



Universidade de São Paulo  
Escola Politécnica  
Departamento de Engenharia de Transportes

## Divisão Modal

**PTR 3431**  
**Planejamento e Operação de Sistemas de**  
**Transportes**

**Prof. Dr. Cassiano A. Isler**  
**2019.1**



- Definições
- Modelagem de Escolha Modal
  - Definições
  - Modelo Logit
  - Agregação
- Função Utilidade
  - Coleta de Dados
  - Pesquisas de Preferência Revelada
  - Pesquisas de Preferência Declarada
  - Estimativa de Parâmetros
- Modelos de Demanda Direta



# Definições

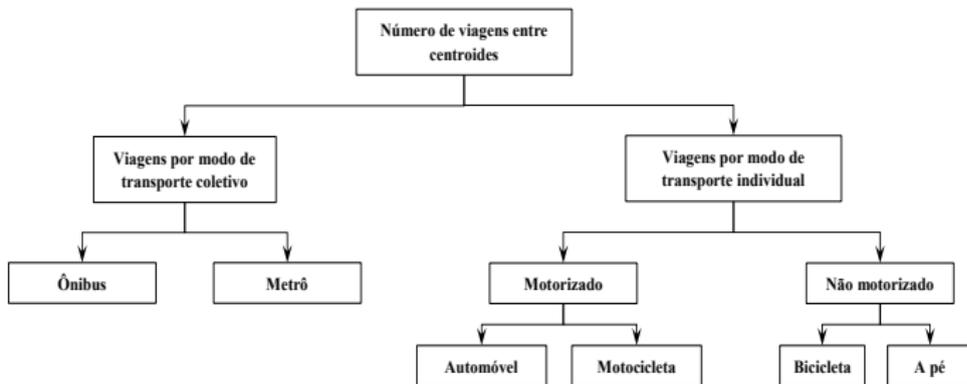


A pergunta que se deseja responder nesta etapa é:

**Qual é o número de viagens ( $T_{ij}^k$ ) entre pares de zonas de tráfego em um dado período de tempo ?**

O objetivo da etapa de “Divisão Modal” é estimar o número de viagens por modo de transporte dada a matriz de deslocamentos totais obtidas na etapa anterior de Distribuição de Viagens.

Em outras palavras, os métodos utilizados nesta etapa buscam dividir a matriz de viagens em matrizes parciais por modo de transporte, em que os elementos de cada matriz OD entre a origem  $i$  e o destino  $j$  para o modo de transporte  $k$  é representado por  $T_{ij}^k$ .





- **Pesquisa**

Questionário aplicado a usuários ou potenciais usuários de um sistema de transporte existente ou em planejamento

- **Alternativa**

Opções de escolha modal para atividade de transporte

- **Atributo**

Característica da viagem (tempo, custo, frequência etc.) e do usuário (renda, posse de automóvel etc.) considerado na análise da escolha modal



- **Nível do atributo**

Valor numérico associado ao atributo

- **Modelagem de escolha modal**

Técnicas para estimativa das escolhas modais provenientes das respostas da pesquisa

- **Modelos de escolha discreta**

Técnica específica para modelagem de de escolha modal a partir das respostas da pesquisa



# Modelagem de Escolha Modal



A **modelagem da escolha modal** consiste na aplicação de métodos matemáticos para estimativa de modelos que permitam representar escolhas de modos de transporte.

Esses modelos são úteis para estimar as escolhas mediante os níveis dos atributos dos modos de transporte e das características dos participantes da pesquisa, e também para extrapolação dessas estimativas para cenários em que os níveis dos atributos não foram apresentados na pesquisa.

O resultado dos **modelos de escolha discreta** é definido por Ortúzar e Wilumsen (2011) como *"a probabilidade dos indivíduos escolherem uma opção entre um conjunto de alternativas preestabelecidos como função das suas características socioeconômicas e da atratividade relativa da opção escolhida"*.



## Definições

A modelagem de escolha discreta baseia-se na **Teoria de Utilidade** e considera as seguintes:

- Os indivíduos de uma população  $\mathbf{Q} = \{1, \dots, q\}$  agem racionalmente e buscam maximizar sua utilidade pessoal segundo restrições físicas e de orçamento;
- Existe um conjunto de alternativas  $\mathbf{K} = \{1, \dots, k\}$  e um conjunto  $\mathbf{X} = \{1, \dots, x_{kqm}\}$  de  $m$  atributos relacionados aos  $q$  indivíduos e às  $k$  alternativas para escolha.
- Cada alternativa correspondente ao indivíduo  $q$  possui uma utilidade  $U_{kq}$  associada.



A utilidade de cada alternativa é definida por uma parcela determinística  $V_{kq}$  e um erro aleatório  $\varepsilon_{kq}$ .

$$U_{kq} = V_{kq} + \varepsilon_{kq}$$

A parcela  $V_{kq}$  pode ser definida como uma função linear ponderada de  $m$  atributos ( $x_{kqm}$ ).

$$V_{kq} = \sum_m \beta_{km} \cdot x_{kqm}$$

O indivíduo seleciona a alternativa que maximiza a sua utilidade, ou seja, a alternativa  $A_j$  é escolhida se

$$U_{kq} \geq U_{lq} \quad \forall \quad k \neq l$$

Como existe um erro aleatório, só é possível estimar a probabilidade de escolha do modo  $k$  pelo indivíduo  $q$ , tal que:

$$P_{kq} = P(U_{kq} \geq U_{lq} \quad \forall \quad k \neq l)$$



Se os resíduos  $\varepsilon_{kq}$  forem representados por uma função de distribuição de probabilidade **Gumbel**, então a probabilidade de escolha do modo  $k$  pelo indivíduo  $q$  é dada pelo modelo **Logit** conforme a equação:

$$P_{kq} = \frac{e^{V_{kq}}}{\sum_a e^{V_{aq}}} \quad \forall \quad k$$

Quando existem apenas duas alternativas de escolha modal, o modelo é denominado **Logit Binomial**.

$$P_{1q} = \frac{e^{V_{1q}}}{e^{V_{1q}} + e^{V_{2q}}} \quad \text{e} \quad P_{2q} = 1 - P_{1q} = \frac{e^{V_{2q}}}{e^{V_{2q}} + e^{V_{1q}}}$$

Para mais de duas alternativas de escolha modal o modelo é denominado **Logit Multinomial (MNL)**.



Os modelos de escolha discreta segundo a Teoria da Utilidade baseiam-se em métodos para cálculo da probabilidade de escolha modal de indivíduos da população, ou seja, em estimativas de natureza desagregada.

Entretanto, as estimativas dessas probabilidades em termos individuais são pouco relevantes para um analista, tal que a maioria das decisões reais sobre planejamento de transporte baseiam-se em estimativas de demanda agregada.

Portanto, é necessário estabelecer algum tipo de relação entre as estimativas dos modelos desagregados de escolha discreta em relação à população de interesse, ou seja, em um nível agregado de previsão de demanda.



# Agregação

O problema de **agregação de escolha modal** consiste em extrapolar as probabilidades de escolhas individuais para um nível agregado da população.

O valor esperado da divisão modal para a população pode ser calculado por um indivíduo caracterizado pelos valores médios dos atributos de todos os indivíduos da população.

Seja  $\bar{x}$  o vetor de valores médios dos atributos da população, então a divisão modal pode ser calculada por:

$$W_k = \sum_{q=1}^N P_{kq}(\bar{x})$$

onde  $P_{kq}$  é a probabilidade de escolha da alternativa  $k$  mediante os valores dos atributos  $\bar{x}$  para o  $q$ -ésimo indivíduo médio da população com  $N$  indivíduos.



Se a dispersão dos valores de utilidade para as alternativas ( $V_{kq}$ ) for grande, então é possível segregar os indivíduos em agrupamentos com características próximas.

Nesse caso a divisão modal é calculada por:

$$W_q = \sum_{g=1}^G \frac{N_g}{N} \cdot P_{kg}(\bar{x}_g)$$

onde  $P_{kq}$  é a probabilidade de escolha da alternativa  $k$  no grupo  $g$  mediante os valores médios dos atributos  $\mathbf{x}_g$ .

Se  $G = 1$  então  $N_g = N$  e a divisão modal por agrupamento é equivalente àquela por valores médios dos atributos.

O zoneamento definido no Modelo 4 Etapas pode ser uma boa estratégia para caracterização desses agrupamentos.



## Exemplo

Considere as seguintes equações para representar as parcelas determinísticas das funções utilidade de um indivíduo  $q$  para estimativa da probabilidade e escolha de dois modos de transporte: automóvel ( $car$ ) e ônibus ( $bus$ ).

$$V_{cq} = 3,5 - 0,25 \cdot t_c - 0,42 \cdot e_c - 0,1 \cdot c_c$$

$$V_{bq} = -0,25 \cdot t_b - 0,42 \cdot e_b - 0,1 \cdot c_b$$

onde  $t$  é o tempo de viagem (minutos),  $e$  é o tempo de acesso (minutos) e  $c$  é o custo da viagem (\$).



## Exemplo

Assumindo os níveis dos atributos abaixo, calcule os itens a seguir considerando modelagem Logit Binomial.

Mode	Variable		
	$t$	$e$	$c$
Car	25	5	140
Bus	40	8	50

- (a) as probabilidades de escolha dos modos de transporte mediante os níveis indicados para os atributos.
- (b) a variação percentual na probabilidade de escolha do automóvel se a tarifa do ônibus aumentar 25%.
- (c) a variação percentual na probabilidade de escolha do automóvel se o custo da viagem por esse modo aumentar em 20 unidades.



- (d) Considere que a seguinte matriz OD foi obtida anteriormente da etapa de Distribuição de Viagens.

Origem	Destino		
	1	2	3
1	0	307	407
2	420	0	212
3	412	246	0

Estime o número de viagens por modo de transporte (automóvel e ônibus) entre cada par Origem Destino.



# Estimativa de Funções Utilidade



Em geral, a coleta de dados a partir de pesquisas é a etapa que absorve a maior parte de recursos disponíveis para conduzir estudos de transportes.

A escolha do método de coleta de dados depende basicamente do tipo de modelo que será utilizado para representar os dados obtidos provenientes de questionários conduzidos com um população de interesse.

Desse modo, a compreensão sobre a utilização do uso dos dados é um dos elementos cruciais na definição do método de qualquer tipo de pesquisa.

Em geral o requisito principal para os dados provenientes de uma pesquisa é a possibilidade de realizar previsões apuradas sobre o comportamento do sistema no futuro sobre condições de alterações específicas desse sistema.



É desejável que conjunto de dados obtidos de uma pesquisa de transporte tenha as seguintes propriedades:

- garanta que as análises possam ser por modo de transporte entre localidades em períodos específicos;
- inclua todos os modos de transportes inclusive os não motorizados;
- forneça informações nos níveis mais desagregados possíveis por motivo de viagem;
- abranja o período de tempo mais amplo possível;
- forneça dados sobre unidades familiares;
- esteja em um sistema integrado de bases de dados.

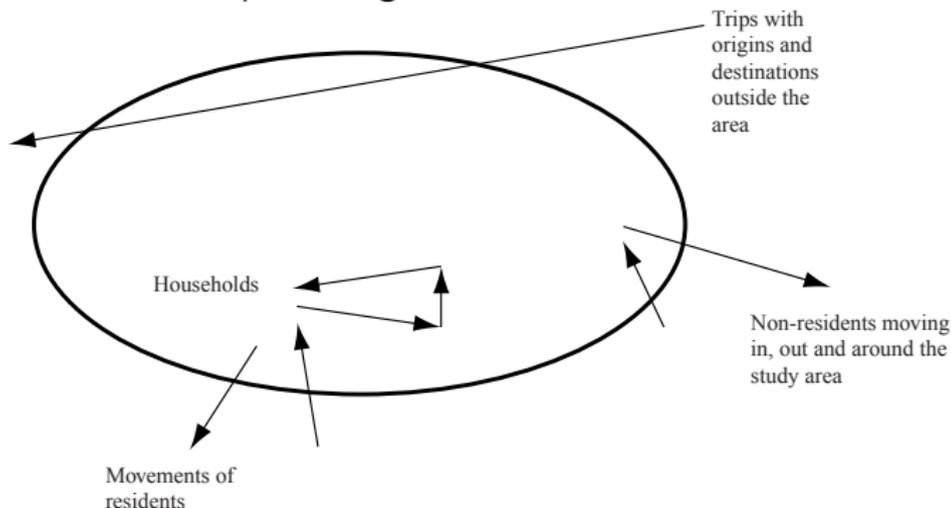


As restrições relacionadas à realização de pesquisas de transporte são:

- duração do estudo em termos de tempo disponível para aquisição de dados e modelagem;
- recursos econômicos e humanos disponíveis;
- horizonte de planejamento a que se propõe o estudo, se em nível do ano-base ou em horizonte estratégico;
- aplicação da pesquisa, como condições climáticas, topográficas, sociais e de tecnologias disponíveis.



Ao definir o escopo da pesquisa, a primeira etapa consiste em delimitar a área de interesse para pesquisa, a qual pode basear-se na ilustração a seguir.



As barreiras físicas naturais (rios, montanhas etc.) ou artificiais (ferrovia, rodovia etc.) que delimitam a área da pesquisa são denominadas *screen lines*.



As pesquisas podem ser:

- **Por unidade familiar:** membros de uma unidade familiar dentro da área de estudo, e na entrada e saída dessas áreas;
- **Nos limites da área:** pessoas que cruzam a fronteira da região (residentes e não residentes);

As pesquisas de demanda de transportes para modelagem de divisão modal podem ser de dois tipos:

## **Pesquisas de Preferência Revelada**

Obtenção de informações sobre escolhas modais em **viagens realizadas** pelos usuários.

## **Pesquisas de Preferência Declarada**

Obtenção de dados sobre o comportamento dos usuários a partir de escolhas em **cenários hipotéticos**.



# Pesquisas de Preferência Revelada

As pesquisas de **preferência revelada** têm o objetivo de coletar informações sobre o comportamento de uma amostra da população em uma região específica a partir de um questionário sobre viagens realizadas em algum período do passado utilizando-se o sistema de transporte existente e disponibilizado aos usuários.

Esse tipo de pesquisa é custoso, porém é a que provê grande quantidade de dados uma vez que medidas empíricas são as principais fontes de informações sobre o comportamento de viagens para planejamento de transportes.





Uma **pesquisa de preferência declarada** é um instrumento de obtenção de respostas de potenciais usuários de um sistema de transporte hipotético, mediante a apresentação de cenários com diferentes alternativas de escolha modal, caracterizadas por atributos sob diferentes níveis.

## Cenário

	Train	Bus		Car
Fare	\$3.00	\$4.00	Petrol Costs	\$1.00
			Toll Cost	\$3.00
			Parking Cost	\$8.00
Access Time	5 mins	10 mins		
In-vehicle Time	35 mins	25 mins	In-vehicle Time	15 mins
Egress Time	15 mins	10 mins	Egress Time	5 mins
I would choose	<input type="radio"/>	<input type="radio"/>	or	<input type="radio"/>



**Alternativa**

**Nível**

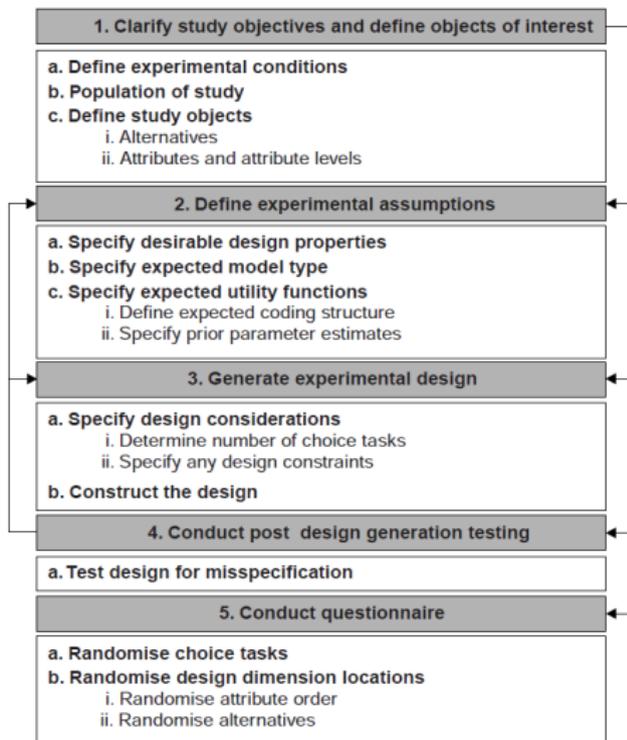
**Atributo**

	Train	Bus		Car
Fare	\$3.00	\$4.00	Petrol Costs	\$1.00
			Toll Cost	\$3.00
			Parking Cost	\$8.00
Access Time	5 mins	10 mins		
In-vehicle Time	35 mins	25 mins	In-vehicle Time	15 mins
Egress Time	15 mins	10 mins	Egress Time	5 mins
I would choose	<input type="radio"/>	<input type="radio"/>	or	<input type="radio"/>

# Pesquisas de Preferência Declarada



As etapas de uma pesquisa de preferência declarada podem ser resumidas pelo fluxograma.





A modelagem de escolha discreta permite calcular a escolha mais provável de modo de transporte sob diferentes hipóteses para a distribuição da probabilidade dos erros aleatórios as funções de idades desses modos.

Portanto, dado que o modelo matemático de cálculo de probabilidade é conhecido (Logit ou Probit), e assumindo que o formato da equação da parcela determinística de cada função utilidade do modo  $k$  para o indivíduo  $q$  ( $V_{kq}$ ) também é previamente definida (geralmente sobre formato aditivo), surge o problema de estimativa dos  $m$  parâmetros ( $\beta_{km}$ ) dessa função matemática.

$$V_{kq} = \sum_m \beta_{km} \cdot x_{kqm}$$



Formalmente, o problema abordado é de estimativa dos parâmetros da função utilidade representada com uma equação em relação aos atributos dos modos de transporte e dos indivíduos envolvidos no processo de escolha modal.

Essa estimativa é realizada a partir das escolhas modais dos indivíduos que responderam a pesquisa de preferência declarada.

Em geral, considera-se uma proporção da amostra de respostas para calibração dos parâmetros  $\beta_{km}$  (por exemplo 75%) e a outra parcela para validação das estimativas (no caso, 25%).



Seja  $P_{kq}$  que a probabilidade da escolha do modo  $k$  pelo indivíduo  $q$ , tal que, por exemplo,  $P_{32}$  é a probabilidade do indivíduo 2 escolher alternativa 3.

Seja  $g_{kq}$  um parâmetro binário tal que:

$$g_{kq} = \begin{cases} 1, & \text{se o indivíduo } q \text{ escolhe a alternativa } k \\ 0, & \text{caso contrário.} \end{cases}$$

Os valores de  $\beta_{km}$  que mais aproximam as estimativas do modelo de escolha discreta às escolhas da pesquisa de preferência declarada são aqueles que maximizam a “função de verossimilhança” dada por:

$$l(\beta) = \sum_q \sum_k g_{kq} \cdot \log(P_{kq})$$

onde  $P_{kq}$  é dado pela equação do modelo Logit.



Por exemplo, considere três cenários de pesquisa de preferência declarada para escolha de dois modos de transportes (1 e 2), cada um caracterizado por um respectivo atributo apresentado ao indivíduo  $q$ , dados por  $x_{1q}$  e  $x_{2q}$ .

Considere ainda a proposição de parcelas determinísticas de funções utilidade com apenas um parâmetro  $\beta$  comum a todos os atributos:

$$V_{1q} = \beta \cdot x_{1q} \qquad V_{2q} = \beta \cdot x_{2q}$$

Sob essas condições, a probabilidade de escolha das alternativas 1 e 2 por um modelo Logit Binomial é dada por:

$$P_{1q} = \frac{e^{V_{1q}}}{e^{V_{1q}} + e^{V_{2q}}} \qquad P_{2q} = 1 - P_{1q} = \frac{e^{V_{2q}}}{e^{V_{1q}} + e^{V_{2q}}}$$



Após coletar a escolha modal de três indivíduos em uma pesquisa de preferência declarada, obteve-se a seguinte tabelas para diferentes níveis de atributos:

Observação ( $q$ )	Escolha Modal	$x_{1q}$	$x_{2q}$
1	1	5	3
2	1	1	2
3	2	3	4

Nesse caso, os parâmetros  $g_{kq}$  podem ser caracterizados por:

		Alternativa $k$	
		1	2
Indivíduo $q$	1	1	0
	2	1	0
	3	0	1



## Estimativa de Parâmetros

A função de verossimilhança linearizada para um modelo Logit Binomial de estimativa das probabilidades de escolha desses modos de transporte, para cálculo do parâmetro  $\hat{\beta}$  que mais aproxima essas probabilidades às escolhas dos três cenários da pesquisa é dado por:

$$l(\beta) = g_{11} \cdot \log(P_{11}) + g_{21} \cdot \log(P_{21}) + g_{12} \cdot \log(P_{12}) \\ + g_{22} \cdot \log(P_{22}) + g_{13} \cdot \log(P_{13}) + g_{23} \cdot \log(P_{23})$$

Substituindo os valores de  $g_{kq}$  na equação tem-se que:

$$l(\beta) = \log(P_{11}) + \log(P_{12}) + \log(P_{23}) \\ = \log \left[ \frac{1}{1 + e^{\beta(x_{21} - x_{11})}} \right] + \log \left[ \frac{1}{1 + e^{\beta(x_{21} - x_{11})}} \right] + \log \left[ \frac{1}{1 + e^{\beta(x_{13} - x_{23})}} \right] \\ = \log \left[ \frac{1}{1 + e^{\beta(3-5)}} \right] + \log \left[ \frac{1}{1 + e^{\beta(2-1)}} \right] + \log \left[ \frac{1}{1 + e^{\beta(3-4)}} \right]$$

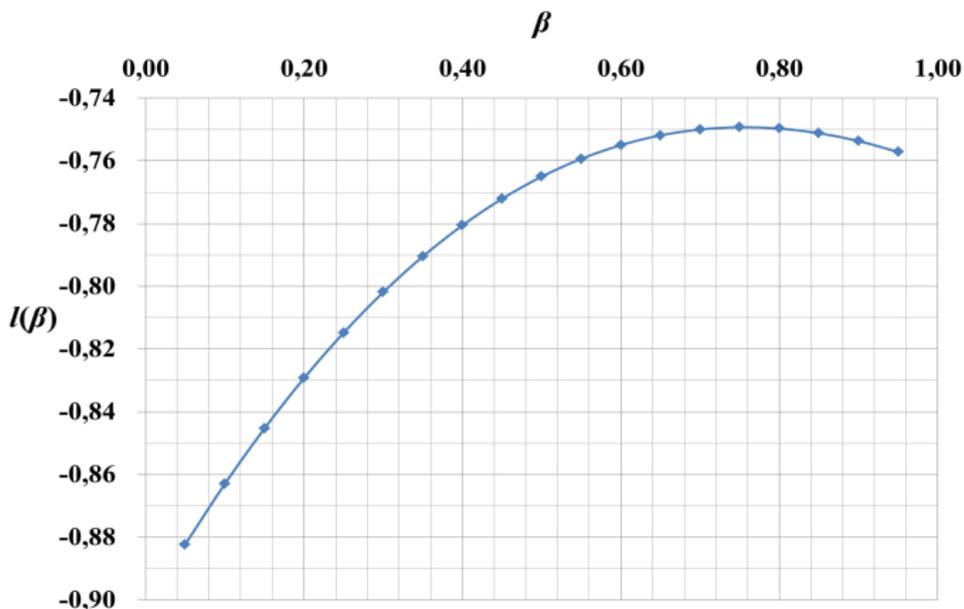
Simplificando a equação de  $l(\beta)$  em função de  $\beta$  tem-se:

$$l(\beta) = 10 \cdot \beta - \log(e^{5 \cdot \beta} + e^{3 \cdot \beta}) - \log(e^{\beta} + e^{2 \cdot \beta}) - \log(e^{3 \cdot \beta} + e^{4 \cdot \beta})$$



# Estimativa de Parâmetros

A função de verossimilhança pode ser representada graficamente para identificação de  $\beta$  que maximiza  $l(\beta)$ .



O valor de  $\beta$  que maximiza a função é 0,756.



Existem softwares específicos para estimativa de múltiplos parâmetros das funções utilidade dos modos de transporte, por exemplo, o Biogeme, Statistica, SPSS etc.

Como exemplo, preencha a pesquisa de preferência declarada no link do *QR Code* com cenários de escolha modal de transporte público para uma viagem a lazer considerando a oferta de transporte ferroviário.

Considerando tempo de viagem e custo, qual modo de transporte você utilizaria para uma viagem individual por motivo de lazer entre Campinas e São Paulo?



Tempo: 87 minutos  
Custo: R\$ 60,00



Tempo: 54 minutos  
Custo R\$ 60,00





# Modelos de Demanda Direta



Os modelos de demanda direta resultam em estimativas do número de viagens entre pares OD por modo de transporte e, por isso, são métodos de estimativa de geração, distribuição de viagens e divisão modal simultaneamente.

$$T_{ij}^k = A_i \cdot O_i \cdot B_j \cdot D_j \cdot \exp^{-\beta \cdot c_{ij}^k}$$

A probabilidade de escolha do  $k$ -ésimo modo de transporte para um modelo Logit Multinomial entre um par OD.

$$P_{ij}^k = \frac{T_{ij}^k}{T_{ij}} = \frac{\exp^{-\beta \cdot c_{ij}^k}}{\sum_m \exp^{-\beta \cdot c_{ij}^m}}$$

Para um modelo Logit Binomial a probabilidade de escolha do modo  $k = 1$  é:

$$P_{ij}^1 = \frac{T_{ij}^1}{T_{ij}} = \frac{\exp^{-\beta \cdot c_{ij}^1}}{\exp^{-\beta \cdot c_{ij}^1} + \exp^{-\beta \cdot c_{ij}^2}}$$



Considere um modelo de escolha binária com modos de transporte Automóvel (*car*) e Ônibus (*bus*), cujos custos generalizados entre pares OD são dados por:

$$V_{car} = -0,15 \cdot T_{car}$$

$$V_{bus} = -0,5 - 0,15 \cdot T_{bus}$$

onde  $T_{car}$  é o tempo de viagem pelo modo Automóvel,  $T_{bus}$  é o tempo de viagem pelo modo Ônibus.

Considere a matriz OD de viagens entre zonas de tráfego na região central de São Paulo obtida a partir da aplicação do modelo gravitacional na etapa de Distribuição de Viagens.

# Exercício



Considere ainda as matrizes de tempos de viagem (minutos) pelo modo Automóvel ( $T_{car}$ ) e Ônibus ( $T_{bus}$ ).

Tempos de Viagem Automóvel

		ZONA DE DESTINO					
		1	2	3	4	5	6
ZONA DE ORIGEM	1	0	13	14	14	21	13
	2	20	0	25	27	27	21
	3	19	18	0	19	27	17
	4	18	13	20	0	11	17
	5	24	24	13	17	0	25
	6	25	19	18	16	13	0

Tempos de Viagem Ônibus

		ZONA DE DESTINO					
		1	2	3	4	5	6
ZONA DE ORIGEM	1	0	16	18	18	27	14
	2	26	0	33	32	35	23
	3	21	20	0	21	32	20
	4	20	14	26	0	12	20
	5	26	31	14	19	0	33
	6	33	21	20	18	14	0

Estime as matrizes OD de número de viagens pelo dois modos de transporte.