

7.4 Procedura CHAID

Le variabili più importanti ai fini della nostra indagine sono quelle relative all'autonomia acquisita e alla situazione generale rispetto al momento dell'ingresso in Comunità- alloggio. Per approfondire le relazioni esistenti tra queste variabili e le altre variabili rilevate sugli ospiti, abbiamo utilizzato la procedura *CHAID* (*Chi-square Automatic Interaction Detection*) contenuta nel programma *SPSS Answer Tree* Versione 2.0.1.

Partendo dalla variabile obiettivo del modello (ad esempio la situazione generale) l'algoritmo di segmentazione della procedura specifica, tra le variabili esplicative, quella che meglio la predice (ad esempio la provenienza da casa) e a cascata le altre collegate con questa, riducendo via via i dati disponibili; il diagramma ad albero che viene creato identifica facilmente i gruppi significativi.

Answer Tree offre quattro diversi metodi di segmentazione, che si adattano a diversi tipi di dati.

- **CHAID**: un algoritmo statistico a più vie per gestire variabili di qualsiasi tipo (quelle continue vengono ripartite automaticamente nel numero di categorie desiderato, prima di iniziare lo studio) che identifica i segmenti nello spazio dei dati suddividendo iterativamente i dati in due o più nodi, partendo da tutti i dati disponibili.

- CHAID esaustivo: per analisi in profondità che fonde iterativamente coppie simili di dati, fino a quando non rimane una sola coppia. L'insieme di categorie con la significatività più alta viene ritenuto il miglior criterio di segmentazione per quella variabile esplicativa.
- C&RT: (Classification and Regression Trees) un algoritmo binario per creare gruppi omogenei. L'albero viene creato suddividendo iterativamente i dati in due nodi, usando tutte le variabili esplicative; il miglior predittore ad ogni livello viene scelto in base a diverse misure di impurità o di dissimilarità. L'obiettivo è la produzione di sottoinsiemi il più possibile omogenei rispetto alla variabile obiettivo; l'albero che si ottiene è binario.
- QUEST: (Quick Unbiased Efficient Statistical Tree) utile per selezionare variabili senza distorsioni; produce un albero binario sul quale sono possibili operazioni quali potatura, regole di interruzione esplicite e suddivisioni sostitutive. Mentre CHAID e C&RT trattano la selezione delle variabili e la selezione dei punti di divisione simultaneamente, QUEST le affronta separatamente.

Abbiamo scelto la procedura CHAID, tra quelle disponibili, poiché meglio si adatta agli obiettivi della nostra indagine e alla natura delle variabili considerate.

L'output ottenuto con tale procedura è un diagramma ad albero che fornisce un'immagine sintetica e intuitiva dei raggruppamenti e delle relazioni tra le variabili esplicative;

dall'alto verso il basso, ogni livello rappresenta il successivo miglior criterio di previsione della variabile obiettivo. Ogni singolo nodo rappresenta un segmento omogeneo dei dati, rendendo facile l'interpretazione dei risultati.

Vediamo ora nel dettaglio il tipo di algoritmo utilizzato dalla procedura CHAID: partendo dall'intera matrice dei dati, la popolazione viene divisa in due o più gruppi in base alla migliore previsione della variabile oggetto di studio. Per determinare la migliore divisione in corrispondenza di ogni nodo, se non c'è differenza statisticamente significativa tra le modalità della variabile predittiva rispetto alla variabile obiettivo, l'algoritmo fonde due o più modalità insieme; il processo viene ripetuto fino a quando viene trovata una coppia non significativa e per tutte le variabili predittive. Il set di modalità delle variabili predittive che viene selezionato rappresenta la migliore segmentazione che rispecchia la variabile obiettivo considerata (vengono scelte le modalità con la maggiore significatività per rappresentare la migliore predizione della variabile obiettivo).

Con la popolazione così suddivisa, viene ripetuto il procedimento per ognuno dei nodi fino a quando non interviene uno dei motivi di arresto e la procedura si ferma.

Le componenti principali dell'analisi CHAID sono:

- una o più variabili predittive che possono essere continue, ordinali o nominali;
- una variabile obiettivo che può essere anch'essa continua, ordinale o nominale (a seconda del tipo di variabile la

procedura utilizza un diverso metodo di segmentazione per la costruzione dell'albero);

- un insieme di parametri che includono i livelli di significatività utilizzati per fondere le modalità e segmentare la popolazione, i criteri di arresto della procedura, gli eventuali pesi dei casi.

L'algoritmo del CHAID si arresta quando:

- viene raggiunta la profondità massima dell'albero fissata a priori;
- non ci sono più variabili predittive statisticamente significative capaci di suddividere ulteriormente il nodo;
- il numero dei casi all'interno dei singoli livelli è inferiore a quello prefissato a priori.

Analizziamo ora i risultati ottenuti per le variabili relative all'autonomia (*Spostamenti esterni, Cura personale, Attività domestiche, Uso del denaro e Autonomia generale*) e per la *Situazione generale*.

Per quanto riguarda gli *Spostamenti esterni* abbiamo che la variabile esplicativa che meglio li predice è il *Livello di istruzione* (vedi grafico 7.2): chi ha un'istruzione media è più autonomo negli spostamenti (la percentuale di chi si sa muovere da solo passa da 36.59 a 50.49) e, successivamente, è il lavoro a rappresentare un incremento di autonomia (88.89 %). Il *Gruppo di appartenenza* ai diversi tipi di Comunità- alloggio emersi dalla Cluster Analysis sulle strutture e il *Tipo di handicap*

intervengono al terzo livello (sono meno autonomi i disabili fisici e coloro che provengono dalle strutture “Istituzionali”).

Nella *Cura personale* interviene innanzi tutto il fatto di lavorare che determina un incremento in chi si gestisce da solo dal 37.56 % al 68.29 %; la seconda variabile esplicativa che meglio predice l'autonomia nella cura personale è, sia per chi lavora sia per chi non lo fa, il *Livello di istruzione* e, successivamente, per chi non lavora e ha un'istruzione bassa, il *Motivo di ingresso*: chi entra in Comunità- alloggio perché lo richiede la famiglia subisce un decremento di autonomia dal 21.13 % al 7.14 % (vedi grafico 7.3).

Anche per quanto riguarda le *Attività domestiche* (grafico 7.4) abbiamo al primo posto il lavoro ma, mentre per chi lavora il secondo fattore di incremento di autonomia è rappresentato dal *Livello di istruzione*, per chi non lavora conta successivamente il *Tipo di handicap* (chi ha un handicap mentale compie attività domestiche da solo nel 37.50 % dei casi, mentre per chi ha un handicap fisico o fisico- mentale la percentuale è del 19.44).

Nella conoscenza dell'uso del denaro abbiamo che interviene maggiormente il Gruppo di appartenenza: sono più indipendenti coloro che provengono dalle strutture “Rigide” e “Autogestite”; tra coloro che appartengono alle Comunità- alloggio “Istituzionali” interviene successivamente la frequenza di un Corso di formazione che incrementa la conoscenza dell'uso del denaro dal 37.32 % al 94.12 %. Per chi appartiene agli altri tipi di strutture interviene, in un secondo momento, la frequenza di un

Centro Socio Educativo che fa decrescere la percentuale di chi conosce il denaro da 76.19 a 50.00. Possiamo notare che nel terzo livello interviene anche il *Sesso* ove i maschi sono più autonomi rispetto alle femmine (grafico 7.5).

Per l'Autonomia generale (che considera tutte le variabili appena considerate) interviene innanzi tutto il fatto di svolgere un lavoro che porta la media del punteggio da 4.31 a 5.63; per chi non lavora, la frequenza di un Corso di formazione incrementa l'autonomia media da 3.98 a 5.07 e, successivamente, i maschi subiscono un ulteriore aumento a 6.23 mentre le femmine diminuiscono a 4.00.

Per la *Situazione generale* abbiamo che la variabile esplicativa che meglio la predice è la *Provenienza da casa* (vedi grafico 7.7); per chi proviene dal domicilio o da un'altra Comunità- alloggio si ha un lieve decremento nei miglioramenti da 83.90 % a 79.59; chi proviene da istituti, ospedali o dalla strada subisce, invece, un miglioramento passando dall'83.90 al 94.83 %. Qualunque sia la provenienza degli utenti, la seconda variabile che meglio predice i miglioramenti nella situazione generale è la presenza o meno di attività compiute (lavoro, scuola CSE, CFP): chi non fa nulla migliora meno in entrambi i casi (da 94.83 % a 87.50 % per chi proviene da casa e da 79.59 % a 68.89 per chi proviene dagli altri luoghi).

Anche nel terzo livello di segmentazione, per chi non compie nessun tipo di attività, interviene la stessa variabile *Tipo di handicap* ma con diverse aggregazioni: nel ramo generato da

coloro che non provengono da casa migliorano di più coloro che hanno disabilità mentali o fisiche (passando da 87.50 % a 91.30); nel ramo generato da coloro che provengono da ambienti familiari migliorano meno coloro che hanno disabilità fisiche.

Alla luce dei risultati emersi possiamo sottolineare ancora una volta l'importanza del lavoro nella vita delle persone disabili (così come nella vita di tutti): abbiamo visto che chi lavora presenta sempre livelli maggiori di autonomia e chi non compie nessuna attività migliora meno degli altri; l'istruzione è un altro fattore importante nell'acquisizione dell'autonomia.

Possiamo quindi individuare nel potenziamento dei livelli di istruzione raggiungibili dai disabili e nell'aumento sempre più cospicuo di persone con handicap inserite nel mondo del lavoro i principali fattori necessari a determinare un incremento di capacità che renda la loro vita sempre più indipendente dall'aiuto degli altri e sempre migliore. Più in generale, ci sembra importante che le persone con handicap abbiano sempre di più la possibilità di accedere e di conquistare livelli di autonomia che li rendano competitivi nel mondo del lavoro e pienamente inseriti nel mondo sociale di cui fanno parte.

Analisi degli ospiti

Analisi degli ospiti