The relationship of the different exogeneity assumptions

November 15, 2006

1 Simple regression, continuos RHS variable

We have an iid sample of n observations $(i = 1, 2, ..., n)$.

The model:

$$y_i = \beta_0 + \beta_1 x_i + u_i$$

We omit the i subscripts from now.

The exogeneity assumptions are the following.

$$E(u) = 0 \tag{1}$$

This is a technical assumption; not meeting the assumption has a consequence on the regression constant only (see the first homework).

The important assumption is exogeneity of x. Consistency of the OLS estimator for β_1 requires the following exogeneity assumption:

$$Cov(u, x) = 0 \tag{2}$$

Note that technical assumption (1) implies that $Cov(u, x) = E(ux) - E(x)E(u) = E(ux)$, and so $Cov(u, x) = 0$ implies that $E(ux) = 0$.

1.1 Mean-independence

u is said to be mean-independent of x if the mean of u is the same for every value of x :

$$E(u|x) = E(u) = 0$$

where the last equality follows from technical assumption (1).

1.1.1 Mean-independence implies zero covariance

Intuitively: a (population) regression of u on x is the linear approximation to $E(u|x)$. The regression would yield a slope coefficient of $Cov(u, x)/V(x)$. If $E(u|x) = E(u)$, the conditional expectation function is a constant. Therefore, a linear approximation to it would yield a slope of zero, which implies that $Cov(u, x) = 0$.

Formally,

$$E(u|x) = E\left(\frac{ux}{x}|x\right) = \frac{1}{x}E(ux|x)$$

and therefore $E(ux|x) = xE(u|x)$. But the law of iterated expectations implies that $E(ux) = E[E(ux|x)]$. Therefore,

$$E(ux) = E[E(ux|x)] = E[xE(u|x)]$$

If $E(u|x) = E(u)$, we have that

$$E(ux) = E[xE(u|x)] = E[xE(u)] = E(x)E(u)$$

Because $Cov(u,x) = E(ux) - E(x)E(u)$, we have shown that $E(u|x) = E(u)$ implies that $Cov(u,x) = 0$.

1.1.2 Zero covariance does not imply mean-independence

Intuition: it may be that $E(u|x)$ is nonconstant but its shape is such that when averaged over x, it becomes zero. Such cases are, of course, more of text book examples than real-life cases.

One such example: $E(u|x) = x^2$, $x \sim N(0,1)$. $E(x) = 0$, $E(u) = E[E(u|x)] = E(x^2) = V(x) = 1$. $E(ux) = E[xE(u|x)] = E(x^3) = 0$. $Cov(u,x) = E(xu) - E(x)E(u) = 0 - 0 = 0$. But $E(u|x) = x^2 \neq E(u)$ in general.

1.2 Independence

u is said to be independent of x if the joint density of u and x is the same as the product of the marginal densities:

$$f(u,x) = f(u)f(x)$$

1.2.1 Independence implies mean-independence

Intuitively: if x and u are independent, then any moment of each cannot be a function of the other.

Formally: independence implies that the conditional density of u given x is equal to the marginal (unconditional) density of u:

$$f(u|x) = \frac{f(u,x)}{f(x)} = \frac{f(u)f(x)}{f(x)} = f(u)$$

As a result, we have that the conditional mean is

$$E(u|x) = \int uf(u|x)\,du = \int uf(u)\,du = E(u).$$

1.2.2 Mean-independence does not imply independence

One such case is mean-independence with conditional heteroskedasticity. Here $V(u|x)$ is a nonconstant function of x, and therefore there is some dependence between u and x. At the same time, we may have mean independence such that $E(u|x)$ is a constant function. An example:

$$u \sim N(0, x^2)$$

So that $E(u|x) = 0$, $V(u|x) = x^2$.

2 Simple regression, dummy RHS variable

Setup is as before, *iid* sample of n observations $(i = 1, 2, ..., n)$, and we omit i subscripts.

The model is as before, but the RHS variable is a dummy (is either zero or one):

$$\begin{aligned} y &= \beta_0 + \beta_1 D + u \\ D &\in \{0, 1\} \end{aligned}$$

In this case the exogeneity assumption is usually stated as mean-independence:

$$E(u|D=0) = E(u|D=1) = E(u) = 0 \tag{3}$$

Of course, mean independence implies zero covariance here as well, because in the derivation above we did not use the fact that the RHS variable (or, for that matter, u) is continuos. But here mean-independence is not stricter than the zero covariance, for the following reason.

2.1 Mean-independence and zero covariance are equivalent

Intuitively: $E(u|D)$ is a two-valued function: $E(u|D=0)$ and $E(u|D=1)$. The two are the same if and only if the line that connects them has zero slope.

Formally:

$$\begin{aligned} E(Du) &= \Pr(D=0) E(Du|D=0) + \Pr(D=1) E(Du|D=1) \\ &= \Pr(D=1) E(u|D=1) = \Pr(D=1) E(u) \\ E(D) &= \Pr(D=1) \\ Cov(u, D) &= E(Du) - E(D) E(u) \\ &= \Pr(D=1) E(u) - \Pr(D=1) E(u) \\ &= 0 \end{aligned}$$

where in the second line we used assumption (3) above, i.e. that $E(u|D=1) = E(u)$.

2.2 Independence

For continuos u, $f(u|D=0) = f(u|D=1)$.

2.2.1 Independence implies mean-independence

Intuitively: if D and u are independent, then any moment of u must be the same for any value of D.

Formally: $E(u|D=j) = \int u f(u|D=j)\, du$ so if $f(u|D=j) = f(u)$, we trivially have that $E(u|D=j) = E(u)$ whether $j = 0$ or 1.

2.2.2 Mean-independence does not imply independence

Similarly to the case before: mean-independence but conditional heteroskedasticity.

$$\begin{aligned} u &\sim N(0,1) \text{ if } D = 0 \\ u &\sim N(0,2) \text{ if } D = 1 \end{aligned}$$

So that $E(u|D=0) = E(u|D=1) = 0$ but $V(u|D=0) = 1 \neq V(u|D=1) = 2$.