201C HW1

John Friedman

April 9, 2024

1. 1: Suppose that X is a continuous random vector that has a density that is everywhere positive, and that ϵ conditional on $X = x$ is $N(\mu(x), \sigma(x)^2)$, where $\mu(x)$ and $\sigma(x)$ are continuous functions of x. Let $Y = g(X) + \epsilon$, where g is a continuous function of x.

   (a) Derive the distribution of Y conditional on $X = x$.

   Solution:

   $$\begin{aligned} f_{Y|X}(y|x) &= P(Y \leq y|X = x) \\ &= P(g(X) + \epsilon \leq y|X = x) \\ &= P(\epsilon \leq y - g(X)|X = x) \\ &= P(\frac{\epsilon - \mu(x)}{\sigma x} \leq \frac{y - g(X) - \mu(x)}{\sigma(x)}|X = x) \\ &= \Phi(\frac{y - g(x) - \mu(x)}{\sigma(x)}) \end{aligned}$$

   (b) Derive the conditional expectation and conditional variance of Y given $X = x$. Explain.

   Solution:

   $$\begin{aligned} E[Y|X = x] &= E[g(X) + \epsilon|X = x] \\ &= g(x) + E[\epsilon|X = x] \\ &= g(x) + \mu(x) \end{aligned}$$

   $$\begin{aligned} Var[Y|X = x] &= Var[g(X) + \epsilon|X = x] \\ &= Var[\epsilon|X = x] \\ &= \sigma(x)^2 \end{aligned}$$

   (c) Suppose that for all values of y, x, you are given an arbitrary conditional distribution function $F_{Y|X}(y)$. You are asked whether this could be the distribution corresponding to the model described above. Under what conditions on the function $F_{Y|X}(y)$ would you answer "yes", and under what conditions would you answer "no"?

   Solution:

   1. The distribution must be normally distributed for all x in the support of X.

2. The conditional mean of Y must be defined by a function $m(x)$ that is continuous in x. This is because both $\mu(x)$ and $g(x)$ are continuous functions of x, so the sum must be continuous.

3. The conditional variance of Y given $X = x$ must be continuous everywhere in x

(d) Is $\mu(x)$ identified? Is $\sigma(x)$ identified? Provide proofs.

Solution:

1. The conditional variance is identified from the observed moments of Y given $X = x$.

$$(\sigma(x))^2 = Var(Y|X = x)$$

2. Consider the composition $g(x)+\mu(x)$. From (a) we have an expression for the conditional distribution of Y given $X = x$ that uses $g(x) + \mu(x)$. If we can construct two different pairs $(g(x), \mu(x))$ and $(g'(x), \mu'(x))$ that yield the same conditional distribution of Y given $X = x$, then $\mu(x)$ is not identified.

   Consider $g'(x) = g(x) + c$ and $\mu'(x) = \mu(x) - c$. Then the conditional distribution of Y given $X = x$ is the same as the conditional distribution of Y given $X = x$ with $g(x)$ and $\mu(x)$.

2. Consider the model:

$$Y = \alpha^* g^*(X) + \epsilon$$

where $X \in R^K$ and $Y \in R$ are observable, $\epsilon \in R$ is unobservable, α^* is a constant, $g^* : R^K \to R$ is continuous, the support of X is R^K, and where ϵ is distributed independently of X with a $N(\mu^*, \sigma^{*2})$ distribution. Suppose that $\alpha^*, g^*, \mu^*, \sigma^{*2}$ are unknown.

(a) What is the conditional distribution of Y given $X = x$?

Solution:

$$\begin{aligned}
f_{Y|X}(y|x) &= P(Y \leq y|X = x) \\
&= P(\alpha^* g^*(X) + \epsilon \leq y|X = x) \\
&= P(\epsilon \leq y - \alpha^* g^*(X)|X = x) \\
&= P(\frac{\epsilon - \mu^*}{\sigma^*} \leq \frac{y - \alpha^* g^*(X) - \mu^*}{\sigma^*}|X = x) \\
&= \Phi(\frac{y - \alpha^* g^*(x) - \mu^*}{\sigma^*})
\end{aligned}$$

The conditional distribution is then $Y|X = x \sim N(\alpha^* g^*(x) + \mu^*, \sigma^{*2})$

(b) Is g^* identified within the set of continuous functions $g : R^K \to R$? Provide a proof of your answer.

Solution: g^* is not identified within the set of continuous functions. Consider any constant

$k > 0$. Let $(\alpha', g'(.)) = (c\alpha^*, c^{-1}g^*(.))$. Then:

$$\begin{aligned}F_{Y|X=x}(y;h') &= \Phi(\frac{y - \alpha'^* g'^*(x) - \mu^*}{\sigma^*}) \\ &= \Phi(\frac{y - c\alpha^* c^{-1} g^*(x) - \mu^*}{\sigma^*}) \\ &= \Phi(\frac{y - \alpha^* g^*(x) - \mu^*}{\sigma^*}) \\ &= F_{Y|X=x}(y;h)\end{aligned}$$

(c) Is $\mu*$ identified in the seat of real numbers? Provide a proof of your answer.

Solution: μ^* is not indentified within the set. Consider a fixed α^*, and $(g'(.), \mu') = (g^*(.) - \alpha^{*-1}k, \mu^* + k)$ where $k \in R$. Then:

$$\begin{aligned}F_{Y|X=x}(y;h') &= \Phi(\frac{y - \alpha^* g'(x) - \mu'}{\sigma^*}) \\ &= \Phi(\frac{y - \alpha^*(g^*(x) - \alpha^{*-1}k) - \mu^* - k}{\sigma^*}) \\ &= \Phi(\frac{y - \alpha^* g^*(x) - \mu^*}{\sigma^*}) \\ &= F_{Y|X=x}(y;h)\end{aligned}$$

$F_{Y|X=x}(y)$ is equivalent with h and h', so the function g is not defined.

(d) Is σ^{*2} identified in the set of positive real numbers? Provide a proof of your answer.

Solution: The variance σ^{*2} is identified $\in R^+$. Note that $Y = \alpha^* g^*(X) + \epsilon$ and $\epsilon \sim N(\mu^*, \sigma^{*2})$. Then the conditional variance:

$$\begin{aligned}Var(Y|X = x) &= Var(\alpha^* g^*(X) + \epsilon | X = x) \\ &= Var(\epsilon | X = x) \\ &= \sigma^{*2}\end{aligned}$$

(e) Suppose that $\alpha^* = 1$ and $\mu^* = 0$. Answer (b) and (d)

Solution: Both are identified.

$$\begin{aligned}F_{Y|X=x}(y;h) &= \Phi(\frac{y - \alpha^* g^*(x) - \mu^*}{\sigma^*}) \\ &= \Phi(\frac{y - g^*(x)}{\sigma^*})\end{aligned}$$

σ^{*2} is identified as in the previous part. g^* is identified as well using the conditional expectation of Y given $X = x$.

$$\begin{aligned}E[Y|X = x] &= E[g^*(X) + \epsilon | X = x] \\ &= g^*(x) + E[\epsilon | X = x] \\ &= g^*(x) + \mu^*\end{aligned}$$

(f) Suppose that for some value of $\bar{x}$ of $X, g(\bar{x}) = 0$. Are $\mu^*, \sigma^{*2}, \alpha^*$ and/or g^* identified? Provide proofs.

Solution: The conditional distribution of Y given $X = \bar{x}$ is $Y|X = \bar{x} \sim N(\mu^*, \sigma^{*2})$. We can identify σ^{*2} using the conditional variance as show in previous parts. We can identify μ^* using the conditional expectation of Y given $X = \bar{x}$.

We cant identify α^* or g^* using the same logic as in part (b)