Methods in Transportation Econometrics and Statistics (Master)

Winter semester 2023/24, Solutions to Tutorial No. 6

Solution to Problem 6.1: Estimating heteroskedastic data: vehicle heading

(a) The GPS sensor, the steering angle sensor, and the gyrosensor are completely independent from each other since they use completely different methods to determine the heading. Thus, their measuring errors are independent.

(b) Assuming as unit for the angle one degree, we have the three variances

$$\sigma_1^2 = \sigma_2^2 = 4, \quad \sigma_3^2 = 16 = 4\sigma_1^2. \tag{1}$$

The variance of a linear combination of two random variables Y_1 and Y_2 is given by (a and b are real numbers)

$$V(aY_1 + bY_2) = a^2 V(Y_1) + b^2 V(Y_2) + 2ab \, \mathrm{Cov}(Y_1, Y_2)$$

For independent variables, we have zero correlation and thus zero covariance $\mathrm{Cov}(Y_1, Y_2)$. Generalizing this to a sum of three variables, we have

$$V(\bar{Y}) = V\left(\frac{Y_1 + Y_2 + Y_3}{3}\right) = \sigma_1^2 \frac{1+1+4}{9} = \frac{24}{9} = 2.67, \sigma = \sqrt{V(\bar{Y})} = \underline{\underline{1.63}}. \tag{2}$$

(c) Since all three measurements are unbiased, we have $E(Y_i) = \mu$, so, a generalized estimator

$$\hat{\mu}(w) = w(Y_1 + Y_2) + (1 - 2w)Y_3$$

is unbiased as well. (For $w = 1/3$ we will obtain the original estimator $\hat{\mu}(1/3) = \bar{Y}$)

With the same variance formula, we obtain

$$V(\hat{\mu}(w)) = w^2(\sigma_1^2 + \sigma_2^2) + (1-2w)^2\sigma_3^2 = \sigma_1^2\left(2w^2 + 4(1-2w)^2\right) = \sigma_1^2(2w^2 + 4 - 16w + 16w^2)$$

A necessary condition for a minimum is setting the first derivative equal to zero (it is a minimum since th quadratic term has apositive prefactor):

$$V'(w) = \sigma_1^2(36w - 16) \stackrel{!}{=} 0 \;\Rightarrow\; w_{\text{opt}} = \frac{4}{9}. \tag{3}$$

So, the efficient estimator (that with the lowest variance) is given by

$$\hat{\mu}_{\text{eff}} = \frac{4}{9}(Y_1 + Y_2) + \frac{1}{9}Y_3 (w_1)_{\text{opt}} = (w_2) \tag{4}$$

The variance and standard deviation of the best estimator is given by

$$V(\hat{\mu}_{\text{eff}} = \frac{16}{81} * 8 + \frac{1}{81} * 16 = \frac{16}{9}, \quad \sqrt{V(\hat{\mu}_{\text{eff}}} = \underline{\underline{\frac{4}{3}}}. \tag{5}$$

Summary:

- Using only one of the two estimators with the lowest variance gives an error of 2 degrees
- Using the arithmetic means of all three gives an error of $\sqrt{24/9} = 1.63$ degrees degrees
- Leaving out the inaccurate GPS-based estimator and averaging the rest gives an error of $\sqrt{2} = 1.41$ degrees
- using all three of them with the optimal weighting gives an error of $\sqrt{16/9} = 4/3 = 1.33$ degrees.

(d) We have now n unbiased measurements Y_i with variances σ_i^2, respectively. How to determine the weightings w_i such that $\sum_i w_i = 1$ (otherwise, the estimator would be biased)?

We want to minimize the variance

$$V(\vec{w}) = \sum_i w_i^2 \sigma_i^2 \quad \text{s.t.} \quad \sum_i w_i = 1$$

with respect to the weighting vector $\vec{w} = (w_1, ..., w_n)'$ (the standard abbreviation "s.t." means "subject to").

Using the technique of the Lagrange multipliers,

(i) we write the restraint in the form $g(\vec{w}) = \sum_i w_i - 1 = 0$ (if there are several restraints, every restraint is written in this form),

(ii) we derive both the function to be minimized and the restraint with respect to $\vec{w}$, multiply the derived restraint with the Lagrange multiplicator λ and set the sum to zero:

$$\frac{\partial V}{\partial w_k} + \lambda \frac{\partial g}{\partial w_k} = 2\sigma_k^2 w_k + \lambda = 0,$$

(iii) we determine the unknown Lagrange multiplicator by applying the result to the restraint. With general λ, we obtain

$$w_k = -\frac{\lambda}{2\sigma_k^2}$$

With the restraint $\sum_i w_i = 1$, we obtain

$$\lambda = \frac{-1}{\sum_i 1/(2\sigma_i^2)}$$

and

$$w_k = \frac{1/\sigma_k^2}{\sum_i 1/\sigma_i^2} \tag{6}$$

This means, the weighting is proportional to the inverse of the variance and the constant denominator only ensures $\sum_i w_i = 1$. Of course, for $n = 3$, $\sigma_1^2 = \sigma_2^2 = 4$, and $\sigma_3^2 = 16$, we obtain the weightings $w_1 = w_2 = 4/9$, $w_3 = 1/9$ already derived earlier.

Solution to Problem 6.2: Survey in the audience

(a)
- β_1: Alternative-specific constant (AC) for pedestrian or bike. A positive value means, there is an "a-priori bonus" for chosing a self-powered mode over the reference mode (motorized mode). Since in discrete-choice theory, only utility differences matter, we need a reference having no own AC. Hence, mode $i = 2$ has no AC. Both signs are plausible for β_1 although generally, motorized modes are favoured leading to $\beta_1 < 0$.
- β_2: Sensitivity to travel times. Since this is a necessary and not a trip for the purpose of itself (such as a sports cycling tour), the less total travel time it needs, the better. Hence, we expect $\beta_2 < 0$.
- β_3: Sensitivity to travel costs. Since everybody wants to spend less money rather than more *ceteris paribus*, $\beta_3 < 0$ is expected.

(b) We make the asymptotic assumption that (i) the estimation errors are Gaussian,[1] (ii) the estimated error variance is assumed to be the true variance. Then, the test statistics for the null hypothesis H_0: $\beta_j = 0$ is given by

$$T = \hat{\beta}_j/\sigma_j \sim N(0,1)$$

with the rejection zone R_α for $\alpha = 5\,\%$:

$$R_\alpha = \left\{t^{\text{data}} \in I\!R : |t^{\text{data}} > z_{1-\alpha/}\right\} = [-1.96, 1.96]$$

We obtain

$$\begin{aligned}
t_1^{\text{data}} &= \frac{|\hat{\beta}_1|}{\sigma_1} = -1.90/0.46 = -4.13 \quad \Rightarrow \ H_0 \text{ rejected,} \\
t_2^{\text{data}} &= \frac{|\hat{\beta}_2|}{\sigma_2} = -0.229/0.041 = -5.59 \quad \Rightarrow \ H_0 \text{ rejected,} \\
t_3^{\text{data}} &= \frac{|\hat{\beta}_3|}{\sigma_3} = -3.67/0.71 = -5.17 \quad \Rightarrow \ H_0 \text{ rejected.}
\end{aligned}$$

This means, all parameters are significantly different from zero.[2]

(c) Choice probabilities of the binary i.i.d. Probit model with random utilities $\epsilon_i \sim N(0,1)$:

$$P_1 = \Phi\left(\frac{V_1 - V_2}{\sqrt{2}}\right), \quad P_2 = 1 - P_1.$$

The deterministic utilities V_{ni} of the first three situations n =1,2, and 3 for the alternatives $i = 1, 2$, the argument z of the standard normal distribution, and the probabilities itself

[1] Notice that, for a Logit model, the estimation errors have a different distribution compared to the random utilities which are Gumbel distributed.

[2] In fact, all p values are below 10^{-4}, so all factors are highly significant.

are given by

$$\begin{aligned} V_{11} &= -8.77,\ V_{12} = -6.87, & \tfrac{V_1 - V_2}{\sqrt{2}} &= -1.34, & P_1 &= 0.089 \\ V_{21} &= -8.77,\ V_{22} = -9.16, & \tfrac{V_1 - V_2}{\sqrt{2}} &= -1.34, & P_1 &= 0.609 \\ V_{31} &= -8.77,\ V_{32} = -11.45, & \tfrac{V_1 - V_2}{\sqrt{2}} &= 0.27, & P_1 &= 0.971 \end{aligned}$$

(d) The binary Logit choice probabilities are

$$P_1 = \frac{e^{V_1}}{e^{V_1} + e^{V_2}} = \frac{1}{1 + e^{-(V_1 - V_2)}}, \quad P_2 = 1 - P_1.$$

Using the estimated Logit parameters $\hat{\beta}_1^{\mathrm{L}} = -2.42$, $\hat{\beta}_2^{\mathrm{L}} = -0.283$, and $\hat{\beta}_3^{\mathrm{L}} = -4.59$, we obtain

$$\begin{aligned} V_{11} &= -10.9,\ V_{12} = -8.5, & P_1 &= 0.082, \\ V_{21} &= -10.9,\ V_{22} = -11.3, & P_1 &= 0.601, \\ V_{31} &= -10.9,\ V_{32} = -14.2, & P_1 &= 0.962. \end{aligned}$$

The probabilities are nearly the same. This is to be expected since the only difference between the Logit and the i.i.d. Probit model is the form of the distribution of the random utilities (RUs). However, since the Gumbel distribution of the Logit RUs looks very similar to the Gaussian ditribution of the Probit model (although it is not symmetric and has a longer tail to the right), this does not make a significant difference except for very low choice probabilities.

(e) As already stated in (d), the form of the standard normal and Gumbel ditributions of the RUs of the Probit and Logit models, respectively, look similar. However, the standard deviation of the Gumbel distribution is $\pi/\sqrt{6}$ while that of the standard normal distribution is =1. As derived in the lecture, the choice probabilities of discrete choice models are translation invariant (not needed here) and scale invariant, i.e., they do not change when multiplying all the deterministic *and* random utilities by a common positive factor. So, in order to make the Probit parameters similar to that of the Logit parameters, the random utilities of the Probit model should have the standard deviation $\pi/\sqrt{6}$ instead of 1. A common multiplication by $\pi/\sqrt{6}$ realizes this. However, since the exogenous factors (travel times etc.) are given, this means, we need to multiply the Probit parameters by this factor:

$$\frac{\pi}{\sqrt{6}}\hat{\beta}_1^{\mathrm{P}} = -2.44, \quad \frac{\pi}{\sqrt{6}}\hat{\beta}_2^{\mathrm{P}} = -0.294, \quad \frac{\pi}{\sqrt{6}}\hat{\beta}_3^{\mathrm{P}} = -4.71$$

which are essentially the same values as the Logit estimates

$$\hat{\beta}_1^{\mathrm{L}} = -2.42, \quad \hat{\beta}_2^{\mathrm{L}} = -0.283, \quad \hat{\beta}_3^{\mathrm{L}} = -4.59.$$

This justifies (in a way) to use Logit instead of Probit models if there are more than two alternatives (multinomial case):

- The expressions for the multinomial Logit probabilities remain simple while that of the multinomial Probit, even if i.i.d., entail calculating a one-dimensional integral.[3]
- The central limit theorem, however, make Gaussian RUs more plausible compared to the somewhat ad-hoc introduced Gumbel distributions of the Logit model. Fortunately, there are hardly any differences except for very many alternatives or very small choice probabilities.

(f) Building fractions of parameters such as the implied value of time (VOT)

$$\text{VOT} = \frac{\hat{\beta}_2}{\hat{\beta}_3} \quad [€/\text{min}]$$

or

$$\text{VOT} = \frac{60\hat{\beta}_2}{\hat{\beta}_3} \quad [€/\text{h}],$$

means, multiplications of all parameters with a common value cancel out each other. From the results of (d) and (e) we therefore expect that we will obtain similar estimates for the implied VOT for both models. Indeed,

$$60\frac{\beta_2^{\text{P}}}{\beta_3^{\text{P}}} = 60\frac{-0.229}{-3.67} = 3.74\,\text{Eur/h}, \quad 60\frac{\beta_2^{\text{L}}}{\beta_3^{\text{L}}} = 60\frac{-0.283}{-4.59} = 3.70\,\text{Eur/h}.$$

[3] Even more complicated calculations are needed if we analyze a Probit model with correlated RUs but, in this case, it is also more general as the Logit model which always has i.i.d. RUs.

Solution to Problem 6.3: Questionnaire design for a conjoint analysis

(a) Survey questionnaire for the *full factorial design*:

$$\begin{aligned}
&\text{Choice Set 1:}(-20\,\text{min}, -1€),\\
&\text{Choice Set 2:}(-20\,\text{min}, 1€),\\
&\text{Choice Set 3:}(0\,\text{min}, -1€),\\
&\text{Choice Set 4:}(0\,\text{min}, 1€),\\
&\text{Choice Set 5:}(20\,\text{min}, -1€),\\
&\text{Choice Set 6:}(20\,\text{min}, 1€).
\end{aligned}$$

(b) In the three choice sets $(-20\,\text{min}, -1€)$, $(20\,\text{min}, -1€)$, and $(20\,\text{min}, 1€)$ given in the problem statement, the travel times and costs are positively correlated. Hence, this is not an orthogonal design.