

*Facoltà di Ingegneria - Università di Bologna*

Anno Accademico: 2010/11

# **TECNICA ED ECONOMIA DEI TRASPORTI**

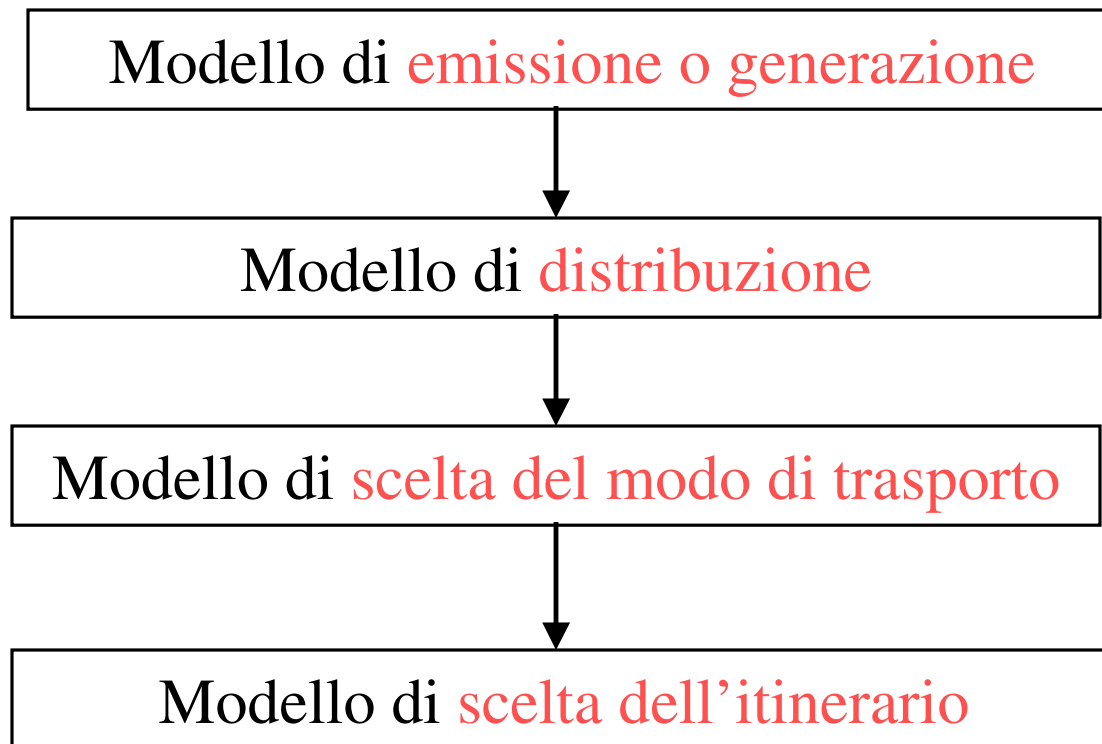
Docente: Marino Lupi

## ***IL MODELLO A 4 STADI***

### ***(Capitolo 3)***

# Modello a 4 stadi per la determinazione della domanda di trasporto in area urbana

Si suddivide la *scelta "totale"* del viaggio ( se muoversi o meno, verso quale destinazione, con quale modo di trasporto, lungo quale itinerario) in una *successione di scelte parziali*.



## Tipo di approccio prevalente nella modellizzazione della domanda nei “vari *stadi*”



$$D^i(h, s, o, d, m, k) = n^i(o) p^i(osh) p^i(d / osh) p^i(m / dosh) p^i(k / mdosh)$$

$i$ : *categoria di utenti* considerata. Per esempio: attivi (lavoratori) residenti in  $o$  per i movimenti casa-lavoro; studenti residenti in  $o$  per i movimenti casa-scuola.

$h$ : *intervallo temporale*. Per esempio: ora di punta del mattino (7-9), oppure giornata feriale media.

$s$ : *motivo* dello spostamento. Per esempio: casa-lavoro; casa- scuola; casa- acquisti.

$od$ : coppia di zone *origine-destinazione*.

$m$ : *modo di trasporto*. Per esempio: autovettura, autobus, bicicletta, a piedi.

$k$ : *percorso*. In generale fra una certa coppia di zone  $od$  ho più itinerari con un certo modo (per esempio se mi sposto con il modo individuale).

$$D^i(h, s, o, d, m, k) = n^i(o) p^i(osh) p^i(d / osh) p^i(m / dosh) p^i(k / mdosh)$$

### *Modello di generazione (o di emissione)*

$$D^i(osh) = n^i(o) p^i(osh)$$

$D^i(osh)$  Numero di utenti della categoria  $i$  che si sposta nell'intervallo di tempo  $h$ , per il motivo  $s$ , dalla zona  $o$ .

$n^i(o)$  Numero di appartenenti alla categoria  $i$  che sono in  $o$ .

$p^i(osh)$  *percentuale di utenti che si sposta*: indica il numero medio di spostamenti di utenti della categoria  $i$ , nell'intervallo  $h$ , per il motivo  $s$ , partendo dalla zona  $o$ . Per esempio: percentuali di attivi (lavoratori), residente in  $o$ , che si sposta, nell'ora di punta del mattino, per il motivo casa- lavoro.

$p^i(osh)$

Ho definito una percentuale, numero minore di 1, ma in realtà se  $h$  è ampio, per esempio tutta una giornata, un utente si può spostare da  $o$  più volte.

$m^i(osh)$

Perciò, più propriamente, parlo di *indice di emissione* per la categoria  $i$  ( $m^i$  può essere maggiore di 1).

Modello di *generazione*



$$D^i(osh) = n^i(o) m^i(osh)$$

## *Modello di distribuzione*

$$D^i(h, s, o, d, m, k) = n^i(o) p^i(osh) p^i(d / osh) p^i(m / dosh) p^i(k / mdosh)$$

$$p^i(d / osh)$$

*Il modello di distribuzione* fornisce la *percentuale di individui* di categoria  $i$  che, spostandosi nel periodo di tempo  $h$ , per il motivo  $s$ , da  $o$ , *si recano nella zona di destinazione  $d$ .*

## *Modello di scelta modale*

$$D^i(h, s, o, d, m, k) = n^i(o) p^i(osh) p^i(d / osh) p^i(m / dosh) p^i(k / mdosh)$$

$$p^i(m / dosh)$$

Il *modello di scelta modale* fornisce la *percentuale di individui* di categoria  $i$  che spostandosi, nel periodo di tempo  $h$ , per il motivo  $s$ , fra  $o$  e  $d$ , *sceglie il modo di trasporto  $m$ .*



## *Modello di scelta modale*

$$D^i(h, s, o, d, m, k) = n^i(o) p^i(osh) p^i(d / osh) p^i(m / dosh) p^i(k / mdosh)$$

$$p^i(k / mdosh)$$

*Il modello di scelta del percorso* fornisce la *percentuale di individui* di categoria  $i$  che spostandosi, nel periodo di tempo  $h$ , per il motivo  $s$ , fra  $o$  e  $d$ , con il modo  $m$ , *sceglie l'itinerario  $k$ .*

## MODELLO DI EMISSIONE

*Indici di emissione* per categoria. Spostamenti medi *giornalieri stimati* in *5 città italiane* di *medie dimensioni*.

Valori stimati su base *indagine campionaria* (molto dettagliata, numerose categorie e motivi considerati).

Motivo	Categoria	Indice di emissione
Casa-lavoro	Attivo settore industria	1,024
	Attivo settore servizi	1,084
	Attivo settore servizi privati	1,245
	Attivo settore servizi pubblici	0,931
Casa-scuola	Alunni scuole elementari	0,84
	Studenti scuole medie inferiori	0,87
	Studenti scuole superiori	0,86
	Studenti istituti professionali	0,88

*Indici di emissione* per categoria. Spostamenti medi *giornalieri stimati* in *5 città italiane* di *medie dimensioni*.

Valori stimati su base *campionaria* (molto dettagliata, numerose categorie e motivi considerati).

Motivo	Categoria	Indice di emissione
Casa-acquisto beni non durevoli	Famiglia	0,25
Casa-acquisto beni durevoli	Famiglia	0,11
Casa-servizi personali	Famiglia	0,16
Casa-svago	Famiglia	0,27
Casa-accompagnamento persone	Famiglia	0,11
Casa-altro	Famiglia	0,13

## Modello di generazione: altro approccio

Invece di stimare direttamente l'indice di generazione per categoria stimo un modello in cui l'indice dipende da alcuni attributi.

modello di *regressione* per *categoria*

$$m_n^i(osh) = \beta_1 x_{n1} + \beta_2 x_{n2} + \beta_3 x_{n3}$$

$= 1$

per esempio:  
reddito di  $n$

$n$ -esimo utente residente in  $o$   
che appartiene alla categoria  $i$   
(per esempio attivo servizi  
privati)

Per esempio: numero di  
auto nel gruppo familiare  
di  $n$

Modello di domanda *descrittivo*, ma *disaggregato* (nello stimarlo considero il singolo utente).

## Modello di *regressione* per *categoria*

$$m_n^i(osh) = \beta_1 x_{n1} + \beta_2 x_{n2} + \beta_{n3} x_{n3} \leftarrow \begin{cases} \beta_1, \beta_2, \beta_3 \text{ coefficienti} \\ \text{del modello da stimare} \end{cases}$$

Intervisto numerosi utenti ( $n=1,2,\dots,T$ ) a cui chiedo il numero di movimenti  $m_n^i$  nel giorno feriale medio, e di cui conosco:

$x_{n2}$  (reddito),  $x_{n3}$  (numero di auto possedute nel gruppo familiare).

Ponendo al solito la variabile *dipendente*  $m_n^i = y_n$  ottengo il solito modello di regressione lineare

$$\left\{ \begin{array}{l} y_1 = \beta_1 x_{11} + \beta_2 x_{12} + \beta_3 x_{13} \\ y_2 = \beta_1 x_{21} + \beta_2 x_{22} + \beta_3 x_{23} \\ \dots \quad \dots \quad \dots \quad \dots \\ y_n = \beta_1 x_{n1} + \beta_2 x_{n2} + \beta_3 x_{n3} \\ \dots \quad \dots \quad \dots \quad \dots \\ y_T = \beta_1 x_{T1} + \beta_2 x_{T2} + \beta_3 x_{T3} \end{array} \right. \begin{array}{l} T \text{ utenti intervistati} \\ \text{Calibrare il modello significa: dati} \\ y_1, \dots, y_T \text{ (spostamenti fatti dagli utenti)} \\ \text{e dati } x_{n1}, x_{n2}, x_{n3} \text{ (attributi dell'utente } n) \\ (\forall n=1, 2, \dots, T), \text{ ricavare i coefficienti} \\ \beta_1, \beta_2 \text{ e } \beta_3 \end{array}$$

## ***Vantaggi* della stima di un modello rispetto ad una stima diretta della domanda (osservazione generale)**

Stimando un modello:  $m_n^i(osh) = \beta_1 x_{n1} + \beta_2 x_{n2} + \beta_3 x_{n3}$

invece di stimare gli indici direttamente  $m_n^i(osh)$  (Indagine campionaria diretta: vedi tabelle sulle 5 città italiane di medie dimensioni) ho i seguenti vantaggi:

- Capisco *da quali attributi*, e *in quale misura*, dipende la domanda di trasporto, (in questo caso la generazione degli spostamenti): reddito, n. auto possedute in famiglia, sesso dell'utente, età, .....

Aspetto molto importante: *comprensione del fenomeno* (considero il punto di vista quantitativo, sempre importante da un punto di vista ingegneristico).

- Posso fare *proiezioni* al *futuro* sulla *domanda* (in questo caso generazione degli spostamenti), sulla base di *previsioni* sugli *attributi*: variazioni del reddito, dell'indice di motorizzazione, dell'età della popolazione.....

# MODELLO DI DISTRIBUZIONE

Modello *gravitazionale*  
(*descrittivo*)

$$D_{od} = k P_o A_d^\beta t_{od}^{-\theta}$$

$P_o$  : Potenziale generativo - Es: popolazione attiva (che lavora) residente in  $o$

$A_d$  : Potenziale attrattivo - Es: addetti della zona  $d$  (persone che lavorano in  $d$ )

$t_{od}$  : Tempo di spostamento da  $o$  a  $d$  – Nella pratica, spesso, semplicemente: distanza in linea d'aria da  $o$  a  $d$  (esprime la “*resistenza*” allo spostamento)

$$D_{od} = k P_o A_d^\beta t_{od}^{-\theta} \longrightarrow \frac{D_{od}}{P_o} = k A_d^\beta t_{od}^{-\theta}$$

Domanda espressa in termini pro capite  
(spostamenti per residente attivo della zona  $o$ )

Per calibrare il modello (determinare i coefficienti incogniti,  $k$ ,  $\beta$  e  $\theta$  )  
linearizzo il modello considerando i logaritmi:

$$\ln \frac{D_{od}}{P_o} = \ln k + \beta \ln A_d - \theta \ln t_{od}$$

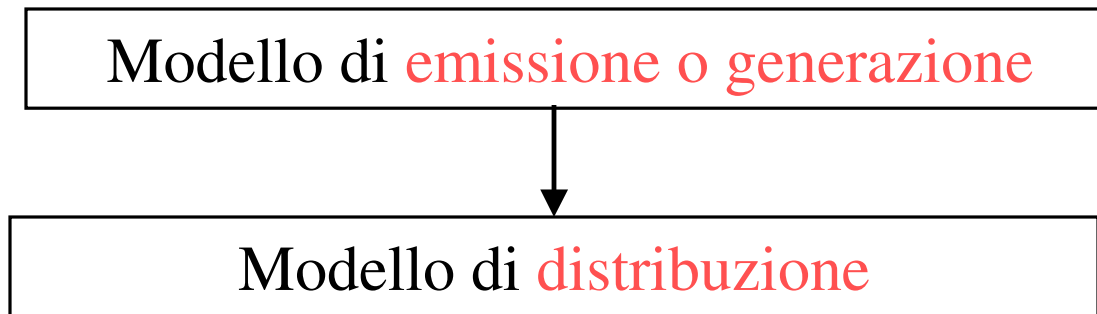
$$y_c = \beta_1 x_{c1} + \beta_2 x_{c2} + \beta_3 x_{c3}$$

↓  
= 1

Mi riduco al solito *modello di regressione lineare* a cui si riducono generalmente i modelli di domanda di tipo *descrittivo*. Utilizzerò il *metodo dei minimi quadrati* per stimare i coefficienti (parametri) incogniti (parte del corso sul *calcolo della domanda*).



Però in realtà  $D_{od} = k P_o A_d^\alpha T_{od}^\gamma$  è un modello di generazione e distribuzione degli spostamento: ossia “copre” *entrambe i primi due stadi del modello a 4 stadi*:



Voglio un modello di sola distribuzione, da applicare a valle di un modello di generazione:



Modello *gravitazionale semplicemente vincolato*

# Modello *gravitazionale semplicemente vincolato*

Dal modello di generazione ho il numero di utenti che escono dalla zona (centroide  $o$ ): lo chiamo:  $O_o$

Impongo il vincolo:

$$\sum_{\forall j} D_{oj} = O_o \quad \longrightarrow \quad \sum_{\forall j} k P_o A_j^\beta t_{oj}^{-\theta} = O_o \quad \longrightarrow \quad k P_o = \frac{O_o}{\sum_{\forall j} A_j^\beta t_{oj}^{-\theta}}$$

Da cui segue:  $D_{od} = k P_o A_d^\beta t_{od}^{-\theta} \quad \longrightarrow \quad D_{od} = O_o \frac{A_d^\beta t_{od}^{-\theta}}{\sum_{\forall j} A_j^\beta t_{oj}^{-\theta}}$

Formulazione *alternativa* del *modello di distribuzione semplicemente vincolato*, basata sulla funzione esponenziale, invece che su quella di potenza (  $t_{od}^{-\theta}$  ):

$$D_{od} = O_o \frac{A_d^\beta e^{-\theta t_{od}}}{\sum_{\forall j} A_j^\beta e^{-\theta t_{oj}}}$$

# Approccio comportamentale al modello di distribuzione

*Insieme di scelta*: l'insieme delle *destinazioni* nella rete

*Modello logit* applicato alla scelta della destinazione  $\longrightarrow p(d / osh) = \frac{e^{\alpha V_d}}{\sum_{\forall j} e^{\alpha V_j}}$

$p(d / osh)$  Probabilità di scegliere la destinazione  $d$ , dato che l'utente esce dalla zona  $o$ , nell'intervallo  $h$ , per il motivo  $s$ .

Scelgo una componente sistematica dell'utilità  $V_d$  avente questa forma:  $V_d = \beta'_1 \ln A_d + \beta'_2 \ln t_{od}$

$$\alpha V_d = \alpha \beta'_1 \ln A_d + \alpha \beta'_2 \ln t_{od} = \beta_1 \ln A_d + \beta_2 \ln t_{od}$$

( $\beta_1 = \alpha \beta'_1$ ;  $\beta_2 = \alpha \beta'_2$ ; ossia la stima di  $\alpha$  viene conglobata in quella dei parametri di  $V_d$ : questo si verifica sempre se la componente sistematica dell'utilità ha una espressione lineare)

$$\begin{aligned}
 p(d / osh) &= \frac{e^{\alpha V_d}}{\sum_{\forall j} e^{\alpha V_j}} = \frac{e^{\beta_1 \ln A_d + \beta_2 \ln t_{od}}}{\sum_{\forall j} e^{\beta_1 \ln A_j + \beta_2 \ln t_{oj}}} = \frac{e^{\beta_1 \ln A_d} e^{\beta_2 \ln t_{od}}}{\sum_{\forall j} e^{\beta_1 \ln A_j} e^{\beta_2 \ln t_{oj}}} = \\
 &= \frac{e^{\ln A_d^{\beta_1}} e^{\ln t_{od}^{\beta_2}}}{\sum_{\forall j} e^{\ln A_j^{\beta_1}} e^{\ln t_{oj}^{\beta_2}}} = \frac{A_d^{\beta_1} t_{od}^{\beta_2}}{\sum_{\forall j} A_j^{\beta_1} t_{oj}^{\beta_2}}
 \end{aligned}$$

Probabilità di scegliere la destinazione  $d$

Totalità di utenti uscenti dalla zona  $o$

La domanda media sarà data da:  $D_{od} = O_o p(d / osh)$

$$D_{od} = O_o p(d / osh) = O_o \frac{A_d^{\beta_1} t_{od}^{\beta_2}}{\sum_{\forall j} A_j^{\beta_1} t_{oj}^{\beta_2}}$$

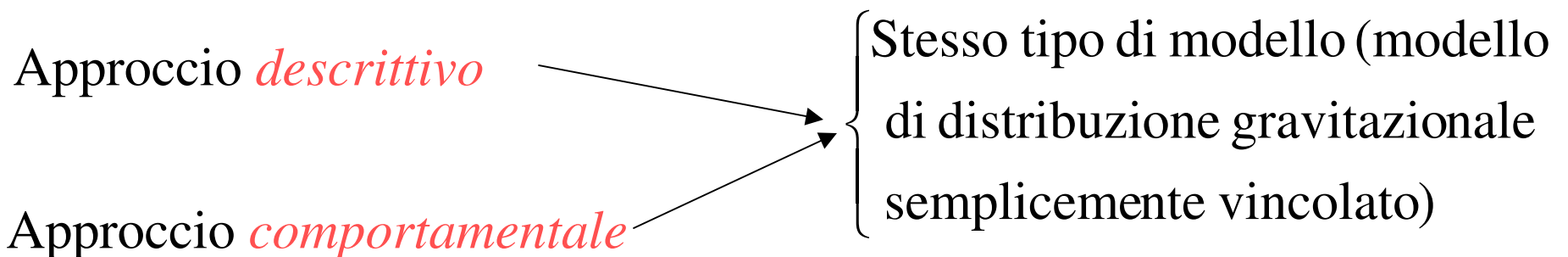
E' *uguale* al modello *gravitazionale semplicemente vincolato* ottenuto con un *approccio descrittivo* ( $\beta_1 = \beta$ ;  $\beta_2 = -\theta$ )

$$\left\{ D_{od} = O_o \frac{A_d^{\beta} t_{od}^{-\theta}}{\sum_{\forall j} A_j^{\beta} t_{oj}^{-\theta}} \right.$$

Per ottenere con un approccio comportamentale il modello gravitazionale semplicemente vincolato con la *funzione esponenziale*:

$$D_{od} = O_o \frac{A_d^\beta e^{-\theta t_{od}}}{\sum_{\forall j} A_j^\beta e^{-\theta t_{oj}}}$$

Basta che scelga una componente sistematica dell'utilità  $V_d$  di questo tipo:  $V_d = \beta'_1 \ln A_d + \beta'_2 t_{od}$



Conferma della *validità* di *entrambi gli approcci*.

Valori stimati per i parametri:  $\beta_1$  e  $\beta_2$

Modello



$$\frac{A_d^{\beta_1} t_{od}^{\beta_2}}{\sum_{\forall j} A_j^{\beta_1} t_{oj}^{\beta_2}}$$

Motivo	$A_d$	$\beta_1$	$\beta_2$
Casa-lavoro (C-L)	Addetti settore industria	1,10	- 0,70
	Addetti settore servizi	0,93	- 0,70
	Addetti settore servizi privati	0,93	- 0,83
Casa-scuola (C-S)	Addetti settore servizi pubblici	0,93	- 0,58
	Alunni scuole elementari	0,90	- 2,52
	Studenti scuole medie inferiori	0,95	- 2,24
Casa-servizi personali (C- SP)	Studenti scuole medie superiori	1,00	- 0,35
	Addetti servizi	0,91	- 0,78
Casa- accompagnamento persone (C- AP)	Stud. scuole elem. e medie inferiori	0,20	- 1,35
Casa-Acquisto Beni (C-AB)	Addetti commercio	1,61	- 2,54

## Modello di scelta **modale**

Approccio comportamentale: i primi modelli di tipo comportamentale utilizzati nel campo dei trasporti sono stati applicati alla scelta modale

Insieme di scelta  $I$   
(caso *semplice*)

{ trasporto pubblico  
trasporto privato

Insieme di scelta  $I$   
(caso *più dettagliato*)

{ auto come conducente  
auto come passeggero  
autobus  
motociclo  
bicycletta  
a piedi

Modello *logit multinomiale*: ho più di due alternative di scelta

$$p_i = \frac{e^{\alpha V_i}}{\sum_{\forall j} e^{\alpha V_j}}$$

$$V_i = \beta'_1 x_{i1} + \beta'_2 x_{i2} + \dots + \beta'_k x_{ik}$$



pongo:  $\beta_1 = \alpha\beta'_1, \beta_2 = \alpha\beta'_2, \dots, \beta_k = \alpha\beta'_k$

$$\alpha V_i = \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik}$$

Nel caso di *forma lineare* per *l'utilità sistematica*,  $V_d$ , la stima del parametro  $\alpha$  della logit è conglobata nella stima dei parametri dell'utilità sistematica

$$p_i = \frac{e^{\beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik}}}{\sum_{\forall j} e^{\beta_1 x_{j1} + \beta_2 x_{j2} + \dots + \beta_k x_{jk}}}$$



Modello *logit binomiale*: ho due alternative di scelta

$$P_{priv} = \frac{e^{\beta_1 x_{priv1} + \beta_2 x_{priv2} + \dots + \beta_k x_{privk}}}{e^{\beta_1 x_{priv1} + \beta_2 x_{priv2} + \dots + \beta_k x_{privk}} + e^{\beta_1 x_{pubb1} + \beta_2 x_{pubb2} + \dots + \beta_k x_{pubbk}}}$$

	$\beta_1$	$\beta_2$	$\beta_3$	$\beta_4$	$\beta_5$	$\beta_6$
$V_{priv} =$ Utilità trasp. privato	$X_{priv,1} =$ Tempo a bordo (min.)	$X_{priv,2} =$ Tempo non a bordo (min.)	$X_{priv,3} =$ Costo viaggio (carb +ped)	$X_{priv,4} =$ Reddito	$X_{priv,5} =$ Numero di auto in famiglia	$X_{priv,6} = 1$ (attributo di pref. modale)
$V_{pubb} =$ Utilità trasp. pubblico	$X_{pubb,1} =$ Tempo a bordo (min.)	$X_{pubb,2} =$ Tempo non a bordo (min.)	$X_{pubb,3} =$ Tariffa del trasp. pubb.	$X_{pubb,4} = 0$	$X_{pubb,5} = 0$	$X_{pubb,6} = 0$

Area metropolitana di *Washington*; modello *logit binomiale*

	$\beta_1$	$\beta_2$	$\beta_3$	$\beta_4$	$\beta_5$	$\beta_6$	$\beta_7$
$V_{priv} =$ Utilità trasp. privato	$X_{priv,1} =$ 1	$X_{priv,2} =$ Tempo a bordo (min.)	$X_{priv,3} =$ Tempo non a bordo (min.)	$X_{priv,4} =$ Costo del viaggio (cents)(o ut-of- pocket)	$X_{priv,5} =$ 0	$X_{priv,6} =$ Numero di auto in famiglia	$X_{priv,7} =$ 1 posto di lavoro in centro; 0 posto di lavoro no centro
$V_{pubb} =$ Utilità trasp. pubblico (autobus)	$X_{pubb,1} =$ 0	$X_{pubb,2} =$ Tempo a Bordo (min.)	$X_{pubb,3} =$ Tempo non a bordo (min.)	$X_{pubb,4} =$ 0	$X_{pubb,5} =$ Tariffa del trasp. Pub. (cents)	$X_{pubb,6} =$ 0	$X_{pubb,7} =$ 0
Stima del coefficie nte	1,454	-0,00897	-0,0308	-0,0115	-0,00708	0,770	-0,561

## Componenti *sistematiche dell'utilità* del modello

$$V_{privato} = -0,00897 t_{priv bordo} - 0,0308 t_{priv no bordo} - 0,0115 c_{priv} + 0,770 na \\ - 0,561 centro + 1,454$$

$$V_{pubblico} = -0,00897 t_{pubb bordo} - 0,0308 t_{pubb no bordo} - 0,00708 c_{pubb}$$

*Logit*  
*Multinomiale*  
 per la *scelta*  
 del *modo* di  
 trasporto -  
 Città di  
*Parma*

<b>PIEDI</b>		
TWL	Tempo [h]	-6.8237
<b>BICI</b>		
TB	Tempo [h]	-8.2718
NB/NAD	Nr. bici possedute in famiglia per adulto	0.6646
BIC	Var. specifica dell'alternativa	-1.5818
<b>CICLOMOTORE E MOTO</b>		
TC	Tempo [h]	-8.2718
GC	Var. ombra (1 ≥ anni 14; 0 altrimenti)	0.6863
NC/NAD	Nr. cicl. posseduti in famiglia per adulto	1.8572
CICLO	Var. specifica dell'alternativa	-2.3789
<b>AUTO</b>		
TA	Tempo [h]	-1.6142
CA	Costo monetario [£.x10 <sup>3</sup> ]	-0.1669
CENTRO	Var. ombra (1 se dest. centro 0 altrim.)	-1.1469
CFAM	Var. ombra (1 se capofamiglia 0 altrim.)	0.4931
NA	Nr. auto in famiglia	6.4014
AUTO	Var. specifica dell'alternativa	-1.7103
<b>BUS</b>		
TB	Tempo totale [h] piedi + attesa + bordo	-1.6142
CB	Costo monetario [£.x10 <sup>3</sup> ]	-0.1669
NINT	Numeri cambi linea	-0.1772
BUS	Var. specifica dell'alternativa	-1.7827
<i>ln L</i> (βott.)		-475
<i>ln L</i> (O)		-697
ro <sup>2</sup>		0.317
% right		0.651

	$\beta_1$	$\beta_2$	$\beta_3$	$\beta_4$	$\beta_5$	$\beta_6$	$\beta_7$	$\beta_8$	$\beta_9$	$\beta_{10}$	$\beta_{11}$	$\beta_{12}$	$\beta_{13}$	$\beta_{14}$	$\beta_{15}$
piedi	<i>TW</i>	0	0	0	0	0	0	0	0	0	0	0	0	0	0
bici	0	<i>TB</i>	0	0	<i>NB</i>	0	0	0	0	0	0	<i>BIC</i>	0	0	0
ciclo moto	0	<i>TC</i>	0	0	0	<i>GC</i>	<i>NC</i>	0	0	0	0	0	<i>CIC</i>	0	0
auto	0	0	<i>TA</i>	<i>CA</i>	0	0	0	<i>CEN</i>	<i>CFA</i>	<i>NA</i>	0	0	0	<i>AU</i>	0
bus	0	0	<i>TB</i>	<i>CB</i>	0	0	0	0	0	0	<i>NINT</i>	0	0	0	<i>BUS</i>
coeff icien te	-6,8237	-8,2718	-1,6142	-0,1669	0,6646	0,6863	1,8572	-1,1469	0,4931	6,4014	-0,1772	-1,5818	-2,3789	-1,7103	-1,7827

## Componenti *sistematiche dell'utilità* del modello

$$V_{\text{piedi}} = -6,8237 \text{ TWL}$$

$$V_{\text{bici}} = -8,2718 \text{ TB} + 0,6646 \text{ NB} - 1,5818$$

$$V_{\text{ciclomotore}} = -8,2718 \text{ TC} + 0,6863 \text{ GC} + 1,8572 \text{ NC} - 2,3789$$

$$V_{\text{auto}} = -1,6142 \text{ TA} - 0,1669 \text{ CA} - 1,1469 \text{ CENTRO} + 0,4931 \text{ CFAM} + \\ + 6,4014 \text{ NA} - 1,7103$$

$$V_{\text{bus}} = -1,6142 \text{ TB} - 0,1669 \text{ CB} - 0,1772 \text{ NINT} - 1,7827$$

# Modello logit multinomiale per la scelta del modo di trasporto - Parma

Insieme di scelta {  
piedi  
bicicletta  
ciclomotore e moto  
automobile  
autobus

Cosa si può ricavare dal valore dei coefficienti?

$$\frac{\beta_{tempo}}{\beta_{costo}} = \frac{1,6142 \text{ utilità / ora}}{0,1669 \text{ utilità / } 1000\text{£(1985)}} = 9672 \frac{\text{£(1985)}}{\text{ora}}$$

Valore monetario  
del *tempo*

$$\frac{\beta_{\text{cambio di linea}}}{\beta_{\text{costo}}} = \frac{0,1772 \text{ utilità/cambio di linea}}{1,6142 \text{ utilità/ora}} = 0,11 \frac{\text{ore}}{\text{cambio di linea}}$$

Un cambio di linea è equivalente ad una disutilità di 0,11 ore di viaggio in più.

A proposito delle variabili *socioeconomiche*:

- Alto valore del *coefficiente* n. di auto in famiglia: 6,4014.  
(Osservazione “in linea” con altri modelli simili)

*L'aumento dell'indice di motorizzazione* favorisce il modo di trasporto *autovettura*. L'osservazione può sembrare banale: bisogna tenere conto che il *modello fornisce* un valore *quantitativo* che può essere confrontato, quantitativamente, con il “peso” di altri attributi.



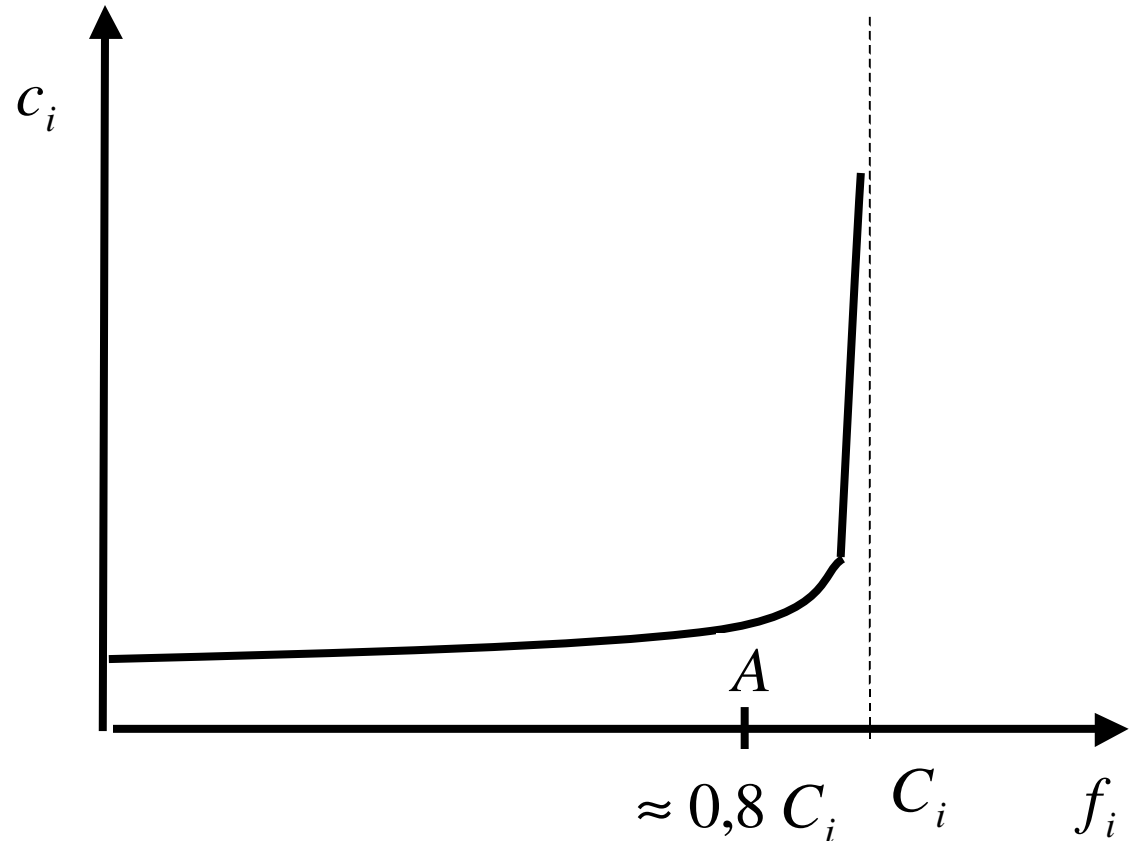
# Modello di scelta dell'itinerario

Andamento del costo (espresso dal tempo) nel caso di un *arco* di una *rete urbana*: trasporto *individuale stradale*

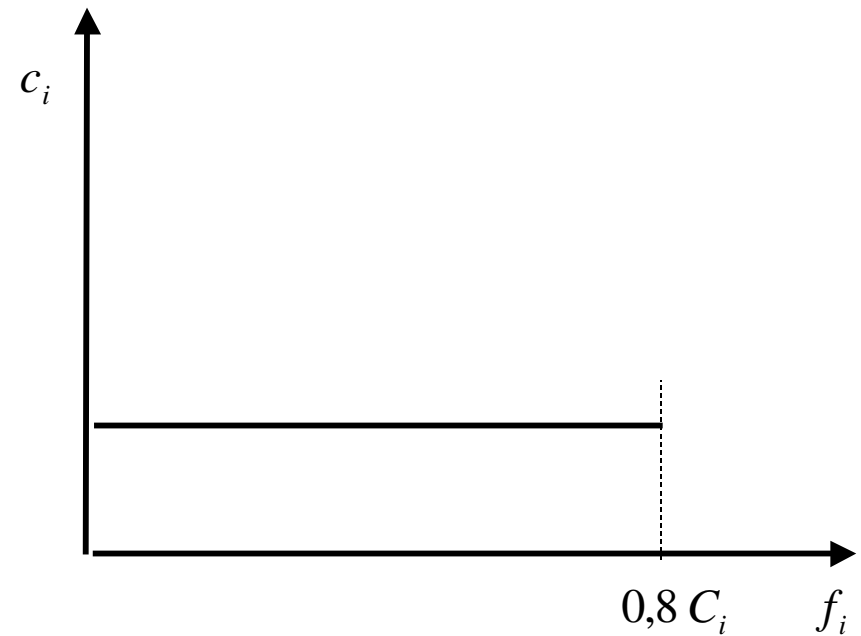
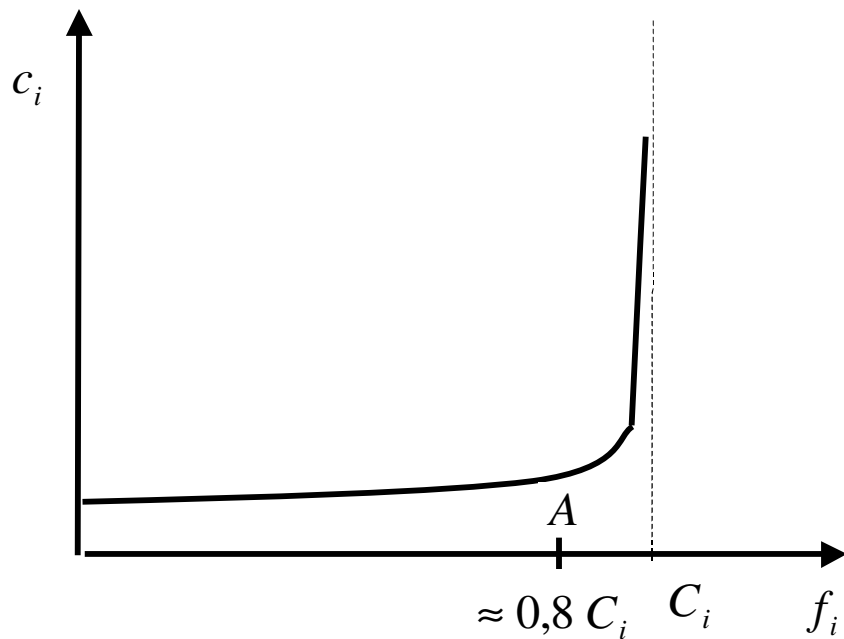
$c_i$  : costo (tempo)  
sull'arco i-esimo

$C_i$  : capacità dell'arco i-  
esimo

$f_i$  : flusso sull'arco  
i-esimo



Fino al punto A (circa 80% capacità arco) posso considerare, approssimativamente, la funzione costante:  $c_i(f_i) = \text{cost}$



*Arco* di una rete di *trasporto individuale stradale, non molto carico*, posso *approssimare* la funzione di costo con una costante: considero il *costo indipendente dal flusso*.

Il *costo* di un arco di una rete di trasporto si considera *indipendente* dal *flusso*

- arco di una rete di trasporto ferroviario, marittima , aerea  
(generalmente)

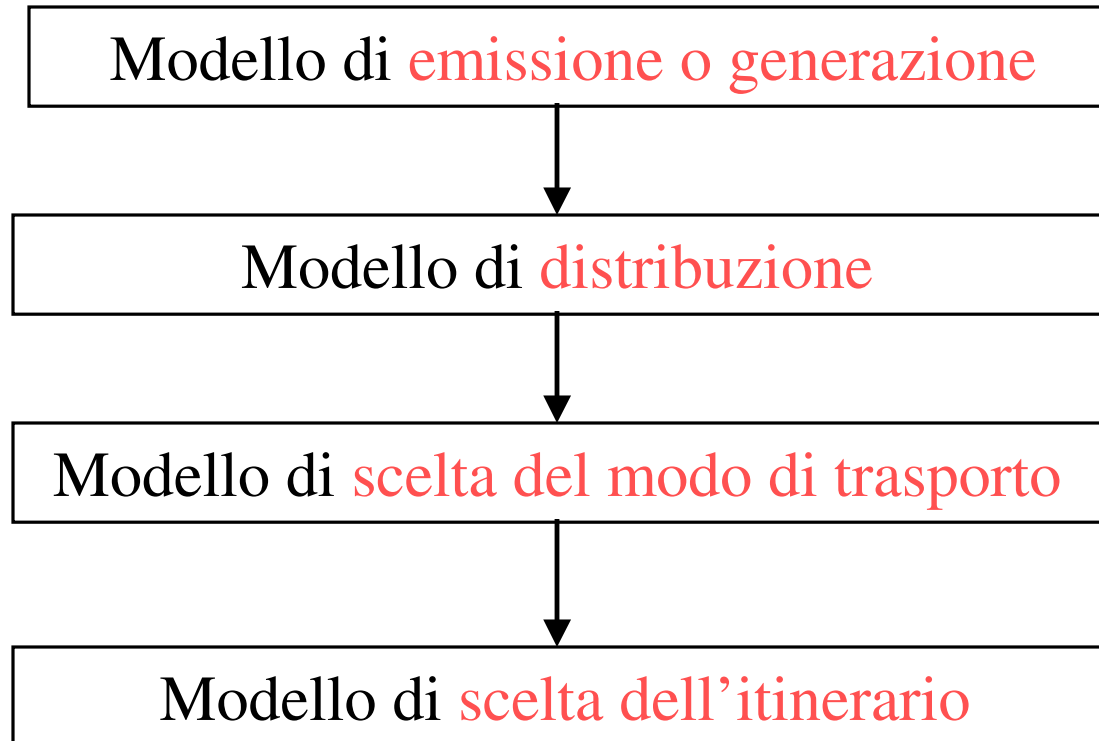
- arco di una rete di trasporto  
collettivo

- su corsia riservata
- anche su corsia non riservata

Pongo il costo uguale ad una costante data da:  $\text{cost} = \frac{L}{V_{\text{comm}}}$

La  $V_{\text{comm}}$  tiene conto della fase di accelerazione, di frenatura e dei tempi di fermata. Attraverso una opportuna scelta della velocità di regime, si tiene conto anche del condizionamento degli autoveicoli in caso di corsia non riservata (questo argomento sarà approfondito quando si parlerà delle tipologie dei *diagrammi di trazione*).

# Modello di scelta dell'itinerario

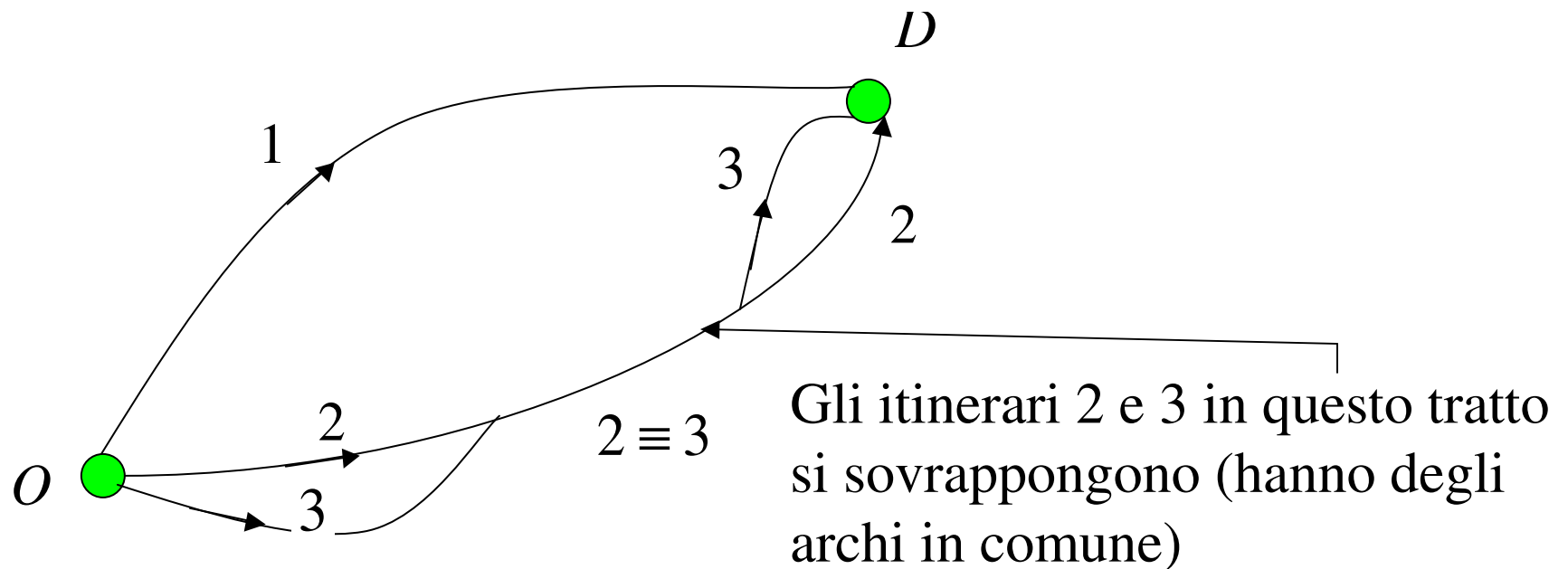


*Ultimo stadio* del modello a quattro stadi

Modello di scelta dell'itinerario: *modello tutto o niente* o *di utilità deterministica*

E' un *modello comportamentale*.

Insieme di scelta: totalità degli itinerari che congiungono una coppia O-D (itinerario: successioni di archi)



Nei modelli di *utilità aleatoria*:  $U_i = V_i + \varepsilon_i$

Nel modello di *utilità deterministica*  $\rightarrow U_i = V_i$  (non c'è lo scarto aleatorio)

Nel modello di utilità deterministico applicato alla *scelta dell'itinerario (modello tutto o niente)*:  $V_i = -C_i$  ( $C_i$  costo sull'itinerario che costituisce l'alternativa i-esima) e perciò  $U_i = -C_i$

L'utente sceglie *l'alternativa di massima utilità*. Ossia l'*itinerario di minimo costo*.

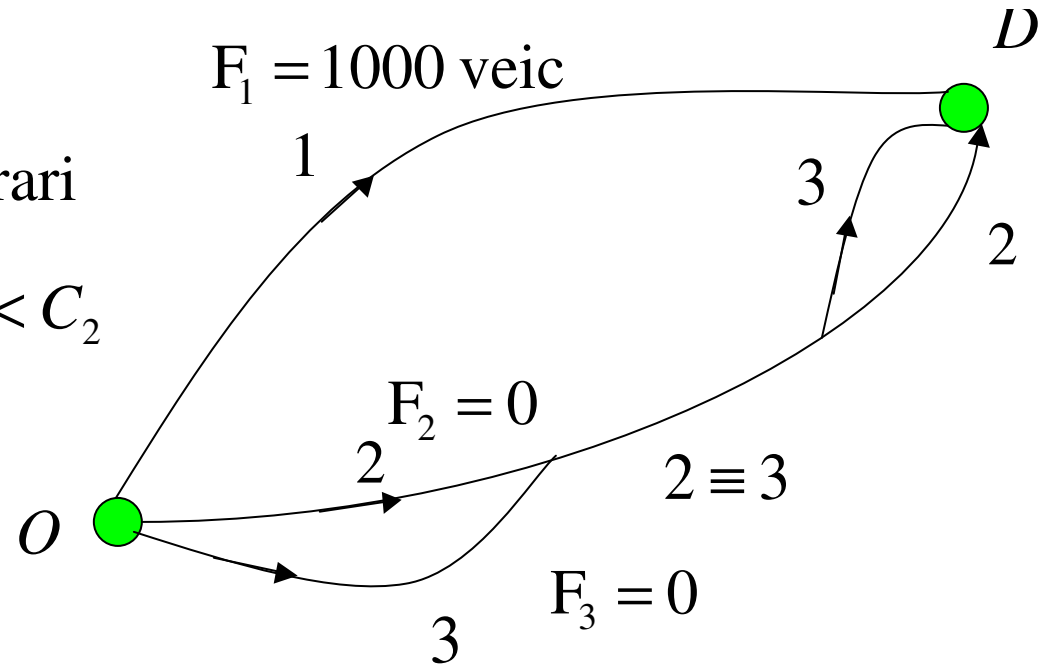
*Assegnazione tutto o niente*: perché tutta la domanda viene assegnata all'itinerario di minimo costo e niente (nessun flusso) agli itinerari alternativi.

$d_{OD} = 1000$  veic

$F_1, F_2, F_3$  : flusso sugli itinerari

Supponiamo che  $C_1 < C_3 < C_2$

tutto il flusso fra O e D  
utilizza l'itinerario 1

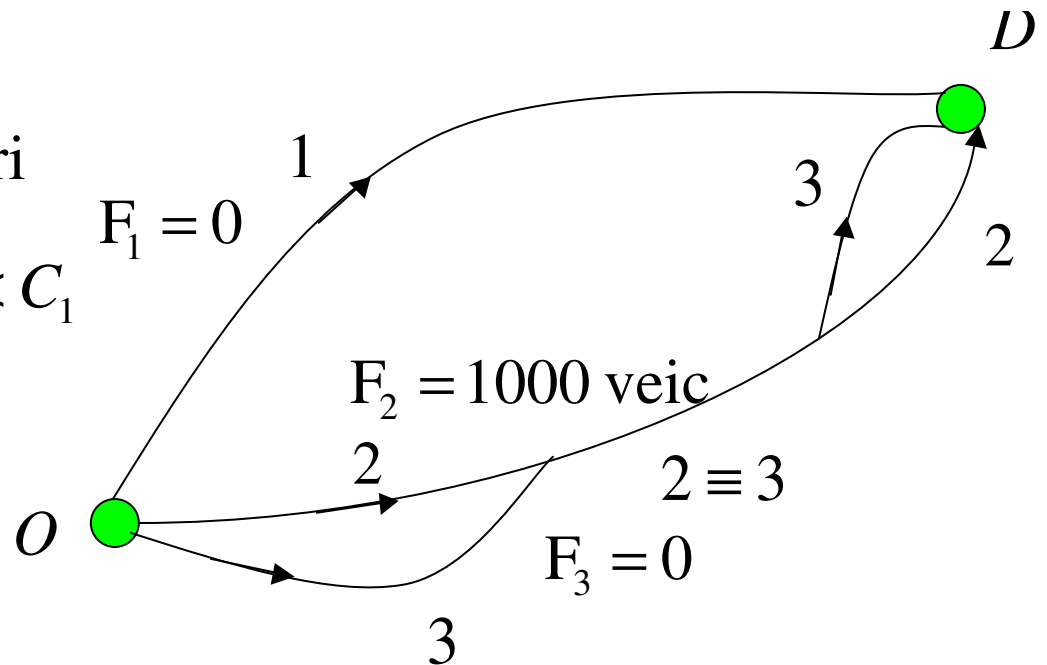


$d_{OD} = 1000$  veic

$F_1, F_2, F_3$  : flusso sugli itinerari

Supponiamo che  $C_2 < C_3 < C_1$

tutto il flusso fra O e D  
utilizza l'itinerario 2



# Algoritmo tutto o niente

Calcolo gli *itinerari* di *minimo costo* fra *ciascuna* coppia O-D ed *assegno tutta la domanda* fra la coppia O-D all'itinerario di minimo costo.

*Opero in questo modo*



Considero ciascun nodo origine (centroide), calcolo l'*albero* degli *itinerari* di *minimo costo* avente come *nodo radice* il *nodo origine*; *assegno la domanda* fra l'origine e tutte le destinazioni nella rete; *ripeto* l'*operazione* per tutti i nodi origine (centroidi) della rete.

Ricordiamo che (elementi di teoria dei grafi):

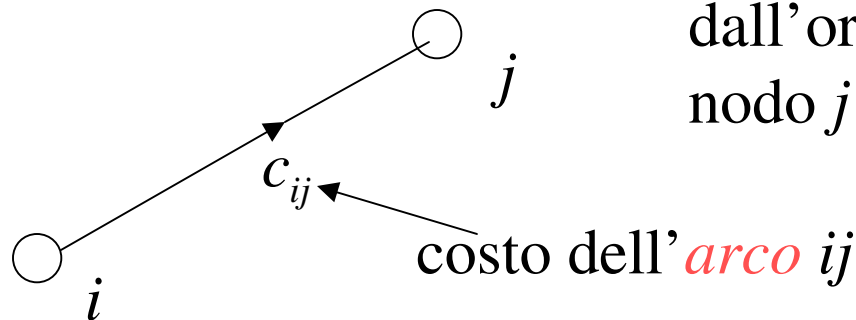
*Un albero di radice  $r$ ,  $T(r)$* : è un *grafo parziale* (ossia ricavato da quello dato eliminando alcuni archi), privo di circuiti, in cui esiste un *unico itinerario* che congiunge la *radice  $r$*  a *ciascun nodo* del grafo.



Base teorica per il calcolo dell'albero degli itinerari di minimo costo è il

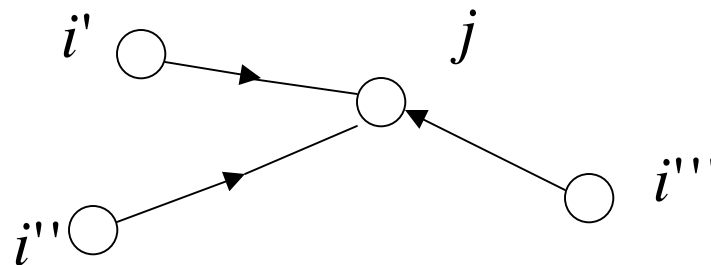
**Teorema di Bellmann:** Condizione necessaria e sufficiente affinché un albero  $T(r)$  sia l'albero degli itinerari di minimo costo è che per ogni coppia di nodi  $ij$ , *collegati da un arco* nel grafo, si verifichi la condizione:  $c_j \leq c_i + c_{ij}$

$c_i$  costo dell'*itinerario* dall'origine (radice) al nodo  $i$



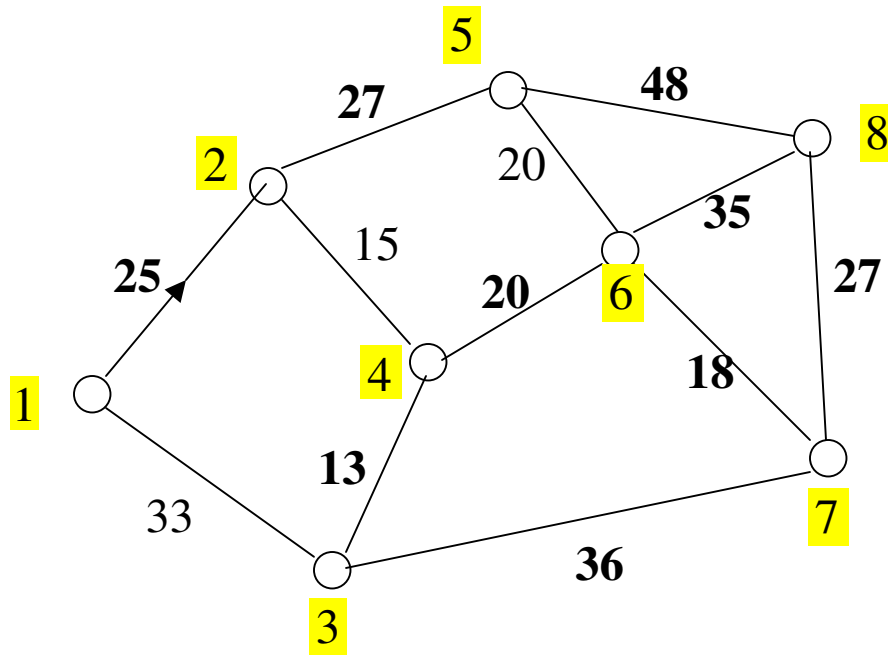
$c_j$  costo dell'*itinerario* dall'origine (radice) al nodo  $j$

Con il segno di uguaglianza valido per almeno uno degli archi che convergono in  $j$ :



# Spiegazione dell'algoritmo di calcolo dell'albero degli itinerari di minimo costo attraverso un esempio

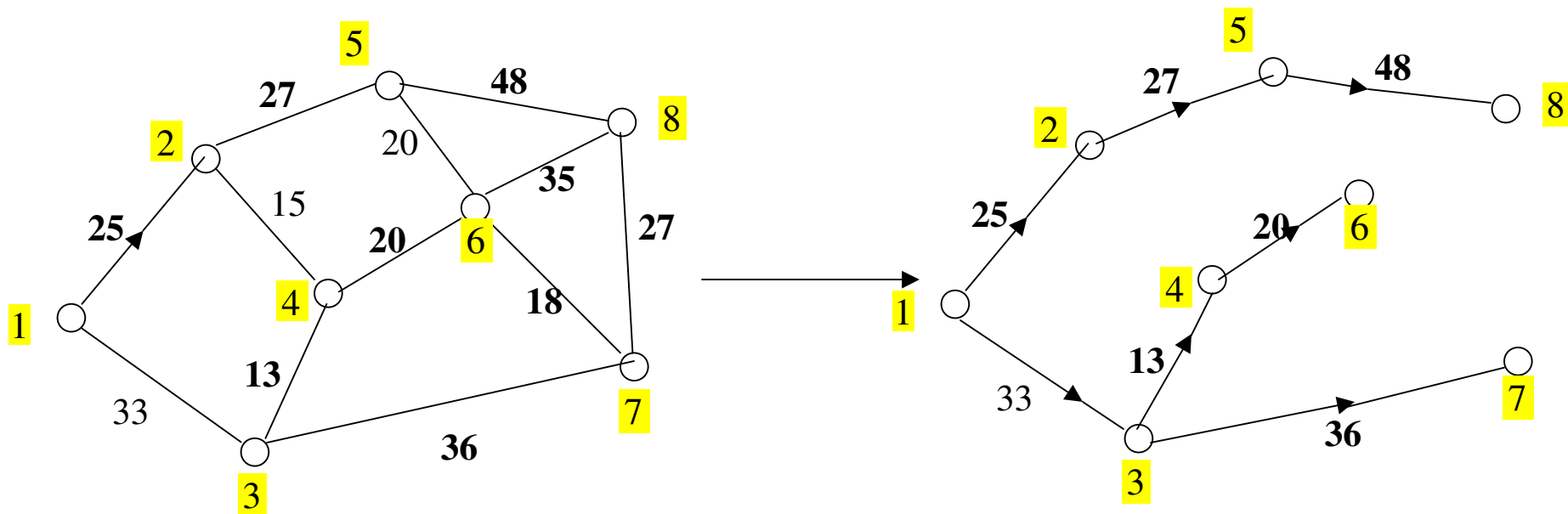
Considero un grafo:



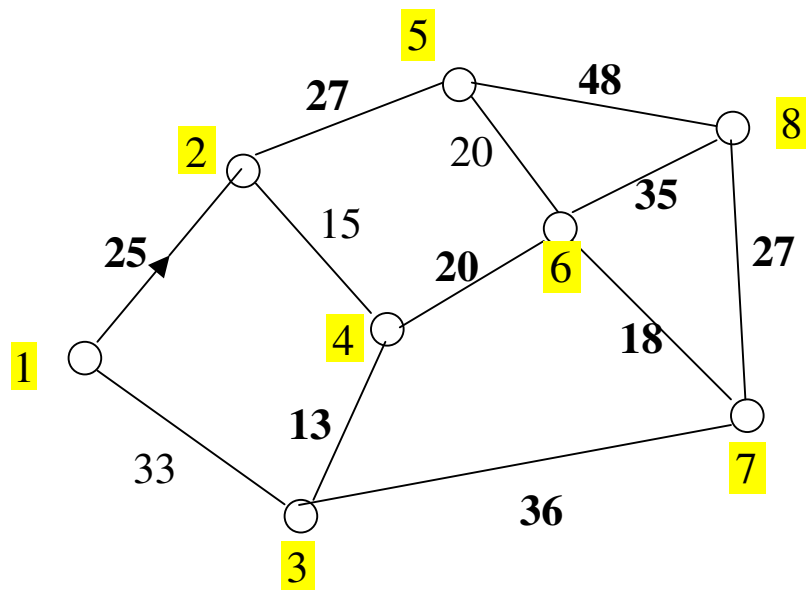
Voglio trovare l'albero degli itinerari di minimo costo avente per radice il nodo 1

L'algoritmo per il calcolo dell'albero degli itinerari di minimo costo opera per *passi successivi*: ad *ogni passo* ho una *albero* che viene via via "migliorato" fino ad arrivare a quello degli itinerari di minimo costo.

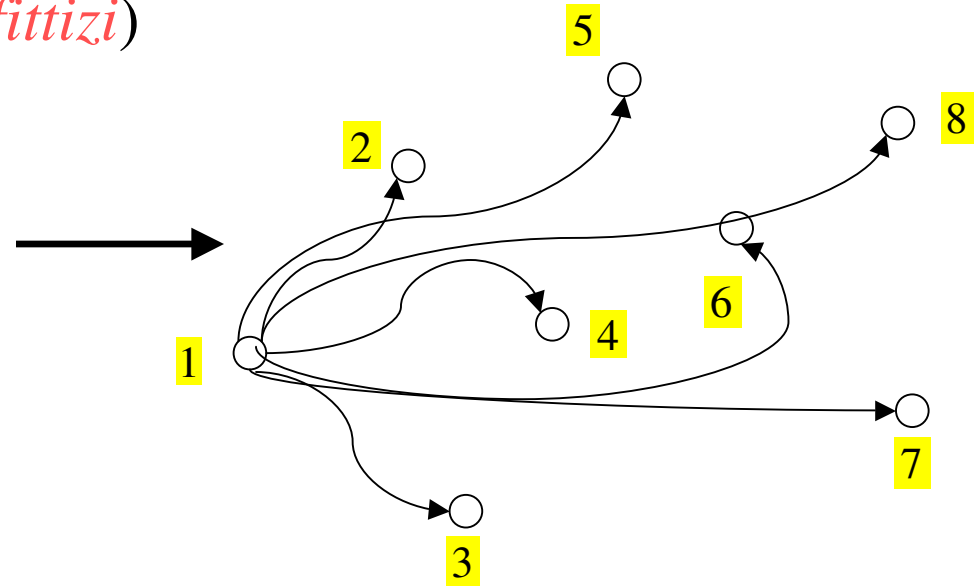
Ho un *problema*: devo avere un albero con cui partire



Nel caso del grafo precedente risolvo il problema facilmente. Però mi serve un metodo che mi permetta di *individuare un albero di partenza*, facilmente, anche nel caso di grafi di *grosse dimensioni* (centinaia di nodi e di archi).



Albero di *partenza* formato da archi *fittizi*)

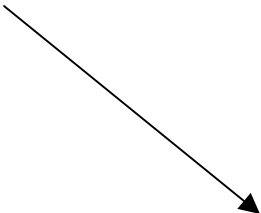


Ipotizzo l'esistenza di archi fittizi (non reali) che collegano direttamente il nodo radice (1) a tutti i nodi della rete.

Assegno a questi *archi fittizi* un *costo molto elevato*, maggiore del costo dei possibili *itinerari* (non archi!) nella rete: nel corso del procedimento di calcolo dell'albero degli itinerari di minimo costo gli *archi fittizi verranno sicuramente eliminati* in quanto, dato il loro costo, non possono fare parte di itinerari di minimo costo.

Esistono più algoritmi per la determinazione dell'albero degli itinerari di minimo costo: vediamo *l'algoritmo di Dijkstra*. Si organizza il calcolo attraverso una matrice a 4 colonne. Ciascuna riga corrisponde ad un nodo del grafo.

La prima colonna non fa parte della matrice di calcolo è stata inserita solo a scopo illustrativo del procedimento



	costo itinerario	codice	nodo che precede	vettore serbatoio
1	0	2	0	1
2	1000	0	1	0
3	1000	0	1	0
4	1000	0	1	0
5	1000	0	1	0
6	1000	0	1	0
7	1000	0	1	0
8	1000	0	1	45 0

1	0	2	0	1
2	1000	0	1	0
3	1000	0	1	0
4	1000	0	1	0
5	1000	0	1	0
6	1000	0	1	0
7	1000	0	1	0
8	1000	0	1	0

*1° colonna: costo di itinerario* per arrivare al dato *nodo* (individuato dal numero di riga) dalla radice.

*2° colonna codice:*

0 se il costo del nodo è quello iniziale.

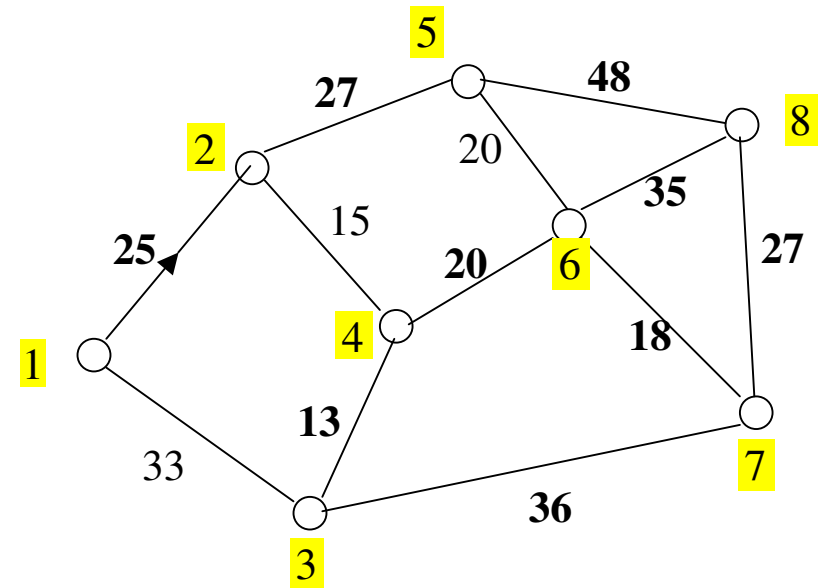
1 se ho estratto il nodo dalla testa del serbatoio.

2 se il nodo è nel serbatoio (4° colonna).

*3° colonna:* nodo che precede il nodo dato nell'itinerario, avente origine nella radice, fino a quel punto del procedimento individuato.

*4° colonna:* vettore serbatoio.

1	0	2	0	1
2	1000	0	1	0
3	1000	0	1	0
4	1000	0	1	0
5	1000	0	1	0
6	1000	0	1	0
7	1000	0	1	0
8	1000	0	1	0

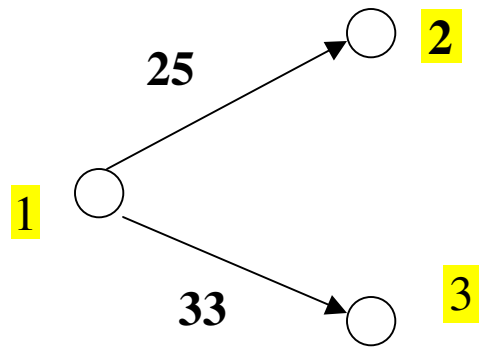


### *Passo 0* del procedimento

Nella prima colonna ho tutti 1000 (eccetto che per il nodo radice): in quanto l'albero iniziale è costituito dagli archi fittizi (tutti i nodi sono preceduti dal nodo 1, terza colonna). Nel serbatoio è contenuto il nodo radice.

*Passo 1* del procedimento

Estraggo 1 dalla “testa” del serbatoio (aggiorno la riga 1) e considero la stella in uscita da esso



Belmann

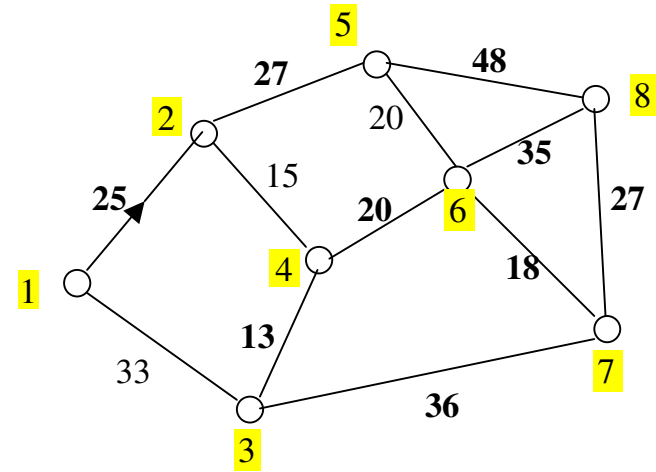
$c_j \leq c_i + c_{ij}$  è verificato?

$1000 \leq 0 + 25$  ?

No!  $\Rightarrow$  2 nel serbatoio

$1000 \leq 0 + 33$  ?

No!  $\Rightarrow$  3 nel serbatoio



Aggiorno i dati della 2° e 3° riga e riordino il serbatoio:

1	0	2	0	1
2	1000	0	1	0
3	1000	0	1	0
4	1000	0	1	0
5	1000	0	1	0
6	1000	0	1	0
7	1000	0	1	0
8	1000	0	1	0

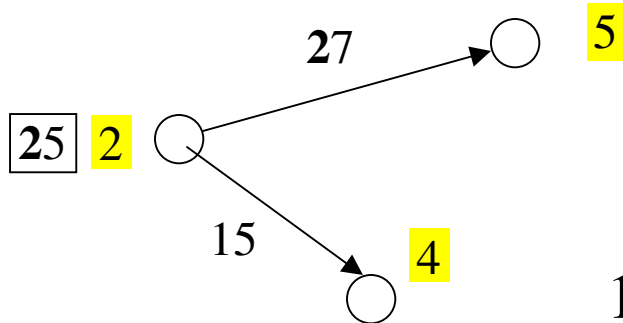


1	0	1	0	2	
2	25	2	1	3	
3	33	2	1	0	
4	1000	0	1	0	
5	1000	0	1	0	
6	1000	0	1	0	
7	1000	0	1	0	
8	1000	0	1	48	0

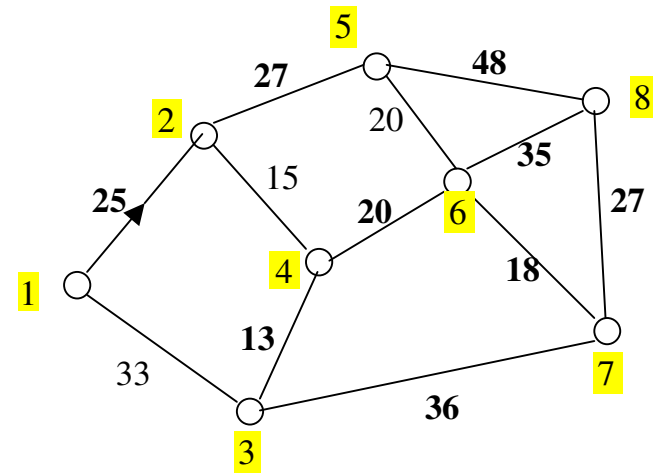


*Passo 2* del procedimento

Estraggo 2 dalla “testa” del serbatoio (aggiorno la riga 2) e considero la stella in uscita da esso



Belmann



$c_j \leq c_i + c_{ij}$  è verificato?

$1000 \leq 25 + 15$  ? No!  $\Rightarrow$  4 nel serbatoio

$1000 \leq 25 + 27$  ? No!  $\Rightarrow$  5 nel serbatoio

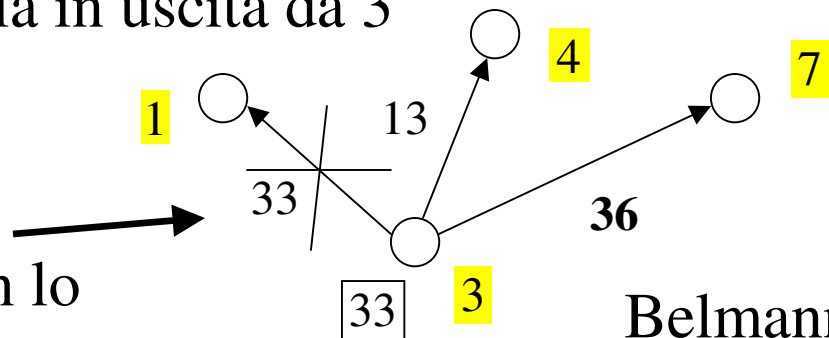
Aggiorno i dati della 4° e 5° riga e riordino il serbatoio

1	0	1	0	2
2	25	2	1	3
3	33	2	1	0
4	1000	0	1	0
5	1000	0	1	0
6	1000	0	1	0
7	1000	0	1	0
8	1000	0	1	0



1	0	1	0	3
2	25	1	1	4
3	33	2	1	5
4	40	2	2	0
5	52	2	2	0
6	1000	0	1	0
7	1000	0	1	0
8	1000	0	1	0

**Passo 3** - Estraggo 3 dalla “testa” del serbatoio (aggiorno la riga 3) e considero la stella in uscita da 3



Non lo considero: il

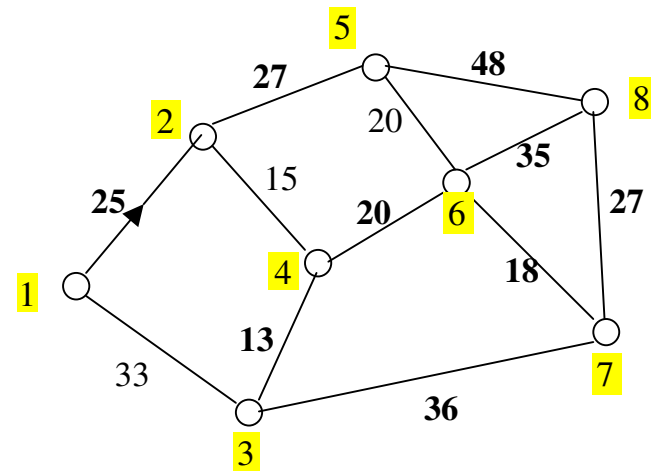
nodo 1 *ha codice*  
*1* l'*itinerario* è  
*definitivo*

Belmann

$$c_j \leq c_i + c_{ij}$$

è verificato?

$40 \leq 33 + 13$  ? si!  $\Rightarrow$  non modifico i dati di 4  
 $1000 \leq 33 + 36$  ? No!  $\Rightarrow$  7 nel serbatoio



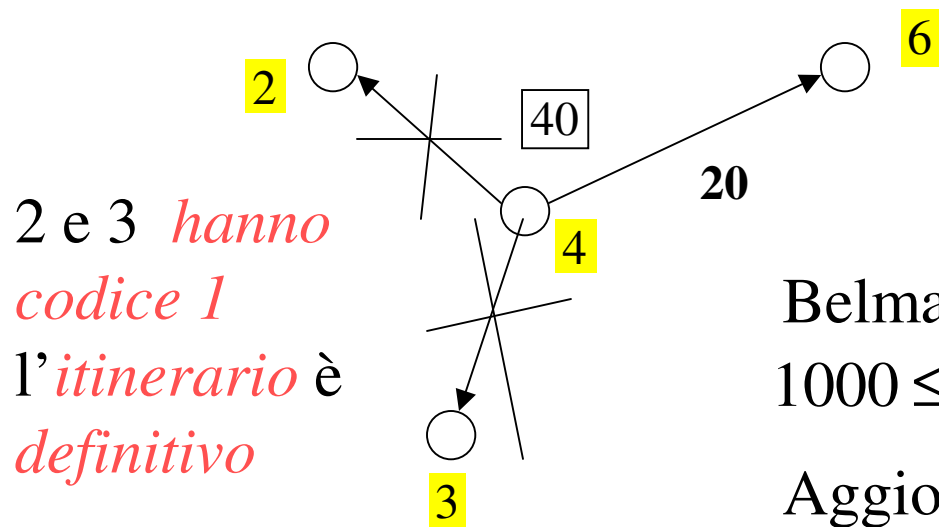
Aggiorno i dati della 7° riga e riordino il serbatoio

1	0	1	0	3
2	25	1	1	4
3	33	2	1	5
4	40	2	2	0
5	52	2	2	0
6	1000	0	1	0
7	1000	0	1	0
8	1000	0	1	0

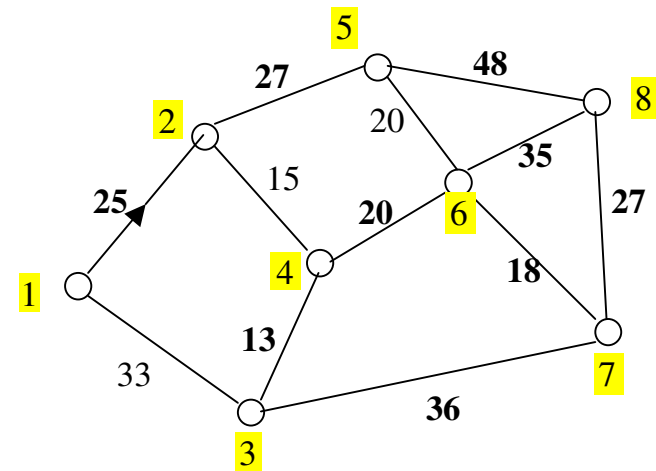


1	0	1	0	4	
2	25	1	1	5	
3	33	1	1	7	
4	40	2	2	0	
5	52	2	2	0	
6	1000	0	1	0	
7	69	2	3	0	
8	1000	0	1	50	0

*Passo 4* - Estraggo 4 dalla “testa” del serbatoio (aggiorno la riga 4) e considero la stella in uscita da 4



2 e 3 *hanno codice 1*  
l'*itinerario* è *definitivo*



Belmann  $c_j \leq c_i + c_{ij}$  è verificato?  
 $1000 \leq 40 + 20$  ? no!  $\Rightarrow$  6 nel serbatoio

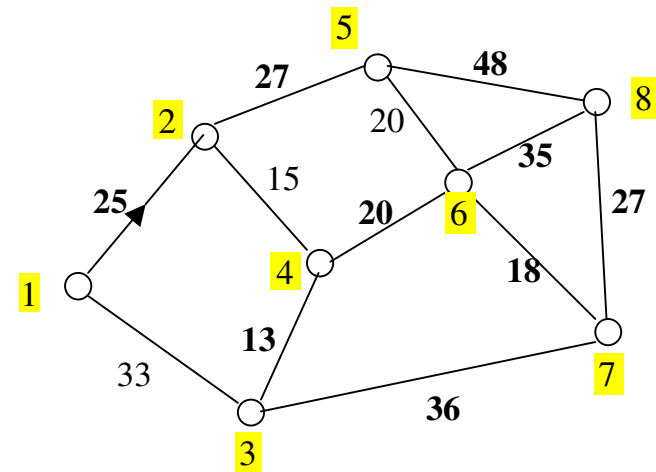
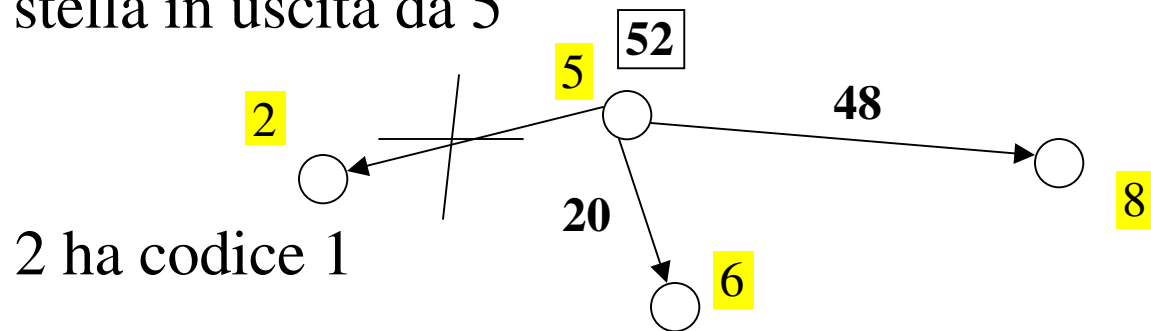
Aggiorno i dati della 6° riga e riordino il serbatoio (in questo caso l'ordine del serbatoio cambia)

1	0	1	0	4
2	25	1	1	5
3	33	1	1	7
4	40	2	2	0
5	52	2	2	0
6	1000	0	1	0
7	69	2	3	0
8	1000	0	1	0



1	0	1	0	5
2	25	1	1	6
3	33	1	1	7
4	40	1	2	0
5	52	2	2	0
6	60	2	4	0
7	69	2	3	0
8	1000	0	1	0

*Passo 5* - Estraggo 5 dalla “testa” del serbatoio (aggiorno la riga 5) e considero la stella in uscita da 5



Belmann  $c_j \leq c_i + c_{ij}$  è verificato?

$60 \leq 52 + 20$  ? si!  $\Rightarrow$  non modifico i dati relativi a 6

$1000 \leq 52 + 48$  ? no!  $\Rightarrow$  8 nel serbatoio

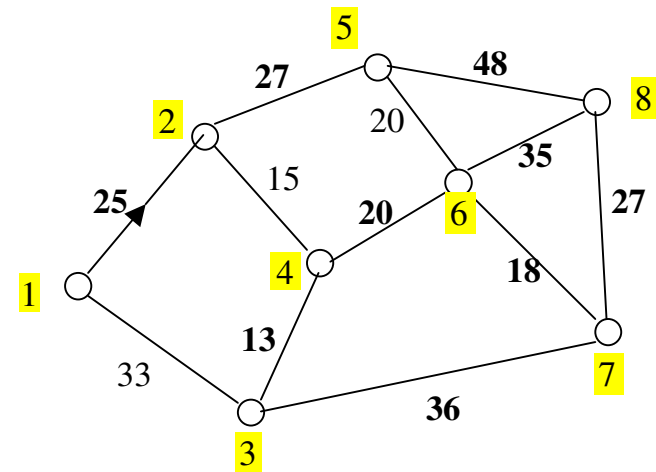
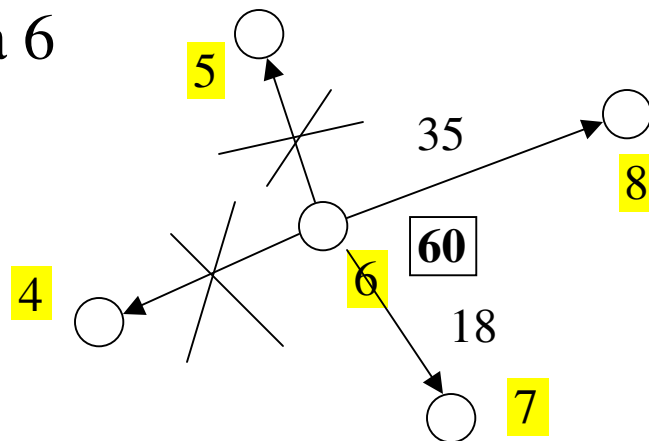
Aggiorno i dati della 8<sup>o</sup> riga e riordino il serbatoio

1	0	1	0	5
2	25	1	1	6
3	33	1	1	7
4	40	1	2	0
5	52	2	2	0
6	60	2	4	0
7	69	2	3	0
8	1000	0	1	0



1	0	1	0	6
2	25	1	1	7
3	33	1	1	8
4	40	1	2	0
5	52	1	2	0
6	60	2	4	0
7	69	2	3	0
8	100	2	5	52

*Passo 6* - Estraggo 6 dalla “testa” del serbatoio (aggiorno la riga 6) e considero la stella in uscita da 6



$69 \leq 60 + 18$  ? si!  $\Rightarrow$  non modifico i dati relativi a 7

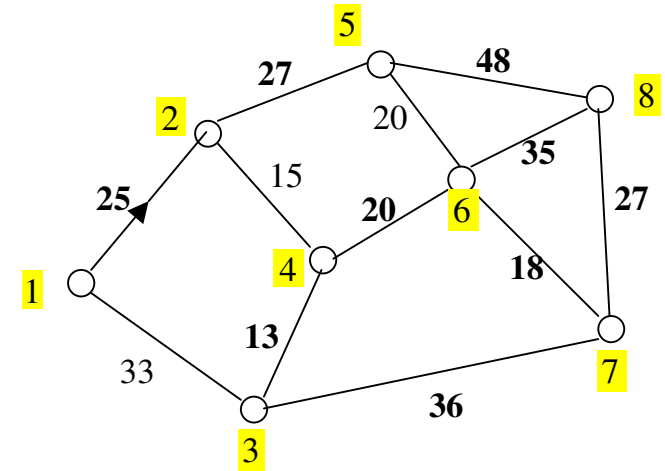
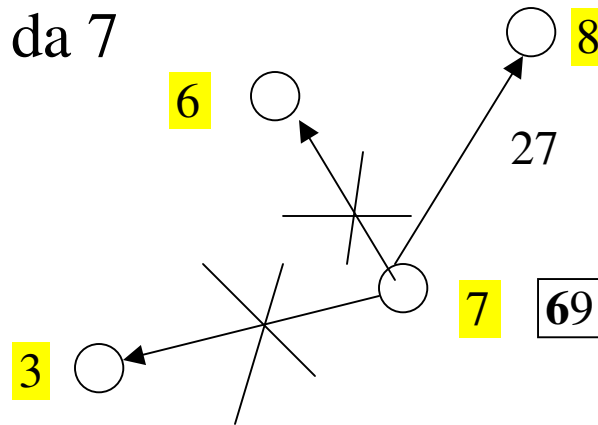
$100 \leq 60 + 35$  ? no!  $\Rightarrow$  modifico i dati di 8 (è già nel serbatoio) e riordino il serbatoio

1	0	1	0	6
2	25	1	1	7
3	33	1	1	8
4	40	1	2	0
5	52	1	2	0
6	60	2	4	0
7	69	2	3	0
8	100	2	5	0



1	0	1	0	7
2	25	1	1	8
3	33	1	1	0
4	40	1	2	0
5	52	1	2	0
6	60	1	4	0
7	69	2	3	0
8	95	2	6	53

*Passo 7* - Estraggo 7 dalla “testa” del serbatoio (aggiorno la riga 7) e considero la stella in uscita da 7



$95 \leq 69 + 27$  ? si!  $\Rightarrow$  non modifico i dati relativi a 8

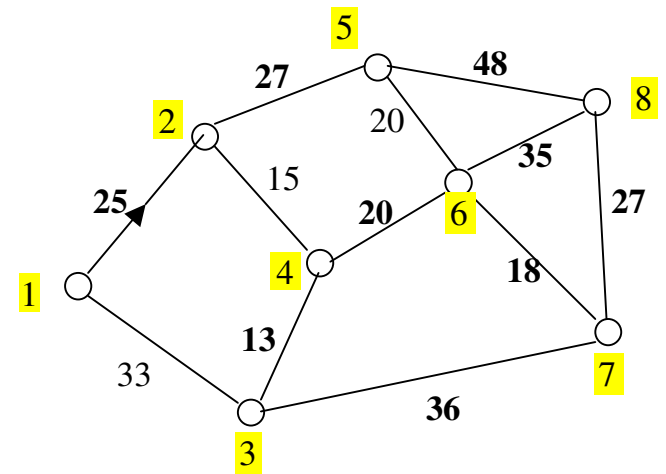
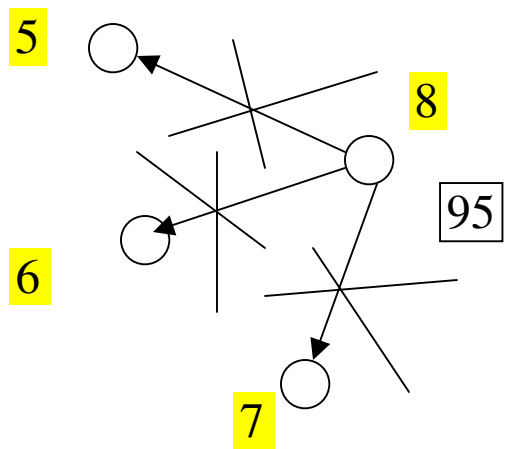
1	0	1	0	7
2	25	1	1	8
3	33	1	1	0
4	40	1	2	0
5	52	1	2	0
6	60	1	4	0
7	69	2	3	0
8	95	2	6	0



1	0	1	0	8
2	25	1	1	0
3	33	1	1	0
4	40	1	2	0
5	52	1	2	0
6	60	1	4	0
7	69	1	3	0
8	95	2	6	0

*Passo 8* - Estraggo 8 dalla “testa” del serbatoio (aggiorno la riga 8) e considero la stella in uscita da 8.

I nodi estremi della stella sono tutti nodi per i quali l’itinerario è stato individuato in modo definitivo (hanno codice 1).



Il serbatoio è vuoto il procedimento è terminato.

1	0	1	0	8
2	25	1	1	0
3	33	1	1	0
4	40	1	2	0
5	52	1	2	0
6	60	1	4	0
7	69	1	3	0
8	95	2	6	0

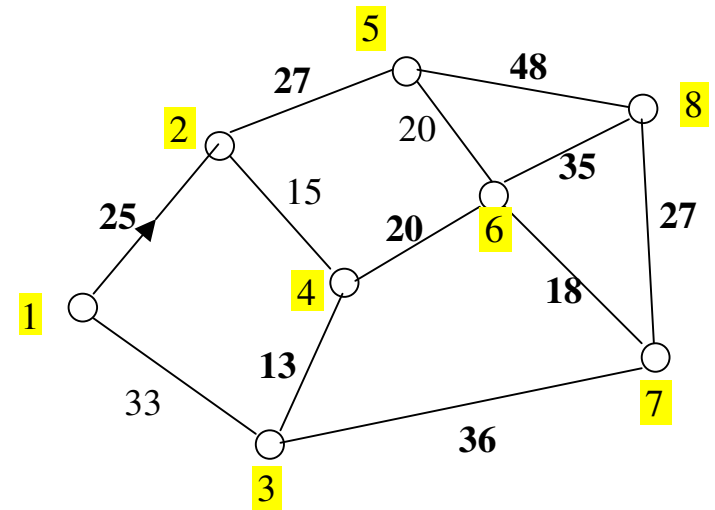


1	0	1	0	0
2	25	1	1	0
3	33	1	1	0
4	40	1	2	0
5	52	1	2	0
6	60	1	4	0
7	69	1	3	0
8	95	1	6	0

55

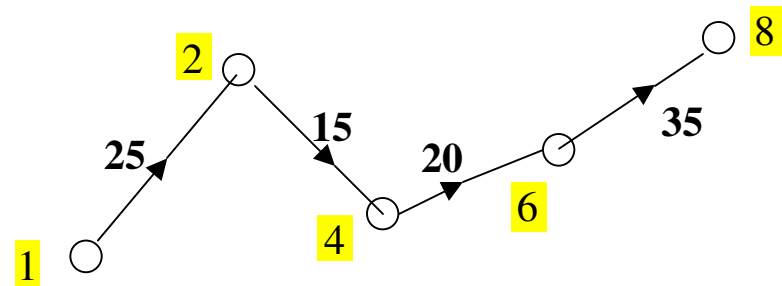
*Matrice finale*

1	0	1	0	0
2	25	1	1	0
3	33	1	1	0
4	40	1	2	0
5	52	1	2	0
6	60	1	4	0
7	69	1	3	0
8	95	1	6	0

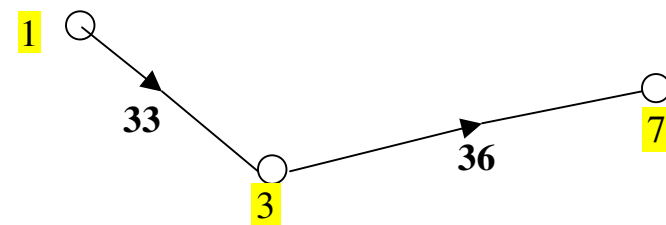


Per ricostruire i percorsi minimi “leggo a ritroso” la 3° colonna della matrice a 4 colonne.

Così l’itinerario dalla radice (nodo 1) al nodo 8 è: 8-6; 6-4; 4-2; 2-1.

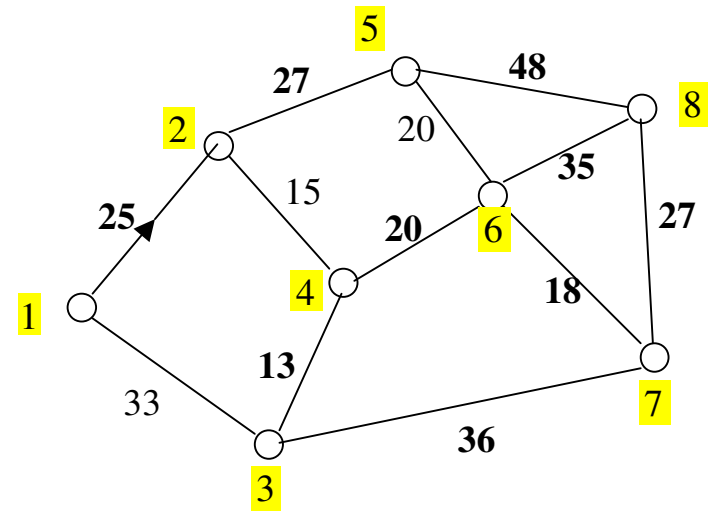


L’itinerario dalla radice (nodo 1) al nodo 7 è: 7-3;3-1.

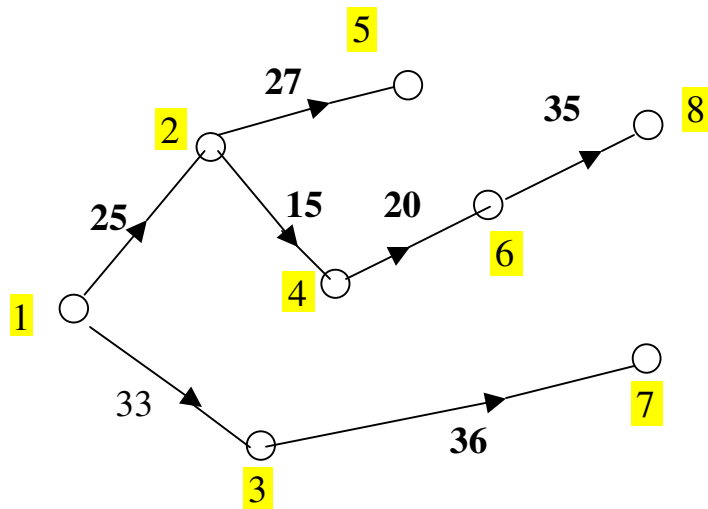




1	0	1	0	0
2	25	1	1	0
3	33	1	1	0
4	40	1	2	0
5	52	1	2	0
6	60	1	4	0
7	69	1	3	0
8	95	1	6	0



*Albero* degli itinerari di *minimo costo*



# Altri algoritmi per determinare l'albero degli itinerari di minimo costo

Caratteristica dell'algoritmo di Dijkstra: alla *fine* di ciascun *passo riordino* il vettore serbatoio in *ordine crescente di costo* (dalla testa del serbatoio).

Questo ha come conseguenza che se un *nodo  $i$*  viene *estratto dal serbatoio* non ci può più rientrare e l'itinerario (dalla radice) per esso individuato è quello *definitivo*. Questo perché:

a) I nodi che sono nel serbatoio hanno costo (dell'itinerario dal nodo radice) superiore al costo di  *$i$* : *perché prima di estrarre si riordinano i nodi del serbatoio in ordine crescente di costo*.

b) I nodi che entreranno nel serbatoio, *dopo l'estrazione di  $i$* , avranno *costi maggiori di  $i$* , perché il costo di questi nodi l'otterrò aggiungendo costi di archi, che sono quantità positive finite, ai costi dei nodi che sono nel serbatoio al momento dell'estrazione di  *$i$* .

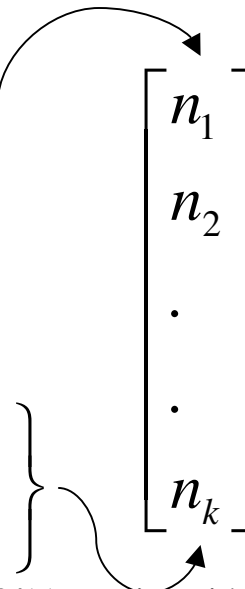
## Algoritmo L-deque (“double queue”: doppia coda )

E' stato però notato che il *riordino del serbatoio nell'algoritmo di Dijkstra* è “*costoso*” in termini di calcolo (dobbiamo pensare ad applicazioni a reti di centinaia, ed anche migliaia, di nodi e di archi).

L'algoritmo di L-deque differisce da Dijkstra per l'*organizzazione del vettore serbatoio*: il vettore serbatoio non viene riordinato in ordine crescente di costo alla fine di ciascun passo, questo fa sì che un nodo può entrare più di una volta nel vettore serbatoio (al contrario di quanto avviene con l'algoritmo di Dijkstra).

Se è già stato nel serbatoio :  
entra dalla testa

Se è la prima volta che entra  
nel vettore serbatoio entra dalla coda



# Algoritmo L-deque

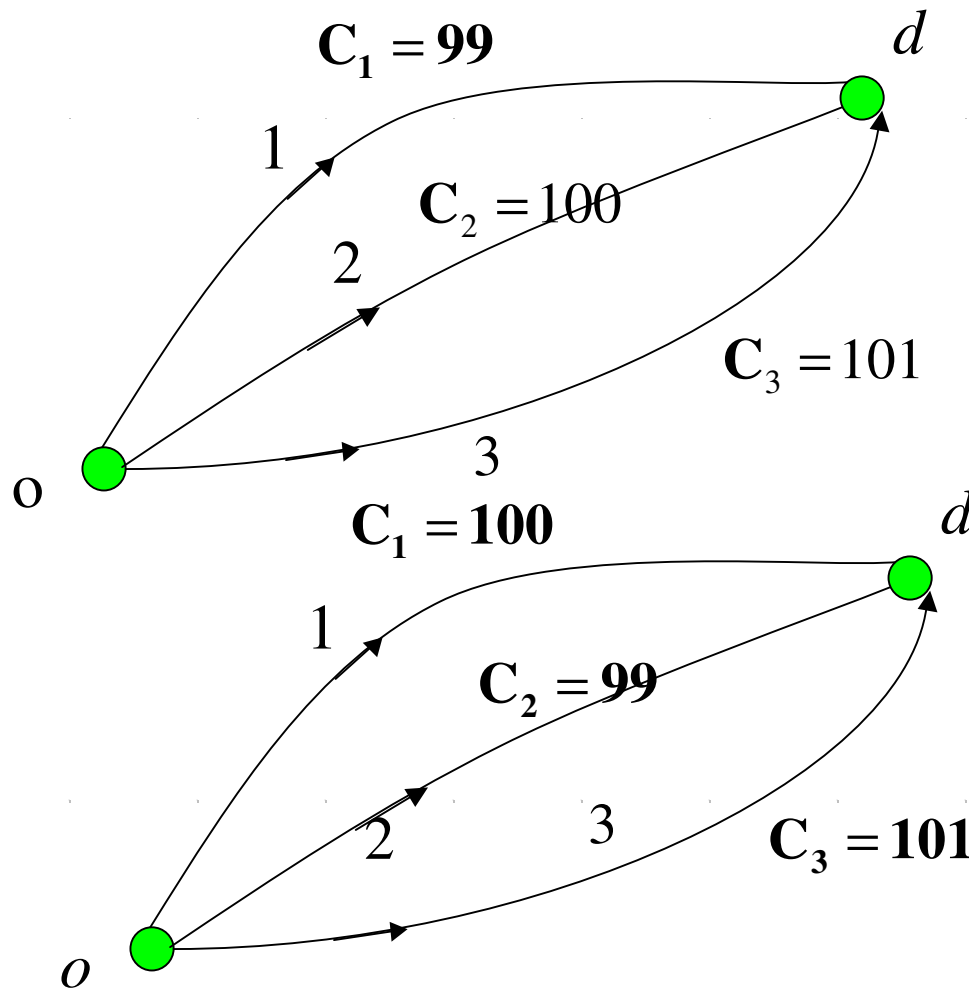
Come faccio a sapere se un nodo è la prima volta che entra nel vettore serbatoio, oppure c'è già stato? Se ha il *costo iniziale molto elevato* (1000 nell'esempio dimostrativo visto) vuol dire che non è mai stato nel vettore serbatoio, altrimenti c'è stato.

E' stato rilevato, sperimentalmente, che l'algoritmo *L-deque è più veloce* di quello di Dijkstra.

Quello di Dijkstra ha però il vantaggio che la *successione dei nodi estratti dalla testa* del vettore serbatoio sono in *ordine crescente di costo* dal nodo radice.

## Problemi dell'assegnazione tutto o niente

L'assegnazione tutto o niente è *instabile*



$$D_{od} = 1000 \text{ veic/h}$$

$$\begin{cases} F_1 = 1000 \\ F_2 = 0 \\ F_3 = 0 \end{cases}$$

$$D_{od} = 1000 \text{ veic/h}$$

$$\begin{cases} F_1 = 0 \\ F_2 = 1000 \\ F_3 = 0 \end{cases}$$

Inoltre *non è realistico* considerare che tutti gli utenti stimino il costo sugli itinerari *allo stesso modo*.

## Assegnazione di tipo probabilistico

Per ovviare all'instabilità del modello tutto o niente (ToN) si ricorre ad un modello comportamentale appartenente alla categoria dei modelli di utilità aleatoria.

*L'insieme di scelta* in questo caso è *l'insieme degli itinerari* che collegano la data coppia o-d: indichiamolo con  $P_{od}$ .

$$U_i = V_i + \varepsilon_i$$

Nei modelli di utilità aleatoria l'utilità è data da una parte deterministica ed una aleatoria (l'utilità è perciò una variabile aleatoria)

Nel caso dei modelli di utilità aleatoria applicati alla scelta dell'itinerario :  $U_i = -\tilde{C}_i$

$\tilde{C}_i$  , il costo di un itinerario, è una variabile aleatoria , (assume valori diversi con diverse probabilità): gli utenti percepiscono ( stimano) il costo di un itinerario in maniera diversa.

# Modelli di utilità aleatoria applicati alla scelta dell'itinerario

$$U_i = V_i + \varepsilon_i \quad e \quad U_i = -\tilde{C}_i$$



$$\tilde{C}_i = C_i + \varepsilon_i$$

*scarto aleatorio a media nulla* ( $E(\varepsilon_i) = 0$ )

*valore medio del costo* sull'itinerario  
(quantità deterministica)

Se  $\varepsilon_i$  è una variabile aleatoria di *Weibull-Gumbel* ho un *modello logit*.

Se  $\varepsilon_i$  è una variabile aleatoria *normale* (di Gauss) ho un *modello probit*.

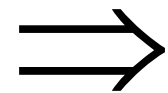
$$p_i = \text{prob}(U_i > U_j) = \text{prob}(\tilde{C}_i < \tilde{C}_j) \quad \forall j \in P_{OD}$$

L'alternativa di *massima utilità* sarà l'*itinerario di minimo costo*. Gli utenti scelgono l'itinerario di minimo costo: poiché però i diversi utenti percepiscono in maniera diversa il costo, posso solo determinare la probabilità di scelta dell'itinerario (alternativa i-esima).

Una volta che determino la probabilità di scelta di un itinerario posso determinare il flusso, medio, sull'itinerario,

$$F_i = D_{od} p_i \quad \longleftarrow \quad \text{Assegno la domanda}$$

Una volta che ho il flusso medio sull'itinerario posso determinare i flussi medi sugli archi.





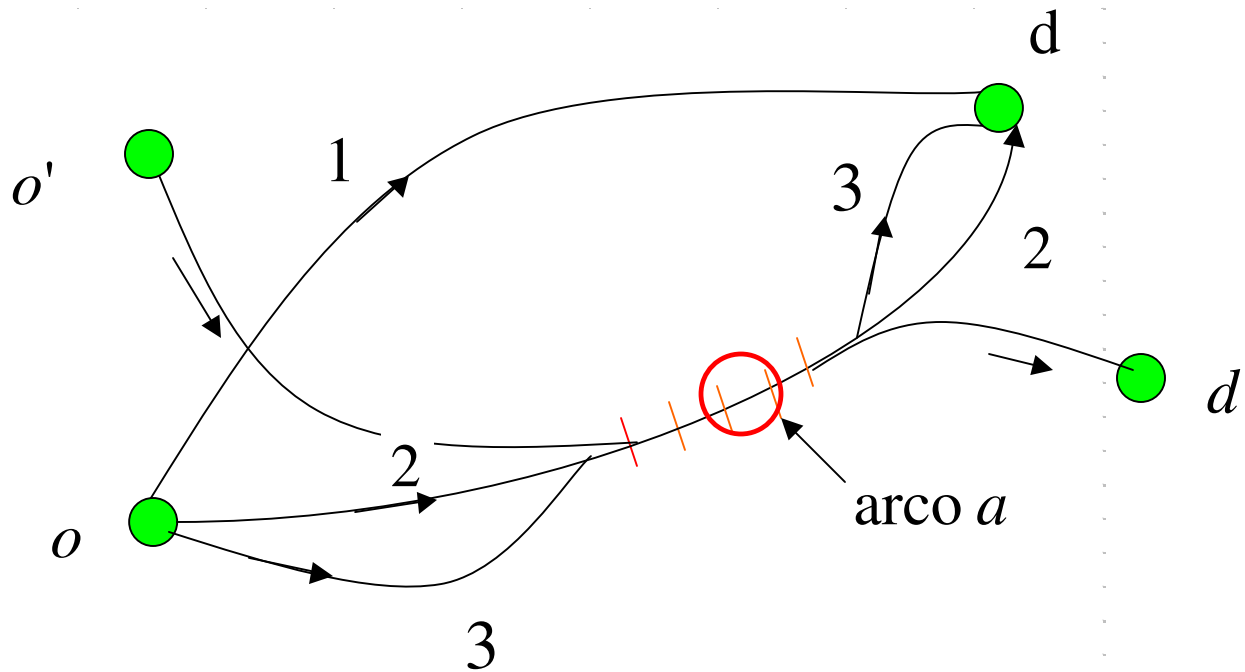
$$f_a = \sum_{\forall od \in I_{od}} \sum_{\forall i \in P_{od}} \delta_{ai} p_i D_{od} = \sum_{\forall od \in I_{od}} \sum_{\forall i \in P_{od}} \delta_{ai} F_i$$

$f_a$  : flusso sul generico arco  $a$

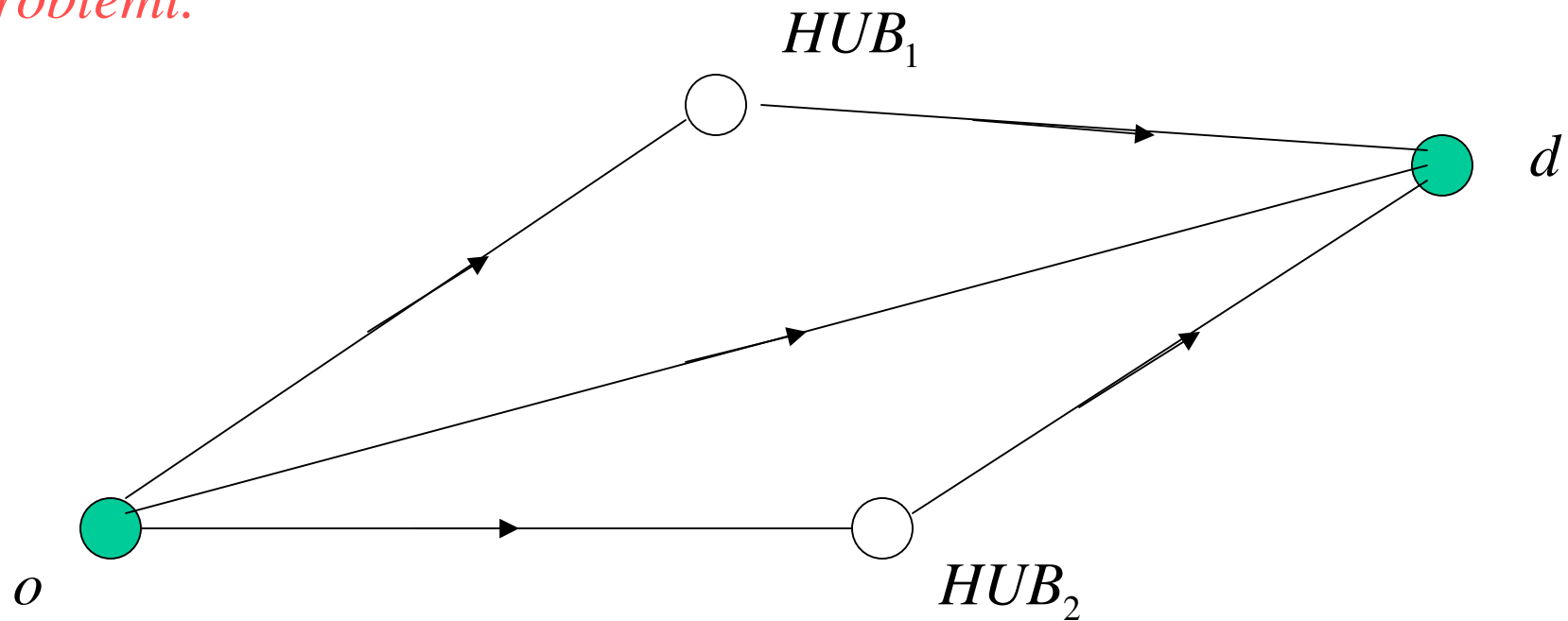
$I_{od}$  : insieme delle coppie  $od$  nella rete

$P_{od}$  : insieme degli itinerari che congiungono la coppia  $od$

$\delta_{ai}$   $\begin{cases} \rightarrow =1 \text{ se l'arco } a \text{ fa parte dell'itinerario } i \\ \rightarrow =0 \text{ se l'arco } a \text{ non fa parte dell'itinerario } i \end{cases}$



Rete di *trasporto aereo*: in questo caso il modello *logit non dà problemi*.



Tre alternative di itinerario che non si sovrappongono

# Esempio: modello di scelta dell'itinerario in una rete di trasporto aereo

$$V_{ir} = \beta_1 t_{ir} + \beta_2 f_{ir} + \beta_3 F_{ir} + \beta_4 D_{ir} + \beta_5 C_{ir}$$

$i$ : coppia od

$r$ : itinerario

$t_{ir}$  : *tempo del viaggio* (in ore)

$f_{ir}$  : *frequenza del servizio* (voli/giorno)

$F_{ir}$  : *tariffa* (dollari)

$D_{ir}$  : variabile legata al *tipo di aereo* (per esempio: 0 se  $N < 30$ , 0,5 se  $30 < N < 50$ , 1 se  $N > 50$ )

$C_{ir}$  : variabile legata al *tipo di servizio* (per esempio: 0 nel caso di volo diretto "non stop", 0,5 nel caso di volo diretto però con una fermata, 1 nel caso di volo indiretto ossia con almeno un cambio di aereo)

Atlanta (Stati Uniti) 19 coppie *od* (1980)

$$V_{ir} = -0.897t_{ir} + 0,239f_{ir} - 0,0278F_{ir} + 1,461D_{ir} - 1,557C_{ir}$$

$$\frac{\beta_{\text{cambio di aereo}}}{\beta_{\text{tempo}}} = \frac{1,557 \text{ utilità/cambio di aereo}}{0,897 \text{ utilità/ora}} \cong 1,7 \frac{\text{ore}}{\text{cambio di aereo}}$$

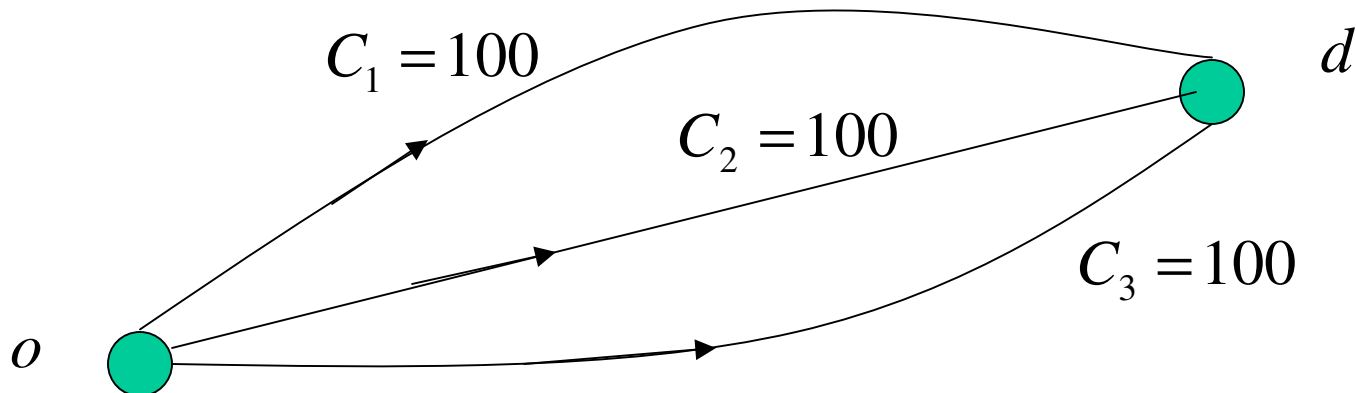
$$\frac{\beta_{\text{tempo}}}{\beta_{\text{tariffa}}} = \frac{0,897 \text{ utilità/ora}}{0,0278 \text{ utilità/\$}} \cong 32 \frac{\$}{\text{ora}}$$

$$\frac{\beta_{\text{frequenza}}}{\beta_{\text{tariffa}}} = \frac{0,239 \text{ utilità/volo aggiunto}}{0,0278 \text{ utilità/\$}} \cong 8,6 \frac{\$}{\text{volo aggiunto}}$$

$$\frac{\beta_{\text{frequenza}}}{\beta_{\text{tempo}}} = \frac{0,239 \text{ utilità/volo aggiunto}}{0,897 \text{ utilità/ore}} \cong 0,27 \frac{\text{ore}}{\text{volo aggiunto}}$$

Anche nel caso stradale, se non ci sono sovrapposizioni di itinerario, il modello funziona bene.

Per esempio tre alternative di itinerario:



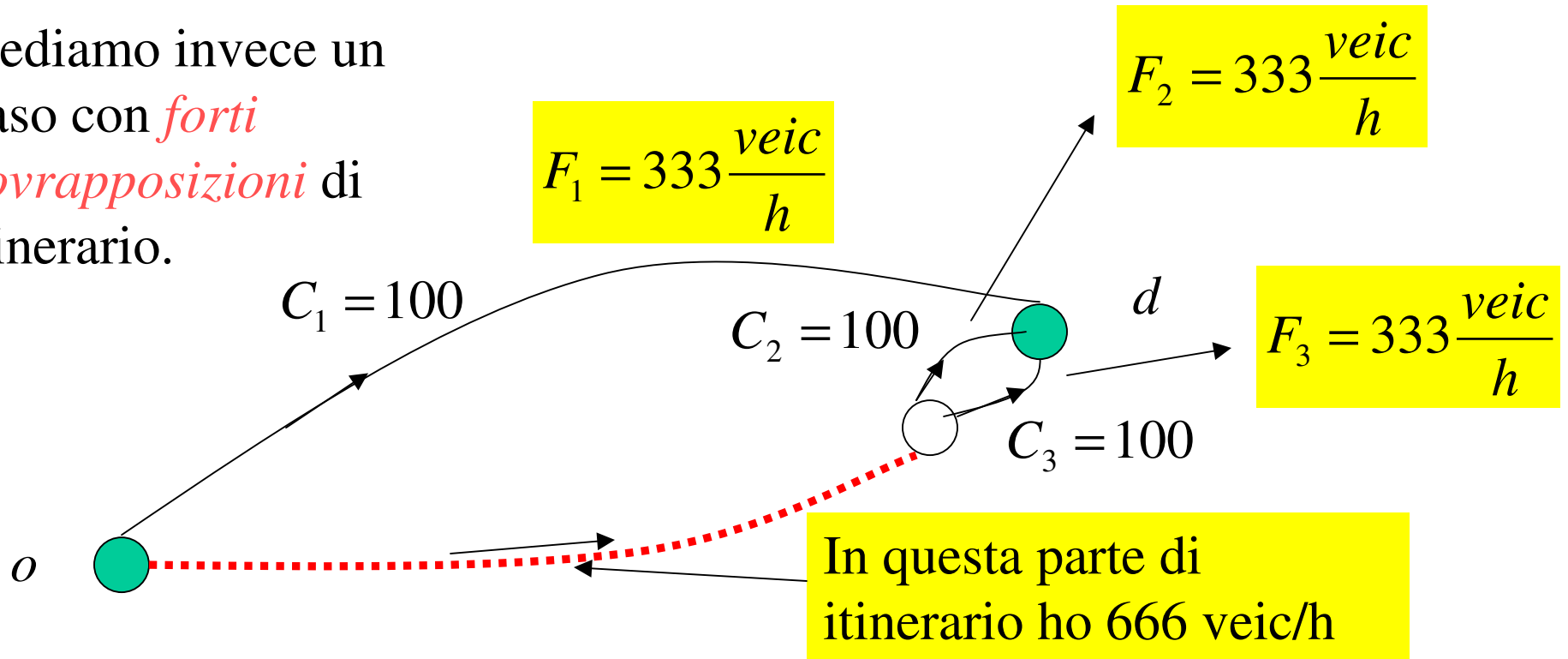
$$F_1 = p_1 D_{od} \approx 333 \frac{veic}{h}$$

$$F_2 = p_2 D_{od} \approx 333 \frac{veic}{h}$$

$$F_3 = p_3 D_{od} \approx 333 \frac{veic}{h}$$

$$p_1 = \frac{e^{-\alpha C_1}}{e^{-\alpha C_1} + e^{-\alpha C_2} + e^{-\alpha C_3}} = \frac{1}{3}$$

Vediamo invece un caso con *forti sovrapposizioni* di itinerario.



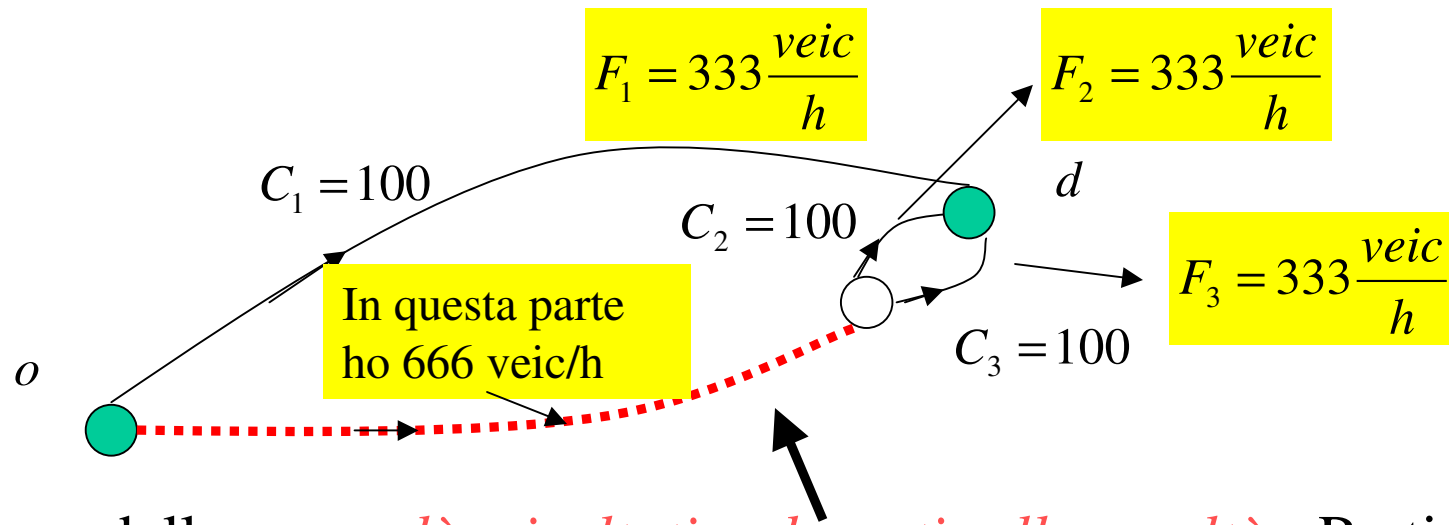
Il modello logit in questo caso dà gli stessi risultati che nel caso precedente

$$p_1 = \frac{e^{-\alpha C_1}}{e^{-\alpha C_1} + e^{-\alpha C_2} + e^{-\alpha C_3}} = \frac{1}{3}$$

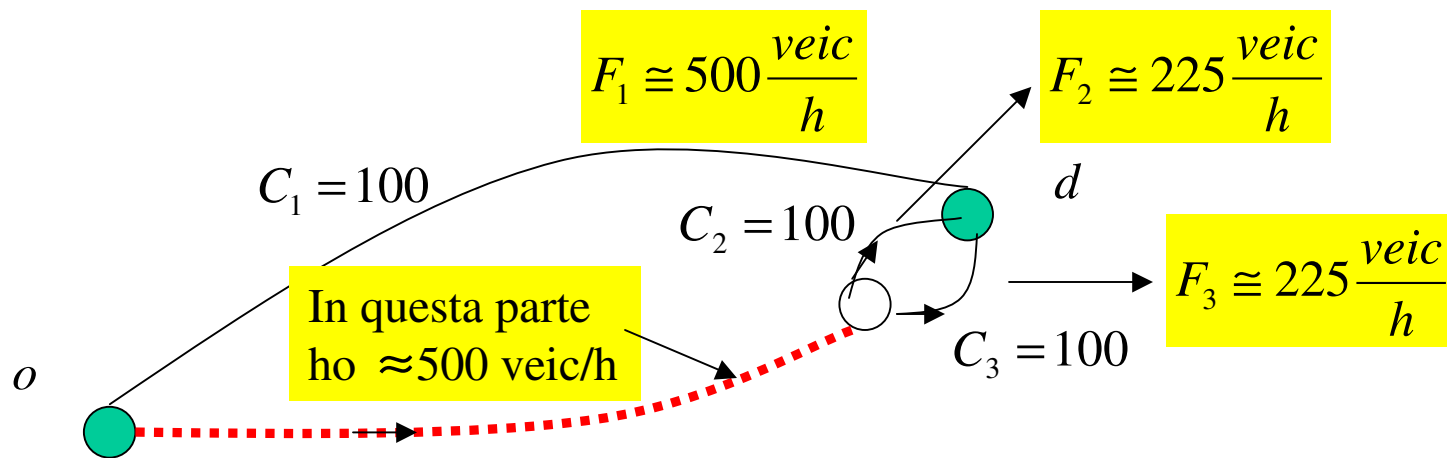
$$F_1 = p_1 D_{od} \approx 333 \frac{veic}{h}$$

$$F_2 = p_2 D_{od} \approx 333 \frac{veic}{h}$$

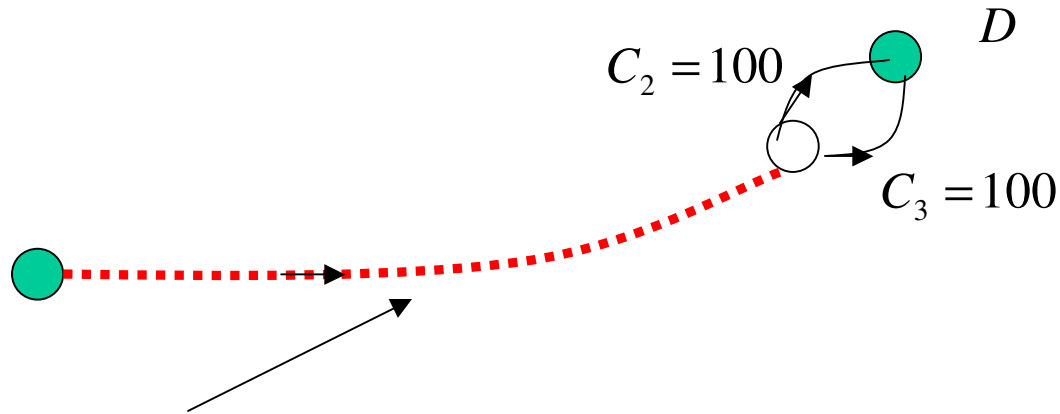
$$F_3 = p_3 D_{od} \approx 333 \frac{veic}{h}$$



Il modello *non dà risultati aderenti alla realtà*. Praticamente l'utente non percepisce 2 e 3 come alternative differenti. Nella realtà ci si aspetta una distribuzione dei flussi del tipo seguente:



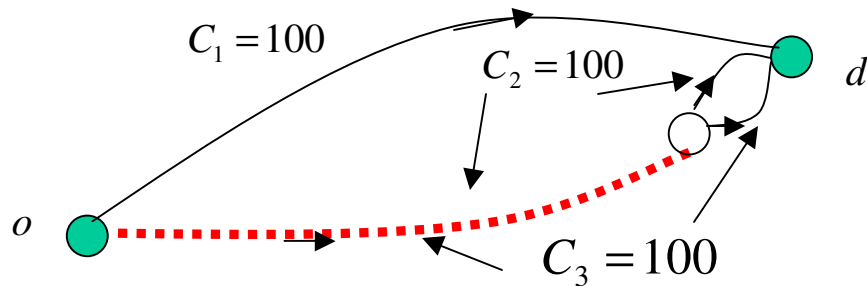
Il modello logit non dà risultati aderenti alla realtà in quanto una ipotesi base del modello logit è che le *utilità* (in questo caso i *costi* delle alternative) siano *variabili aleatorie indipendenti*.



In questo caso l'*ipotesi* che i costi sugli itinerari siano *variabili aleatorie indipendenti* non è vicina alla realtà perché gli itinerari si sovrappongono fortemente: un utente che associa all'itinerario 2 un costo superiore (inferiore) alla media, assocerà, molto probabilmente, anche all'itinerario 3 un costo superiore (inferiore) alla media.



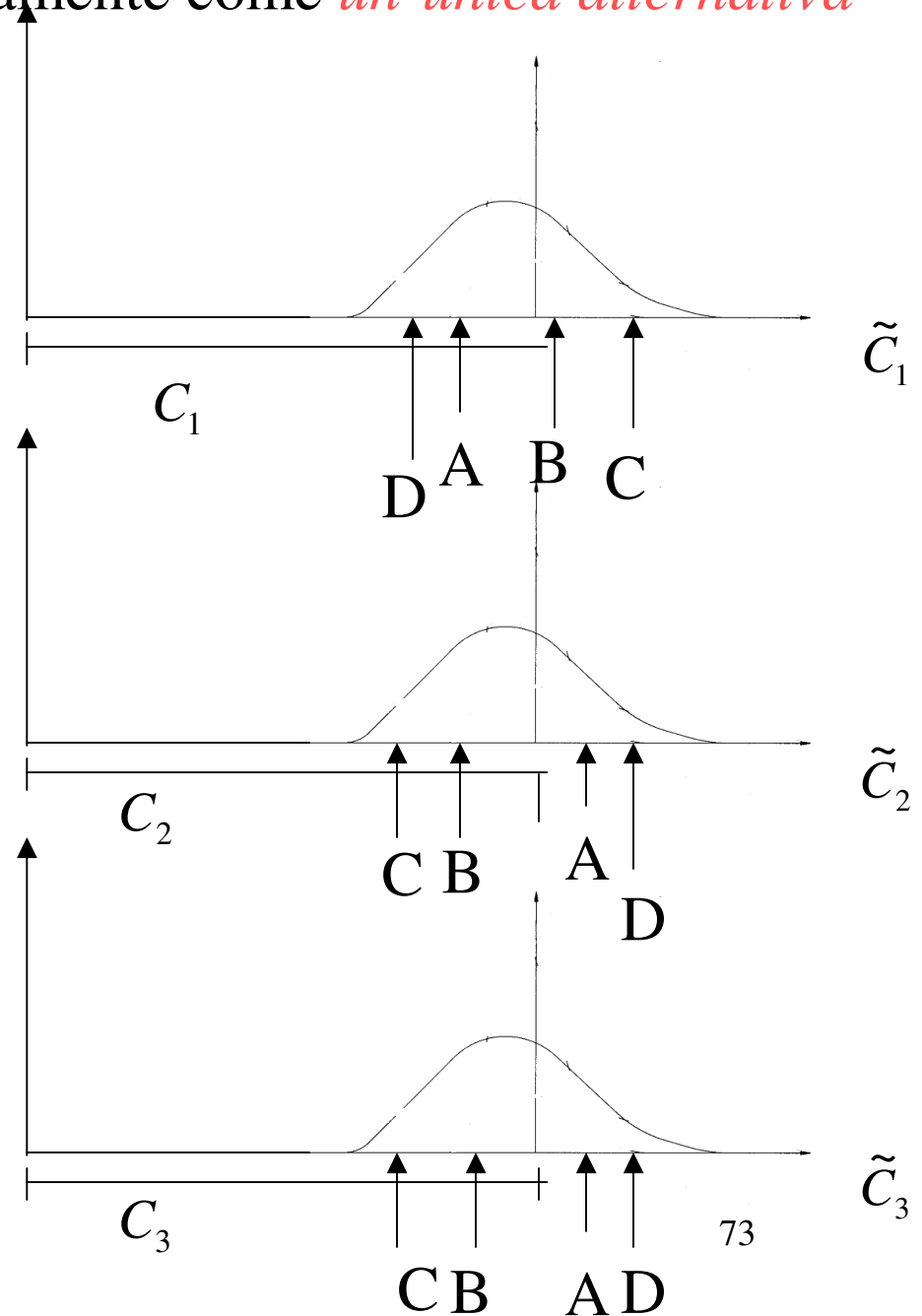
Le alternative 2 e 3 funzionano praticamente come *un'unica alternativa*



Tutti gli utenti, A, B, C e D percepiscono i costi delle alternative 2 e 3 in modo molto simile: in quanto le alternative si sovrappongono fortemente.

Gli utenti A e D scelgono l'alternativa 1.

Gli utenti B e C sceglieranno l'alternativa 2 o 3. L'insieme delle due alternative funziona *in pratica* come *un'unica alternativa*.



# Modello probit

Voglio un modello che tenga conto della dipendenza delle utilità (costi nella scelta dell'itinerario) fra le alternative (itinerari).

Il modello logit non va bene: introduco il *modello probit*

$$U_i = -\tilde{C}_i + \varepsilon_i \quad \text{Al solito suppongo che sia} \quad E[\varepsilon_i] = 0$$

$$E[U_i] = -E[\tilde{C}_i] = -C_i$$

Nel modello probit i costi *sugli archi* sono variabili aleatorie *indipendenti* e distribuiti secondo una variabile aleatoria normale (Gauss).

La variabile  $\tilde{c}_a$  è una variabile aleatoria normale

$$c_a = E[\tilde{c}_a] \quad \leftarrow \text{Media del costo su } a$$

$$(\alpha c_a)^2 = \text{var}[\tilde{c}_a] \quad \leftarrow \text{Varianza del costo su } a$$

$$(\text{Altra ipotesi } \text{var}[\tilde{c}_a] = \beta c_a \quad )$$

$$\tilde{c}_a \sim N(c_a, (\alpha c_a)^2)$$

media

varianza

$\tilde{c}_a$  variabile aleatoria *normale*:

Funzione di *densità di probabilità*:

$$f_{\tilde{c}_a}(t) = \frac{1}{\sqrt{2\pi \text{var}(\tilde{c}_a)}} e^{-\frac{(t-\mu_{\tilde{c}_a})^2}{2 \text{var}(\tilde{c}_a)}}$$

Costo su un itinerario: sommatoria dei costi degli archi che lo compongono:

$$\tilde{C}_P = \sum_{\forall a \in A} \delta_{ap} \tilde{c}_a$$

$$\delta_{ap} = \begin{cases} 1 & \text{se } a \in P \\ 0 & \text{se } a \notin P \end{cases}$$

Insieme di archi della rete

Costo sull'itinerario è una variabile aleatoria.

Questa è la media

$$E[\tilde{C}_P] = C_P = \sum_{\forall a \in A} \delta_{ap} E[\tilde{c}_a] = \sum_{\forall a \in A} \delta_{ap} c_a$$

*Costo medio* su un *itinerario* è uguale alla *sommatoria* dei *costi medi* degli *archi* che lo compongono.

$$\text{var}(\tilde{C}_P) = \sum_{\forall a \in A} \delta_{aP} \text{var}(\tilde{c}_a) = \sum_{\forall a \in A} \delta_{aP} (\alpha c_a)^2$$

*Varianza* del costo *su un itinerario* è uguale alla *somma* delle *varianze* dei costi sugli archi che compongono l'itinerario: questa eguaglianza è valida perché ho supposto i costi *sugli archi* variabili aleatorie *indipendenti*.

*Attenzione*: nel modello *probit* i costi sugli *archi* sono variabili aleatorie (normali) *indipendenti, non i costi sugli itinerari* (come nel modello logit: ed era l'indipendenza dei costi *sugli itinerari*, alternative, che mi creava dei problemi nel logit).



I costi sugli *itinerari* sono variabili aleatorie *dipendenti*?  
 Posso individuare un “*grado*” di dipendenza?

A questo fine studio la covarianza fra i costi di due itinerari che congiungono una stessa coppia O-D.

Definizione di covarianza  
 fra due variabili aleatorie  
 generiche  $X$  e  $Y$

$$Cov(X, Y) = E[(X - \mu_X)(Y - \mu_Y)]$$

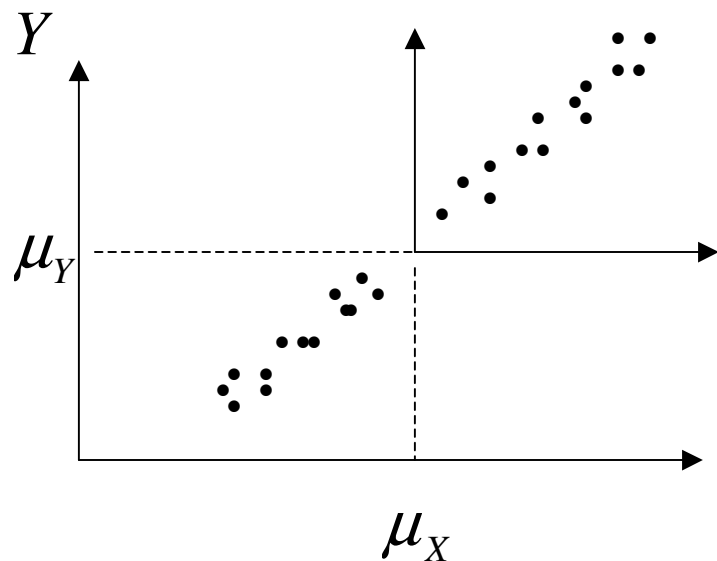
Ammettiamo che  $X$  e  $Y$  siano variabili aleatorie *indipendenti*

$$Cov(X, Y) = E[(X - \mu_X)(Y - \mu_Y)] = E[XY - X\mu_Y - \mu_X Y + \mu_X \mu_Y] =$$

$$E[XY] - \mu_Y E[X] - \mu_X E[Y] + \mu_X \mu_Y = \mu_X \mu_Y - \mu_X \mu_Y - \mu_X \mu_Y + \mu_X \mu_Y = 0$$

E' valida solo se  $X$  e  $Y$  sono  
 indipendenti

$$Cov(X, X) = E[(X - \mu_X)(X - \mu_X)] = E[(X - \mu_X)^2] = var(X)$$



Cosa vuol dire se ho una varianza *positiva alta*?

Se X è superiore alla media anche Y è superiore alla media. Se X è inferiore alla media, anche Y è inferiore alla media. Le due variabili non sono indipendenti.

Covarianza positiva:

$$Cov(X, Y) > 0$$

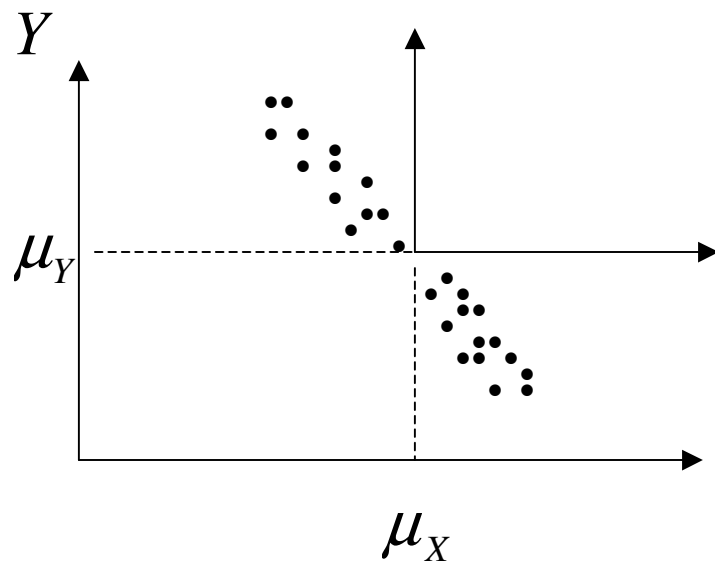
$$E[(X - \mu_X)(Y - \mu_Y)]$$

+

+

-

-



Covarianza **negativa** ed alta (in valore assoluto).

Se  $X$  è superiore alla media,  $Y$  è inferiore alla media. Se  $X$  è inferiore alla media,  $Y$  è superiore alla media. Le due variabili non sono  $X$  indipendenti.

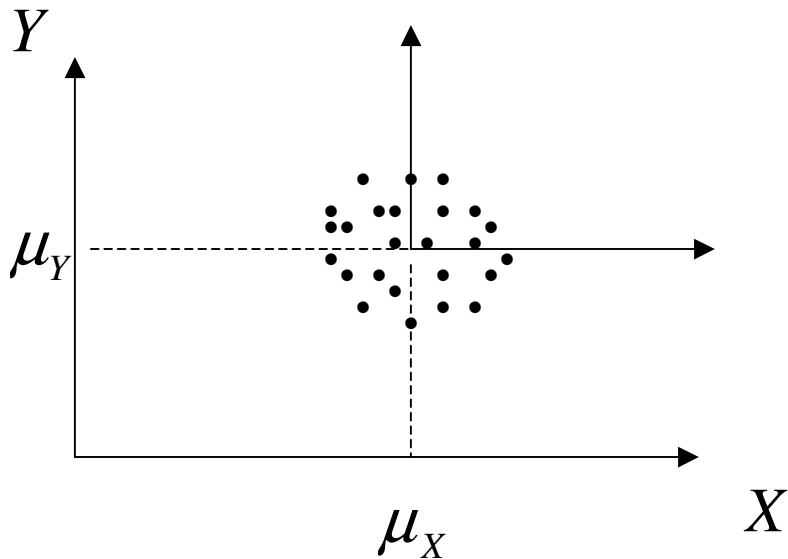
Covarianza negativa:

$$\text{Cov}(X, Y) < 0$$

$$E[(X - \mu_X)(Y - \mu_Y)]$$

<b>+</b>	<b>-</b>
<b>-</b>	<b>+</b>





Covarianza  $\cong 0$

Sapere qualcosa sulla X (per esempio che è superiore alla media) non mi dà informazioni sulla Y. In media i prodotti positivi si elidono con quelli negativi.

Covarianza  $\cong 0$

$$Cov(X, Y) \cong 0$$

$$E[(X - \mu_X)(Y - \mu_Y)]$$

+	-
+	+
-	-
-	+

*Modello probit*: per sapere se i costi di due itinerari sono variabili aleatorie indipendenti o dipendenti, ed in quale misura, *studio la covarianza*.

Definizione di covarianza:

Costo su un itinerario uguale a somma dei costi sugli archi

$$Cov(\tilde{C}_p, \tilde{C}_q) = E[(\tilde{C}_p - C_p)(\tilde{C}_q - C_q)] = E\left[\left(\sum_{a \in A} \delta_{ap} (\tilde{c}_a - c_a)\right) \left(\sum_{l \in A} \delta_{lq} (\tilde{c}_l - c_l)\right)\right]$$

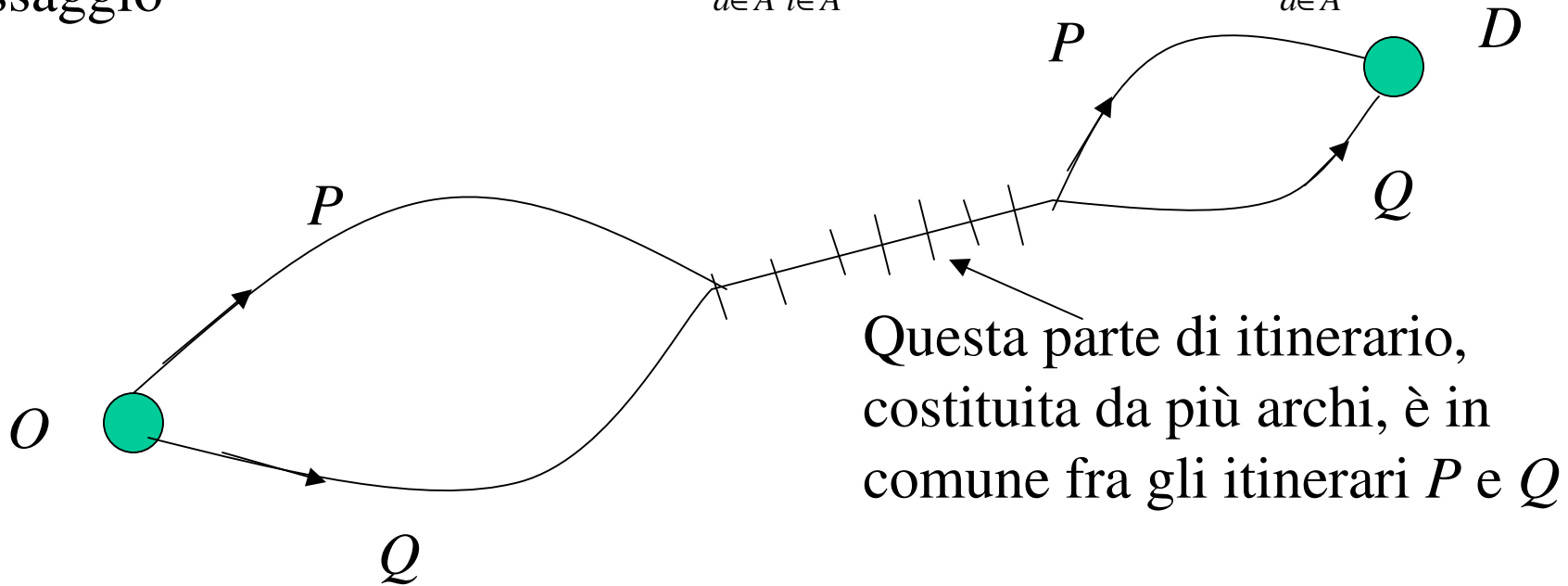
La media è un operatore lineare

$$= E\left[\sum_{a \in A} \sum_{l \in A} \delta_{ap} \delta_{lq} (\tilde{c}_a - c_a)(\tilde{c}_l - c_l)\right] = \sum_{a \in A} \sum_{l \in A} \delta_{ap} \delta_{lq} E[(\tilde{c}_a - c_a)(\tilde{c}_l - c_l)] =$$

Covarianza dei costi sugli archi  $a$  e  $l$

$$= \sum_{a \in A} \sum_{l \in A} \delta_{ap} \delta_{lq} Cov(\tilde{c}_a, \tilde{c}_l) = \sum_{a \in A} \delta_{ap} \delta_{aq} var(\tilde{c}_a) = \sum_{a \in A} \delta_{ap} \delta_{aq} (\alpha c_a)^2$$

Vediamo di capire il perché del passaggio  $\sum_{a \in A} \sum_{l \in A} \delta_{ap} \delta_{lq} Cov(\tilde{c}_a, \tilde{c}_l) = \sum_{a \in A} \delta_{ap} \delta_{aq} var(\tilde{c}_a)$

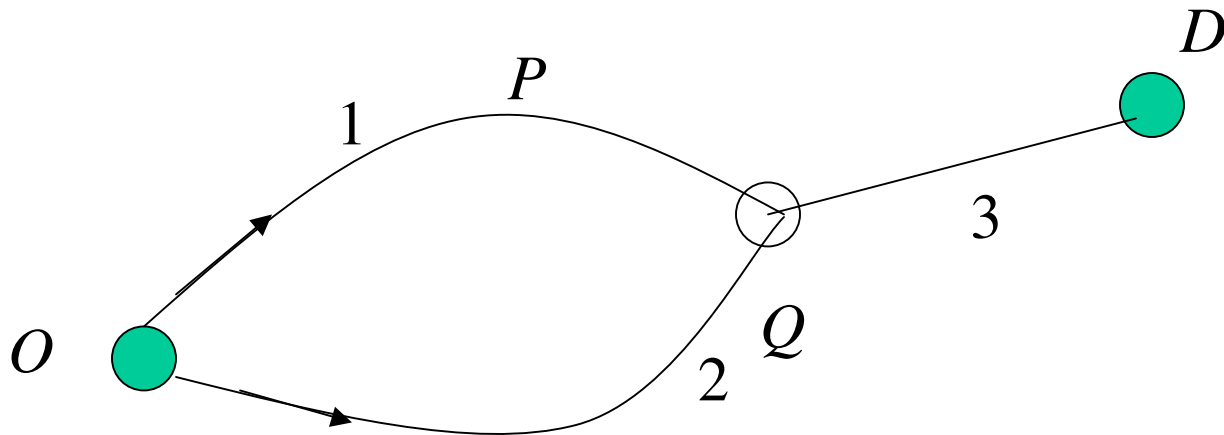


Poiché nel modello probit i costi sugli archi sono variabili indipendenti sarà:  $Cov(\tilde{c}_a, \tilde{c}_l) = 0$  se  $a \neq l$  ;  $Cov(\tilde{c}_a, \tilde{c}_l) = var(\tilde{c}_a)$  se  $a = l$

$\sum_{a \in A} \sum_{l \in A} \delta_{ap} \delta_{lq} Cov(\tilde{c}_a, \tilde{c}_l) = \sum_{a \in A} \delta_{ap} \delta_{aq} var(\tilde{c}_a)$  Somma delle varianze degli archi a comune dei due itinerari.

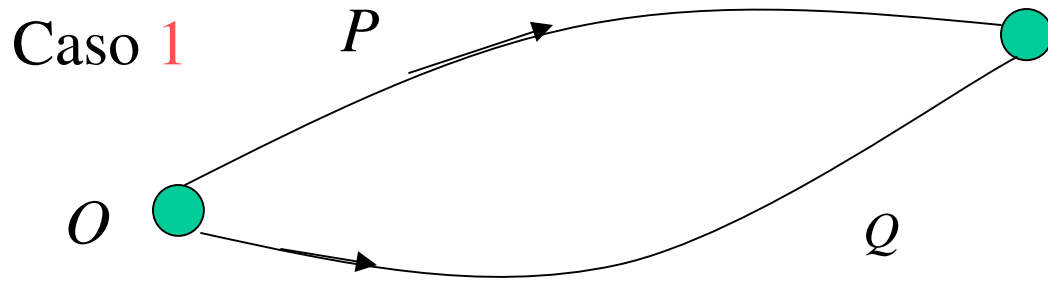
Il prodotto  $\delta_{ap} \delta_{aq}$  è diverso da 0 solo per gli archi che sono in comune ai due itinerari

Esempio: itinerario  $P$ , archi 1 e 3; itinerario  $Q$ , archi 2 e 3

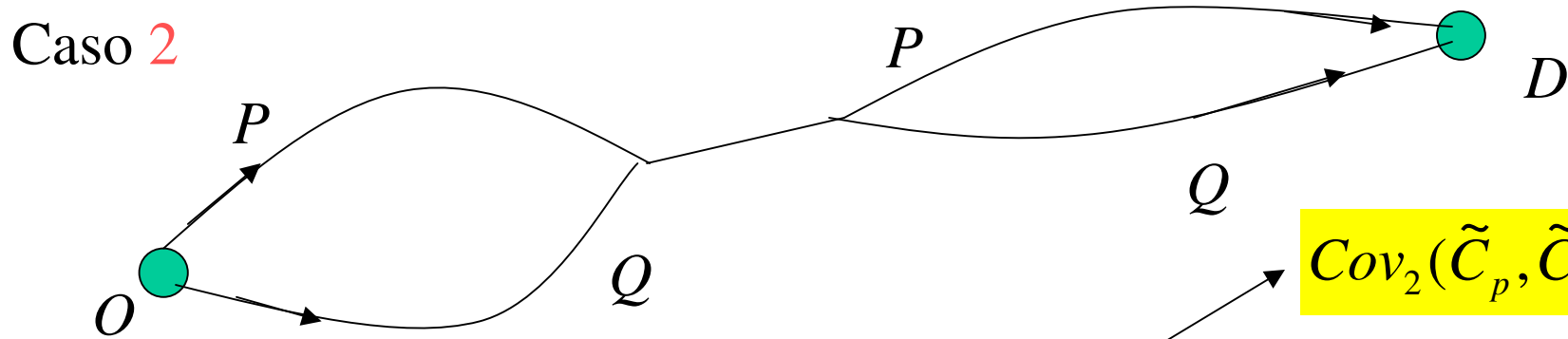


$$\begin{aligned}
 \sum_{a \in A} \sum_{l \in A} \delta_{ap} \delta_{lq} \text{Cov}(\tilde{c}_a, \tilde{c}_l) &= \delta_{1p} \delta_{1q} \text{Cov}(\tilde{c}_1, \tilde{c}_1) + \delta_{1p} \delta_{2q} \text{Cov}(\tilde{c}_1, \tilde{c}_2) + \delta_{1p} \delta_{3q} \text{Cov}(\tilde{c}_1, \tilde{c}_3) \\
 &+ \delta_{2p} \delta_{1q} \text{Cov}(\tilde{c}_2, \tilde{c}_1) + \delta_{2p} \delta_{2q} \text{Cov}(\tilde{c}_2, \tilde{c}_2) + \delta_{2p} \delta_{3q} \text{Cov}(\tilde{c}_2, \tilde{c}_3) + \\
 &+ \delta_{3p} \delta_{1q} \text{Cov}(\tilde{c}_3, \tilde{c}_1) + \delta_{3p} \delta_{2q} \text{Cov}(\tilde{c}_3, \tilde{c}_2) + \delta_{3p} \delta_{3q} \text{Cov}(\tilde{c}_3, \tilde{c}_3) \\
 &= 1 \quad = 1 \quad = \text{var}(\tilde{c}_3)
 \end{aligned}$$

$\nearrow = 0$        $\nearrow = 0$        $\nearrow = 0$   
 $\nearrow = 0$        $\nearrow = 0$        $\nearrow = 0$   
 $\nearrow = 0$        $\nearrow = 0$



$D$   $Cov_1(\tilde{C}_p, \tilde{C}_q) = 0$   
 I costi sugli itinerari sono *indipendenti*



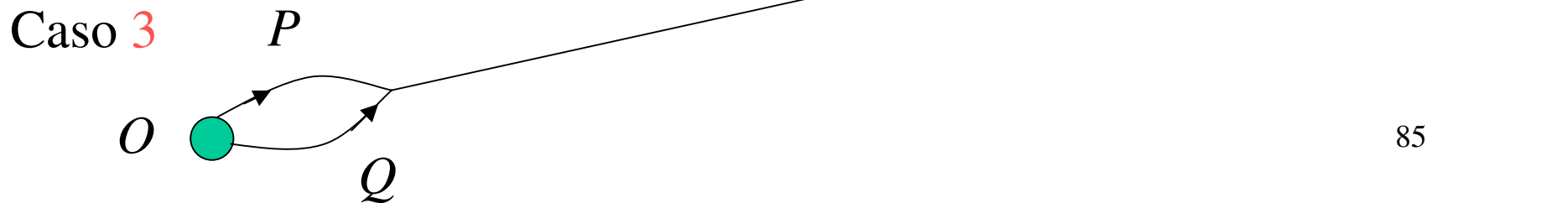
I costi sugli itinerari, nei casi 2 e 3, non sono *indipendenti*

$Cov_2(\tilde{C}_p, \tilde{C}_q) \neq 0$

$Cov_3(\tilde{C}_p, \tilde{C}_q) \neq 0$

Tanto più dipendenti,  $Cov_3(\tilde{C}_p, \tilde{C}_q) > Cov_2(\tilde{C}_p, \tilde{C}_q)$

quanto più sono gli archi a comune



Come faccio a determinare la *probabilità di scelta di un'alternativa* con il modello probit?

Nel caso del logit avevo una forma esplicita piuttosto semplice:

$$p_i = \underset{\forall j \neq i}{\text{prob}}(U_i > U_j) = \frac{e^{\alpha V_i}}{\sum_{\forall j} e^{\alpha V_j}}$$

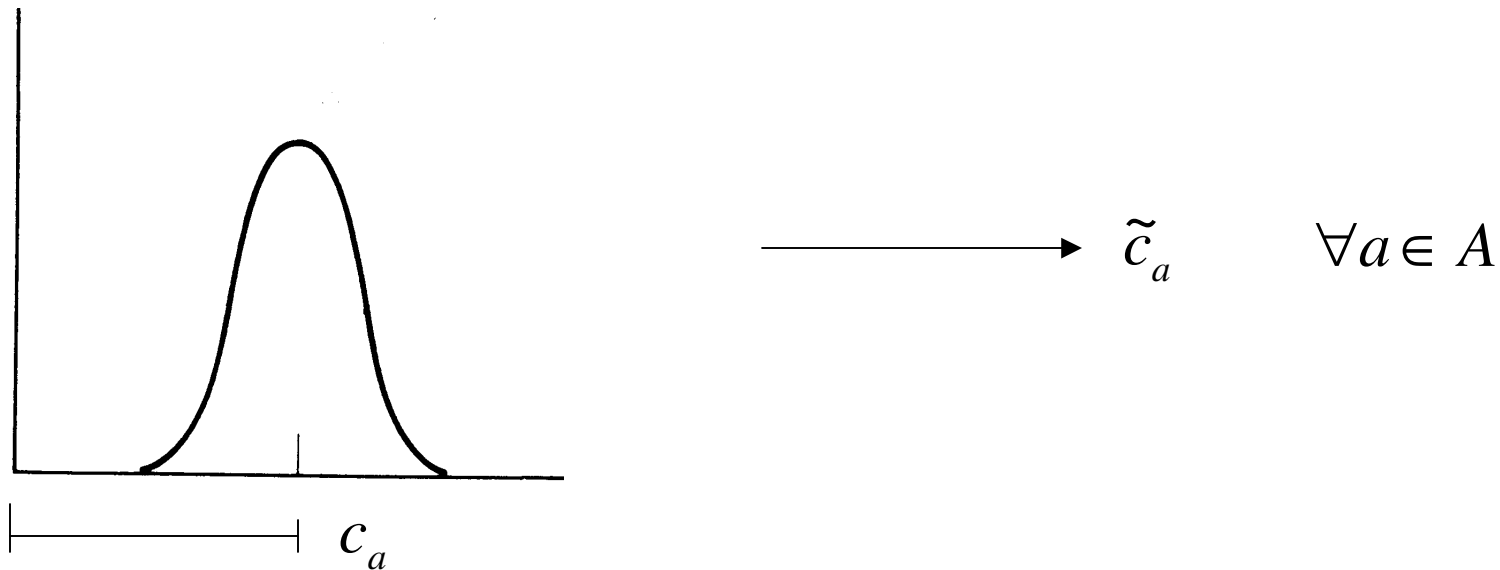
Il problema del modello probit è che *non ho una forma esplicita* di questo tipo.

Inoltre ho il problema di *enumerare* le alternative, ossia gli *itinerari* fra ciascuna coppia O-D che, in *reti di grandi dimensioni*, non è un problema di facile soluzione (questo problema è comune anche al modello logit applicato a reti di grandi dimensioni).

Ricorro allora alla *simulazione*.

# Modello probit – Simulazione di Montecarlo

- Estraggo i costi sugli archi.



$$\tilde{c}_a \sim N(c_a, (\alpha c_a)^2)$$

(questi costi estratti per ciascun arco possono essere pensati come i costi percepiti da un certo utente)

I costi estratti possono essere utilizzati per una *assegnazione tutto o niente di tipo deterministico*

La domanda  $D_{od}$ , fra una determinata coppia od, viene tutta assegnata all'itinerario di minimo costo (in base ai costi estratti in una certa iterazione) : ottengo un vettore  $\mathbf{f}^1$  di flussi sugli archi.

- Ripeto *l'estrazione dei costi sugli archi*

- Ripeto l'assegnazione tutto o niente  $\xrightarrow{\text{ottengo}}$   $\mathbf{f}^k$  (generica iterazione  $k$ )

I flussi stimati alla fine di una determinata iterazione  $k$  sono calcolati come *medie dei flussi* sugli archi ottenuti in *tutte* le iterazioni fino a quel momento compiute:

$$\bar{f}_a^k = \frac{[(k-1)\bar{f}_a^{k-1} + f_a^k]}{k}$$



# Passi dell'algoritmo probit

Passo 1: *Campionamento dei costi percepiti*

Per ogni arco della rete viene ottenuto un costo percepito  $\tilde{c}_a^k$

$$N[c_a, (\alpha c_a)^2] \rightarrow \tilde{c}_a^k$$

Passo 2: *Assegnazione tutto o niente*

Faccio una assegnazione tutto o niente con i costi costanti  $\tilde{c}_a^k$

Otengo un vettore dei flussi sugli archi:  $\mathbf{f}^k$

### Passo 3: *Media dei flussi*

$$\bar{f}_a^k = \frac{[(k-1)\bar{f}_a^{k-1} + f_a^k]}{k}$$

$\forall a \in A$  (insieme degli archi della rete)

### Passo 4: *Test di arresto*

$$\max_{\forall a \in A} \frac{|\bar{f}_a^k - \bar{f}_a^{k-1}|}{\bar{f}_a^{k-1}} < \varepsilon_1$$

$$\frac{1}{n_a} \sum_{a \in A} \frac{|\bar{f}_a^k - \bar{f}_a^{k-1}|}{\bar{f}_a^{k-1}} < \varepsilon_2$$

( per esempio:  $\varepsilon_1 = 0,05$ ;  $\varepsilon_2 = 0,02$  )

Se il *test non è verificato*, si aumenta il contatore  $k = k + 1$  e si ritorna al passo 1.

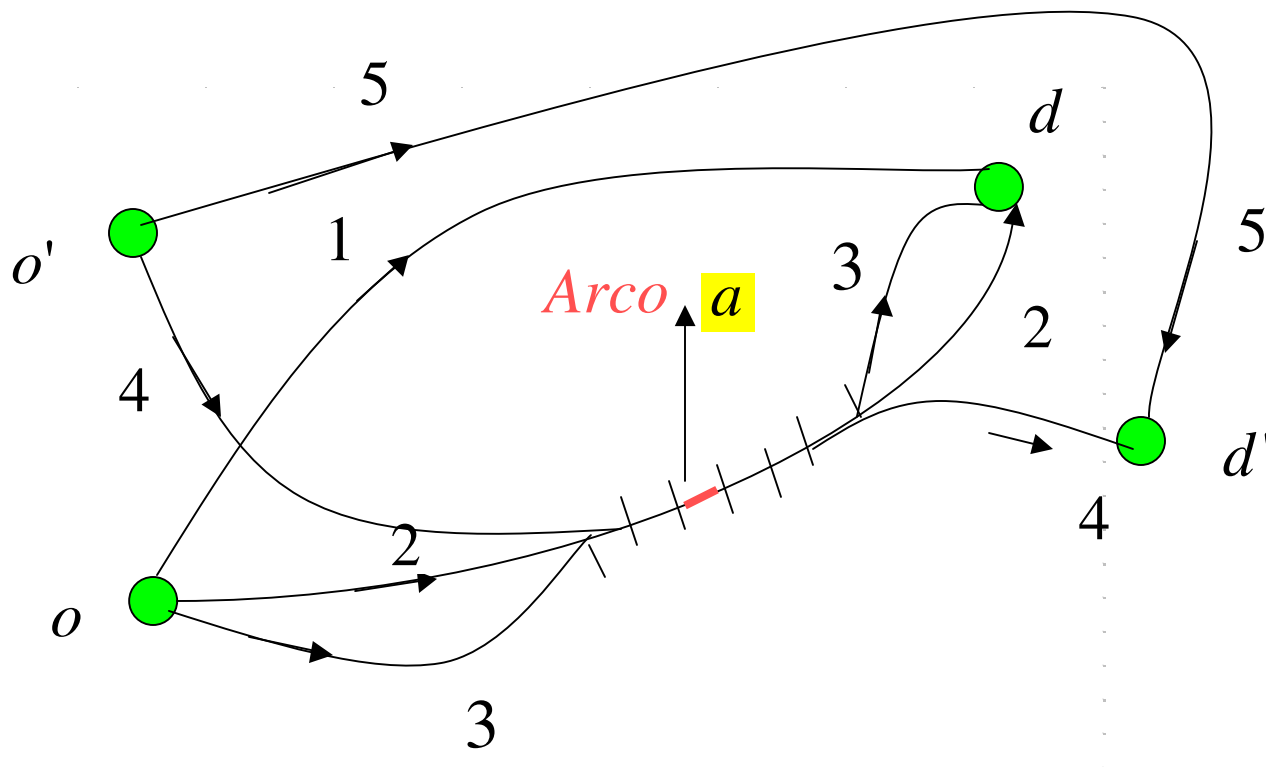
Vediamo di capire come mai agendo in questo modo ottengo implicitamente una assegnazione di tipo probit.

$$\sum_{h=1}^k f_a^h = \sum_{od \in I_{od}} \sum_{j \in P_{od}} \delta_{aj} k_j D_{od}$$

Questo è il flusso totale che assegno all'arco  $a$  in  $k$  iterazioni

$$\delta_{aj} k_j D_{od}$$

Numero di volte nel quale l'itinerario  $j$  risulta di minimo costo in  $k$  iterazioni

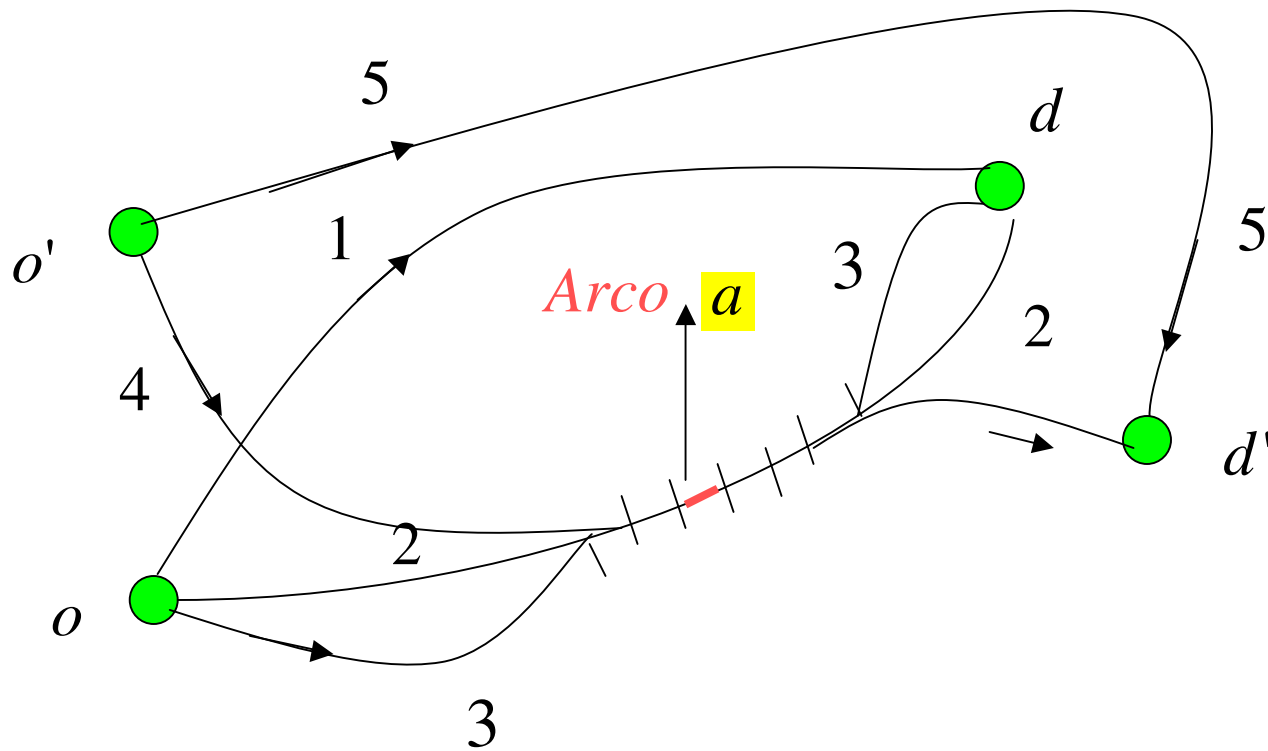


$$\sum_{h=1}^k f_a^h = \sum_{od \in I_{od}} \sum_{j \in P_{od}} \delta_{aj} k_j D_{od} = \delta_{a1} k_1 D_{od} + \delta_{a2} k_2 D_{od} + \delta_{a3} k_3 D_{od} + \delta_{a4} k_4 D_{o'd'}$$

$\downarrow$   $\downarrow$   $\downarrow$   $\downarrow$   
 ~~$=0$~~   $=1$   $=1$   $=1$

$$+ \delta_{a5} k_5 D_{o'd'}$$

$\downarrow$   
 ~~$=0$~~



La media del flusso sull'arco  $a$  in  $k$  iterazioni è data da

$$\bar{f}_a^k = \frac{1}{k} \sum_{h=1}^k f_a^h = \sum_{od \in I_{od}} \sum_{j \in P_{od}} \delta_{aj} \frac{k_j}{k} D_{od}$$

$\frac{k_j}{k}$  \_\_\_\_\_ E' la *frequenza* per la quale l'itinerario  $j$  risulta di minimo costo

Se  $k \rightarrow \infty$   $\frac{k_j}{k} \rightarrow p_j$  Probabilità di utilizzo di un itinerario  $j$  in un modello probit

$\frac{k_j}{k} D_{od} \cong p_j D_{od}$  E' la stima del *flusso medio*  $F_j$  sull'*itinerario*  $j$

$\bar{f}_a^k = \frac{1}{k} \sum_{h=1}^k f_a^h = \sum_{od \in I_{od}} \sum_{j \in P_{od}} \delta_{aj} F_j$  è quindi una stima della *media del flusso sull'arco*  $a$  derivante da un *modello di scelta dell'itinerario di tipo probit.*