HOMEWORK 6

Exercise are from Freedman.

7.A.12, 7.D (all, but turn in 2, 7, 13), 7.E (all, but turn in 6), Discussion Questions: 6, 7

Also, submit the "Class Laboratory" labelled `lab2.pdf`, answering the questions there. (It is on Canvas under "Pages/Class Laboratories").

Problem Freedman 7A12. Suppose X, Y, Z are independent normal random variables, each having variance 1. The means are $\alpha+\beta, \alpha+2\beta, 2\alpha+\beta$, respectively: α, β are parameters to be estimated. Show that maximum likelihood and OLS give the same estimates. Note: this won't usually be true – the result depends on the normality assumption.

Solution to Problem Freedman 7A12. If

$$\boldsymbol{D} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 2 & 1 \end{bmatrix},$$

then

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \boldsymbol{D} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} + \boldsymbol{\varepsilon},$$

where $\boldsymbol{\varepsilon} \sim N(\mathbf{0}, I_3)$. The OLS estimate is

$$\begin{bmatrix} \hat{\alpha} \\ \hat{\beta} \end{bmatrix} = (\boldsymbol{R}'\boldsymbol{R})^{-1}\boldsymbol{R}' \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \frac{1}{11} \begin{bmatrix} 1 & -4 & 7 \\ 1 & 7 & -4 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \frac{1}{11} \begin{bmatrix} X - 4Y + 7Z \\ X + 7Y - 4Z \end{bmatrix}.$$

On the other hand, we have

$$f(x, y, z; \alpha, \beta) = \frac{1}{(2\pi)^{3/2}} e^{-\frac{1}{2}[(x-(\alpha+\beta^2)+(y-(\alpha+2\beta))^2+(z-(2\alpha+\beta))^2]},$$

and so the log-likelihood is

$$\ell(\alpha, \beta) = -\frac{1}{2}\left[(X - (\alpha+\beta))^2 + (Y - (\alpha+2\beta))^2 + (Z - (2\alpha+\beta))^2\right] - \frac{3}{2}\log(2\pi).$$

Thus

$$\frac{\partial \ell}{\partial \alpha} = (X - (\alpha+\beta)) + (Y - (\alpha+2\beta)) + 2(Z - (2\alpha+\beta))$$
$$\frac{\partial \ell}{\partial \beta} = (X - (\alpha+\beta)) + 2(Y - (\alpha+2\beta)) + (Z - (2\alpha+\beta))$$

Setting to 0 and solving two linear equations for two unknowns yields

$$\hat{\alpha} = \frac{1}{11}(X - 4Y + 7Z)$$
$$\hat{\beta} = \frac{1}{11}(X + 7Y - 4Z)$$

□

Problem Freedman 7D2. Conversely, if U is uniform on $[0,1]$, show that $F^{-1}(U)$ has distribution function F.

Solution to Problem Freedman 7D2. Let $Y = F^{-1}(U)$. Then

$$F_Y(y) = \mathbb{P}(Y \le y) = \mathbb{P}(F^{-1}(U) \le y) = \mathbb{P}(U \le F(y)) = F_U((F(y)) = F(y)\,,$$

since $F(u) = u$ for $u \in (0,1)$. Thus, Y has distribution function F. $\square$

Problem Freedman 7D7. What is the distribution of $\log U - \log(1-U)$, where U is uniform on $[0,1]$?

Solution to Problem Freedman 7D7. Let $g(u) = \log(u/1-u)$. Then

$$g'(u) = \frac{1-u}{u}\frac{1}{(1-u)^2} = \frac{1}{u(1-u)} > 0$$

for $u \in (0,1)$; whence g is increasing in u. In particular, it is $1-1$ from $(0,1)$ to $(-\infty,\infty)$, with inverse

$$g^{-1}(x) = \frac{e^x}{1+e^x}\,.$$

Thus, the cdf $F_{g(U)}$ of $g(U)$ is given by

$$F_{g(U)}(t) = \mathbb{P}(g(U) \le t) = \mathbb{P}(U \le g^{-1}(t)) = F_U(g^{-1}(t)) = g^{-1}(t) = \frac{e^x}{1+e^x}\,.$$

$\square$

Problem Freedman 7D13. Show that the log likelihood for the probit model is concave, and strictly concave if X has full rank.

Solution to Problem Freedman 7D13. Let $\boldsymbol{x}_k$ be the k-th row of $\boldsymbol{X}$.

$$L(\boldsymbol{\beta} \mid \boldsymbol{y}, \boldsymbol{x}) = \prod_{k=1}^{n} \Phi(\boldsymbol{x}_k\boldsymbol{\beta})^{y_k}(1-\Phi(\boldsymbol{x}_k\boldsymbol{\beta})^{1-y_k})$$

$$\ell(\boldsymbol{\beta} \mid \boldsymbol{y}, \boldsymbol{x}) = \sum_{k=1}^{n} y_k \log\Phi(\boldsymbol{x}_k\boldsymbol{\beta}) + (1-y_k)\log(1-\Phi(\boldsymbol{x}_k\boldsymbol{\beta}))$$

Let $g = \log\Phi$ and $h = \log(1-\Phi)$; we know from Exercise 12 that $g'' < 0$ and $h'' < 0$ everywhere. Then

$$\begin{aligned}
\frac{\partial \ell}{\partial \beta_i}(\boldsymbol{\beta} \mid \boldsymbol{y}, \boldsymbol{x}) &= \sum_{k=1}^{n} y_k g'(\boldsymbol{x}_k\boldsymbol{\beta})x_{k,i} + (1-y_k)h'(\boldsymbol{x}_k\boldsymbol{\beta})x_{k,i} \\
\frac{\partial^2 \ell}{\partial \beta_j \partial \beta_i}(\boldsymbol{\beta} \mid \boldsymbol{y}, \boldsymbol{x}) &= \sum_{k=1}^{n} y_k g''(\boldsymbol{x}_k\boldsymbol{\beta})x_{k,i}x_{k,j} + (1-y_k)h''(\boldsymbol{x}_k\boldsymbol{\beta})x_{k,i}x_{k,j} \\
&= \sum_{k=1}^{n} x'_{i,k}\left(y_k g''(\boldsymbol{x}_k\boldsymbol{\beta}) + (1-y_k)h''(\boldsymbol{x}_k\boldsymbol{\beta})\right)x_{k,j} \\
&= -\boldsymbol{X}'\boldsymbol{D}\boldsymbol{X}\,,
\end{aligned}$$

where $\boldsymbol{D}$ is the diagonal matrix with k-th diagonal entry

$$d_k = y_k(-g''(\boldsymbol{x}_k\boldsymbol{\beta})) + (1-y_k)(-h''(\boldsymbol{x}_k\boldsymbol{\beta}))\,.$$

Note that $d_k > 0$ since it is impossible that both y_k and $1-y_k$ equal 0, and $-g''$ and $-h''$ are strictly positive everywhere.

Let $\boldsymbol{u}$ be any p column vector. Then

$$\boldsymbol{u}'\boldsymbol{X}\boldsymbol{D}\boldsymbol{X}\boldsymbol{u} = (\sqrt{\boldsymbol{D}}\boldsymbol{X}\boldsymbol{u})'(\sqrt{\boldsymbol{D}}\boldsymbol{X}\boldsymbol{u}) = \|\sqrt{\boldsymbol{D}}\boldsymbol{X}\boldsymbol{u}\|^2 \geq 0\,,$$

whence $-\ell''$ is non-negative definite. This implies that ℓ is concave.

Suppose now that $\boldsymbol{X}$ has full rank. Let $\boldsymbol{u} \neq 0$. Then $\boldsymbol{X}\boldsymbol{u} \neq 0$, since $\boldsymbol{X}$ has full rank. Since $\boldsymbol{D}$ is a strictly positive diagonal matrix, $\sqrt{\boldsymbol{D}}\boldsymbol{X}\boldsymbol{u} \neq 0$. We conclude that $-\boldsymbol{u}'\ell''\boldsymbol{u} > 0$, and $-\ell''$ is positive definite. This implies that ℓ is strictly concave. □

Problem Freedman 7E6. Student #77 is Presbyterian, went to public school, and graduated. What does this subject contribute to the likelihood function? Write your answer using ϕ in equation (15).

Solution to Problem Freedman 7E6. We have $C_i = 0$ and $Y_i = 1$. This student contributed the factor

$$\begin{aligned}\mathbb{P}(C_i = 0, Y_i = 1 \mid X_i, IsCat_i = 0) &= \mathbb{P}(U_i < -X_i b,\ V_i > -X_i\beta) \\ &= \iint_{\substack{u<-X_i b \\ v>-X_i\beta}} \phi(u,v)\,du dv\end{aligned}$$

to the likelihood function. □

Problem Freedman 7Dis6. Powers and Rock (1999) consider a two-equation model for the effect of coaching on SAT scores:

$$\begin{aligned} X_i &= \begin{cases} 1 & \text{if } U_i\alpha + \delta_i > 0 \\ 0 & \text{otherwise.}\end{cases} \\ Y_i &= cX_i + V_i\beta + \sigma\varepsilon_i\,. \end{aligned}$$

Here, $X_i = 1$ if subject i is coached, else $X_i = 0$. The response variable Y_i is subject i's SAT score; U_i and V_i are vectors of personal characteristics for subject i, treated as data. The latent variables $(\delta_i, \varepsilon_i)$ are IID bivariate normal with mean 0, variance 1, and correlation ρ; they are independent of the U's and V's. (In this problem, U and V are observable, δ and ε are latent.)

(a) Which parameter measures the effect of coaching? How would you estimate it?
(b) State the assumptions carefully (including a response schedule, if one is needed). Do you find the assumptions plausible?
(c) Why do Powers and Rock need two equations, and why do they need ρ?
(d) Why can they assume that the disturbance terms have variance 1?

Solution to Problem Freedman 7Dis6. (a) The parameter c measures the effect of coaching. The parameters can be estimated via maximum likelihood.

(b) The vector (U_i, V_i) needs to be independent of $(\delta_i, \varepsilon_i)$, the variable X_i is generated from the first equation, then put in the response schedule

$$Y_{i,x} = cx + V_i\beta + \sigma\varepsilon_i\,.$$

Why is the effect c of coaching the same for all individuals?

(c) X_i may be related to unmeasured variables represented in the error term. Thus ρ allows for correlation between X_i and the error ε_i.

(d) The parameter σ scales the error term in the second equation. Parameters are not identifiable in the second if δ_i can have arbitrary variance.

$\square$