Econometrics 710
Midterm Exam
March 8, 2016
Sample Answers

1. Since $\frac{\partial}{\partial x}m(x) = c_1 + 2c_2x$, then $\theta = E\left[\frac{\partial}{\partial x}m(x_i)\right] = \theta = E\left[c_1 + 2c_2x_i\right] = c_1 + 2c_2\mu_x$

2. By the formula for the best linear predictor, we know that

$$\beta_1 = \frac{\operatorname{cov}(x,y)}{\operatorname{var}(x)} = \frac{\operatorname{cov}(x,m(x))}{\operatorname{var}(x)} = \frac{\operatorname{cov}\left(x, c_0 + c_1x + c_2x^2\right)}{\operatorname{var}(x)} = c_1 + c_2\frac{\operatorname{cov}\left(x,x^2\right)}{\sigma_x^2}$$

We can write the cov term in terms of the uncentered moments as

$$\operatorname{cov}\left(x,x^2\right) = E\left((x - Ex)\left(x^2 - Ex^2\right)\right) = E\left((x-\mu_x)x^2\right) = E\left(x^3\right) - \mu_x Ex^2.$$

Making the substitutions $Ex^2 = \sigma_x^2 + \mu_x^2$ and $Ex^3 = s_x + 3\mu_x\sigma_x^2 + \mu_x^3$ we find

$$\operatorname{cov}\left(x,x^2\right) = s_x + 3\mu_x\sigma_x^2 + \mu_x^3 - \mu_x\sigma_x^2 - \mu_x^3 = s_x + 2\mu_x\sigma_x^2.$$

Thus

$$\beta_1 = c_1 + c_2\frac{s_x + 2\mu_x\sigma_x^2}{\sigma_x^2}$$

Since $\theta = E\left[c_1 + 2c_2x_i\right] = c_1 + 2c_2\mu_x$ we find

$$\beta_1 - \theta = c_2\left(\frac{s_x + 2\mu_x\sigma_x^2}{\sigma_x^2} - 2\mu_x\right) = c_2\frac{s_x}{\sigma_x^2}$$

3. $\beta_1 = \theta$ if either $c_2 = 0$ or $s_x = 0$.

First, $c_2 = 0$ occurs when the true regression is linear. Thus, as seems natural, the linear approximation will equal the average derivative when the true regression is linear.

Second, $s_x = 0$ occurs when the third centered moment of x_i is zero. This occurs when the distribution of x_i is symmetric about its mean. Thus the linear approximation will equal the average derivative when the true regression is quadratic and x_i is symmetric about its mean. This is not so obvious. Roughly, the bias on the two sides of μ_x cancel out.

In general, however, $\beta_1 \neq \theta$

4. $\widehat{\theta} = \widehat{\beta}_1 + 2\widehat{\beta}_2\mu_x$

5. The conditional mean is correctly specified, so the least-squares estimators $\widehat{\beta} = (\widehat{\beta}_0, \widehat{\beta}_1, \widehat{\beta}_2)$ are unbiased for $\beta = (c_0, c_1, c_2)$. $\widehat{\theta}$ is a linear function of $\widehat{\beta}$. Since expectation is a linear operator $\widehat{\theta}$ is therefore unbiased. More formally,

$$E\left(\widehat{\theta} - \theta\right) = E\left(\left(\widehat{\beta}_1 + 2\widehat{\beta}_2\mu_x\right) - (c_1 + 2c_2\mu_x)\right) = E\left(\widehat{\beta}_1 - c_1\right) + 2\mu_x E\left(\widehat{\beta}_2 - c_2\right) = 0$$

Thus $\widehat{\theta}$ is unbiased for θ.

6. Since this is a correctly specified regression, $\widehat{\beta} = (\widehat{\beta}_0, \widehat{\beta}_1, \widehat{\beta}_2) \to_p \beta = (c_0, c_1, c_2)$. By the continuous mapping theorem, $\widehat{\theta} = \widehat{\beta}_1 + 2\widehat{\beta}_2\mu_x \to_p c_1 + 2c_2\mu_x = \theta$. Thus $\widehat{\theta}$ is consistent for θ.

7. Since this is a correctly specified rgression

$$\sqrt{n}\left(\widehat{\beta} - \beta\right) \to_d N(0, V)$$

where $V = Q^{-1}\Omega Q^{-1}$, $Q = E\left(\widetilde{x}_i\widetilde{x}_i'\right)$, $\Omega = E\left(\widetilde{x}_i\widetilde{x}_i'e_i^2\right)$, and $\widetilde{x}_i = (1, x_i, x_i^2)'$. Set $R = (0, 1, 2\mu_x)'$. Note that $\theta = R'\beta$ and $\widehat{\theta} = R'\widehat{\beta}$. Thus

$$\sqrt{n}\left(\widehat{\theta} - \theta\right) = \sqrt{n}R'\left(\widehat{\beta} - \beta\right) \to_d N(0, R'VR)$$

8. The natural estimator is $\widehat{\theta} = \widehat{\beta}_1 + 2\widehat{\beta}_2\widehat{\mu}_x$ where $\widehat{\mu}_x = n^{-1}\sum_{i=1}^n x_i$

9. By the WLLN, $\widehat{\mu}_x \to_p \mu_x$. Thus by the continuous mapping theorem $\widehat{\theta} = \widehat{\beta}_1 + 2\widehat{\beta}_2\widehat{\mu}_x \to_p c_1 + 2c_2\mu_x = \theta$. Thus it is consistent.

10.

(a) $\widehat{\mu}_x$ is random. So sampling error in the estimate of $\widehat{\mu}_x$ for μ_x needs to be taken into account.

(b) $\widehat{\theta}$ is a nonlinear function of $(\widehat{\beta}_0, \widehat{\beta}_1, \widehat{\beta}_2, \widehat{\mu}_x)$. To use the delta method to find the asymptotic distribution, we need the joint asymptotic distribution of $(\widehat{\beta}_0, \widehat{\beta}_1, \widehat{\beta}_2, \widehat{\mu}_x)$. Thus, we define the augmented parameter vector $\gamma = (c_0, c_1, c_2, \mu_x)$ and calculate the derivative

$$R_\gamma = \frac{\partial}{\partial\gamma}\theta = \frac{\partial}{\partial\gamma}(c_1 + 2c_2\mu_x) = \begin{pmatrix} R \\ 2c_2 \end{pmatrix}$$

with $R = (0, 1, 2\mu_x)'$ defined earlier. Then we find the asymptotic distribution

$$\sqrt{n}\,(\widehat{\gamma} - \gamma) \to_d N(0, V_\gamma)$$

By the delta method we deduce

$$\sqrt{n}\left(\widehat{\theta} - \theta\right) \to_d N(0, R_\gamma' V_\gamma R_\gamma)$$

The relatively new & challenging part is the asymptotic distribution of $\widehat{\gamma}$.

(c) The method for finding the asymptotic distribution of $\widehat{\gamma}$ is to stack the estimators. We have

$$\begin{aligned}
\sqrt{n}\,(\widehat{\gamma} - \gamma) &= \sqrt{n}\begin{pmatrix} \widehat{\beta} - \beta \\ \widehat{\mu}_x - \mu_x \end{pmatrix} \\
&= \begin{pmatrix} \left(\frac{1}{n}\sum_{i=1}^n \widetilde{x}_i\widetilde{x}_i'\right)^{-1}\left(\frac{1}{\sqrt{n}}\sum_{i=1}^n \widetilde{x}_i e_i\right) \\ \frac{1}{\sqrt{n}}\sum_{i=1}^n (x_i - \mu_x) \end{pmatrix} \\
&= \begin{bmatrix} \left(\frac{1}{n}\sum_{i=1}^n \widetilde{x}_i\widetilde{x}_i'\right)^{-1} & 0 \\ 0 & 1 \end{bmatrix} \frac{1}{\sqrt{n}}\sum_{i=1}^n u_i
\end{aligned}$$

where

$$u_i = \begin{pmatrix} \widetilde{x}_i e_i \\ x_i - \mu_x \end{pmatrix}$$

Notice that u_i is iid, mean zero, and has variance

$$V_u = E(u_i u_i') = \begin{pmatrix} E\left(\widetilde{x}_i\widetilde{x}_i' e_i^2\right) & E\left(\widetilde{x}_i' e_i (x_i - \mu_x)\right) \\ E\left(\widetilde{x}_i e_i (x_i - \mu_x)\right) & E(x_i - \mu_x)^2 \end{pmatrix} = \begin{pmatrix} \Omega & 0 \\ 0 & \sigma_x^2 \end{pmatrix}$$

The off-diagonal terms are zero since $E(e_i|x_i) = 0$ under the assumption that the true conditional mean is quadratic, so the regression is correctly specified. By the CLT

$$\frac{1}{\sqrt{n}}\sum_{i=1}^n u_i \to_d N(0, V_u)$$

and by the CMT

$$\sqrt{n}\,(\widehat{\gamma} - \gamma) \to_d \begin{bmatrix} Q^{-1} & 0 \\ 0 & 1 \end{bmatrix} N(0, V_u) = N(0, V_\gamma)$$

where

$$V_\gamma = \begin{bmatrix} Q^{-1} & 0 \\ 0 & 1 \end{bmatrix}\begin{pmatrix} \Omega & 0 \\ 0 & \sigma_x^2 \end{pmatrix}\begin{bmatrix} Q^{-1} & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} V & 0 \\ 0 & \sigma_x^2 \end{bmatrix}$$

which is conveniently block-diagonal. Putting this together with the delta method we find

$$\sqrt{n}\left(\widehat{\theta} - \theta\right) \to_d N(0, V_\theta)$$

where

$$V_\theta = R_\gamma' V_\gamma R_\gamma = R'VR + 4c_2^2\sigma_x^2$$

This is similar to the distribution in part 7, but the asymptotic variance has been increased by the factor $4c_2^2\sigma_x^2$. (Notice that the change is non-negative.) Notice that the formulas from part 7 and 10 are identical when $c_2 = 0$. Thus when the true model is linear but we estimate a quadratic model and estimate the average derivative using the estimate $\widehat{\mu}_x$, then the asymptotic variance of $\widehat{\theta}$ is not increased; there is no penalty from estimation of $\widehat{\mu}_x$. In addition, the simple standard errors calculated as in part 7 will still be valid. However, when the true model is quadratic (e.g. $c_2 \neq 0$) then the asymptotic variances are different and it is necessary to use the second formula to obtain a correct standard error for $\widehat{\theta}$.