42178 – Transportation System Analysis E21
Portfolio exercise II

These exercises will be part of the oral exam. Hence, you do not need to hand in your solution. However, if you would like feedback, you are free to upload your solutions by the 12/10. In this case, I will pick a random of the two exercises for each group and give feedback. This feedback is only for your own learning and does not influence the overall final grade.

In case you decide to hand in, the page limit of the full exercise is 12 pages. Feel free to put additional material in appendices but please keep essential results within the text so that your answers are understandable without appendices.

Exercise 1

Consider the data set "Trips2020". These data include the following variables

Variable	Description
$Trips$	Number of trips in a day
$Income$	Income in DKK
$Cars$	Number of cars in household
Age	Age
$Male$	Dummy indicating if the respondent is male
$Trips1$	Number of trips in a day censored at 5 trips

1) Analyse the data using descriptive statistics, e.g. summary statistics and a correlation matrix. What is the average number of trips using $Trips$ and $Trips1$? Also plot the distribution of trips using the $Trips1$ variable. Comment your results.

2) Assume you have estimated a Poisson model with the number of trips as dependent variable based on $Trips1$ with the following specification for the expected number of tours:

$$\ln(E(Trips1_n)) = k + \alpha_1\, ln(income_n + 1) + \alpha_2 Cars_n + \alpha_3 Age_n + \alpha_4 Male_n$$

where 1 is added to income to allow for ln(). We get the following results

Parameter	Estimate	Std. err.	z/t stat	Pr > |z|
k	1.001	0.06	16.8	<.0001
α_1	0.058	0.01	4.7	<.0001
α_2	0.043	0.02	2.8	0.004
α_3	-0.005	0.00	-4.5	<.0001

α_4	-0.099	0.04	-2.6	0.009

NoObs	1000
No. Parameters	5
Final LL	-1902.1
LL(c)	-1923.1

What can you say about the results and what is the expected number of trips for your sample according to the model? In addition, try to analyse the expected number of individuals with 0,1,2,3,4,5+ trips based on your model. Finally, use your estimated model to simulate the effect from a 20% increase in income on the number of trips. You can use either the income elasticity or simulation (the latter is similar to what you have to do for the logit model in step 5).

3) Specify a trip generation model based on cross-classification where you use at least two of the four variables in the Trips2020 data. You can use the quartiles of income to divide income in four intervals. Use the data to determine rates for different groups of individuals and use these rates when income is increased by 20% in a scenario. Discuss the difference between the cross-classification results and the results of the Poisson model.

4) On the basis of the data, suppose that you have estimated a multinomial logit model for the number of trips with the following utility specifications:

$$V_{i=0,n} = 0$$
$$V_{i=1,n} = k_1 + \beta_4 Age_n$$
$$V_{i=2,n} = k_2 + \beta_1 income_n + \beta_4 Age_n + \beta_6 Male_n$$
$$V_{i=3,n} = k_3 + \beta_1 income_n + \beta_3 Cars_n + \beta_5 Age_n + \beta_6 Male_n$$
$$V_{i=4,n} = k_4 + \beta_1 income_n + \beta_3 Cars_n + \beta_5 Age_n + \beta_6 Male_n$$
$$V_{i=5,n} = k_5 + \beta_2 income_n + \beta_3 Cars_n + \beta_5 Age_n + \beta_6 Male_n$$

and that you get the following results

Parameter	**Estimate**	**Std. Err.**	**T statistic**	**Approx. Pr > |t|**
k_1	-0.919	0.31	-3.0	0.003
k_2	1.440	0.28	5.1	<.0001
k_3	0.232	0.31	0.8	0.4501
k_4	1.011	0.30	3.4	0.0007
k_5	0.812	0.30	2.7	0.0071
β_1	0.001	0.00	2.8	0.0049
β_2	0.002	0.00	3.9	0.0001
β_3	0.213	0.08	2.6	0.0084
β_4	-0.011	0.01	-2.2	0.0290
β_5	-0.018	0.01	-3.6	0.0003

β_6	-0.586	0.18	-3.3	0.0012

NoObs	1000
No. Parameters	11
Final LL	-1591
LL(c)	-1614
LL(0)	-1792

Comment the results.

5) Apply the estimated model in 4) to the data. ❶
   a. Based on the estimated parameters, predict the total number of trips for the data.
   b. Analyse the expected number of individuals with 0,1,2,3,4,5+ trips based on your model. Compare to the sample shares.
   c. In addition, simulate the total number of trips when income is increased by 20%.❷ Based on this simulation, calculate the elasticity of the total number of trips with respect to income.

6) Discuss the difference between results from the logit and the Poisson model.

Hint ❶: Here you will need to simulate using the model based on the parameters from the previous step.
Hint ❷: You need to create a new dataset with income changed according to the specification. This new dataset is now your input to the simulation.

Exercise 2

You are asked to model trip distribution within the three central zones of Labtown. These are numbered 1, 2, 3. You have information on marginal totals, i.e. O_i, and D_j, both in the base year as well as the scenario year. In addition to this, you also have information on the generalised cost, gc_{ij}, and the initial matrix t_{ij} based on a sample of the trips in the base year.

O / D	Base		Scenario	
Zone	O_i	D_j	O_i	D_j
1	50	100	60	110
2	90	60	120	100
3	100	80	120	90

	Base			Scenario		
gc_{ij}	$j=1$	$j=2$	$j=3$	$j=1$	$j=2$	$j=3$
$i=1$	10	21	22	12	21	22
$i=2$	21	12	27	21	14	23
$i=3$	23	27	12	23	23	16

t_{ij}	$j=1$	$j=2$	$j=3$
$i=1$	7	2	4
$i=2$	8	9	6
$i=3$	10	5	11

1. Assume a model with independence among origins and destinations, i.e. $T_{ij} = A_i B_j O_i D_j$. Calculate the base matrix, T_{ij}^B, using this model. Do you think this is a useful model assuming independence?

2. Now consider a model based on the initial solution, i.e. $T_{ij} = A_i B_j O_i D_j t_{ij}$. Calculate the base matrix, T_{ij}^B, using this model. Do you think this is a useful model?

3. Now consider a gravity model, i.e. $T_{ij} = A_i B_j O_i D_j f(gc_{ij})$. A first question is what cost function to use. To find a suitable cost function, you have additional data on travel to/from zones 1, 2, and 3. These data are in the file: Distribution_data_pf2.xlsx.

One way to estimate a cost function is to regress $\ln(T)$ on generalised cost, gc. You are welcome to use other approaches. Following this you should calculate the base matrix, T_{ij}^{B}, using the gravity model. Do you think this is a useful model?

4. Assume one of your colleagues has found the correct trip matrix for the base year

T_{ij}^{C}	$j=1$	$j=2$	$j=3$
$i=1$	31	4	15
$i=2$	32	36	22
$i=3$	37	20	43

Try to evaluate the methods you applied in exercises 1-3 against the correct trip matrix. One way to compare them could be using the root mean square error (RMSE), i.e.

$$RMSE = \sqrt{\frac{1}{9}\sum_{i,j}\left(T_{ij}^{C} - \hat{T}_{ij}\right)^2}$$

Argue which of the three methods, you prefer considering that you in question 5 will be asked to investigate what happens in Labtown when a cross-city tunnel is opened between zones 2 and 3.

5. Use your preferred model from 1-3 to predict the number of trips in the future scenario where a cross-city tunnel is opened between zones 2 and 3 leading to lower generalised cost between these two zones. Note that some of the other generalised cost values have been increased due to rising congestion in the scenario. Comment on your results.