|U_i| = \alpha + \beta X^{1/2}$$

$$|U_i| = \alpha + \beta X^{-1/2}$$

اقتصاد سنجی

۵ر۳ر۹۱

معادلات همزمان
- ۱ـ تشخیص
- ۲ـ حل معادله به روش ILS و 2SLS
- ۳ـ نمونه‌ای مثال بزنید

فرم ساختاری ـ

$c = a + by$

$y = c + \bar{I}$

c و y درونزا

$\bar{I}$ برونزا

فرم حل شده

$$c = \pi_{11} + \pi_{12}\bar{I}$$

$$\hat{c} = \frac{a}{1-b} + \frac{b}{1-b}\bar{I}$$

$$\hat{ج} = \frac{a}{1-b} + \frac{1}{1-b}\bar{I}$$

$$Y. = \pi_{12} + \pi_{22}\bar{I}$$

حالات ممکن
- ۱ـ بعد از تخمین فرم حل شده نمی‌توان به فرم ساختاری رسید = غیر قابل تشخیص
- ۲ـ بعد از تخمین فرم حل شده می‌توان به فرم ساختاری رسید = قابل تشخیص
  - ۱. فقط جوابهای منحصر به فرد وجود دارد = Justify دقیقاً قابل تشخیص
  - ۲. بیش از یک جواب دارد = over estimate

$$\begin{cases} Q_t^d = a + bP_t + fY_t \\ Q_t^s = c + dP_t + gW_t \end{cases}$$

$$Q_t = 0.5a + 0.5bP_t + 0.5c + 0.5dP_t$$

$$Q_t = 0.5(a+c) + 0.5(b+d)P_t + 0.5fY_t + 0.5gW_t$$

PAPCO

مثال ۲: تخمین معادلات

C	Y	I	$\hat{C}$	$\hat{Y}$
۱	۲	۱	$\frac{۲۴}{۲۱}$	$\frac{۳۳}{۲۱}$
۲	۳	۲	$\frac{۳۸}{۲۱}$	$\frac{۳۳}{۲۱}$
۳	۴	۱	$\frac{۲۴}{۲۱}$	$\frac{۳۳}{۲۱}$
۴	۳	۳	$\frac{۴۲}{۲۱}$	$\frac{۳۳}{۲۱}$

$$C = \alpha + \beta Y \qquad Y = C + I$$

$$\begin{cases} C = \pi_{11} + \pi_{12} I \\ Y = \pi_{11} + \pi_{12} \bar{I} \end{cases}$$

$$XX'\beta = X'Y$$

$$\begin{bmatrix} ۴ & \Sigma \bar{I} \\ \Sigma \bar{I} & \Sigma \bar{I}^۲ \end{bmatrix} \begin{bmatrix} \pi_{11} \\ \pi_{12} \end{bmatrix} = \begin{bmatrix} \Sigma C \\ \Sigma CI \end{bmatrix} \qquad \begin{bmatrix} N & \Sigma I \\ \Sigma I & \Sigma I^۲ \end{bmatrix} \begin{bmatrix} \pi_{21} \\ \pi_{22} \end{bmatrix} = \begin{bmatrix} \Sigma Y \\ \Sigma IY \end{bmatrix}$$

$$\begin{bmatrix} ۴ & ۷ \\ ۷ & ۱۵ \end{bmatrix} \begin{bmatrix} \pi_{11} \\ \pi_{12} \end{bmatrix} = \begin{bmatrix} ۱۰ \\ ۲۰ \end{bmatrix} \qquad \begin{cases} \hat{C} = \frac{۱۰}{۲۱} + \frac{۱۴}{۲۱}\bar{I} \\ \hat{Y} = \frac{۳۳}{۲۱} + \cancel{\frac{۰}{۲۱}\bar{I}} \end{cases}$$

$$\begin{bmatrix} ۴ & ۷ \\ ۷ & ۱۵ \end{bmatrix} \begin{bmatrix} \pi_{21} \\ \pi_{22} \end{bmatrix} = \begin{bmatrix} ۱۲ \\ ۲۱ \end{bmatrix}$$

روش 2SLS: باید جدول را برای $\hat{C}$ و $\hat{Y}$ ادامه دهیم.

$$Y = \hat{C} + \bar{I}$$

$$C = \alpha + \beta \hat{Y} \longrightarrow \begin{bmatrix} N & \Sigma \hat{Y} \\ \Sigma \hat{Y} & \Sigma \hat{Y}^۲ \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} \Sigma C \\ \Sigma C\hat{Y} \end{bmatrix}$$

$$\pi_{11} = \frac{\alpha}{1-\beta} \Rightarrow \frac{۱۰}{۲۱} = \frac{\alpha}{1-\beta} \longrightarrow ۲۱\alpha + ۱۰\beta = ۱۰ \longrightarrow \alpha = \frac{۲}{۷}$$

$$\pi_{12} = \frac{\beta}{1-\beta} \Rightarrow \frac{۱۴}{۲۱} = \frac{\beta}{1-\beta} \longrightarrow ۱۴ - ۱۴\beta = ۲۱\beta \longrightarrow ۱۴ = ۳۵\beta \longrightarrow \beta = \frac{۱۴}{۳۵}$$

$$\beta = \frac{۲}{۵}$$

$$\frac{۳۳}{۲۱} = \frac{\alpha}{1-\beta} \Rightarrow ۳۳ - ۳۳\beta = ۲۱\alpha \longrightarrow$$

مدل قابل تشخیص است.

اگر بتوان برای α و β جواب بدست آورد مدل قابل تشخیص است در غیر این صورت قابل تشخیص نمی‌باشد.

PAPCO

۲۵۱

Subject
Date

* سوال: شیب برای رفع خودهمبستگی ۲ روش وجود دارد را دوربین واتسون را ۲ کاکرین - اورکات

۱. روش‌های خودهمبستگی - علل و عوامل حقیقی

* مدل ۳ متغیره
مفاهیم واریانس - کوواریانس - ضریب تعیین و ...

* مدل با چه شکلی روبرو است حل کنید:

$$Y = \alpha + \beta x + U_i$$

واریانس ناهمسانی است

$$V(U_i) = \frac{\sigma^2}{X_i^2}$$

واریانس همسان ⟵

$$Y_i x_i = \alpha x_i + \beta x_i^2 + x_i U_i$$

$$V(x_i U_i) = x_i^2 V(U_i) = x_i^2 \cdot \frac{\sigma^2}{x_i^2} = \sigma^2$$

* مثال عددی از وقفه زمانی
سری‌های زمانی

$$C = \bar{C} + cY$$
$$Y = C + I$$

$$C = 200 + 0.9Y$$
$$Y = C + I$$

۰٫۹ ۰٫۸۱ ۰٫۷۲۹

$\Delta C_1 = 90$
$\Delta C_2 = 81$
$\Delta C_3 = 721$

مجموع تصاعد هندسی در صورتی وجود دارد که قدر مطلق ضریب هندسی کوچکتر از یک باشد در این صورت

$$\text{مجموع} = \frac{\text{جمله اول}}{1 - \text{ضریب}} = \frac{90}{1 - 0.91} = 900$$

PAPCO

مدلهای سری زمانی

$$C_t = 0.9\, C_{t-1} + 0.3\, C_{t-2} + U_t$$

مدلهای AR:

$$C_t = \alpha_0 + \alpha_1 C_{t-1} + \alpha_2 C_{t-2} + \alpha_3 C_{t-3} + \cdots$$

$$C_t = \alpha_0 + \alpha_1 C_{t-1} \qquad AR(1)$$

$$C_t = \alpha_0 + \alpha_1 C_{t-1} + \alpha_2 C_{t-2} \qquad AR(2)$$

$$C_t = \alpha + \beta_0 Y_t + \beta_1 Y_{t-1} \qquad MA(1)$$

$$C_t = \alpha_0 + \alpha_1 C_{t-1} + \beta_0 Y_t + \beta_1 Y_{t-1} \qquad ARMA(1,1)$$

خود همبستگی

کدام یک از فروض روش OLS را نقض می‌کند؟

$$\begin{cases} Y = \alpha + \beta X + \underbrace{U_i}_{\text{خود همبستگی}} \\ U_i = \rho U_{i-1} + \underbrace{\varepsilon_i}_{\text{تمام شرایط را دارد}} \end{cases}$$

عدم وجود همبستگی بین U_iها را نقض می‌کند.

بین U_iها همبستگی وجود دارد.

$$\varepsilon_i = U_i - \rho U_{i-1}$$

حل ε را برای Y_i بیان می‌کنیم

$$Y_i - \rho Y_{i-1} = \alpha + \beta(X_i - \rho X_{i-1}) + (U_i - \rho U_{i-1})$$

همه متغیرها را طبق فرم بالا تغییر متغیر می‌دهیم:

$$Y_i - \rho Y_{i-1} = \alpha + \beta(X_i - \rho X_{i-1}) + \varepsilon$$

خود همبستگی از بین رفت.

PAPCO

۳۱