240D Cameron Fall 2004
Department of Economics, U.C.–Davis

Final Exam: December 14 2004

Compulsory. Closed book. 2 hours. Worth 50% of course grade.
Read question carefully so you answer the question.

Question scores (total 50 points and 50% of course grade)

Question	1	2	3	4	5
Points	14	10	10	6	10

1. Estimation.
Consider the random variable y with density

$$f(y) = \frac{y^{\alpha-1}\exp(-y/\lambda)}{\lambda^{\alpha}\Gamma(\alpha)}, \quad y > 0, \alpha > 0, \lambda > 0,$$

where $\Gamma(\alpha)$ is the gamma function, i.e. $\Gamma(\alpha) = \int_0^\infty e^{-t}t^{\alpha-1}dt$. The derivative of the gamma function is the digamma function $\psi(\alpha)$, i.e. $\Gamma'(\alpha) = \psi(\alpha)$. Unless α is an integer there are no closed form solutions for $\Gamma(\alpha)$ and $\psi(\alpha)$, which are evaluated by numerical methods.
The first two moments of y can be shown to be

$$\begin{aligned} \mathrm{E}[y] &= \alpha\lambda \\ \mathrm{V}[y] &= \alpha\lambda^2. \end{aligned}$$

To form a regression model we suppose y_i given $\mathbf{x}_i$ has the density $f(y)$ given above, with

$$\begin{aligned} \lambda_i &= \exp(\mathbf{x}_i'\boldsymbol{\beta}) \\ \alpha &= \alpha. \end{aligned}$$

(a) Give the log-likelihood function for $\boldsymbol{\beta}$ and α.

(b) Give the first-order conditions for the MLE of $\boldsymbol{\beta}$ and α.

(c) Give the complete formula for an estimate of the variance matrix of the MLE of $\boldsymbol{\beta}$ and α that is consistent assuming that the density is correctly specified.
[Hint: Think carefully. There are several possible valid estimates. Give one that is easy to derive].

(d) Suppose the density is misspecified, but it is still the case that $\mathrm{E}[y_i|\mathbf{x}_i] = \alpha\lambda_i$ and $\mathrm{V}[y_i|\mathbf{x}_i] = \alpha\lambda_i^2$. Will the part (b) estimates of $\boldsymbol{\beta}$ and α still be consistent? Explain.

(e) Suppose α is known. Say $\alpha = 1$ for simplicity. Give a way to obtain a consistent estimator of $\boldsymbol{\beta}$ based on the conditional moment condition $\mathrm{E}[y_i|\mathbf{x}_i] = \exp(\mathbf{x}_i'\boldsymbol{\beta})$.

(f) Suppose α is known. Say $\alpha = 1$ for simplicity. Suppose also that there is an endogeneity problem, so that $\mathrm{E}[y_i|\mathbf{x}_i] \neq \exp(\mathbf{x}_i'\boldsymbol{\beta})$. Instead there are m instruments $\mathbf{z}_i$, where $m > k = \dim(\mathbf{x}_i)$, such that $\mathrm{E}[y_i - \exp(\mathbf{x}_i'\boldsymbol{\beta})|\mathbf{z}_i] = \mathbf{0}$. Give the objective function for a consistent estimator of $\boldsymbol{\beta}$.

2. Discrete and multinomial choice.
(a) Consider a linear regression model based on the latent variable

$$y_i^* = \mathbf{x}_i'\boldsymbol{\beta} + u_i,$$

where the errors u_i are logistic distributed with density $f(u) = e^{-u}/(1+e^{-u})^2$ and cdf $F(u) = 1/(1+e^{-u})$. Suppose we observe only

$$y_i = \begin{cases} 1 & \text{if } y_i^* > 0 \\ 0 & \text{if } y_i^* \leq 0. \end{cases}$$

(a) Find $\Pr[y_i = 1|\mathbf{x}_i]$.

(b) Provide details for ML estimation of $\boldsymbol{\beta}$.

(c) Suppose that we observe three mutually exclusive outcomes, with utility of outcome j for individual i given by

$$U_{ij} = \mathbf{x}_{ij}'\boldsymbol{\beta} + \varepsilon_{ij},$$

where ε_{ij} are iid extreme value distributed. We observe $y_i = j$

$$y_i = j \quad \text{if } U_{ij} \geq U_{ik},\ k = 1, 2, 3,$$

and define the indicator variables

$$y_{ij} = \begin{cases} 1 & \text{if } y_i = j \\ 0 & \text{if } y_i \neq j. \end{cases}$$

Give the log-likelihood function for this model.
[Hint: There is no need to formally derive the choice probabilities (and it will take a long time). You should be able to state these from memory].

(d) State in a few lines a well-known weakness of the model of part (c).

3. Censoring, sample selection and truncation.
Consider the latent variable model

$$\begin{aligned} y_1^* &= \mathbf{x}_1'\boldsymbol{\beta}_1 + \varepsilon_1 \\ y_2^* &= \mathbf{x}_2'\boldsymbol{\beta}_2 + \varepsilon_2, \end{aligned}$$

where we observe

$$y_1 = \begin{cases} 1 & \text{if } y_1^* > 0 \\ 0 & \text{if } y_1^* \leq 0, \end{cases}$$

and

$$y_2 = \begin{cases} y_2^* & \text{if } y_1^* > 0 \\ - & \text{if } y_1^* \leq 0. \end{cases}$$

(a) Find a general expression for $\mathrm{E}[y_2|\mathbf{x}, y_1^* > 0]$, where $\mathbf{x}$ is composed of the distinct elements of $\mathbf{x}_1$ and $\mathbf{x}_2$.

(b) Specialize this expression when $\varepsilon_2 = \delta\varepsilon_1 + \xi$, where ξ is independent of ε_1 and has mean zero.

(c) If ε_1 is standard normal distributed, specialize further.
[Hint: For $z \sim \mathcal{N}[0,1]$: $\mathrm{E}[z|z > c] = \phi(c)/[1 - \Phi(c)] = \phi(-c)/\Phi(-c)$].

(d) Hence state how to consistently estimate $\boldsymbol{\beta}_2$ given only probit and OLS regression commands.

4. Hypothesis testing
Stata estimation of the probit model for regression of y = dwork on an intercept and the scalar regressor x = age (with coefficient β) yields the following output.
. probit dwork age
Probit estimates Number of obs = 4486
LR chi2(1) = 3.36
Prob > chi2 = 0.0666
Log likelihood = -2400.7413 Pseudo R2 = 0.0007

dwork	Coef.	Std. Err.	z	P>|z|
age	.00684	.00373	1.83	0.067
_cons	.48677	.1444	3.37	0.001

(a) Test $H_0 : \beta = 0$ against $H_a : \beta \neq 0$ at level 0.05 using a Wald test.
If there is not enough information to do so then state this and also state what extra things you would do to perform a Wald test.

(b) Test $H_0 : \beta = 0$ against $H_a : \beta \neq 0$ at level 0.05using a likelihood ratio test.
If there is not enough information to do so then state this and also state what extra things you would do to perform an LR test.

(c) Test $H_0 : \beta = 0$ against $H_a : \beta \neq 0$ at level 0.05 using a score (or LM) test.
If there is not enough information to do so then state this and also state what extra things you would do to perform a score test.

5. Consider the linear panel data model

$$y_{it} = \alpha_i + \mathbf{x}'_{it}\boldsymbol{\beta} + \varepsilon_{it}, \quad i = 1, ..., N, \ t = 1, ..., T,$$

where $\boldsymbol{\beta}$ are parameters to be estimated, α_i, $i = 1, ..., N$ are individual specific effects, ε_{it} are iid $[0, \sigma_\varepsilon^2]$ errors and $N \to \infty$ while T is small.

(a) Use an appropriate differencing transformation applied to the y_{it} equation that leads to the within estimator and give the formula for the within estimator.

(b) Explain the Hausman test for fixed versus random effects.
State clearly the null and alternative hypotheses, the test statistic and the critical region.

(c) How, if at all, will your answer in (b) will change if ε_{it} is not i.i.d $(0, \sigma_\varepsilon^2)$. Give details.

(d) Stack the regression model appropriately and show that premultiplication by the matrix $\mathbf{Q} = \mathbf{I}_T - \frac{1}{T}\mathbf{e}\mathbf{e}'$ leads to elimination of α_i.