

说明: 闭卷考试。答题时间: 100 分钟。选择题请在试卷顶端答题区域填写答案。问答题请在题目下方对应空白处填写答案。

题号	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
答案																		

题号	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35
答案																	

1 选择题 (每题只有一个正确选项, 每题 2 分, 共 70 分)

如无其他说明, 下列问题涉及的变量由以下模型定义:

$$y_i = \mathbf{x}_i' \boldsymbol{\beta} + \varepsilon_i, \quad (i = 1, \dots, n) \quad (1)$$

其中, $\mathbf{x}_i \equiv (1 \ x_{i2} \dots x_{iK})'$, $\boldsymbol{\beta} \equiv (\beta_1 \ \beta_2 \dots \beta_K)'$. 叠放形式:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon},$$

$\mathbf{y} \equiv (y_1 \ y_2 \dots y_n)'$, 数据矩阵 $\mathbf{X} \equiv (\mathbf{x}_1 \ \mathbf{x}_2 \dots \mathbf{x}_n)'$, $\boldsymbol{\varepsilon} \equiv (\varepsilon_1 \ \varepsilon_2 \dots \varepsilon_n)'$.

1. 假设模型(1)满足

- 1) 线性假定 (linearity)
- 2) 数据矩阵 \mathbf{X} 满列秩, 也即 $\text{rank}(\mathbf{X}) = K$,
- 3) 严格外生性 (strict exogeneity), 也即 $E(\varepsilon_i \mid \mathbf{x}_1, \dots, \mathbf{x}_n) = 0, \forall i = 1, \dots, n$.
- 4) 球形扰动项 (spherical error), 也即 $\text{Var}(\boldsymbol{\varepsilon} \mid \mathbf{X}) = \sigma^2 \mathbf{I}_n$, 其中 \mathbf{I}_n 是 n 阶单位矩阵.

现有 $\boldsymbol{\beta}$ 的 OLS 估计量 \mathbf{b} 和另一估计量 $\hat{\boldsymbol{\beta}}$, 以下哪个结论是正确的:

- A. $\text{Var}(\mathbf{b} \mid \mathbf{X}) - \text{Var}(\hat{\boldsymbol{\beta}} \mid \mathbf{X})$ 是半正定矩阵
- B. $\text{Var}(\hat{\boldsymbol{\beta}} \mid \mathbf{X}) - \text{Var}(\mathbf{b} \mid \mathbf{X})$ 是半正定矩阵
- C. $\text{Var}(\hat{\boldsymbol{\beta}} \mid \mathbf{X}) - \text{Var}(\mathbf{b} \mid \mathbf{X}) = \mathbf{0}$
- D. 条件不足, 无法判断

2. 继续上题模型假设. 此时:

- A. $E(\mathbf{b} \mid \mathbf{X}) = \boldsymbol{\beta}$
- B. $E(\mathbf{b} \mid \mathbf{X}) \neq \boldsymbol{\beta}$
- C. $\text{Var}(\mathbf{b} \mid \mathbf{X}) = \mathbf{0}$
- D. 条件不足, 无法判断

3. 继续上题模型假设. 以下哪个说法是**错误**的?

- A. 无条件期望 $E(\mathbf{b}) = \boldsymbol{\beta}$
- B. $\mathbf{X}'\mathbf{e} = \mathbf{0}$, 其中 $\mathbf{e} \equiv \mathbf{y} - \mathbf{X}\mathbf{b}$
- C. $\hat{\mathbf{y}}'\mathbf{e} = 0, \hat{\mathbf{y}} = \mathbf{X}\mathbf{b}$
- D. $\mathbf{b} \mid \mathbf{X}$ 服从正态分布, $\mathbf{b} \mid \mathbf{X} \sim N(\boldsymbol{\beta}, \sigma^2(\mathbf{X}'\mathbf{X})^{-1})$

4. 假设模型(1)中 ε_i 存在异方差. 这会导致:
- OLS 估计量方差 $\text{Var}(\mathbf{b}|\mathbf{X}) \neq \sigma^2(\mathbf{X}'\mathbf{X})^{-1}$
 - OLS 估计量不一致 (not consistent)
 - OLS 的渐进分布不是正态分布
 - OLS 仍是 Best Linear Unbiased Estimator
5. 回归模型 $\log(y_i) = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \varepsilon_i$, 若 β_1 的估计值是 0.02, 则可以解释为:
- X_1 增加 1 单位, y 增大 0.02 单位
 - X_1 增加 1%, y 增大 0.02 单位
 - X_1 增加 1%, y 增大 2%
 - X_1 增加 1 单位, y 增大 2%
6. 假设线性回归方程中有常数项. 预测值 \hat{y} 的均值 $\bar{\hat{y}} \equiv \frac{1}{n} \sum_{i=1}^n \hat{y}_i$, $\mathbf{M} = \mathbf{I}_n - \mathbf{P}$ 为消灭矩阵, \mathbf{e} 是残差. 以下哪个式子是正确的:
- $\bar{\hat{y}} = 0$
 - $\bar{\hat{y}} = \frac{1}{n} \sum_{i=1}^n e_i$
 - $\mathbf{e} = \mathbf{M}\boldsymbol{\varepsilon}$
 - $\mathbf{M}\mathbf{y} = \mathbf{0}$
7. 令 $\mathbf{P} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$ 为投影矩阵, $\mathbf{M} = \mathbf{I}_n - \mathbf{P}$ 为消灭矩阵, 其中 \mathbf{I}_n 是 n 阶单位矩阵. 以下等式哪个是错误的?
- $\mathbf{M}^4 = \mathbf{M}^2$
 - $\mathbf{P}^2\mathbf{X} = \mathbf{X}$
 - $\mathbf{P}\mathbf{e} = \mathbf{0}$
 - $\mathbf{M}^2\mathbf{X} = \mathbf{X}$
8. 有常数项时, 关于拟合优度 R^2 , 以下哪个说法是错误的:
- $R^2 = [\text{Corr}(y_i, \hat{y}_i)]^2$
 - $R^2 = 1 - \frac{\sum_{i=1}^n e_i^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$
 - 当模型拟合程度很差时, R^2 可能小于 0
 - R^2 不能超过 1
9. 如果在给定的显著性水平 α 下检验原假设 $H_0: \beta_2 = 0$ 时, 关于估计值 $\hat{\beta}_2$ 的 t 统计量 t_2 的值如图 1 所示, 则此时我们根据 t_2 判断:

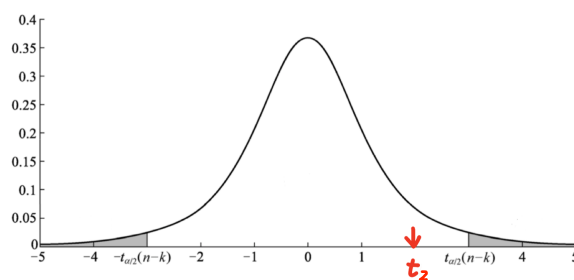


图 1

- 拒绝原假设
- 不能拒绝原假设
- 取决于 α 的具体取值
- 不能确定

10. 在上题中, 如果真实的 $\beta_2 \neq 0$, 则我们根据 t_2 作出的判断:
- 没有犯错误
 - 犯了第 I 类错误 (弃真, False Positive)
 - 犯了第 II 类错误 (存伪, False Negative)
 - 不能确定
11. 假设上题中的 β_2 出现在回归方程 $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon$ 中. 若想根据 $\hat{\beta}_2$ 的 t 值判断 β_2 是否为 0, 并据此决定是否把 x_2 排除在回归方程之外, 以下哪种说法是正确的:
- 应当在 5% 显著性水平下检验
 - 应当在 $<5\%$ 显著性水平下检验, 以减少犯第 II 类错误的概率
 - 应当在 $>5\%$ 显著性水平下检验, 以减少犯第 II 类错误的概率
 - 应当在 $>5\%$ 显著性水平下检验, 以减少犯第 I 类错误的概率
12. 当“被解释变量滞后项”被用作“解释变量”时, 以下哪项经典 OLS 假设不成立?
- 球形扰动项 (spherical disturbance)
 - 不存在“严格多重共线性” (no strict multicollinearity)
 - 线性模型 (linearity)
 - 严格外生性 (strict exogeneity)
13. 如果 $\hat{\beta}_n$ 是 β 的一致估计量 (consistent estimator), 则以下哪个说法是**错误的**:
- $\hat{\beta}_n$ 依分布收敛于 β (Converge in distribution)
 - $\hat{\beta}_n$ 依均方收敛于 β (Converge in mean square)
 - $\hat{\beta}_n$ 依概率收敛于 β (Converge in probability)
 - $\hat{\beta}_n$ 不一定是 β 的无偏估计量
14. 随机过程 $\{x_t\}_{t=1}^{\infty} = \{x_1, x_1, x_1, \dots\}$. 则
- $\{x_t\}_{t=1}^{\infty}$ 不存在序列相关 (serial correlation)
 - $\{x_t\}_{t=1}^{\infty}$ 是平稳过程 (stationary)
 - $\{x_t\}_{t=1}^{\infty}$ 不是平稳过程
 - 以上答案都不对
15. 一般而言, 用统计量 $\frac{1}{n} \sum_{i=1}^n x_i$ 估计 $E(x_i)$ 的依据是:
- 某种形式的中心极限定理 (Central Limit Theorem)
 - 某种形式的大数定律 (Law of Large Numbers)
 - 斯拉斯基定理 (Slutsky Theorem)
 - 以上答案都不对
16. 在图 2 中, β 是真实待估计参数, β^* 和 $\hat{\beta}$ 是两个不同的估计量, $f(\cdot)$ 是对应的抽样分布密度函数 (sampling density)。以下那个说法是正确的:
- β^* 是无偏的 (unbiased)
 - $\hat{\beta}$ 是无偏的
 - β^* 比 $\hat{\beta}$ 更有效 (more efficient)
 - $\hat{\beta}$ 比 β^* 更有效
17. 符号定义与上题相同。在图 3 中, 以下那个说法是正确的:
- β^* 是有偏的 (biased)
 - $\hat{\beta}$ 是有偏的
 - β^* 比 $\hat{\beta}$ 更有效 (more efficient)
 - $\hat{\beta}$ 比 β^* 更有效

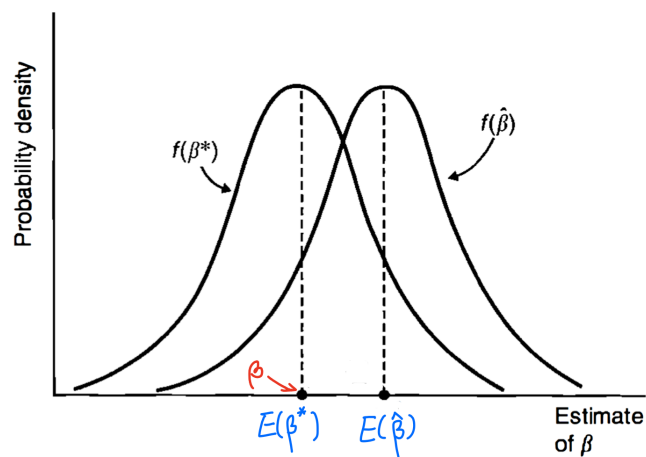


图 2

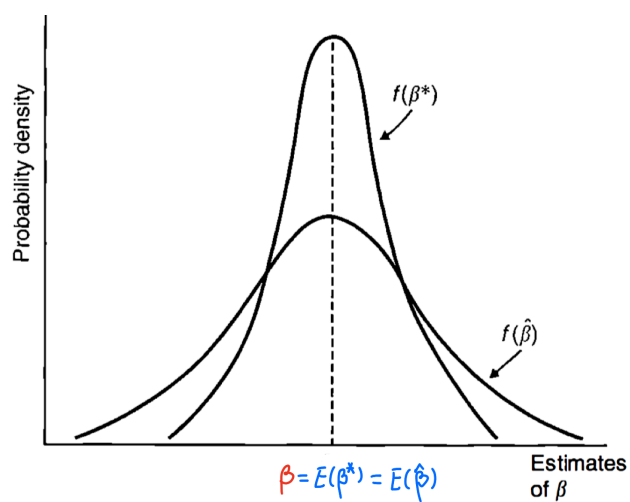


图 3

18. 假设 $\hat{\theta}$ 是对真实参数 θ_0 的任意无偏估计, 则在一定的正则条件下, $\hat{\theta}$ 的方差不会小于“克莱默-劳下限”(Cramer-Rao Lower Bound), 也即:

$$\begin{aligned} \text{A. } \text{Var}(\hat{\theta}) &\geq -E \left[\frac{\partial^2 \ln L(\theta; \mathbf{y})}{\partial \theta \partial \theta'} \right] \Big|_{\theta=\theta_0} \\ \text{B. } \text{Var}(\hat{\theta}) &\geq \left(-E \left[\frac{\partial^2 \ln L(\theta; \mathbf{y})}{\partial \theta \partial \theta'} \right] \Big|_{\theta=\theta_0} \right)^{-1} \\ \text{C. } \text{Var}(\hat{\theta}) &\geq E \left[\frac{\partial^2 \ln L(\theta; \mathbf{y})}{\partial \theta \partial \theta'} \right] \Big|_{\theta=\theta_0} \\ \text{D. } \text{Var}(\hat{\theta}) &\geq \left(E \left[\frac{\partial^2 \ln L(\theta; \mathbf{y})}{\partial \theta \partial \theta'} \right] \Big|_{\theta=\theta_0} \right)^{-1} \end{aligned}$$

19 - 20题模型设定如下: 对于线性回归模型, 检验原假设 $H_0: \beta = \beta_0$, 其中 $\beta_{K \times 1}$ 为未知参数, β_0 已知, 共有 K 个约束。定义 $\hat{\beta}_U$ 为关于参数 β 的无约束估计量:

$$\hat{\beta}_U \equiv \arg \max_{\beta} \ln L(\beta),$$

$\hat{\beta}_R$ 为关于参数 β 的有约束估计量:

$$\hat{\beta}_R \equiv \arg \max_{\beta} \ln L(\beta) \quad \text{s.t.} \quad \beta = \beta_0.$$

其中 $\ln L(\cdot)$ 为对数似然函数。相关变量如图 4 所示。直线 AB 与 $\ln L(\beta)$ 在 $\ln L(\beta) = \ln L(\hat{\beta}_R)$ 处相切。

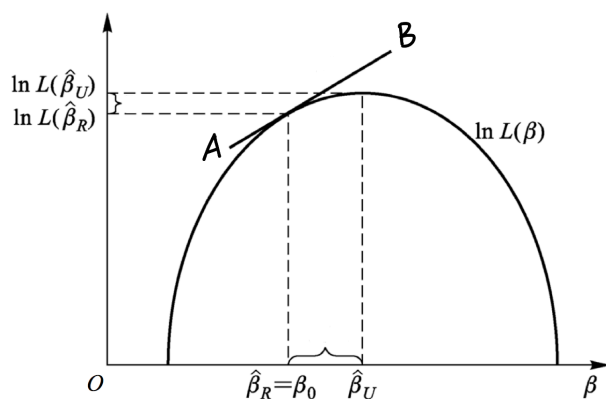


图 4

19. 沃尔德检验 (Wald test) 的原理是:

- A. 在 H_0 成立时, $\ln L(\hat{\beta}_U) - \ln L(\hat{\beta}_R) \approx 0$
- B. 在 H_0 成立时, $|\hat{\beta}_U - \hat{\beta}_R| \approx 0$
- C. 在 H_0 成立时, 图 4 中直线 AB 的斜率接近 0
- D. 以上答案都不对

20. 拉格朗日乘子检验 (Lagrange Multiplier test) 的原理是:

- A. 在 H_0 成立时, $\ln L(\hat{\beta}_U) - \ln L(\hat{\beta}_R) \approx 0$
- B. 在 H_0 成立时, $|\hat{\beta}_U - \hat{\beta}_R| \approx 0$
- C. 在 H_0 成立时, 图 4 中直线 AB 的斜率接近 0
- D. 以上答案都不对

21. 若真实模型为 $y_i = x_{i1}\beta_1 + x_{i2}\beta_2 + \epsilon_i$, $\text{Cov}(x_{i1}, \epsilon_i) = \text{Cov}(x_{i2}, \epsilon_i) = 0$ 。而实际估计的模型为 $y_i = x_{i1}\beta_1 + u_i$ 。如果 $\text{Cov}(x_{i1}, x_{i2}) \neq 0$, 此时

- A. β_1 的估计量 $\hat{\beta}_1$ 是一致的
- B. β_1 的估计量 $\hat{\beta}_1$ 不一致
- C. β_1 的估计量 $\hat{\beta}_1$ 的样本方差变小
- D. 条件不足无法判断

22. 考虑如下结构变动模型:

$$\begin{aligned} \mathbf{y}_1 &= \mathbf{X}_1\beta_1 + \epsilon_1 \\ \mathbf{y}_2 &= \mathbf{X}_2\beta_2 + \epsilon_2 \end{aligned}$$

其中，下标 1 表示变量在第 1 阶段的样本值，下标 2 表示同一变量在第 2 阶段的样本值。需要检验经济结构在第 1 阶段和第 2 阶段是否发生变化，也即 $H_0: \beta_1 = \beta_2$ 。考虑如下方案：

1. 对整个样本进行回归，得到残差平方和 $\mathbf{e}'\mathbf{e}$
2. 对第 1 部分子样本进行回归，得到残差平方和 $\mathbf{e}_1'\mathbf{e}_1$
3. 对第 2 部分子样本进行回归，得到残差平方和 $\mathbf{e}_2'\mathbf{e}_2$
4. 构建 F 统计量， $F = \frac{(\mathbf{e}'\mathbf{e} - \mathbf{e}_1'\mathbf{e}_1 - \mathbf{e}_2'\mathbf{e}_2)/K}{(\mathbf{e}_1'\mathbf{e}_1 + \mathbf{e}_2'\mathbf{e}_2)/(n-2K)} \approx F(K, n-2K)$ ，其中 n 为样本容量， K 为回归方程中解释变量的个数

这种检验是：

- A. Hausman test
 - B. Chow test
 - C. Hansen test
 - D. Wald test
23. 在回归模型 $y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \epsilon_i$ 中， x_{i2} 和 x_{i3} 是内生解释变量。为识别模型，不在方程中出现的工具变量的个数最少应为：
- A. 1
 - B. 2
 - C. 3
 - D. 4
24. 在回归模型 $y = \beta_1 x_1 + \beta_2 x_2 + \epsilon$ 中， x_2 为内生解释变量。对于变量 z_2 ，已知 $E(z_2 \epsilon) = 0$ ，为使 z_2 为 x_2 的工具变量， z_2 仍需要满足：
- A. $\text{Cov}(x_2, z_2) = 0$
 - B. $\text{Cov}(x_1, z_2) \neq 0$
 - C. 在线性投影 (Linear projection) $x_2 = \delta_1 x_1 + \theta_2 z_2 + u$ 中，参数 $\theta_2 \neq 0$
 - D. 以上答案都不对
25. “工具变量的个数与内生解释变量的个数一样多时，无法定量判断工具变量是否外生。”以上说法：
- A. 正确
 - B. 错误
 - C. 条件不足无法判断
26. 当工具变量的个数 L 与内生解释变量的个数 K 一样多时 ($K = L$)，考虑两组 GMM 估计量。一组选择权重矩阵 $\widehat{\mathbf{W}}_1 = \mathbf{I}_L$ ，其中 \mathbf{I}_L 为 L 阶单位矩阵，一组选择 $\widehat{\mathbf{W}}_2 = \widehat{\mathbf{S}}^{-1} \equiv \frac{1}{n} \sum_{i=1}^n e_i^2 \mathbf{z}_i \mathbf{z}_i'$ 是 $E(\epsilon_i \mathbf{z}_i \mathbf{z}_i')$ 的一致估计量。记

$$\hat{\beta} \equiv \arg \min_{\beta} J(\hat{\beta}, \widehat{\mathbf{W}}), \quad J(\hat{\beta}, \widehat{\mathbf{W}}) \equiv n \left(\mathbf{g}_n(\hat{\beta}) \right)' \widehat{\mathbf{W}} \left(\mathbf{g}_n(\hat{\beta}) \right)$$

其中 $\mathbf{g}_n(\hat{\beta}) \equiv \frac{1}{n} \sum_{i=1}^n \mathbf{z}_i (y_i - \mathbf{x}_i' \hat{\beta})$ ，则

- A. $J(\hat{\beta}_1, \widehat{\mathbf{W}}_1) \neq 0$
 - B. $J(\hat{\beta}_1, \widehat{\mathbf{W}}_1) = J(\hat{\beta}_2, \widehat{\mathbf{W}}_2) = 0$
 - C. $J(\hat{\beta}_1, \widehat{\mathbf{W}}_1) < J(\hat{\beta}_2, \widehat{\mathbf{W}}_2)$
 - D. $J(\hat{\beta}_1, \widehat{\mathbf{W}}_1) > J(\hat{\beta}_2, \widehat{\mathbf{W}}_2)$
27. 当工具变量的个数 L 大于内生解释变量的个数 K ($L > K$) 且存在异方差 (Heteroskedasticity) 时，做过度识别检验 (overidentification test)，结果如图 5 所示。以下说法哪个是正确的：

- A. 根据 Sargan statistic, $p\text{-value} > 0.05$ ，无法拒绝原假设，因此所有工具变量都是外生的

Sargan statistic (overidentification test of all instruments):		0.130
Chi-sq(1)	P-val =	0.7185
Hansen J statistic (overidentification test of all instruments):		0.151
Chi-sq(1)	P-val =	0.6972

图 5

- B. 根据 Sargan statistic, $p\text{-value} > 0.05$, 拒绝原假设, 因此存在部分工具变量是内生的
 C. 根据 Hansen J statistic, $p\text{-value} > 0.05$, 无法拒绝原假设, 因此所有工具变量都是外生的
 D. 根据 Hansen J statistic, $p\text{-value} > 0.05$, 拒绝原假设, 因此存在部分工具变量是内生的

面板模型设定如下: $y_{it} = \mathbf{x}_{it}'\boldsymbol{\beta} + \mathbf{z}_i'\boldsymbol{\delta} + u_i + \varepsilon_{it}$, $i = 1, \dots, n$, $t = 1, \dots, T$. 其中 \mathbf{z}_i 为不随时间而变的个体特征, \mathbf{x}_{it} 可随个体及时间而变。 u_i 为不可观测且不随时间变化的个体特征, 如无特别说明, u_i 与 x_{it} 的某个分量相关。 ε_{it} 为不可观测、随个体及时间变化的扰动项。假设 $\{\varepsilon_{it}\}$ 为 i.i.d., 且与 u_i 不相关。

28. 对面板模型进行固定效应 (fixed effect) 估计, 也即对给定的个体 i , 将方程两边对时间平均, 再减去均值:

$$y_{it} - \bar{y}_i = (\mathbf{x}_{it} - \bar{\mathbf{x}}_i)'\boldsymbol{\beta} + (\varepsilon_{it} - \bar{\varepsilon}_i)$$

为了使参数 $\boldsymbol{\beta}$ 的估计是一致的 (consistent), 需要以下哪个条件:

- A. $E(\varepsilon_{it}|\mathbf{x}_{it}) = 0$
 B. $E(\varepsilon_{it}|\mathbf{x}_{i1}, \dots, \mathbf{x}_{iT}) = 0$
 C. $E(u_i|\mathbf{x}_{it}) = 0$
 D. $E(u_i|\mathbf{x}_{i1}, \dots, \mathbf{x}_{iT}) = 0$
29. 随机效应模型 (random effects) 和固定效应模型 (fixed effects) 的主要区别是:
- A. 随机效应模型不包含 u_i
 B. 随机效应模型中 u_i 和 ε_{it} 不相关
 C. 随机效应模型中 u_i 和 \mathbf{z}_i 不相关
 D. 随机效应模型中 u_i 和 $\{\mathbf{x}_{it}, \mathbf{z}_i\}$ 均不相关
30. 考虑动态面板模型, $y_{it} = \alpha + \rho y_{i,t-1} + \mathbf{x}_{it}'\boldsymbol{\beta} + \mathbf{z}_i'\boldsymbol{\delta} + u_i + \varepsilon_{it}$, $t = (2, \dots, T)$. 用 Arellano-Bond “差分 GMM” 进行估计的前提条件是:
- A. ε_{it} 关于时间 t 是平稳的
 B. ε_{it} 关于个体 i 同方差
 C. ε_{it} 关于时间 t 不存在自相关
 D. ε_{it} 关于个体 i 独立同分布
31. 假设 x_i 服从均值为 μ , 方差为 σ^2 的正态分布。现从 x_i 的分布中随机抽取 n 个样本 x_1, \dots, x_n , 用 $\hat{\mu} \equiv \frac{1}{n} \sum_{i=1}^n x_i$ 估计 μ , $\hat{\sigma}^2 \equiv \frac{1}{n-1} \sum_{i=1}^n (x_i - \hat{\mu})^2$ 估计 σ^2 。则样本均值 $\hat{\mu}$ 的方差

$$\text{Var}(\hat{\mu}) = \sigma^2/n,$$

其估计值为

$$\widehat{\text{Var}}(\hat{\mu}) = \hat{\sigma}^2/n.$$

考虑用 “自助法”(Bootstrap) 估计 $\widehat{\text{Var}}(\hat{\mu})$: 从 $\{x_1, \dots, x_n\}$ 中以 $1/n$ 的概率有放回抽样 n 个数 $\{x_1^*, \dots, x_n^*\}$ 并计算样本均值 $\mu_{nb}^* \equiv \frac{1}{n} \sum_{i=1}^n x_i^*$, 重复 B 次, 计算

$$\widehat{\text{Var}}_{\text{boot}} = \frac{1}{B-1} \sum_{b=1}^B \left(\mu_{nb}^* - \frac{1}{B} \sum_{b=1}^B \mu_{nb}^* \right)^2$$

当 $n = 10$, $B = 100000$ 时, 以下哪种情况最有可能发生:

- A. $\text{Var}(\hat{\mu}) \approx \widehat{\text{Var}}(\hat{\mu})$
- B. $\text{Var}(\hat{\mu}) \approx \widehat{\text{Var}}_{\text{boot}}$
- C. $\widehat{\text{Var}}(\hat{\mu}) \approx \widehat{\text{Var}}_{\text{boot}}$
- D. 以上答案都不对

32. 考虑以下两期面板模型,

$$y_{it} = \beta_0 + \beta_1 G_i \cdot D_t + \beta_2 G_i + \gamma D_t + \epsilon_{it}, \quad (i = 1, \dots, n; t = 1, 2)$$

用“双重差分估计量”(Difference-in-Difference estimator) 估计实验组与控制组在实验后的差异。则 β_1 的估计量在图 6 中为:

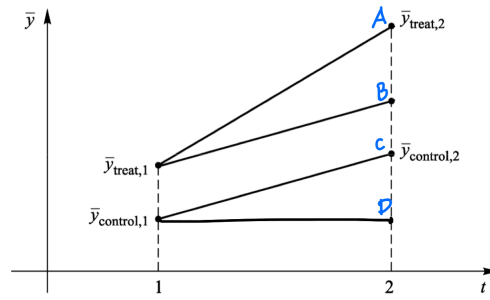


图 6

- A. AB
 - B. BC
 - C. AC
 - D. AD
33. 用倾向得分 (Propensity Score) 匹配处理组 and 对照组时, 重叠性假定 (overlap assumption) 指的是:
- A. 对于 \mathbf{x} 的任何取值, 都有 $0 < p(\mathbf{x}) < 1$, $p(\cdot)$ 是进入实验组的概率
 - B. 给定 \mathbf{x}_i , (y_{0i}, y_{1i}) 独立于 D_i , D_i 是进入实验组的虚拟变量
 - C. $E(y_{0i} | \mathbf{x}_i, D_i) = E(y_{0i} | \mathbf{x}_i)$, 且 $E(y_{1i} | \mathbf{x}_i, D_i) = E(y_{1i} | \mathbf{x}_i)$
 - D. 存在 x_i , D_i 可完全由 x_i 是否超过某值所决定
34. 选择在 $x - c$ 两侧带宽为 h 的断点回归, 不考虑 x 的高次项, 以下哪个回归模型是正确的:
- A. $y_i = \alpha + \beta(x_i - c) + \delta D_i + \gamma(x_i - c) D_i + \epsilon_i \quad (i = 1, \dots, n)$
 - B. $y_i = \alpha + \beta(x_i - c) + \delta D_i + \gamma_1(x_i - c) D_i + \epsilon_i \quad (c - h < x_i < c + h)$
 - C. $y_i = \alpha + \beta x_i + \delta D_i + \gamma x_i D_i + \epsilon_i \quad (i = 1, \dots, n)$
 - D. $y_i = \alpha + \beta x_i + \delta D_i + \gamma x_i D_i + \epsilon_i \quad (c - h < x_i < c + h)$
35. 利用多因子模型做股票收益率预测, 备选模型有 3 因子、4 因子模型。将数据划分为训练集 (training set) 和验证集 (validation set), 在验证集上 3 因子和 4 因子模型的均方误差 (Mean squared error) 分别为 0.463 和 0.570。以下哪个说法是正确的:
- A. 根据均方误差, 应当选择 3 因子模型
 - B. 在新的测试集 (test set) 上, 3 因子模型一定比 4 因子模型表现更好
 - C. 根据均方误差, 应当选择 4 因子模型
 - D. 在新的测试集 (test set) 上, 4 因子模型一定比 3 因子模型表现更好

2 问答题 (共 30 分)

1. 有回归模型

$$y = \beta_1 x_1 + \beta_2 x_2 + u.$$

其中 x_1 是外生变量, x_2 是内生变量. 现有 x_2 的工具变量 z_1 , 并做 2SLS 的第一阶段回归:

$$x_2 = \pi_1 x_1 + \pi_2 z_1 + v,$$

得到残差项 \hat{v} . 继续做如下回归:

$$y = \delta_1 x_1 + \delta_2 x_2 + \rho \hat{v} + r$$

- 1) (3 分) x_1 对 \hat{v} 回归的残差是什么? 为什么?
- 2) (3 分) x_2 对 \hat{v} 回归的残差是什么? 为什么?
- 3) (4 分) 用分块回归的结果说明, (1)式的回归系数估计 $\hat{\delta}_1, \hat{\delta}_2$ 等于用 2SLS 估计(1)式得到的 $\hat{\beta}_1, \hat{\beta}_2$.

2. 考虑如下关于计量考试成绩的模型:

$$score = \beta_0 + \beta_1 hours + \beta_2 prmEcon + \beta_3 mathScore + \beta_4 male + \beta_5 mentorHour + \beta_6 exercise + u_1, \quad (2)$$

其中 $score$ 是计量考试成绩的某个量化指标; $hours$ 是每周学习小时数; $prmEcon$ 是是否学习过初级计量经济学的虚拟变量; $mathScore$ 是本科数学课程的平均分; $male$ 是性别虚拟变量; $mentorHour$ 是每周和导师讨论的小时数; $exercise$ 是每周运动小时数.

- 1) (2 分) 误差项 u_1 是否可能包含了遗漏变量? 如有遗漏变量, 可能和哪些回归变量相关?
- 2) (2 分) 回归方程中是否可能包含多余变量? 多余变量对回归参数估计的影响是什么?
- 3) (2 分) 假设你收集了另外一个变量, $distSport$, 表示“住处和运动馆之间的距离”. 这个变量是否和 u_1 相关? 为什么?
- 4) (2 分) 现假设 $distSport$ 和 u_1 不相关. 同时在(2)式中, $exercise$ 和 u_1 相关, 但其他变量和 u_1 不相关. 写出 2SLS 中第一阶段回归的式子, 并说明还需要什么条件, (2)式中参数的估计是一致的.
- 5) (2 分) 上一问中你给出的条件可以用什么办法进行检验?

3. 考虑如下回归, x 是内生变量, z 是 x 的工具变量.

$$y = \alpha + \beta x + \varepsilon,$$

$$E(\varepsilon) = 0, \quad \text{Cov}(z, \varepsilon) = 0, \quad \text{Cov}(z, x) \neq 0, \quad E(\varepsilon^2|z) = \sigma^2$$

1) (4 分) 证明 OLS 回归时, 估计量的渐进方差

$$\text{Avar}(\sqrt{N}(\hat{\beta}_{OLS} - \beta)) = \frac{\sigma^2}{\sigma_x^2},$$

其中 $\sigma_x^2 = \text{Var}(x)$.

2) (4 分) 证明 IV 估计量的渐进方差

$$\text{Avar}(\sqrt{N}(\hat{\beta}_{IV} - \beta)) = \frac{\sigma^2}{\rho_{zx}^2 \sigma_x^2},$$

其中 $\rho_{zx} = \text{Corr}(z, x)$. (提示: 工具变量回归系数的表达式为 $\hat{\beta}_{IV} = (\mathbf{Z}'\mathbf{X})^{-1}\mathbf{Z}'\boldsymbol{\varepsilon}$.)

3) (2) 比较两个渐进方差, 说明 σ^2 , σ_x^2 , ρ_{zx} 对工具变量估计量方差的影响。