Methods in Transportation Econometrics and Statistics (Master)

Winter semester 2021/22, Solutions to Tutorial No. 8

Solution to Problem 8.1: Elasticities

(a) Exogenous variables: approach times T_{ni} and costs C_{ni}. Both are characteristica since they are attributes of the alternatives (airports), not the decision makers. The specification in the utilities is generic. This is appropriate for the problem at hand for both times and costs since the times relate to the approach by whatever mode, so there is no reason to assume different time sensitivities. For costs, a generic modelling is reasonable, anyway.

(b)
- β_1: Global preference BER vs. DRS. The marginally positive value $\hat{\beta}_1 = 0.195$ (although probably not significant for a sample of this size) means that, given the same airfares and approach times, BER is marginally preferred by 0.195 utility units (UU).
- β_2: Global preference FRA vs. DRS in UU. The higher preference is plausible since FRA, but not DRS, is a hub. So, the observed preference probably arises from the opportunity to use connecting flights in FRA, which is not modelled explicitely.
- β_3: Time sensitivity which, as expected, is negative.
- β_4: sensitivity to costs; negative, as expected.

Why are, compared to inner-city modal-split analyses, the absolute values of $\hat{\beta}_3$ and $\hat{\beta}_4$ smaller by one order of magnitude? The reason is the higher standard deviation of the random utility. Specifically, the time equivalent of one UU is $1/|\hat{\beta}_3| = 77\,\text{Minuten}$ while it is of the order of 5 minutes for inner-city trips. Since one UU denotes (up to a factor $\pi/\sqrt{6}$) the standard deviation of the random utility (see below), this means the random utility is greater by about a factor of 10. This is to be expected: Obviously, for a multi-hour journey, there are more acumulated uncertainties than for an inner-city trip.

(c)
- AC Berlin-Dresden in minutes: $-\beta_1/\beta_3 = 15.0$
- AC Berlin-Dresden in Euro: $-\beta_1/\beta_4 = 7.98$
- AC Frankfurt-Dresden in minutes: $-\beta_2/\beta_3 = 44.7$
- AC Frankfurt-Dresden in Euro: $-\beta_2/\beta_4 = 23.80$
- VoT in Euro/h: $60\beta_3/\beta_4 = 32\,\text{Euro/h}$
- Utility unit (UU) in minutes: $1\,\text{NE} = 1/\beta_3 = 77\,\text{Min}$
- Random utility standard deviation in minutes: $\sigma_\epsilon = -\pi/\sqrt{6}\,1/\beta_3 = 99\,\text{Min}$.

The VoT of 32 Euro/h appears comparatively elevated. However, one needs to bear in mind that the airfare includes the *round trip* while the approach time is for the single trip. Hence, for the complete journey, we have twice the approach time reducing the actual VoT to 16 €/h.

(d) Deterministic utilities:

$$V_{61} = -6.19, \quad V_{62} = -7.47, \quad V_{63} = -8.31.$$

Denominator:

$$\sum_{i=1}^{3} e^{V_{6i}} = 0.00287.$$

Choice probabilities:

Dresden: $P_{61} = 71.6\,\%$, Berlin: $P_{62} = 19.8\,\%$, Frankfurt: $P_{63} = 8.6\,\%$.

Observed percentaged frequencies:

Dresden: $h_{61} = 37/53 = 69.8\,\%$, Berlin: $h_{62} = 12/53 = 22.6\,\%$ Frankfurt: $h_{63} = 4/53 = 7.5\,\%$.

(e) Microscopic proper price elasticities:

$$\epsilon_{nii}^{(\text{mic,C})} = \frac{C_{ni}}{P_{ni}} \frac{\partial P_{ni}}{\partial C_{ni}} = \beta_4 C_{ni} \left(1 - P_{ni}\right).$$

For group 6:

$$\begin{aligned}
\epsilon_{611}^{(\text{mic,C})} &= \beta_4 C_{61} \left(1 - P_{61}\right) = -1.4, \\
\epsilon_{622}^{(\text{mic,C})} &= \beta_4 C_{62} \left(1 - P_{62}\right) = -4.9, \\
\epsilon_{633}^{(\text{mic,C})} &= \beta_4 C_{63} \left(1 - P_{63}\right) = -6.7
\end{aligned}$$

Similarly, the microscopic proper approach-time elasticities:

$$\begin{aligned}
\epsilon_{611}^{(\text{mic,T})} &= \beta_3 T_{61} \left(1 - P_{61}\right) = -0.37, \\
\epsilon_{622}^{(\text{mic,T})} &= \beta_3 T_{62} \left(1 - P_{62}\right) = -1.3, \\
\epsilon_{633}^{(\text{mic,T})} &= \beta_3 T_{63} \left(1 - P_{63}\right) = -1.4
\end{aligned}$$

(f) The microscopic cross elasticities ϵ_{nij} denotes the increase of the booking at airport i in percent if airport $j \neq i$ raises its airfares by 1 %: $\epsilon_{nij} = -\hat{\beta}_4 K_{nj} P_{nj}$. For person group $n = 6$, we obtain, if FRA ($j = 3$) changes its fares:

$$\epsilon_{613}^{(\text{mic,C})} = \epsilon_{623}^{(\text{mic,C})} = -\beta_4 C_{63} P_{63} = 0.63$$

The cross elasticity is positive: If FRA increases its fares, the demand increases in BER *and* DRS by 0.63 %. This is due to the IIA property: The relative preference of DRS over BER does not change if a third airport (FRA) changes its conditions. In order that the ratio of the bookings at DRS and BER remains constant, both need to obtain a share from the former FRA customers which is proportional to the customers these airports *already have.*

Sum condition explicit for the MNL:

$$\begin{aligned}\sum_i P_{ni}\epsilon_{ni3} &= -\beta_4 P_{n1} C_{n3} P_{n3}) - \beta_4 P_{n2} C_{n3} P_{n3}) + \beta_4 P_{n3} C_{n3}(1 - P_{n3}) \\ &= \beta_4 C_{n3} P_{n3}(-P_{n1} - P_{n2} + 1 - P_{n3}) \\ &= 0 \qquad\qquad [\text{da} \sum_i P_{ni} = 1].\end{aligned}$$

Sum condition generally from the definition of the elasticities (e.g., price easticities):

$$\sum_i P_{ni}\epsilon_{ni3} = \sum_i P_{ni} \frac{C_{n3}}{P_{ni}} \frac{\partial P_{ni}}{\partial C_{n3}} = C_{n3} \frac{\mathrm{d}}{\mathrm{d}C_{n3}} \left(\sum_i P_{ni} \right) = 0$$

since the sum of the choice probabilities =1 (because the choice set must be complete).

Intuitively, the sum relation reflects a zero-sum game. Because of the completeness of the alternative set, what is added to one alternative must be taken from another one.

(g) General expression for the macroscopic price proper elasticity:

$$\epsilon_{ii}^{(\mathrm{mac,C})} = \frac{C_i}{N_i} \frac{\partial N_i}{\partial C_i}. \tag{1}$$

- $N_i = \sum_{n=1}^N P_{ni}$: Estimated total demand for flights at airport i if there are N potential customers.
- $C_i = \sum_n C_{ni}$ sum of the airfares at airport i

and, after inserting the choice probabilities,

$$\epsilon_{ii}^{(\mathrm{mac,C})} = \frac{C_i}{N_i} \sum_n \frac{\partial P_{ni}}{\partial C_i} \tag{2}$$

Case 1: all prices are changed by a fixed *absolute* amount

The, we have $\mathrm{d}C_i = N\ \mathrm{d}C_{ni}$, hence

$$\frac{\partial P_{ni}}{\partial C_i} = \frac{\partial P_{ni}}{\partial C_{ni}} = \frac{P_{ni}}{C_{ni}} \epsilon_{nii}$$

and finally

$$\epsilon_{ii}^{(\mathrm{mac,abs,C})} = \frac{C_i}{N_i} \sum_n \frac{P_{ni}}{C_{ni}} \epsilon_{nii} \tag{3}$$

Case 2: All prices are changed by a fixed *relative* amount

Then, we have for each customer $\mathrm{d}C_{ni}/C_{ni} = \mathrm{d}C_i/C_i$, hence

$$\frac{\partial P_{ni}}{\partial C_i} = \frac{C_{ni}}{C_i}\frac{\partial P_{ni}}{\partial C_{ni}} = \frac{P_{ni}}{C_i}\epsilon_{nii}$$

and finally

$$\epsilon_{ii}^{(\mathrm{mac,rel,C})} = \frac{1}{N_i}\sum_n P_{ni}\epsilon_{nii} \tag{4}$$

This can be written as a weighted mean of the micro elasticities:

$$\epsilon_{ii}^{(\mathrm{mac,rel,C})} = \sum_n w_{ni}\epsilon_{nii}, \quad w_{ni} = \frac{P_{ni}}{N_i} = \frac{P_{ni}}{\sum_j P_{nj}}. \tag{5}$$

This means, the weight is given by the expected share that customer group n contributes to the total ticket demand N_i at this airport.

Notice that the macro elasticities are the quantities that are relevant for the airport managers while the micro elasticities can be directly calculated by the disrete-choice model.

Interpretation of the numerical values given for the macro-elasticities:

If either one of the three airports DRS, BER, and FRA increased its fares by 1 %, the consequences would be

- 2.6 % loss in DRS,
- 1.4 % loss in BER,
- and 1.3 % loss in FRA.

Notice that, if the macroscopic price elasticity is greater than (less negative than) -1, the airport managers will obtain a greater cash inflow when increasing the airfares.[1]

(h) Assuming the same choice probabilities in DRS and BER,

$$\frac{P_1}{P_2} = e^{V_1-V_2} = 1 \quad \Rightarrow \quad V_1 = V_2,$$

we can formulate a conditions for this equality in terms of times and costs at DRS:

$$\ln\left(\frac{P_1}{P_2}\right) = V_1 - V_2 = -\beta_1 + \beta_3(T_1 - T_2) + \beta_4(C_1 - C_2)$$

Solving for the approach time T_1 to DRS, we obtain

$$T_1(C_1) = \frac{\ln\left(\frac{P_1}{P_2}\right) + \beta_1 + \beta_3 T_2 - \beta_4(C_1 - C_2)}{\beta_3}$$

[1] Probably, higher *earnings* are also possible at a slightly more negative elasticity because of the reduced variable (per-customer) costs.

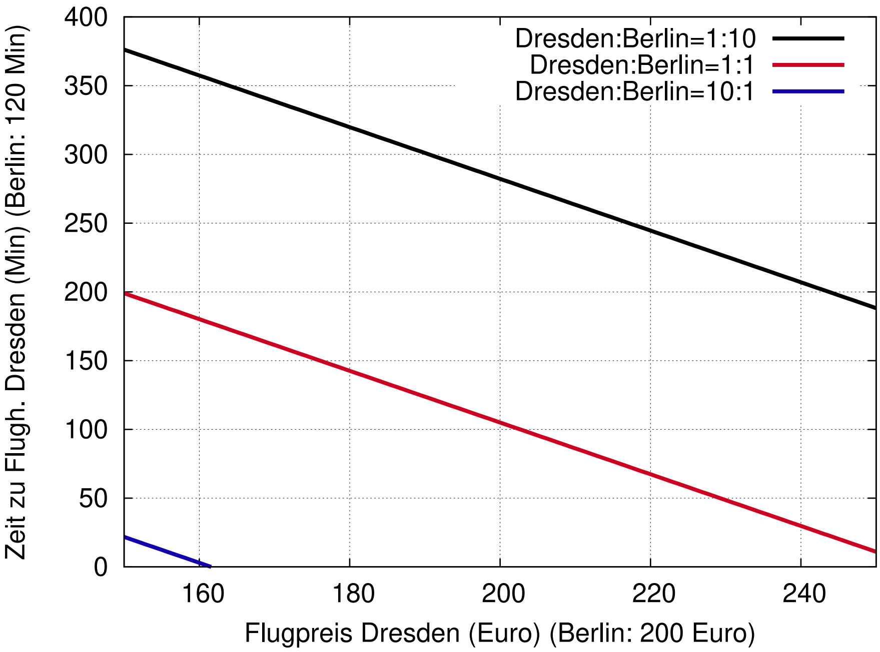
Particularly, at the same costs, we obtain for the approach times to DRS and BER the relation

$$T_1 = T_2 + \frac{\beta_1}{\beta_3} = T_2 - 15\,\mathrm{Min.}$$

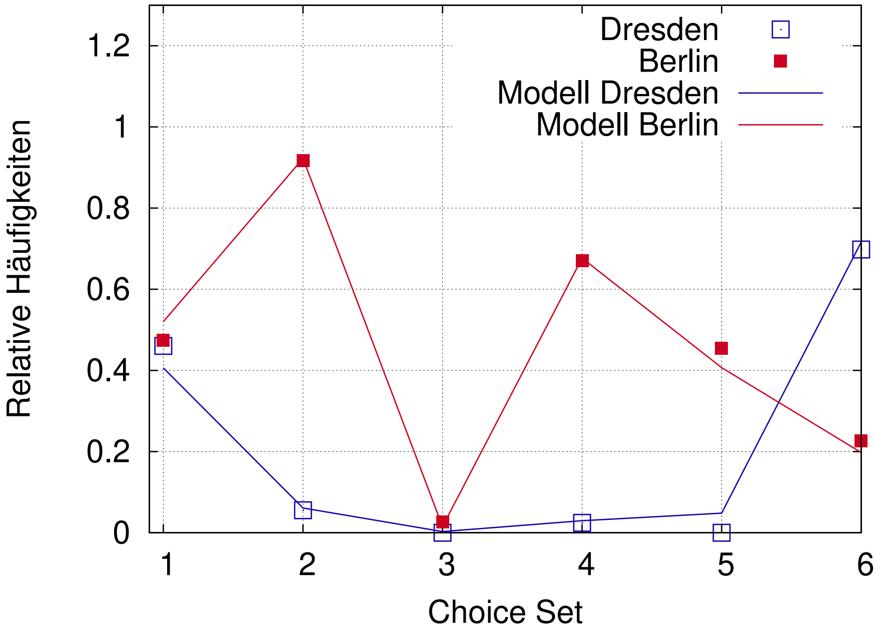
This difference is exactly the global *malus* of DRS with respect to BER, cf. Part (c). The slope of the curve represents the inverse of the implicit VoT, $-\beta_4/\beta_3 = -1.88\,\mathrm{Minuten/Euro}$.

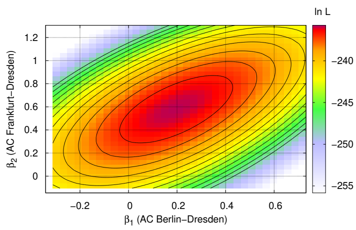
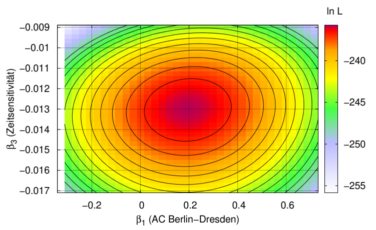
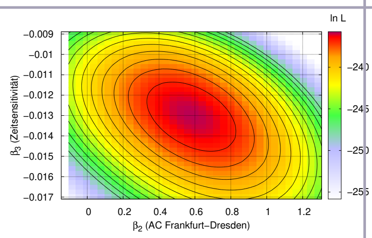
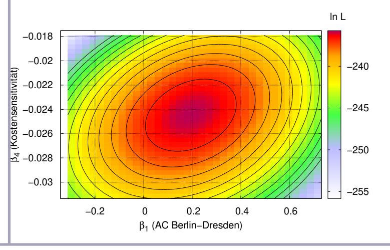
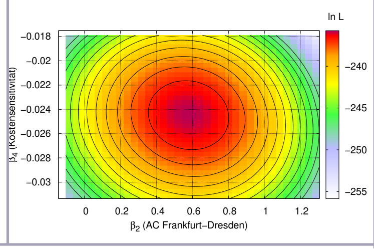
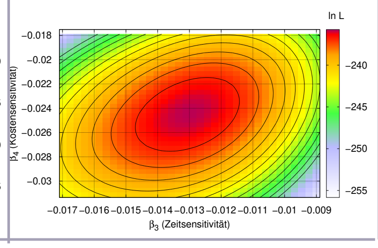
Specifically, if $T_2 = 120\,\mathrm{Min}$ and $C_2 = 200\,\mathrm{Euro}$, the indifference line for $P_1 = P_2$ is given by

$$T_1(C_1) = \frac{\beta_1}{\beta_3} + 120\,\mathrm{Min} - \frac{\beta_4}{\beta_1}(C_1 - 200\,\mathrm{Euro}).$$



Further results from the parameter estimation



L(.,.)	$\hat{\beta}_2$	$\hat{\beta}_4$	$\hat{\beta}_5$
$\hat{\beta}_4$			
$\hat{\beta}_5$			
$\hat{\beta}_6$			

Solution to Problem 8.2: Maximum-Likelihood-Method: the Basics

(a) **Binomial distribution**

(i) By definition, the binomial distribution $B(N, \theta)$ gives the number of "trues" in N independent binary true/false decisions ("Bernoulli experiment") provided that, for each experiment, the probability for "true"is constant= θ. In our case, the outcome "true" is just "alternative 1".

(ii) Finding the maximum of the second ($y = 2$) and sixth ($y = 6$) data line of the table gives $\hat{\theta} = 0.2$ and 0.6, respectively.[2]

(iii) Likelihood:

$$L(\theta|y) = \begin{pmatrix} 10 \\ y \end{pmatrix} \theta^y (1-\theta)^{10-y}$$

Log-Likelihood:

$$\tilde{L}(\theta|y) = \ln\left[\begin{pmatrix} 10 \\ y \end{pmatrix}\right] + y\ln\theta + (10-y)\ln(1-\theta)$$

Maximizing the not-logarithmized likelihood $L(\theta)$:

$$\begin{aligned}
\frac{\mathrm{d}L}{\mathrm{d}\theta} &= \begin{pmatrix} 10 \\ y \end{pmatrix} \left[y\theta^{y-1}(1-\theta)^{10-y} + \theta^y(10-y)(1-\theta)^{10-y-1}(-1)\right] \\
&= \begin{pmatrix} 10 \\ y \end{pmatrix} \theta^{y-1}(1-\theta)^{10-y-1}\left[y(1-\theta) - \theta(10-y)\right] \\
&\stackrel{!}{=} 0.
\end{aligned}$$

Since the prefactor before the prackets [...] is always > 0, the factor inside the brackets must be equal to zero:

$$[y(1-\theta) - \theta(10-y)] = 0 \ \Rightarrow \underline{\underline{\theta = \frac{y}{10}}}.$$

Maximizing the log-likelihood $\tilde{L}(\theta)$:

$$\frac{\mathrm{d}\tilde{L}(\theta)}{\mathrm{d}\theta} = \frac{y}{\theta} + \frac{10-y}{1-\theta}(-1) \stackrel{!}{=} 0,$$

i.e.,

$$y(1-\theta) = \theta(10-y) \Rightarrow \theta = \frac{y}{10}.$$

As expected, we otain the same result (and the verification that the maximum from the table lines are, in fact, exact).

[2]Of course, the true argument of the maximum may lay somewhere in between, particularly since the likelihood as a function of θ is non-symmetric. Interestingly, see part (iii), these values are exact.

(b) **Poisson distribution**

The Poisson distribution is valid if the probability of receiviung a call does not depend on history ("memory-less"), particularly not on the time difference to the last call. Thus, one needs several phones and the possibility to "hold the line" since, otherwise, calls are blocked during a consveration violating the requirement of no memory.

Furthermore, the two employees must receive different kinds of calls with different frequencies since, otherwise, the expectation value μ of the number of calls in a certain time is the same for both lines and one could average them to estimate μ instead of a separate esimation.

(i) Read off the table: $y = 2$ calls $\to$ 2nd data line of the table, $\hat{\mu} = 2$; $y = 4$ calls $\to$ 4nd data line, $\hat{\mu} = 4$;

(ii) Likelihood function:

$$L(\mu) = \frac{\mu^y e^{-\mu}}{y!}$$

Maximization:

$$\begin{aligned}
\frac{\mathrm{d}L}{\mathrm{d}\mu} &= \frac{\mathrm{d}}{\mathrm{d}\mu}\left(\frac{e^{y\ln\mu}e^{-\mu}}{y!}\right) \\
&= \frac{\mathrm{d}}{\mathrm{d}\mu}\left(\frac{e^{y\ln\mu-\mu}}{y!}\right) \\
&= \frac{1}{y!}e^{y\ln\mu-\mu}\left(\frac{y}{\mu}-1\right)
\end{aligned}$$

Since the prefactor of the brackets is always > 0, the factor in the brackets must vanish $\Rightarrow \underline{\underline{\mu = y}}$.

Alternatively by using the product rule of differentiation:

$$\begin{aligned}
\frac{\mathrm{d}L}{\mathrm{d}\mu} &= \frac{1}{y!}\left[y\mu^{y-1}e^{-\mu} + \mu^y(-1)e^{-\mu}\right] \\
&= \frac{\mu^{y-1}e^{-\mu}}{y!}\left[y-\mu\right]
\end{aligned}$$

(iii) Log-Likelihood:

$$\begin{aligned}
\tilde{L}(\mu) &= -\ln(y!) + y\ln\mu - \mu, \\
\frac{\mathrm{d}\tilde{L}(\mu)}{\mathrm{d}\mu} &= \frac{y}{\mu} - 1 \Rightarrow \underline{\underline{\mu = y}}.
\end{aligned}$$

(c) **Normal distribution**:

- Given: data of sales y_i in week i,
- Model: Gaussian distribution of the sales (central-limit theorem): $Y_i \sim N(\mu, \sigma^2)$.

- Likelihood-Funktion:
$$L(\mu, \sigma^2) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(y_i-\mu)^2}{2\sigma^2}}.$$

- Log-Likelihood-Funktion:
$$\begin{aligned}\tilde{L}(\mu, \sigma^2) = \ln\left(L(\mu, \sigma^2)\right) &= \sum_{i=1}^{n} \left(\frac{-\ln 2\pi - \ln\sigma^2}{2} - \frac{(y_i-\mu)^2}{2\sigma^2} \right) \\ &= \text{const.} - \frac{n}{2}\ln\sigma^2 - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i-\mu)^2\end{aligned}$$

- Setting zero the partial derivatives $\frac{\partial\tilde{L}}{\partial\mu}$ and $\frac{\partial\tilde{L}}{\partial\sigma^2}$ (watch out, we derive with respect to σ^2, not σ!):
$$\begin{aligned}\frac{\partial l}{\partial\mu} &= -\frac{1}{2}\sum_{i=1}^{n}\left(-2\frac{y_i-\mu}{\sigma^2}\right) \stackrel{!}{=} 0, \\ \frac{\partial l}{\partial\sigma^2} &= -\frac{n}{2\sigma^2} + \frac{1}{2}\sum_{i=1}^{n}\left(\frac{(y_i-\mu)^2}{\sigma^4}\right) \stackrel{!}{=} 0.\end{aligned}$$

  From $\frac{\partial\tilde{L}}{\partial\mu}$, we obtain
$$\sum_{i=1}^{n} y_i = n\mu \ \Rightarrow \hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} y_i = \underline{\underline{\bar{y}}}$$

  and from $\frac{\partial\tilde{L}}{\partial\sigma^2}$:
$$n = \sum_{i=1}^{n}\frac{(y_i-\mu)^2}{\sigma^2} \Rightarrow \hat{\sigma}^2 = \underline{\underline{\frac{1}{n}\sum_{i=1}^{n}(y_i-\mu)^2)}}.$$

- compare the result with the least-squared errors estimate for the trivial regression model:
$$y = \mu + \epsilon, \ \epsilon \sim N(0, \sigma^2)$$

  The OLS estimate for μ gives
$$\hat{\mu} = \bar{y} \quad \Rightarrow \text{the same result!}$$

  The OLS estimator for the variance is equal to the variance of the random term since no exogenous variables are included, thus all of the variance is the unexplained variance. We obtain
$$\hat{\sigma}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(y_i-\bar{y})^2.$$

  Here, the ML estimator and the unbiaed OLS estimator are *not* identical unless the expectation μ is known a-priori (in this unrealistic case, $\bar{y}$ is replaced by the true μ and $n-1$ by n for an unbiased estimate).