Statistical Theory Homework 1

Bohao Tang

1. $Y = 2 + 0.5Z + X + ZX + \varepsilon$

(a) $E[Y \mid Z=z, X=x] = E_\varepsilon[2 + 0.5z + x + zx + \varepsilon]$

$= 2 + 0.5z + x + zx$

(b) $Var(Y \mid Z=z, X=x) = Var_\varepsilon[2 + 0.5z + x + zx + \varepsilon]$

$= Var_\varepsilon[\varepsilon] = \sigma_\varepsilon^2$

(c) $Y \mid X=x, Z=z \sim N(2 + 0.5z + x + zx, \sigma_\varepsilon^2)$

$Pr(Y=y \mid Z=z, X=x) = \frac{1}{\sqrt{2\pi}\sigma_\varepsilon} e^{-(y-2-0.5z-x-zx)^2/2\sigma_\varepsilon^2}$

(d) $E(Y \mid Z=1, X=x) - E[Y \mid Z=0, X=x] = 0.5 + x$

therefore $E_X\{E[Y \mid Z=1, X] - E[Y \mid Z=0, X]\} = 0.5 + EX$

$= 0.5.$

(e) $E[Y \mid Z=1] = E_{X,\varepsilon}[2 + 0.5 + X + X + \varepsilon]$

$= 2.5$

$E[Y \mid Z=0] = E_{X,\varepsilon}[2 + X + \varepsilon] = 2$

$\Rightarrow E[Y \mid Z=1] - E[Y \mid Z=0] = 0.5$

2:

(a): $E[Y|z,x] = 2+0.5z+x+zx = \beta_0+\beta_1 z+\beta_2 x$

$$\Rightarrow \begin{cases} z=1: & 2.5+2x = \beta_0+\beta_1+\beta_2 x \Rightarrow \beta_2 = 2 \\ z=0: & 2+x = \beta_0+\beta_2 x \Rightarrow \beta_2 = 1 \end{cases}$$

Contradiction

Therefore this model misspecified the distribution

(b) Then $2+0.5z+x+zx = \beta_0+\beta_1 z+\beta_2 zx+\beta_3 x^2$

$$\Rightarrow \begin{cases} z=1: & 2.5+2x = \beta_0+\beta_1+\beta_2 x+\beta_3 x^2 \Rightarrow \beta_3=0,\ \beta_2=2 \\ z=0: & 2+x = \beta_0+\beta_3 x^2 \text{ can't hold.} \end{cases}$$

So the model misspecified the distribution.

3:

We do $Y = 1+2A+3\sin(B)+5\sin(AB)+\varepsilon$

where A, B, ε mutal independent and $B, \varepsilon \sim N(0,1)$

Then since $N(0,1)$ is symmetric and $\sin(x)$ is odd function

we have $E\sin(B)=0$

Then $E(Y|A=1) - E(Y|A=0) = 2$

See the coding part.