

# IBM SPSS Conjoint 20



*Nota:* Prima di utilizzare queste informazioni e il relativo prodotto, leggere le informazioni generali disponibili in Note legali a pag. 41.

Questa versione si applica a IBM® SPSS® Statistics 20 e a tutte le successive versioni e modifiche fino a eventuali disposizioni contrarie indicate in nuove versioni.

Le schermate dei prodotti Adobe sono state ristampate su autorizzazione di Adobe Systems Incorporated.

Le schermate dei prodotti Microsoft sono state ristampate su autorizzazione di Microsoft Corporation.

Materiali concessi in licenza - Proprietà di IBM

© **Copyright IBM Corporation 1989, 2011.**

Tutti i diritti riservati.

---

# Prefazione

IBM® SPSS® Statistics è un sistema completo per l'analisi dei dati. Il modulo aggiuntivo opzionale Conjoint include le tecniche di analisi aggiuntive descritte nel presente manuale. Il modulo aggiuntivo Conjoint deve essere usato con il modulo Core SPSS Statistics in cui è completamente integrato.

## **Informazioni su Business Analytics di IBM**

Il software IBM Business Analytics fornisce informazioni complete, coerenti e accurate a cui i responsabili delle decisioni possono affidarsi per ottimizzare le prestazioni dell'azienda. Un ampio portafoglio di applicazioni di [business intelligence](#), [analisi predittiva](#), [gestione delle prestazioni e delle strategie finanziarie](#) e [analisi](#) offre una panoramica chiara, istantanea e interattiva delle prestazioni attuali e la possibilità di prevedere i risultati futuri. Utilizzato in combinazione con potenti soluzioni di settore, prassi consolidate e servizi professionali, questo software consente alle aziende di tutte le dimensioni di ottimizzare la produttività, automatizzare le decisioni senza problemi e fornire risultati migliori.

Come parte di questo portafoglio, il software IBM SPSS Predictive Analytics consente alle aziende di prevedere gli eventi futuri e di agire tempestivamente in modo da migliorare i risultati delle attività aziendali. Le aziende, gli enti governativi e le università di tutto il mondo si affidano alla tecnologia IBM SPSS perché rappresenta un vantaggio concorrenziale in termini di attrazione, retention e aumento dei clienti, riducendo al tempo stesso le frodi e limitando i rischi. Incorporando il software IBM SPSS nelle attività quotidiane, le aziende diventano imprese in grado di effettuare previsioni e di gestire e automatizzare le decisioni, per raggiungere gli obiettivi aziendali e vantaggi tangibili sulla concorrenza. Per ulteriori informazioni o per contattare un rappresentante, visitare il sito <http://www.ibm.com/spss>.

## **Supporto tecnico**

Ai clienti che richiedono la manutenzione, viene messo a disposizione un servizio di supporto tecnico. I clienti possono contattare il supporto tecnico per richiedere assistenza per l'utilizzo dei prodotti IBM Corp. o per l'installazione di uno degli ambienti hardware supportati. Per contattare il supporto tecnico, visitare il sito Web di IBM Corp. all'indirizzo <http://www.ibm.com/support>. Ricordare che durante la richiesta di assistenza sarà necessario fornire i dati di identificazione personali, i dati relativi alla propria società e il numero del contratto di manutenzione.

## **Supporto tecnico per studenti**

Gli studenti che utilizzano una versione accademica o grad pack di qualsiasi prodotto software IBM SPSS sono pregati di utilizzare le apposite pagine online per studenti [Solutions for Education](#) (<http://www.ibm.com/spss/rd/students/>). Gli studenti che utilizzano una copia del software IBM SPSS fornita dall'università, sono pregati di contattare il coordinatore del prodotto IBM SPSS presso l'università.

## ***Servizio clienti***

Per eventuali chiarimenti in merito alla spedizione o al proprio conto, rivolgersi alla sede locale. Tenere presente che sarà necessario fornire il numero di serie.

## ***Corsi di formazione***

IBM Corp. organizza corsi di formazione pubblici e onsite che includono esercitazioni pratiche. Tali corsi si terranno periodicamente nelle principali città. Per ulteriori informazioni su questi seminari, andare a <http://www.ibm.com/software/analytics/spss/training>.

## ***Pubblicazioni aggiuntive***

I documenti *SPSS Statistics: Guide to Data Analysis*, *SPSS Statistics: Statistical Procedures Companion* e *SPSS Statistics: Advanced Statistical Procedures Companion*, scritti da Marija Norušis e pubblicati da Prentice Hall sono disponibili come materiale supplementare consigliato. Queste pubblicazioni descrivono le procedure statistiche nei moduli SPSS Statistics Base, Advanced Statistics e Regression. Utili sia come guida iniziale all'analisi dei dati che per applicazioni avanzate, questi manuali consentono di ottimizzare l'utilizzo delle funzionalità presenti nell'offerta IBM® SPSS® Statistics. Per ulteriori informazioni, inclusi contenuti delle pubblicazioni e capitoli di esempio, visitare il sito Web dell'autrice: <http://www.norusis.com>

---

# Contenuto

## **1 Introduzione all'analisi Conjoint 1**

L'approccio con profilo completo .....	2
Array ortogonale .....	2
Stimoli sperimentali .....	2
Raccolta e analisi dei dati .....	3

## **Parte I: Manuale dell'utente**

## **2 Generazione di un disegno ortogonale 5**

Definizione di valori per un disegno ortogonale .....	7
Opzioni della finestra di dialogo Disegno ortogonale .....	7
Funzioni aggiuntive del comando ORTHOPLAN .....	8

## **3 Visualizzazione di un disegno 9**

Visualizza disegno: Titoli .....	10
Opzioni aggiuntive del comando PLANCARDS .....	10

## **4 Esecuzione di un'analisi Conjoint 11**

Requisiti .....	11
Specifica del file del piano e del file di dati .....	12
Specifica della modalità di registrazione dei dati .....	12
Sottocomandi facoltativi .....	13

## **Parte II: Esempi**

## **5 Utilizzo dell'analisi Conjoint per creare un modello di preferenze per battitappeto 17**

Generazione di un disegno ortogonale .....	17
--	----

Creazione di stimoli sperimentali: Visualizzazione del disegno . . . . .	21
Esecuzione dell'analisi . . . . .	24
Punteggi di utilità . . . . .	26
Coefficienti . . . . .	27
Importanza relativa . . . . .	27
Correlazioni . . . . .	28
Inversioni . . . . .	28
Esecuzione delle simulazioni. . . . .	29
Probabilità di preferenza delle simulazioni . . . . .	30

## ***Appendici***

<b><i>A File di esempio</i></b>	<b>31</b>
---------------------------------	-----------

<b><i>B Note legali</i></b>	<b>41</b>
-----------------------------	-----------

<b><i>Bibliografia</i></b>	<b>44</b>
----------------------------	-----------

<b><i>Indice</i></b>	<b>45</b>
----------------------	-----------

# Introduzione all'analisi Conjoint

L'analisi Conjoint è uno strumento di ricerca marketing utile per sviluppare disegni di prodotto efficaci. Questo tipo di analisi permette di rispondere a domande quali: Quali sono gli attributi dei prodotti che interessano di più ai consumatori? Quali livelli di attributi prodotto sono più o meno desiderabili per i consumatori? Qual è la quota di mercato di preferenza dei prodotti ' leader rispetto a quella del prodotto esistente o proposto?

L'analisi Conjoint permette di chiedere agli intervistati di fare delle scelte, al pari di quello che fa ciascun consumatore — quando decide di acquistare un prodotto anziché un altro.

Si supponga, ad esempio, di dover acquistare un biglietto aereo. Si può scegliere tra un posto spazioso e ristretto. Se questa fosse l'unica considerazione non ci sarebbero dubbi, perché la maggior parte di noi preferirebbe sicuramente un posto spazioso. Oppure supponiamo di dover scegliere tra due prezzi diversi: \$225 o \$800. Se dovessimo tener conto solo del prezzo, probabilmente opteremo per quello più basso. Adesso supponiamo di dover scegliere tra un volo diretto che dura 2 ore e un volo con sosta che dura cinque ore. La maggior parte delle persone sceglierebbe sicuramente il volo diretto.

Lo svantaggio di questo approccio deriva dal fatto che le alternative di scelta possono essere riferite ad un solo attributo alla volta. L'analisi Conjoint presenta opzioni di scelta tra prodotti definiti da insiemi di attributi. Questo concetto è illustrato da quanto segue: preferireste un posto ristretto, al prezzo di \$225 con una sosta oppure uno spazio comodo, al prezzo di \$800 e con volo diretto? Se la comodità, il prezzo e la durata sono attributi importanti, le alternative di prodotto sono potenzialmente otto:

Prodotto	Comodità	Prezzo	Durata
1	Spazio ristretto	\$225	2 ore
2	Spazio ristretto	\$225	5 ore
3	Spazio ristretto	\$800	2 ore
4	Spazio ristretto	\$800	5 ore
5	Posto spazioso	\$225	2 ore
6	Posto spazioso	\$225	5 ore
7	Posto spazioso	\$800	2 ore
8	Posto spazioso	\$800	5 ore

Se queste sono le alternative, il prodotto 4 è probabilmente quello preferito mentre il prodotto 5 quello meno appetibile. Le preferenze degli utenti relativamente ai prodotti in offerta sono implicitamente determinata da ciò che loro considerano importante.

L'analisi Conjoint permette di determinare l'importanza relativa di ciascun attributo nonché quali sono i livelli preferiti di ciascun attributo. Se il prodotto preferito non è disponibile per qualche motivo, ad esempio perché ha un costo elevato, è possibile sapere subito quale sarà la seconda alternativa preferita. Se si dispone di altre informazioni sugli intervistati, come dati demografici, è possibile anche identificare segmenti di mercato a cui destinare i singoli prodotti. Ad esempio una persona che viaggia per lavoro e uno studente possono avere preferenze diverse che possono essere soddisfatte con tipologie di prodotti diverse.

## ***L'approccio con profilo completo***

Conjoint utilizza il **profilo completo** (o concetto completo) **in modo da poter distribuire**, ordinare o assegnare una priorità a un insieme di profili o schede, in base alle preferenze degli intervistati stessi. Ciascun profilo descrive un prodotto o un servizio completo ed è costituito da più combinazioni di livelli di fattore per tutti i fattori (attributi) di interesse.

## ***Array ortogonale***

Uno dei possibili problemi che derivano dall'approccio con profilo completo diventa subito evidente se ci sono più fattori coinvolti e se ciascun fattore ha più di una coppia di livelli. Il numero totale di profili risultanti da tutte le possibili combinazioni di livelli diventa eccessivo e tale da impedire alle persone intervistate di classificare i livelli o di assegnare un punteggio significativo. Per ovviare a questo problema, l'approccio con profilo completo utilizza il cosiddetto **disegno fattoriale frazionario** che presenta una frazione idonea di tutte le possibili combinazioni di livelli di fattori. L'insieme risultante, chiamato **matrice ortogonale** ha lo scopo di identificare gli effetti principali per ciascun livello di fattore. Le interazioni tra un fattore e i livelli di un altro fattore sono considerate trascurabili.

La procedura Genera disegno ortogonale permette di generare una matrice ortogonale, che costituisce generalmente la base di partenza per l'analisi Conjoint. Questa procedura permette anche di generare combinazioni di fattori di livello o **casi di controllo**, che vengono valutati dagli intervistati ma non vengono usati per creare il modello di preferenze. Questi casi vengono infatti solo usati per verificare la validità del modello.

## ***Stimoli sperimentali***

Ciascun insieme di livelli di fattori di un disegno ortogonale rappresenta una versione diversa del prodotto esaminato e devono essere presentati agli interessati come profili di prodotto singoli. Questo approccio fa sì che gli intervistati si concentrino solo sul prodotto in esame. Gli stimoli devono essere standardizzati, verificando che i profili siano fisicamente simili, fatta eccezione per la combinazione di caratteristiche.

La procedura Visualizza disegno semplifica la creazione di profili prodotto. Questa procedura utilizza i disegni creati dalla procedura Genera disegno ortogonale o immessi dall'utente per creare un insieme di profili prodotto che possano essere immediatamente utilizzati.

## **Raccolta e analisi dei dati**

Considerata le sostanziali differenze tra le preferenze dei singoli soggetti, gran parte dell'analisi Conjoint prende principalmente in esame i singoli soggetti. Per ottenere risultati generalizzati, è necessario selezionare un campione casuale di soggetti della popolazione di destinazione ed esaminare i risultati del gruppo in questione.

La dimensione dei campioni utilizzati nelle analisi Conjoint variano considerevolmente. In un report (Cattin e Wittink, 1982), gli autori affermano che le dimensioni del campione in studi Conjoint commerciali sono generalmente comprese tra 100 e 1.000, con un intervallo tipico compreso tra 300 e 550. In un altro studio (Akaah e Korgaonkar, 1988) è stato invece rilevato che sono più comuni campioni di dimensioni minori (inferiori a 100). Come in altri casi, la dimensione del campione deve sempre essere tale da poter essere ritenuto affidabile.

Dopo aver selezionato il campione, il ricercatore deve inviare l'insieme di profili o schede a ciascun intervistato. La procedura Conjoint offre tre metodi per la registrazione dei dati. Nel primo metodo viene chiesto agli intervistati di assegnare un punteggio di preferenza a ciascun profilo. Il primo metodo viene generalmente utilizzato quando si impiega una scala di Likert o quando viene chiesto ai soggetti di indicare la propria preferenza con un numero compreso tra 1 e 100. Nel secondo metodo viene chiesto ai soggetti di assegnare un ordine a ciascun profilo compreso tra 1 e il numero totale di profili. Nel terzo metodo si chiede ai soggetti di ordinare i profili in termini di preferenze. Quest'ultimo metodo permette al ricercatore di registrare i numeri di profili nell'ordine dato da ciascun soggetto.

L'analisi dei dati viene effettuata con la procedura Conjoint (disponibile tramite la sintassi di comando) e i risultati vengono registrati in un punteggio di utilità chiamato **valore delle parti** per ciascun livello di fattore. Questi punteggi, analoghi ai coefficienti di regressione, offrono una misura quantitativa delle preferenze per ciascun livello di fattore. I valori maggiori corrispondono a preferenze più marcate. I valori delle parti sono espressi come unità comune e possono quindi essere sommati per ottenere l'utilità complessiva o preferenza complessiva per qualsiasi combinazione di livelli di fattore. I valori delle parti permettono di stimare la preferenza per qualsiasi profilo di prodotto, compresi i cosiddetti **casì di simulazione** non inclusi nell'esperimento.

Le informazioni risultanti dall'analisi Conjoint possono essere applicate a una vasta gamma di domande relative a studi di mercato. Ad esempio possono essere usate per analizzare caratteristiche progettuali di prodotti, quote di mercato, strategie pubblicitarie, costi e vantaggi e segmentazioni di mercato.

Benché questo manuale sia dedicato ad applicazioni destinate alle ricerche di mercato, va rilevato che l'analisi Conjoint può essere utilizzata in qualsiasi settore scientifico o commerciale in cui è necessario misurare le percezioni o i giudizi delle persone.

***Parte I:***  
***Manuale dell'utente***

## ***Generazione di un disegno ortogonale***

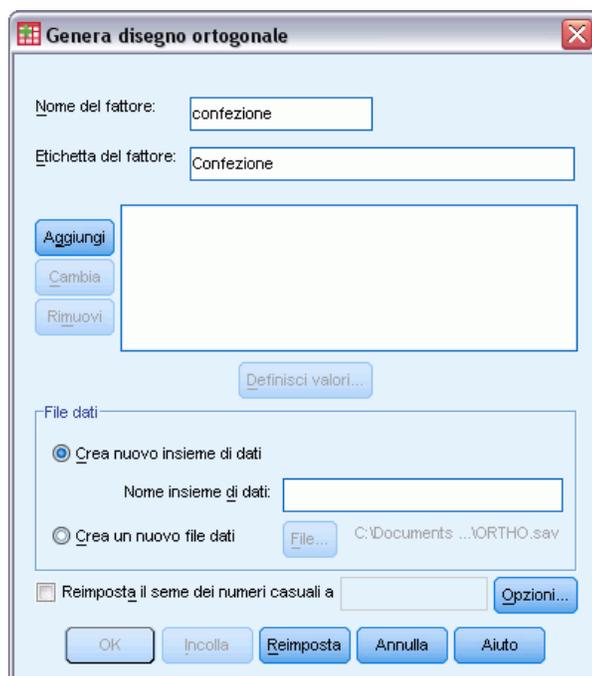
La procedura Genera disegno ortogonale consente di creare un file di dati contenente un disegno ortogonale con effetti principali, grazie al quale è possibile eseguire test statistici di numerosi fattori senza che sia necessario verificare tutte le combinazioni dei livelli dei fattori. È possibile visualizzare il disegno mediante la procedura Visualizza disegno e utilizzare il file di dati in altre procedure, ad esempio l'analisi Conjoint.

**Esempio.** Una compagnia a basso costo desidera individuare l'importanza che i potenziali clienti attribuiscono ai vari fattori che incidono sull'offerta. Benché sia evidente che il prezzo è il fattore principale, la compagnia area è interessata anche ad acquisire informazioni su altri fattori quali lo spazio del posto a sedere, il numero di soste e la disponibilità di un servizio ristorazione a bordo. Chiedere a un campione di potenziali clienti di ordinare i profili di prodotto che rappresentano tutte le possibili combinazioni di fattori è praticamente impossibile visto l'elevato numero di profili. La procedura Genera disegno ortogonale permette di creare un insieme di profili prodotto abbastanza ridotto da poter essere incluso in una ricerca di mercato, ma abbastanza grande da consentire la valutazione di ciascun fattore.

### ***Per generare un disegno ortogonale***

- Dai menu, scegliere:  
Dati > Disegno ortogonale > Genera...

Figura 2-1  
*Finestra di dialogo Genera disegno ortogonale*



- ▶ Definire almeno un fattore. Immettere un nome nella casella di testo Nome fattore. I nomi dei fattori possono essere nomi di variabili validi, ad eccezione di *status\_* o *card\_*. È inoltre possibile assegnare un'etichetta di fattore opzionale.
- ▶ Per aggiungere il nome del fattore e un'etichetta facoltativa, fare clic su Aggiungi. Per eliminare un fattore, selezionarlo nell'elenco e fare clic su Rimuovi. Per modificare il nome o l'etichetta di un fattore, selezionarlo nell'elenco, modificare il nome o l'etichetta, quindi fare clic su Modifica.
- ▶ Per definire i valori di ogni fattore, selezionare il fattore desiderato e fare clic su Definisci valori.

**File dati.** Permette di controllare la destinazione del disegno ortogonale. È possibile salvare il disegno in un nuovo insieme di dati nella sessione corrente oppure in un file di dati esterno.

- **Crea nuovo insieme di dati.** Crea un nuovo insieme di dati nella sessione corrente contenente i fattori e i casi generati dal piano.
- **Crea nuovo file di dati.** Crea un nuovo file di dati esterno contenente i fattori e i casi generati dal piano. Per impostazione predefinita, questo file viene denominato *ortho.sav* e salvato nella directory corrente. Fare clic su File per specificare un nome e una destinazione diversi per il file.

**Reimposta il seme dei numeri casuali a.** Consente di reimpostare il seme dei numeri casuali al valore specificato. Il seme può essere qualsiasi valore intero compreso tra 0 e 2.000.000.000. All'interno di una sessione, ogni volta che si genera un insieme di numeri casuali viene utilizzato un seme diverso e si ottengono pertanto risultati diversi. Se si desidera duplicare gli stessi numeri casuali, è consigliabile impostare il valore del seme prima di generare il primo disegno e quindi reimpostare lo stesso valore per il seme a ogni successiva generazione del disegno.

Se lo si desidera, è possibile:

- Fare clic su Opzioni per specificare il numero minimo di casi del disegno ortogonale e per selezionare i casi di controllo.

## Definizione di valori per un disegno ortogonale

Figura 2-2  
Finestra di dialogo Genera disegno ortogonale

Valore	Etichetta
1: 1	A*
2: 2	B*
3: 3	C*
4:	
5:	
6:	
7:	
8:	
9:	

È necessario assegnare i valori a ogni livello del fattore o dei fattori selezionati. Il nome del fattore verrà visualizzato dopo Valori ed etichette per.

Immettere ciascun valore del fattore. È possibile assegnare ai valori etichette descrittive. Se non si assegnano etichette ai valori, vengono assegnate automaticamente etichette corrispondenti ai valori stessi, ad esempio l'etichetta 1 per il valore 1, l'etichetta 3 per il valore 3 e così via.

**Riempi.** Consente di compilare automaticamente le caselle Valore con valori consecutivi che iniziano con 1. Immettere il valore massimo, quindi fare clic su Riempi per inserire i valori.

## Opzioni della finestra di dialogo Disegno ortogonale

Figura 2-3  
Finestra di dialogo Opzioni di Genera disegno ortogonale

Numero minimo di casi da generare: 18

Casi di controllo

Numero casi di controllo: 4

Combina casualmente con altri casi

**Numero minimo di casi da generare.** Specifica un numero minimo di casi per il piano. Selezionare un intero positivo inferiore o uguale al numero totale dei casi che possono essere creati in base a tutte le combinazioni possibili dei livelli dei fattori. Se non si specifica in modo esplicito il numero minimo dei casi da generare, viene generato il numero minimo di casi necessari per il disegno ortogonale. Se la procedura per il disegno ortogonale non è in grado di generare almeno il numero minimo di profili richiesto, verrà generato il numero maggiore di profili possibile più appropriato per i fattori e i livelli specificati. Si noti che il disegno non include necessariamente il numero esatto di casi specificati, ma include piuttosto il numero minimo di casi possibile nel disegno ortogonale utilizzando il valore come minimo.

**Casi di controllo.** È possibile definire casi di controllo classificati in base ai soggetti, ma non inclusi nell'analisi Conjoint.

- **Numero di casi di controllo.** Consente di creare i casi di controllo in aggiunta ai casi del piano normale. I casi di controllo vengono stimati in base ai soggetti, ma non vengono utilizzati durante la stima delle utilità eseguita dalla procedura Analisi Conjoint. È possibile specificare qualsiasi intero positivo inferiore o uguale al numero totale di casi che possono essere creati in base a tutte le combinazioni possibili dei livelli dei fattori. I casi di controllo vengono generati da un altro piano casuale e non dal piano sperimentale per gli effetti principali. I casi di controllo non sono duplicati dei profili sperimentali né degli altri casi di controllo. Per impostazione predefinita, non vengono generati casi di controllo.
- **Combina casualmente con altri casi.** Consente di combinare casualmente i casi di controllo con i casi sperimentali. Se questa opzione è deselezionata, i casi di controllo vengono visualizzati separatamente dopo i casi sperimentali.

## ***Funzioni aggiuntive del comando ORTHOPLAN***

Il linguaggio della sintassi dei comandi consente inoltre di:

- Aggiungere il disegno ortogonale al file di dati attivo anziché crearne uno nuovo.
- Specificare casi di simulazione prima di generare il disegno ortogonale anziché dopo la creazione dello stesso.

Per informazioni dettagliate sulla sintassi, vedere *Command Syntax Reference*.

## Visualizzazione di un disegno

La procedura Visualizza disegno consente di stampare un disegno sperimentale nel formato di una bozza di lista o come profili da sottoporre ai soggetti in uno studio congiunto. È inoltre possibile visualizzare i disegni creati con la procedura Genera disegno ortogonale o qualsiasi disegno visualizzato in un file dati attivo.

### Per visualizzare un disegno ortogonale

- Dai menu, scegliere:  
Dati > Disegno ortogonale > Visualizza...

Figura 3-1  
Finestra di dialogo Proprietà visualizzazione disegno



- Spostare uno o più fattori nell'elenco Fattori.
- Selezionare un formato per la visualizzazione dei profili nell'output.

**Formato.** È possibile scegliere una o più delle seguenti opzioni relative al formato:

- **Lista per lo sperimentatore.** Consente di visualizzare il disegno in un formato di bozza nel quale i profili di controllo sono diversi dai profili sperimentali e di elencare i profili di simulazione separatamente dopo quelli sperimentali e di controllo.
- **Profili per i soggetti.** Consente la creazione di profili da presentare ai soggetti. Questo formato non prevede la differenziazione dei profili di controllo né la creazione di profili di simulazione.

Se lo si desidera, è possibile:

- Fare clic su Titoli per definire le intestazioni e i piè di pagina dei profili.

## Visualizza disegno: Titoli

Figura 3-2  
Finestra di dialogo Titoli visualizzazione disegno



**Titolo.** Specificare un titolo di profilo lungo al massimo 80 caratteri. I titoli vengono visualizzati nella parte iniziale dell'output se è stata selezionata l'opzione Lista per lo sperimentatore e nella parte superiore di ogni nuovo profilo se è stata selezionata l'opzione Profili per i soggetti nella finestra di dialogo principale. Per Profili per i soggetti se la sequenza di caratteri speciali )CARD è specificata in qualunque punto del titolo, la procedura la sostituisce con il numero di profilo sequenziale. Questa sequenza di caratteri non viene convertita per Lista per lo sperimentatore.

**Piè di pagina del profilo.** Specificare un piè di pagina di profilo lungo al massimo 80 caratteri. I piè di pagina vengono visualizzati nella parte finale dell'output se è stata selezionata l'opzione Lista per lo sperimentatore e nella parte inferiore di ogni profilo se è stata selezionata l'opzione Profili per i soggetti nella finestra di dialogo principale. Per Profili per i soggetti se la sequenza di caratteri speciali )CARD è specificata in qualunque punto del piè di pagina, la procedura la sostituisce con il numero di profilo sequenziale. Questa sequenza di caratteri non viene convertita per Lista per lo sperimentatore.

## Opzioni aggiuntive del comando PLANCARDS

Il linguaggio della sintassi dei comandi consente inoltre di:

- Scrivere profili per soggetti in un file esterno (con il sottocomando `OUTFILE`).

Per informazioni dettagliate sulla sintassi, vedere *Command Syntax Reference*.

## ***Esecuzione di un'analisi Conjoint***

Per la procedura Conjoint non è ancora disponibile un'interfaccia utente grafica. Per eseguire un'analisi Conjoint, in una finestra di sintassi è necessario immettere una sintassi dei comandi relativa a un comando `CONJOINT` e quindi eseguirlo.

- Per un esempio di sintassi dei comandi relativa a un comando `CONJOINT` nel contesto di un'analisi Conjoint completa, incluse generazione e visualizzazione di un disegno ortogonale, vedere [il capitolo 5](#).
- Per informazioni dettagliate sulla sintassi dei comandi, per il comando `CONJOINT`, vedere *Command Syntax Reference*.

### ***Per eseguire un comando da una finestra di sintassi***

Dai menu, scegliere:

File > Nuovo > Sintassi...

Viene aperta una finestra di sintassi.

- ▶ Immettere la sintassi dei comandi relativa al comando `CONJOINT`.
- ▶ Evidenziare il comando nella finestra di sintassi e fare clic sul pulsante Esegui (il triangolo rivolto verso destra) nella barra degli strumenti dell'Editor della sintassi.

Per ulteriori informazioni sull'esecuzione dei comandi nella finestra di sintassi, vedere il *Manuale dell'utente del modulo Core*.

## ***Requisiti***

La procedura Conjoint richiede due file—un file di dati e un file del piano—e la specifica relativa alla modalità di registrazione dei dati (ad esempio, ogni punto di dati rappresenta un punteggio di preferenza compreso tra 1 e 100). Il file del piano è formato da un insieme di profili prodotto da classificare secondo i soggetti e da generare utilizzando la procedura [Genera disegno ortogonale](#). Il file di dati contiene le classificazioni o i punteggi di preferenza dei profili raccolti dai soggetti. I file di dati e del piano possono essere specificati ricorrendo ai sottocomandi `PLAN` e `DATA`, rispettivamente. Il metodo di registrazione dei dati viene specificato utilizzando i sottocomandi `SEQUENCE`, `RANK` o `SCORE`. La seguente sintassi dei comandi indica una specifica minima:

```
CONJOINT PLAN='CPLAN.SAV' /DATA='RUGRANKS.SAV'  
/SEQUENCE=PREF1 TO PREF22.
```

## **Specifica del file del piano e del file di dati**

Il comando `CONJOINT` fornisce varie opzioni che consentono di specificare il file del piano e il file di dati.

- È possibile specificare esplicitamente i nomi di file per due file. Ad esempio:

```
CONJOINT PLAN='CPLAN.SAV' /DATA='RUGRANKS.SAV'
```

- Se si specifica solo un file del piano o un file di dati, il comando `CONJOINT` legge il file specificato e utilizza il file di dati attivo per l'altro file. Ad esempio, se si specifica un file di dati ma si omette il file del piano (non è possibile ometterli entrambi), il file di dati attivo viene utilizzato come file del piano, come dimostrato nel seguente esempio:

```
CONJOINT DATA='RUGRANKS.SAV'
```

- Per indicare il file di dati attivo, è possibile ricorrere a un asterisco (\*) in sostituzione del nome di file, come dimostrato nel seguente esempio:

```
CONJOINT PLAN='CPLAN.SAV' /DATA=*
```

I file di dati attivi vengono utilizzati come dati di preferenza. Si noti che non è possibile utilizzare l'asterisco (\*) sia per il file del piano che per il file di dati.

## **Specifica della modalità di registrazione dei dati**

È necessario specificare la modalità con cui i dati di preferenza sono stati registrati. I dati possono essere registrati in base a uno dei tre seguenti metodi: in modo sequenziale, come classificazioni o come punteggi di preferenza. Questi tre metodi sono indicati dai sottocomandi `SEQUENCE`, `RANK` e `SCORE`. Uno, e solo uno, di questi sottocomandi può essere specificato nell'ambito del comando `CONJOINT`.

### **Sottocomando SEQUENCE**

Il sottocomando `SEQUENCE` specifica che i dati sono stati registrati sequenzialmente, in modo che ogni punto di dati del file di dati rappresenti un numero di profilo, che inizia con il profilo favorito e termina con quello meno preferito. È questa la modalità con cui i dati sono stati registrati se al soggetto è stato richiesto di ordinare i profili dal favorito al meno preferito. Il ricercatore registra qual era il primo numero di profilo, qual era il secondo e così via.

```
CONJOINT PLAN=* /DATA='RUGRANKS.SAV'  
/SEQUENCE=PREF1 TO PREF22.
```

- La variabile `PREF1` contiene il numero di profilo relativo al profilo favorito rispetto ai 22 profili presenti nel piano ortogonale. La variabile `PREF22` contiene il numero di profilo relativo al profilo meno preferito del piano.

### **Sottocomando RANK**

Il sottocomando `RANK` indica che ogni punto di dati rappresenta una classificazione, a partire dalla classificazione del profilo 1, quindi del profilo 2 e così via. È questa la modalità con cui i dati sono stati registrati se al soggetto è stato richiesto di assegnare ai singoli profili una

classificazione da 1 a  $n$ , dove  $n$  rappresenta il numero di profili. Una classificazione minore implica una maggiore preferenza.

```
CONJOINT PLAN=* /DATA='RUGRANKS.SAV'  
/RANK=RANK1 TO RANK22.
```

- La variabile *RANK1* contiene la classificazione del profilo 1 rispetto al totale dei 22 profili presenti nel piano ortogonale. La variabile *RANK22* contiene la classificazione del profilo 22.

### **Sottocomando SCORE**

Il sottocomando *SCORE* indica che ogni punto di dati rappresenta un punteggio di preferenza assegnato ai profili, a partire dal punteggio del profilo 1, quindi del profilo 2 e così via. Questo tipo di dati può essere generato, ad esempio, quando si richiede ai soggetti di assegnare un numero compreso tra 1 e 100 per quantificare l'apprezzamento del profilo. Un punteggio superiore implica una maggiore preferenza.

```
CONJOINT PLAN=* /DATA='RUGRANKS.SAV'  
/SCORE=SCORE1 TO SCORE22.
```

- La variabile *SCORE1* contiene il punteggio relativo al profilo 1, mentre *SCORE22* quello relativo al profilo 22.

## **Sottocomandi facoltativi**

Il comando *CONJOINT* offre numerosi sottocomandi facoltativi che forniscono funzionalità aggiuntive e un livello di controllo ulteriore rispetto al necessario.

### **Sottocomando SUBJECT**

Il sottocomando *SUBJECT* consente di specificare una variabile dal file di dati da utilizzare come identificatore per i soggetti. Se non si specifica una variabile soggetto, il comando *CONJOINT* presume che tutti i casi presenti nel file di dati provengano da un unico soggetto. Nell'esempio seguente viene specificato che, come identificatore per i soggetti, deve essere utilizzato l'*ID* della variabile, dal file *rugranks.sav*.

```
CONJOINT PLAN=* /DATA='RUGRANKS.SAV'  
/SCORE=SCORE1 TO SCORE22 /SUBJECT=ID.
```

### **Sottocomando FACTORS**

Il sottocomando *FACTORS* consente di specificare il modello che descrive la relazione prevista tra i fattori e le classificazioni o i punteggi. Se per un fattore non si specifica alcun modello, *CONJOINT* assume un modello discreto. È possibile specificare uno dei quattro modelli seguenti:

**DISCRETE.** Il modello *DISCRETE* indica che i livelli di fattore sono categoriali e non è stata fatta alcuna ipotesi sulla relazione tra il fattore e le classificazioni o i punteggi. È l'impostazione predefinita.

**LINEAR.** Il modello `LINEAR` indica una relazione lineare prevista tra il fattore e le classificazioni o i punteggi. È possibile specificare la direzione prevista della relazione lineare utilizzando le parole chiave `MORE` e `LESS`. `MORE` indica che si prevede che siano preferiti i livelli più alti di un fattore, mentre `LESS` indica che la preferenza verrà accordata ai livelli più bassi. La specifica di `MORE` o `LESS` non incide sulle stime delle utilità. Queste parole chiave vengono semplicemente utilizzate per identificare i soggetti le cui stime non corrispondono alla direzione prevista.

**IDEAL.** Il modello `IDEAL` indica una relazione quadratica prevista tra le classificazioni o i punteggi e il fattore. Si suppone che esista un livello ideale per il fattore e la distanza da questo punto (in entrambe le direzioni) è associata a una riduzione della preferenza. I fattori descritti con questo modello devono possedere almeno tre livelli.

**ANTIIDEAL.** Il modello `ANTIIDEAL` indica una relazione quadratica prevista tra le classificazioni o i punteggi e il fattore. Si suppone che esista un livello peggiore per il fattore e la distanza da questo punto (in entrambe le direzioni) è associata a un aumento della preferenza. I fattori descritti con questo modello devono possedere almeno tre livelli.

Nella seguente sintassi dei comandi viene fornito un esempio che utilizza il sottocomando `FACTORS`:

```
CONJOINT PLAN=* /DATA='RUGRANKS.SAV'
/RANK=RANK1 TO RANK22 /SUBJECT=ID
/FACTORS=PACKAGE BRAND (DISCRETE) PRICE (LINEAR LESS)
SEAL (LINEAR MORE) MONEY (LINEAR MORE).
```

- Si noti che sia *package* sia *brand* sono definiti come modelli discreti.

### **Sottocomando PRINT**

Il sottocomando `PRINT` consente di controllare il contenuto dell'output tabellare. Ad esempio, se si dispone di un elevato numero di soggetti, è possibile scegliere di limitare l'output ai soli risultati riassuntivi, omettendo l'output dettagliato relativo ai singoli soggetti, come dimostrato nell'esempio seguente:

```
CONJOINT PLAN=* /DATA='RUGRANKS.SAV'
/RANK=RANK1 TO RANK22 /SUBJECT=ID
/PRINT=SUMMARYONLY.
```

È inoltre possibile stabilire se l'output deve includere l'analisi dei dati sperimentali, i risultati dei vari casi di simulazione inclusi nel file del piano, entrambi o nessuno. I casi di simulazione non vengono classificati dai soggetti, ma rappresentano i profili dei prodotti che potrebbero risultare interessanti. La procedura Conjoint utilizza l'analisi dei dati sperimentali per eseguire previsioni della preferenza relativa per ognuno dei profili di simulazione. Nell'esempio seguente l'output dettagliato per i singoli soggetti viene eliminato e l'output è limitato ai soli risultati delle simulazioni:

```
CONJOINT PLAN=* /DATA='RUGRANKS.SAV'
/RANK=RANK1 TO RANK22 /SUBJECT=ID
/PRINT=SIMULATION SUMMARYONLY.
```

**Sottocomando PLOT**

Il sottocomando `PLOT` controlla se i grafici sono inclusi nell'output. Come nell'output tabellare (sottocomando `PRINT`), è possibile verificare se l'output è limitato ai risultati riassuntivi o se include anche i risultati relativi ai singoli soggetti. Per impostazione predefinita, non viene generato alcun grafico. Nell'esempio seguente l'output include tutti i grafici disponibili:

```
CONJOINT PLAN=* /DATA='RUGRANKS.SAV'  
/RANK=RANK1 TO RANK22 /SUBJECT=ID  
/PLOT=ALL.
```

**Sottocomando UTILITY**

Il sottocomando `UTILITY` scrive un file di dati in formato IBM® SPSS® Statistics che contiene informazioni dettagliate sui singoli soggetti. Include inoltre le utilità per i fattori `DISCRETE`, l'inclinazione e le funzioni quadratiche per i fattori `LINEAR`, `IDEAL` e `ANTIIDEAL`, la costante di regressione e i punteggi di preferenza previsti. Questi valori possono quindi essere utilizzati in ulteriori analisi o per eseguire grafici aggiuntivi con altre procedure. Nell'esempio seguente viene creato un file di utilità denominato *rugutil.sav*:

```
CONJOINT PLAN=* /DATA='RUGRANKS.SAV'  
/RANK=RANK1 TO RANK22 /SUBJECT=ID  
/UTILITY='RUGUTIL.SAV' .
```

# ***Parte II: Esempi***

# Utilizzo dell'analisi Conjoint per creare un modello di preferenze per battitappeto

Come tipico esempio di un'analisi Conjoint (Green e Wind, 1973), un'azienda interessata alla commercializzazione di un nuovo battitappeto desidera esaminare l'influenza di cinque fattori sulle preferenze del consumatore, ovvero design della confezione, marca, prezzo, la presenza di un *marchio di qualità* e una garanzia "Soddisfatti o rimborsati". Esistono tre livelli di fattore per il design della confezione, che differiscono per la posizione della spazzola dell'applicatore; tre marchi (*K2R*, *Glory* e *Bissell*); tre livelli di prezzo e due livelli (no o sì) per ciascuno degli ultimi due fattori. La tabella seguente mostra le variabili utilizzate nello studio relativo al battitappeto, con relative etichette e valori.

Tabella 5-1  
Variabili nello studio relativo al battitappeto

Nome di variabile	Etichetta di valore	Etichetta del valore
<i>confezione</i>	design confezione	A*, B*, C*
<i>marca</i>	nome marca	K2R, Glory, Bissell
<i>prezzo</i>	prezzo	\$1.19, \$1.39, \$1.59
<i>marchio di qualità</i>	Presenza di un marchio di qualità	no, sì
<i>garanzia</i>	garanzia "Soddisfatti o rimborsati"	no, sì

Potrebbero esservi altri fattori e livelli di fattori che caratterizzano i battitappeto, ma la direzione nutre interesse solo per quelli sopraelencati. Si tratta di un punto fondamentale nell'analisi congiunta poiché si devono scegliere solo i fattori (variabili indipendenti) che si ritiene possano influenzare maggiormente le preferenze del soggetto (variabile dipendente). Mediante l'uso dell'analisi congiunta, si svilupperà un modello delle preferenze del cliente basate sui suddetti cinque fattori.

Nell'esempio si utilizzano le informazioni dei seguenti file di dati: *carpet\_prefs.sav* contiene i dati raccolti dai soggetti, *carpet\_plan.sav* contiene i profili dei prodotti esaminati e *conjoint.sps* contiene la sintassi dei comandi necessaria per eseguire l'analisi. [Per ulteriori informazioni, vedere l'argomento File di esempio in l'appendice A in IBM SPSS Conjoint 20.](#)

## Generazione di un disegno ortogonale

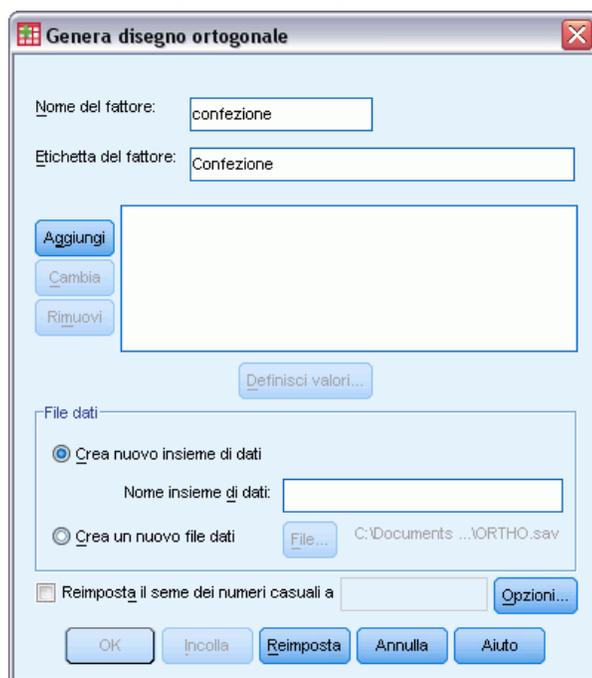
Il primo passo in un'analisi congiunta è di creare le combinazioni dei livelli di fattore che vengono presentati ai soggetti come profili di prodotti. Poiché perfino un numero ridotto di fattori e pochi livelli per ogni fattore conducono a un numero ingestibile di potenziali profili di prodotti, è necessario generare un sottoinsieme rappresentativo noto come **matrice ortogonale**.

La procedura Genera disegno ortogonale crea una matrice ortogonale, denominata anche **disegno ortogonale**, e memorizza le informazioni in un file di dati. Diversamente dalla maggior parte delle procedure, prima dell'esecuzione di Genera disegno ortogonale non è richiesto un file di dati attivo. Se non si dispone di un file di dati attivo, è possibile crearne uno, generare i nomi delle variabili, le etichette delle variabili e le etichette dei valori dalle opzioni disponibili nelle finestre di dialogo. Se un file di dati attivo è già disponibile, è possibile sostituirlo o salvare il disegno ortogonale come un file di dati distinto.

Per creare un disegno ortogonale:

- Dai menu, scegliere:  
Dati > Disegno ortogonale > Genera...

Figura 5-1  
Finestra di dialogo Genera disegno ortogonale



- Immettere confezione nella casella di testo Nom fattore e design della confezione nella casella di testo Etichetta del fattore.
- Fare clic su Aggiungi.  
In tal modo si crea un elemento etichettato *confezione* 'design confezione' (?). Selezionare l'elemento.
- Fare clic su Definisci valori.

Figura 5-2  
Finestra di dialogo Genera disegno ortogonale

- ▶ Immettere i valori 1, 2 e 3 per rappresentare i design della confezione  $A^*$ ,  $B^*$  e  $C^*$ . Immettere anche le etichette  $A^*$ ,  $B^*$  e  $C^*$ .
- ▶ Fare clic su Continua.

Ora, ripetere il processo per i fattori restanti: *marca*, *prezzo*, *marchio di qualità* e *garanzia*. Utilizzare i valori e le etichette della tabella seguente, che contiene i valori immessi precedentemente per *confezione*.

Nome del fattore	Etichetta del fattore	Valori	Etichette
confezione	design confezione	1, 2, 3	$A^*$ , $B^*$ , $C^*$
marca	nome marca	1, 2, 3	K2R, Glory, Bissell
prezzo	prezzo	1.19, 1.39, 1.59	\$1.19, \$1.39, \$1.59
marchio di qualità	Presenza di un marchio di qualità	1, 2	no, sì
garanzia	garanzia "Soddisfatti o rimborsati"	1, 2	no, sì

Dopo aver completato le specifiche dei fattori

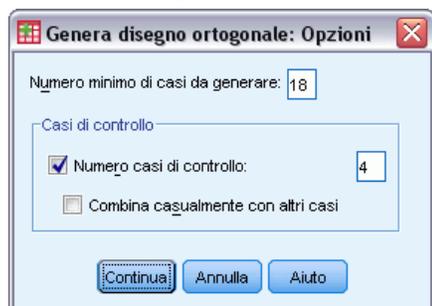
- ▶ Nel gruppo File di dati, lasciare il valore predefinito Crea nuovo insieme di dati e immettere il nome dell'insieme di dati. Il disegno generato verrà salvato in un nuovo insieme di dati, nella sessione corrente, con il nome specificato.
- ▶ Selezionare Reimposta il seme dei numeri casuali a e immettere il valore 2000000.

Per la generazione di un disegno ortogonale è richiesto un insieme di numeri casuali. Per duplicare un disegno, che in questo caso è quello utilizzato per lo studio di casi presente, è necessario impostare il valore di seme prima di generare il disegno e reimpostarlo allo stesso valore tutte le volte successive che si genera il disegno. Il disegno utilizzato per questo studio di casi è stato generato con il valore di seme 2000000.

- Fare clic su Opzioni.

Figura 5-3

Finestra di dialogo Opzioni di Genera disegno ortogonale



- Nella casella di testo Numero minimo di casi da generare, digitare 18.

Per impostazione predefinita, viene generato il numero minimo di casi necessari per una matrice ortogonale. La procedura determina il numero di casi da amministrare per eseguire una stima delle utilità. È inoltre possibile specificare un numero minimo di casi da generare, come è stato fatto qui. Tale azione è necessaria se, per esempio, il numero minimo di casi predefinito è troppo basso o se è necessario un determinato numero minimo di casi per delle considerazioni sul disegno sperimentali.

- Selezionare Numero di casi di controllo e digitare 4.

I casi di controllo vengono stimati in base ai soggetti, ma non vengono utilizzati dall'analisi congiunta per eseguire la stima delle utilità. Essi vengono utilizzati per controllare la validità delle utilità stimate. I casi di controllo vengono generati da un altro piano casuale e non dal piano ortogonale sperimentale.

- Fare clic su Continua nella finestra di dialogo Genera disegno ortogonale: Opzioni.
- Fare clic su OK nella finestra di dialogo Genera disegno ortogonale.

Figura 5-4  
Esempio di disegno ortogonale per il battitappeto

	pacchetto	marca	prezzo	guarnizione	soldi	STATUS_	CARD_
14	A*	Glory	1	Si	No	Disegno	14
15	B*	K2R	1	Si	Si	Disegno	15
16	A*	K2R	1	No	No	Disegno	16
17	A*	Bissell	2	No	Si	Disegno	17
18	B*	Bissell	1	No	No	Disegno	18
19	A*	Bissell	2	Si	No	Controllo	19
20	C*	K2R	1	Si	No	Controllo	20
21	A*	Glory	2	No	No	Controllo	21
22	A*	Bissell	1	No	No	Controllo	22

Il disegno ortogonale viene visualizzato nell'Editor dei dati e se si visualizzano le etichette dei valori al posto dei valori dei dati effettivi si ottiene un risultato migliore. A tal fine selezionare Etichette dei valori dal menu Visualizza.

Le variabili nel file di dati sono i fattori utilizzati per specificare il disegno. Ogni caso rappresenta un profilo del prodotto nel disegno. Si noti che nel file di dati sono visualizzate le due variabili aggiuntive *CARD\_* e *STATUS\_*. *CARD\_* assegna un numero progressivo a ogni profilo che viene utilizzato come identificativo. *STATUS\_* indica se il profilo fa parte del disegno sperimentale (i primi 18 casi), di un caso di controllo (gli ultimi 4 casi) o di un caso di simulazione (è un argomento che verrà trattato successivamente nel presente studio di casi).

Per l'analisi dei dati è richiesto il disegno ortogonale. È quindi opportuno salvare il disegno in un file di dati. Per comodità, il disegno corrente è stato salvato in *carpet\_plan.sav* (i disegni ortogonali vengono denominati anche **piani**).

## Creazione di stimoli sperimentali: Visualizzazione del disegno

Una volta creato il disegno ortogonale, lo si utilizza per creare i profili dei prodotti da classificare secondo i soggetti. È possibile ottenere un elenco di profili in una tabella unica oppure visualizzare ogni profilo in una tabella distinta.

Per visualizzare un disegno ortogonale:

- Dai menu, scegliere:  
Dati > Disegno ortogonale > Visualizza...

Figura 5-5  
Finestra di dialogo Proprietà visualizzazione disegno



- ▶ Per i fattori selezionare *confezione*, *marca*, *prezzo*, *marchio di qualità* e *garanzia*.  
Le informazioni contenute nelle variabili *STATUS\_* e *CARD\_* vengono incluse automaticamente nell'output, quindi non è necessario selezionarle.
- ▶ Selezionare *Lista per lo sperimentatore* nel gruppo *Formato*. In tal modo si visualizza l'intero disegno ortogonale in una tabella unica.
- ▶ Fare clic su *OK*.

**Figura 5-6**  
Visualizzazione di un disegno ortogonale: Layout a tabella singola

	ID scheda	Confezione	Marca	Prezzo	Sigillo di garanzia	Soddisfatti o rimborsati
1	1	A*	Glory	1	Sì	No
2	2	B*	K2R	1	No	No
3	3	B*	Glory	1	No	Sì
4	4	C*	Glory	2	No	No
5	5	C*	Bissell	1	No	No
6	6	A*	Bissell	1	No	No
7	7	B*	Bissell	2	Sì	No
8	8	A*	K2R	2	No	Sì
9	9	C*	K2R	1	No	No
10	10	C*	Glory	1	No	Sì
11	11	C*	K2R	2	Sì	No
12	12	B*	Glory	2	No	No
13	13	C*	Bissell	1	Sì	Sì
14	14	A*	Glory	1	Sì	No
15	15	B*	K2R	1	Sì	Sì
16	16	A*	K2R	1	No	No
17	17	A*	Bissell	2	No	Sì
18	18	B*	Bissell	1	No	No
19 <sup>a</sup>	19	A*	Bissell	2	Sì	No
20 <sup>a</sup>	20	C*	K2R	1	Sì	No
21 <sup>a</sup>	21	A*	Glory	2	No	No
22 <sup>a</sup>	22	A*	Bissell	1	No	No

<sup>a</sup>. Controllo

Come illustrato nell'Editor dei dati, l'output è simile assomiglia al disegno ortogonale, con una riga per ogni profilo e i fattori come colonne. Si noti tuttavia, che le intestazioni delle colonne corrispondono alle etichette delle variabili, non ai nomi delle variabili visualizzati nell'Editor dei dati. Inoltre, i casi di controllo vengono identificati con un'annotazione. Questo tipo di informazione è utile allo sperimentatore, ma i soggetti non devono sapere quali sono gli eventuali casi di controllo.

In base a come si creano e si consegnano i profili dei prodotti finali, è opportuno salvare la tabella come file HTML, Word/RTF, Excel o PowerPoint. Per compiere tale operazione, è sufficiente selezionare la tabella nel Viewer, fare clic con il pulsante destro del mouse e selezionare Esporta. Se inoltre si utilizza una versione esportata per creare i profili dei prodotti finali, accertarsi di rimuovere le annotazioni relative ai casi di controllo.

È probabile che ai fini dell'indagine sia meglio generare una tabella distinta per ogni profilo di prodotti. In questo caso risulta ideale l'esportazione in PowerPoint, poiché ogni tabella (profilo dei prodotti) viene collocata su una diapositiva separata.

Per visualizzare ogni profilo in una tabella distinta:

- ▶ Fare clic sullo strumento Richiama finestra per selezionare Visualizza disegno.
- ▶ Deselezionare Lista per lo sperimentatore e selezionare Profili per i soggetti.
- ▶ Fare clic su OK.

**Figura 5-7**  
Visualizzazione di un disegno ortogonale: layout a più tabelle

Numero del profilo 1					
ID scheda	Confezione	Marca	Prezzo	Sigillo di garanzia	Soddisfatti o rimborsati
1	A*	Glory	1	Sì	No

Numero del profilo 2					
ID scheda	Confezione	Marca	Prezzo	Sigillo di garanzia	Soddisfatti o rimborsati
2	B*	K2R	1	No	No

Le informazioni di ogni profilo di prodotti vengono visualizzate in una tabella separata. I casi di controllo inoltre, sono indistinguibili dal resto dei casi, quindi non è necessario rimuovere gli identificatori dei casi di controllo come per il layout a tabella singola.

## Esecuzione dell'analisi

Finora si è generato un disegno ortogonale e si è appreso come visualizzare i profili dei prodotto associati. A questo punto si è pronti per eseguire un'analisi congiunta.

**Figura 5-8**  
Esempio di dati di preferenza per il battitappeto

	ID	PREF1	PREF2	PREF3	PREF4	PREF5	PREF6	PREF7
1	1	13	15	1	20	14	7	11
2	2	15	7	18	2	12	3	11
3	3	2	18	14	16	22	13	20
4	4	13	10	20	14	2	18	16
5	5	13	18	2	10	20	15	9
6	6	15	2	3	12	18	7	20
7	7	13	7	15	18	2	3	10
8	8	15	7	13	4	6	16	8
9	9	20	9	10	11	4	5	13
10	10	8	21	19	17	4	11	12

I dati di preferenza raccolti dai soggetti sono memorizzati nel file *carpet\_prefs.sav*. I dati consistono nelle risposte di 10 soggetti, ognuno dei quali è identificato con un valore unico della variabile *ID*. Ai soggetti è stato chiesto di classificare 22 profili di prodotti in ordine di preferenza. Nelle variabili *PREF1* fino a *PREF22* sono contenuti gli ID dei profili di prodotti associati, ovvero gli ID scheda provenienti da *carpet\_plan.sav*. Il soggetto 1, ad esempio, ha preferito più di tutti il profilo 13, quindi *PREF1* ha il valore 13.

L'analisi dei dati richiede l'uso della sintassi dei comandi, in particolare il comando `CONJOINT`. Nel file *conjoint.sps* è disponibile la sintassi del comando necessaria.

```
CONJOINT PLAN='specifica di file'
  /DATA='specifica di file'
  /SEQUENCE=PREF1 TO PREF22
  /SUBJECT=ID
  /FACTORS=PACKAGE BRAND (DISCRETE)
  PRICE (LINEAR LESS)
```

```
SEAL (LINEAR MORE) MONEY (LINEAR MORE)
/PRINT=SUMMARYONLY.
```

- Il sottocomando `PLAN` specifica il file che contiene il disegno ortogonale, nell'esempio è il file `carpet_plan.sav`.
- Il sottocomando `DATA` specifica il file che contiene i dati di preferenza, nell'esempio è il file `carpet_prefs.sav`. Se si selezionano i dati di preferenza come file di dati attivo, è possibile sostituire la specifica del file con un asterisco (\*), senza le virgolette.
- Il sottocomando `SEQUENCE` specifica che ogni punto di dati nei dati di preferenza è un numero di profilo, che inizia con il profilo favorito e termina con quello meno preferito.
- Il sottocomando `SUBJECT` specifica che la variabile `ID` identifica il soggetto.
- Il sottocomando `FACTORS` specifica un modello che descrive la relazione prevista tra i dati di preferenza e i livelli di fattore. I fattori specificati si riferiscono alle variabili definite nel file del piano denominato nel sottocomando `PLAN`.
- Si utilizza la parola chiave `DISCRETE` quando i livelli di fattore sono categoriali e non è stata fatta alcuna ipotesi sulla relazione tra i livelli e i dati. È il caso dei fattori *confezione* e *marca* che rappresentano rispettivamente il design della confezione e il nome della marca. `DISCRETE` viene assunta se un fattore non è etichettato con una delle quattro alternative (`DISCRETE`, `LINEAR`, `IDEAL`, `ANTIIDEAL`) o se non è incluso nel sottocomando `FACTORS`.
- La parola chiave `LINEAR`, utilizzata per il resto dei fattori, indica che si presume che i dati siano correlati in modo lineare al fattore. Per esempio, generalmente si presume che la preferenza sia correlata in modo lineare al prezzo. Con le parole chiave `IDEAL` e `IDEAL` si possono specificare anche dei modelli quadratici (non utilizzati nell'esempio).
- Le parole chiave `MORE` e `LESS` che seguono `LINEAR`, indicano la direzione prevista per la relazione. Poiché si presume una preferenza maggiore per i prezzi inferiori, la parola chiave `LESS` viene utilizzata per il *prezzo*. Si presume tuttavia che vi sia una preferenza maggiore per un *marchio di qualità* o per una garanzia "Soddisfatti o rimborsati", quindi si utilizza la parola chiave `MORE` per *marchio di qualità* e *garanzia* (tenere presente che i livelli di entrambi i fattori sono stati impostati a 1 per *no* e a 2 per *si*).  
 Specificando `MORE` o `LESS` non si cambiano i segni dei coefficienti né si incide sulle stime delle utilità. Queste parole chiave vengono utilizzate solo per l'identificazione dei soggetti le cui stime non corrispondono alla direzione prevista. Analogamente, la selezione di `IDEAL` invece di `ANTIIDEAL` o viceversa, non influisce sui coefficienti né sulle utilità.
- Il sottocomando `PRINT` specifica che l'output contiene le informazioni solo per il gruppo di soggetti nel suo complesso (parola chiave `SUMMARYONLY`). Le informazioni individuali di ogni soggetto vengono eliminate.

Provare a eseguire questa sintassi dei comandi. Accertarsi di aver incluso i percorsi corretti di `carpet_prefs.sav` e `carpet_plan.sav`. Per ottenere una descrizione completa di tutte le opzioni, vedere il comando `CONJOINT` in *Command Syntax Reference*.

## Punteggi di utilità

Figura 5-9  
Punteggi di utilità

		Valuta. utilità	Err. standard
confezione	A*	-2,233	,192
	B*	1,867	,192
	C*	,367	,192
marca	K2R	,367	,192
	Glory	-,350	,192
	Bissell	-,017	,192
prezzo	€1,19	-6,595	,988
	€1,39	-7,703	1,154
	€1,59	-8,811	1,320
garanzia	no	2,000	,287
	sì	4,000	,575
soddrimb	no	1,250	,287
	sì	2,500	,575
(Costante)		12,870	1,282

La tabella mostra i punteggi di utilità (con valore parziale) e i rispettivi errori standard relativi a ogni livello di fattore. I valori di utilità più alti indicano una maggiore preferenza. Come previsto, vi è una relazione inversa tra il prezzo e l'utilità in cui i prezzi più alti corrispondono all'utilità più bassa (valori negativi più elevati indicano un'utilità inferiore). Come anticipato, la presenza di un marchio di qualità o di una garanzia "Soddisfatti o rimborsati" corrisponde a un'utilità più elevata.

Poiché le utilità vengono espresse tutte in un'unità comune, è possibile sommarle per fornire l'**utilità totale** di qualsiasi combinazione. Ad esempio, per un battitappeto con un design della confezione *B\**, di marca *K2R*, prezzo di \$1,19 e privo di marchio di qualità o di garanzia, l'utilità totale è:

$$\begin{aligned} & \text{utilità(confezione B*)} + \text{utilità(K2R)} + \text{utilità(\$1,19)} \\ & + \text{utilità(nessun marchio di qualità)} + \text{utilità(nessuna garanzia)} + \text{costante} \end{aligned}$$

o

$$1.867 + 0.367 + (-6.595) + 2.000 + 1.250 + 12.870 = 11.759$$

per un battitappeto con un design della confezione *C\**, di marca *Bissell*, prezzo di \$1,59, con marchio di qualità e garanzia "Soddisfatti o rimborsati", l'utilità totale è:

$$0.367 + (-0.017) + (-8.811) + 4.000 + 2.500 + 12.870 = 10.909$$

## Coefficienti

Figura 5-10  
Coefficienti

	B valutazione
prezzo	-5.542
guar.	2.000
soldi	1.250

La tabella mostra i coefficienti di regressione lineare relativi ai fattori specificati come LINEAR (per i modelli IDEAL e ANTIIDEAL ci sarebbe anche un termine quadratico). Per determinare l'utilità per un livello di fattore particolare, moltiplicare il livello per il coefficiente. Ad esempio, l'utilità prevista per un prezzo di \$1,19 era catalogata come -6,595 nella tabella delle utilità. Si tratta semplicemente del valore del livello di prezzo, 1,19, moltiplicato per il coefficiente di prezzo -5,542.

## Importanza relativa

L'intervallo dei valori di utilità (dai più alti ai più bassi) di ogni fattore fornisce una misura dell'importanza del fattore nella preferenza complessiva. I fattori con intervalli di utilità maggiori giocano un ruolo più significativo rispetto a quelli con intervalli minori.

Figura 5-11  
Valori di importanza

pacchetto	35.635
marca	14.911
prezzo	29.410
guar.	11.172
soldi	8.872

segno important medio

La seguente tabella fornisce una misura dell'importanza relativa di ogni fattore, nota anche come un punteggio o un valore di **importanza**. Il valori vengono calcolati prendendo separatamente l'intervallo di utilità di ogni fattore e dividendo la somma degli intervalli di utilità per tutti i fattori. I valori rappresentano quindi le percentuali e sono caratterizzati dal fatto che sommati equivalgono a 100. Tenere presente che per ogni soggetto vengono effettuati dei calcoli separati e che con i risultati viene calcolata una media di tutti i soggetti.

Si noti che mentre le utilità globali o di riepilogo e i coefficienti di regressione dei disegni ortogonali sono gli stessi con o senza un sottocomando SUBJECT, le importanze generalmente differiscono. Per i risultati riassuntivi senza un sottocomando SUBJECT, è possibile calcolare le importanze direttamente dalle utilità di riepilogo, proprio come per i soggetti individuali. Tuttavia, quando si utilizza un sottocomando SUBJECT, viene fatta una media delle importanze relative ai soggetti individuali e generalmente le importanze che ne risultano non corrispondono a quelle calcolate mediante l'uso delle utilità di riepilogo.

I risultati mostrano che il design della confezione influisce maggiormente sulla preferenza globale. Vale a dire che vi è una gran differenza di preferenza tra i profili di prodotti che contengono le confezioni più apprezzate e quelli che contengono le confezioni meno apprezzate. I risultati mostrano inoltre che la garanzia "Soddisfatti o rimborsati" ha il ruolo meno importante

nella definizione della preferenza globale. Il prezzo, pur di fondamentale importanza, non è altrettanto rilevante quanto il design della confezione, forse a causa del ridotto intervallo di prezzi.

## Correlazioni

Figura 5-12  
Coefficienti di correlazione

	Valore	Sig.
Pearson's R	.982	.000
Kendall's tau	.892	.000
Kendall's tau for Holdouts	.667	.087

Nella tabella sono visualizzate due statistiche,  $R$  di Pearson e tau di Kendall, mediante le quali si ottengono le misurazioni della correlazione tra le preferenze osservate e quelle stimate.

Nella tabella è riportata anche la statistica tau di Kendall esclusivamente per i profili di controllo. Tenere presente che i profili di controllo (quattro nell'esempio presente) sono stati classificati dai soggetti, ma non sono stati utilizzati dalla procedura Analisi congiunta per la stima delle utilità. L'Analisi congiunta calcola invece le correlazioni tra i punteggi di rango osservati e previsti relativi ai profili per controllare la validità delle utilità.

In numerose analisi congiunte, il numero di parametri è vicino al numero dei profili classificati, gonfiando in modo artificiale la correlazione tra i punteggi osservati e quelli stimati. In questi casi, le correlazioni per i profili di controllo possono fornire un'indicazione migliore dell'adattamento del modello. Tenere comunque presente che i controlli produrranno sempre dei coefficienti di correlazione più bassi.

## Inversioni

Quando si specificano i modelli `LINEAR` per *prezzo*, *marchio di qualità* e *garanzia*, si sceglie una direzione prevista (`LESS` o `MORE`) per la relazione lineare tra il valore della variabile e la preferenza a esso relativa. La procedura di Analisi congiunta tiene traccia del numero di soggetti le cui preferenze si sono rivelate opposte alla relazione prevista, per esempio, una preferenza maggiore per i prezzi più elevati o una preferenza minore per la garanzia "Soddisfatti o rimborsati". Tali casi vengono denominati **inversi**.

**Figura 5-13**  
Numero di inversi per fattore e soggetto

Factor	prezzo		3
	soldi		2
	guar.		2
	marca		0
	pacchetto		0
oggetto	1	oggetto 1	1
	2	oggetto 2	2
	3	oggetto 3	0
	4	oggetto 4	0
	5	oggetto 5	0
	6	oggetto 6	1
	7	oggetto 7	0
	8	oggetto 8	0
	9	oggetto 9	1
	10	oggetto 10	2

Nella seguente tabella è riportato il numero di inversi per ogni fattore e ogni soggetto. Se per esempio, tre soggetti hanno mostrato un inverso per il *prezzo*, significa che essi hanno preferito dei profili di prodotto con prezzi più elevati.

## **Esecuzione delle simulazioni**

Il vero punto di forza dell'analisi congiunta è costituito dalla capacità di prevedere la preferenza dei profili di prodotti che non sono stati classificati dai soggetti. Tali casi vengono denominati **casi di simulazione**. I casi di simulazione fanno parte del piano, insieme ai profili provenienti dal disegno ortogonale e a qualsiasi profilo di controllo.

Il modo più semplice per immettere dei casi di simulazione è dall'Editor dei dati, utilizzando le etichette dei valori create durante la generazione del disegno sperimentale.

Per immettere un caso di simulazione nel file del piano:

- ▶ In una riga nuova della finestra dell'Editor dei dati, selezionare una cella e selezionare il valore desiderato dall'elenco (è possibile visualizzare le etichette selezionando Etichette dei valori dal menu Visualizza). Ripetere l'operazione per tutte le variabili (fattori).
- ▶ Selezionare Simulazione per il valore della variabile *STATUS\_*.
- ▶ Immettere un valore intero da utilizzare come identificatore per la variabile *CARD\_*. I casi di simulazione devono essere numerati separatamente dagli altri casi.

Figura 5-14  
Dati del battitappeto che includono i casi di simulazione

	pacchetto	marca	prezzo	guarnizione	soldi	STATUS_	CARD_
14	A*	Glory	€1.59	Sì	No	Disegno	14
15	B*	K2R	€1.19	Sì	Sì	Disegno	15
16	A*	K2R	€1.59	No	No	Disegno	16
17	A*	Bissell	€1.19	No	Sì	Disegno	17
18	B*	Bissell	€1.59	No	No	Disegno	18
19	A*	Bissell	€1.19	Sì	No	Controllo	19
20	C*	K2R	€1.19	Sì	No	Controllo	20
21	A*	Glory	€1.19	No	No	Controllo	21
22	A*	Bissell	€1.19	No	No	Controllo	22

La figura mostra una parte del file del piano relativo allo studio sul battitappeto, nel quale sono stati aggiunti due casi di simulazione. Per comodità, i casi sono stati inclusi in *carpet\_plan.sav*.

Per effettuare l'analisi dei casi di simulazione si utilizza la stessa sintassi dei comandi utilizzata precedentemente, ovvero la sintassi del file *conjoint.sps*. Infatti, se è stata eseguita la sintassi descritta precedentemente, si sarà notato che l'output contiene anche i risultati dei casi di simulazione, poiché questi sono inclusi nel file *carpet\_plan.sav*.

È possibile scegliere di eseguire le simulazioni insieme all'analisi iniziale, come nell'esempio, oppure in un momento successivo. Per compiere tale operazione è sufficiente includere i casi di simulazione nel file del piano ed eseguire di nuovo `CONJOINT`. Per ulteriori informazioni, vedere il comando `CONJOINT` in *Command Syntax Reference*.

## Probabilità di preferenza delle simulazioni

Figura 5-15  
Risultati della simulazione

numero di scheda	ID	utilità massima <sup>a</sup>	Bradley-Terry-Luce	Logit
1	1	30.0%	43.1%	30.9%
2	2	70.0%	56.9%	69.1%

Nella tabella sono riportate le probabilità previste relative alla scelta di ogni caso di simulazione come preferito, con tre diversi probabili modelli da scegliere. Il **modello di utilità massima** calcola la probabilità come numero di rispondenti che si prevede scelgano il profilo diviso per il numero totale di rispondenti. Per ciascun rispondente, la scelta prevista è semplicemente il profilo con l'utilità totale più alta. Il modello **BTL (Bradley-Terry-Luce) model** calcola la probabilità come rapporto di un'utilità di profilo per tutti i profili di simulazione, basato sulla media di tutti i rispondenti. Il **modello logit** è simile al modello di Bradley-Terry-Luce, ma usa il logaritmo naturale delle utilità invece delle utilità. Nei 10 soggetti dello studio, tutti e tre i modelli hanno indicato che il profilo di simulazione 2 è quello preferito.

## *File di esempio*

Il file di esempio installato con il prodotto si trova nella sottodirectory *Samples* della directory di installazione. La sottodirectory *Samples* contiene cartelle separate per ciascuna delle seguenti lingue: Inglese, Francese, Tedesco, Italiano, Giapponese, Coreano, Polacco, Russo, Cinese semplificato, Spagnolo e Cinese tradizionale.

Non tutti i file di esempio sono disponibili in tutte le lingue. Se un file di esempio non è disponibile in una lingua, la cartella di tale lingua contiene una versione inglese del file.

### **Descrizioni**

Questa sezione contiene brevi descrizioni dei file di esempio utilizzati negli esempi riportati in tutta la documentazione.

- **accidents.sav.** File di dati ipotetici che prende in esame una compagnia di assicurazioni impegnata nello studio dei fattori di rischio correlati all'età e al sesso per gli incidenti automobilistici che si verificano in una determinata regione. Ciascun caso corrisponde a una classificazione incrociata della categoria relativa età e del sesso.
- **adl.sav.** File di dati ipotetici che prende in esame l'impegno richiesto per determinare i vantaggi di un tipo di terapia proposto per i pazienti con problemi di cuore. I medici hanno assegnato in modo casuale i pazienti con problemi di cuore di sesso femminile a uno di due gruppi. Al primo gruppo è stata assegnata la terapia fisica standard; al secondo gruppo, un'ulteriore terapia di supporto psicologico. Dopo tre mesi di trattamenti, a ciascuna capacità dei pazienti che consente di riprendere le normali attività giornaliere è stato assegnato un punteggio come variabile ordinale.
- **advert.sav.** File di dati ipotetici che prende in esame l'impegno di un rivenditore al dettaglio che desidera esaminare la relazione tra il denaro speso per la pubblicità e le vendite risultanti. Finora sono stati raccolti i dati delle vendite precedenti e i relativi costi pubblicitari.
- **afatoxin.sav.** File di dati ipotetici che prende in esame il test di raccolti di mais con presenza di Aflatossina, un veleno la cui concentrazione varia notevolmente nei raccolti. Una macchina per la lavorazione dei cereali ha ricevuto 16 campioni da ciascuno degli otto raccolti di mais e ha misurato i livelli di Aflatossina in parti per miliardo (PPB).
- **anorectic.sav.** Per trovare una sintomatologia standardizzata del comportamento anoressico/bulimico, i ricercatori (Van der Ham, Meulman, Van Strien, e Van Engeland, 1997) hanno condotto uno studio basato su 55 adolescenti affetti da disordini alimentari conosciuti. Ogni paziente è stato visitato quattro volte in quattro anni, per un totale di 220 visite. Durante ogni visita, ai pazienti sono stati assegnati punteggi per ciascuno dei 16 sintomi. I punteggi relativi ai sintomi sono assenti per il paziente 71 alla visita 2, il paziente 76 alla visita 2 e il paziente 47 alla visita 3, con 217 osservazioni valide.
- **bankloan.sav.** File di dati ipotetici che prende in esame l'impegno di una banca nel tentativo di ridurre il tasso di inadempienza nel rimborso di un prestito. Il file contiene informazioni finanziarie e demografiche su 850 vecchi e potenziali clienti. I primi 700 casi riguardano i

clienti a cui sono stati concessi dei prestiti precedentemente. Gli ultimi 150 casi riguardano i potenziali clienti che la banca deve classificare come rischi di credito positivi o negativi.

- **bankloan\_binning.sav.** File di dati ipotetici che contiene informazioni finanziarie e demografiche su 5000 vecchi clienti.
- **behavior.sav.** In un classico esempio (Prezzo e Bouffard, 1974), è stato chiesto a 52 studenti di classificare una combinazione di 15 situazioni e 15 comportamenti utilizzando una scala da 0=“molto appropriato” a 9=“molto inadeguato”. I valori medi riferiti ai partecipanti sono stati considerati dissimilarità.
- **behavior\_ini.sav.** Questo file di dati contiene la configurazione iniziale di una soluzione a due dimensioni per *behavior.sav*.
- **brakes.sav.** File di dati ipotetici che prende in esame il controllo di qualità di un'industria che produce freni a disco per automobili con elevate prestazioni. Il file di dati contiene le misurazioni del diametro di 16 dischi da ciascuna delle otto macchine di produzione. L'obiettivo finale è ottenere un diametro dei dischi pari a 322 millimetri.
- **breakfast.sav.** In uno studio classico (Green e Rao, 1972), è stato chiesto a 21 studenti MBA della Wharton School e ai loro consorti di classificare 15 cibi da colazione in ordine di preferenza, dove il valore 1 corrispondeva all'alimento preferito in assoluto e il valore 15 a quello meno preferito. Le loro preferenze sono state registrate per sei diversi scenari, che comprendevano tutti gli scenari compresi tra “Preferenza generale” e “Solo snack con bibita”.
- **breakfast-overall.sav.** Questo file contiene le preferenze degli alimenti della colazione solo per il primo scenario, “Preferenza generale”.
- **broadband\_1.sav.** File di dati ipotetici che contiene il numero di sottoscrittori, per area, di un provider di servizi a banda larga nazionale. Il file di dati contiene il numero dei sottoscrittori mensili di 85 aree in un periodo di quattro anni.
- **broadband\_2.sav.** Questo file è identico al file *broadband\_1.sav*, ma contiene i dati per ulteriori tre mesi.
- **car\_insurance\_claims.sav.** Un insieme di dati presentato e analizzato altrove (McCullagh e Nelder, 1989) riguarda le richieste di risarcimento auto. La quantità media di richieste di risarcimento può essere adattata come avente una distribuzione gamma, utilizzando una funzione di collegamento inverso per correlare la media della variabile dipendente a una combinazione lineare di età del contraente della polizza e tipo e anni del veicolo. Il numero delle richieste di risarcimento specificato può essere utilizzato come peso scalato.
- **car\_sales.sav.** Questo file di dati ipotetici contiene le stime sulle vendite, i prezzi di listino e le specifiche fisiche di numerose marche e modelli di veicoli. I prezzi di listino e le specifiche fisiche sono state ottenute dal sito *edmunds.com* e dai siti dei produttori.
- **car\_sales\_uprepared.sav.** Questa è una versione modificata di *car\_sales.sav* che non comprende versioni trasformate dei campi.
- **carpet.sav.** Come esempio tipico (Green e Wind, 1973), un'azienda interessata alla commercializzazione di un nuovo battitappeto desidera esaminare l'influenza di cinque fattori sulle preferenze del consumatore, ovvero design della confezione, marca, prezzo, la presenza di un *marchio di qualità* e una garanzia “Soddisfatti o rimborsati”. Esistono tre livelli di fattore per il design della confezione, che differiscono per la posizione della spazzola dell'applicatore; tre marchi (*K2R*, *Glory* e *Bissell*); tre livelli di prezzo e due livelli (no o sì) per ciascuno degli ultimi due fattori. Dieci consumatori sono classificati in 22 profili definiti

da questi fattori. La variabile *Preferenza* include il rango delle classificazioni medie per ogni profilo. Classificazioni basse corrispondono a una preferenza elevata. La variabile riflette una misura globale della preferenza per ogni profilo.

- **carpet\_prefs.sav.** Questo file di dati si basa sullo stesso esempio del file *carpet.sav*, ma contiene le classificazioni effettive raccolte da ciascuno dei 10 clienti. Ai clienti è stato chiesto di classificare 22 profili di prodotti in ordine di preferenza. Le variabili da *PREF1* a *PREF22* contengono gli ID dei profili associati, come definito nel file *carpet\_plan.sav*.
- **catalog.sav.** File di dati ipotetico che contiene le cifre sulle vendite mensili di tre prodotti venduti da una società di vendita per corrispondenza. Il file include anche i dati di cinque possibili variabili predittore.
- **catalog\_seasfac.sav.** Questo file di dati è uguale al file *catalog.sav* con l'eccezione che contiene un insieme di fattori stagionali calcolati dalla procedura Decomposizionale stagionale insieme a variabili di dati.
- **cellular.sav.** File di dati ipotetici che prende in esame l'impegno di un'azienda di telefonia cellulare nel tentativo di ridurre il churn, ovvero l'abbandono dei clienti. Agli account vengono applicati i punteggi relativi alla propensione al churn, con valori compresi tra 0 e 100. Gli account con punteggio pari a 50 o superiore è probabile che stiano cercando nuovi provider.
- **ceramics.sav.** File di dati ipotetici che prende in esame l'impegno di un produttore che desidera stabilire se una nuova lega premium ha una maggiore resistenza al calore rispetto alla lega standard. Ciascun caso rappresenta il test separato di una delle leghe. È indicata la temperatura massima alla quale può essere sottoposto il cuscinetto.
- **cereal.sav.** File di dati ipotetici che prende in esame le preferenze relative agli alimenti della colazione di un campione di 880 persone. Il file riporta anche l'età, il sesso e lo stato civile del campione e se le persone conducono uno stile di vita attivo (in base a un'attività sportiva con frequenza di due volte alla settimana). Ogni caso rappresenta un rispondente separato.
- **clothing\_defects.sav.** File di dati ipotetici che prende in esame il processo di controllo di qualità di un'industria di abbigliamento. Per ciascun lotto prodotto nella fabbrica, gli ispettori prelevano un campione di abiti per contare il numero dei capi che non sono accettabili per la vendita.
- **coffee.sav.** Questo file di dati contiene informazioni sulle immagini percepite di sei marche di caffè freddo (Kennedy, Riquier, e Sharp, 1996). Per ciascuno dei 23 attributi dell'immagine del caffè freddo, sono state selezionate tutte le marche descritte da tale attributo. Le sei marche sono indicate dalle sigle AA, BB, CC, DD, EE e FF per tutelare la confidenzialità dei dati.
- **contacts.sav.** File di dati ipotetici che prende in esame l'elenco dei contatti di un gruppo di rappresentanti di vendita di computer aziendali. Ciascun contatto è classificato in base al reparto della società in cui lavora e dalle relative categorie aziendali. Il file riporta anche l'importo dell'ultima vendita effettuata, il tempo trascorso dall'ultima vendita e le dimensioni della società del contatto.
- **creditpromo.sav.** File di dati ipotetici che prende in esame l'impegno di un grande magazzino nel tentativo di valutare l'efficacia di una recente promozione con carta di credito. A tale scopo, sono stati selezionati 500 titolari di carta in modo casuale. Alla metà di questi è stato inviato un annuncio promozionale che comunica la riduzione del tasso d'interesse nel caso di acquisti effettuati entro i tre mesi successivi. All'altra metà è stato inviato un annuncio stagionale standard.

- **customer\_dbase.sav.** File di dati ipotetico che prende in esame l'impegno di una società nel tentativo di utilizzare le informazioni contenute nel proprio database dei dati per creare offerte speciali per i clienti che più probabilmente risponderanno all'offerta. È stato selezionato in modo casuale un sottoinsieme della base dei clienti a cui è stata inviata l'offerta speciale e sono state registrate le risposte ricevute.
- **customer\_information.sav.** File di dati ipotetici contenente le informazioni postali del cliente, ad esempio il nome e l'indirizzo.
- **customer\_subset.sav.** Un sottoinsieme di 80 casi da *customer\_dbase.sav*.
- **debate.sav.** File di dati ipotetici che prende in esame le risposte appaiate a un'indagine da parte dei partecipanti a un dibattito politico prima e dopo il dibattito. Ogni caso rappresenta un rispondente separato.
- **debate\_aggregate.sav.** File di dati ipotetici che aggrega le risposte contenute nel file *debate.sav*. Ciascun caso corrisponde a una classificazione incrociata della preferenza prima e dopo il dibattito.
- **demo.sav.** File di dati ipotetici che prende in esame un database di clienti che hanno fatto acquisti al fine di inviare offerte mensili tramite il metodo del direct mailing. Viene registrata la risposta dei clienti, sia che abbiano aderito all'offerta o meno, insieme a diverse informazioni demografiche.
- **demo\_cs\_1.sav.** File di dati ipotetici che prende in esame il primo passo che una società intraprende per compilare un database con informazioni ricavate dai sondaggi. Ogni caso rappresenta una diversa città. Sono registrate anche le informazioni sulla regione, provincia, distretto e città.
- **demo\_cs\_2.sav.** File di dati ipotetici che prende in esame il secondo passo che una società intraprende per compilare un database con informazioni ricavate dai sondaggi. Ogni caso rappresenta una diversa unità di abitazione, ricavata dalle città selezionate nel primo passo. Sono registrate anche le informazioni sulla regione, provincia, distretto, città, suddivisione e unità. Il file include inoltre informazioni sul campionamento ottenute dai primi due stadi del disegno.
- **demo\_cs.sav.** File di dati ipotetici che contiene informazioni sulle indagini raccolte utilizzando un disegno di campionamento complesso. Ogni caso rappresenta una diversa unità di abitazione. Sono registrate diverse informazioni demografiche e sul campionamento.
- **dmdata.sav.** File di dati ipotetici che contiene informazioni demografiche e di acquisto di una società di direct marketing. *dmdata2.sav* contiene informazioni su un sottoinsieme di contatti che hanno ricevuto un mailing di prova e *dmdata3.sav* contiene informazioni sui contatti rimanenti che non hanno ricevuto il mailing di prova.
- **dietstudy.sav.** File di dati ipotetici che contiene il risultato di uno studio ipotetico sulla dieta chiamato "Stillman diet" (Rickman, Mitchell, Dingman, e Dalen, 1974). Ogni caso rappresenta un diverso soggetto e ne riporta il peso prima e dopo la dieta in libbre e i livelli dei trigliceridi in mg/100 ml.
- **dvdplayer.sav.** File di dati ipotetici che prende in esame lo sviluppo di un nuovo lettore DVD. Utilizzando un prototipo, il personale addetto al marketing ha raccolto dati sui gruppi di interesse. Ogni caso rappresenta un diverso utente che è stato sottoposto all'indagine e include informazioni demografiche personali dell'utente e sulle risposte che ha fornito riguardo al prototipo.

- **german\_credit.sav.** Questo file di dati contiene informazioni ricavate dall'insieme di dati "German Credit" del Repository of Machine Learning Databases (Blake e Merz, 1998) presso la University of California, Irvine.
- **grocery\_1month.sav.** Questo file di dati ipotetici corrisponde al file di dati *grocery\_coupons.sav* con gli acquisti settimanali organizzati in modo che ogni caso corrisponda a un cliente separato. Alcune delle variabili che cambiano settimanalmente non vengono riportate nei risultati; l'importo speso registrato corrisponde ora alla somma degli importi spesi durante le quattro settimane dello studio.
- **grocery\_coupons.sav.** File di dati ipotetici che contiene i dati sui sondaggi raccolti da una catena di drogherie interessata alle abitudini di acquisto dei suoi clienti. Ciascun cliente viene seguito per quattro settimane e ciascun caso corrisponde a una settimana per cliente con informazioni sul luogo degli acquisti e i tipi di acquisti, incluso l'importo speso nelle drogherie durante la settimana.
- **guttman.sav.** Bell (Bell, 1961) ha presentato una tabella per illustrare i possibili gruppi sociali. Guttman (Guttman, 1968) ha utilizzato una parte di tale tabella, in cui cinque variabili che descrivono elementi come l'interazione sociale, i sentimenti di appartenenza a un gruppo, la vicinanza fisica dei membri e il grado di formalità della relazione, sono state incrociate con cinque gruppi sociali teorici, compresi folla (ad esempio, le persone presenti a una partita di calcio), uditorio (ad esempio, di uno spettacolo teatrale o di una lezione universitaria), pubblico (ad esempio televisivo), calca (come una folla, ma con un'interazione molto maggiore), gruppi primari (intimi), gruppi secondari (volontari) e la comunità moderna (unione non stretta derivante da una vicinanza fisica elevata e dall'esigenza di servizi specializzati).
- **health\_funding.sav.** File di dati ipotetici che contiene i dati sui fondi di assistenza sanitaria (importo per 100 persone), sui tassi di malattie (tasso per 10.000 persone) e sulle visite ai fornitori di assistenza sanitaria (tasso per 10.000 persone). Ogni caso rappresenta una diversa città.
- **hivassay.sav.** File di dati ipotetici che prende in esame l'impegno di un'industria farmaceutica nel tentativo di sviluppare un'analisi che riesca a rilevare in tempi brevi l'infezione da virus HIV. I risultati dell'analisi sono otto sfumature di colore rosso sempre più intenso; le sfumature più intense indicano la maggiore probabilità di infezione. Un esperimento di laboratorio è stato condotto su 2000 campioni di sangue. La metà di questi è risultata infetta al virus HIV, l'altra metà non è risultata infetta.
- **hourlywagedata.sav.** File di dati ipotetici che prende in esame la paga oraria degli infermieri occupati presso uffici e ospedali e in base ai diversi livelli di esperienza.
- **insurance\_claims.sav.** File di dati ipotetici che prende in esame una compagnia di assicurazioni impegnata nella creazione di un modello per contrassegnare le richieste di risarcimento sospette e potenzialmente fraudolente. Ogni caso rappresenta una richiesta di risarcimento separata.
- **insure.sav.** File di dati ipotetici che prende in esame una compagnia di assicurazioni impegnata nello studio dei fattori di rischio, che indicano l'eventualità che un cliente presenti una domanda di indennizzo in un contratto assicurativo sulla vita della durata di dieci anni. Ogni caso nel file di dati rappresenta una coppia di contratti. In un contratto sono contenute informazioni su una richiesta di risarcimento, l'altro sull'età e sul sesso.

- **judges.sav.** File di dati ipotetici che prende in esame il punteggio assegnato, da giurie qualificate (più un appassionato) a 300 prestazioni sportive. Ciascuna riga rappresenta una diversa prestazione; i giudici hanno esaminato le stesse prestazioni.
- **kinship\_dat.sav.** Rosenberg e Kim (Rosenberg e Kim, 1975) si prefiggono di analizzare 15 termini indicanti parentela (zia, fratello, cugino, padre, nipote femmina, di nonni, nonno, nonna, nipote maschio di nonni, madre, nipote maschio di zii), nipote femmina di zii, sorella, figlio, zio). Hanno richiesto a quattro gruppi di studenti universitari (due composti da femmine e due da maschi) di ordinare questi termini in base alla similitudine. A due gruppi (uno femminile e uno maschile) è stato richiesto di effettuare l'ordinamento due volte, con il secondo ordinamento basato su un criterio diverso rispetto al primo. Di conseguenza, sono state ottenute sei "sorgenti" in totale. Ogni sorgente corrisponde a una matrice di prossimità  $15 \times 15$ , le cui celle sono uguali al numero delle persone in una sorgente meno il numero di volte in cui gli oggetti sono stati ripartiti insieme nella sorgente.
- **kinship\_ini.sav.** Questo file di dati contiene la configurazione iniziale di una soluzione a tre dimensioni per *kinship\_dat.sav*.
- **kinship\_var.sav.** Questo file di dati contiene variabili indipendenti relative a *sexo*, *generazione* e *grado* di separazione che possono essere utilizzate per interpretare le dimensioni di una soluzione per *kinship\_dat.sav*. In modo specifico, tali variabili possono essere utilizzate per limitare lo spazio della soluzione a una combinazione lineare di tali variabili.
- **marketvalues.sav.** File di dati che prende in esame le vendite di abitazioni in un nuovo centro abitato in Algonquin, Ill., durante gli anni 1999–2000. Tali vendite sono una questione di dominio pubblico.
- **nhis2000\_subset.sav.** Il National Health Interview Survey (NHIS) è un sondaggio di grandi dimensioni condotto sulla popolazione civile americana. Le interviste vengono realizzate di persona e si basano su un campione rappresentativo di famiglie a livello nazionale. Per ogni membro di una famiglia vengono raccolte osservazioni e informazioni di carattere demografico relative allo stato di salute. Questo file di dati contiene un sottoinsieme delle informazioni ottenute dall'indagine del 2000. National Center for Health Statistics. National Health Interview Survey, 2000. File di dati e documentazione di dominio pubblico. [ftp://ftp.cdc.gov/pub/Health\\_Statistics/NCHS/Datasets/NHIS/2000/](ftp://ftp.cdc.gov/pub/Health_Statistics/NCHS/Datasets/NHIS/2000/). Accesso 2003.
- **ozone.sav** I dati includono 330 osservazioni basate su sei variabili meteorologiche per quantificare la concentrazione dell'ozono dalle variabili rimanenti. I precedenti ricercatori, (Breiman e Friedman, 1985) e (Hastie e Tibshirani, 1990), hanno rilevato non linearità tra queste variabili, che impediscono un approccio di regressione standard.
- **pain\_medication.sav.** File di dati ipotetici che contiene i risultati di un test clinico per stabilire la cura antinfiammatoria per il trattamento del dolore generato dall'artrite cronica. Di particolare interesse, il test ha evidenziato il tempo che impiega il farmaco ad avere effetto e il confronto con altri farmaci esistenti.
- **patient\_los.sav.** File di dati ipotetici che contiene informazioni sul trattamento dei pazienti ricoverati per sospetto di infarto del miocardio. Ogni caso corrisponde a un diverso paziente e contiene diverse variabili correlate alla degenza nell'ospedale.
- **patlos\_sample.sav.** File di dati ipotetici che contiene informazioni sul trattamento di un campione di pazienti curato con trombolitici durante la degenza per infarto del miocardio. Ogni caso corrisponde a un diverso paziente e contiene diverse variabili correlate alla degenza nell'ospedale.

- **poll\_cs.sav.** File di dati ipotetici che prende in esame i sondaggi per stabilire il livello di sostegno pubblico nei confronti di un disegno di legge prima che diventi una legge vera e propria. I casi corrispondono ai votanti registrati. Ciascun caso riporta informazioni sulla contea, sul comune e sul quartiere in cui vive il votante.
- **poll\_cs\_sample.sav.** File di dati ipotetici che contiene un campione dei votanti elencati nel file *poll\_cs.sav*. Il campione è stato selezionato in base al disegno specificato nel file di piano *poll.csplan* e questo file di dati contiene le probabilità di inclusione e i pesi del campione. Tuttavia, notare che poiché fa uso del metodo PPS (probability-proportional-to-size, probabilità proporzionale alla dimensione), esiste anche un file contenente le probabilità di selezione congiunte (*poll\_jointprob.sav*). Le ulteriori variabili corrispondenti ai dati demografici dei votanti e alla loro opinione sul disegno di legge, sono state raccolte e aggiunte al file di dati dopo aver acquisito il campione.
- **property\_assess.sav.** File di dati ipotetici che prende in esame l'impegno di un perito di una contea nel tentativo di mantenere gli accertamenti sui valori delle proprietà aggiornati in base alle risorse limitate. I casi rappresentano le proprietà vendute nella contea nello scorso anno. Ogni caso nel file di dati contiene informazioni sul comune in cui si trova la proprietà, il perito che per ultimo ha visitato la proprietà, il tempo trascorso dall'accertamento, la valutazione fatta in tale momento e il valore di vendita della proprietà.
- **property\_assess\_cs.sav.** File di dati ipotetici che prende in esame l'impegno di un perito di uno stato nel tentativo di mantenere aggiornati gli accertamenti sui valori delle proprietà in base alle risorse limitate. I casi corrispondono alle proprietà nello stato. Ogni caso nel file di dati include informazioni sulla contea, il comune e il quartiere in cui risiede la proprietà, la data dell'ultimo accertamento e la valutazione fatta in tale data.
- **property\_assess\_cs\_sample.sav.** File di dati ipotetici che contiene un campione delle proprietà elencate nel file *property\_assess\_cs.sav*. Il campione è stato selezionato in base al disegno specificato nel file di piano *property\_assess.csplan* e questo file di dati contiene le probabilità di inclusione e i pesi del campione. L'ulteriore variabile *Valore corrente* è stata raccolta e aggiunta al file di dati dopo aver acquisito il campione.
- **recidivism.sav.** File di dati ipotetici che prende in esame l'impegno delle Forze dell'Ordine nel tentativo di valutare il tasso di recidività nella propria area di giurisdizione. Ogni caso corrisponde a un precedente trasgressore e include le informazioni demografiche, alcuni dettagli sul primo crimine, il tempo trascorso fino al secondo arresto e se tale arresto è avvenuto entro due anni dal primo.
- **recidivism\_cs\_sample.sav.** File di dati ipotetici che prende in esame l'impegno delle Forze dell'Ordine nel tentativo di valutare il tasso di recidività nella propria area di giurisdizione. Ogni caso corrisponde a un trasgressore precedente, rilasciato dopo il primo arresto durante il mese di giugno del 2003 e registra le relative informazioni demografiche, alcuni dettagli sul primo crimine commesso e i dati del secondo arresto, se si è verificato prima della fine di giugno del 2006. I trasgressori sono stati selezionati dai dipartimenti sottoposti a campione in base al piano di campionamento specificato nel file *recidivism\_cs.csplan*. Poiché viene utilizzato un metodo PPS (Probability-Proportional-to-Size, probabilità proporzionale alla dimensione), esiste anche un file contenente le probabilità di selezione congiunte (*recidivism\_cs\_jointprob.sav*).
- **rfm\_transactions.sav.** File di dati ipotetici contenente i dati delle transazioni di acquisto, inclusa la data di acquisto, gli articoli acquistati e il valore monetario di ciascuna transazione.

- **salesperformance.sav.** File di dati ipotetici che prende in esame la valutazione di due nuovi corsi di formazione alle vendite. Sessanta dipendenti, divisi in tre gruppi, ricevono tutti la formazione standard. In più, al gruppo 2 viene assegnato un corso di formazione tecnica e al gruppo 3 un'esercitazione pratica. Alla fine del corso di formazione, ciascun dipendente viene sottoposto a un esame e il punteggio conseguito viene registrato. Ciascun caso nel file di dati rappresenta un diverso partecipante. Il file di dati include il gruppo a cui è assegnato il partecipante e il punteggio conseguito all'esame finale.
- **satisf.sav.** File di dati ipotetico che prende in esame un'indagine sulla soddisfazione dei clienti condotta da una società di vendita al dettaglio presso 4 negozi. Sono stati intervistati 582 clienti e ciascun caso rappresenta le risposte ottenute da un singolo cliente.
- **screws.sav.** Questo file di dati contiene informazioni sulle caratteristiche di viti, bulloni, dadi e puntine (Hartigan, 1975).
- **shampoo\_ph.sav.** File di dati ipotetici che prende in esame il processo di controllo di qualità di un'industria di prodotti per capelli. A intervalli di tempo regolari, vengono misurati sei diversi lotti prodotti e ne viene registrato il relativo pH. I valori accettati sono compresi tra 4,5 e 5,5.
- **ships.sav.** Ad esempio, un insieme di dati presentato e analizzato altrove (McCullagh et al., 1989) riguarda i danni subiti dalle navi da carico a causa delle onde. I conteggi degli incidenti possono essere presentati con un tasso di Poisson in base al tipo di nave, al periodo di costruzione e al periodo di servizio. I mesi di servizio aggregati di ciascuna cella della tabella generata dalla classificazione incrociata dei fattori fornisce i valori di esposizione al rischio.
- **site.sav.** File di dati ipotetici che prende in esame l'impegno di una società nella scelta di nuovi siti in cui espandere la propria presenza. La società ha incaricato due consulenti separati che, oltre a valutare i siti e presentare un report completo, devono classificarli come potenzialmente "molto adatti", "adatti" o "poco adatti".
- **smokers.sav.** Questo file di dati è un estratto del 1998 National Household Survey of Drug Abuse e rappresenta un campione probabile di famiglie americane. (<http://dx.doi.org/10.3886/ICPSR02934>) Il primo passo nell'analisi di questo file di dati consiste quindi nel pesare i dati per rispecchiare le tendenze della popolazione.
- **stocks.sav** Questo file di dati ipotetici contiene i prezzi e i volumi delle scorte riferiti a un anno.
- **stroke\_clean.sav.** File di dati ipotetici che riporta lo stato di un database medico dopo averne eseguito la pulizia utilizzando le procedure del modulo Data Preparation.
- **stroke\_invalid.sav.** File di dati ipotetici che riporta lo stato iniziale di un database medico e contiene numerosi errori di immissione dati.
- **stroke\_survival.** Questo file di dati ipotetici riguarda i tempi di sopravvivenza per i pazienti che, dopo avere completato un programma riabilitativo in seguito a un ictus postischemico, affrontano alcune sfide. Dopo l'attacco, viene annotata l'occorrenza dell'infarto miocardico, dell'ictus ischemico o emorragico e viene registrata l'ora dell'evento. Questo campione viene troncato a sinistra perché include solo i pazienti che sono sopravvissuti fino alla fine del programma riabilitativo post-ictus.
- **stroke\_valid.sav.** File di dati ipotetici che riporta lo stato di un database medico dopo il controllo dei valori eseguito con la procedura Convalida i dati. Il database contiene comunque casi potenzialmente anomali.

- **survey\_sample.sav.** File di dati che contiene i dati dell'indagine, compresi i dati demografici e varie misure dell'atteggiamento. Si basa su un sottoinsieme di variabili tratte dal 1998 NORC General Social Survey, benché i valori di alcuni dati siano stati modificati e siano state aggiunte variabili fittizie a scopo dimostrativo.
- **telco.sav.** File di dati ipotetici che prende in esame l'impegno di un'azienda di telecomunicazioni nel tentativo di ridurre il churn, ovvero l'abbandono dei propri clienti. Ciascun caso rappresenta un cliente separato e riporta diverse informazioni demografiche e sull'uso del servizio.
- **telco\_extra.sav.** Questo file di dati è simile al file *telco.sav*, ma le variabili "tenure" e spesa del cliente trasformata tramite logaritmo sono state sostituite dalle variabili di spesa del cliente trasformata tramite logaritmo standardizzate.
- **telco\_missing.sav.** Questo file di dati è un sottoinsieme del file di dati *telco.sav*, ma alcuni dei valori di dati demografici sono stati sostituiti con valori mancanti.
- **testmarket.sav.** File di dati ipotetici che prende in esame i piani di una catena di fast food per aggiungere un nuovo prodotto al proprio menu. Sono previste tre campagne promozionali del nuovo prodotto. Il prodotto viene introdotto in diversi mercati selezionati in modo casuale. Per ogni sede viene utilizzata una promozione differente registrando le vendite settimanali della nuova voce per le prime quattro settimane. Ogni caso rappresenta un luogo e una settimana diversi.
- **testmarket\_1month.sav.** Questo file di dati ipotetici corrisponde al file *testmarket.sav* con le vendite settimanali organizzate in modo che ogni caso corrisponda a un luogo separato. Alcune delle variabili che cambiano settimanalmente non vengono riportate nei risultati; le vendite registrate corrispondono ora alla somma delle vendite conseguite durante le quattro settimane dello studio.
- **tree\_car.sav.** File di dati ipotetici che contiene dati demografici e sul prezzo di acquisto dei veicoli.
- **tree\_credit.sav.** File di dati ipotetici che contiene dati demografici e sulla cronologia dei mutui di una banca.
- **tree\_missing\_data.sav.** File di dati ipotetici che contiene dati demografici e sulla cronologia dei mutui di una banca con un numero elevato di valori mancanti.
- **tree\_score\_car.sav.** File di dati ipotetici che contiene dati demografici e sul prezzo di acquisto dei veicoli.
- **tree\_textdata.sav.** File di dati semplice con due variabili destinato principalmente per mostrare lo stato predefinito delle variabili prima dell'assegnazione dei livelli di misurazione e delle etichette dei valori.
- **tv-survey.sav.** File di dati ipotetici che prende in esame un sondaggio condotto da una emittente televisiva che deve stabilire se estendere la durata di un programma di successo. A un campione di 906 intervistati è stato chiesto se preferisce guardare il programma con diverse condizioni. Ciascuna riga rappresenta un diverso intervistato e ciascuna colonna una diversa condizione.
- **ulcer\_recurrence.sav.** Questo file contiene informazioni parziali su uno studio svolto per mettere a confronto l'efficacia di due terapie preventive per la recidiva delle ulcere. Fornisce un ottimo esempio di dati acquisiti a intervalli ed è stato presentato e analizzato in altri luoghi (Collett, 2003).

- **ulcer\_recurrence\_recoded.sav.** In questo file sono contenute le informazioni del file *ulcer\_recurrence.sav* riorganizzate per consentire di presentare la probabilità degli eventi per ciascun intervallo dello studio, anziché solo alla fine. È stato presentato e analizzato in altri luoghi (Collett et al., 2003).
- **verd1985.sav.** Questo file di dati prende in esame un'indagine (Verdegaal, 1985). Sono state registrate le risposte di quindici soggetti a otto variabili. Le variabili di interesse sono suddivise in tre insiemi. L'insieme 1 include *età* e *statociv*, l'insieme 2 include *andom* e *giornale* e l'insieme 3 include *musica* e *vicinato*. *Andom* viene scalata come nominale multipla ed *età* come ordinale; tutte le altre variabili vengono scalate come nominali singole.
- **virus.sav.** File di dati ipotetici che prende in esame l'impegno di un ISP (Internet Service Provider) nel tentativo di determinare gli effetti che un virus può generare nelle sue reti. Si è tenuta traccia della percentuale (approssimativa) di traffico e-mail infettato da virus sulla rete in un lasso di tempo, dal momento dell'individuazione fino alla soppressione della minaccia.
- **wheeze\_steubenville.sav.** Questo file è un sottoinsieme di uno studio longitudinale degli effetti che l'inquinamento provoca sulla salute dei bambini (Ware, Dockery, Spiro III, Speizer, e Ferris Jr., 1984). I dati contengono misure binarie ripetute del livello di asma dei bambini della città di Steubenville, Ohio, di 7, 8, 9 e 10 anni. I dati indicano anche se la mamma dei bambini era fumatrice durante il primo anno dello studio.
- **workprog.sav.** File di dati ipotetici che prende in esame un programma di lavoro governativo il cui obiettivo è fornire attività più adatte alle persone diversamente abili. È stato seguito un campione di potenziali partecipanti al programma, alcuni dei quali sono stati selezionati in modo casuale e altri no. Ogni caso rappresenta un diverso partecipante al programma.
- **worldsales.sav** Questo file di dati ipotetici contiene i ricavi suddivisi per continenti e prodotti.

## ***Note legali***

Queste informazioni sono state preparate per prodotti e servizi offerti in tutto il mondo.

IBM potrebbe non offrire i prodotti, i servizi o le funzionalità di cui si tratta nel presente documento in altri paesi. Contattare il rappresentante IBM locale per informazioni sui prodotti e i servizi attualmente disponibili nella propria zona. Qualsiasi riferimento a un prodotto, programma o servizio IBM non intende dichiarare o implicare che sia possibile utilizzare esclusivamente tale prodotto, programma o servizio IBM. Potrà invece essere utilizzato qualsiasi prodotto, programma o servizio con funzionalità equivalente e che non violi i diritti di proprietà intellettuale di IBM. Tuttavia, è responsabilità dell'utente valutare e verificare il funzionamento di qualsiasi prodotto, programma o servizio non IBM.

IBM può essere titolare di brevetti o domande di brevetto relativi alla materia oggetto del presente documento. La consegna del presente documento non conferisce alcuna licenza rispetto a questi brevetti. Rivolgere per iscritto i quesiti sulle licenze a:

*IBM Director of Licensing, IBM Corporation, North Castle Drive, Armonk, NY 10504-1785, U.S.A.*

Per richieste di informazioni sulle licenze riguardanti il set di caratteri a byte doppio (DBCS), contattare l'Intellectual Property Department di IBM del proprio paese, oppure inviare le richieste in forma scritta all'indirizzo:

*Intellectual Property Licensing, Legal and Intellectual Property Law, IBM Japan Ltd., 1623-14, Shimotsuruma, Yamato-shi, Kanagawa 242-8502 Giappone.*

**Il seguente paragrafo non si applica per il Regno Unito o altri paesi in cui le presenti disposizioni non sono conformi alle leggi locali:** INTERNATIONAL BUSINESS MACHINES FORNISCE QUESTA PUBBLICAZIONE “COSÌ COM'È” SENZA GARANZIA DI ALCUN TIPO, SIA ESSA ESPRESSA O IMPLICITA, INCLUSE, MA NON LIMITATE A, LE GARANZIE IMPLICITE DI NON VIOLAZIONE, COMMERCIALIZZABILITÀ O IDONEITÀ A UNO SCOPO SPECIFICO. Alcuni stati non consentono limitazioni di garanzie espresse o implicite in determinate transazioni, pertanto quanto sopra potrebbe non essere applicabile.

Le presenti informazioni possono includere imprecisioni tecniche o errori tipografici. Le modifiche periodiche apportate alle informazioni contenute in questa pubblicazione verranno inserite nelle nuove edizioni della pubblicazione. IBM può apportare miglioramenti e/o modifiche al/ai prodotto/i e/o al/ai programma/i descritti nella presente pubblicazione in qualsiasi momento senza preavviso.

Qualsiasi riferimento nelle presenti informazioni a siti Web non IBM viene fornito esclusivamente per facilitare la consultazione e non rappresenta in alcun modo un'approvazione o sostegno da parte nostra di tali siti Web. I materiali contenuti in tali siti Web non fanno parte dei materiali di questo prodotto IBM e il loro utilizzo è esclusivamente a rischio dell'utente.

IBM può utilizzare o distribuire eventuali informazioni fornite dall'utente nei modi che ritiene appropriati senza incorrere in alcun obbligo nei confronti dell'utente.

I licenziatari del programma che desiderassero informazioni su di esso allo scopo di abilitare: (i) lo scambio di informazioni tra programmi creati indipendentemente e altri programmi (questo compreso) e (ii) l'utilizzo in comune delle informazioni scambiate, dovranno rivolgersi a:

*IBM Software Group, All'attenzione di: Licensing, 233 S. Wacker Dr., Chicago, IL 60606, USA.*

Tali informazioni saranno fornite in conformità ai termini e alle condizioni in vigore e, in alcuni casi, dietro pagamento.

Il programma concesso in licenza descritto nel presente documento e tutto il materiale correlato disponibile sono forniti da IBM in base ai termini del contratto di licenza cliente IBM, del contratto di licenza internazionale IBM o del contratto equivalente esistente tra le parti.

le informazioni relative a prodotti non IBM sono state ottenute dai fornitori di tali prodotti, da loro annunci pubblicati e da altre fonti disponibili al pubblico. IBM non ha verificato tali prodotti e non può confermare l'accuratezza delle prestazioni, la compatibilità o qualsiasi altra dichiarazione relativa a prodotti non IBM. Eventuali domande in merito alle funzionalità dei prodotti non IBM vanno indirizzate ai fornitori di tali prodotti.

Le presenti informazioni includono esempi di dati e report utilizzati in operazioni aziendali quotidiane. Per fornire una descrizione il più possibile esaustiva, gli esempi includono nomi di persone, società, marchi e prodotti. Tutti questi nomi sono fittizi e ogni somiglianza a nomi e indirizzi utilizzati da aziende reali è puramente casuale.

Per chi visualizza queste informazioni a video: le fotografie e le illustrazioni a colori potrebbero non essere disponibili.

### ***Marchi commerciali***

IBM, il logo IBM, ibm.com e SPSS sono marchi di IBM Corporation, registrati in numerose giurisdizioni nel mondo. Un elenco aggiornato dei marchi IBM è disponibile sul Web all'indirizzo <http://www.ibm.com/legal/copytrade.shtml>.

Adobe, il logo Adobe, PostScript e il logo PostScript sono marchi o marchi registrati di Adobe Systems Incorporated negli Stati Uniti e/o in altri paesi.

Intel, il logo Intel, Intel Inside, il logo Intel Inside, Intel Centrino, il logo Intel Centrino, Celeron, Intel Xeon, Intel SpeedStep, Itanium e Pentium sono marchi o marchi registrati di Intel Corporation o delle sue consociate negli Stati Uniti e in altri paesi.

Java e tutti i marchi e i logo basati su Java sono marchi di Sun Microsystems, Inc. negli Stati Uniti e/o negli altri paesi.

Linux è un marchio registrato di Linus Torvalds negli Stati Uniti e/o negli altri paesi.

Microsoft, Windows, Windows NT e il logo Windows sono marchi di Microsoft Corporation negli Stati Uniti e/o negli altri paesi.

UNIX è un marchio registrato di The Open Group negli Stati Uniti e in altri paesi.

Questo prodotto utilizza WinWrap Basic, Copyright 1993-2007, Polar Engineering and Consulting, <http://www.winwrap.com>.

Altri nomi di prodotti e servizi possono essere marchi commerciali di IBM o di altre aziende.

Le schermate dei prodotti Adobe sono state ristampate su autorizzazione di Adobe Systems Incorporated.

Le schermate dei prodotti Microsoft sono state ristampate su autorizzazione di Microsoft Corporation.



---

# Bibliografia

- Akaah, I. P., e P. K. Korgaonkar. 1988. A conjoint investigation of the relative importance of risk relievers in direct marketing. *Journal of Advertising Research*, 28:4, .
- Bell, E. H. 1961. *Social foundations of human behavior: Introduction to the study of sociology*. New York: Harper & Row.
- Blake, C. L., e C. J. Merz. 1998. "UCI Repository of machine learning databases." Available at <http://www.ics.uci.edu/~mllearn/MLRepository.html>.
- Breiman, L., e J. H. Friedman. 1985. Estimating optimal transformations for multiple regression and correlation. *Journal of the American Statistical Association*, 80, .
- Cattin, P., e D. R. Wittink. 1982. Commercial use of conjoint analysis: A survey. *Journal of Marketing*, 46:3, .
- Collett, D. 2003. *Modelling survival data in medical research*, 2 ed. Boca Raton: Chapman & Hall/CRC.
- Green, P. E., e V. Rao. 1972. *Applied multidimensional scaling*. Hinsdale, Ill.: Dryden Press.
- Green, P. E., e Y. Wind. 1973. *Multiattribute decisions in marketing: A measurement approach*. Hinsdale, Ill.: Dryden Press.
- Guttman, L. 1968. A general nonmetric technique for finding the smallest coordinate space for configurations of points. *Psychometrika*, 33, .
- Hartigan, J. A. 1975. *Clustering algorithms*. New York: John Wiley and Sons.
- Hastie, T., e R. Tibshirani. 1990. *Generalized additive models*. London: Chapman and Hall.
- Kennedy, R., C. Riquier, e B. Sharp. 1996. Practical applications of correspondence analysis to categorical data in market research. *Journal of Targeting, Measurement and Analysis for Marketing*, 5, .
- McCullagh, P., e J. A. Nelder. 1989. *Generalized Linear Models*, 2nd ed. London: Chapman & Hall.
- Prezzo, R. H., e D. L. Bouffard. 1974. Behavioral appropriateness and situational constraints as dimensions of social behavior. *Journal of Personality and Social Psychology*, 30, .
- Rickman, R., N. Mitchell, J. Dingman, e J. E. Dalen. 1974. Changes in serum cholesterol during the Stillman Diet. *Journal of the American Medical Association*, 228, .
- Rosenberg, S., e M. P. Kim. 1975. The method of sorting as a data-gathering procedure in multivariate research. *Multivariate Behavioral Research*, 10, .
- Van der Ham, T., J. J. Meulman, D. C. Van Strien, e H. Van Engeland. 1997. Empirically based subgrouping of eating disorders in adolescents: A longitudinal perspective. *British Journal of Psychiatry*, 170, .
- Verdegaal, R. 1985. *Meer sets analyse voor kwalitatieve gegevens (in Dutch)*. Leiden: Department of Data Theory, University of Leiden.
- Ware, J. H., D. W. Dockery, A. Spiro III, F. E. Speizer, e B. G. Ferris Jr.. 1984. Passive smoking, gas cooking, and respiratory health of children living in six cities. *American Review of Respiratory Diseases*, 129, .

---

# Indice

approccio con profilo completo, 2  
array ortogonale, 2

## *J*CARD

nella visualizzazione disegno, 10  
casi di controllo, 2  
in Genera disegno ortogonale, 7  
casi di simulazione, 3, 14, 29  
in Genera disegno ortogonale, 8  
coefficienti, 27  
coefficienti di correlazione, 28

dimensione campione, 3  
disegni ortogonali  
casi di controllo, 7  
generazione, 5, 17  
numero minimo di casi, 7  
visualizzazione, 9, 21

fattori, 2, 17  
file di dati  
in Genera disegno ortogonale, 5  
file di esempio  
posizione, 31  
formato a profilo singolo  
nella visualizzazione disegno, 9  
formato di elenco  
nella visualizzazione disegno, 9

Genera disegno ortogonale, 2, 5, 17  
casi di controllo, 7  
casi di simulazione, 8  
definizione di nomi di fattori, etichette e valori, 7  
file di dati, 5  
numero minimo di casi, 7  
seme dei numeri casuali, 5

inversioni, 28

livelli di fattore, 2, 17

marchi commerciali, 42  
modello antiideale, 25  
Modello Bradley-Terry-Luce (BTL), 30  
modello di utilità massima, 30  
modello discreto, 25  
modello ideale, 25  
modello lineare, 25  
modello logit, 30

note legali, 41

piè di pagina  
nella visualizzazione disegno, 10  
punteggi di importanza, 27  
punteggi di utilità, 26  
punteggi di utility, 3

*R* di Pearson, 28  
risultati della simulazione  
Modello Bradley-Terry-Luce (BTL), 30  
modello di utilità massima, 30  
modello logit, 30

seme dei numeri casuali  
in Genera disegno ortogonale, 5  
sintassi  
comando CONJOINT, 24  
sintassi dei comandi  
comando CONJOINT, 24

tau di Kendall, 28  
titoli  
nella visualizzazione disegno, 10

utilità totale, 26

valori delle parti, 3  
valori di importanza, 27  
*variabile\_scheda*  
in Genera disegno ortogonale, 5  
*variabile\_stato*  
in Genera disegno ortogonale, 5  
Visualizza disegno, 2, 9, 21  
*J*CARD, 10  
formato a profilo singolo, 9  
formato di elenco, 9  
piè di pagina, 10  
salvataggio dei profili, 10  
titoli, 10