

ps4

rafehq

November 2018

Question 3

a

\Rightarrow

Take $\mathbb{A}\mathbb{X} = \mathbb{I}$

$$E(\tilde{\beta}|X_1, \dots, X_n) = E(\mathbb{A}\mathbb{Y}|X_1, \dots, X_n) \quad (1)$$

$$= \mathbb{A}E(\tilde{\beta}|X_1, \dots, X_n) \quad (2)$$

Note, this follows as $E(f(x)y|x) = E(f(x))E(y|x)$ Now, we have that this becomes

$$\begin{aligned} &= \mathbb{A}'(E(Y_1|X_1, \dots, X_n) \dots E(Y_n|X_1, \dots, X_n))' \\ &= \mathbb{A}'(E(Y_1|X_1) \dots E(Y_n|X_n))' \\ &= \mathbb{A}'(X_1\beta \dots X_n\beta)' \\ &= \mathbb{A}'\mathbb{X}\beta \\ &= \mathbb{I}\beta = \beta \end{aligned}$$

The second equality above follows from the fact that (X_i, Y_i) is iid so Y_i is independent of all of the X_j for $j \neq i$.

\Leftarrow

Suppose $E(\tilde{\beta}|X_1, \dots, X_n) = \beta$. Then, we have, from the equalities above, that

$$E(\tilde{\beta}|X_1, \dots, X_n) = \mathbb{A}'\mathbb{X}\beta$$

Thus,

$$\begin{aligned} \mathbb{A}'\mathbb{X}\beta - \beta &= 0 \Rightarrow \\ (\mathbb{A}'\mathbb{X} - \mathbb{I})\beta &= 0 \end{aligned}$$

As this must hold true for all β , we must have that $\mathbb{A}\mathbb{X} = \mathbb{I}$ (i.e the only eigenvalue is $\lambda = 1$ for all $\beta \in \mathbb{R}^{k+1}$)

b

$$\begin{aligned} \text{Var}(\tilde{\beta}|X_1, \dots, X_n) &= E((\mathbb{A}\mathbb{Y})^2|X_1, \dots, X_n) - E(\mathbb{A}\mathbb{Y}|X_1, \dots, X_n)^2 \\ &= \mathbb{A}'(E(\mathbb{Y})^2|X_1, \dots, X_n)\mathbb{A} - \mathbb{A}'E(\mathbb{Y}|X_1, \dots, X_n)\mathbb{A} \end{aligned}$$

Note, again, (X_i, Y_i) is iid so Y_i is independent of all of the X_j for $j \neq i$ and thus $E(Y_i|X_1, \dots, X_n) = E(Y_i|X_i)$. Furthermore, we have that Y_i is independent of all of the Y_j for $j \neq i$. Thus, $E(Y_j Y_i) = 0$ for $i \neq j$ and thus:

$$\begin{aligned} &= \mathbb{A}'(\text{diag}(E(Y_1^2|X_1), \dots, E(Y_n^2|X_n)) - E(Y_1|X_1)^2 \dots E(Y_n|X_n)^2)\mathbb{A} \\ &= \mathbb{A}'(\text{diag}(\text{Var}(Y_1|X_1) \dots \text{Var}(Y_n|X_n))\mathbb{A} \\ &= \mathbb{A}'\text{diag}(\sigma^2(X_1) \dots \sigma^2(X_n))\mathbb{A} \\ &= \mathbb{A}'\mathbb{D}\mathbb{A} \end{aligned}$$

c

We take

$$\begin{aligned} \mathbb{X}'\mathbb{D}^{-1}\mathbb{X} &= (X_1 \dots X_n) \text{diag}\left(\frac{1}{\sigma^2(X_1)}, \dots, \frac{1}{\sigma^2(X_n)}\right) (X_1 \dots X_n)' \\ &= \frac{X_1' X_1}{\sigma^2(X_1)} + \dots + \frac{X_n' X_n}{\sigma^2(X_n)} \end{aligned}$$

Now, note that as \mathbb{X} has all its columns linearly independent, $\mathbb{X}a \neq 0 \Leftrightarrow a \neq 0$. Take such an $a \neq 0$:

$$a^T \mathbb{X}'\mathbb{D}^{-1}\mathbb{X}a = \frac{a_1^2 X_1' X_1}{\sigma^2(X_1)} + \dots + \frac{a_n^2 X_n' X_n}{\sigma^2(X_n)}$$

We have that $a_i^2 \geq 0$ (and, by definition, $a_i^2 > 0$ for some i). Thus, the above sum is strictly positive. This shows that $\mathbb{X}'\mathbb{D}^{-1}\mathbb{X}$ is positive definite, which in turn establishes that it is invertible.

c

Take¹

$$\begin{aligned} \text{Var}(\tilde{\beta}|X_1, \dots, X_n) &= \mathbb{A}'\mathbb{D}\mathbb{A} \\ &= (\mathbb{X}'\mathbb{D}^{-1}\mathbb{X})^{-1} \mathbb{X}'\mathbb{D}^{-1}\mathbb{D}\mathbb{D}^{-1}\mathbb{X}(\mathbb{X}\mathbb{D}'^{-1}\mathbb{X})^{-1} \\ &= (\mathbb{X}'\mathbb{D}^{-1}\mathbb{X})^{-1} \mathbb{X}'\mathbb{D}^{-1}\mathbb{X}(\mathbb{X}\mathbb{D}'^{-1}\mathbb{X})^{-1} \\ &= (\mathbb{X}'\mathbb{D}^{-1}\mathbb{X})^{-1} \end{aligned}$$

¹for a diagonal matrix \mathbb{D} , $\mathbb{D}' = \mathbb{D}$

Also, clearly:

$$((\mathbb{X}'\mathbb{D}^{-1}\mathbb{X})^{-1}\mathbb{X}'\mathbb{D}^{-1})\mathbb{X} = (\mathbb{X}'\mathbb{D}^{-1}\mathbb{X})^{-1}(\mathbb{X}'\mathbb{D}^{-1}\mathbb{X}) = \mathbb{I}$$

By (a), the estimator $\tilde{\beta}_n$ is then unbiased

e

Take $\tilde{\mathbb{A}}\mathbb{Y}$ as another unbiased estimator of β wherein $\gamma_n \equiv \tilde{\mathbb{A}}\mathbb{Y}$. As $E(\gamma_n|X_1, \dots, X_n) = \beta$, we have that $\tilde{\mathbb{A}}'\mathbb{X} = \mathbb{I}$. Now, we run through the argument used in the Gauss-Markov theorem: Take $C = \mathbb{A} - \tilde{\mathbb{A}}$, then:

$$\begin{aligned} Var(\gamma_n|X_1, \dots, X_n) - Var(\tilde{\beta}_n|X_1, \dots, X_n) &= (C + \mathbb{A})'\mathbb{D}(C + \mathbb{A}) - \mathbb{A}'\mathbb{D}\mathbb{A} \\ &= C'\mathbb{D}C + \mathbb{A}'\mathbb{D}C + C'\mathbb{D}\mathbb{A} \\ &= C'\mathbb{D}C + (\mathbb{X}'\mathbb{D}^{-1}\mathbb{X})^{-1}\mathbb{X}'\mathbb{D}^{-1}\mathbb{D}C + C'\mathbb{D}\mathbb{D}^{-1}\mathbb{X}(\mathbb{X}'\mathbb{D}^{-1}\mathbb{X}) \\ &= C'\mathbb{D}C + (\mathbb{X}'\mathbb{D}^{-1}\mathbb{X})^{-1}\mathbb{X}'C + C'\mathbb{X}(\mathbb{X}'\mathbb{D}^{-1}\mathbb{X}) \\ &= C'\mathbb{D}C + (\mathbb{X}'\mathbb{D}^{-1}\mathbb{X})^{-1}\mathbb{X}'(\mathbb{A} - \tilde{\mathbb{A}}) + (\mathbb{A} - \tilde{\mathbb{A}})\mathbb{X}(\mathbb{X}'\mathbb{D}^{-1}\mathbb{X}) \\ &= C'\mathbb{D}C \end{aligned}$$

Again, using the above argument in (c), we see that the differences in the variances is a positive semidefinite matrix s.t

$$a^T C' \mathbb{D} C a = \sum_{i=1}^n \frac{a_i^2}{\sigma^2(X_i)} \geq 0$$

This establishes that the estimator in the previous part is *best* as the variance of any other predictor is the same or greater.

Question 7

a

Note that this follows simply from the random assignment. Because individuals are not aware of their assignment before the experiment and have equal likelihood of being assigned to treatment or control groups, their probability of being assigned to the treatment is independent of their α_i and β_i . Thus, D_i is independent of (α_i, β_i)

b

Note, that we can write down the β (i.e from class) as $Var(D_i, D_i')^{-1}Cov(D_i, Y_i)$ for the special case of a bivariate regression.

$$\begin{aligned} \beta &= Var(D_i D_i')^{-1} Cov(D_i Y_i) \\ &= Var(D_i D_i')^{-1} Cov(D_i \alpha_i + \beta_i D_i) \\ &= Var(D_i D_i')^{-1} (Cov(D_i \alpha_i) + Cov(\beta_i D_i D_i)) \end{aligned}$$

Since (α_i, β_i) are independent of D_i , the first term is just 0, and the second term is $E(\beta_i)E(D_i D_i')$. Thus, we have:

$$\begin{aligned} &= \text{Var}(D_i D_i')^{-1} E(\beta_i) \text{Cov}(D_i D_i') \\ &= \text{Var}(D_i D_i')^{-1} E(D_i D_i') E(\beta_i) \\ &= E(\beta_i) \end{aligned}$$

Using the above, note we can also solve for α :

$$\alpha = E(y - \beta D_i)$$

$$\begin{aligned} \alpha &= E(Y_i) - \beta E(D_i) \\ &= E(\alpha_i + \beta_i D_i) - \beta E(D_i) \\ &= E(\alpha_i) + E(\beta_i D_i) - \beta E(D_i) \end{aligned}$$

Again, by the independence of β_i and D_i , we get that this equals

$$\begin{aligned} &= E(\alpha_i) + E(\beta_i)E(D_i) - \beta E(D_i) \\ &= E(\alpha_i) + \beta E(D_i) - \beta E(D_i) \\ &= E(\alpha_i) \end{aligned}$$

C

Note that since we aren't given homoskedasticity², we can use robust standard errors:

$$\begin{aligned} \hat{\Omega} &= \left(\frac{1}{n} \sum_{i=1}^n X_i X_i' \right)^{-1} \left(\frac{1}{n} \sum_{i=1}^n X_i X_i' \hat{\epsilon}_i \right) \left(\frac{1}{n} \sum_{i=1}^n X_i X_i' \right)^{-1} \\ &\rightarrow \Omega \end{aligned}$$

Now, note that the errors in β are just the component of Ω in position (2,2). As $\hat{\Omega}$ is a consistent estimator of Ω (from class and delineated above), we can construct the confidence region:

$$C_n = [\hat{\beta}_n - \Phi^{-1}(1 - \alpha/2) \times \sqrt{\frac{\hat{\Omega}_{2,2}}{n}}, \hat{\beta}_n + \Phi^{-1}(1 - \alpha/2) \times \sqrt{\frac{\hat{\Omega}_{2,2}}{n}}]$$

²In fact, in this setup it is probable that $\text{Var}(\epsilon_i)$ is not constant as α_i and β_i depend on i , and indeed, $\text{Var}(\epsilon_i | D_i) = (1 - D_i)\text{Var}(\alpha_i) + D_i\text{Var}(\alpha_i + \beta_i)$