

Alcune considerazioni circa l'offerta formativa e la soddisfazione nelle prime esperienze lavorative dei laureati¹

Franca Crippa*, Ralph M. Feldberg**

**Università di Milano – Bicocca; **Valuepartners*

Riassunto. Nella transizione Università – lavoro eventi favorevoli scaturiscono da diverse determinanti, indagabili mediante i modelli di regressione logistica, non scervri tuttavia da limiti (Porcu e Muggeo, 2005). La ricerca di strumenti che potenzino l'analisi conduce, fra gli altri, ai modelli di Regressione Logistica Booleana (Braumoeller, 2003) e di Regressione Strutturale Generalizzata (Popescul e Ungar, 2004). Questi ultimi esplorano le strutture di legami fra le osservazioni, vagliandole in seguito mediante il modello logistico. L'utilità di tale strumento è saggiata, in via preliminare, sulla situazione occupazionale dei neolaureati, verificando *ex-post* i legami stimati mediante analisi di raggruppamento. Questi ultimi risultati di consentono un'interpretazione in chiave dinamica della relazione fra la soddisfazione per le prospettive di carriera e le competenze statistiche e informatiche, relazione distorta nel modello logistico tradizionale.

Parole chiave: regressione logistica, struttura dei legami, analisi di raggruppamento.

1. Introduzione

Nella transizione Università – lavoro, eventi favorevoli quali l'inserimento nella realtà occupazionale o la soddisfazione per la posizione occupata sono le risultanti, in diversa misura, di una combinazione 'ottimale' fra le caratteristiche della formazione del candidato neolaureato e quelle della posizione aziendale offertagli. L'effetto esercitato da tale combinazione sulla probabilità di successo negli esiti delle prime e-

¹ Il presente lavoro è stato realizzato nell'ambito del progetto: "Outcomes. Transizioni Università-Lavoro e valorizzazione delle competenze professionali dei laureati: modelli e metodi di analisi multidimensionale delle determinanti", cofinanziato dal MIUR; coordinatore nazionale è Luigi Fabbris, coordinatore del gruppo dell'Università di Milano-Bicocca è Marisa Civardi. La nota è stata redatta per i parr. 2 e 3 da F. Crippa, per il par. 4 da Ralph M. Feldberg, per i parr. 1 e 5 congiuntamente.

sperienze lavorative è inquadrabile nella cornice teorica dei modelli di regressione logistica, impiegati per la loro natura discriminante ai fini della classificazione dei dati.

L'elevata dinamicità propria dei fenomeni in esame, suscettibili di cambiamenti netti nel tempo e nei diversi contesti, solleva tuttavia l'esigenza di supportare con strumenti esplorativi dei dati e delle loro relazioni il rigore metodologico dei modelli statistici atti a vagliare le relazioni stesse. I progressi della ricerca informatica nell'ultimo ventennio conducono, come è noto, alla specificazione di tecniche di *Data Mining* per l'individuazione di gruppi (*Group Detection*) e dell'analisi dei legami (*Link Analysis*) sulla base dei dati osservazionali, tecniche che alcuni autori hanno inteso unificare alla metodologia statistica, dando luogo fra gli altri metodi alla cosiddetta Regressione Logistica Strutturale Generalizzata per l'Analisi dei Legami (*Structural Generalized Logistic Regression (SGRL) for Link Analysis*, Popescul e Ungar, 2004).

La gestione di basi di dati propria di tale impostazione e l'esplorazione di legami fra gli stessi dati si fondano su di un impianto teorico articolato, generalmente nella forma di basi di dati relazionali, ovvero di più tabelle di dati (in luogo dell'unica matrice individui per variabili), alcune delle quali esprimono le relazioni in esame. Nel caso dell'incontro fra neolaureati ed aziende, alle due matrici descrittive delle caratteristiche della formazione universitaria dei primi e dell'occupazione offerta dalla seconda, si aggiungerebbe la matrice che associa ad ogni neolaureato l'azienda d'accoglienza.

Questo lavoro si propone di saggiare se vi siano evidenze che inducano a considerare utile un percorso analitico quale quello sinora succintamente esposto. A partire da uno studio di caso, già oggetto di precedenti elaborazioni (Crippa e Mariani, 2005), si vuole in primo luogo applicare un modello di regressione logistica che metta in relazione la soddisfazione del neolaureato (evento favorevole) con le caratteristiche del suo percorso formativo e con la forma contrattuale offerta dall'azienda. Successivamente, s'intendono esplorare i legami fra le caratteristiche del neolaureato e quelle dell'azienda, mediante l'applicazione dell'analisi di raggruppamento o *cluster analysis* per lo studio delle relazioni (Popescul e Ungar, 2004).

Lo scopo è essenzialmente quello di verificare se e come le connessioni emerse dal modello statistico siano conclusive oppure suscettibili di ulteriori specificazioni.

L'analisi dei raggruppamenti viene condotta tradizionalmente sui dati osservazionali. Il ricorso ai *database* relazionali e quindi ai modelli SGLR offre possibilità di ulteriori approfondimenti e, contemporaneamente, consente di misurare i guadagni in accuratezza nello studio delle dinamiche occupazionali dei neolaureati.

2. La percezione soggettiva dell'occupazione

I dati impiegati per l'analisi provengono dalla rilevazione, mediante intervista telefonica nel novembre 2003, delle caratteristiche formative e professionali dei 300 diplomati e laureati dall'A.A. 1988/89 all'A.A. 2002/2003 presso la Facoltà di Scienze Statistiche dell'Ateneo di Milano - Bicocca, in base alla lista fornita dalle Segreterie degli Studenti².

In ragione dell'elevata velocità d'inserimento nel mercato del lavoro degli statistici dell'Ateneo milanese (Crippa e Mariani, *op. cit.*), ai fini dell'applicazione si sono considerati i 270 soggetti occupati al momento dell'intervista. Su ciascuno d'essi, oltre al titolo, all'indirizzo di studi, alla votazione di laurea conseguita, si sono rilevate alcune qualifiche conseguibili durante il percorso formativo ed ipotizzate

Tabella 1. Distribuzione dei 270 laureati della Facoltà di Scienze Statistiche dell'Università di Milano-Bicocca, dall'A.A. 1988/89 all'A.A. 2002/2003 e occupati nel novembre 2003, rispetto alle caratteristiche del corso di studi universitario e dell'attuale posizione lavorativa, inizialmente considerate nell'analisi di regressione (valori percentuali).

<u>Percorso formativo universitario</u>			<u>Percorso lavorativo in azienda</u>		
<i>Titolo</i>	Diploma	46,4	<i>Tipologia lavorativa</i>	Dipendente	90,7
	Laurea triennale	24,1		Autonoma	9,3
	Laurea quadriennale	29,5			
<i>Stage richiesto da:</i>	Università	57,4	<i>Ente di diritto</i>	Privato	82,0
	Altri	5,2		Pubblico	18,0
<i>Formazione post lauream</i>	Master	3,5	<i>Posizione nella professione</i>	Dirigente	0,6
	Scuola specializzazione	1,5		Quadro	3,8
				Impiegata/o	95,6
<i>Indirizzo</i>	Statistico-economico	15,6	<i>Contratto a tempo</i>	Determinato	37,9
	Demografico-sociale	25,4		Indeterminato	61,6
	Statistico	12,6		Occasionale	0,5
	Diploma	46,4			
<i>Metodi statistici utilizzati</i>	Solo descrittivi	45,3	<i>Regime</i>	Tempo parziale	6,2
	Multivariati e/o inferenziali	54,7		Tempo pieno	93,8
			<i>Software applicativo</i>	SAS/SPSS	40,37
			Office/simili	59,63	

² La rilevazione, svolta retroattivamente nel novembre 2003, coinvolge oltre dieci coorti diverse e attraversa un arco temporale nel quale le condizioni di mercato subiscono modificazioni. Il permanere negli anni considerati di condizioni favorevoli all'occupazione nel territorio della Lombardia coinvolto nell'indagine attenua, almeno parzialmente, tali fonti di eterogeneità.

rilevanti nell'inserimento nel lavoro, quali lo svolgimento di un tirocinio formativo/*stage*, la frequenza a master di primo livello o corsi di specializzazione. Riguardo all'azienda d'accoglienza, si sono rilevate le caratteristiche dell'attività del laureato, se dipendente o autonoma, se con contratto a tempo indeterminato o altro, l'impiego del patrimonio di competenze specifiche acquisite negli anni universitari, in particolare quelle informatiche ed il grado di complessità della metodologia statistica impiegata (Tab. 1).

Sebbene si nutra l'impressione che proprio la dinamica dell'assunzione di neolaureati si avvantaggerebbe di riflessioni sul modello di SGRL, l'indagine sugli esiti positivi della ricerca d'occupazione non è qui espressiva perché riguarda la quasi totalità del collettivo. Si è pertanto focalizzata l'attenzione su indicatori soggettivi, relativi alla soddisfazione per il tipo d'attività intrapresa rispetto al trattamento economico ed alle possibilità di carriera, che conducono ad esiti positivi rispettivamente per il 70,4% ed il 75,4% dei neolaureati occupati. Si è condotta l'analisi delle variabili esplicative mediante due modelli di regressione logistica binaria, ciascuno dei quali colleghi la probabilità di una fra le due fonti di soddisfazione indagate alle caratteristiche, da un lato, del percorso formativo e, dall'altro, dell'attività in azienda.

3. La soddisfazione dei neolaureati come incontro fra offerta universitaria e richieste aziendali

Fra i molteplici contesti applicativi dell'analisi logistica, interessano in questa sede l'assegnazione dell'unità a gruppi ai fini classificatori e la capacità previsiva (Fabbris, 1997). Ciò consente di far emergere, nell'ambito delle prime esperienze lavorative dei neolaureati, sia le caratteristiche del percorso formativo e della condizione in azienda che generano situazioni d'accertata soddisfazione individuale, sia lo scenario nel mercato del lavoro conseguente a percorsi e/o scelte universitarie.

Come noto, il modello logistico binario è esprimibile come:

$$\text{Logit}(\pi) = \log \frac{\pi}{1-\pi} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k \quad (1)$$

dove la probabilità π d'esito favorevole di un determinato evento è data da:

$$\pi = \frac{\exp(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k)}{1 + \exp(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k)} \quad (2)$$

Le relazioni fra i differenti percorsi formativi universitari, dal lato della domanda occupazionale, e le caratteristiche del posto di lavoro, dal lato dell'offerta, possono dar luogo a condizioni giudicate dal neoassunto favorevoli o inadatte.

Nell'applicare il modello ai dati in oggetto, si assume così che, per un neolaureato in statistica, la probabilità di essere soddisfatto dell'attuale posizione lavorativa sia funzione degli attributi dei due ambienti, quello formativo universitario e quello aziendale (Tab.1), attributi che costituiscono i due sottoinsiemi di variabili esplicative:

$$\pi(\mathbf{x}) = \frac{\exp(\beta_0 + \sum_{\text{Università}} \beta_j X_j + \sum_{\text{Azienda}} \beta_j X_j)}{1 + \exp(\beta_0 + \sum_{\text{Università}} \beta_j X_j + \sum_{\text{Azienda}} \beta_j X_j)} \quad j_{\text{Università}}=1,\dots,h, j_{\text{Azienda}}=h+1,\dots,k; \quad (3)$$

L'intensità della 'superiorità' in termini di soddisfazione, da parte di un soggetto (denotato con B) che presenti una data caratteristica, rispetto a chi ne sia sprovvisto (A), è espresso come noto dal rapporto fra i due indici di coesistenza $\pi/(1-\pi)$ dei rispettivi individui (*odds ratio* (OR_{j^*})) relativamente alla covariata $j=j^*$, mantenendo il valore delle altre covariate fisso ed identico per i due soggetti A e B:

$$OR_{j^*} = \frac{\pi_{j^*}^B}{1 - \pi_{j^*}^B} \bigg/ \frac{\pi_{j^*}^A}{1 - \pi_{j^*}^A} \quad (4)$$

Esso è una misura d'associazione; un suo valore maggiore all'unità per una specifica covariata indica l'effetto incrementale di quest'ultima sulla probabilità di conseguire un esito positivo.

Fra tutte le possibili determinanti della soddisfazione per le condizioni economiche (Tab. 2), il procedimento di stima indica come variabili significative, al livello $\alpha=0,05$, lo stage/tirocinio nell'iter universitario e l'impiego, in azienda, di programmi a contenuto informativo-statistico più avanzato, segnatamente SAS e SPSS, oltre che il regime lavorativo a tempo pieno.

La bontà di adattamento del modello saggia, come noto, l'ipotesi di nullità di tutti i coefficienti espressi nell'equazione sottoposta a verifica, rispetto all'alternativa che almeno uno fra essi sia diverso da zero. I valori stimati, in relazione all'effetto delle tre variabili anzidette sulla soddisfazione, segnalano l'ottimo adattamento del modello (Tab.3).

Tali variabili agiscono tutte nella direzione d'incremento della soddisfazione per le condizioni economiche; al regime di tempo pieno in luogo di quello definito è associato l'effetto più elevato, mentre allo svolgimento di uno stage all'interno del proprio iter formativo e all'uso del software applicativo avanzato per le indagini statistiche corrispondono effetti sostanzialmente analoghi.

Lo stesso modello, relativo alla soddisfazione per le prospettive di carriera offerte dall'occupazione attuale, conduce a risultati che suggeriscono come l'attività a tempo pieno sia molto più favorevole a sviluppi professionali rispetto a quella parziale e come la formazione *post lauream*, in apparente controtendenza rispetto agli esiti di altre indagini sulla popolazione universitaria totale, apra più strade ai neolaureati.

Tabella 2. *Stima dei parametri e della bontà di adattamento del modello finale di regressione logistica per la soddisfazione per le condizioni economiche nell'attuale professione dei 270 laureati della Facoltà di Scienze Statistiche dell'Università di Milano-Bicocca dall'A.A. 1988/89 all'A.A. 2002/2003 e occupati nel novembre 2003.*

<u>Parametro</u>					<u>Odds ratio</u>		
	<i>Stima</i>	<i>Errore</i>	<i>Chi quadro</i>	<i>Pr > Chi quadro</i>	<i>Stima</i>	<i>Inf. (IC Wald 95%)</i>	<i>Sup. (IC Wald 95%)</i>
<i>Intercetta</i>	-1,24	0,61	4,18	0,040			
<i>Uso di SAS Spss</i>	0,92	0,43	4,45	0,035	2,50	1.07	5.85
<i>Tempo pieno</i>	1,38	0,59	5,45	0,029	3,98	1.25	12.72
<i>Stage</i>	0,90	0,42	4,65	0,030	2,46	1.09	5.58
<u>Bontà di adattamento del modello</u>							
<i>Test</i>	<i>Chi-q</i>	<i>g,d,l.</i>	<i>Pr > Chi-quadro</i>				
<i>Likelihood Ratio</i>	17,75	3	0,0005				
<i>Score</i>	17,55	3	0,0005				
<i>Wald</i>	14,80	3	0,0020				

La lettura di tale risultato richiede cautela, in ragione del ridotto numero di soggetti che hanno optato per tale opportunità formativa; questa considerazione induce anche a ripiegare su un livello di significatività per i diversi test pari ad $\alpha=0,10$, con valore orientativo. Occorre poi osservare come i cosiddetti Master di Primo Livello istituiti dalla Facoltà milanese nascano da bisogni specifici del territorio, quale

Tabella 3. *Stima dei parametri e della bontà di adattamento del modello finale di regressione logistica per la soddisfazione per le prospettive di carriera dei 270 laureati della Facoltà di Scienze Statistiche dell'Università di Milano-Bicocca dall'A.A. 1988/89 all'A.A. 2002/2003 e occupati nel novembre 2003.*

<u>Parametro</u>					<u>Odds ratio</u>		
	<i>Stima</i>	<i>Errore</i>	<i>Chi quadro</i>	<i>Pr > Chi quadro</i>	<i>Stima</i>	<i>Inf. (IC Wald 95%)</i>	<i>Sup. (IC Wald 95%)</i>
<i>Intercetta</i>	-2,12	0,47	20,16	0,0001			
<i>Uso di SAS/ Spss</i>	-0,91	0,58	2,49	0,1146	0.40	1.07	5.85
<i>Tempo pieno</i>	1,38	0,56	2,65	0,1033	2.50	1.25	12.73
<i>Stage</i>	1,68	0,83	4,13	0,0422	5.35	1.06	27.02
<u>Bontà di adattamento del modello</u>							
<i>Test</i>	<i>Chi-q</i>	<i>g,d,l.</i>	<i>Pr > Chi-quadro</i>				
<i>Likelihood Ratio</i>	7,47	3	0,0582				
<i>Score</i>	8,02	3	0,0456				
<i>Wald</i>	6,98	3	0,0723				

la ricerca biostatistica, e trovano pertanto una risposta favorevole in un mercato del lavoro che ne è almeno in parte fautore. Il segno negativo del parametro (ed il conseguente valore dell'*odds ratio* inferiore all'unità) associato all'impiego di *software* applicativi a più elevato contenuto specialistico crea perplessità interpretative: si potrebbe ritenere che ai neolaureati con maggiori competenze informatiche spettino condizioni economiche favorevoli, come indicato dal modello precedente, ma che siano assegnati sin dall'inizio a mansioni computazionali dalle quali difficilmente evolvono. Tale spiegazione, ad ogni modo, si attaglierebbe solo ad alcuni casi e la generalizzazione pare forzata.

4. L'incontro delle realtà universitaria ed aziendale: esplorare le relazioni

Nel paragrafo precedente sono emerse alcune difficoltà interpretative per il segno del legame fra soddisfazione per le prospettive di carriera e le competenze informatiche. Inoltre, nonostante l'analisi consideri inizialmente un congruo numero di variabili esplicative, tre solamente si sono rivelate significative in ciascuno dei modelli stimati. La ridotta numerosità dei soggetti in esame, oltre ad essere una possibile causa di questo risultato, inibisce anche la stratificazione rispetto a variabili d'interesse, stratificazione atta all'analisi differenziale della soddisfazione per diversi 'tipi' di neolaureato alle prime esperienze professionali. Sembra desiderabile quindi poter inserire la modellazione statistica classica in prospettive d'indagine suscettibili di evidenziare, preliminarmente alla stima del modello, i possibili legami fra variabili, che si suppongono articolati in considerazione anche della dinamicità degli ambiti coinvolti, in particolare quello aziendale, sottoposto a pressioni sovente non governabili dall'interno.

La limitatezza dell'impostazione cosiddetta classica, volta alla stima di un modello di regressione, emerge peraltro negli studi politici, i quali sottolineano come il concetto di 'complessità causale' imponga specificazioni statistiche appropriate (Breumoeller, 2003). Il modello logistico booleiano emerso in tale contesto si rivela adeguato anche alle indagini sugli studenti universitari, perché capace di evidenziare causalità non colte dal modello tradizionale; quest'ultimo, infatti, pone sovente il ricercatore di fronte all'esito di non significatività dell'interazione fra variabili esplicative, esito nettamente stridente con le teorie consolidate nel caso, fra gli altri, di causalazione congiunturale multipla (Porcu e Muggeo, 2005). In ambito informatico, lo sviluppo di metodologie indirizzate al cosiddetto 'apprendimento informativo' (*machine learning*) si rivolge allo studio delle relazioni fra le osservazioni rilevate, così da unificare la generazione di pattern derivanti dai dati di partenza e dalle relazioni

logiche. La validazione statistica delle relazioni emerse consente di integrare i modelli di regressione con i legami evidenziati da procedure informatiche automatizzate, come sopra accennato, nell'intento di costruire modelli più robusti partendo da fonti di dati complesse.

Nella transizione dall'Università al lavoro, la ricerca di relazioni fra i soggetti o 'entità' del processo, i neolaureati e le aziende, consiste nel trovare le coppie neolaureato-azienda, intense come relazioni fra le rispettive caratteristiche - o 'attributi' - che danno luogo a tali associazioni. Per generare i legami dai dati si ricorre all'analisi dei gruppi (*cluster analysis*), appropriata per l'identificazione di tipologie (Fabbris, 1997) e per derivare relazioni dai dati di base (Popescul e Hungar, 2004). Si è scelto, in questa sede, di sottoporre *ex post* ad analisi di raggruppamento i legami emersi nel precedente modello logistico relativo alla soddisfazione per le prospettive di carriera. Operando sulla base di dati iniziale 'unità statistica per variabili', si sono considerate, quali variabili di classificazione, la soddisfazione, il regime lavorativo a tempo parziale o pieno, il tipo di software in uso, lo svolgimento di stage e la formazione *post lauream*. Le osservazioni sono state classificate in quattro *cluster* distinti per accertare l'esistenza delle associazioni già emerse nel modello logistico, utilizzando la procedura *k-means*, una fra le tecniche adatte allo scopo (Popescul e Ungar, 2004). I risultati sono esprimibili come frequenza relativa 'attesa' di soggetti che manifestano la modalità 'successo' di una specifica variabile dicotomica in uno fra i cluster individuati; tali esiti consentono di approfondire alcuni fenomeni parsi anomali nella regressione logistica, in particolare l'impatto negativo che l'*expertise* statistico-informatica sembra esercitare sulla soddisfazione (Tab. 4). La *cluster analysis* evidenzia come la negatività del segno sia legata a due tipologie di neolaureati: un gruppo per il quale la buona conoscenza dei programmi statistici si associa a lauree triennali e a contratti a tempo determinato o atipici (*cluster 2*) ed un altro nel quale una bassa conoscenza di SAS e SPSS si lega ad un'elevata frequenza di contratti a tempo indeterminato (*cluster 4*).

Tabella 4. Composizione dei cluster, per variabile di classificazione, individuati dall'analisi di raggruppamento fra i 270 laureati della Facoltà di Scienze Statistiche dell'Università di Milano-Bicocca dall'A.A. 1988/89 all'A.A. 2002/2003 occupati nel novembre 2003 (valori relativi).

C*	Soddisfazione per le prospettive di carriera	Corsi di formazione post lauream	Laurea quadriennale	Laurea triennale	Tempo indeterminato	Conoscenza di SAS/SPS.
1	0,041	0,041	0,449	0,000	0,408	0,000
2	0,050	0,050	0,000	1,000	0,250	0,550
3	0,087	0,087	0,478	0,000	0,652	1,000
4	1,000	0,091	0,181	0,182	0,909	0,091

*C è l'acronimo di cluster o raggruppamento

Il segno negativo stimato nell'equazione (3) è pertanto interpretabile alla luce dell'evoluzione del mercato del lavoro negli anni più recenti, che vedono l'istituzione delle cosiddette lauree brevi e l'introduzione forme contrattuali lontane dal lungo termine, a fronte di maggiori richieste tecnico-professionali.

5. Considerazioni conclusive

La semplice verifica effettuata suggerisce indicazioni d'utilità dell'integrazione fra metodologia informatica e modellazione statistica, in particolare nella trattazione di basi di dati complesse, di dimensioni rilevanti e in continuo aggiornamento, quali quelle concernenti i neolaureati e secondo alcuni autori quest'approccio rappresenta una delle future sfide della statistica. La costruzione di basi di dati relazionali a partire dai dati originari e l'aggregazione delle unità statistiche mediante l'analisi di raggruppamento, nelle formulazioni proprie dell'analisi dei legami o *link analysis*, sono effettuate sulla base delle relazioni fra unità stesse. L'onerosità computazionale derivante trova compensazione in una lettura duttile ed espressiva di una realtà complessa e instabile, qual è il mercato del lavoro, con l'ulteriore vantaggio della possibilità di arricchire il procedimento con conoscenze teoriche a priori, come avviene nei modelli booleani.

L'integrazione fra *computer sciences* e tecniche statistiche, delle quali la Regressione Logistica Strutturale Generalizzata è una fra le molte espressioni, conosce un notevole sviluppo nell'indagine informatica negli anni recenti, con ambiti applicativi molteplici, da quello dell'analisi testuale a quello biologico, medico ed altri ancora. Queste ultime non possono che trarre un proficuo vantaggio dal vaglio critico e dall'elaborazione sistematica da parte della ricerca statistica, così fortemente implicata sul piano metodologico, in uno sforzo d'interdisciplinarietà.

Riferimenti bibliografici

- BRAUMOELLER B. F. (2003) Causal Complexity and the Study of Politics, *Political Annals*, **11**: 209-233.
- CRIPPA F., MARIANI P. Strategie di scelta e destinazioni nel mercato del lavoro. In D'OVIDIO F. (a cura di) *Professioni e competenze nel lavoro dei laureati*, CLEUP, Padova: 155-166.
- FABBRIS L. (1997) *Statistica Multivariata*, McGraw-Hill, Milano.

- POPESCU A., UNGAR L. H. (2003) Structural Relational Learning For Link Analysis. In *Workshop on Multi-Relational Data Mining at Knowledge Discovery and Data Mining (KDD)*.
- POPESCU A., UNGAR L. H. (2004) Cluster-based Concept Invention for Statistical Relational Learning. In *Proceedings of ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*: pp. 665-670
- POPESCU A., UNGAR L. H. (2004) Dynamic Feature Generation for Relational Learning. In *Proceedings of Multi-Relational Data Mining Workshop (MRDM 2004)*.
- POPESCU A., UNGAR L. H. (2004) Dynamic Feature Generation for Relational Learning. In *Proceedings of ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*.
- PORCU M., MUGGEO V. M. R. (2004) Factors that cause university students to drop out. An alternative modelling of interaction terms in logistic regression models. In: (Società Italiana di Statistica) *Atti della XLII Riunione della Società Italiana di Statistica, Università di Bari, 9-11 giugno 2004, Sessioni Spontanee*, Cleup, Padova: 511-514.

Some considerations on the university educational offering and graduates' satisfaction with their first work experiences

Summary. *In the transition from University to the job market, positive outcomes are due to different causes, that can be analyzed using logistic regression models. These, however, are not free of limitations (Porcu e Muggeo, 2005). Possible methodological refinements are, among others, Boolean Logistic Regression Models (Braumoeller, 2003) and Generalized Structural Regression Models for Link Analysis (Popescu e Ungar, 2004). These start by investigating the link structure between observations and continue by estimating the links by means of logistic regression. The usefulness of this type of approach is preliminarily evaluated on graduates' first work experience. Significant relationships in the logistic model are verified ex-post by means of cluster analysis. This assessment provides a dynamic understanding of the connection between statistical software and job satisfaction, while the traditional logistic model provides a biased reading.*

Keywords: *Logistic regression, Link analysis, Cluster analysis.*