INSTITUTO SUPERIOR TÉCNICO

Universidade do Porto FEUP Faculdade de Engenharia

PhD Program in Transportation
Transport Demand Modeling
Written Exam
12th of January 2011

Notes:

The total duration 1h30m; each question should be answered in a different paper sheet

The questions are to be answered without any type of consultation; there is no need for a calculation machine

1. A developer who specializes in summer cottage properties is considering purchasing several lots of land. The developer wants to forecast the value of each lot. From previous experience, she knows that the most important factors affecting the price of the lot are size ('000 m^2), number of mature trees, distance to road infrastructure (km) and distance to closest town (km). She gathers the relevant data for 60 recently sold lots in the region where she wants to buy the properties. A multiple regression analysis was performed and the following (incomplete) statistics were obtained.

SUMMARY OUTPUT-MODEL 1

Regression Statistics	
Multiple R	0.7845
R Square	0.6155
Adjusted R Square	0.5711
Standard Error	27.1230
Observations	60

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	3	73686	24562	30	1.15E-11
Residual	56	46038	822		
Total	59	119724			

	Coefficients	*Standard Error*	*Collinearity statistics*	
			Tolerance	*VIF*
Intercept	6.9495	3.5704	0.8940	1.1186
Lot size ('000 m2)	2.0950	0.2616	0.9550	1.0471
Trees	0.4644	0.1495	0.9130	1.0953
Distance to road infrastructure (km)	0.8772	0.4283	0.9880	1.0121

INSTITUTO SUPERIOR TÉCNICO

Universidade do Porto FEUP Faculdade de Engenharia

SUMMARY OUTPUT-MODEL 2

Regression Statistics	
Multiple R	0.779
R Square	0.607
Adjusted R Square	0.561
Standard Error	29.320
Observations	60

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	3	72656	24219	29	2.12E-11
Residual	56	47068	841		
Total	59	119724			

	Coefficients	*Standard Error*	*Collinearity statistics*	
			Tolerance	*VIF*
Intercept	22.4336	8.0531	0.9560	1.0460
Lot size ('000 m2)	2.2348	0.2794	0.8010	1.2484
Distance to road infrastructure (km)	1.3062	0.6543	0.2219	4.5057
Distance to closest town (km)	-0.1364	0.0737	0.1670	5.9880

a) Compare the two models and present your arguments to choose one of them. (3 points)

b) How do you interpret the value of the intercept calibrated in this model (explanatory or predictive meaning)? (2 points)

2. The 'Poisson' distribution and the 'Negative Binomial' distribution are probability distributions normally used in 'Transport Demand Modeling', such as in 'Accident Prediction Models'. Comment this statement, distinguishing both probability distributions applications (to 'Accident Prediction Models') and referring which of it tend to present 'over dispersion'. Explain carefully all your statements. (3 points)

3. Explain the main differences between the Fixed and Random Effects specification of panel data models, and the sampling procedures that underlie their main differences (theoretical and practical). You should also identify the type of studies more suited for each type of model (exploratory vs. prediction or forecasting) and indentify the available statistical tests to assess the most suitable specification of the model, presenting their null hypothesis. (4 points)

4. Explain the parallel slopes assumption mechanism and why is an important assumption in ordered response models. (3 points)

5. Suppose that for studying the modal split between Car, Bus and Subway of a given city a consulting company which was assigned that task has opted to conduct a stated preference survey in the region.

a) Supposing that the attributes which were used in the survey for characterizing the transport modes were: travel time, cost and distance, write down the utility functions of each mode considering different values of time and a special preference of mode (define all variables used). (2 points)

b) One of the municipality's technicians has criticized the estimated value of travel time of the car mode stating that it seems too large and that there are great differences amongst the population. He also added that the average household income should affect greatly this value. Redefine the utility functions in order to test the technician's hypothesis. Which estimation process would you use for this new model? (3 points).