Lecture 8: Logit and Probit Models

8.1 Logit Models: Definition

All Logit models are defined by **Gumbel-distributed** random utilities.

- The standard **Multinomial-Logit model (MNL)** has RUs distributed according to $\epsilon_i \sim \text{i.i.d Gumbel}(0, 1)$
- Distribution:

$$F_{\text{Gu}}^{(\eta,\lambda)}(x) = \exp\left[-e^{-\lambda(x-\eta)}\right]$$

- Density:

$$f_{\text{Gu}}^{(\eta,\lambda)}(x) = \frac{\mathrm{d}F_{\text{Gu}}^{(\eta,\lambda)}(x)}{\mathrm{d}x} = \lambda e^{-\lambda(x-\eta)} \exp\left[-e^{-\lambda(x-\eta)}\right].$$

- Statistical properties:

$$\epsilon_{\text{mode}} = \eta, \quad E(\epsilon) = \eta + \gamma/\lambda \text{ with } \gamma = 0.5772, \quad V(\epsilon) = \frac{\pi^2}{6\lambda^2}$$

8.1 Logit Models: Definition

All Logit models are defined by **Gumbel-distributed** random utilities.

- The standard **Multinomial-Logit model (MNL)** has RUs distributed according to $\epsilon_i \sim$ i.i.d Gumbel$(0, 1)$

- Distribution:

$$F_{\text{Gu}}^{(\eta,\lambda)}(x) = \exp\left[-e^{-\lambda(x-\eta)}\right]$$

- Density:

$$f_{\text{Gu}}^{(\eta,\lambda)}(x) = \frac{\mathrm{d}F_{\text{Gu}}^{(\eta,\lambda)}(x)}{\mathrm{d}x} = \lambda e^{-\lambda(x-\eta)} \exp\left[-e^{-\lambda(x-\eta)}\right].$$

- Statistical properties:

$$\epsilon_{\text{mode}} = \eta, \quad E(\epsilon) = \eta + \gamma/\lambda \text{ with } \gamma = 0.5772, \quad V(\epsilon) = \frac{\pi^2}{6\lambda^2}$$

8.1 Logit Models: Definition

All Logit models are defined by **Gumbel-distributed** random utilities.

- The standard **Multinomial-Logit model (MNL)** has RUs distributed according to $\epsilon_i \sim$ i.i.d Gumbel$(0,1)$
- Distribution:

$$F_{\text{Gu}}^{(\eta,\lambda)}(x) = \exp\left[-e^{-\lambda(x-\eta)}\right]$$

- Density:

$$f_{\text{Gu}}^{(\eta,\lambda)}(x) = \frac{\mathrm{d}F_{\text{Gu}}^{(\eta,\lambda)}(x)}{\mathrm{d}x} = \lambda e^{-\lambda(x-\eta)} \exp\left[-e^{-\lambda(x-\eta)}\right].$$

- Statistical properties:

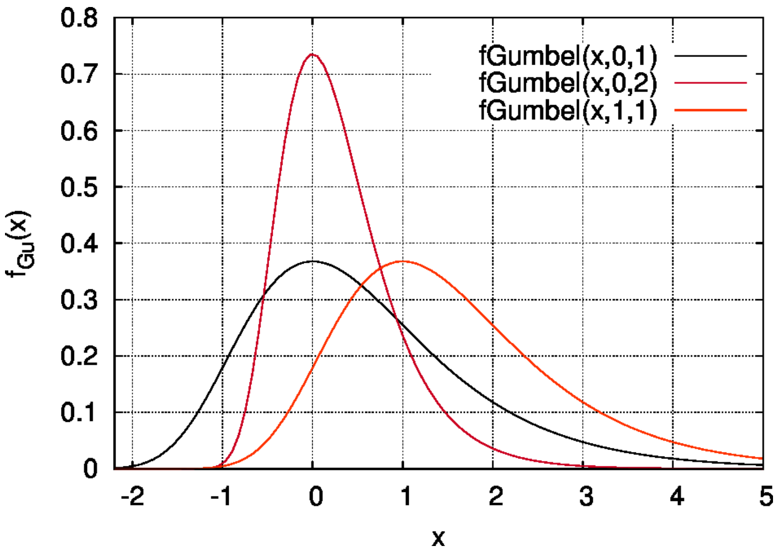
$$\epsilon_{\text{mode}} = \eta, \quad E(\epsilon) = \eta + \gamma/\lambda \text{ with } \gamma = 0.5772, \quad V(\epsilon) = \frac{\pi^2}{6\lambda^2}$$

8.1 Logit Models: Definition

All Logit models are defined by **Gumbel-distributed** random utilities.

- The standard **Multinomial-Logit model (MNL)** has RUs distributed according to $\epsilon_i \sim$ i.i.d Gumbel$(0, 1)$
- Distribution:
$$F_{\text{Gu}}^{(\eta,\lambda)}(x) = \exp\left[-e^{-\lambda(x-\eta)}\right]$$
- Density:
$$f_{\text{Gu}}^{(\eta,\lambda)}(x) = \frac{\mathrm{d}F_{\text{Gu}}^{(\eta,\lambda)}(x)}{\mathrm{d}x} = \lambda e^{-\lambda(x-\eta)} \exp\left[-e^{-\lambda(x-\eta)}\right].$$
- Statistical properties:
$$\epsilon_{\text{mode}} = \eta, \quad E(\epsilon) = \eta + \gamma/\lambda \text{ with } \gamma = 0.5772, \quad V(\epsilon) = \frac{\pi^2}{6\lambda^2}$$

Density functions of some Gumbel distributions



$\Rightarrow$ not symmetric; expectation $! = \eta$, particularly $E(\epsilon) = \gamma = 0.5772$ if $\epsilon \sim \text{Gu}(0, 1)$

Questions

? The numerical values of the deterministic utilities V_i are $\pi/\sqrt{6} \approx 1.28$ times as large as if the RU variance $V(\epsilon)$ were $= 1$. Why?

Because of the **scaling invariance** of discrete-choice models: The choice probability remains unchanged if *both* the random and deterministic utilities are multiplied by a factor $\lambda > 0$

? The nonzero $E(\epsilon_i) = 0.5772$ is irrelevant. Why?

This is due to the **translation invariance** of discrete-choice models: When adding a real-valued constant to the utilities of all alternatives, nothing changes. Here, a common $E(\epsilon_i) = 0.5772$ (remember, $\epsilon \sim i.i.d.$!) is just such a common constant.

Questions

? The numerical values of the deterministic utilities V_i are $\pi/\sqrt{6} \approx 1.28$ times as large as if the RU variance $V(\epsilon)$ were $= 1$. Why?
Because of the **scaling invariance** of discrete-choice models: The choice probability remains unchanged if *both* the random and deterministic utilities are multiplied by a factor $\lambda > 0$

? The nonzero $E(\epsilon_i) = 0.5772$ is irrelevant. Why?
This is due to the **translation invariance** of discrete-choice models: When adding a real-valued constant to the utilities of all alternatives, nothing changes. Here, a common $E(\epsilon_i) = 0.5772$ (remember, $\epsilon \sim i.i.d.!$) is just such a common constant.

Questions

? The numerical values of the deterministic utilities V_i are $\pi/\sqrt{6} \approx 1.28$ times as large as if the RU variance $V(\epsilon)$ were $= 1$. Why?
Because of the **scaling invariance** of discrete-choice models: The choice probability remains unchanged if *both* the random and deterministic utilities are multiplied by a factor $\lambda > 0$

? The nonzero $E(\epsilon_i) = 0.5772$ is irrelevant. Why?
This is due to the **translation invariance** of discrete-choice models: When adding a real-valued constant to the utilities of all alternatives, nothing changes. Here, a common $E(\epsilon_i) = 0.5772$ (remember, $\epsilon \sim i.i.d.!$) is just such a common constant.

Questions

? The numerical values of the deterministic utilities V_i are $\pi/\sqrt{6} \approx 1.28$ times as large as if the RU variance $V(\epsilon)$ were $= 1$. Why?
Because of the **scaling invariance** of discrete-choice models: The choice probability remains unchanged if *both* the random and deterministic utilities are multiplied by a factor $\lambda > 0$

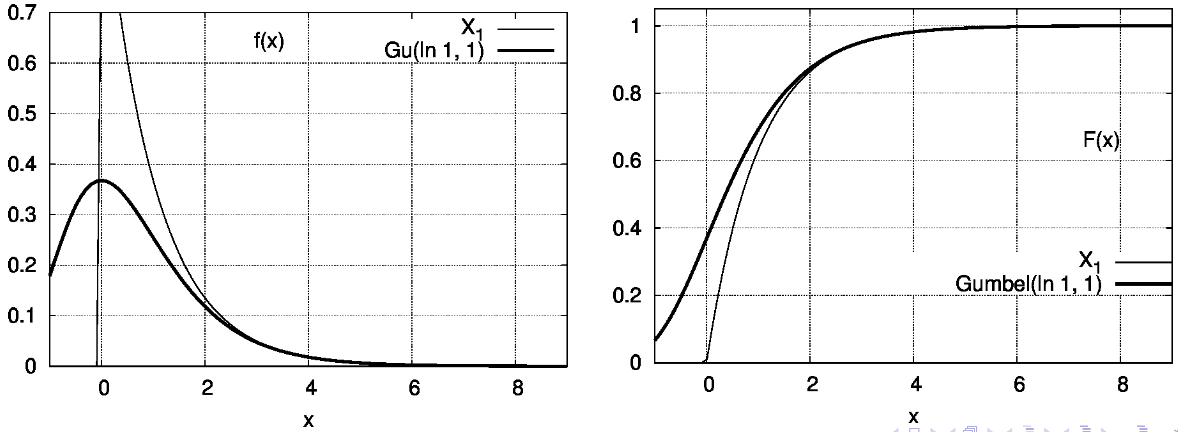
? The nonzero $E(\epsilon_i) = 0.5772$ is irrelevant. Why?
This is due to the **translation invariance** of discrete-choice models: When adding a real-valued constant to the utilities of all alternatives, nothing changes. Here, a common $E(\epsilon_i) = 0.5772$ (remember, $\epsilon \sim i.i.d.!$) is just such a common constant.

Gumbel distribution as a limit distribution for max(.)

The maximum of many i.i.d. random variables X_i with exponential tails $\propto \exp(-\lambda x)$ approaches a Gumbel or Generalized Extreme Value Type-I distribution:

$$\max(X_1, ..., X_n) \overset{\text{asympt.}}{\sim} \text{Gu}(\ln n, \lambda)$$

Example 1: Maximum of i.i.d. exponentially distributed RUs

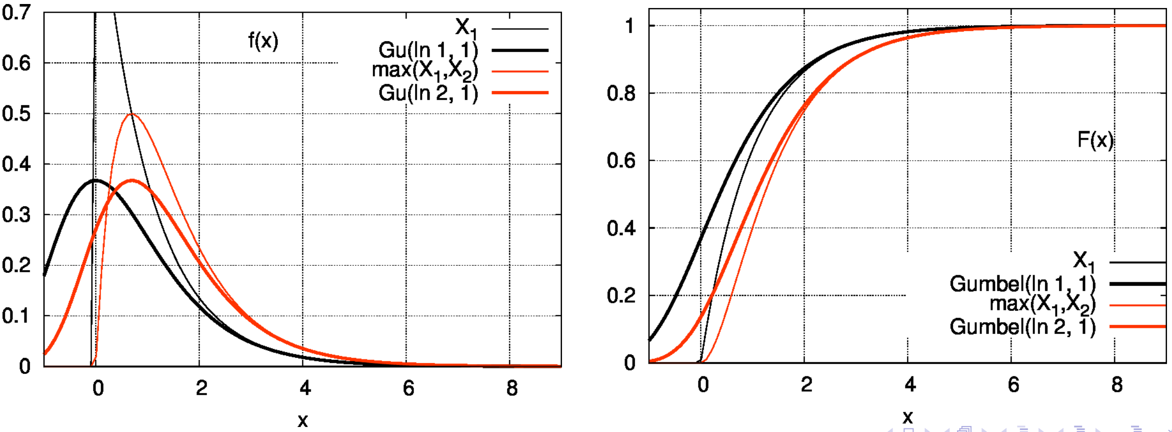


Gumbel distribution as a limit distribution for max(.)

The maximum of many i.i.d. random variables X_i with exponential tails $\propto \exp(-\lambda x)$ approaches a Gumbel or Generalized Extreme Value Type-I distribution:

$$\max(X_1, ..., X_n) \overset{\text{asympt.}}{\sim} \text{Gu}(\ln n, \lambda)$$

Example 1: Maximum of i.i.d. exponentially distributed RUs

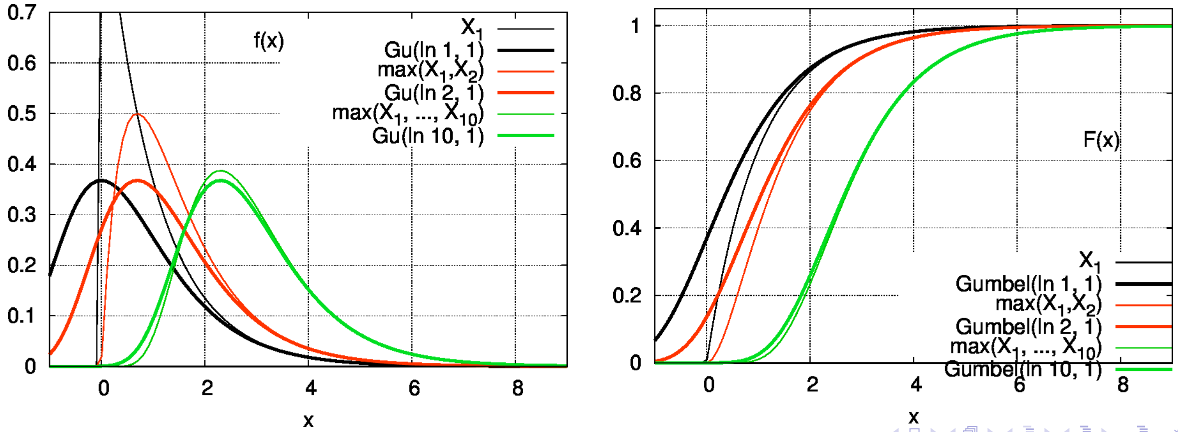


Gumbel distribution as a limit distribution for max(.)

The maximum of many i.i.d. random variables X_i with exponential tails $\propto \exp(-\lambda x)$ approaches a Gumbel or Generalized Extreme Value Type-I distribution:

$$\max(X_1, ..., X_n) \overset{\text{asympt.}}{\sim} \text{Gu}(\ln n, \lambda)$$

Example 1: Maximum of i.i.d. exponentially distributed RUs

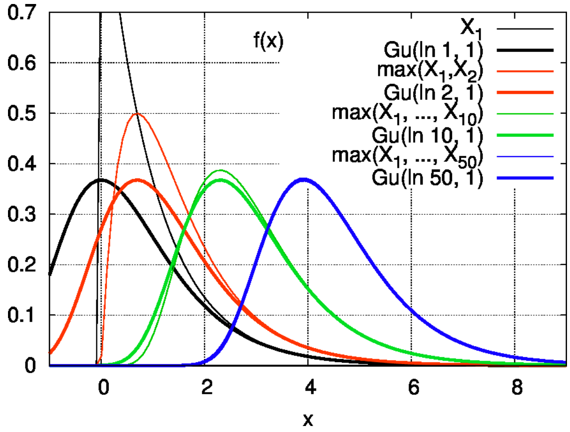


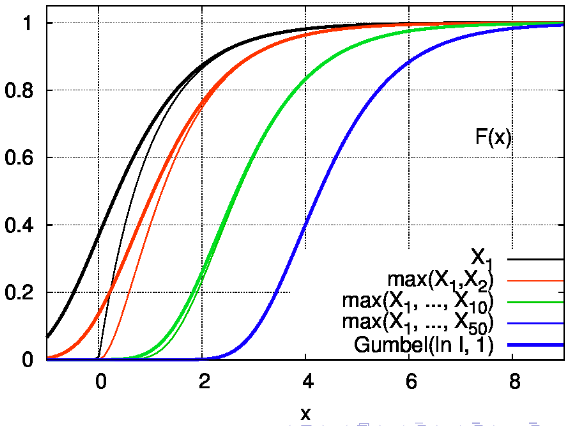
Gumbel distribution as a limit distribution for max(.)

The maximum of many i.i.d. random variables X_i with exponential tails $\propto \exp(-\lambda x)$ approaches a Gumbel or Generalized Extreme Value Type-I distribution:

$$\max(X_1, ..., X_n) \overset{\text{asympt.}}{\sim} \text{Gu}(\ln n, \lambda)$$

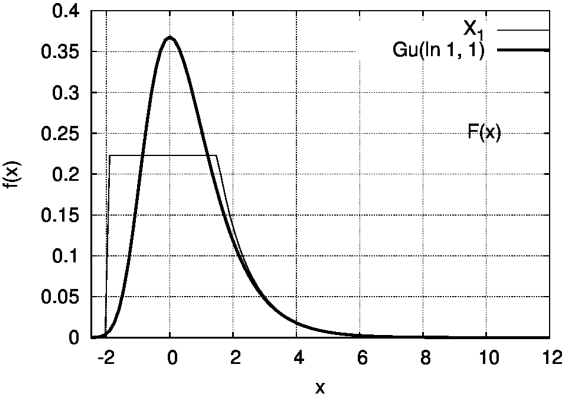
Example 1: Maximum of i.i.d. exponentially distributed RUs

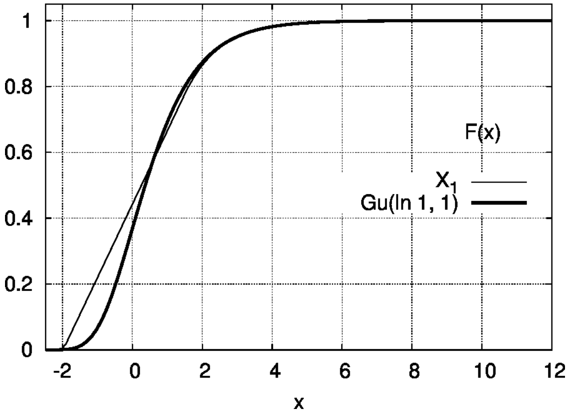




Gumbel distribution as a limit distribution for max(.)

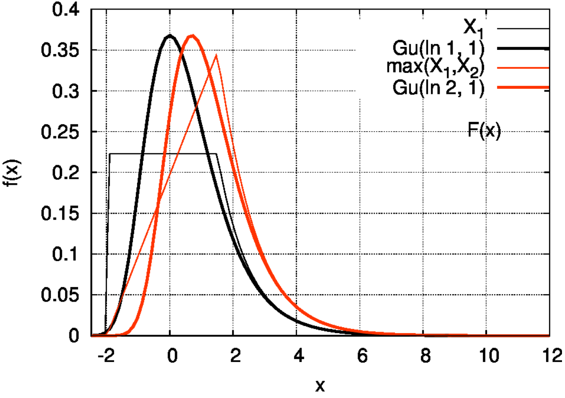
Example 2: Maximum of i.i.d. combined uniform-exponential RUs

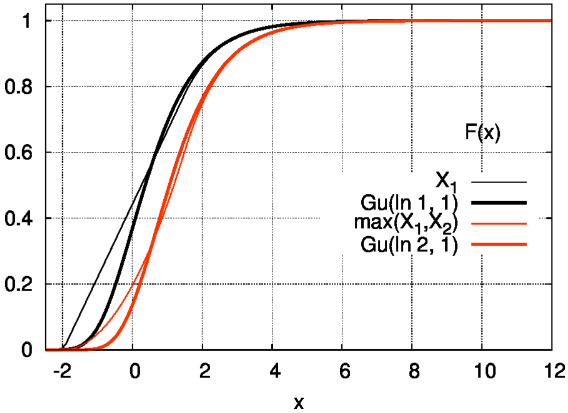




Gumbel distribution as a limit distribution for max(.)

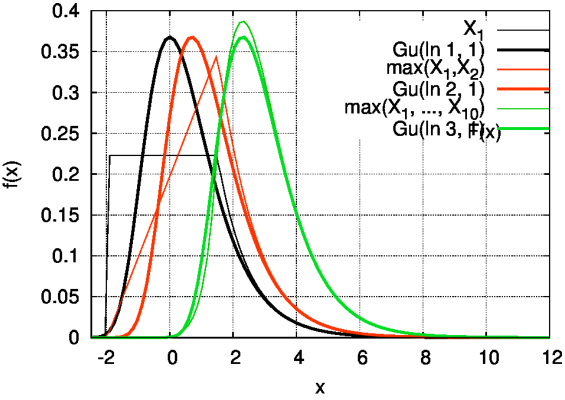
Example 2: Maximum of i.i.d. combined uniform-exponential RUs

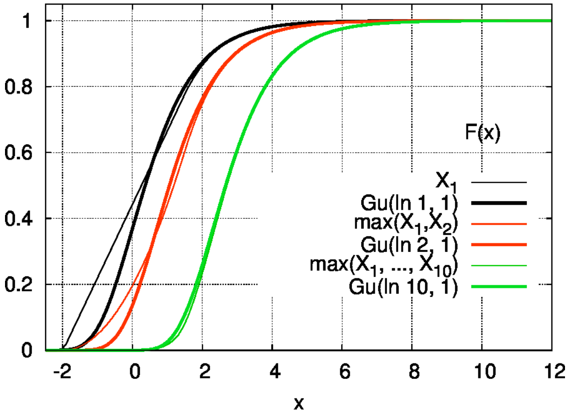




Gumbel distribution as a limit distribution for max(.)

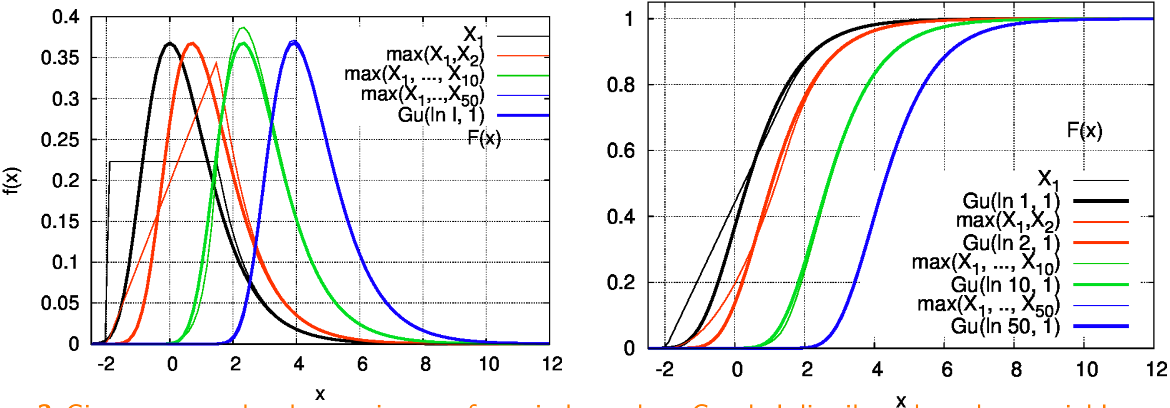
Example 2: Maximum of i.i.d. combined uniform-exponential RUs





Gumbel distribution as a limit distribution for max(.)

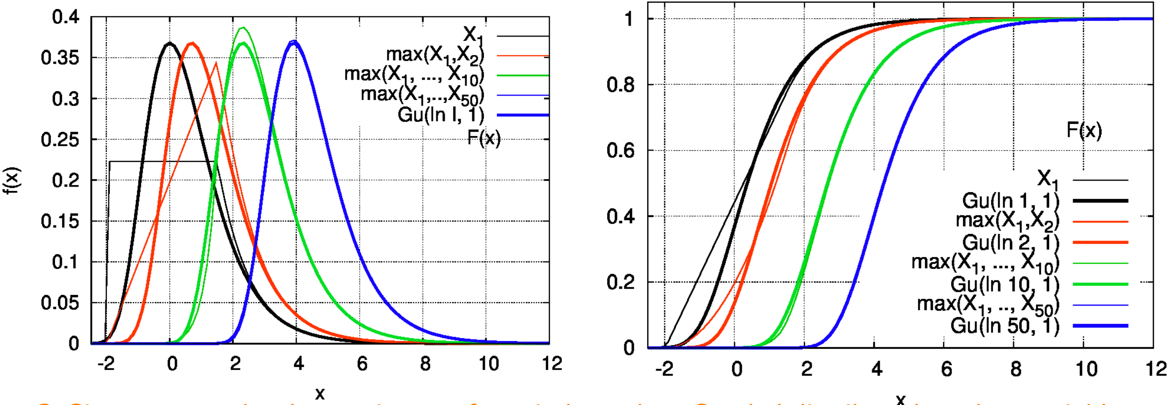
Example 2: Maximum of i.i.d. combined uniform-exponential RUs



? Give reasons why the maximum of two independent Gumbel distributed random variables of the same scale parameter is Gumbel distributed as well

Gumbel distribution as a limit distribution for max(.)

Example 2: Maximum of i.i.d. combined uniform-exponential RUs



? Give reasons why the maximum of two independent Gumbel distributed random variables of the same scale parameter is Gumbel distributed as well

Since $\max(\max(x_1, x_2), \max(x_3, x_4)) = \max(x_1, x_2, x_3, x_4)$

Properties of the Multinomial-Logit Model (MNL)

- Models of the Logit family (MNL, nested Logit, GEV models) are the only ones with explicit expressions for the choice probabilities for the multinomial case $I > 2$. For the MNL itself, we have

$$P_i^{\text{MNL}} = \frac{\exp(V_i)}{\sum_j \exp(V_j)}$$

- Besides the translational and scale invariance of all simple discrete-choice models, the MNL has the **Independence of Irrelevant Alternatives (IIA)** property:

> IIA property: The relative preference of Alternative i over j as defined by the choice probability ratio P_i/P_j does not depend on other alternatives $k \neq i, j$

- The IIA property is exclusively true for the MNL. In fact, the MNL can be equivalently defined by the IIA property instead of i.i.d. Gumbel RUs.

Properties of the Multinomial-Logit Model (MNL)

- Models of the Logit family (MNL, nested Logit, GEV models) are the only ones with explicit expressions for the choice probabilities for the multinomial case $I > 2$. For the MNL itself, we have

$$P_i^{\text{MNL}} = \frac{\exp(V_i)}{\sum_j \exp(V_j)}$$

- Besides the translational and scale invariance of all simple discrete-choice models, the MNL has the **Independence of Irrelevant Alternatives (IIA)** property:

> IIA property: The relative preference of Alternative i over j as defined by the choice probability ratio P_i/P_j does not depend on other alternatives $k \neq i, j$

- The IIA property is exclusively true for the MNL. In fact, the MNL can be equivalently defined by the IIA property instead of i.i.d. Gumbel RUs.

Properties of the Multinomial-Logit Model (MNL)

- Models of the Logit family (MNL, nested Logit, GEV models) are the only ones with explicit expressions for the choice probabilities for the multinomial case $I > 2$. For the MNL itself, we have

$$P_i^{\text{MNL}} = \frac{\exp(V_i)}{\sum_j \exp(V_j)}$$

- Besides the translational and scale invariance of all simple discrete-choice models, the MNL has the **Independence of Irrelevant Alternatives (IIA)** property:

> IIA property: The relative preference of Alternative i over j as defined by the choice probability ratio P_i/P_j does not depend on other alternatives $k \neq i, j$

- The IIA property is exclusively true for the MNL. In fact, the MNL can be equivalently defined by the IIA property instead of i.i.d. Gumbel RUs.

Questions

? Show that the MNL choice probabilities satisfy translational invariance.
Just devide the Logit choice probability formula by, e.g., $\exp(V_1)$:

$$P_i^{\text{MNL}} = \frac{\exp(V_i - V_1)}{\sum_j \exp(V_j - V_1)}$$

? For a comparison with another model, we want $V(\epsilon_i) = 1$ instead of $\pi^2/6$. In which way the model parameters must be changed?
This means, the standard deviation of ϵ is now given by 1 rather than by $\pi/\sqrt{6} \approx 1.28$, i.e., multiplied by $\lambda = \sqrt{6}/\pi$. The choice probabilities (no longer given by the Logit formula!) will be unchanged if the determinsitic utilities, i.e., the parameters, are multiplied by λ as well

? Derive the IIA property from the choice probability formula.
The IIA says that the relative preference of an alternative i over j, i.e. P_i/P_j, does not depend on any V_k, $k \neq i, j$. Just calculate this ratio:

$$\frac{P_i}{P_j} = \frac{\exp(V_i)}{\sum_k \exp(V_k)} \frac{\sum_l \exp(V_l)}{\exp(V_j)} = \exp(V_i - V_j)$$

Questions

? Show that the MNL choice probabilities satisfy translational invariance.
Just devide the Logit choice probability formula by, e.g., $\exp(V_1)$:

$$P_i^{\text{MNL}} = \frac{\exp(V_i - V_1)}{\sum_j \exp(V_j - V_1)}$$

? For a comparison with another model, we want $V(\epsilon_i) = 1$ instead of $\pi^2/6$. In which way the model parameters must be changed?
This means, the standard deviation of ϵ is now given by 1 rather than by $\pi/\sqrt{6} \approx 1.28$, i.e., multiplied by $\lambda = \sqrt{6}/\pi$. The choice probabilities (no longer given by the Logit formula!) will be unchanged if the determinsitic utilities, i.e., the parameters, are multiplied by λ as well

? Derive the IIA property from the choice probability formula.
The IIA says that the relative preference of an alternative i over j, i.e. P_i/P_j, does not depend on any V_k, $k \neq i, j$. Just calculate this ratio:

$$\frac{P_i}{P_j} = \frac{\exp(V_i)}{\sum_k \exp(V_k)} \frac{\sum_l \exp(V_l)}{\exp(V_j)} = \exp(V_i - V_j)$$

Questions

? Show that the MNL choice probabilities satisfy translational invariance.
Just devide the Logit choice probability formula by, e.g., $\exp(V_1)$:

$$P_i^{\text{MNL}} = \frac{\exp(V_i - V_1)}{\sum_j \exp(V_j - V_1)}$$

? For a comparison with another model, we want $V(\epsilon_i) = 1$ instead of $\pi^2/6$. In which way the model parameters must be changed?
This means, the standard deviation of ϵ is now given by 1 rather than by $\pi/\sqrt{6} \approx 1.28$, i.e., multiplied by $\lambda = \sqrt{6}/\pi$. The choice probabilities (no longer given by the Logit formula!) will be unchanged if the determinsitic utilities, i.e., the parameters, are multiplied by λ as well

? Derive the IIA property from the choice probability formula.
The IIA says that the relative preference of an alternative i over j, i.e. P_i/P_j, does not depend on any V_k, $k \neq i, j$. Just calculate this ratio:

$$\frac{P_i}{P_j} = \frac{\exp(V_i)}{\sum_k \exp(V_k)} \frac{\sum_l \exp(V_l)}{\exp(V_j)} = \exp(V_i - V_j)$$

Questions

? Show that the MNL choice probabilities satisfy translational invariance.
Just devide the Logit choice probability formula by, e.g., $\exp(V_1)$:

$$P_i^{\text{MNL}} = \frac{\exp(V_i - V_1)}{\sum_j \exp(V_j - V_1)}$$

? For a comparison with another model, we want $V(\epsilon_i) = 1$ instead of $\pi^2/6$. In which way the model parameters must be changed?
This means, the standard deviation of ϵ is now given by 1 rather than by $\pi/\sqrt{6} \approx 1.28$, i.e., multiplied by $\lambda = \sqrt{6}/\pi$. The choice probabilities (no longer given by the Logit formula!) will be unchanged if the determinsitic utilities, i.e., the parameters, are multiplied by λ as well

? Derive the IIA property from the choice probability formula.
The IIA says that the relative preference of an alternative i over j, i.e. P_i/P_j, does not depend on any V_k, $k \neq i, j$. Just calculate this ratio:

$$\frac{P_i}{P_j} = \frac{\exp(V_i)}{\sum_k \exp(V_k)} \frac{\sum_l \exp(V_l)}{\exp(V_j)} = \exp(V_i - V_j)$$

Questions

? Show that the MNL choice probabilities satisfy translational invariance.
Just devide the Logit choice probability formula by, e.g., $\exp(V_1)$:

$$P_i^{\text{MNL}} = \frac{\exp(V_i - V_1)}{\sum_j \exp(V_j - V_1)}$$

? For a comparison with another model, we want $V(\epsilon_i) = 1$ instead of $\pi^2/6$. In which way the model parameters must be changed?
This means, the standard deviation of ϵ is now given by 1 rather than by $\pi/\sqrt{6} \approx 1.28$, i.e., multiplied by $\lambda = \sqrt{6}/\pi$. The choice probabilities (no longer given by the Logit formula!) will be unchanged if the determinsitic utilities, i.e., the parameters, are multiplied by λ as well

? Derive the IIA property from the choice probability formula.
The IIA says that the relative preference of an alternative i over j, i.e. P_i/P_j, does not depend on any V_k, $k \neq i, j$. Just calculate this ratio:

$$\frac{P_i}{P_j} = \frac{\exp(V_i)}{\sum_k \exp(V_k)} \frac{\sum_l \exp(V_l)}{\exp(V_j)} = \exp(V_i - V_j)$$

Questions

? Show that the MNL choice probabilities satisfy translational invariance.
Just devide the Logit choice probability formula by, e.g., $\exp(V_1)$:

$$P_i^{\text{MNL}} = \frac{\exp(V_i - V_1)}{\sum_j \exp(V_j - V_1)}$$ ✓

? For a comparison with another model, we want $V(\epsilon_i) = 1$ instead of $\pi^2/6$. In which way the model parameters must be changed?
This means, the standard deviation of ϵ is now given by 1 rather than by $\pi/\sqrt{6} \approx 1.28$, i.e., multiplied by $\lambda = \sqrt{6}/\pi$. The choice probabilities (no longer given by the Logit formula!) will be unchanged if the determinsitic utilities, i.e., the parameters, are multiplied by λ as well

? Derive the IIA property from the choice probability formula.
The IIA says that the relative preference of an alternative i over j, i.e. P_i/P_j, does not depend on any V_k, $k \neq i, j$. Just calculate this ratio:

$$\frac{P_i}{P_j} = \frac{\exp(V_i)}{\sum_k \exp(V_k)} \frac{\sum_l \exp(V_l)}{\exp(V_j)} = \exp(V_i - V_j)$$ ✓

Questions (2)

? The choice probabilities of three alternatives are given by $P_1 = 0.2$, $P_2 = 0.4$, and $P_3 = 0.4$. Now, Alternative 3 is no longer available. Give the new Logit choice probabilities.

No need to re-calculate using the Logit probability formula. Just use the IIA property:

$$\frac{P_1}{P_2} = 1/2 = \text{const.} \Rightarrow P_1 = 1/3, \; P_2 = 2/3$$

? The Gumbel distribution is the limit distribution of the maximum of exponentially-tailed random variables. Is there really a justification for this sort of distribution if the RUs are the result of many unknown/not considered effects?

Not really. If there are many unknown/not considered effects, the chance is high that they are not correlated and the **central limit theorem** can be applied (even if there are correlations, this theorem is quite robust). Hence, there would be a justification for Gaussian rather than Gumbel RUs. The fact that the maximum of exponentially-tailed distributions is Gumbel distributed has no real relevance here.

Questions (2)

? The choice probabilities of three alternatives are given by $P_1 = 0.2$, $P_2 = 0.4$, and $P_3 = 0.4$. Now, Alternative 3 is no longer available. Give the new Logit choice probabilities.
No need to re-calculate using the Logit probability formula. Just use the IIA property:

$$\frac{P_1}{P_2} = 1/2 = \text{const.} \ \Rightarrow \ P_1 = 1/3, \ P_2 = 2/3$$

? The Gumbel distribution is the limit distribution of the maximum of exponentially-tailed random variables. Is there really a justification for this sort of distribution if the RUs are the result of many unknown/not considered effects?
Not really. If there are many unknown/not considered effects, the chance is high that they are not correlated and the **central limit theorem** can be applied (even if there are correlations, this theorem is quite robust). Hence, there would be a justification for Gaussian rather than Gumbel RUs. The fact that the maximum of exponentially-tailed distributions is Gumbel distributed has no real relevance here.

Questions (2)

? The choice probabilities of three alternatives are given by $P_1 = 0.2$, $P_2 = 0.4$, and $P_3 = 0.4$. Now, Alternative 3 is no longer available. Give the new Logit choice probabilities.

No need to re-calculate using the Logit probability formula. Just use the IIA property:

$$\frac{P_1}{P_2} = 1/2 = \text{const.} \Rightarrow P_1 = 1/3, \; P_2 = 2/3$$

? The Gumbel distribution is the limit distribution of the maximum of exponentially-tailed random variables. Is there really a justification for this sort of distribution if the RUs are the result of many unknown/not considered effects?

Not really. If there are many unknown/not considered effects, the chance is high that they are not correlated and the **central limit theorem** can be applied (even if there are correlations, this theorem is quite robust). Hence, there would be a justification for Gaussian rather than Gumbel RUs. The fact that the maximum of exponentially-tailed distributions is Gumbel distributed has no real relevance here.

Questions (2)

? The choice probabilities of three alternatives are given by $P_1 = 0.2$, $P_2 = 0.4$, and $P_3 = 0.4$. Now, Alternative 3 is no longer available. Give the new Logit choice probabilities.

No need to re-calculate using the Logit probability formula. Just use the IIA property:

$$\frac{P_1}{P_2} = 1/2 = \text{const.} \;\Rightarrow\; P_1 = 1/3,\; P_2 = 2/3$$

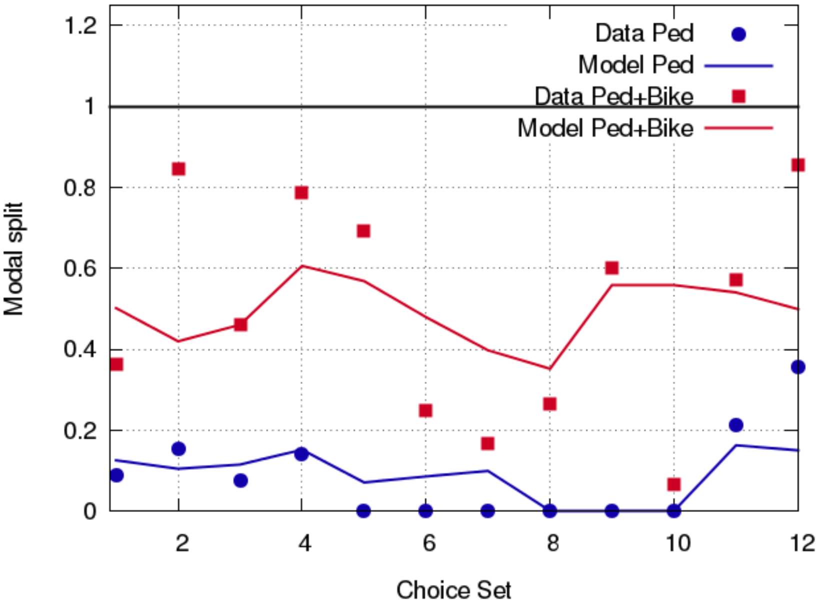
? The Gumbel distribution is the limit distribution of the maximum of exponentially-tailed random variables. Is there really a justification for this sort of distribution if the RUs are the result of many unknown/not considered effects?

Not really. If there are many unknown/not considered effects, the chance is high that they are not correlated and the **central limit theorem** can be applied (even if there are correlations, this theorem is quite robust). Hence, there would be a justification for Gaussian rather than Gumbel RUs. The fact that the maximum of exponentially-tailed distributions is Gumbel distributed has no real relevance here.

8.1.1 Example: SP Survey in the Audience WS18/19 (red: bad weather, $W = 1$)

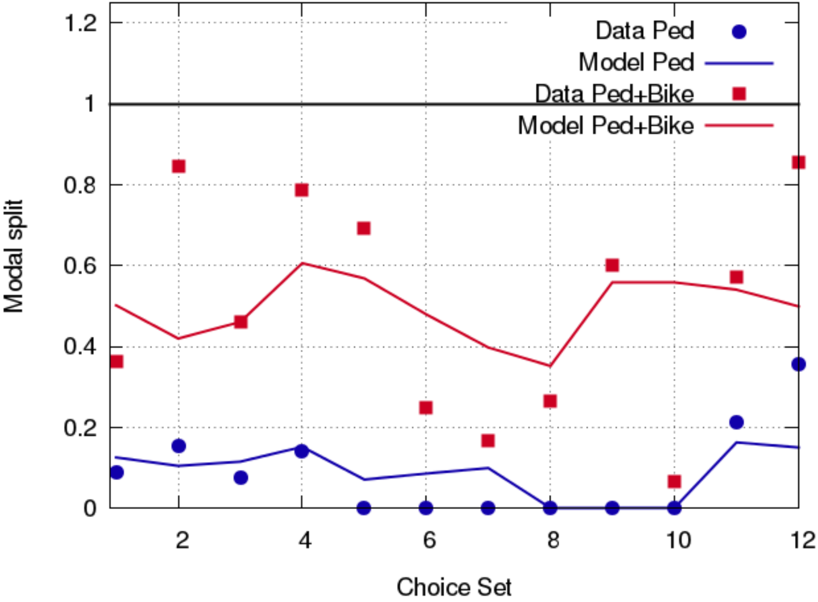
Choice Set	Alt. 1: Ped	Alt. 2: Bike	Alt. 3: PT/Car	Alt 1	Alt 2	Alt 3
1	30 min	20 min	20 min+0€	1	3	7
2	30 min	20 min	20 min+2€	2	9	2
3	30 min	20 min	20 min+1€	1	5	7
4	30 min	20 min	30 min+0€	2	9	3
5	50 min	20 min	30 min+0€	0	9	4
6	50 min	30 min	30 min+0€	0	3	9
7	50 min	40 min	30 min+0€	0	2	10
8	180 min	60 min	60 min+2€	0	4	11
9	180 min	40 min	60 min+2€	0	9	6
10	180 min	40 min	60 min+2€	0	1	14
11	12 min	8 min	10 min+0€	3	5	6
12	12 min	8 min	10 min+1€	5	7	2

Model 1: generic times and costs, no weather



$$
\begin{aligned}
V_i &= \beta_0 \delta_{i1} + \beta_1 \delta_{i2} \\
&+ \beta_2 K_i + \beta_3 T_i
\end{aligned}
$$

Model 1: generic times and costs, no weather

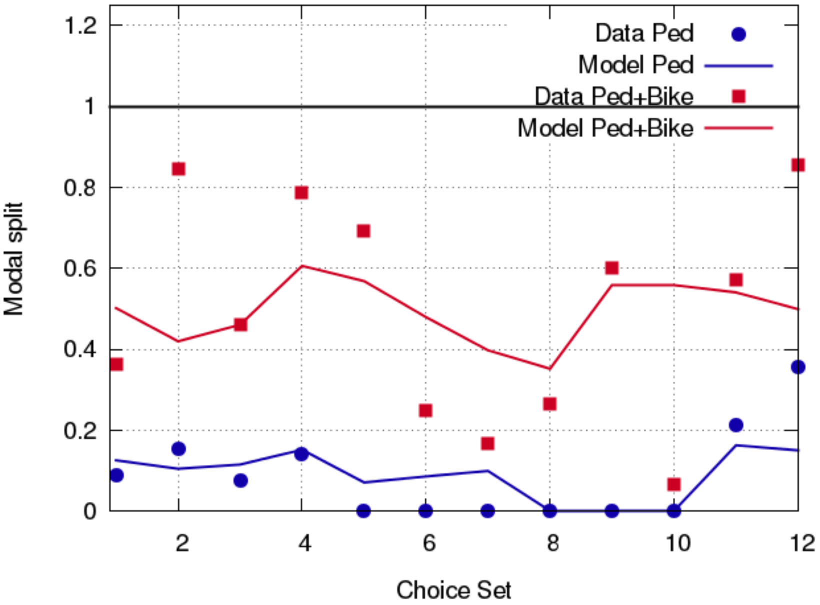


$$V_i = \beta_0\delta_{i1} + \beta_1\delta_{i2} + \beta_2 K_i + \beta_3 T_i$$

or

$$V_1 = \beta_0 + \beta_2 K_1 + \beta_3 T_1,$$
$$V_2 = \beta_1 + \beta_2 K_2 + \beta_3 T_2,$$
$$V_3 = \beta_2 K_3 + \beta_3 T_3$$

Model 1: generic times and costs, no weather



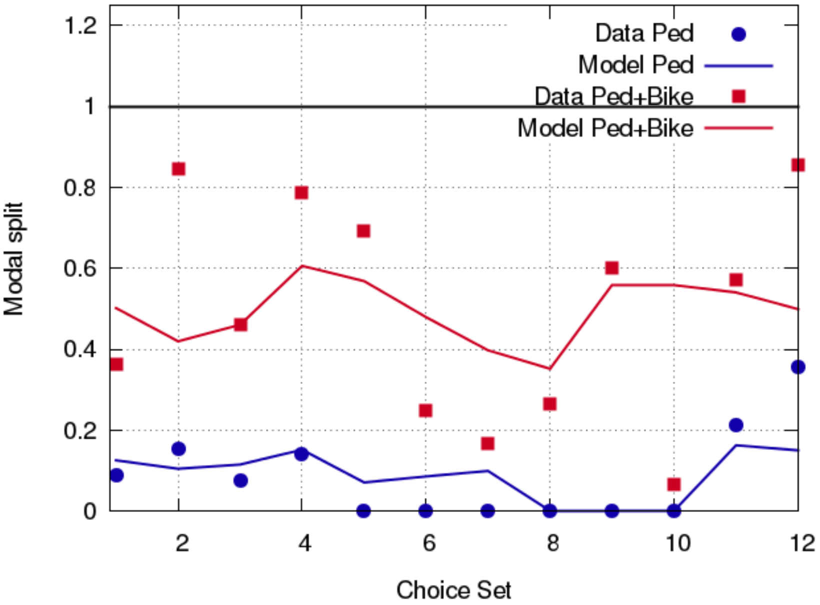
$$V_i = \beta_0 \delta_{i1} + \beta_1 \delta_{i2} + \beta_2 K_i + \beta_3 T_i$$

or

$$V_1 = \beta_0 + \beta_2 K_1 + \beta_3 T_1,$$
$$V_2 = \beta_1 + \beta_2 K_2 + \beta_3 T_2,$$
$$V_3 = \beta_2 K_3 + \beta_3 T_3$$

$$\beta_0 = -0.95 \pm 0.37,$$
$$\beta_1 = -0.28 \pm 0.24,$$
$$\beta_2 = +0.17 \pm 0.19,$$
$$\beta_3 = -0.04 \pm 0.02$$

Model 1: generic times and costs, no weather



$$V_i = \beta_0\delta_{i1} + \beta_1\delta_{i2} + \beta_2 K_i + \beta_3 T_i$$

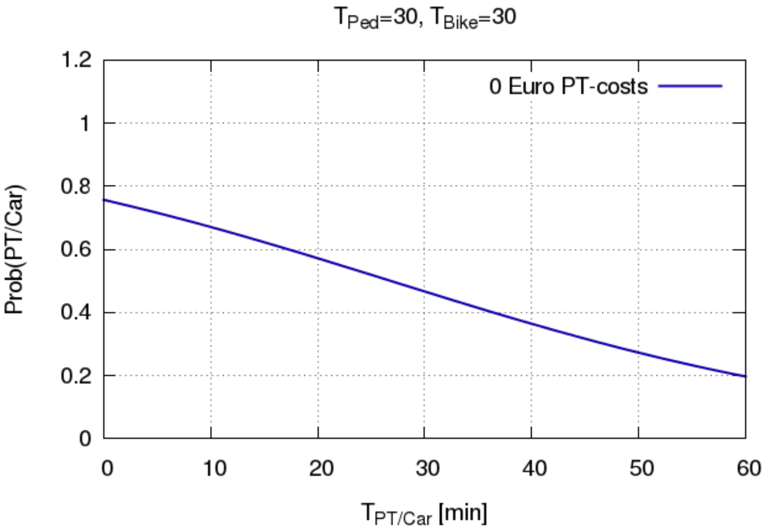
or

$$V_1 = \beta_0 + \beta_2 K_1 + \beta_3 T_1,$$
$$V_2 = \beta_1 + \beta_2 K_2 + \beta_3 T_2,$$
$$V_3 = \beta_2 K_3 + \beta_3 T_3$$

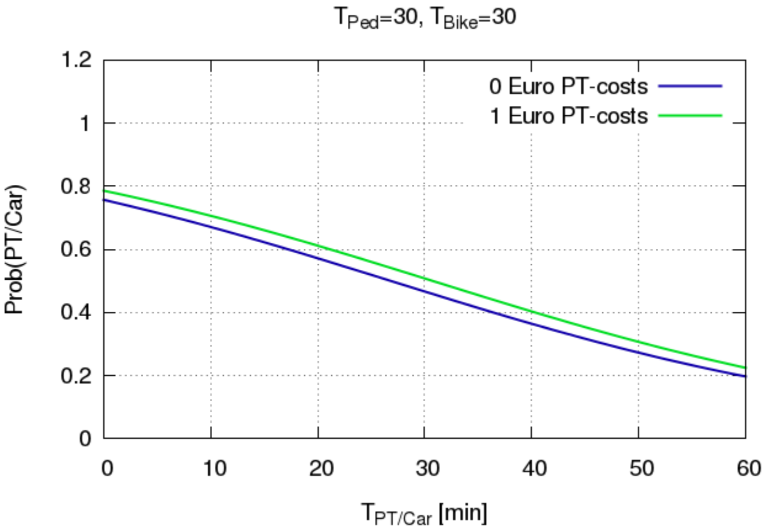
$$\beta_0 = -0.95 \pm 0.37,$$
$$\beta_1 = -0.28 \pm 0.24,$$
$$\beta_2 = +0.17 \pm 0.19,$$
$$\beta_3 = -0.04 \pm 0.02$$

$$\frac{\beta_0}{-\beta_3} = -22.4\,\text{min},$$
$$\frac{\beta_1}{-\beta_3} = -6.6\,\text{min},$$
$$\frac{60\beta_3}{\beta_2} = -15\,€/\text{h}$$

Dependence of the modal split on the PT attributes

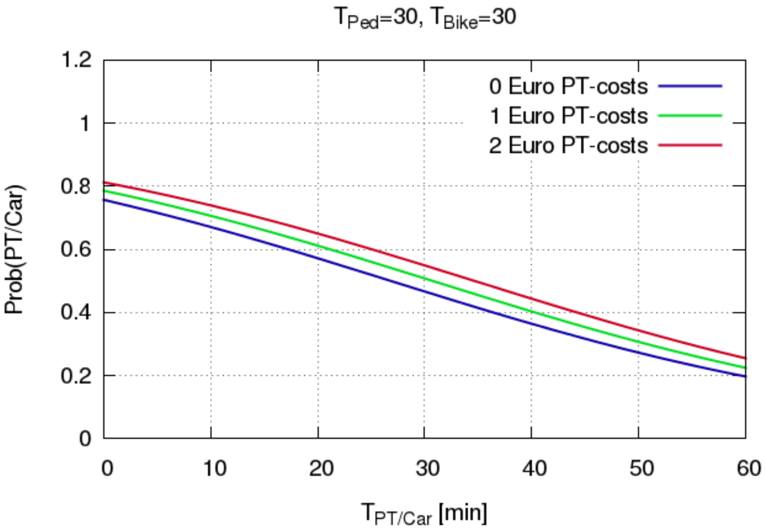


Dependence of the modal split on the PT attributes



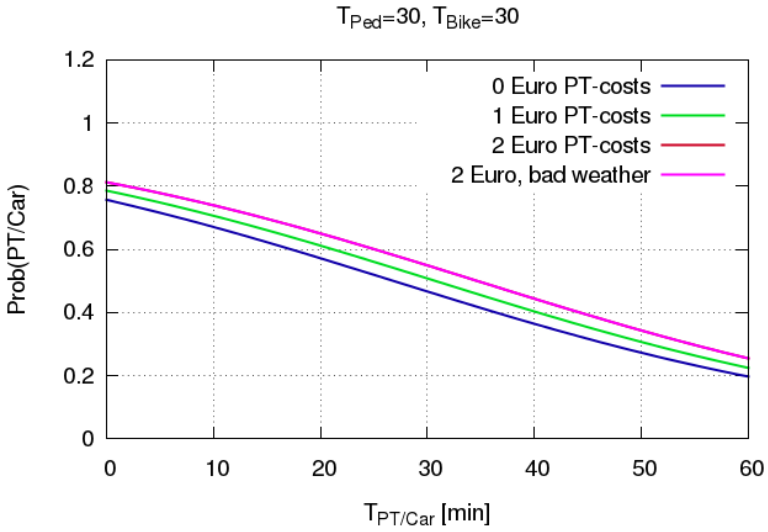
Wrong sign for cost sensitivity, too low time sensitivity!

Dependence of the modal split on the PT attributes



Wrong sign for cost sensitivity, too low time sensitivity!

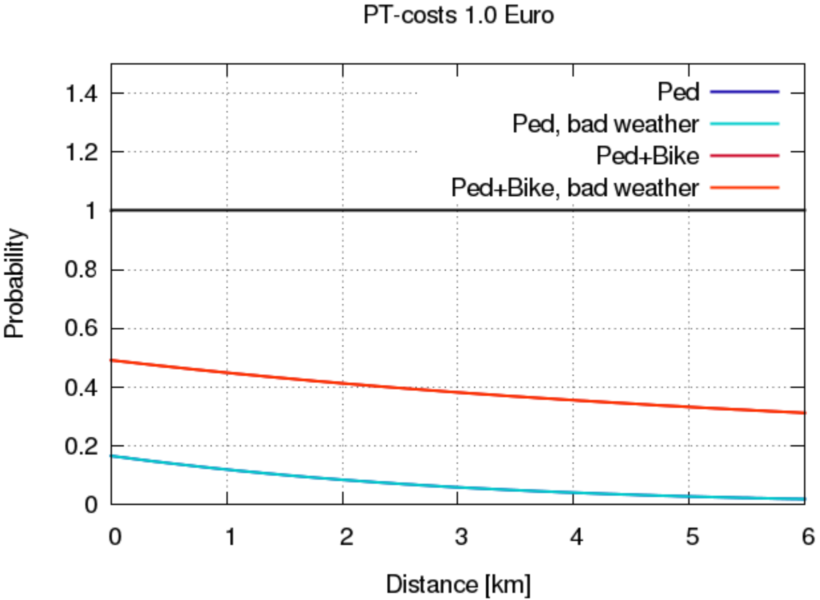
Dependence of the modal split on the PT attributes



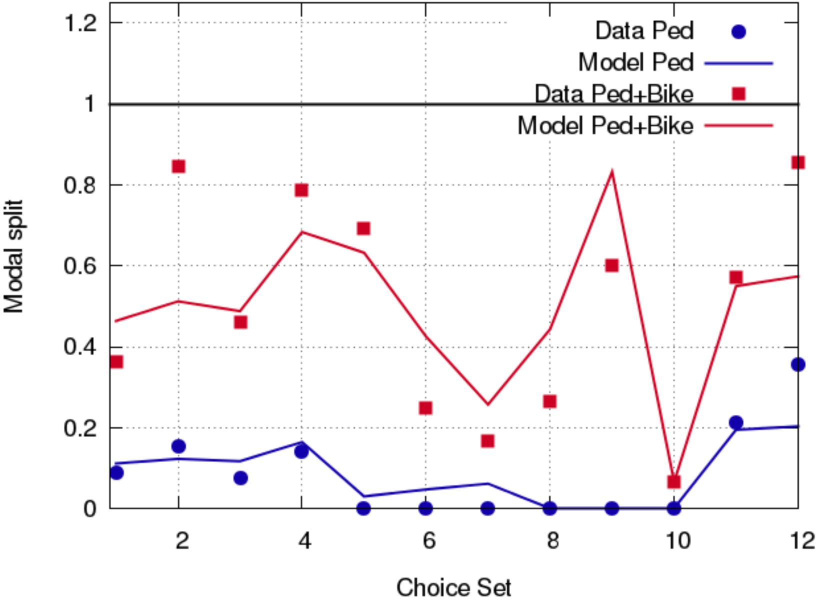
Wrong sign for cost sensitivity, too low time sensitivity!

Dependence on the distance

assuming plausible speeds 5, 15, and 25 km/h for each mode, respectively



Model 2: generic times and costs plus weather factor (bad weather, $W=1$)

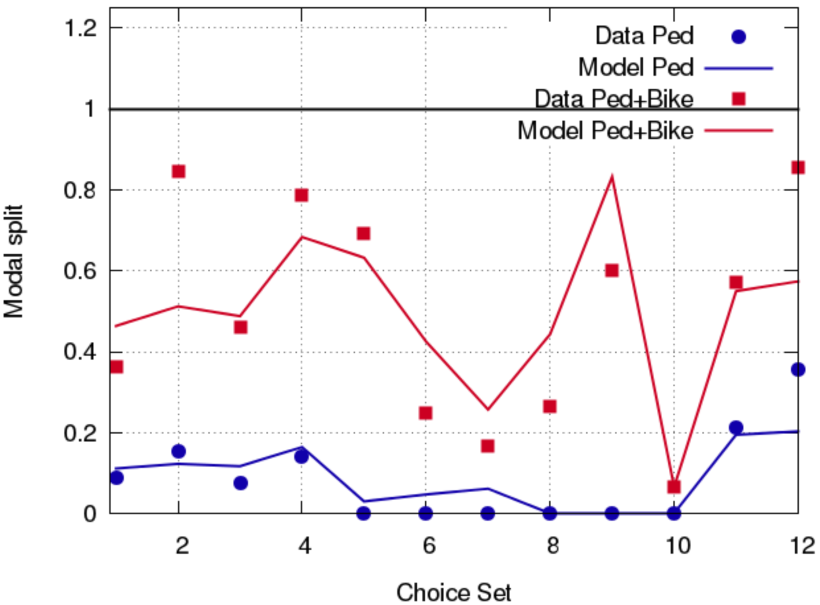


$$
\begin{aligned}
V_i &= \beta_0\delta_{i1} + \beta_1\delta_{i2} \\
&+ \beta_2 K_i + \beta_3 T_1 \\
&+ \beta_4 W \delta_{i3}
\end{aligned}
$$

$\beta_0 = -0.65 \pm 0.37,$
$\beta_1 = -0.42 \pm 0.25,$
$\beta_2 = -0.10 \pm 0.20,$
$\beta_3 = -0.09 \pm 0.02,$
$\beta_4 = 4.2 \pm 1.1$

$\frac{\beta_0}{-\beta_3} = -7.1\,\mathrm{min},$
$\frac{\beta_1}{-\beta_3} = -4.6\,\mathrm{min},$
$\frac{\beta_0}{-\beta_2} = -6.7\,€,$
$\frac{\beta_1}{-\beta_2} = -4.3\,€,$
$\frac{60\beta_3}{\beta_2} = +57\,€/\mathrm{h},$
$\frac{\beta_4}{-\beta_2} = +44\,€$

Model 2: generic times and costs plus weather factor (bad weather, $W=1$)

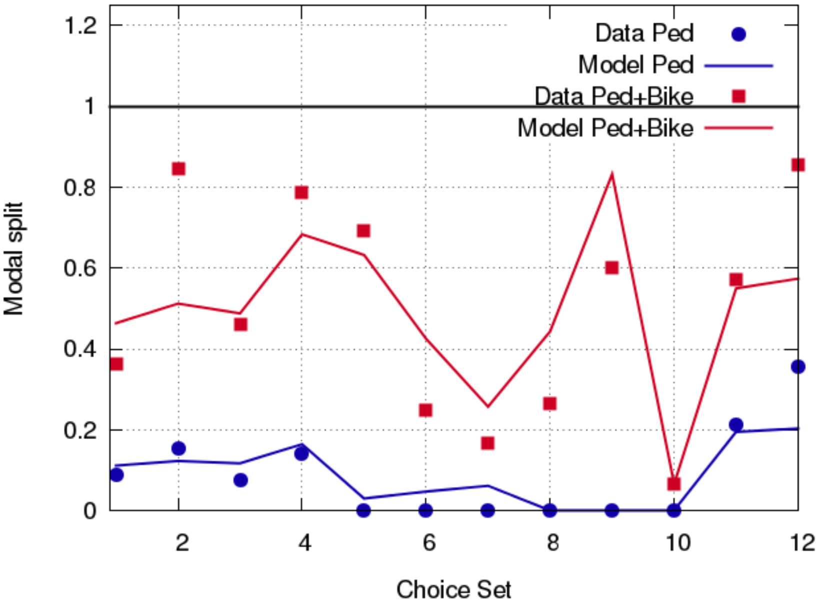


$$\begin{aligned} V_i &= \beta_0 \delta_{i1} + \beta_1 \delta_{i2} \\ &+ \beta_2 K_i + \beta_3 T_1 \\ &+ \beta_4 W \delta_{i3} \end{aligned}$$

$$\begin{aligned} \beta_0 &= -0.65 \pm 0.37, \\ \beta_1 &= -0.42 \pm 0.25, \\ \beta_2 &= -0.10 \pm 0.20, \\ \beta_3 &= -0.09 \pm 0.02, \\ \beta_4 &= 4.2 \pm 1.1 \end{aligned}$$

$$\begin{aligned} \frac{\beta_0}{-\beta_3} &= -7.1\,\text{min}, \\ \frac{\beta_1}{-\beta_3} &= -4.6\,\text{min}, \\ \frac{\beta_0}{-\beta_2} &= -6.7\,€, \\ \frac{\beta_1}{-\beta_2} &= -4.3\,€, \\ \frac{60\beta_3}{\beta_2} &= +57\,€/\text{h}, \\ \frac{\beta_4}{-\beta_2} &= +44\,€ \end{aligned}$$

Model 2: generic times and costs plus weather factor (bad weather, $W = 1$)

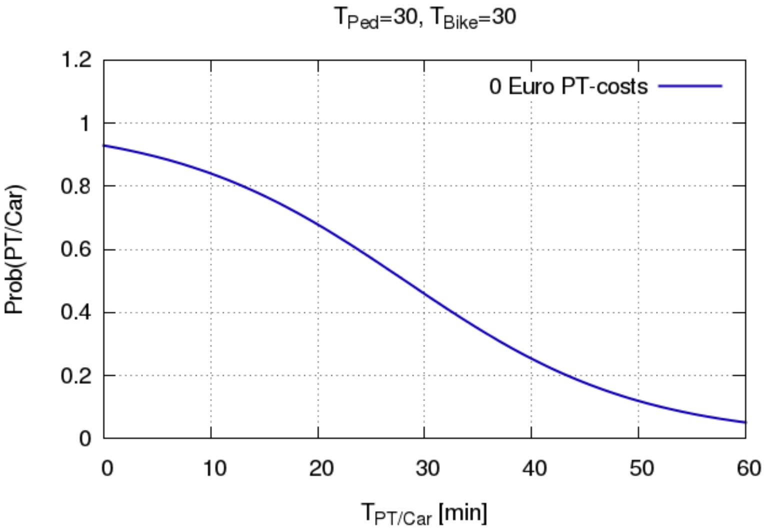


$$
\begin{aligned}
V_i &= \beta_0 \delta_{i1} + \beta_1 \delta_{i2} \\
&+ \beta_2 K_i + \beta_3 T_1 \\
&+ \beta_4 W \delta_{i3}
\end{aligned}
$$

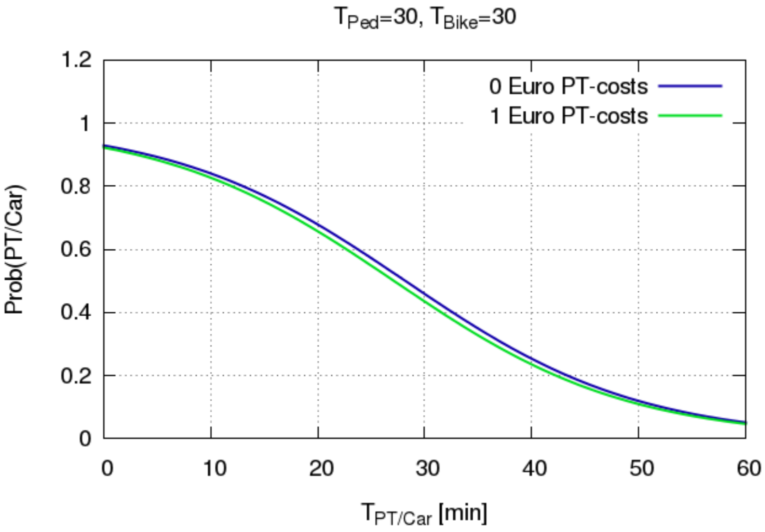
$$
\begin{aligned}
\beta_0 &= -0.65 \pm 0.37, \\
\beta_1 &= -0.42 \pm 0.25, \\
\beta_2 &= -0.10 \pm 0.20, \\
\beta_3 &= -0.09 \pm 0.02, \\
\beta_4 &= 4.2 \pm 1.1
\end{aligned}
$$

$$
\begin{aligned}
\frac{\beta_0}{-\beta_3} &= -7.1\,\text{min}, \\
\frac{\beta_1}{-\beta_3} &= -4.6\,\text{min}, \\
\frac{\beta_0}{-\beta_2} &= -6.7\,€, \\
\frac{\beta_1}{-\beta_2} &= -4.3\,€, \\
\frac{60\beta_3}{\beta_2} &= +57\,€/\text{h}, \\
\frac{\beta_4}{-\beta_2} &= +44\,€
\end{aligned}
$$

Dependence of the modal split on the PT attributes

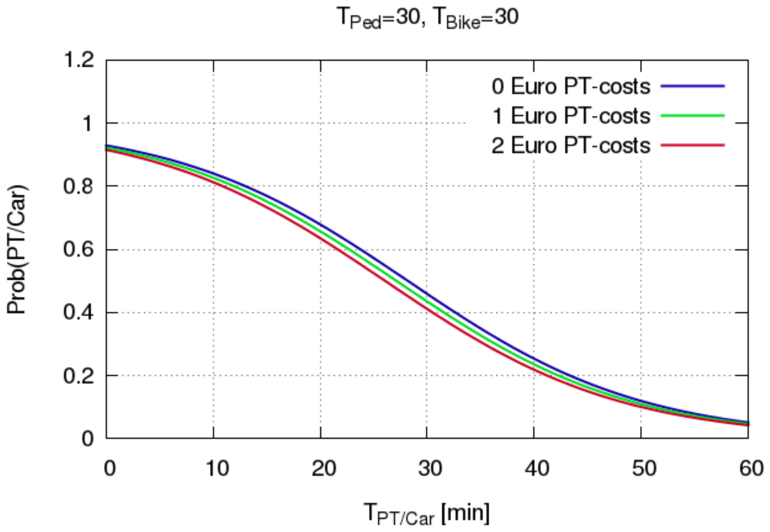


Dependence of the modal split on the PT attributes



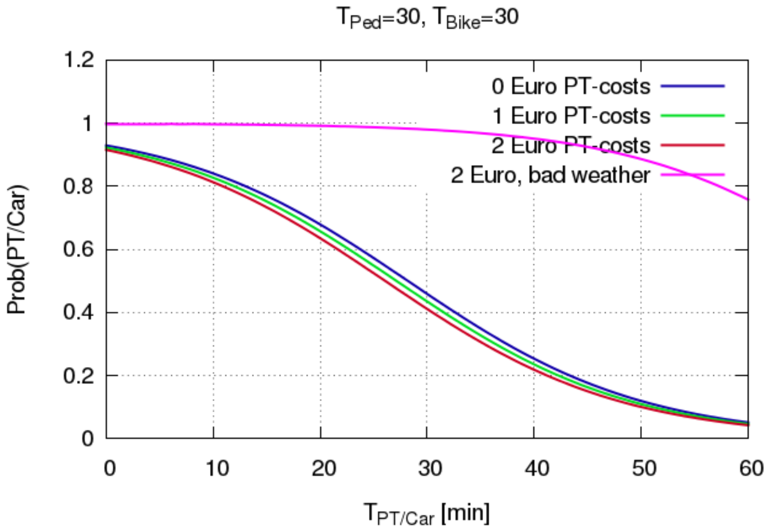
Too low cost sensitivity!

Dependence of the modal split on the PT attributes



Too low cost sensitivity!

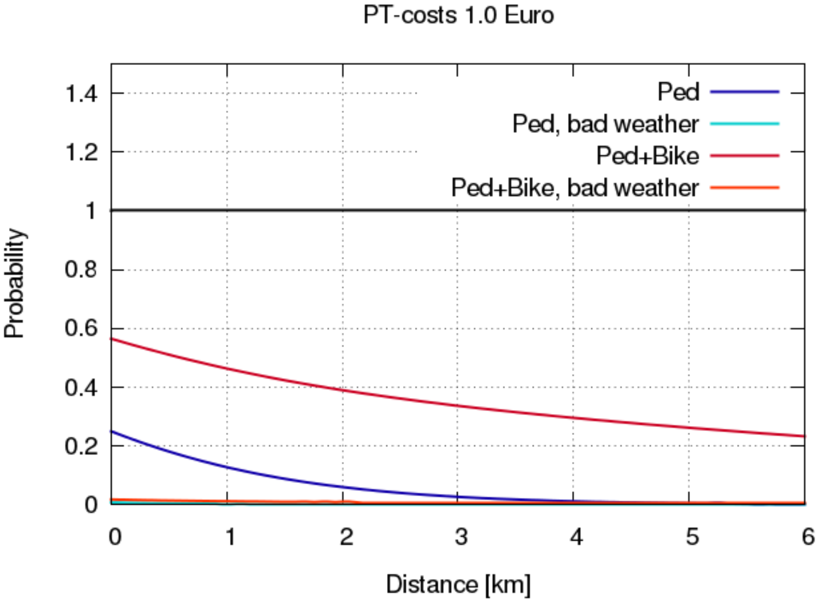
Dependence of the modal split on the PT attributes



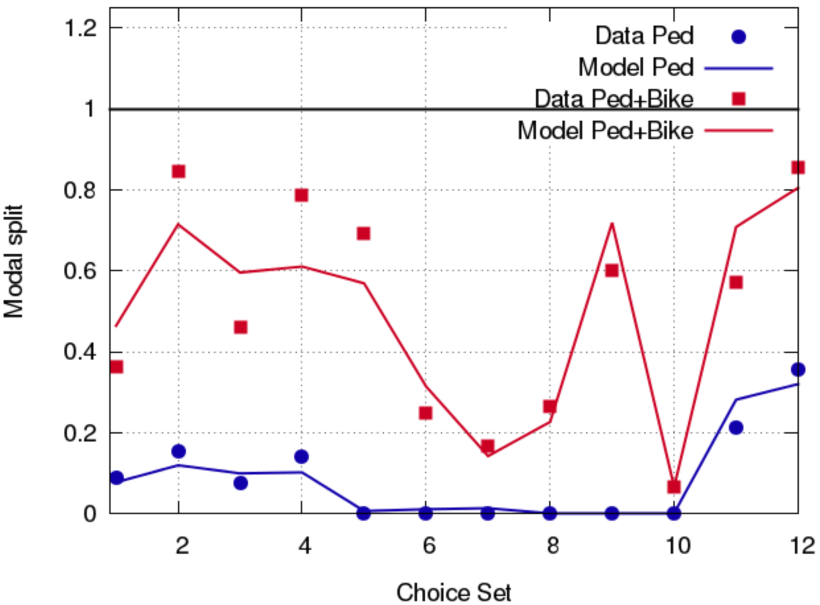
Too low cost sensitivity!

Dependence on the distance

assuming plausible speeds 5, 15, and 25 km/h for each mode, respectively



Model 3: alt-spec time sensitivities plus weather factor

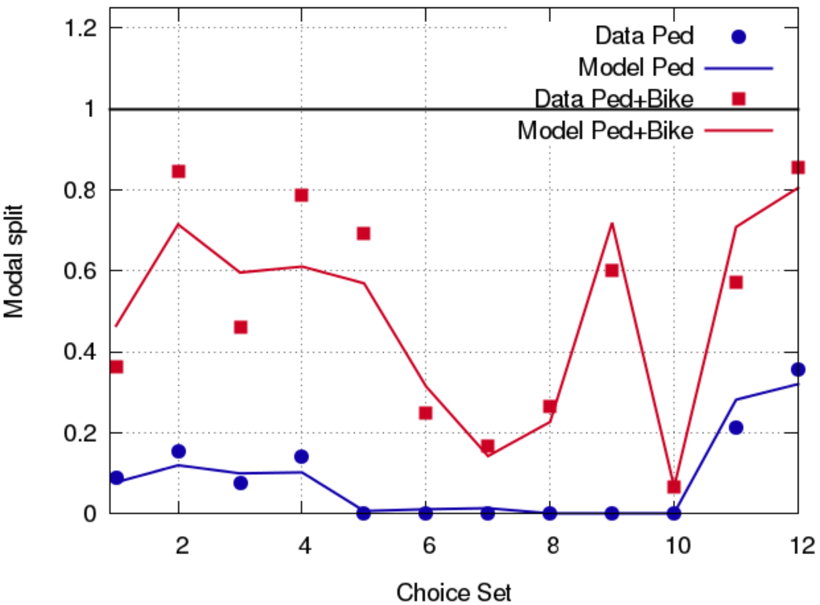


$$
\begin{aligned}
V_i &= \beta_0\delta_{i1} + \beta_1\delta_{i2} \\
&+ \beta_2 K + \beta_3 T_1 \delta_{i1} \\
&+ \beta_4 T_2 \delta_{i2} + \beta_5 T_3 \delta_{i3} \\
&+ \beta_6 W \delta_{i3}
\end{aligned}
$$

$$
\begin{aligned}
\beta_0 &= +1.03 \pm 0.74, \\
\beta_1 &= +0.66 \pm 0.40, \\
\beta_2 &= -0.53 \pm 0.25, \\
\beta_3 &= -0.14 \pm 0.03, \\
\beta_4 &= -0.11 \pm 0.03, \\
\beta_5 &= -0.06 \pm 0.03, \\
\beta_6 &= +3.6 \pm 1.1
\end{aligned}
$$

$$
\begin{aligned}
\frac{\beta_0}{-\beta_3} &= +7.5\,\text{min}, \\
\frac{\beta_1}{-\beta_3} &= +4.7\,\text{min}, \\
\frac{\beta_0}{-\beta_2} &= +1.9\,€, \\
\frac{\beta_1}{-\beta_2} &= +4.7\,€, \\
\frac{60\beta_5}{\beta_2} &= +6.7\,€/\text{h}, \\
\frac{\beta_4}{-\beta_2} &= +6.7\,€
\end{aligned}
$$

Model 3: alt-spec time sensitivities plus weather factor

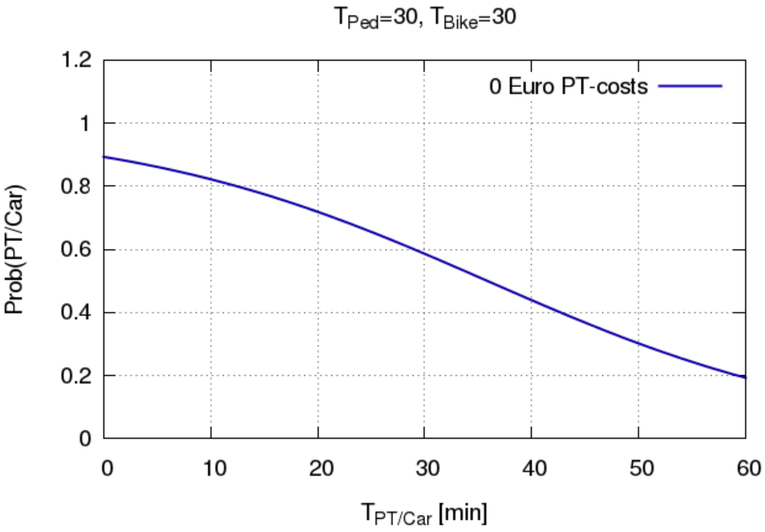


$$\begin{aligned} V_i &= \beta_0\delta_{i1} + \beta_1\delta_{i2} \\ &+ \beta_2 K + \beta_3 T_1\delta_{i1} \\ &+ \beta_4 T_2\delta_{i2} + \beta_5 T_3\delta_{i3} \\ &+ \beta_6 W\delta_{i3} \end{aligned}$$

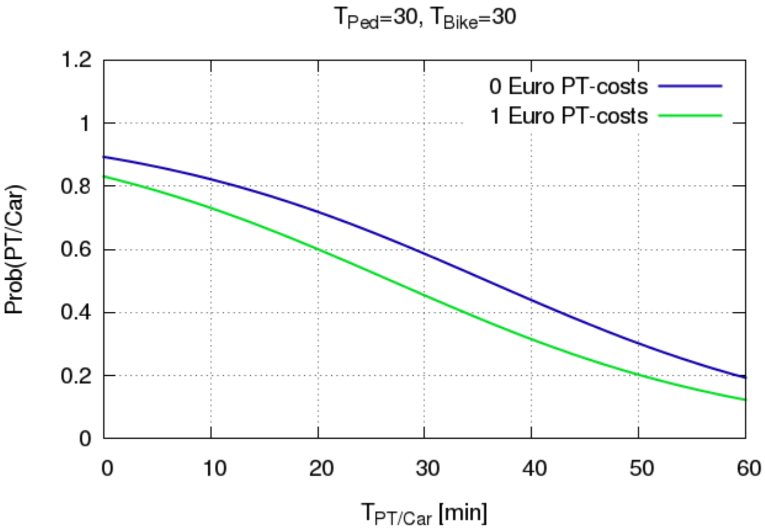
$$\begin{aligned} \beta_0 &= +1.03 \pm 0.74, \\ \beta_1 &= +0.66 \pm 0.40, \\ \beta_2 &= -0.53 \pm 0.25, \\ \beta_3 &= -0.14 \pm 0.03, \\ \beta_4 &= -0.11 \pm 0.03, \\ \beta_5 &= -0.06 \pm 0.03, \\ \beta_6 &= +3.6 \pm 1.1 \end{aligned}$$

$$\begin{aligned} \frac{\beta_0}{-\beta_3} &= +7.5\,\text{min}, \\ \frac{\beta_1}{-\beta_3} &= +4.7\,\text{min}, \\ \frac{\beta_0}{-\beta_2} &= +1.9\,€, \\ \frac{\beta_1}{-\beta_2} &= +4.7\,€, \\ \frac{60\beta_5}{\beta_2} &= +6.7\,€/\text{h}, \\ \frac{\beta_4}{-\beta_2} &= +6.7\,€ \end{aligned}$$

Dependence of the modal split on the PT attributes

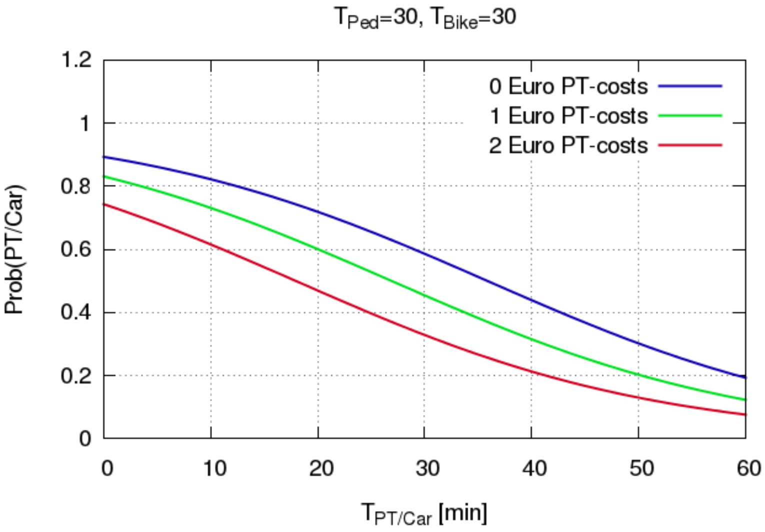


Dependence of the modal split on the PT attributes



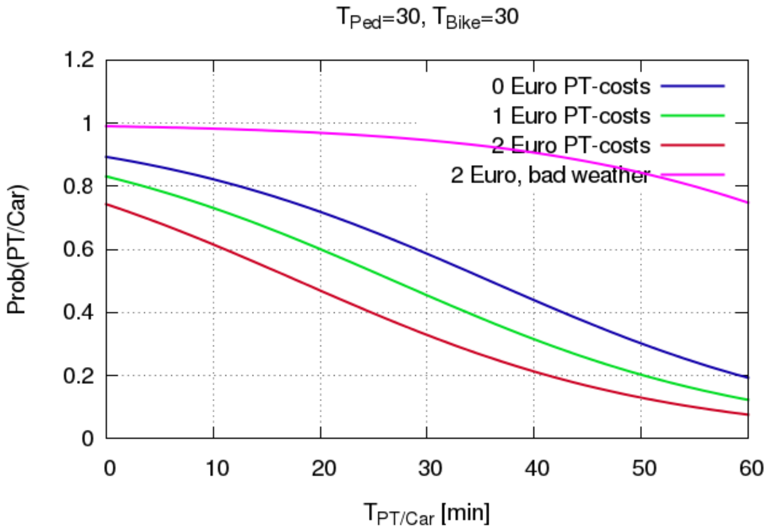
Everything plausible

Dependence of the modal split on the PT attributes



Everything plausible

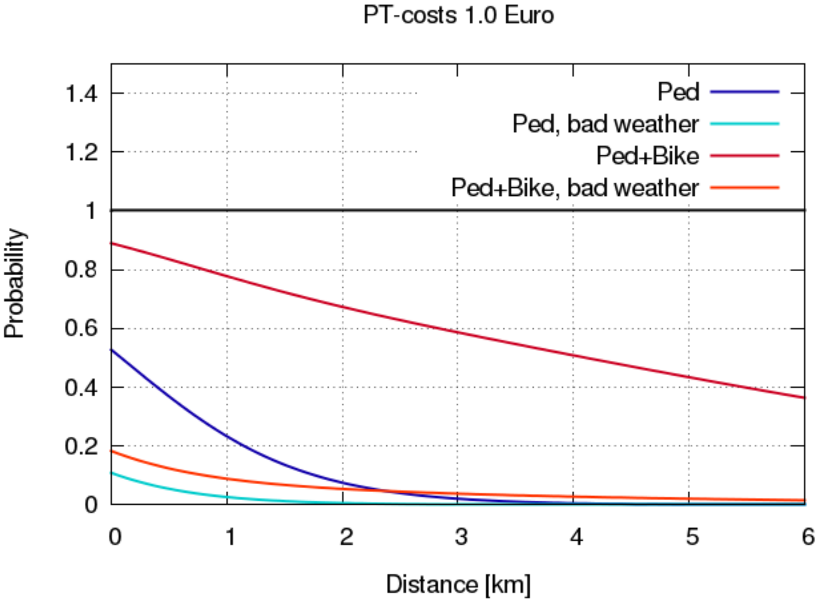
Dependence of the modal split on the PT attributes



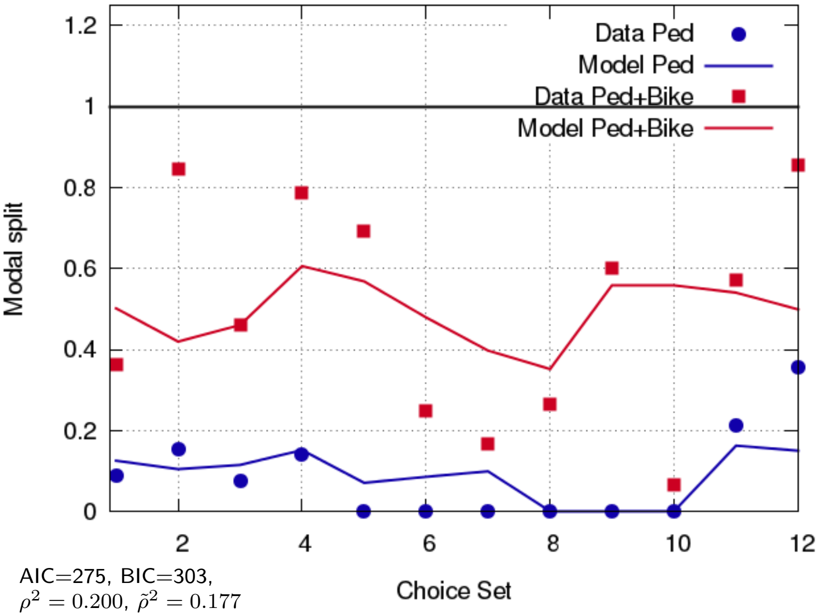
Everything plausible

Dependence on the distance

assuming plausible speeds 5, 15, and 25 km/h for each mode, respectively



Comparison: Model 1



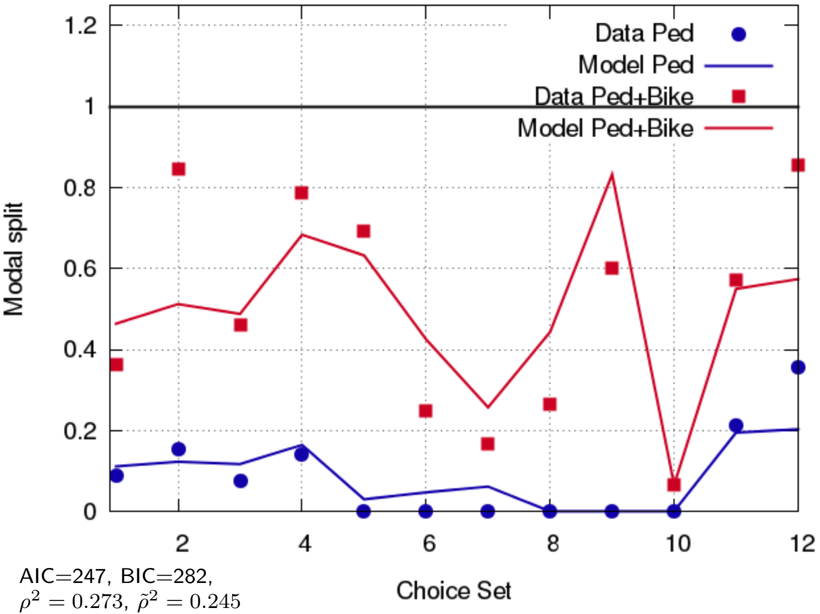
$$V_i = \beta_0\delta_{i1} + \beta_1\delta_{i2} + \beta_2 K_i + \beta_3 T_i$$

$$\beta_0 = -0.95 \pm 0.37,\quad \beta_1 = -0.28 \pm 0.24,\quad \beta_2 = +0.17 \pm 0.19,\quad \beta_3 = -0.04 \pm 0.02$$

$$\frac{\beta_0}{-\beta_3} = -22.4\,\text{min},\quad \frac{\beta_1}{-\beta_3} = -6.6\,\text{min},\quad \frac{60\beta_3}{\beta_2} = -15\,€/\text{h}$$

AIC=275, BIC=303, $\rho^2 = 0.200$, $\tilde{\rho}^2 = 0.177$

Model 2



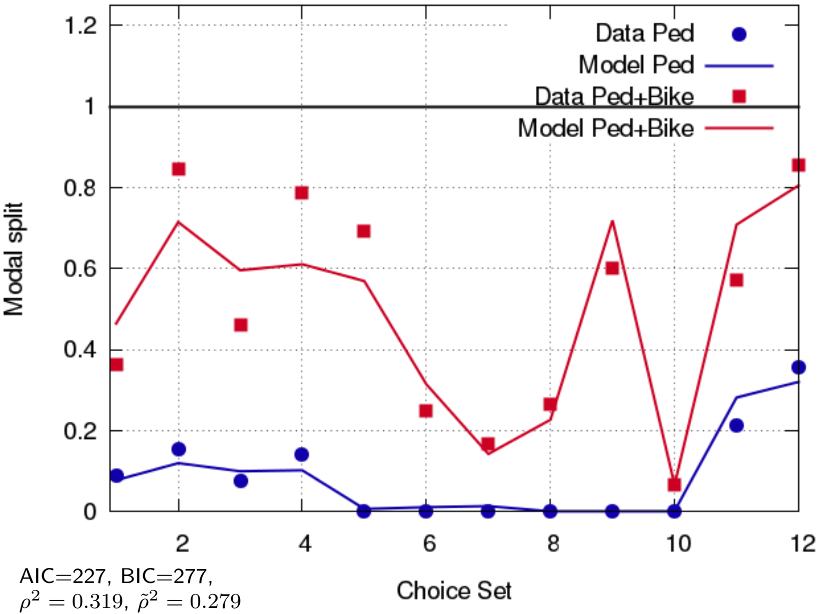
AIC=247, BIC=282,
$\rho^2 = 0.273$, $\tilde{\rho}^2 = 0.245$

$$
\begin{aligned}
V_i &= \beta_0 \delta_{i1} + \beta_1 \delta_{i2} \\
&+ \beta_2 K_i + \beta_3 T_i \\
&+ \beta_4 W \delta_{i3}
\end{aligned}
$$

$$
\begin{aligned}
\beta_0 &= -0.65 \pm 0.37, \\
\beta_1 &= -0.42 \pm 0.25, \\
\beta_2 &= -0.10 \pm 0.20, \\
\beta_3 &= -0.09 \pm 0.02, \\
\beta_4 &= 4.2 \pm 1.1
\end{aligned}
$$

$$
\begin{aligned}
\frac{\beta_0}{-\beta_3} &= -7.1\,\mathrm{min}, \\
\frac{\beta_1}{-\beta_3} &= -4.6\,\mathrm{min}, \\
\frac{\beta_0}{-\beta_2} &= -6.7\,€, \\
\frac{\beta_1}{-\beta_2} &= -4.3\,€, \\
\frac{60\beta_3}{\beta_2} &= +57\,€/\mathrm{h}, \\
\frac{\beta_4}{-\beta_2} &= +44\,€
\end{aligned}
$$

Model 3



AIC=227, BIC=277,
$\rho^2 = 0.319$, $\tilde{\rho}^2 = 0.279$

$$
\begin{aligned}
V_i &= \beta_0\delta_{i1} + \beta_1\delta_{i2} \\
&+ \beta_2 K + \beta_3 T_1\delta_{i1} \\
&+ \beta_4 T_2\delta_{i2} + \beta_5 T_3\delta_{i3} \\
&+ \beta_6 W\delta_{i3}
\end{aligned}
$$

$$
\begin{aligned}
\beta_0 &= +1.03 \pm 0.74, \\
\beta_1 &= +0.66 \pm 0.40, \\
\beta_2 &= -0.53 \pm 0.25, \\
\beta_3 &= -0.14 \pm 0.03, \\
\beta_4 &= -0.11 \pm 0.03, \\
\beta_5 &= -0.06 \pm 0.03, \\
\beta_6 &= +3.6 \pm 1.1
\end{aligned}
$$

$$
\begin{aligned}
\frac{\beta_0}{-\beta_3} &= +7.5\,\text{min}, \\
\frac{\beta_1}{-\beta_3} &= +4.7\,\text{min}, \\
\frac{\beta_0}{-\beta_2} &= +1.9\,€, \\
\frac{\beta_1}{-\beta_2} &= +4.7\,€, \\
\frac{60\beta_5}{\beta_2} &= +6.7\,€/\text{h}, \\
\frac{\beta_4}{-\beta_2} &= +6.7\,€
\end{aligned}
$$

8.2 Probit Models

The **Probit Model** class is defined by (generally correlated) Gaussian RUs.

- The general multinomial Probit model (MNP) has random utilities $\boldsymbol{\epsilon} \sim N(\mathbf{0}, \boldsymbol{\Sigma})$ with the variance-covariance matrix $\boldsymbol{\Sigma}$ of the RUs
- The special case of the i.i.d. MNP with $\Sigma = \mathbf{1}$ (unit matrix), i.e., $\epsilon_i \sim$ i.i.d. $N(0, 1)$ has similar properties as the MNL (but not the IIA property!). However, for $I \geq 3$, the MNP needs integrals (1d, if there are no correlations) to calculate the choice probabilities.
- Often, i.i.d Gaussian RUs can be motivited by the central-limit theorem while Gumbel distributed ones cannot. However, since the MNL behaves similarly and has explicit choice probabilities and a simpler calibration, it is often favoured over the i.i.d. MNP.

? Why one can set the variance-covariance matrix to be the unit matrix (i.e. setting all variances=1) in case of the i.i.d MNP?

Because of the Scaling invariance of all Discrete-choice models with additive random utilities. If we had $\epsilon_i \sim$ i.i.d.$N(0, 1/\lambda^2)$, just multiply the deterministic and random utilities by λ to have an equivalent Probit model with $\epsilon_i \sim$ i.i.d.$N(0, 1)$

8.2 Probit Models

The **Probit Model** class is defined by (generally correlated) Gaussian RUs.

- The general multinomial Probit model (MNP) has random utilities $\boldsymbol{\epsilon} \sim N(\mathbf{0}, \boldsymbol{\Sigma})$ with the variance-covariance matrix $\boldsymbol{\Sigma}$ of the RUs
- The special case of the i.i.d. MNP with $\boldsymbol{\Sigma} = \mathbf{1}$ (unit matrix), i.e., $\epsilon_i \sim$ i.i.d. $N(0,1)$ has similar properties as the MNL (but not the IIA property!). However, for $I \geq 3$, the MNP needs integrals (1d, if there are no correlations) to calculate the choice probabilities.
- Often, i.i.d Gaussian RUs can be motivited by the central-limit theorem while Gumbel distributed ones cannot. However, since the MNL behaves similarly and has explicit choice probabilities and a simpler calibration, it is often favoured over the i.i.d. MNP.

? Why one can set the variance-covariance matrix to be the unit matrix (i.e. setting all variances=1) in case of the i.i.d MNP?

Because of the Scaling invariance of all Discrete-choice models with additive random utilities. If we had $\epsilon_i \sim$ i.i.d.$N(0, 1/\lambda^2)$, just multiply the deterministic and random utilities by λ to have an equivalent Probit model with $\epsilon_i \sim$ i.i.d.$N(0,1)$

8.2 Probit Models

The **Probit Model** class is defined by (generally correlated) Gaussian RUs.

- The general multinomial Probit model (MNP) has random utilities $\boldsymbol{\epsilon} \sim N(\mathbf{0}, \boldsymbol{\Sigma})$ with the variance-covariance matrix $\boldsymbol{\Sigma}$ of the RUs
- The special case of the i.i.d. MNP with $\boldsymbol{\Sigma} = \mathbf{1}$ (unit matrix), i.e., $\epsilon_i \sim$ i.i.d. $N(0, 1)$ has similar properties as the MNL (but not the IIA property!). However, for $I \geq 3$, the MNP needs integrals (1d, if there are no correlations) to calculate the choice probabilities.
- Often, i.i.d Gaussian RUs can be motivited by the central-limit theorem while Gumbel distributed ones cannot. However, since the MNL behaves similarly and has explicit choice probabilities and a simpler calibration, it is often favoured over the i.i.d. MNP.

? Why one can set the variance-covariance matrix to be the unit matrix (i.e. setting all variances=1) in case of the i.i.d MNP?

Because of the Scaling invariance of all Discrete-choice models with additive random utilities. If we had $\epsilon_i \sim$ i.i.d.$N(0, 1/\lambda^2)$, just multiply the deterministic and random utilities by λ to have an equivalent Probit model with $\epsilon_i \sim$ i.i.d.$N(0, 1)$

8.2 Probit Models

The **Probit Model** class is defined by (generally correlated) Gaussian RUs.

- The general multinomial Probit model (MNP) has random utilities $\boldsymbol{\epsilon} \sim N(\mathbf{0}, \boldsymbol{\Sigma})$ with the variance-covariance matrix $\boldsymbol{\Sigma}$ of the RUs
- The special case of the i.i.d. MNP with $\boldsymbol{\Sigma} = \mathbf{1}$ (unit matrix), i.e., $\epsilon_i \sim$ i.i.d. $N(0, 1)$ has similar properties as the MNL (but not the IIA property!). However, for $I \geq 3$, the MNP needs integrals (1d, if there are no correlations) to calculate the choice probabilities.
- Often, i.i.d Gaussian RUs can be motivited by the central-limit theorem while Gumbel distributed ones cannot. However, since the MNL behaves similarly and has explicit choice probabilities and a simpler calibration, it is often favoured over the i.i.d. MNP.

? Why one can set the variance-covariance matrix to be the unit matrix (i.e. setting all variances=1) in case of the i.i.d MNP?

Because of the Scaling invariance of all Discrete-choice models with additive random utilities. If we had $\epsilon_i \sim$ i.i.d.$N(0, 1/\lambda^2)$, just multiply the deterministic and random utilities by λ to have an equivalent Probit model with $\epsilon_i \sim$ i.i.d.$N(0, 1)$

8.2 Probit Models

The **Probit Model** class is defined by (generally correlated) Gaussian RUs.

- The general multinomial Probit model (MNP) has random utilities $\boldsymbol{\epsilon} \sim N(\mathbf{0}, \boldsymbol{\Sigma})$ with the variance-covariance matrix $\boldsymbol{\Sigma}$ of the RUs
- The special case of the i.i.d. MNP with $\boldsymbol{\Sigma} = \mathbf{1}$ (unit matrix), i.e., $\epsilon_i \sim$ i.i.d. $N(0, 1)$ has similar properties as the MNL (but not the IIA property!). However, for $I \geq 3$, the MNP needs integrals (1d, if there are no correlations) to calculate the choice probabilities.
- Often, i.i.d Gaussian RUs can be motivited by the central-limit theorem while Gumbel distributed ones cannot. However, since the MNL behaves similarly and has explicit choice probabilities and a simpler calibration, it is often favoured over the i.i.d. MNP.

? Why one can set the variance-covariance matrix to be the unit matrix (i.e. setting all variances=1) in case of the i.i.d MNP?
Because of the Scaling invariance of all Discrete-choice models with additive random utilities. If we had $\epsilon_i \sim$ i.i.d.$N(0, 1/\lambda^2)$, just multiply the deterministic and random utilities by λ to have an equivalent Probit model with $\epsilon_i \sim$ i.i.d.$N(0, 1)$

Choice probabilities of the binary Probit model I

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

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

? Derive the choice probabilities for the correlated binary Probit model. *Hint:* a linear combination of Gaussians is again a Gaussian

Assume without loss of generality zero expectations and use the general rules for the variance of two random variables X_1, X_2 $(a, b \in \mathbb{R})$:

$$V(aX_1 + bX_2) = a^2 V(X_1) + b^2 V(X_2) + 2ab\,\mathrm{Cov}(X_1, X_2)$$

? The Probit time and cost sensitivities are $\hat{\beta}_T = -0.1\,\mathrm{min}^{-1}$ and $\hat{\beta}_C = -0.6\,€^{-1}$. Give the implied value of time (VOT). Give also the approximate parameter values and the VOT for the corresponding Probit model

The VOT in €/min is just the ratio of the time and cost sensitivities, VOT $= \hat{\beta}_T/\hat{\beta}_C = 1/6\,€/\mathrm{min} = 10\,€/\mathrm{h}$. The Logit parameters are approximately the Probit parameters multiplied by the standard deviation $\lambda = \pi/\sqrt{6}$ of the Gumbel distributed Logit RUs. The VOT is essentially unchanged.

Choice probabilities of the binary Probit model I

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

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

? Derive the choice probabilities for the correlated binary Probit model. *Hint:* a linear combination of Gaussians is again a Gaussian

Assume without loss of generality zero expectations and use the general rules for the variance of two random variables X_1, X_2 $(a, b \in \mathbb{R})$:

$$V(aX_1 + bX_2) = a^2 V(X_1) + b^2 V(X_2) + 2ab\,\mathrm{Cov}(X_1, X_2)$$

? The Probit time and cost sensitivities are $\hat{\beta}_T = -0.1\,\text{min}^{-1}$ and $\hat{\beta}_C = -0.6\,€^{-1}$. Give the implied value of time (VOT). Give also the approximate parameter values and the VOT for the corresponding Probit model

The VOT in €/min is just the ratio of the time and cost sensitivities, VOT $= \hat{\beta}_T/\hat{\beta}_C = 1/6\,€/\text{min} = 10\,€/\text{h}$. The Logit parameters are approximately the Probit parameters multiplied by the standard deviation $\lambda = \pi/\sqrt{6}$ of the Gumbel distributed Logit RUs. The VOT is essentially unchanged.

Choice probabilities of the binary Probit model I

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

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

? Derive the choice probabilities for the correlated binary Probit model. *Hint:* a linear combination of Gaussians is again a Gaussian

Assume without loss of generality zero expectations and use the general rules for the variance of two random variables X_1, X_2 $(a, b \in I\!R)$:

$$V(aX_1 + bX_2) = a^2 V(X_1) + b^2 V(X_2) + 2ab\,\text{Cov}(X_1, X_2)$$

? The Probit time and cost sensitivities are $\hat{\beta}_T = -0.1\,\text{min}^{-1}$ and $\hat{\beta}_C = -0.6\,€^{-1}$. Give the implied value of time (VOT). Give also the approximate parameter values and the VOT for the corresponding Probit model

The VOT in €/min is just the ratio of the time and cost sensitivities, VOT $= \hat{\beta}_T / \hat{\beta}_C = 1/6\,€/\text{min} = 10\,€/\text{h}$. The Logit parameters are approximately the Probit parameters multiplied by the standard deviation $\lambda = \pi/\sqrt{6}$ of the Gumbel distributed Logit RUs. The VOT is essentially unchanged.

Choice probabilities of the binary Probit model I

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

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

? Derive the choice probabilities for the correlated binary Probit model. *Hint:* a linear combination of Gaussians is again a Gaussian
Assume without loss of generality zero expectations and use the general rules for the variance of two random variables X_1, X_2 $(a, b \in \mathbb{R})$:

$$V(aX_1 + bX_2) = a^2 V(X_1) + b^2 V(X_2) + 2ab\,\text{Cov}(X_1, X_2)$$

? The Probit time and cost sensitivities are $\hat{\beta}_T = -0.1\,\text{min}^{-1}$ and $\hat{\beta}_C = -0.6\,€^{-1}$. Give the implied value of time (VOT). Give also the approximate parameter values and the VOT for the corresponding Probit model

The VOT in €/min is just the ratio of the time and cost sensitivities, VOT $= \hat{\beta}_T/\hat{\beta}_C = 1/6\,€/\text{min} = 10\,€/\text{h}$. The Logit parameters are approximately the Probit parameters multiplied by the standard deviation $\lambda = \pi/\sqrt{6}$ of the Gumbel distributed Logit RUs. The VOT is essentially unchanged.

Choice probabilities of the binary Probit model I

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

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

? Derive the choice probabilities for the correlated binary Probit model. *Hint:* a linear combination of Gaussians is again a Gaussian

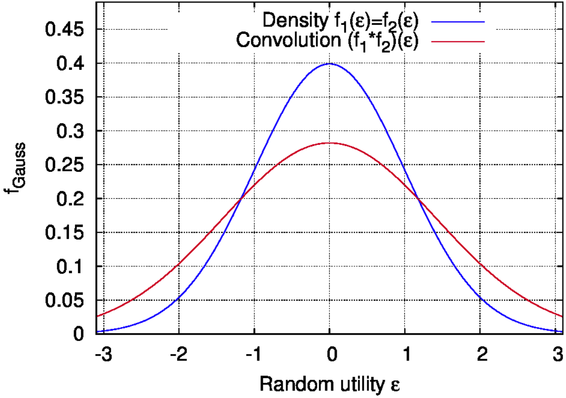
Assume without loss of generality zero expectations and use the general rules for the variance of two random variables X_1, X_2 $(a, b \in \mathbb{R})$:

$$V(aX_1 + bX_2) = a^2 V(X_1) + b^2 V(X_2) + 2ab\,\text{Cov}(X_1, X_2)$$

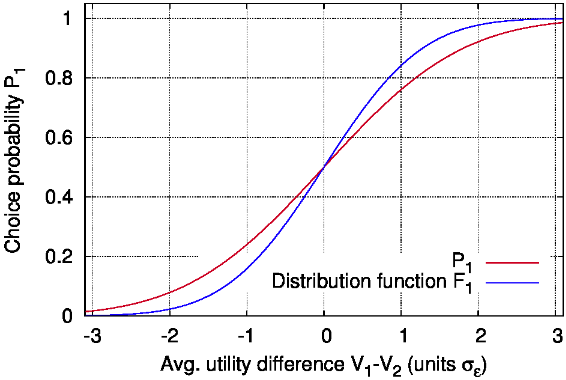
? The Probit time and cost sensitivities are $\hat{\beta}_T = -0.1\,\text{min}^{-1}$ and $\hat{\beta}_C = -0.6\,€^{-1}$. Give the implied value of time (VOT). Give also the approximate parameter values and the VOT for the corresponding Probit model

The VOT in €/min is just the ratio of the time and cost sensitivities, VOT $= \hat{\beta}_T/\hat{\beta}_C = 1/6\,€/\text{min} = 10\,€/\text{h}$. The Logit parameters are approximately the Probit parameters multiplied by the standard deviation $\lambda = \pi/\sqrt{6}$ of the Gumbel distributed Logit RUs. The VOT is essentially unchanged.

Choice probabilities of the binary Probit model II

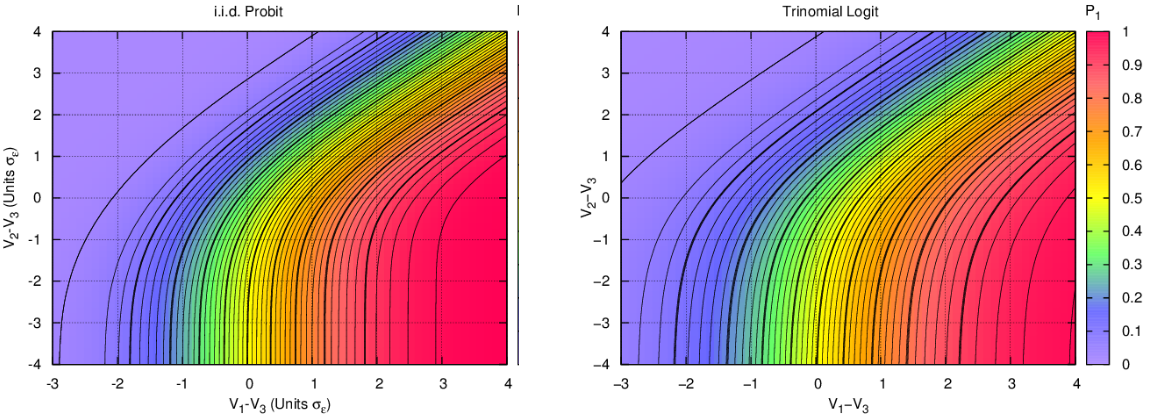


Densities of the standardnormal distributed random utilities ϵ_1 and ϵ_2 and of the utility difference $\epsilon_1 - \epsilon_2$



Distribution functions of the random utilities and the utility difference as a function of the deterministic utility difference $V_1 - V_2$

Choice probabilities of trinomial i.i.d. Probit and Logit



Symmetrie considerations:

$$P_2(V_2 - V_3, V_1 - V_3) = P_1(V_1 - V_3, V_2 - V_3),$$
$$P_3(V_2 - V_3, V_1 - V_3) = 1 - P_1 - P_2$$

8.3 Elasticities

- *General definition:*

  Elasticities denote the percentaged change of endogenous variables y_i per small percentaged change of exogenous variables x_j for an average situation

$$\epsilon_{ij} = \frac{\bar{x}_j}{\bar{y}_i} \frac{\partial y_i}{\partial x_j}\bigg|_{\boldsymbol{x}=\bar{\boldsymbol{x}}, \boldsymbol{y}=\bar{\boldsymbol{y}}}$$

- *Regression:*

$$y = \sum_j \beta_j x_j + \epsilon, \quad \epsilon_j = \frac{\bar{x}_j}{\bar{y}} \frac{\partial y}{\partial x_j} = \frac{\bar{x}_j}{\bar{y}} \hat{\beta}_j$$

- *Discrete-choice models:*
  Generally, with several endogenous variables, one distinguishes between
  - **Substitution** vs, **full/ordinary** elastities,
  - **Microscopic** vs, **macroscopic** elastities,
  - **proper elasticity** vs. **cross-elasticity**

? Why there are only substitution elastities in discrete-choice models?

! Because of the exclusivity condition on the alternatives

8.3 Elasticities

- *General definition:*

  Elasticities denote the percentaged change of endogenous variables y_i per small percentaged change of exogenous variables x_j for an average situation

  $$\epsilon_{ij} = \frac{\bar{x}_j}{\bar{y}_i} \frac{\partial y_i}{\partial x_j}\bigg|_{\boldsymbol{x}=\bar{\boldsymbol{x}},\boldsymbol{y}=\bar{\boldsymbol{y}}}$$

- *Regression:*

  $$y = \sum_j \beta_j x_j + \epsilon, \quad \epsilon_j = \frac{\bar{x}_j}{\bar{y}} \frac{\partial y}{\partial x_j} = \frac{\bar{x}_j}{\bar{y}} \hat{\beta}_j$$

- *Discrete-choice models:*
  Generally, with several endogenous variables, one distinguishes between
  - **Substitution** vs, **full/ordinary** elastities,
  - **Microscopic** vs, **macroscopic** elastities,
  - **proper elasticity** vs. **cross-elasticity**

? Why there are only substitution elastities in discrete-choice models?

! Because of the exclusivity condition on the alternatives

8.3 Elasticities

- *General definition:*

  Elasticities denote the percentaged change of endogenous variables y_i per small percentaged change of exogenous variables x_j for an average situation
  $$\epsilon_{ij} = \frac{\bar{x}_j}{\bar{y}_i} \frac{\partial y_i}{\partial x_j}\bigg|_{\boldsymbol{x}=\bar{\boldsymbol{x}},\boldsymbol{y}=\bar{\boldsymbol{y}}}$$

- *Regression:*
  $$y = \sum_j \beta_j x_j + \epsilon, \quad \epsilon_j = \frac{\bar{x}_j}{\bar{y}} \frac{\partial y}{\partial x_j} = \frac{\bar{x}_j}{\bar{y}} \hat{\beta}_j$$

- *Discrete-choice models:*
  Generally, with several endogenous variables, one distinguishes between
  - **Substitution** vs, **full/ordinary** elastities,
  - **Microscopic** vs, **macroscopic** elastities,
  - **proper elasticity** vs. **cross-elasticity**

? Why there are only substitution elastities in discrete-choice models?

! Because of the exclusivity condition on the alternatives

8.3 Elasticities

- *General definition:*

  Elasticities denote the percentaged change of endogenous variables y_i per small percentaged change of exogenous variables x_j for an average situation

  $$\epsilon_{ij} = \frac{\bar{x}_j}{\bar{y}_i} \frac{\partial y_i}{\partial x_j}\bigg|_{\boldsymbol{x}=\bar{\boldsymbol{x}}, \boldsymbol{y}=\bar{\boldsymbol{y}}}$$

- *Regression:*

  $$y = \sum_j \beta_j x_j + \epsilon, \quad \epsilon_j = \frac{\bar{x}_j}{\bar{y}} \frac{\partial y}{\partial x_j} = \frac{\bar{x}_j}{\bar{y}} \hat{\beta}_j$$

- *Discrete-choice models:*
  Generally, with several endogenous variables, one distinguishes between
  - **Substitution** vs, **full/ordinary** elastities,
  - **Microscopic** vs, **macroscopic** elastities,
  - **proper elasticity** vs. **cross-elasticity**

? Why there are only substitution elastities in discrete-choice models?

! Because of the exclusivity condition on the alternatives

8.3 Elasticities

- *General definition:*

  Elasticities denote the percentaged change of endogenous variables y_i per small percentaged change of exogenous variables x_j for an average situation
  $$\epsilon_{ij} = \frac{\bar{x}_j}{\bar{y}_i} \frac{\partial y_i}{\partial x_j}\bigg|_{\boldsymbol{x}=\bar{\boldsymbol{x}},\boldsymbol{y}=\bar{\boldsymbol{y}}}$$

- *Regression:*
  $$y = \sum_j \beta_j x_j + \epsilon, \quad \epsilon_j = \frac{\bar{x}_j}{\bar{y}} \frac{\partial y}{\partial x_j} = \frac{\bar{x}_j}{\bar{y}} \hat{\beta}_j$$

- *Discrete-choice models:*
  Generally, with several endogenous variables, one distinguishes between
  - **Substitution** vs, **full/ordinary** elastities,
  - **Microscopic** vs, **macroscopic** elastities,
  - **proper elasticity** vs. **cross-elasticity**

? Why there are only substitution elastities in discrete-choice models?

! Because of the exclusivity condition on the alternatives

8.3.1 Microscopic Logit elasticities

Since elasticities describe average aspects, we take the choice probabilities P_i rather than the discrete actual choices as endogenous variables. For the general deterministic utilities

$$V_{ni} = \sum_m \beta_{mi} x_{mni}$$

we derive the

- **Proper (substitution) elasticities:** The attribute (characteristic) m of an alternative i feeds back on the demand for this alternative:

$$\epsilon_{nii}^{(\text{mic,m})} = \frac{x_{mni}}{P_{ni}} \frac{\partial P_{ni}}{\partial x_{mni}} = \beta_m x_{mni} (1 - P_{ni})$$

- **Cross elasticities:** The attribute (characteristic) m of an alternative j feeds back on the demand for another alternative $i \neq j$:

$$\epsilon_{nij}^{(\text{mic,m})} = \frac{x_{nmj}}{P_{ni}} \frac{\partial P_{ni}}{\partial x_{nmj}} = -\beta_m x_{nmj} P_{nj}$$

8.3.1 Microscopic Logit elasticities

Since elasticities describe average aspects, we take the choice probabilities P_i rather than the discrete actual choices as endogenous variables. For the general deterministic utilities

$$V_{ni} = \sum_m \beta_{mi} x_{mni}$$

we derive the

- **Proper (substitution) elasticities:** The attribute (characteristic) m of an alternative i feeds back on the demand for this alternative:

$$\epsilon_{nii}^{(\text{mic,m})} = \frac{x_{mni}}{P_{ni}} \frac{\partial P_{ni}}{\partial x_{mni}} = \beta_m x_{mni}(1 - P_{ni})$$

- **Cross elasticities:** The attribute (characteristic) m of an alternative j feeds back on the demand for another alternative $i \neq j$:

$$\epsilon_{nij}^{(\text{mic,m})} = \frac{x_{nmj}}{P_{ni}} \frac{\partial P_{ni}}{\partial x_{nmj}} = -\beta_m x_{nmj} P_{nj}$$

Some questions on micro-elasticities

? Derive the formulas for the proper and cross elasticities

! We start with the normal MNL choice probability $P_{ni} = e^{V_{ni}} / \sum_k e^{V_{nk}}$ and first calculate the sensitivities in terms of the derivatives of V_{ni} with respect to x_{nmj}:

$$\frac{\partial P_{ni}}{\partial x_{nmj}} = \frac{e^{V_{ni}}}{\sum_k e^{V_{nk}}} \frac{\partial V_{ni}}{\partial x_{nmj}} - \frac{e^{V_{ni}}}{\left(\sum_k e^{V_{nk}}\right)^2} \frac{\partial}{\partial x_{nmj}} \left(\sum_l e^{V_{nl}} \right)$$

Some questions on micro-elasticities

? Derive the formulas for the proper and cross elasticities

! We start with the normal MNL choice probability $P_{ni} = e^{V_{ni}} / \sum_k e^{V_{nk}}$ and first calculate the sensitivities in terms of the derivatives of V_{ni} with respect to x_{nmj}:

$$\frac{\partial P_{ni}}{\partial x_{nmj}} = \frac{e^{V_{ni}}}{\sum_k e^{V_{nk}}} \frac{\partial V_{ni}}{\partial x_{nmj}} - \frac{e^{V_{ni}}}{\left(\sum_k e^{V_{nk}}\right)^2} \frac{\partial}{\partial x_{nmj}} \left(\sum_l e^{V_{nl}} \right)$$

Some questions on micro-elasticities

? Derive the formulas for the proper and cross elasticities

! We start with the normal MNL choice probability $P_{ni} = e^{V_{ni}} / \sum_k e^{V_{nk}}$ and first calculate the sensitivities in terms of the derivatives of V_{ni} with respect to x_{nmj}:

$$\begin{aligned}
\frac{\partial P_{ni}}{\partial x_{nmj}} &= \frac{e^{V_{ni}}}{\sum_k e^{V_{nk}}} \frac{\partial V_{ni}}{\partial x_{nmj}} - \frac{e^{V_{ni}}}{\left(\sum_k e^{V_{nk}}\right)^2} \frac{\partial}{\partial x_{nmj}} \left(\sum_l e^{V_{nl}} \right) \\
&= P_{ni} \frac{\partial V_{ni}}{\partial x_{nmj}} - \sum_l \frac{e^{V_{ni}} e^{V_{nl}}}{\left(\sum_k e^{V_{nk}}\right)^2} \frac{\partial V_{nl}}{\partial x_{nmj}}
\end{aligned}$$

Some questions on micro-elasticities

? Derive the formulas for the proper and cross elasticities

! We start with the normal MNL choice probability $P_{ni} = e^{V_{ni}} / \sum_k e^{V_{nk}}$ and first calculate the sensitivities in terms of the derivatives of V_{ni} with respect to x_{nmj}:

$$\begin{aligned}
\frac{\partial P_{ni}}{\partial x_{nmj}} &= \frac{e^{V_{ni}}}{\sum_k e^{V_{nk}}} \frac{\partial V_{ni}}{\partial x_{nmj}} - \frac{e^{V_{ni}}}{\left(\sum_k e^{V_{nk}}\right)^2} \frac{\partial}{\partial x_{nmj}} \left(\sum_l e^{V_{nl}} \right) \\
&= P_{ni} \frac{\partial V_{ni}}{\partial x_{nmj}} - \sum_l \frac{e^{V_{ni}} e^{V_{nl}}}{\left(\sum_k e^{V_{nk}}\right)^2} \frac{\partial V_{nl}}{\partial x_{nmj}} \\
&= P_{ni} \frac{\partial V_{ni}}{\partial x_{nmj}} - \sum_l P_{ni} P_{nl} \frac{\partial V_{nl}}{\partial x_{nmj}}
\end{aligned}$$

Some questions on micro-elasticities

? Derive the formulas for the proper and cross elasticities

! We start with the normal MNL choice probability $P_{ni} = e^{V_{ni}} / \sum_k e^{V_{nk}}$ and first calculate the sensitivities in terms of the derivatives of V_{ni} with respect to x_{nmj}:

$$\begin{aligned}\frac{\partial P_{ni}}{\partial x_{nmj}} &= \frac{e^{V_{ni}}}{\sum_k e^{V_{nk}}} \frac{\partial V_{ni}}{\partial x_{nmj}} - \frac{e^{V_{ni}}}{(\sum_k e^{V_{nk}})^2} \frac{\partial}{\partial x_{nmj}} \left(\sum_l e^{V_{nl}} \right) \\ &= P_{ni} \frac{\partial V_{ni}}{\partial x_{nmj}} - \sum_l \frac{e^{V_{ni}} e^{V_{nl}}}{(\sum_k e^{V_{nk}})^2} \frac{\partial V_{nl}}{\partial x_{nmj}} \\ &= P_{ni} \frac{\partial V_{ni}}{\partial x_{nmj}} - \sum_l P_{ni} P_{nl} \frac{\partial V_{nl}}{\partial x_{nmj}}\end{aligned}$$

where

$$\frac{\partial V_{ni}}{\partial x_{nmj}} = \beta_m \delta_{ij}, \quad \frac{\partial V_{nl}}{\partial x_{nmj}} = \beta_m \delta_{lj}$$

Some questions on micro-elasticities

? Derive the formulas for the proper and cross elasticities

! We start with the normal MNL choice probability $P_{ni} = e^{V_{ni}} / \sum_k e^{V_{nk}}$ and first calculate the sensitivities in terms of the derivatives of V_{ni} with respect to x_{nmj}:

$$\begin{aligned}
\frac{\partial P_{ni}}{\partial x_{nmj}} &= \frac{e^{V_{ni}}}{\sum_k e^{V_{nk}}} \frac{\partial V_{ni}}{\partial x_{nmj}} - \frac{e^{V_{ni}}}{(\sum_k e^{V_{nk}})^2} \frac{\partial}{\partial x_{nmj}} \left(\sum_l e^{V_{nl}} \right) \\
&= P_{ni} \frac{\partial V_{ni}}{\partial x_{nmj}} - \sum_l \frac{e^{V_{ni}} e^{V_{nl}}}{(\sum_k e^{V_{nk}})^2} \frac{\partial V_{nl}}{\partial x_{nmj}} \\
&= P_{ni} \frac{\partial V_{ni}}{\partial x_{nmj}} - \sum_l P_{ni} P_{nl} \frac{\partial V_{nl}}{\partial x_{nmj}}
\end{aligned}$$

where

$$\frac{\partial V_{ni}}{\partial x_{nmj}} = \beta_m \delta_{ij}, \quad \frac{\partial V_{nl}}{\partial x_{nmj}} = \beta_m \delta_{lj}$$

Hence

$$\frac{\partial P_{ni}}{\partial x_{nmj}} = \beta_m P_{ni} \left(\delta_{ij} - P_{nj} \right), \quad \epsilon_{nij}^{(\text{mic,m})} = \frac{x_{nmj}}{P_{ni}} \frac{\partial P_{ni}}{\partial x_{nmj}} = \beta_m x_{nmj} \left(\delta_{ij} - P_{nj} \right)$$

Some questions on micro-elasticities

? Derive the formulas for the proper and cross elasticities

! We start with the normal MNL choice probability $P_{ni} = e^{V_{ni}} / \sum_k e^{V_{nk}}$ and first calculate the sensitivities in terms of the derivatives of V_{ni} with respect to x_{nmj}:

$$\begin{aligned}
\frac{\partial P_{ni}}{\partial x_{nmj}} &= \frac{e^{V_{ni}}}{\sum_k e^{V_{nk}}} \frac{\partial V_{ni}}{\partial x_{nmj}} - \frac{e^{V_{ni}}}{\left(\sum_k e^{V_{nk}}\right)^2} \frac{\partial}{\partial x_{nmj}} \left(\sum_l e^{V_{nl}} \right) \\
&= P_{ni} \frac{\partial V_{ni}}{\partial x_{nmj}} - \sum_l \frac{e^{V_{ni}} e^{V_{nl}}}{\left(\sum_k e^{V_{nk}}\right)^2} \frac{\partial V_{nl}}{\partial x_{nmj}} \\
&= P_{ni} \frac{\partial V_{ni}}{\partial x_{nmj}} - \sum_l P_{ni} P_{nl} \frac{\partial V_{nl}}{\partial x_{nmj}}
\end{aligned}$$

where

$$\frac{\partial V_{ni}}{\partial x_{nmj}} = \beta_m \delta_{ij}, \quad \frac{\partial V_{nl}}{\partial x_{nmj}} = \beta_m \delta_{lj}$$

Hence

$$\frac{\partial P_{ni}}{\partial x_{nmj}} = \beta_m P_{ni} \left(\delta_{ij} - P_{nj} \right), \quad \epsilon_{nij}^{(\text{mic,m})} = \frac{x_{nmj}}{P_{ni}} \frac{\partial P_{ni}}{\partial x_{nmj}} = \beta_m x_{nmj} \left(\delta_{ij} - P_{nj} \right)$$

$j = i$: $\epsilon_{nii} = \beta_m x_{nmi} (1 - P_{ni})$, $\quad j \neq i$: $\epsilon_{nij} = -\beta_m x_{nmj} P_{nj}$

Questions (2)

? Derive and motivate the "null sum" condition $\sum_i P_{ni}\epsilon_{nij}^{(m)} = 0$

Questions (2)

? Derive and motivate the "null sum" condition $\sum_i P_{ni}\epsilon_{nij}^{(\mathrm{m})} = 0$

$$\sum_i P_{ni}\epsilon_{nij}^{(\mathrm{m})} = \sum_{i \neq j} P_{ni}\epsilon_{nij}^{(\mathrm{m})} + P_{ni}\epsilon_{nii}^{(\mathrm{m})}$$

Questions (2)

? Derive and motivate the "null sum" condition $\sum_i P_{ni}\epsilon_{nij}^{(\mathrm{m})} = 0$

$$\sum_i P_{ni}\epsilon_{nij}^{(\mathrm{m})} = \sum_{i\neq j} P_{ni}\epsilon_{nij}^{(\mathrm{m})} + P_{ni}\epsilon_{nii}^{(\mathrm{m})}$$

$$= -\sum_{i\neq j} P_{ni}\beta_m x_{nmj} P_{nj} + P_{ni}\beta_m x_{nmi}\left(1 - P_{ni}\right)$$

Questions (2)

? Derive and motivate the "null sum" condition $\sum_i P_{ni}\epsilon_{nij}^{(m)} = 0$

$$\sum_i P_{ni}\epsilon_{nij}^{(m)} = \sum_{i\neq j} P_{ni}\epsilon_{nij}^{(m)} + P_{ni}\epsilon_{nii}^{(m)}$$
$$= -\sum_{i\neq j} P_{ni}\beta_m x_{nmj} P_{nj} + P_{ni}\beta_m x_{nmi}\left(1 - P_{ni}\right)$$
$$= \beta_m\left(-\sum_i P_{ni}P_{nj}x_{nmj} + P_{ni}x_{nmi}\right) = 0$$

(Notice $\sum_i P_{ni} = 1$ in the last step!)

Questions (3)

? The cross elasticities do not depend on i, i.e., on the target alternative for the changing demand. Motivate this by the IIA condition

According to the IIA, if the utility of an alternative j changes, the changes of the relative preferences with respect to all other alternatives are the same. Moreover, the relative preferences are the probability ratios and their changes are the cross elasticities

? Given are three airports i from which person n can book flights to a desired destination at cost C_{ni} (because of revenue management, C depends on n), so

$$V_{ni} = \beta_{01}\delta_{01} + \beta_{02}\delta_{02} + \beta_1 C_{ni}$$

Show that the proper elasticities are negative while the cross elasticities are positive. Proper elasticity $\epsilon_{nii}^{(C)} = \beta_1 C_{ni}(1 - P_{ni}) < 0$ since $P_{ni} < 1$ $C_{ni} > 0$, and the price sensitivity $\beta_1 < 0$. The cross elasticities $\epsilon_{nji}^{(C)} = -\beta_1 C_{nj} P_{nj}$ are therefore positive.

Questions (3)

? The cross elasticities do not depend on i, i.e., on the target alternative for the changing demand. Motivate this by the IIA condition
According to the IIA, if the utility of an alternative j changes, the changes of the relative preferences with respect to all other alternatives are the same. Moreover, the relative preferences are the probability ratios and their changes are the cross elasticities

? Given are three airports i from which person n can book flights to a desired destination at cost C_{ni} (because of revenue management, C depends on n), so

$$V_{ni} = \beta_{01}\delta_{01} + \beta_{02}\delta_{02} + \beta_1 C_{ni}$$

Show that the proper elasticities are negative while the cross elasticities are positive.
Proper elasticity $\epsilon^{(C)}_{nii} = \beta_1 C_{ni}(1 - P_{ni}) < 0$ since $P_{ni} < 1$ $C_{ni} > 0$, and the price sensitivity $\beta_1 < 0$. The cross elasticities $\epsilon^{(C)}_{nii} = -\beta_1 C_{nj} P_{nj}$ are therefore positive.

Questions (3)

? The cross elasticities do not depend on i, i.e., on the target alternative for the changing demand. Motivate this by the IIA condition
According to the IIA, if the utility of an alternative j changes, the changes of the relative preferences with respect to all other alternatives are the same. Moreover, the relative preferences are the probability ratios and their changes are the cross elasticities

? Given are three airports i from which person n can book flights to a desired destination at cost C_{ni} (because of revenue management, C depends on n), so

$$V_{ni} = \beta_{01}\delta_{01} + \beta_{02}\delta_{02} + \beta_1 C_{ni}$$

Show that the proper elasticities are negative while the cross elasticities are positive.
Proper elasticity $\epsilon_{nii}^{(C)} = \beta_1 C_{ni}(1 - P_{ni}) < 0$ since $P_{ni} < 1$ $C_{ni} > 0$, and the price sensitivity $\beta_1 < 0$. The cross elasticities $\epsilon_{nii}^{(C)} = -\beta_1 C_{nj} P_{nj}$ are therefore positive.

Questions (3)

? The cross elasticities do not depend on i, i.e., on the target alternative for the changing demand. Motivate this by the IIA condition
According to the IIA, if the utility of an alternative j changes, the changes of the relative preferences with respect to all other alternatives are the same. Moreover, the relative preferences are the probability ratios and their changes are the cross elasticities

? Given are three airports i from which person n can book flights to a desired destination at cost C_{ni} (because of revenue management, C depends on n), so

$$V_{ni} = \beta_{01}\delta_{01} + \beta_{02}\delta_{02} + \beta_1 C_{ni}$$

Show that the proper elasticities are negative while the cross elasticities are positive.
Proper elasticity $\epsilon_{nii}^{(C)} = \beta_1 C_{ni}(1 - P_{ni}) < 0$ since $P_{ni} < 1$ $C_{ni} > 0$, and the price sensitivity $\beta_1 < 0$. The cross elasticities $\epsilon_{nii}^{(C)} = -\beta_1 C_{nj} P_{nj}$ are therefore positive.

8.3.2 Macroscopic elasticities

For a company, the relative probability increase of single customers chosing their products is not relevant but the aggregate over all customers. Hence, the macroscopic elasticity

$$\epsilon_{ij}^{(\text{mac,m})} = \frac{X_{mj}}{N_i}\frac{\partial N_i}{\partial X_{mj}}, \quad X_{mj} = \sum_{n=1}^{N} x_{nmj}, \quad N_i = \sum_{n=1}^{N} P_{ni}$$

gives the percentage increase of people chosing alternative i when the sum of attributes m increases at alternative j by one percent.

(i) Same absolute changes for all persons, $\mathrm{d}x_{nmj} = \mathrm{d}X_{mj}/N$:

$$\epsilon_{ij}^{(\text{mac,abs,m})} = \frac{X_{mj}}{N_i}\frac{1}{N}\sum_n \frac{P_{ni}}{x_{nmj}}\epsilon_{nij}^{(\text{mic,m})}$$

(ii) Same relatives changes for all, $\mathrm{d}x_{nmj}/x_{nmj} = \mathrm{d}X_{mj}/X_{mj}$:

$$\epsilon_{ij}^{(\text{mac,rel,m})} = \sum_n w_{ni}\,\epsilon_{nij}^{(\text{mic,m})}, \quad w_{ni} = \frac{P_{ni}}{N_i} = \frac{P_{ni}}{\sum_n P_{ni}}$$

8.3.2 Macroscopic elasticities

For a company, the relative probability increase of single customers chosing their products is not relevant but the aggregate over all customers. Hence, the macroscopic elasticity

$$\epsilon_{ij}^{\text{(mac,m)}} = \frac{X_{mj}}{N_i}\frac{\partial N_i}{\partial X_{mj}}, \quad X_{mj} = \sum_{n=1}^{N} x_{nmj}, \quad N_i = \sum_{n=1}^{N} P_{ni}$$

gives the percentage increase of people chosing alternative i when the sum of attributes m increases at alternative j by one percent.

(i) Same absolute changes for all persons, $\mathrm{d}x_{nmj} = \mathrm{d}X_{mj}/N$:

$$\epsilon_{ij}^{\text{(mac,abs,m)}} = \frac{X_{mj}}{N_i}\frac{1}{N}\sum_n \frac{P_{ni}}{x_{nmj}}\epsilon_{nij}^{\text{(mic,m)}}$$

(ii) Same relatives changes for all, $\mathrm{d}x_{nmj}/x_{nmj} = \mathrm{d}X_{mj}/X_{mj}$:

$$\epsilon_{ij}^{\text{(mac,rel,m)}} = \sum_n w_{ni}\,\epsilon_{nij}^{\text{(mic,m)}}, \quad w_{ni} = \frac{P_{ni}}{N_i} = \frac{P_{ni}}{\sum_n P_{ni}}$$