2016 PRELIMINARY EXAMINATION
241B Questions

1)

You are asked to determine how the conditional mean of a variable y depends on a (continuous) conditioning variable x. Because x is continuous we typically cannot determine the exact form of the conditional expectation and so use the linear approximation:

$$y = x^{\mathrm{T}}\beta_0 + u. \tag{1}$$

a) Why study the mean of the conditional distribution of y given x? Write the components of u.

b) Let β_0 be the value of β that minimizes the mean-square approximation error. Derive β_0.

c) Consider two values β and β'. State the requirement for these two values to be separately identified. In light of your answer to part b, state the requirement for β_0 to be identified and state the condition for the approximation to the conditional mean, $x^{\mathrm{T}}\beta_0$, to be identified.

d) Prove that $\mathbb{E}(xu) = 0$.

2)

Let y and x be two scalar, continuous random variables. Interest centers on the conditional mean of y given x and, in particular, on how this mean changes as x changes. As the mean is unknown, it is approximated with a regression of the form

$$y = \beta x + \alpha + u.$$

The actual approximation, which is termed the linear projection, is

$$\begin{aligned} \mathcal{P}(y|x) &= \beta_0 x + \alpha_0, \\ y &= \mathcal{P}(y|x) + u. \end{aligned}$$

where β_0 and α_0 are the values of β and α that minimize the mean-square approximation error.

a) Derive β_0 and α_0.

For the remainder of the problem, assume that

$$y = x + x^2 + x^3,$$

so that $e \equiv 0$, and let $x \sim \mathcal{N}(0,1)$.

b) Derive β_0 and α_0 and write the components of u.

c) Compare, with a table of values, $\mathbb{E}(y|x)$ and $\mathcal{P}(y|x)$ for $x = \{1,2,3,4\}$.

d) Prove that $\mathbb{E}(xu) = 0$.

Proposed Answer

Question 1

a) One purpose of econometrics is to understand the relation between a dependent variable y and a covariate x. We may be interested in determining the causal impact of x on the distribution of y or we may be interested in predicting the value of y for a given value of x. In either case, we wish to understand how the distribution of y depends on x. One natural measure of the distribution is the mean, which corresponds to an average causal effect or the predictor that minimizes mean squared error.

To determine the components of u, write

$$\begin{aligned} y &= \mathbb{E}(y|x) + e \\ &= x^{\mathrm{T}}\beta + u \quad \text{where } u = e + \left[\mathbb{E}(y|x) - x^{\mathrm{T}}\beta\right]. \end{aligned}$$

The error u consists of the random component e and the error from approximating the conditional mean with a linear function, $\mathbb{E}(y|x) - x^{\mathrm{T}}\beta$.

b) The mean-square approximation error is

$$d(\beta) = \int_{\mathbb{R}^k} \left(m(x) - x^{\mathrm{T}}\beta\right)^2 f_x(x)\, dx,$$

where $m(x) := \mathbb{E}(y|x)$. Expanding the quadratic term

$$\begin{aligned} d(\beta) &= \mathbb{E}m(x)^2 - 2\beta^{\mathrm{T}}\mathbb{E}(xm(x)) + \beta^{\mathrm{T}}\mathbb{E}\left(xx^{\mathrm{T}}\right)\beta \\ &= \mathbb{E}m(x)^2 - 2\beta^{\mathrm{T}}\mathbb{E}(xy) + \beta^{\mathrm{T}}\mathbb{E}\left(xx^{\mathrm{T}}\right)\beta \qquad \mathit{LIE} \end{aligned}$$

The first-order condition to minimize $d(\beta)$ yields

$$-2\mathbb{E}(xy) + 2\mathbb{E}\left(xx^{\mathrm{T}}\right)\beta_0 = 0.$$

Thus

$$\mathbb{E}\left(xx^{\mathrm{T}}\right)\beta_0 = \mathbb{E}(xy),$$

defines β_0 as the minimizer of $d(\beta)$. (The uniqueness of β_0 is a question of identification and is not a question of definition.)

c) In general, β and β' are separately identified if

$$\mathbb{P}_\beta \neq \mathbb{P}_{\beta'} \Rightarrow \beta \neq \beta'.$$

For the specific case of the linear approximation to the conditional mean, the condition is β and β' are separately identified if

$$\left(\mathbb{E}\left(xx^{\mathrm{T}}\right)\right)^{-1}\mathbb{E}\left(xy\right) \text{ from } \beta$$

does not equal

$$\left(\mathbb{E}\left(xx^{\mathrm{T}}\right)\right)^{-1}\mathbb{E}\left(xy\right) \text{ from } \beta',$$

which is equivalent to, there is a unique solution

$$\beta_0 = \left(\mathbb{E}\left(xx^{\mathrm{T}}\right)\right)^{-1}\mathbb{E}\left(xy\right).$$

Thus, if $\mathbb{E}\left(xx^{\mathrm{T}}\right)$ is invertible, then β_0 is identified. That is, the linear projection coefficients are unique. It is important to note that $\mathbb{E}\left(xu\right) = 0$ is not needed for identification, but follows directly from the unique solution for β_0.
If $\mathbb{E}\left(xx^{\mathrm{T}}\right)$ is not invertible, then there are many values of

$$\beta_0 = \left(\mathbb{E}\left(xx^{\mathrm{T}}\right)\right)^{-}\mathbb{E}\left(xy\right),$$

but all solutions yield an equivalent value of $x^{\mathrm{T}}\beta_0$, so that the best linear prediction of y is always identified.

d) Note, by definition

$$u = y - x^{\mathrm{T}}\beta_0,$$

and

$$\begin{aligned} \mathbb{E}\left(xu\right) &= \mathbb{E}\left(xy\right) - \mathbb{E}\left(xx^{\mathrm{T}}\right)\left(\mathbb{E}\left(xx^{\mathrm{T}}\right)\right)^{-1}\mathbb{E}\left(xy\right) \\ &= 0. \end{aligned}$$

The value of β that minimizes the mean-square approximation error automatically yields an error that is uncorrelated with the covariates.

Question 2

a) The mean-square approximation error is

$$d(\alpha,\beta)=\int_{\mathbb{R}^k}(m(x)-\beta x-\alpha)^2 f_x(x)\,dx,$$

where $m(x):=\mathbb{E}(y|x)$. Expanding the quadratic term

$$\begin{aligned}
d(\alpha,\beta) &= \mathbb{E}m(x)^2-2\left[\beta\mathbb{E}(xm(x))+\alpha\mathbb{E}(m(x))\right]+\left(\beta^2\mathbb{E}x^2+2\alpha\beta\mathbb{E}(x)+\alpha^2\right)\\
&= \mathbb{E}m(x)^2-2\left[\beta\mathbb{E}(xy)+\alpha\mathbb{E}(y)\right]+\left(\beta^2\mathbb{E}x^2+2\alpha\beta\mathbb{E}(x)+\alpha^2\right) \qquad LIE
\end{aligned}$$

The first-order condition to minimize $d(\alpha,\beta)$ yields:

$$\begin{aligned}
-2\mathbb{E}(y)+2\beta_0\mathbb{E}(x)+2\alpha_0 &= 0 \qquad \text{for } \alpha\\
-2\mathbb{E}(xy)+2\beta_0\mathbb{E}x^2+2\alpha_0\mathbb{E}(x) &= 0 \qquad \text{for } \beta
\end{aligned}$$

Thus

$$\begin{aligned}
\alpha_0 &= \mathbb{E}(y)-\beta_0\mathbb{E}(x)\\
\beta_0 &= \frac{\mathbb{E}(xy)-\mathbb{E}(x)\mathbb{E}(y)}{\mathbb{E}x^2-(\mathbb{E}(x))^2}
\end{aligned}$$

defines (α_0,β_0) as the minimizers of $d(\alpha,\beta)$. (The uniqueness of β_0 is a question of identification (β_0 is uniquely defined if x is not constant) and is not a question of definition.)

b) Under the assumption that $x\sim\mathcal{N}(0,1)$

$$\alpha_0=1 \qquad \beta_0=4,$$

so the linear projection of y onto x is

$$\mathcal{P}(y|x)=4x+1.$$

To determine the components of u, write

$$\begin{aligned}
y &= \mathbb{E}(y|x)+e\\
&= \mathcal{P}(y|x)+u \quad \text{where } u=e+\left[\mathbb{E}(y|x)-\mathcal{P}(y|x)\right]\\
u &= x+x^2+x^3-(4x+1)\\
&= x^2+x^3-3x-1.
\end{aligned}$$

c) The conditional expectation $\mathbb{E}(y|x)=x+x^2+x^3$ and the linear projection $\mathcal{P}(y|x)=4x+1$ differ sharply. For the given values of x

	$x=1$	$x=2$	$x=3$	$x=4$	
$\mathbb{E}(y\vert x)$	3	14	39	84	.
$\mathcal{P}(y\vert x)$	5	9	13	17	

d) Note, by construction

$$u = x^2 + x^3 - 3x - 1,$$

and

$$\begin{aligned}
\mathbb{E}(xu) &= \mathbb{E}\left(x^3\right) + \mathbb{E}\left(x^4\right) - 3\mathbb{E}\left(x^2\right) - \mathbb{E}(x) \\
&= 0.
\end{aligned}$$

The value of (α, β) that minimizes the mean-square approximation error automatically yields an error that is uncorrelated with the covariates.