Solution to Exercise 11.22

$\star$**11.22** Explain how to use the DCAR (11.42) to test the IIA assumption for the conditional logit model (11.36). This involves testing it against the nested logit model (11.40) with the $\boldsymbol{\beta}^j$ constrained to be the same. Do this for the special case in which $J = 2$, $A_1 = \{0, 1\}$, $A_2 = \{2\}$. **Hint:** Use the results proved in the preceding exercise.

Recall that, in general, the DCAR may be written as

$$\Pi_{tj}^{-1/2}(\boldsymbol{\theta})\big(d_{tj} - \Pi_{tj}(\boldsymbol{\theta})\big) = \Pi_{tj}^{-1/2}(\boldsymbol{\theta})\boldsymbol{T}_{tj}(\boldsymbol{\theta})\boldsymbol{b} + \text{residual}, \tag{11.42}$$

for $t = 1, \ldots, n$ and $j = 0, \ldots, J$. Here, the vector $\boldsymbol{\theta}$, intended to denote all the parameters of the alternative hypothesis, is to be replaced by a vector whose first k components are those of the vector $\boldsymbol{\beta}$ in the conditional logit model (11.36) and whose last m components are the θ_k of the nested logit model (11.40). In this particular case, $m = 2$.

For the test of the conditional logit model, we must specify all the ingredients of regression (11.42) for that model. It is to be understood that we wish to test the specification (11.36) against the alternative specification (11.40), where we impose the constraints of the conditional logit model, that is, we require $\boldsymbol{\beta}^j = \boldsymbol{\beta}$ for all $j = 0, \ldots, J$. Thus, under the null hypothesis, we have

$$\Pi_{tj} = \frac{\exp(\boldsymbol{W}_{tj}\boldsymbol{\beta})}{\sum_{l=0}^{J}\exp(\boldsymbol{W}_{tl}\boldsymbol{\beta})}.$$

The derivatives with respect to the k components of $\boldsymbol{\beta}$ are obtained by summing the derivatives

$$\frac{\partial \Pi_{tj}}{\partial \boldsymbol{\beta}^l} = \Pi_{tj}\boldsymbol{W}_{tl}(\delta_{jl} - \Pi_{tl}),$$

which were derived in the solution to Exercise 11.21, over $l = 0, \ldots, J$ and setting all the $\boldsymbol{\beta}^l$ equal to the common $\boldsymbol{\beta}$. We obtain, for $h = 1, \ldots, k$, that

$$\frac{\partial \Pi_{tj}}{\partial \beta_h} = \Pi_{tj}\sum_{l=0}^{J}(\boldsymbol{W}_{tl})_h(\delta_{jl} - \Pi_{tl}),$$

where $(\boldsymbol{W}_{tl})_h$ denotes the h^{th} element of $\boldsymbol{W}_{tl}$. Since there are just two subsets of outcomes, the index i of the A_i takes on just two values, 1 and 2. For $j = 0, 1$, we have that $i(j) = 1$, and so

$$\frac{\partial \Pi_{tj}}{\partial \theta_1} = \Pi_{tj}(v_{tj} - \Pi_{t0}v_{t0} - \Pi_{t1}v_{t1}) \;\text{ and }\; \frac{\partial \Pi_{tj}}{\partial \theta_2} = -\Pi_{tj}\Pi_{t2}v_{t2}.$$

For $j = 2$, $i(j) = 2$, so that

$$\frac{\partial \Pi_{t2}}{\partial \theta_1} = -\Pi_{t2}(\Pi_{t0}v_{t0} + \Pi_{t1}v_{t1}) \;\text{ and }\; \frac{\partial \Pi_{t2}}{\partial \theta_2} = \Pi_{t2}(v_{t2} - \Pi_{t2}v_{t2}).$$

The inclusive values, h_{ti}, are given by (11.39). With a common parameter vector $\boldsymbol{\beta}$ and with $\theta_k = 1$, they are

$$h_{t1} = \log\Big(\exp(\boldsymbol{W}_{t0}\boldsymbol{\beta}) + \exp(\boldsymbol{W}_{t1}\boldsymbol{\beta})\Big) \;\text{ and }\; h_{t2} = \log\exp(\boldsymbol{W}_{t2}\boldsymbol{\beta}) = \boldsymbol{W}_{t2}\boldsymbol{\beta}.$$

Thus the quantities v_{tj}, $j = 0, 1, 2$, are

$$v_{t0} = h_{t1} - \boldsymbol{W}_{t0}\boldsymbol{\beta}, \quad v_{t1} = h_{t1} - \boldsymbol{W}_{t1}\boldsymbol{\beta}, \quad\text{and}\quad v_{t2} = h_{t2} - \boldsymbol{W}_{t2}\boldsymbol{\beta} = 0.$$

This last result, that $v_{t2} = 0$, implies that $\partial \Pi_{tj}/\partial \theta_2 = 0$ for $j = 0, 1, 2$. As we could have suspected, the fact that the second group is a singleton means that θ_2 cannot be identified, and so we cannot test its value.

Thus, in this case, the artificial regression (11.42) has $3n$ observations and just one testing regressor. For observation t, the regressand is

$$\begin{bmatrix} \Pi_{t0}^{-1/2}(d_{t0} - \Pi_{t0}) \\ \Pi_{t1}^{-1/2}(d_{t1} - \Pi_{t1}) \\ \Pi_{t2}^{-1/2}(d_{t2} - \Pi_{t2}) \end{bmatrix},$$

the regressors that correspond to $\boldsymbol{\beta}$ are

$$\begin{bmatrix} \Pi_{t0}^{1/2}\big(\boldsymbol{W}_{t0}(1 - \Pi_{t0}) - \boldsymbol{W}_{t1}\Pi_{t1} - \boldsymbol{W}_{t2}\Pi_{t2}\big) \\ \Pi_{t1}^{1/2}\big(-\boldsymbol{W}_{t0}\Pi_{t0} + \boldsymbol{W}_{t1}(1 - \Pi_{t1}) - \boldsymbol{W}_{t2}\Pi_{t2}\big) \\ \Pi_{t2}^{1/2}\big(-\boldsymbol{W}_{t0}\Pi_{t0} - \boldsymbol{W}_{t1}\Pi_{t1} + \boldsymbol{W}_{t2}(1 - \Pi_{t2})\big) \end{bmatrix},$$

and the regressor that corresponds to θ_1 is

$$\begin{bmatrix} \Pi_{t0}^{1/2}\big(v_{t0}(1 - \Pi_{t0}) - v_{t1}\Pi_{t1}\big) \\ \Pi_{t1}^{1/2}\big(-v_{t0}\Pi_{t0} + v_{t1}(1 - \Pi_{t1})\big) \\ \Pi_{t2}^{1/2}\big(-v_{t0}\Pi_{t0} - v_{t1}\Pi_{t1}\big) \end{bmatrix}.$$

Of course, all of the Π_{tj} and v_{tj} here are to be evaluated at the ML estimates of the conditional logit model. The easiest test statistic is the explained sum of squares, which should be distributed as $\chi^2(1)$. The t statistic on the testing regressor is *not* suitable, unless it is multiplied by the standard error of the regression, in which case it is exactly equal to the square root of the explained sum of squares.