

*IBM SPSS Complex Samples 31*



**Nota**

Prima di utilizzare queste informazioni e il prodotto che supportano, leggere le informazioni in [“Informazioni particolari” a pagina 41.](#)

**Informazioni sul prodotto**

Questa edizione si applica alla versione 31, release 0, modifica 1 di IBM® SPSS Statistica e a tutte le release e modifiche successive se non diversamente indicato nelle nuove edizioni.

© Copyright International Business Machines Corporation .

---

# Indice

<b>Capitolo 1. Campioni complessi.....</b>	<b>1</b>
Introduzione alle procedure di Campionamento Complesso.....	1
Proprietà dei Campioni Complessi.....	1
Utilizzo delle procedure di Campionamento Complesso.....	2
Ulteriori Letture.....	2
Campionamento da un Complesso Design.....	3
Creazione di un nuovo piano di campionamento.....	3
Procedura guidata di campionamento: Variabili di progettazione.....	3
Procedura guidata di campionamento: Metodo di campionamento.....	4
Procedura guidata di campionamento: Dimensione di esempio.....	5
Procedura guidata di campionamento: Variabili di output.....	6
Procedura guidata di campionamento: Riepilogo piano.....	6
Procedura guidata di campionamento: Opzioni di selezione di esempio Draw.....	6
Procedura guidata di campionamento: File di output di esempio Draw.....	6
Procedura guidata di campionamento: Fine.....	7
Modifica di un piano di esempio esistente.....	7
Procedura guidata di campionamento: Riepilogo piano.....	7
Esecuzione di un piano di campionamento esistente.....	7
Comandi aggiuntivi CSPLAN e CSSELECT.....	8
Preparazione di un esempio Complesso per l'analisi.....	8
Creazione di un nuovo piano di analisi.....	8
Procedura guidata di Preparazione analisi: Variabili di progettazione.....	8
Procedura guidata di Preparazione analisi: Metodo di stima.....	9
Procedura guidata Preparazione analisi: Dimensioni.....	9
Procedura guidata di Preparazione analisi: Riepilogo piano.....	10
Procedura guidata di Preparazione analisi: Finish.....	10
Modifica di un piano di analisi esistente.....	11
Procedura guidata di Preparazione analisi: Riepilogo piano.....	11
Piano campioni complessi.....	11
Frequenze campioni complessi.....	11
Complesso Statistiche Frequenze Statistiche.....	12
Complessi Campioni Mancanti.....	12
Opzioni Campioni Complessi.....	13
Descrittive campioni complessi.....	13
Statistiche Descrittive Campioni Statistici.....	13
Complessi Campioni Descrittivi Mancanti Valori.....	14
Opzioni Campioni Complessi.....	14
Tabelle di contingenza campioni complessi.....	14
Complessi Campioni Crosstabs statistiche.....	15
Complessi Campioni Mancanti.....	16
Opzioni Campioni Complessi.....	16
Rapporti campioni complessi.....	16
Statistiche Rapporti di Campioni Complessi.....	17
Complessi Campioni Rapporti Mancanti.....	17
Opzioni Campioni Complessi.....	17
Modello lineare generalizzato campioni complessi.....	18
Modello lineare generalizzato campioni complessi.....	18
Complesso Campioni General Linear Model Statistics.....	19
Test Ipotesi di Campioni Complessi.....	19
Complesso Campioni General Linear Model Stimato Mezzo.....	20
Complesso Campioni General Linear Model Save.....	20

Complesso Campioni General Linear Model Options.....	21
Funzioni Aggiuntive di comando CSGLM.....	21
Regressione logistica campioni complessi.....	21
Complesso Campioni Logistic Regression Reference Categoria.....	22
Modello Complesso di regressione logistica Modello.....	22
Statistiche Complesse di regressione logistica.....	23
Test Ipotesi di Campioni Complessi.....	24
Rapporti Complessi Logistic Regression Odds Ratio.....	24
Complesso Campionamento Logistic Regression Save.....	24
Opzioni Complesse di regressione logistica.....	25
CSLOGISTIC Comando funzioni aggiuntive.....	25
Regressione ordinale campioni complessi.....	26
Complesso Campioni Ordinale Regressione Risposta.....	27
Modello Complesso di regressione ordinale.....	27
Complesso Campioni Ordinali Statistiche di regressione.....	27
Test Ipotesi di Campioni Complessi.....	28
Complesso Campioni Ordinali Regression Odds Ratio.....	29
Complesso Campioni Ordinale Regression Save.....	29
Opzioni Ordinali di regressione ordinale.....	30
Funzioni Aggiuntive di comando CSORDINAL.....	30
Regressione di Cox di campioni complessi.....	31
Definire l'evento .....	32
Predittori .....	32
Sottogruppi .....	33
Modello .....	33
Statistica .....	34
Grafici .....	35
Test di ipotesi .....	35
Salva .....	36
Esportazione .....	37
Opzioni .....	38
Funzioni Aggiuntive Comando CSCOXREG.....	38
<b>Informazioni particolari.....</b>	<b>41</b>
Marchi.....	42
<b>Indice analitico.....</b>	<b>45</b>

# Capitolo 1. Campioni complessi

Le seguenti funzioni di campioni complessi sono incluse in SPSS Statistiche Premium Edition o l'opzione Complex Samples.

## Introduzione alle procedure di Campionamento Complesso

Un'assunzione intrinseca di procedure analitiche nei pacchetti software tradizionali è che le osservazioni in un file di dati rappresentino un semplice campione casuale dalla popolazione di interesse. Questo presupposto è insostenibile per un numero crescente di aziende e ricercatori che lo trovano sia conveniente che conveniente per ottenere campioni in modo più strutturato.

L'Opzione Complex Campioni consente di selezionare un campione secondo un design complesso e di incorporare le specifiche di progettazione nell'analisi dei dati, garantendo così che i vostri risultati siano validi.

## Proprietà dei Campioni Complessi

Un campione complesso può differire da un semplice campione casuale in molti modi. In un semplice campione casuale, le singole unità di campionamento vengono selezionate a caso con pari probabilità e senza sostituzione (WOR) direttamente da tutta la popolazione. Al contrario, un determinato campione complesso può avere alcune o tutte le seguenti caratteristiche:

**Stratificazione.** Il campionamento stratificato comporta la selezione dei campioni in modo indipendente all'interno di sottogruppi non sovrapposti della popolazione, o strati. Ad esempio, gli strati possono essere gruppi socio - economici, categorie di lavoro, gruppi di età o gruppi etnici. Con la stratificazione è possibile garantire adeguate dimensioni del campione per i sottogruppi di interesse, migliorare la precisione delle stime complessive e utilizzare diversi metodi di campionamento da stratum a stratum.

**Raggruppamento.** Il campionamento del cluster comporta la selezione di gruppi di unità di campionamento, o cluster. Ad esempio, i cluster possono essere scuole, ospedali o aree geografiche, e le unità di campionamento possono essere studenti, pazienti o cittadini. Il clustering è comune nei disegni multistage e nei campioni di area (geografica).

**Più fasi.** Nel campionamento in multistage, si seleziona un campione di prima fase basato su cluster. Poi si crea un campione di seconda fase disegnando sottocampioni dai cluster selezionati. Se il campione di seconda fase è basato su sottocluster, è quindi possibile aggiungere una terza fase al campione. Ad esempio, nella prima fase di un sondaggio potrebbe essere tratto un campione di città. Poi, dalle città selezionate, le famiglie potrebbero essere campionate. Infine, dalle famiglie selezionate, gli individui potrebbero essere interpellati. le wizardi di Preparazione e analisi di analisi consentono di specificare tre fasi in un design.

**Campi non casuali.** Se è difficile ottenere la selezione casuale, è possibile campionare le unità in modo sistematico (a un intervallo fisso) o sequenziale.

**Probabilità di selezione non uguale.** Quando i cluster di campionamento che contengono numeri non uguali di unità, è possibile utilizzare il campionamento di probabilità - proporzionale (PPS) per rendere una probabilità di selezione del cluster uguale alla proporzione di unità che contiene. Il campionamento SPA può anche utilizzare schemi di ponderazione più generali per selezionare le unità.

**Campi di campionamento illimitati.** Il campionamento senza restrizioni seleziona unità con sostituzione (WR). Così, una singola unità può essere selezionata per il campione più di una volta.

**Pesi di campionamento.** I pesi di campionamento vengono calcolati automaticamente mentre disegnano un campione complesso e corrispondono idealmente alla "frequenza" che ogni unità di campionamento rappresenta nella popolazione target. Pertanto, la somma dei pesi sul campione dovrebbe stimare la dimensione della popolazione. Complesse Procedure di analisi dei campioni richiedono pesi di campionamento al fine di analizzare correttamente un campione complesso. Si noti che questi pesi dovrebbero essere utilizzati interamente all'interno dell'opzione Complex Campioni e non devono essere

utilizzati con altre procedure analitiche tramite la procedura di Weight Cases, che tratta i pesi come repliche di caso.

## Utilizzo delle procedure di Campionamento Complesso

Il tuo utilizzo delle procedure Complesse Campioni dipende dalle tue particolari esigenze. I tipi primari di utenti sono coloro che:

- Pianificare e realizzare sondaggi secondo disegni complessi, eventualmente analizzando il campione in seguito. Lo strumento principale per gli sorveglianti è il [Wizard di campionamento](#).
- Analizzare file di dati di esempio precedentemente ottenuti in base a disegni complessi. Prima di utilizzare le procedure di analisi dei Campioni Complessi, potrebbe essere necessario utilizzare il [Wizard preparazione analisi](#).

Indipendentemente dal tipo di utente che si è, è necessario fornire informazioni di progettazione alle procedure Complesse Campioni. Queste informazioni vengono memorizzate in un **file di piano** per un facile riutilizzo.

### File di piano

Un file di piano contiene specifiche di esempio complesse. Esistono due tipi di file di piano:

**Piano di campionamento.** Le specifiche riportate nella procedura guidata di Sampling definiscono un design di esempio utilizzato per disegnare un campione complesso. Il file del piano di campionamento contiene quelle specifiche. Il file del piano di campionamento contiene anche un piano di analisi predefinito che utilizza metodi di stima adatti al design di esempio specificato.

**Piano di analisi.** Questo file di piano contiene le informazioni necessarie alle procedure di analisi di Complex Campioni per le stime di varianza correttamente compatte per un campione complesso. Il piano include la struttura del campione, i metodi di stima per ogni fase e i riferimenti alle variabili richieste, come ad esempio i pesi campione. Il Wizard di preparazione dell'analisi consente di creare e modificare i piani di analisi.

Ci sono diversi vantaggi per salvare le tue specifiche in un file di piano, tra cui:

- Un sorvegliante può specificare la prima fase di un piano di campionamento multistadio e disegnare ora unità di prima fase, raccogliere informazioni sulle unità di campionamento per la seconda fase, quindi modificare il piano di campionamento per includere la seconda fase.
- Un analista che non ha accesso al file del piano di campionamento può specificare un piano di analisi e fare riferimento a tale piano da ogni procedura di analisi dei Campioni Complessi.
- Un progettista di campioni di uso pubblico su larga scala può pubblicare il file del piano di campionamento, che semplifica le istruzioni per gli analisti ed evita la necessità di ogni analista di specificare i propri piani di analisi.

## Ulteriori Letture

Per maggiori informazioni sulle tecniche di campionamento, consultare i seguenti testi:

W. G. Cochran 1977. *Tecniche di campionamento*, 3rd ed. New York: John Wiley e Sons.

L. Kish 1965. *Survey Sampling*. New York: John Wiley and Sons.

L. Kish 1987. *Statistical Design for Research*. New York: John Wiley and Sons.

M. N. Murthy 1967. *Sampling Theory and Methods*. Calcutta, India: Statistical Publishing Society.

Särndal, C., B. Swensson, e J. Wretman. 1992. *Model Assisted Survey Sampling*. New York: Springer-Verlag.

# Campionamento da un Complesso Design

---

La Procedura Guidata di campionamento ti guida attraverso le fasi per la creazione, la modifica o l'esecuzione di un file di piano di campionamento. Prima di utilizzare il Wizard, si dovrebbe avere una popolazione target ben definita, un elenco di unità di campionamento e un adeguato design di campione in mente.

## Creazione di un nuovo piano di campionamento

1. Dai menu, scegliere:

**Analizzare > Campioni complessi > Selezionare un esempio ...**

2. Selezionare **Design a sample** e scegliere un nomefile di piano per salvare il piano di esempio.
3. Clicca su **Avanti** per continuare attraverso il Wizard.
4. Opzionalmente, nella fase Design Variables, è possibile definire i pesi di esempio di strati, cluster e input. Dopo aver definito questi, fare clic su **Avanti**.
5. Opzionalmente, nella fase di Sampling Method, è possibile scegliere un metodo per la selezione degli articoli.

Se si seleziona **PPS Brewer** o **PPS Murthy**, è possibile fare clic su **Fine** per disegnare il campione. In caso contrario, fare clic su **Avanti** e quindi:

6. Nel passo Dimensioni di esempio specificare il numero o la proporzione di unità da campionare.
7. Ora è possibile fare clic su **Fine** per disegnare il campione.

Opzionalmente, in ulteriori passi è possibile:

- Scegliere le variabili di output da salvare.
- Aggiungere una seconda o terza fase al disegno.
- Impostare varie opzioni di selezione, tra cui le fasi per disegnare i campioni da, il seme del numero casuale e se trattare i valori mancanti dell'utente come valori validi delle variabili di progettazione.
- Scegli dove salvare i dati di output.
- Incollare le selezioni come sintassi del comando.

## Procedura guidata di campionamento: Variabili di progettazione

Questo passo consente di selezionare le variabili di stratificazione e clustering e di definire i pesi del campione di input. È anche possibile specificare un'etichetta per la fase.

**Stratifica per.** La classificazione incrociata delle variabili di stratificazione definisce sottopopolazioni distinte, o strati. Campioni separati sono ottenuti per ogni strato. Per migliorare la precisione delle vostre stime, le unità all'interno degli strati dovrebbero essere il più possibile omogenee per le caratteristiche di interesse.

**Cluster.** Le variabili cluster definiscono gruppi di unità osservazionali, o cluster. I cluster sono utili quando le unità osservative di campionamento direttamente dalla popolazione sono costose o impossibili; si possono invece assaggiare i cluster dalla popolazione e poi campionare unità osservative dai cluster selezionati. Tuttavia, l'utilizzo di cluster può introdurre correlazioni tra le unità di campionamento, con conseguente perdita di precisione. Per minimizzare questo effetto, le unità all'interno di cluster dovrebbero essere il più eterogenee possibile per le caratteristiche di interesse. È necessario definire almeno una variabile cluster al fine di pianificare un design multistage. I cluster sono inoltre necessari nell'utilizzo di diversi metodi di campionamento differenti. Per ulteriori informazioni, consultare l'argomento ["Procedura guidata di campionamento: Metodo di campionamento"](#) a pagina 4.

**Peso del campione di input.** Se l'attuale design del campione fa parte di un design di esempio più grande, si possono avere pesi di esempio da una fase precedente del design più grande. È possibile specificare una variabile numerica contenente questi pesi nella prima fase del design attuale. I pesi di esempio vengono calcolati automaticamente per le fasi successive del design attuale.

**Etichetta di stage.** È possibile specificare un'etichetta di stringa opzionale per ogni fase. Questo viene utilizzato nell'output per aiutare a identificare le informazioni stagne.

*Nota:* l'elenco di variabili di origine ha lo stesso contenuto in tutte le fasi della procedura guidata. In altre parole, le variabili eliminate dall'elenco di origine in un determinato passo vengono eliminate dall'elenco in tutti i passaggi. Le variabili restituite all'elenco di origine appaiono nella lista in tutti i passaggi.

## Albero Controlli per Navigare il Mago di Sampling

Sul lato sinistro di ogni passo del Mago di Sampling è uno schema di tutti i passaggi. È possibile navigare il Wizard cliccando sul nome di un passo abilitato nello schema. I passaggi sono abilitati fino a quando tutti i passaggi precedenti sono validi - vale a dire, se ad ogni passaggio precedente sono state date le specifiche minime richieste per quel passaggio. Consultare la Guida per i singoli passaggi per ulteriori informazioni sul motivo per cui un determinato passo potrebbe essere non valido.

## Procedura guidata di campionamento: Metodo di campionamento

Questo passo consente di specificare come selezionare i casi dal dataset attivo.

**Metodo.** I controlli in questo gruppo sono utilizzati per scegliere un metodo di selezione. Alcuni tipi di campionamento consentono di scegliere se sample con sostituzione (WR) o senza sostituzione (WOR). Consultare le descrizioni di tipo per ulteriori informazioni. Si noti che alcuni tipi di probabilità - proporzionale (PPS) sono disponibili solo quando i cluster sono stati definiti e che tutti i tipi di PPS sono disponibili solo nella prima fase di un design. Inoltre i metodi WR sono disponibili solo nell'ultima fase di un design.

- **Campionamento casuale semplice.** Le unità sono selezionate con uguale probabilità. Possono essere selezionati con o senza sostituzione.
- **Semplice sistematico.** Le unità vengono selezionate ad un intervallo fisso in tutta la cornice di campionamento (o strati, se sono state specificate) ed estratte senza sostituzione. Un'unità selezionata casualmente all'interno del primo intervallo viene scelta come punto di partenza.
- **Sequenziale semplice.** Le unità vengono selezionate sequenzialmente con pari probabilità e senza sostituzione.
- **SPA.** Si tratta di un metodo di prima fase che seleziona unità a caso con probabilità proporzionale alle dimensioni. Eventuali unità possono essere selezionate con sostituzione; solo i cluster possono essere campionati senza sostituzione.
- **PPS Sistematica.** Si tratta di un metodo di prima fase che seleziona sistematicamente unità con probabilità proporzionale alle dimensioni. Sono selezionati senza sostituzione.
- **Sequenziale di PPS.** Si tratta di un metodo di prima fase che seleziona sequenzialmente le unità con probabilità proporzionale alla dimensione del cluster e senza sostituzione.
- **PPS Brewer.** Si tratta di un metodo di prima fase che seleziona due cluster da ogni strato con probabilità proporzionale alla dimensione del cluster e senza sostituzione. Una variabile cluster deve essere specificata per utilizzare questo metodo.
- **PPS Murthy.** Si tratta di un metodo di prima fase che seleziona due cluster da ogni strato con probabilità proporzionale alla dimensione del cluster e senza sostituzione. Una variabile cluster deve essere specificata per utilizzare questo metodo.
- **Sampford SPA.** Si tratta di un metodo di prima fase che seleziona più di due cluster da ogni strato con probabilità proporzionale alla dimensione del cluster e senza sostituzione. Si tratta di un'estensione del metodo di Brewer. Una variabile cluster deve essere specificata per utilizzare questo metodo.
- **Utilizza stima WR per l'analisi.** Per impostazione predefinita, viene specificato un metodo di stima nel file di piano che è coerente con il metodo di campionamento selezionato. Ciò consente di utilizzare la stima con sostituzione anche se il metodo di campionamento implica la stima WOR. Questa opzione è disponibile solo nella fase 1.

**Misura di Dimensione (MOS).** Se viene selezionato un metodo PPS, è necessario specificare una misura di dimensione che definisce le dimensioni di ciascuna unità. Queste dimensioni possono essere esplicitamente definite in una variabile o possono essere calcolate dai dati. Facoltativamente, è possibile

impostare limiti inferiori e superiori sul MOS, sovrascrivendo qualsiasi valore trovato nella variabile MOS o calcolato dai dati. Queste opzioni sono disponibili solo nella fase 1.

## Procedura guidata di campionamento: Dimensione di esempio

Questo passaggio consente di specificare il numero o la proporzione di unità da campionare all'interno della fase corrente. La dimensione del campione può essere fissa o può variare tra gli strati. Ai fini della specifica dimensione del campione, i cluster scelti nelle fasi precedenti possono essere utilizzati per definire strati.

**Unità.** È possibile specificare una dimensione di esempio esatta o una proporzione di unità da campionare.

- **Valore.** Un valore unico viene applicato a tutti gli strati. Se **Conti** viene selezionato come metrica di unità, si dovrebbe inserire un intero positivo. Se **Proporzioni** è selezionata, è necessario inserire un valore non negativo. A meno che il campionamento con sostituzione, anche i valori di proporzione non dovrebbero essere superiori a 1.
- **Valori non uguali per strati.** Consente di inserire valori di dimensione su base per strato tramite la finestra di dialogo Definisci dimensioni non uguali.
- **I valori di lettura da variabile.** Consente di selezionare una variabile numerica che contenga valori di dimensione per strati.

Se **Proporzioni** è selezionata, si ha l'opzione di impostare limiti inferiori e superiori sul numero di unità campionate.

### Definisci dimensioni non uguali

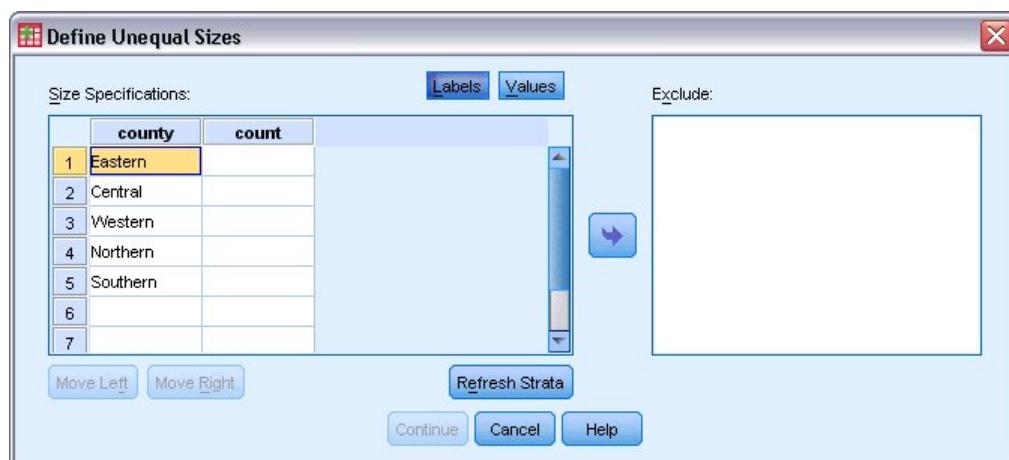


Figura 1. Definire Finestre di dialogo Dimensioni unuguali

La Finestra di dialogo Define Unuguali Dimensioni consente di inserire dimensioni su base per strati.

**Griglia delle specifiche di dimensione.** La griglia visualizza le classificazioni incrociate fino a cinque strati o variabili cluster - una combinazione stratum/cluster per riga. Le variabili di griglia idonee includono tutte le variabili di stratificazione dalle fasi attuali e precedenti e tutte le variabili di cluster dalle fasi precedenti. Le variabili possono essere riordinate all'interno della griglia o spostate nella lista Escludi. Inserire dimensioni nella colonna più a destra. Clicca su **Labels** o **Valori** per alternare la visualizzazione delle etichette di valore e dei valori dei dati per le variabili di stratificazione e cluster nelle celle della griglia. Le celle che contengono valori non etichettati mostrano sempre valori. Clicca su **Strata Aggiorna** per ripopolare la griglia con ogni combinazione di valori di dati etichettati per le variabili nella griglia.

**Escludi.** Per specificare le dimensioni per un sottoinsieme di combinazioni strato / cluster, spostare una o più variabili nella lista Escludi. Queste variabili non vengono utilizzate per definire le dimensioni del campione.

## Procedura guidata di campionamento: Variabili di output

Questo passo consente di scegliere le variabili da salvare quando il campione viene disegnato.

**Dimensione della popolazione.** Il numero stimato di unità nella popolazione per una determinata fase. Il nome della variabile salvata è *PopulationSize\_*.

**Proporzione di esempio.** La frequenza di campionamento in una determinata fase. Il nome della variabile salvata è *SamplingRate\_*.

**Dimensione campione.** Il numero di unità tracciate in una determinata fase. Il nome della variabile salvata è *SampleSize\_*.

**Peso di esempio.** L'inverso delle probabilità di inclusione. Il nome della variabile salvata è *SampleWeight\_*.

Alcune variabili stagewise vengono generate automaticamente. Queste release includono:

**Probabilità di inclusione.** La proporzione di unità disegnate in una determinata fase. Il nome della variabile salvata è *InclusionProbability\_*.

**Peso cumulativo.** Il peso del campione cumulativo su fasi precedenti a e inclusa quella attuale. Il nome della variabile salvata è *SampleWeightCumulative\_*.

**Indice.** Identifica le unità selezionate più volte all'interno di una determinata fase. Il rootname per la variabile salvata è *Index\_*.

*Nota:* i rootnomi variabili salvati includono un suffisso intero che riflette il numero di stage - ad esempio, *PopulationSize\_1\_* per la dimensione della popolazione salvata per la fase 1.

## Procedura guidata di campionamento: Riepilogo piano

Questo è l'ultimo passo all'interno di ogni fase, fornendo un riepilogo delle specifiche di progettazione del campione attraverso la fase attuale. Da qui è possibile procedere alla fase successiva (crearla, se necessario) o impostare le opzioni per il disegno del campione.

## Procedura guidata di campionamento: Opzioni di selezione di esempio Draw

Questo passo consente di scegliere se disegnare un campione. È anche possibile controllare altre opzioni di campionamento, come il seme casuale e la gestione del valore mancante.

**Prelievo campione.** Oltre a scegliere se disegnare un campione, è anche possibile scegliere di eseguire parte del disegno di campionamento. Le tappe devono essere tracciate in ordine - vale a dire la fase 2 non può essere disegnata a meno che non sia tracciata anche la fase 1. Durante la modifica o l'esecuzione di un piano, non è possibile ricampionare gli stage bloccati.

**Seme.** Questo consente di scegliere un valore di seme per la generazione di numeri casuali.

**Include i valori mancanti dell'utente.** Questo determina se i valori mancanti dell'utente sono validi. In caso affermativo, i valori mancanti dell'utente sono trattati come una categoria separata.

**Dati già ordinati.** Se il vostro frame di esempio è presorto dai valori delle variabili di stratificazione, questa opzione consente di accelerare il processo di selezione.

## Procedura guidata di campionamento: File di output di esempio Draw

Questo passo consente di scegliere dove indirizzare i casi campionati, le variabili di peso, le probabilità congiunte e le regole di selezione dei casi.

**Dati di esempio.** Queste opzioni consentono di determinare dove viene scritto l'output di esempio. Può essere aggiunto al dataset attivo, scritto ad un nuovo dataset oppure salvato in un file di dati IBM SPSS Statistica esterno. I Datasets sono disponibili durante la sessione corrente ma non sono disponibili nelle sessioni successive a meno che non si salvino esplicitamente come file di dati. I nomi di dataset devono aderire a regole di denominazione variabili. Se viene specificato un file esterno o un nuovo

dataset, vengono scritte le variabili di output del campionamento e le variabili nel dataset attivo per i casi selezionati.

**probabilità congiunte.** Queste opzioni consentono di stabilire dove sono scritte le probabilità congiunte. Vengono salvati in un file di dati IBM SPSS Statistica esterno. Le probabilità congiunte sono prodotte se il metodo PPS WOR, PPS Brewer, PPS Sampford o PPS Murthy è selezionato e la stima WR non è specificata.

**Norme di selezione del caso.** Se si sta costruendo il proprio campione uno stage alla volta, si potrebbe voler salvare le regole di selezione dei casi in un file di testo. Sono utili per costruire il sottotelaio per fasi successive.

## Procedura guidata di campionamento: Fine

Questo è il passo finale. È possibile salvare il file del piano e disegnare ora il campione o incollare le selezioni in una finestra di sintassi.

Quando si effettuano modifiche alle fasi nel file del piano esistente, è possibile salvare il piano modificato in un nuovo file o sovrascrivere il file esistente. Quando si aggiungono fasi senza apportare modifiche alle fasi esistenti, il Wizard sovrascrive automaticamente il file del piano esistente. Se si desidera salvare il piano in un nuovo file, selezionare **Incolla la sintassi generata dal Wizard in una finestra di sintassi** e modificare il filename nei comandi di sintassi.

## Modifica di un piano di esempio esistente

1. Dai menu, scegliere:

**Analizzare > Campioni complessi > Selezionare un esempio ...**

2. Selezionare **Modifica un design di esempio** e scegliere un file di piano da modificare.

3. Clicca su **Avanti** per continuare attraverso il Wizard.

4. Rivedere il piano di campionamento nella fase di riepilogo del piano, quindi fare clic su **Avanti**.

I passi successivi sono in gran parte gli stessi di un nuovo design. Consultare la Guida per i singoli passi per ulteriori informazioni.

5. Navigare al passo Fine e specificare un nuovo nome per il file di piano modificato o scegliere di sovrascrivere il file di piano esistente.

Facoltativamente, è possibile:

- Specificare fasi che sono già state campionate.
- Rimuovere fasi dal piano.

## Procedura guidata di campionamento: Riepilogo piano

Questo passo consente di rivedere il piano di campionamento e indicare fasi già campionate. Se si modifica un piano, è possibile rimuovere anche le fasi dal piano.

**stadi precedentemente campionati.** Se un frame di campionamento esteso non è disponibile, dovrà eseguire una progettazione di campionamento multistage una fase alla volta. Selezionare quali stage sono già stati campionati dall'elenco a discesa. Eventuali stage che sono stati eseguiti sono bloccati; non sono disponibili nel passo Opzioni di selezione di esempio Draw e non possono essere alterati quando si modifica un piano.

**Rimuovi stage.** È possibile rimuovere gli stage 2 e 3 da un design multistadio.

## Esecuzione di un piano di campionamento esistente

1. Dai menu, scegliere:

**Analizzare > Campioni complessi > Selezionare un esempio ...**

2. Selezionare **Draw a sample** e scegliere un file di piano da eseguire.

3. Clicca su **Avanti** per continuare attraverso il Wizard.

4. Rivedere il piano di campionamento nella fase di riepilogo del piano, quindi fare clic su **Avanti**.
  5. Le singole fasi contenenti le informazioni di stage vengono saltate durante l'esecuzione di un piano di esempio. Ora puoi passare al passo Finish in qualsiasi momento.
- Facoltativamente, è possibile specificare stage già campionati.

## Comandi aggiuntivi CSPLAN e CSSELECT

Il linguaggio della sintassi dei comandi consente inoltre di:

- Specificare nomi personalizzati per le variabili di output.
- Controllare l'output nel Viewer. Ad esempio, è possibile sopprimere il riepilogo stagewise del piano che viene visualizzato se un campione viene progettato o modificato, sopprimere il riepilogo della distribuzione dei casi campionati da strati che viene mostrato se il disegno del campione viene eseguito e richiedere un riepilogo di elaborazione del caso.
- Scegli un sottoinsieme di variabili nel dataset attivo per scrivere su un file di esempio esterno o su un dataset diverso.

Vedere *Command Syntax Reference* per informazioni dettagliate sulla sintassi.

## Preparazione di un esempio Complesso per l'analisi

La Procedura guidata di Preparazione analisi ti guida attraverso le fasi per la creazione o la modifica di un piano di analisi per l'utilizzo con le varie procedure di analisi Complesse Campioni. Prima di utilizzare il Wizard, si dovrebbe avere un campione disegnato secondo un design complesso.

La creazione di un nuovo piano è più utile quando non si ha accesso al file del piano di campionamento utilizzato per disegnare il campione (ricordando che il piano di campionamento contiene un piano di analisi predefinito). Se si ha accesso al file del piano di campionamento utilizzato per disegnare il campione, è possibile utilizzare il piano di analisi predefinito contenuto nel file del piano di campionamento o sovrascrivere le specifiche di analisi predefinite e salvare le modifiche ad un nuovo file.

### Creazione di un nuovo piano di analisi

1. Dai menu, scegliere:

**Analizzare > Campioni complessi > Prepararsi per l'analisi ...**

2. Selezionare **Crea un file di piano** e scegliere un nomefile di piano a cui salverete il piano di analisi.
3. Clicca su **Avanti** per continuare attraverso il Wizard.
4. Specificare la variabile contenente pesi di esempio nel passo Design Variables, optionalmente definendo strati e cluster.
5. Ora è possibile fare clic su **Fine** per salvare il piano.

Opzionalmente, in ulteriori passi è possibile:

- Selezionare il metodo per stimare gli errori standard nella fase di stima del metodo.
- Specificare il numero di unità campionate o la probabilità di inclusione per unità nel passo Dimensioni.
- Aggiungere una seconda o terza fase al disegno.
- Incollare le selezioni come sintassi del comando.

### Procedura guidata di Preparazione analisi: Variabili di progettazione

Questo passo consente di identificare le variabili di stratificazione e clustering e definire i pesi di esempio. È inoltre possibile fornire un'etichetta per la fase.

**Strata.** La classificazione incrociata delle variabili di stratificazione definisce sottopopolazioni distinte, o strati. Il tuo campione totale rappresenta la combinazione di campioni indipendenti da ogni strato.

**Cluster.** Le variabili cluster definiscono gruppi di unità osservazionali, o cluster. I campioni disegnati in più fasi selezionano cluster nelle fasi precedenti e poi sottocampionamenti dai cluster selezionati. Quando si analizza un file di dati ottenuto da cluster di campionamento con sostituzione, è necessario includere l'indice di duplicazione come variabile cluster.

**Peso del campione.** È necessario fornire pesi di esempio nella prima fase. I pesi di esempio vengono calcolati automaticamente per le fasi successive del design attuale.

**Etichetta di stage.** È possibile specificare un'etichetta di stringa opzionale per ogni fase. Questo viene utilizzato nell'output per aiutare a identificare le informazioni stagne.

*Nota:* l'elenco di variabili di origine ha lo stesso contenuto a tutte le fasi della procedura guidata. In altre parole, le variabili eliminate dall'elenco di origine in un determinato passo vengono eliminate dall'elenco in tutti i passaggi. Le variabili restituite alla lista di origine si presentano in tutti i passaggi.

## Controlli Albero per la navigazione della procedura guidata di analisi

Al lato sinistro di ogni passo della Mago di analisi è presente uno schema di tutti i passaggi. È possibile navigare il Wizard cliccando sul nome di un passo abilitato nello schema. I passaggi sono abilitati fino a quando tutti i passaggi precedenti sono validi - vale a dire, a patto che ogni fase precedente sia stata fornita le specifiche minime richieste per quel passaggio. Per ulteriori informazioni sul motivo per cui un determinato passo potrebbe essere non valido, consultare la Guida per i singoli passi.

## Procedura guidata di Preparazione analisi: Metodo di stima

Questo passo consente di specificare un metodo di stima per la fase.

**WR (campionamento con sostituzione).** La stima WR non include una correzione per il campionamento da una popolazione finita (FPC) quando stimola la varianza sotto il disegno di campionamento complesso. È possibile scegliere di includere o escludere la FPC quando si stima la varianza sotto semplice campionamento casuale (SRS).

Scegliere di non includere la stima della varianza FPC per SRS è consigliabile quando i pesi di analisi sono stati scalati in modo da non aggiungere alla dimensione della popolazione. La stima della varianza SRS viene utilizzata nelle statistiche di calcolo come l'effetto di progettazione. La stima WR può essere specificata solo nella fase finale di un disegno; il Wizard non permetterà di aggiungere un'altra fase se si seleziona la stima WR.

**Equal WOR (campionamento di probabilità uguale senza sostituzione).** La stima WOR uguale include la correzione della popolazione finita e presuppone che le unità siano campionate con uguale probabilità. Uguale WOR può essere specificato in qualsiasi fase di un design.

**Unequal WOR (campionamento di probabilità non uguale senza sostituzione).** Oltre ad utilizzare la correzione della popolazione finita, Unequal WOR rappresenta le unità di campionamento (di solito cluster) selezionate con una probabilità ineguagliabile. Questo metodo di stima è disponibile solo nella prima fase.

## Procedura guidata Preparazione analisi: Dimensioni

Questo passo viene utilizzato per specificare le probabilità di inclusione o le dimensioni della popolazione per la fase corrente. Le dimensioni possono essere fisse o possono variare tra gli strati. Ai fini della specifica delle dimensioni, i cluster specificati nelle fasi precedenti possono essere utilizzati per definire strati. Da notare che questo passo è necessario solo quando Parco WOR viene scelto come Metodo di stima.

**Unità.** È possibile specificare le dimensioni esatte della popolazione o le probabilità con cui le unità sono state campionate.

- Valore.** Un valore unico viene applicato a tutti gli strati. Se **Popolazione** viene selezionata come metrica di unità, è necessario inserire un intero non negativo. Se **Probabilità di inclusione** viene selezionata, è necessario inserire un valore compreso tra 0 e 1, inclusivo.

- **Valori non uguali per strati.** Consente di inserire valori di dimensione su base per strato tramite la finestra di dialogo Definisci dimensioni non uguali.
- **I valori di lettura da variabile.** Consente di selezionare una variabile numerica che contenga valori di dimensione per strati.

## Definisci dimensioni non uguali

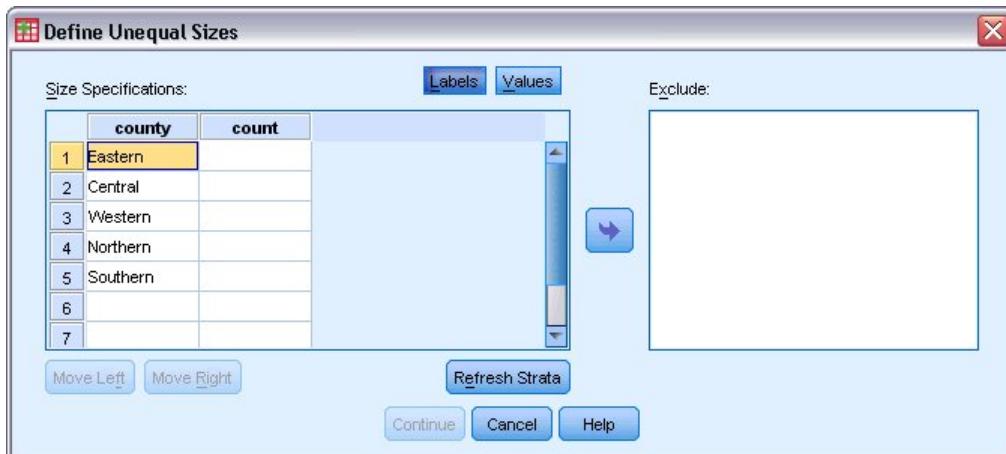


Figura 2. Definire Finestre di dialogo Dimensioni unuguali

La Finestra di dialogo Define Unuguali Dimensioni consente di inserire dimensioni su base per strati.

**Griglia delle specifiche di dimensione.** La griglia visualizza le classificazioni incrociate fino a cinque strati o variabili cluster - una combinazione stratum/cluster per riga. Le variabili di griglia idonee includono tutte le variabili di stratificazione dalle fasi attuali e precedenti e tutte le variabili di cluster dalle fasi precedenti. Le variabili possono essere riordinate all'interno della griglia o spostate nella lista Escludi. Inserire dimensioni nella colonna più a destra. Clicca su **Labels** o **Valori** per alternare la visualizzazione delle etichette di valore e dei valori dei dati per le variabili di stratificazione e cluster nelle celle della griglia. Le celle che contengono valori non etichettati mostrano sempre valori. Clicca su **Strata Aggiorna** per ripopolare la griglia con ogni combinazione di valori di dati etichettati per le variabili nella griglia.

**Escludi.** Per specificare le dimensioni per un sottoinsieme di combinazioni strato / cluster, spostare una o più variabili nella lista Escludi. Queste variabili non vengono utilizzate per definire le dimensioni del campione.

## Procedura guidata di Preparazione analisi: Riepilogo piano

Questo è l'ultimo passo all'interno di ogni fase, fornendo un riepilogo delle specifiche di progettazione dell'analisi attraverso la fase attuale. Da qui è possibile procedere alla fase successiva (crearla se necessario) o salvare le specifiche di analisi.

Se non è possibile aggiungere un'altra fase, è probabile perché:

- Non è stata specificata alcuna variabile cluster nel passo Design Variables.
- Hai selezionato la stima WR nella fase di stima del metodo.
- Questa è la terza fase dell'analisi e il Wizard supporta un massimo di tre tappe.

## Procedura guidata di Preparazione analisi: Finish

Questo è il passo finale. È possibile salvare il file del piano ora o incollare le selezioni ad una finestra di sintassi.

Quando si effettuano modifiche alle fasi nel file del piano esistente, è possibile salvare il piano modificato in un nuovo file o sovrascrivere il file esistente. Quando si aggiungono fasi senza apportare modifiche alle fasi esistenti, il Wizard sovrascrive automaticamente il file del piano esistente. Se si desidera salvare il

piano in un nuovo file, scegliere di **Incollare la sintassi generata dal Wizard in una finestra di sintassi** e modificare il filename nei comandi di sintassi.

## Modifica di un piano di analisi esistente

1. Dai menu, scegliere:

**Analizzare > Campioni complessi > Prepararsi per l'analisi ...**

2. Selezionare **Modifica un file di piano** e scegliere un nomefile di piano a cui salverete il piano di analisi.
3. Clicca su **Avanti** per continuare attraverso il Wizard.
4. Rivedere il piano di analisi nella fase di riepilogo del piano, quindi fare clic su **Avanti**.

I passi successivi sono in gran parte gli stessi di un nuovo design. Per ulteriori informazioni, consultare la Guida per i singoli passi.

5. Navigare al passo Fine e specificare un nuovo nome per il file di piano modificato oppure scegliere di sovrascrivere il file di piano esistente.

Facoltativamente, è possibile rimuovere fasi dal piano.

## Procedura guidata di Preparazione analisi: Riepilogo piano

Questo passo consente di rivedere il piano di analisi e rimuovere fasi dal piano.

**Rimuovere Fasi.** È possibile rimuovere gli stage 2 e 3 da un design multistadio. Dal momento che un piano deve avere almeno una fase, è possibile modificare ma non rimuovere la fase 1 dal design.

## Piano campioni complessi

Complesse Procedure di analisi dei campioni richiedono specifiche di analisi da un file di analisi o piano di esempio al fine di fornire risultati validi.

**Pianificazione.** Specificare il percorso di un file di analisi o di piano di esempio.

**Probabilità congiunte.** Per utilizzare la stima Unequal WOR per i cluster tracciati utilizzando un metodo PPS WOR, è necessario specificare un file separato o un dataset aperto contenente le probabilità congiunte. Questo file o dataset viene creato dal Mago di Sampling durante il campionamento.

## Frequenze campioni complessi

La Procedura Complessa Campioni Frequenze produce tabelle di frequenza per variabili selezionate e visualizza le statistiche univariate. Facoltativamente, è possibile richiedere statistiche per sottogruppi, definite da una o più variabili categoriali.

**Esempio.** Utilizzando la procedura Frequenze campioni complessi, è possibile ottenere statistiche tabulari univariate per l'utilizzo di vitamine tra i cittadini degli U.S, in base ai risultati del National Health Interview Survey (NHIS) e con un piano di analisi appropriato per questi dati di uso pubblico.

**Statistiche.** La procedura produce stime delle dimensioni della popolazione cellulare e delle percentuali di tabella, più gli errori standard, gli intervalli di confidenza, i coefficienti di variazione, gli effetti di design, le radici quadrate degli effetti di design, i valori cumulativi e i conteggi non ponderati per ogni stima. Inoltre, le statistiche di chi - quadrato e di probabilità - ratio sono calcolate per il test di proporzioni cellulari uguali.

Complessi Campioni di frequenze Dati sulle frequenze

**Dati.** Le variabili per le quali vengono prodotte tabelle di frequenza dovrebbero essere categoriali. Le variabili di sottopopolazione possono essere stringa o numerica ma devono essere categoriali.

**Ipotesi.** I casi nel file dati rappresentano un campione da un design complesso che dovrebbe essere analizzato in base alle specifiche nel file selezionato nella finestra di dialogo Complex Campioni Plan.

Ottenere Frequenze di Campioni Complessi

1. Dai menu, scegliere:

**Analizzare > Campioni complessi > frequenze ...**

2. Selezionare un file del piano. Facoltativamente, selezionare un file di probabilità congiunta personalizzato.

3. Fare clic su **Continua**.

4. Nella finestra di dialogo Complex Campioni Frequenze, selezionare almeno una variabile di frequenza.

Facoltativamente, è possibile specificare le variabili per definire le sottopopolazioni. Le statistiche sono calcolate separatamente per ogni sottopopolazione.

## Complesso Statistiche Frequenze Statistiche

**Celle.** Questo gruppo consente di richiedere preventivi delle dimensioni della popolazione cellulare e delle percentuali di tabella.

**Statistiche.** Questo gruppo produce statistiche associate alla dimensione della popolazione o alla percentuale di tabella.

- **Errore standard.** Errore standard della stima.
- **Intervallo di confidenza.** Un intervallo di confidenza per la stima, utilizzando il livello specificato.
- **CV.** Il rapporto dell'errore standard della stima alla stima.
- **Concount non ponderato.** Il numero di unità utilizzate per calcolare la stima.
- **Effetto di progettazione.** Il rapporto della varianza della stima alla varianza ottenuta ipotizzando che il campione sia un semplice campione casuale. È una misura dell'effetto derivato dalla specificazione di un disegno complesso, dove valori più piccoli indicano effetti maggiori.
- **radice di piazza dell'effetto design.** È una misura dell'effetto derivato dalla specificazione di un disegno complesso, dove valori più piccoli indicano effetti maggiori.
- **Valori cumulativi.** La stima cumulativa attraverso ogni valore della variabile.

**Test di proporzioni cellulari uguali.** Questo produce test di chi - quadrato e di verosimiglianza dell'ipotesi che le categorie di una variabile abbiano frequenze uguali. I test separati vengono eseguiti per ogni variabile.

## Complessi Campioni Mancanti

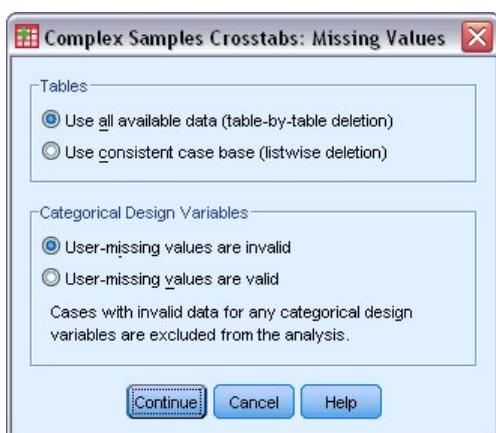


Figura 3. Finestra di dialogo Valori mancanti

**Tabelle.** Questo gruppo determina quali casi sono utilizzati nell'analisi.

- **Utilizza tutti i dati disponibili.** I valori mancanti sono determinati su base da tavolo. Così, i casi utilizzati per calcolare le statistiche possono variare tra le tabelle di frequenza o di stabulazione.
- **Utilizza base di case coerenti.** I valori mancanti sono determinati in tutte le variabili. Così, i casi utilizzati per calcolare le statistiche sono coerenti tra le tabelle.

**Variabili di progettazione categoriale.** Questo gruppo determina se i valori mancanti dell'utente sono validi o non validi.

## Opzioni Campioni Complessi

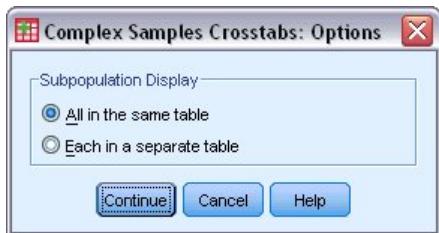


Figura 4. Finestra di dialogo Opzioni

**Visualizza sottopopolazione.** È possibile scegliere di avere sottopopolazioni visualizzate nella stessa tabella o in tabelle separate.

## Descrittive campioni complessi

La Procedura Complessa Campioni Descrittivi visualizza le statistiche di riepilogo univariate per diverse variabili. Facoltativamente, è possibile richiedere statistiche per sottogruppi, definite da una o più variabili categoriali.

**Esempio.** Utilizzando la procedura Descrittive campioni complessi, è possibile ottenere statistiche descrittive univariate per i livelli di attività dei cittadini U.S. in base ai risultati del National Health Interview Survey (NHIS) e con un piano di analisi appropriato per questi dati di uso pubblico.

**Statistiche.** La procedura produce mezzi e somme, più  $t$  test, errori standard, intervalli di confidenza, coefficienti di variazione, conteggi non ponderati, dimensioni della popolazione, effetti di progettazione e radici quadrate degli effetti di progettazione per ogni stima.

Complessi Campioni Descrittivi Dati Considerazioni

**Dati.** Le misure dovrebbero essere variabili di scala. Le variabili di sottopopolazione possono essere stringa o numerica ma devono essere categoriali.

**Ipotesi.** I casi nel file dati rappresentano un campione da un design complesso che dovrebbe essere analizzato in base alle specifiche nel file selezionato nella finestra di dialogo [Complex Campioni Plan](#).

Ottenere Dettagli Campioni Complessi

1. Dai menu, scegliere:

**Analizzare > Campioni complessi > descrittivi ...**

2. Selezionare un file del piano. Facoltativamente, selezionare un file di probabilità congiunta personalizzato.
3. Fare clic su **Continua**.
4. Nella finestra di dialogo Complex Campioni Descriptives, selezionare almeno una variabile di misura.

Facoltativamente, è possibile specificare le variabili per definire le sottopopolazioni. Le statistiche sono calcolate separatamente per ogni sottopopolazione.

## Statistiche Descrittive Campioni Statistici

**Riepiloghi.** Questo gruppo consente di richiedere preventivi dei mezzi e delle somme delle variabili di misura. Inoltre, è possibile richiedere i test  $t$  delle stime su un valore specificato.

**Statistiche.** Questo gruppo produce statistiche associate alla media o alla somma.

- **Errore standard.** Errore standard della stima.
- **Intervallo di confidenza.** Un intervallo di confidenza per la stima, utilizzando il livello specificato.

- **CV.** Il rapporto dell'errore standard della stima alla stima.
- **Concount non ponderato.** Il numero di unità utilizzate per calcolare la stima.
- **Dimensione della popolazione.** Il numero stimato di unità nella popolazione.
- **Effetto di progettazione.** Il rapporto della varianza della stima alla varianza ottenuta ipotizzando che il campione sia un semplice campione casuale. È una misura dell'effetto derivato dalla specificazione di un disegno complesso, dove valori più piccoli indicano effetti maggiori.
- **radice di piazza dell'effetto design.** È una misura dell'effetto derivato dalla specificazione di un disegno complesso, dove valori più piccoli indicano effetti maggiori.

## Complessi Campioni Descrittivi Mancanti Valori

**Statistiche per Variabili di misura.** Questo gruppo determina quali casi sono utilizzati nell'analisi.

- **Utilizza tutti i dati disponibili.** I valori mancanti sono determinati su base variabile - per variabile, quindi i casi utilizzati per calcolare le statistiche possono variare tra le variabili di misura.
- **Garantire la base del caso coerente.** I valori mancanti sono determinati in tutte le variabili, quindi i casi utilizzati per calcolare le statistiche sono coerenti.

**Variabili di progettazione categoriale.** Questo gruppo determina se i valori mancanti dell'utente sono validi o non validi.

## Opzioni Campioni Complessi

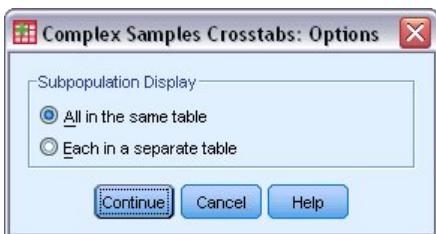


Figura 5. Finestra di dialogo Opzioni

**Visualizza sottopopolazione.** È possibile scegliere di avere sottopopolazioni visualizzate nella stessa tabella o in tabelle separate.

## Tabelle di contingenza campioni complessi

La Procedura Complessa Campioni Crosstabs produce tavoli di stabulazione per coppie di variabili selezionate e visualizza le statistiche a due vie. Facoltativamente, è possibile richiedere statistiche per sottogruppi, definite da una o più variabili categoriali.

**Esempio.** Utilizzando la procedura Tabelle di contingenza campioni complessi, è possibile ottenere statistiche di classificazione incrociata per la frequenza di fumo in base all'utilizzo di vitamine dei cittadini U.S., in base ai risultati del National Health Interview Survey (NHIS) e con un piano di analisi appropriato per questi dati di uso pubblico.

**Statistiche.** La procedura produce stime delle dimensioni della popolazione cellulare e della riga, della colonna e delle percentuali di tabella, più errori standard, intervalli di confidenza, coefficienti di variazione, valori previsti, effetti di progettazione, radici quadrate degli effetti di progettazione, residui, residui regolati e conteggi non ponderati per ogni stima. Il rapporto di probabilità, il rischio relativo e la differenza di rischio sono calcolati per le tabelle 2-by-2. Inoltre, le statistiche di Pearson e di probabilità - ratio sono calcolate per la prova di indipendenza delle variabili di fila e colonne.

Complessi Campioni Crosstabs Dati Considerazioni

**Dati.** Le variabili di riga e colonna dovrebbero essere categoriali. Le variabili di sottopopolazione possono essere stringa o numerica ma devono essere categoriali.

**Ipotesi.** I casi nel file dati rappresentano un campione da un design complesso che dovrebbe essere analizzato in base alle specifiche nel file selezionato nella finestra di dialogo Complex Campioni Plan.

Ottenimento Campioni Complessi Crosstabs

1. Dai menu, scegliere:

**Analizzare > Campioni complessi > Crosstabs ...**

2. Selezionare un file del piano. Facoltativamente, selezionare un file di probabilità congiunta personalizzato.
3. Fare clic su **Continua**.
4. Nella finestra di dialogo Complex Campioni Crosstabs, selezionare almeno una variabile di riga e una variabile di colonna.

Facoltativamente, è possibile specificare le variabili per definire le sottopopolazioni. Le statistiche sono calcolate separatamente per ogni sottopopolazione.

## Complessi Campioni Crosstabs statistiche

**Celle.** Questo gruppo consente di richiedere le stime della dimensione della popolazione cellulare e della riga, della colonna e delle percentuali di tabella.

**Statistiche.** Questo gruppo produce statistiche associate alla dimensione della popolazione e alla riga, alla colonna e alle percentuali di tabella.

- **Errore standard.** Errore standard della stima.
- **Intervallo di confidenza.** Un intervallo di confidenza per la stima, utilizzando il livello specificato.
- **Acronimo di Coefficient of Variation, ossia coefficiente di variazione.** Il rapporto dell'errore standard della stima alla stima.
- **Valori attesi.** Il valore atteso della stima, nell'ipotesi di indipendenza della variabile di fila e colonna.
- **Concount non ponderato.** Il numero di unità utilizzate per calcolare la stima.
- **Effetto di progettazione.** Il rapporto della varianza della stima alla varianza ottenuta ipotizzando che il campione sia un semplice campione casuale. È una misura dell'effetto derivato dalla specificazione di un disegno complesso, dove valori più piccoli indicano effetti maggiori.
- **radice di piazza dell'effetto design.** È una misura dell'effetto derivato dalla specificazione di un disegno complesso, dove valori più piccoli indicano effetti maggiori.
- **Residui.** Il valore previsto è il numero di casi che ci si aspetterebbe in cella se non ci fossero rapporti tra le due variabili. Un residuo positivo indica che ci sono più casi nella cella di quanto ci sarebbe se le variabili di riga e colonna fossero indipendenti.
- **Residui corretti.** Il residuo per una cella (osservato meno valore previsto) diviso per una stima del suo errore standard. Il residuo standardizzato risultante è espresso in unità di deviazione standard sopra o sotto la media.

**Sintesi per 2-by-2 Tabelle.** Questo gruppo produce statistiche per tabelle in cui la variabile riga e colonna ognuna ha due categorie. Ognuna è una misura della forza dell'associazione tra la presenza di un fattore e la ricorrenza di un evento.

- **Rapporto di odds.** Il rapporto di probabilità può essere utilizzato come stima del rischio relativo quando la ricorrenza del fattore è rara.
- **Rischio relativo.** Il rapporto rischio di un evento in presenza del fattore a rischio dell'evento in assenza del fattore.
- **Differenza di rischio.** La differenza tra il rischio di un evento in presenza del fattore e il rischio dell'evento in assenza del fattore.

**Prova di indipendenza di righe e colonne.** Questo produce test di chi - quadrato e di verosimiglianza dell'ipotesi che una variabile riga e colonna siano indipendenti. Vengono eseguiti dei test separati per ogni coppia di variabili.

## Complessi Campioni Mancanti

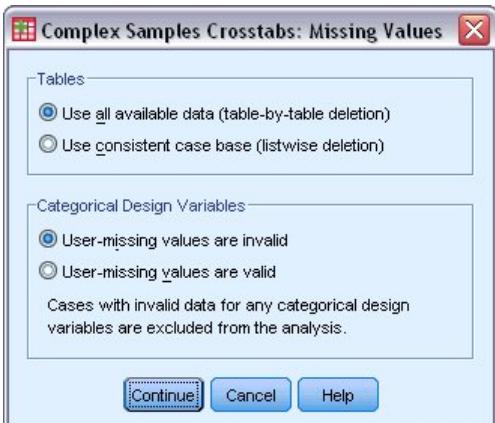


Figura 6. Finestra di dialogo Valori mancanti

**Tabelle.** Questo gruppo determina quali casi sono utilizzati nell'analisi.

- **Utilizza tutti i dati disponibili.** I valori mancanti sono determinati su base da tavolo. Così, i casi utilizzati per calcolare le statistiche possono variare tra le tabelle di frequenza o di stabulazione.
- **Utilizza base di case coerenti.** I valori mancanti sono determinati in tutte le variabili. Così, i casi utilizzati per calcolare le statistiche sono coerenti tra le tabelle.

**Variabili di progettazione categoriale.** Questo gruppo determina se i valori mancanti dell'utente sono validi o non validi.

## Opzioni Campioni Complessi

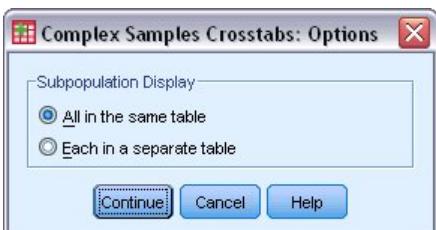


Figura 7. Finestra di dialogo Opzioni

**Visualizza sottopopolazione.** È possibile scegliere di avere sottopopolazioni visualizzate nella stessa tabella o in tabelle separate.

## Rapporti campioni complessi

La Procedura Complex Campioni Ratio visualizza le statistiche di riepilogo univariate per i rapporti delle variabili. Facoltativamente, è possibile richiedere statistiche per sottogruppi, definite da una o più variabili categoriali.

**Esempio.** Utilizzando la procedura Complex Campioni Ratio, è possibile ottenere statistiche descrittive per il rapporto tra valore di proprietà corrente e ultimo valore valutato, in base ai risultati di un'indagine statewide effettuata secondo un design complesso e con un adeguato piano di analisi per i dati.

**Statistiche.** La procedura produce stime di rapporto,  $t$  test, errori standard, intervalli di confidenza, coefficienti di variazione, conteggi non ponderati, dimensioni della popolazione, effetti di progettazione e radici quadrate degli effetti di progettazione.

Complessi Campioni Rapporti di Campioni Rapporti

**Dati.** Numeratori e denominatori dovrebbero essere variabili di scala valorizzate. Le variabili di sottopopolazione possono essere stringa o numerica ma devono essere categoriali.

**Ipotesi.** I casi nel file dati rappresentano un campione da un design complesso che dovrebbe essere analizzato in base alle specifiche nel file selezionato nella finestra di dialogo Complex Campioni Plan.

Ottenere Rapporti di campioni complessi

1. Dai menu, scegliere:

**Analizzare > Campioni Complessi > Rapporti ...**

2. Selezionare un file del piano. Facoltativamente, selezionare un file di probabilità congiunta personalizzato.
3. Fare clic su **Continua**.
4. Nella finestra di dialogo Complex Campioni Ratio, selezionare almeno una variabile di numeratore e denominatore.

Facoltativamente, è possibile specificare le variabili per definire i sottogruppi per i quali vengono prodotte statistiche.

## Statistiche Rapporti di Campioni Complessi

**Statistiche.** Questo gruppo produce statistiche associate al rapporto stima.

- **Errore standard.** Errore standard della stima.
- **Intervallo di confidenza.** Un intervallo di confidenza per la stima, utilizzando il livello specificato.
- **CV.** Il rapporto dell'errore standard della stima alla stima.
- **Concount non ponderato.** Il numero di unità utilizzate per calcolare la stima.
- **Dimensione della popolazione.** Il numero stimato di unità nella popolazione.
- **Effetto di progettazione.** Il rapporto della varianza della stima alla varianza ottenuta ipotizzando che il campione sia un semplice campione casuale. È una misura dell'effetto derivato dalla specificazione di un disegno complesso, dove valori più piccoli indicano effetti maggiori.
- **radice di piazza dell'effetto design.** È una misura dell'effetto derivato dalla specificazione di un disegno complesso, dove valori più piccoli indicano effetti maggiori.

**prova T.** È possibile richiedere i test *t* delle stime a fronte di un valore specificato.

## Complessi Campioni Rapporti Mancanti

**Rapporti.** Questo gruppo determina quali casi sono utilizzati nell'analisi.

- **Utilizza tutti i dati disponibili.** I valori mancanti sono determinati su base rapporto per rapporto. Così, i casi utilizzati per calcolare le statistiche possono variare tra le coppie di numeratori - denominatori.
- **Garantire la base del caso coerente.** I valori mancanti sono determinati in tutte le variabili. Così, i casi utilizzati per calcolare le statistiche sono coerenti.

**Variabili di progettazione categoriale.** Questo gruppo determina se i valori mancanti dell'utente sono validi o non validi.

## Opzioni Campioni Complessi

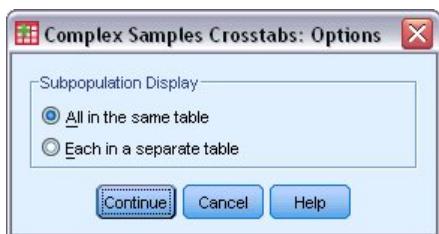


Figura 8. Finestra di dialogo Opzioni

**Visualizza sottopopolazione.** È possibile scegliere di avere sottopopolazioni visualizzate nella stessa tabella o in tabelle separate.

## Modello lineare generalizzato campioni complessi

---

Il Complesso Campioni General Linear Model lineare (CSGLM) esegue analisi di regressione lineare, nonché analisi di varianza e covarianza, per i campioni disegnati da metodi di campionamento complessi. Facoltativamente, è possibile richiedere analisi per una sottopopolazione.

**Esempio.** Una catena di negozi di alimentari ha sorvegliato una serie di clienti riguardanti le loro abitudini di acquisto, secondo un design complesso. Dati i risultati dell'indagine e quanto ogni cliente ha speso nel mese precedente, il negozio vuole vedere se la frequenza con cui i clienti shop è legata all'importo che spendono in un mese, controllando per il genere del cliente e incorporando il disegno di campionamento.

**Statistiche.** La procedura produce preventivi, errori standard, intervalli di confidenza, test  $t$ , effetti di progettazione e radici quadrate degli effetti di progettazione per i parametri del modello, nonché le correlazioni e le covarianze tra le stime dei parametri. Sono disponibili anche misure di inquadratura del modello e statistiche descrittive per le variabili dipendenti e indipendenti. Inoltre, è possibile richiedere mezzi marginali stimati per livelli di fattori di modello e interazioni dei fattori.

Complessi Campioni Generali di Linear Model Data Considerazioni

**Dati.** La variabile dipendente è quantitativa. I fattori sono categoriali. Le covariate sono variabili quantitative correlate alla variabile dipendente. Le variabili di sottopopolazione possono essere stringa o numerica ma devono essere categoriali.

**Ipotesi.** I casi nel file dati rappresentano un campione da un design complesso che dovrebbe essere analizzato in base alle specifiche nel file selezionato nella finestra di dialogo [Complex Campioni Plan](#).

Ottenimento di un modello lineare Complex Linear Model

1. Dai menu, scegliere:

**Analizzare > Campioni complessi > Modello lineare generale ...**

2. Selezionare un file del piano. Facoltativamente, selezionare un file di probabilità congiunta personalizzato.

3. Fare clic su **Continua**.

4. Nella finestra di dialogo Complex Campioni General Linear Model, selezionare una variabile dipendente.

Facoltativamente, è possibile:

- Selezionare le variabili per i fattori e le covariate, a seconda dei dati.
- Specificare una variabile per definire una sottopopolazione. L'analisi viene eseguita solo per la categoria selezionata della variabile sottopopolazione.

## Modello lineare generalizzato campioni complessi

**Specifiche effetti del modello.** Per impostazione predefinita, la procedura costruisce un modello di effetti principali utilizzando i fattori e le covariate specificate nella finestra di dialogo principale. In alternativa, è possibile costruire un modello personalizzato che include effetti di interazione e termini nidificati.

Termini Non Nidificati

Per i fattori e le covariate selezionati:

**Interazione.** Crea il termine di interazione ad alto livello per tutte le variabili selezionate.

**Effetti principali.** Consente di creare un termine di effetti principali per ciascuna variabile selezionata.

**Tutti - 2 vie** Consente di creare tutte le possibili interazioni a due vie delle variabili selezionate.

**Tutti - 3 vie** Consente di creare tutte le possibili interazioni a tre vie delle variabili selezionate.

**Tutti - 4 vie** Consente di creare tutte le possibili interazioni a quattro vie delle variabili selezionate.

**Tutti - 5 vie** Consente di creare tutte le possibili interazioni a cinque vie delle variabili selezionate.

#### Termini Nidificati

Questa procedura consente di costruire termini nidificati per il modello. I termini nidificati sono utili per modellare l'effetto di un fattore o di una covariata i cui valori non interagiscono con i livelli di un altro fattore. Per esempio, una catena di supermercati può seguire le abitudini di spesa dei propri clienti in più negozi. Poiché ogni cliente frequenta un solo negozio, l'effetto *Cliente* può definirsi **nidificato** all'interno dell'effetto *Negozio*.

È inoltre possibile includere effetti di interazione, ad esempio termini polinomiali che interessano la stessa covariata, o aggiungere più livelli di nidificazione al termine nidificato.

**Limitazione.** I termini nidificati sono sottoposti alle seguenti restrizioni:

- Tutti i fattori compresi in un'interazione devono essere univoci. Di conseguenza, se  $A$  è un fattore, non è possibile specificare  $A^*A$ .
- Tutti i fattori compresi in un effetto nidificato devono essere univoci. Di conseguenza, se  $A$  è un fattore, non è possibile specificare  $A(A)$ .
- Nessun effetto può essere nidificato all'interno di una covariata. Di conseguenza, se  $A$  è un fattore e  $X$  è una covariata, non è possibile specificare  $A(X)$ .

**Intercetta.** L'intercetta viene in genere inclusa nel modello. Se è possibile presumere che i dati passino attraverso l'origine, l'intercettazione può essere esclusa. Anche se si include l'intercettazione nel modello, è possibile scegliere di sopprimere le statistiche relative ad esso.

## Complesso Campioni General Linear Model Statistics

**Parametri del modello.** Questo gruppo consente di controllare la visualizzazione delle statistiche relative ai parametri del modello.

- **Stima.** Visualizza le stime dei coefficienti.
- **Errore standard.** Visualizza l'errore standard per ogni stima del coefficiente.
- **Intervallo di confidenza.** Visualizza un intervallo di confidenza per ogni stima di coefficiente. Il livello di confidenza per l'intervallo è impostato nella finestra di dialogo Opzioni.
- **prova T.** Visualizza un test  $t$  di ogni stima di coefficiente. L'ipotesi nulla per ogni prova è che il valore del coefficiente sia di 0.
- **Le covarianze delle stime dei parametri.** Visualizza una stima della matrice di covarianza per i coefficienti di modello.
- **Correlazioni delle stime dei parametri.** Visualizza una stima della matrice di correlazione per i coefficienti di modello.
- **Effetto di progettazione.** Il rapporto della varianza della stima alla varianza ottenuta ipotizzando che il campione sia un semplice campione casuale. È una misura dell'effetto derivato dalla specificazione di un disegno complesso, dove valori più piccoli indicano effetti maggiori.
- **radice di piazza dell'effetto design.** È una misura dell'effetto derivato dalla specificazione di un disegno complesso, dove valori più piccoli indicano effetti maggiori.

**Adattamento del modello.** Visualizza le statistiche di errore quadrati  $R^2$  e root.

**Popolazione media di variabile dipendente e covariate.** Visualizza le informazioni di riepilogo sulla variabile dipendente, le covariate e i fattori.

**Informazioni di progettazione di esempio.** Visualizza le informazioni di riepilogo sul campione, incluso il conteggio non ponderato e la dimensione della popolazione.

## Test Ipotesi di Campioni Complessi

**Statistica del test.** Questo gruppo consente di selezionare il tipo di statistica utilizzata per le ipotesi di prova. È possibile scegliere tra  $F$ , rettificato  $F$ , chi - quadrato e il chi quadrato rettificato.

**Gradi di campionamento della libertà.** Questo gruppo ti dà il controllo sui gradi di progettazione del campionamento di libertà utilizzati per calcolare i valori  $p$  per tutte le statistiche di prova. Se in base al disegno di campionamento, il valore è la differenza tra il numero di unità di campionamento primarie e il numero di strati nella prima fase di campionamento. In alternativa, è possibile impostare un grado di libertà personalizzato specificando un intero positivo.

**Regolazione per Confronti multipli.** Quando si eseguono test di ipotesi con più contrasti, il livello di significatività generale può essere adattato in base ai livelli di significatività dei contrasti inclusi. Questo gruppo consente di scegliere il metodo di regolazione.

- **Differenza meno significativa.** Questo metodo non controlla la probabilità generale di rifiuto delle ipotesi che alcuni contrasti lineari siano diversi dai valori di ipotesi null.
- *Sidak sequenziale.* Una procedura di Sidak con scarti sequenzialmente decrescenti, molto meno conservativa in termini di rifiuto di singole ipotesi, ma che mantiene lo stesso livello di significatività globale.
- *Bonferroni sequenziale.* Una procedura di Bonferroni con scarti sequenzialmente decrescenti, molto meno conservativa in termini di rifiuto di singole ipotesi, ma che mantiene lo stesso livello di significatività globale.
- *Sidak.* Questo metodo fornisce limiti più stretti rispetto all'approccio di Bonferroni.
- *Bonferroni.* Questo metodo corregge il livello di significatività osservato tenendo conto del fatto che si stanno verificando confronti multipli.

## Complesso Campioni General Linear Model Stimato Mezzo

La Finestra di dialogo Medie Stimate consente di visualizzare i mezzi marginali stimati per i livelli di fattori e le interazioni dei fattori specificati nella finestra di sottodialogo Modello. È anche possibile richiedere che la popolazione complessiva media venga visualizzata.

**Termine.** Le medie stimate vengono calcolate per i fattori e le interazioni dei fattori selezionati.

**Contrasto.** Il contrasto determina la modalità di impostazione dei test di ipotesi per il confronto delle medie stimate.

- *Semplice.* Consente di confrontare la media di ciascun livello con la media di un livello specifico. Questo tipo di contrasto risulta utile quando è disponibile un gruppo di controllo.
- *Deviazione.* Consente di confrontare la media di ciascun livello, a eccezione di una categoria di riferimento, con la media di tutti i livelli (media principale). I livelli del fattore possono essere in qualsiasi ordine.
- *Differenza.* Consente di confrontare la media di ciascun livello (a eccezione del primo) con la media dei livelli precedenti. Sono a volte definiti contrasti inversi di Helmert.
- *Helmert.* Confronta la media di ciascun livello del fattore (eccetto l'ultimo) con la media dei livelli successivi.
- *Ripetuto.* Consente di confrontare la media di ciascun livello (a eccezione dell'ultimo) con la media del livello successivo.
- *Polinomiale.* Consente di confrontare l'effetto lineare, quadratico, cubico e così via. Tutte le categorie del primo grado di libertà includono l'effetto lineare, quelle del secondo includono l'effetto quadratico e così via. Questi contrasti sono spesso usati per stimare le tendenze polinomiali.

**CATEGORIA DI RIFERIMENTO.** I contrasti semplici e di deviazione richiedono una categoria di riferimento o un livello di fattore rispetto al quale si confronta gli altri.

## Complesso Campioni General Linear Model Save

**Salva Variabili.** Questo gruppo consente di salvare i valori previsti e i residui come nuove variabili nel file di lavoro.

**Esporta modello come dati IBM SPSS Statistica**. Scrive un dataset in formato IBM SPSS Statistica contenente la correlazione dei parametri o la matrice di covarianza con le stime dei parametri, gli errori standard, i valori di significatività e i gradi di libertà. L'ordine delle variabili nel file di matrice è il seguente.

- **rowtype\_**. I valori di Takes (e le etichette di valore), COV (Covarianze), CORR (Correlazioni), EST (stime dei parametri), SE (errori standard), SIG (Livelli di Significanza) e DF (Sampling design gradi di libertà). Esiste un caso separato con tipo di riga COV (o CORR) per ogni parametro modello, più un caso separato per ciascuno degli altri tipi di riga.
- **varname\_**. Valori di Takes P1, P2, ..., corrispondente ad un elenco ordinato di tutti i parametri del modello, per i tipi di righe COV o CORR, con etichette di valore corrispondenti alle stringhe di parametro mostrate nella tabella delle stime dei parametri. Le celle sono vuote per altri tipi di righe.
- **P1, P2, ...** Queste variabili corrispondono ad un elenco ordinato di tutti i parametri del modello, con etichette variabili corrispondenti alle stringhe di parametri mostrate nella tabella delle stime dei parametri e assumono valori in base al tipo di riga. Per i parametri ridondanti, tutte le covarianze sono impostate a zero; le correlazioni sono impostate sul valore mancante di sistema; tutte le stime dei parametri sono impostate a zero; e tutti gli errori standard, i livelli di significatività e i gradi residui di libertà sono impostati sul valore mancante di sistema.

*Nota:* questo file non è immediatamente utilizzabile per ulteriori analisi in altre procedure che leggono un file di matrice a meno che tali procedure non accettino tutti i tipi di riga esportati qui.

**Modello di esportazione come XML.** Salva le stime dei parametri e la matrice di covarianza dei parametri, se selezionata, in formato XML (PMML). È possibile utilizzare questo file modello per applicare le informazioni del modello ad altri file di dati per il calcolo del punteggio.

## Complesso Campioni General Linear Model Options

**Valori mancanti definiti dall'utente.** Tutte le variabili di progettazione, così come la variabile dipendente e le eventuali covariate, devono avere dati validi. I casi con dati non validi per una di queste variabili vengono cancellati dall'analisi. Questi controlli consentono di decidere se i valori mancanti degli utenti sono trattati come validi tra le variabili strati, cluster, sottopopolazione e fattore.

**Intervallo di confidenza.** Questo è il livello di intervallo di confidenza per stime di coefficiente e medie marginali stimate. Specificare un valore maggiore o uguale a 50 e inferiore a 100.

## Funzioni Aggiuntive di comando CSGLM

Il linguaggio della sintassi dei comandi consente inoltre di:

- Specificare test personalizzati di effetti contro una combinazione lineare di effetti o un valore (utilizzando il comando CUSTOM ).
- Fissare covariate a valori diversi dai loro mezzi quando si calcolano i mezzi marginali stimati (utilizzando il comando EMMEANS ).
- Specificare una metrica per i contrasti polinomiali (utilizzando il comando EMMEANS ).
- Specificare un valore di tolleranza per la verifica della singolarità (utilizzando il comando CRITERIA ).
- Creare nomi specificati dall'utente per le variabili salvate (utilizzando il comando secondario SAVE ).
- Produrre una tabella di funzioni generali stimabili (utilizzando il sottocomando PRINT ).

Vedere *Command Syntax Reference* per informazioni dettagliate sulla sintassi.

## Regressione logistica campioni complessi

---

La Complessa Procedura Logistic Regression esegue analisi di regressione logistica su una variabile dipendente binario o multinomiale per i campioni tracciati da metodi di campionamento complessi. Facoltativamente, è possibile richiedere analisi per una sottopopolazione.

**Esempio.** Un addetto ai prestiti ha raccolto record di clienti passati dati i prestiti a diversi rami diversi, secondo un design complesso. Pur incorporando il design del campione, l'agente vuole vedere se la

probabilità con cui un cliente default è legata all'età, alla storia dell'occupazione e all'ammontare del debito di credito.

**Statistiche.** La procedura produce stime, stime esponenziate, errori standard, intervalli di confidenza, test  $t$ , effetti di progettazione e radici quadrate degli effetti di progettazione per i parametri del modello, nonché le correlazioni e le covarianze tra le stime dei parametri. Sono disponibili anche le statistiche di pseudo  $R^2$ , tabelle di classificazione e statistiche descrittive per le variabili dipendenti e indipendenti.

Considerazioni Complesse di Regressione Logistica Considerazioni di dati

**Dati.** La variabile dipendente è categoriale. I fattori sono categoriali. Le covariate sono variabili quantitative correlate alla variabile dipendente. Le variabili di sottopopolazione possono essere stringa o numerica ma devono essere categoriali.

**Ipotesi.** I casi nel file dati rappresentano un campione da un design complesso che dovrebbe essere analizzato in base alle specifiche nel file selezionato nella finestra di dialogo Complex Campioni Plan.

Ottenimento Campioni Complessi Logistic Regression

1. Dai menu, scegliere:

**Analizzare > Campioni Complessi > Regressione Logistica ...**

2. Selezionare un file del piano. Facoltativamente, selezionare un file di probabilità congiunta personalizzato.
3. Fare clic su **Continua**.
4. Nella finestra di dialogo Complex Campioni Logistic Regression, selezionare una variabile dipendente.  
Facoltativamente, è possibile:
  - Selezionare le variabili per i fattori e le covariate, a seconda dei dati.
  - Specificare una variabile per definire una sottopopolazione. L'analisi viene eseguita solo per la categoria selezionata della variabile sottopopolazione.

## Complesso Campioni Logistic Regression Reference Categoria

Per impostazione predefinita, la procedura Complex Samples Logistic Regression rende la categoria di riferimento più elevata la categoria di riferimento. Questa finestra di dialogo consente di specificare il valore più alto, il valore più basso o una categoria personalizzata come categoria di riferimento.

## Modello Complesso di regressione logistica Modello

**Specifico effetti del modello.** Per impostazione predefinita, la procedura costruisce un modello di effetti principali utilizzando i fattori e le covariate specificate nella finestra di dialogo principale. In alternativa, è possibile costruire un modello personalizzato che include effetti di interazione e termini nidificati.

Termini Non Nidificati

Per i fattori e le covariate selezionati:

**Interazione.** Crea il termine di interazione ad alto livello per tutte le variabili selezionate.

**Effetti principali.** Consente di creare un termine di effetti principali per ciascuna variabile selezionata.

**Tutti - 2 vie** Consente di creare tutte le possibili interazioni a due vie delle variabili selezionate.

**Tutti - 3 vie** Consente di creare tutte le possibili interazioni a tre vie delle variabili selezionate.

**Tutti - 4 vie** Consente di creare tutte le possibili interazioni a quattro vie delle variabili selezionate.

**Tutti - 5 vie** Consente di creare tutte le possibili interazioni a cinque vie delle variabili selezionate.

Termini Nidificati

Questa procedura consente di costruire termini nidificati per il modello. I termini nidificati sono utili per modellare l'effetto di un fattore o di una covariata i cui valori non interagiscono con i livelli di un altro fattore. Per esempio, una catena di supermercati può seguire le abitudini di spesa dei propri clienti in più

negozi. Poiché ogni cliente frequenta un solo negozio, l'effetto *Cliente* può definirsi **nidificato** all'interno dell'effetto *Negozi*.

È inoltre possibile includere effetti di interazione, ad esempio termini polinomiali che interessano la stessa covariata, o aggiungere più livelli di nidificazione al termine nidificato.

**Limitazione.** I termini nidificati sono sottoposti alle seguenti restrizioni:

- Tutti i fattori compresi in un'interazione devono essere univoci. Di conseguenza, se  $A$  è un fattore, non è possibile specificare  $A^*A$ .
- Tutti i fattori compresi in un effetto nidificato devono essere univoci. Di conseguenza, se  $A$  è un fattore, non è possibile specificare  $A(A)$ .
- Nessun effetto può essere nidificato all'interno di una covariata. Di conseguenza, se  $A$  è un fattore e  $X$  è una covariata, non è possibile specificare  $A(X)$ .

**Intercetta.** L'intercetta viene in genere inclusa nel modello. Se è possibile presumere che i dati passino attraverso l'origine, l'intercettazione può essere esclusa. Anche se si include l'intercettazione nel modello, è possibile scegliere di sopprimere le statistiche relative ad esso.

## Statistiche Complesse di regressione logistica

**Adattamento modello.** Controlla la visualizzazione delle statistiche che misurano le prestazioni generali del modello.

- **Pseudo R-quadrato.** La statistica  $R^2$  dalla regressione lineare non ha una controparte esatta tra i modelli di regressione logistica. Ci sono, invece, più misure che tentano di imitare le proprietà della statistica  $R^2$ .
- **Tabella di classificazione.** Visualizza le classificazioni incrociate tabulati della categoria osservata dalla categoria predittivo sulla variabile dipendente.

**Parametri.** Questo gruppo consente di controllare la visualizzazione delle statistiche relative ai parametri del modello.

- **Stima.** Visualizza le stime dei coefficienti.
- **stima esponenziata.** Visualizza la base del logaritmo naturale alzato al potere delle stime dei coefficienti. Mentre la stima ha delle belle proprietà per i test statistici, la stima esponenziale, o  $\exp(B)$ , è più facile da interpretare.
- **Errore standard.** Visualizza l'errore standard per ogni stima del coefficiente.
- **Intervallo di confidenza.** Visualizza un intervallo di confidenza per ogni stima di coefficiente. Il livello di confidenza per l'intervallo è impostato nella finestra di dialogo Opzioni.
- **prova T.** Visualizza un test  $t$  di ogni stima di coefficiente. L'ipotesi nulla per ogni prova è che il valore del coefficiente sia di 0.
- **Le covarianze delle stime dei parametri.** Visualizza una stima della matrice di covarianza per i coefficienti di modello.
- **Correlazioni delle stime dei parametri.** Visualizza una stima della matrice di correlazione per i coefficienti di modello.
- **Effetto di progettazione.** Il rapporto della varianza della stima alla varianza ottenuta ipotizzando che il campione sia un semplice campione casuale. È una misura dell'effetto derivato dalla specificazione di un disegno complesso, dove valori più piccoli indicano effetti maggiori.
- **radice di piazza dell'effetto design.** È una misura dell'effetto derivato dalla specificazione di un disegno complesso, dove valori più piccoli indicano effetti maggiori.

**Statistiche di riepilogo per le variabili del modello.** Visualizza le informazioni di riepilogo sulla variabile dipendente, le covariate e i fattori.

**Informazioni di progettazione di esempio.** Visualizza le informazioni di riepilogo sul campione, incluso il conteggio non ponderato e la dimensione della popolazione.

## Test Ipotesi di Campioni Complessi

**Statistica del test.** Questo gruppo consente di selezionare il tipo di statistica utilizzata per le ipotesi di prova. È possibile scegliere tra *F*, rettificato *F*, chi - quadrato e il chi quadrato rettificato.

**Gradi di campionamento della libertà.** Questo gruppo ti dà il controllo sui gradi di progettazione del campionamento di libertà utilizzati per calcolare i valori *p* per tutte le statistiche di prova. Se in base al disegno di campionamento, il valore è la differenza tra il numero di unità di campionamento primarie e il numero di strati nella prima fase di campionamento. In alternativa, è possibile impostare un grado di libertà personalizzato specificando un intero positivo.

**Regolazione per Confronti multipli.** Quando si eseguono test di ipotesi con più contrasti, il livello di significatività generale può essere adattato in base ai livelli di significatività dei contrasti inclusi. Questo gruppo consente di scegliere il metodo di regolazione.

- **Differenza meno significativa.** Questo metodo non controlla la probabilità generale di rifiuto delle ipotesi che alcuni contrasti lineari siano diversi dai valori di ipotesi null.
- **Sidak sequenziale.** Una procedura di Sidak con scarti sequenzialmente decrescenti, molto meno conservativa in termini di rifiuto di singole ipotesi, ma che mantiene lo stesso livello di significatività globale.
- **Bonferroni sequenziale.** Una procedura di Bonferroni con scarti sequenzialmente decrescenti, molto meno conservativa in termini di rifiuto di singole ipotesi, ma che mantiene lo stesso livello di significatività globale.
- **Sidak.** Questo metodo fornisce limiti più stretti rispetto all'approccio di Bonferroni.
- **Bonferroni.** Questo metodo corregge il livello di significatività osservato tenendo conto del fatto che si stanno verificando confronti multipli.

## Rapporti Complessi Logistic Regression Odds Ratio

La Finestra di dialogo Odds Ratio consente di visualizzare i rapporti di probabilità stimati per i fattori specificati e le covariate. Viene calcolato un insieme separato di odds ratio per ogni categoria della variabile dipendente tranne la categoria di riferimento.

**Fattori.** Per ogni fattore selezionato, visualizza il rapporto delle probabilità in ogni categoria del fattore alle probabilità nella categoria di riferimento specificata.

**Covariati.** Per ciascuna covariata selezionata, visualizza il rapporto delle probabilità al valore medio di covariato più le unità specificate di variazione alle probabilità alla media.

Quando si calcolano i rapporti di probabilità di calcolo per un fattore o covariato, la procedura corregge tutti gli altri fattori ai loro livelli più alti e tutte le altre covariate ai loro mezzi. Se un fattore o covariato interagisce con altri predittori nel modello, allora i rapporti di probabilità dipendono non solo dalla variazione della variabile specificata ma anche dai valori delle variabili con cui interagisce. Se una covariata specificata interagisce con se stessa nel modello (ad esempio, *age\*age*), allora i rapporti di probabilità dipendono sia dal cambiamento della covariata che dal valore della covariata.

## Complesso Campionamento Logistic Regression Save

**Salva Variabili.** Questo gruppo consente di salvare la categoria di tipo previsto e le probabilità previste come nuove variabili nel dataset attivo.

**Esporta modello come dati IBM SPSS Statistica .** Scrive un dataset in formato IBM SPSS Statistica contenente la correlazione dei parametri o la matrice di covarianza con le stime dei parametri, gli errori standard, i valori di significatività e i gradi di libertà. L'ordine delle variabili nel file di matrice è il seguente.

- **rowtype\_.** I valori di Takes (e le etichette di valore), COV (Covarianze), CORR (Correlazioni), EST (stime dei parametri), SE (errori standard), SIG (Livelli di Significanza) e DF (Sampling design gradi di libertà). Esiste un caso separato con tipo di riga COV (o CORR) per ogni parametro modello, più un caso separato per ciascuno degli altri tipi di riga.

- **varname\_**. Valori di Takes P1, P2, ..., corrispondente ad un elenco ordinato di tutti i parametri del modello, per i tipi di righe COV o CORR, con etichette di valore corrispondenti alle stringhe di parametro mostrate nella tabella delle stime dei parametri. Le celle sono vuote per altri tipi di righe.
- **P1, P2, ...** Queste variabili corrispondono ad un elenco ordinato di tutti i parametri del modello, con etichette variabili corrispondenti alle stringhe di parametri mostrate nella tabella delle stime dei parametri e assumono valori in base al tipo di riga. Per i parametri ridondanti, tutte le covarianze sono impostate a zero; le correlazioni sono impostate sul valore mancante di sistema; tutte le stime dei parametri sono impostate a zero; e tutti gli errori standard, i livelli di significatività e i gradi residui di libertà sono impostati sul valore mancante di sistema.

*Nota:* questo file non è immediatamente utilizzabile per ulteriori analisi in altre procedure che leggono un file di matrice a meno che tali procedure non accettino tutti i tipi di riga esportati qui.

**Modello di esportazione come XML.** Salva le stime dei parametri e la matrice di covarianza dei parametri, se selezionata, in formato XML (PMML). È possibile utilizzare questo file modello per applicare le informazioni del modello ad altri file di dati per il calcolo del punteggio.

## Opzioni Complesse di regressione logistica

**Stima.** Questo gruppo ti dà il controllo di vari criteri utilizzati nella stima del modello.

- **Numero massimo di iterazioni.** Numero massimo di iterazioni eseguite dall'algoritmo. Specificare un intero non negativo.
- **Massimo numero di dimezzamenti.** Ad ogni iterazione, la dimensione di passo viene ridotta di un fattore di 0,5 finché la verosimiglianza log non aumenta o non si raggiunge il numero massimo di dimezzamenti. Specificare un intero positivo.
- **Limitare le iterazioni in base alla modifica delle stime dei parametri.** Quando viene selezionato, l'algoritmo si arresta dopo un'iterazione in cui la variazione assoluta o relativa nelle stime dei parametri è inferiore al valore specificato, che deve essere non negativo.
- **Limitare le iterazioni in base al cambio di logaritmo.** Quando viene selezionato, l'algoritmo si arresta dopo un'iterazione in cui la modifica assoluta o relativa nella funzione di verosimiglianza è inferiore al valore specificato, che deve essere non negativo.
- **Verifica completa separazione dei punti dati.** Quando è selezionato, l'algoritmo esegue dei test per garantire che le stime dei parametri abbiano valori univoci. La separazione si verifica quando la procedura può generare un modello che classifica correttamente ogni caso.
- **Visualizza cronologia di iterazione.** Visualizza le stime del parametro e le statistiche ad ogni  $n$  iterazioni a partire dalla<sup>th</sup> iterazione (le stime iniziali). Se si sceglie di stampare la cronologia dell'iterazione, l'ultima iterazione viene sempre stampata a prescindere dal valore  $n$ .

**Valori mancanti definiti dall'utente.** Tutte le variabili di progettazione, così come la variabile dipendente e le eventuali covariate, devono avere dati validi. I casi con dati non validi per una di queste variabili vengono cancellati dall'analisi. Questi controlli consentono di decidere se i valori mancanti degli utenti sono trattati come validi tra le variabili strati, cluster, sottopopolazione e fattore.

**Intervallo di confidenza.** Questo è il livello di intervallo di confidenza per le stime dei coefficienti, stime di coefficiente esponenziali e rapporti di probabilità. Specificare un valore maggiore o uguale a 50 e inferiore a 100.

## CSLOGISTIC Comando funzioni aggiuntive

Il linguaggio della sintassi dei comandi consente inoltre di:

- Specificare test personalizzati di effetti contro una combinazione lineare di effetti o un valore (utilizzando il comando CUSTOM ).
- Correggere i valori di altre variabili del modello quando si calcolano i rapporti di probabilità per fattori e covariate (utilizzando il comando ODDSRATIOS ).
- Specificare un valore di tolleranza per la verifica della singolarità (utilizzando il comando CRITERIA ).
- Creare nomi specificati dall'utente per le variabili salvate (utilizzando il comando secondario SAVE ).

- Produrre una tabella di funzioni generali stimabili (utilizzando il sottocomando PRINT ).

Vedere *Command Syntax Reference* per informazioni dettagliate sulla sintassi.

## Regressione ordinale campioni complessi

---

La Procedura Complessa Campionaria Ordinale Regressione esegue l'analisi di regressione su una variabile dipendente o ordinale per i campioni tracciati da metodi di campionamento complessi. Facoltativamente, è possibile richiedere analisi per una sottopopolazione.

**Esempio.** I rappresentanti che valutano un disegno di legge prima della legislatura sono interessati a che ci sia un sostegno pubblico per il disegno di legge e come il sostegno al disegno di legge sia legato alla demografia dei votanti. I sondaggisti progettano e conducono interviste secondo un design di campionamento complesso. Utilizzando Complessi Campioni Ordinali Regression, è possibile adattare un modello per il livello di supporto al disegno di legge basato sulla demografia dei votanti.

Considerazioni Complesse di Regressione Dati ordinali

**Dati.** La variabile dipendente è ordinale. I fattori sono categoriali. Le covariate sono variabili quantitative correlate alla variabile dipendente. Le variabili di sottopopolazione possono essere stringa o numerica ma devono essere categoriali.

**Ipotesi.** I casi nel file dati rappresentano un campione da un design complesso che dovrebbe essere analizzato in base alle specifiche nel file selezionato nella finestra di dialogo Complex Campioni Plan.

Ottenere Campioni Complessi Ordinale Regressione

1. Dai menu, scegliere:

**Analizzare > Campioni Complessi > Regressione Ordinale ...**

2. Selezionare un file del piano. Facoltativamente, selezionare un file di probabilità congiunta personalizzato.
3. Fare clic su **Continua**.
4. Nella finestra di dialogo Complex Samples Ordinal Regression, selezionare una variabile dipendente.

Facoltativamente, è possibile:

- Selezionare le variabili per i fattori e le covariate, a seconda dei dati.
- Specificare una variabile per definire una sottopopolazione. L'analisi viene eseguita solo per la categoria selezionata della variabile sottopopolazione, sebbene le varianze siano ancora correttamente stimate in base all'intero dataset.
- Selezionare una funzione di collegamento.

**Funzione di collegamento.** La funzione di collegamento è la trasformazione delle probabilità cumulative che permette di stimare il modello. Sono disponibili le cinque funzioni di collegamento riportate di seguito.

- **Logit.**  $f(x) = \log(x / (1 - x))$ . Generalmente utilizzato per categorie distribuite in modo uniforme.
- **Log-log complementare.**  $f(x) = \log(-\log(1-x))$ . In genere, viene utilizzata quando le categorie più alte sono più probabili.
- **Log-log negativa.**  $f(x) = -\log(-\log(x))$ . In genere, viene utilizzata quando le categorie più basse sono più probabili.
- **Probit.**  $f(x) = \Phi^{-1}(x)$ . In genere, viene utilizzata quando la variabile latente viene distribuita normalmente.
- **Cauchit (Cauchy inverso).**  $f(x) = \tan(\pi(x-0.5))$ . In genere, viene utilizzata quando la variabile latente ha molti valori estremi.

## Complesso Campioni Ordinale Regressione Risposta

La finestra di dialogo Probabilità di risposta consente di specificare se la probabilità cumulativa di una risposta (cioè la probabilità di appartenere a una determinata categoria della variabile dipendente) aumenta con valori crescenti o decrescenti della variabile dipendente.

## Modello Complesso di regressione ordinale

**Specifico effetti del modello.** Per impostazione predefinita, la procedura costruisce un modello di effetti principali utilizzando i fattori e le covariate specificate nella finestra di dialogo principale. In alternativa, è possibile costruire un modello personalizzato che include effetti di interazione e termini nidificati.

Termini Non Nidificati

Per i fattori e le covariate selezionati:

**Interazione.** Crea il termine di interazione ad alto livello per tutte le variabili selezionate.

**Effetti principali.** Consente di creare un termine di effetti principali per ciascuna variabile selezionata.

**Tutti - 2 vie** Consente di creare tutte le possibili interazioni a due vie delle variabili selezionate.

**Tutti - 3 vie** Consente di creare tutte le possibili interazioni a tre vie delle variabili selezionate.

**Tutti - 4 vie** Consente di creare tutte le possibili interazioni a quattro vie delle variabili selezionate.

**Tutti - 5 vie** Consente di creare tutte le possibili interazioni a cinque vie delle variabili selezionate.

Termini Nidificati

Questa procedura consente di costruire termini nidificati per il modello. I termini nidificati sono utili per modellare l'effetto di un fattore o di una covariata i cui valori non interagiscono con i livelli di un altro fattore. Per esempio, una catena di supermercati può seguire le abitudini di spesa dei propri clienti in più negozi. Poiché ogni cliente frequenta un solo negozio, l'effetto *Cliente* può definirsi **nidificato** all'interno dell'effetto *Negozio*.

È inoltre possibile includere effetti di interazione, ad esempio termini polinomiali che interessano la stessa covariata, o aggiungere più livelli di nidificazione al termine nidificato.

**Limitazione.** I termini nidificati sono sottoposti alle seguenti restrizioni:

- Tutti i fattori compresi in un'interazione devono essere univoci. Di conseguenza, se *A* è un fattore, non è possibile specificare *A\*A*.
- Tutti i fattori compresi in un effetto nidificato devono essere univoci. Di conseguenza, se *A* è un fattore, non è possibile specificare *A(A)*.
- Nessun effetto può essere nidificato all'interno di una covariata. Di conseguenza, se *A* è un fattore e *X* è una covariata, non è possibile specificare *A(X)*.

## Complesso Campioni Ordinali Statistiche di regressione

**Adattamento modello.** Controlla la visualizzazione delle statistiche che misurano le prestazioni generali del modello.

• **Pseudo R-quadrato.** La statistica  $R^2$  dalla regressione lineare non ha una controparte esatta tra i modelli di regressione ordinale. Ci sono, invece, più misure che tentano di imitare le proprietà della statistica  $R^2$ .

• **Tabella di classificazione.** Visualizza le classificazioni incrociate tabulati della categoria osservata dalla categoria predittivo sulla variabile dipendente.

**Parametri.** Questo gruppo consente di controllare la visualizzazione delle statistiche relative ai parametri del modello.

• **Stima.** Visualizza le stime dei coefficienti.

- **stima esponenziata.** Visualizza la base del logaritmo naturale alzato al potere delle stime dei coefficienti. Mentre la stima ha delle belle proprietà per i test statistici, la stima esponenziale, o  $\exp(B)$ , è più facile da interpretare.
- **Errore standard.** Visualizza l'errore standard per ogni stima del coefficiente.
- **Intervallo di confidenza.** Visualizza un intervallo di confidenza per ogni stima di coefficiente. Il livello di confidenza per l'intervallo è impostato nella finestra di dialogo Opzioni.
- **prova T.** Visualizza un test  $t$  di ogni stima di coefficiente. L'ipotesi nulla per ogni prova è che il valore del coefficiente sia di 0.
- **Le covarianze delle stime dei parametri.** Visualizza una stima della matrice di covarianza per i coefficienti di modello.
- **Correlazioni delle stime dei parametri.** Visualizza una stima della matrice di correlazione per i coefficienti di modello.
- **Effetto di progettazione.** Il rapporto della varianza della stima alla varianza ottenuta ipotizzando che il campione sia un semplice campione casuale. È una misura dell'effetto derivato dalla specificazione di un disegno complesso, dove valori più piccoli indicano effetti maggiori.
- **radice di piazza dell'effetto design.** Si tratta di una misura, espressa in unità paragonabili a quelle dell'errore standard, dell'effetto di specificare un design complesso, dove i valori più lontani dal 1 indicano effetti maggiori.

**Linee parallele.** Questo gruppo consente di richiedere statistiche associate ad un modello con linee non parallele dove è montata una linea di regressione separata per ogni categoria di risposta (tranne l'ultima).

- **Test Wald.** Produce una prova dell'ipotesi nulla che i parametri di regressione siano uguali per tutte le risposte cumulative. Si stima il modello con linee non parallele e viene applicato il test Wald dei parametri uguali.
- **Stime dei parametri.** Visualizza le stime dei coefficienti e degli errori standard per il modello con linee non parallele.
- **Le covarianze delle stime dei parametri.** Visualizza una stima della matrice di covarianza per i coefficienti del modello con linee non parallele.

**Statistiche di riepilogo per le variabili del modello.** Visualizza le informazioni di riepilogo sulla variabile dipendente, le covariate e i fattori.

**Informazioni di progettazione di esempio.** Visualizza le informazioni di riepilogo sul campione, incluso il conteggio non ponderato e la dimensione della popolazione.

## Test Ipotesi di Campioni Complessi

**Statistica del test.** Questo gruppo consente di selezionare il tipo di statistica utilizzata per le ipotesi di prova. È possibile scegliere tra  $F$ , rettificato  $F$ , chi - quadrato e il chi quadrato rettificato.

**Gradi di campionamento della libertà.** Questo gruppo ti dà il controllo sui gradi di progettazione del campionamento di libertà utilizzati per calcolare i valori  $p$  per tutte le statistiche di prova. Se in base al disegno di campionamento, il valore è la differenza tra il numero di unità di campionamento primarie e il numero di strati nella prima fase di campionamento. In alternativa, è possibile impostare un grado di libertà personalizzato specificando un intero positivo.

**Regolazione per Confronti multipli.** Quando si eseguono test di ipotesi con più contrasti, il livello di significatività generale può essere adattato in base ai livelli di significatività dei contrasti inclusi. Questo gruppo consente di scegliere il metodo di regolazione.

- **Differenza meno significativa.** Questo metodo non controlla la probabilità generale di rifiuto delle ipotesi che alcuni contrasti lineari siano diversi dai valori di ipotesi null.
- **Sidak sequenziale.** Una procedura di Sidak con scarti sequenzialmente decrescenti, molto meno conservativa in termini di rifiuto di singole ipotesi, ma che mantiene lo stesso livello di significatività globale.

- *Bonferroni sequenziale*. Una procedura di Bonferroni con scarti sequenzialmente decrescenti, molto meno conservativa in termini di rifiuto di singole ipotesi, ma che mantiene lo stesso livello di significatività globale.
- *Sidak*. Questo metodo fornisce limiti più stretti rispetto all'approccio di Bonferroni.
- *Bonferroni*. Questo metodo corregge il livello di significatività osservato tenendo conto del fatto che si stanno verificando contrasti multipli.

## Complesso Campioni Ordinali Regression Odds Ratio

La Finestra di dialogo Odds Ratio consente di visualizzare i rapporti di odds cumulative stimati per i fattori specificati e le covariate. Questa funzione è disponibile solo per i modelli che utilizzano la funzione di collegamento Logit. Viene calcolato un unico rapporto di odds cumulativo per tutte le categorie della variabile dipendente tranne l'ultimo; il modello di odds proporzionale si postulato che sono tutti uguali.

**Fattori.** Per ogni fattore selezionato, visualizza il rapporto tra le probabilità cumulative in ogni categoria del fattore alle probabilità nella categoria di riferimento specificata.

**Covariati.** Per ciascuna covariata selezionata, visualizza il rapporto tra le probabilità cumulative al valore medio di covariato più le unità specificate di variazione alle probabilità alla media.

Quando si calcolano i rapporti di probabilità di calcolo per un fattore o covariato, la procedura corregge tutti gli altri fattori ai loro livelli più alti e tutte le altre covariate ai loro mezzi. Se un fattore o covariato interagisce con altri predittori nel modello, allora i rapporti di probabilità dipendono non solo dalla variazione della variabile specificata ma anche dai valori delle variabili con cui interagisce. Se una covariata specificata interagisce con se stessa nel modello (ad esempio, *age\*age*), allora i rapporti di probabilità dipendono sia dal cambiamento della covariata che dal valore della covariata.

## Complesso Campioni Ordinale Regression Save

**Salva Variabili.** Questo gruppo consente di salvare la categoria predite del modello, la probabilità di categoria prevista, la probabilità della categoria osservata, le probabilità cumulative e le probabilità previste come nuove variabili nel dataset attivo.

**Esporta modello come dati IBM SPSS Statistica**. Scrive un dataset in formato IBM SPSS Statistica contenente la correlazione dei parametri o la matrice di covarianza con le stime dei parametri, gli errori standard, i valori di significatività e i gradi di libertà. L'ordine delle variabili nel file di matrice è il seguente.

- **rowtype\_**. I valori di Takes (e le etichette di valore), COV (Covarianze), CORR (Correlazioni), EST (stime dei parametri), SE (errori standard), SIG (Livelli di Significanza) e DF (Sampling design gradi di libertà). Esiste un caso separato con tipo di riga COV (o CORR) per ogni parametro modello, più un caso separato per ciascuno degli altri tipi di riga.
- **varname\_**. Valori di Takes P1, P2, ..., corrispondente ad un elenco ordinato di tutti i parametri del modello, per i tipi di righe COV o CORR, con etichette di valore corrispondenti alle stringhe di parametro mostrate nella tabella delle stime dei parametri. Le celle sono vuote per altri tipi di righe.
- **P1, P2, ...** Queste variabili corrispondono ad un elenco ordinato di tutti i parametri del modello, con etichette variabili corrispondenti alle stringhe di parametri mostrate nella tabella delle stime dei parametri e assumono valori in base al tipo di riga. Per i parametri ridondanti, tutte le covarianze sono impostate a zero; le correlazioni sono impostate sul valore mancante di sistema; tutte le stime dei parametri sono impostate a zero; e tutti gli errori standard, i livelli di significatività e i gradi residui di libertà sono impostati sul valore mancante di sistema.

*Nota:* questo file non è immediatamente utilizzabile per ulteriori analisi in altre procedure che leggono un file di matrice a meno che tali procedure non accettino tutti i tipi di riga esportati qui.

**Modello di esportazione come XML.** Salva le stime dei parametri e la matrice di covarianza dei parametri, se selezionata, in formato XML (PMML). È possibile utilizzare questo file modello per applicare le informazioni del modello ad altri file di dati per il calcolo del punteggio.

## Opzioni Ordinali di regressione ordinale

**Metodo di stima.** È possibile selezionare un metodo di stima dei parametri; scegliere tra Newton - Raphson, Fisher scoring o un metodo ibrido in cui vengono eseguite le iterazioni di punteggio Fisher prima di passare al metodo Newton - Raphson. Se durante la fase di Fisher-scoring del metodo ibrido, prima del raggiungimento del numero massimo di iterazioni di Fisher, viene raggiunta la convergenza, l'algoritmo continua con il metodo di Newton-Raphson.

**Stima.** Questo gruppo ti dà il controllo di vari criteri utilizzati nella stima del modello.

- **Numero massimo di iterazioni.** Numero massimo di iterazioni eseguite dall'algoritmo. Specificare un intero non negativo.
- **Massimo numero di dimezzamenti.** Ad ogni iterazione, la dimensione di passo viene ridotta di un fattore di 0,5 finché la verosimiglianza log non aumenta o non si raggiunge il numero massimo di dimezzamenti. Specificare un intero positivo.
- **Limitare le iterazioni in base alla modifica delle stime dei parametri.** Quando viene selezionato, l'algoritmo si arresta dopo un'iterazione in cui la variazione assoluta o relativa nelle stime dei parametri è inferiore al valore specificato, che deve essere non negativo.
- **Limitare le iterazioni in base al cambio di logaritmo.** Quando viene selezionato, l'algoritmo si arresta dopo un'iterazione in cui la modifica assoluta o relativa nella funzione di verosimiglianza è inferiore al valore specificato, che deve essere non negativo.
- **Verifica completa separazione dei punti dati.** Quando è selezionato, l'algoritmo esegue dei test per garantire che le stime dei parametri abbiano valori univoci. La separazione si verifica quando la procedura può generare un modello che classifica correttamente ogni caso.
- **Visualizza cronologia di iterazione.** Visualizza le stime del parametro e le statistiche ad ogni  $n$  iterazioni a partire dalla<sup>th</sup> iterazione (le stime iniziali). Se si sceglie di stampare la cronologia dell'iterazione, l'ultima iterazione viene sempre stampata a prescindere dal valore  $n$ .

**Valori mancanti definiti dall'utente.** Le variabili di design di scala, così come la variabile dipendente e le eventuali covariate, dovrebbero avere dati validi. I casi con dati non validi per una di queste variabili vengono cancellati dall'analisi. Questi controlli consentono di decidere se i valori mancanti degli utenti sono trattati come validi tra le variabili strati, cluster, sottopopolazione e fattore.

**Intervallo di confidenza.** Questo è il livello di intervallo di confidenza per le stime dei coefficienti, stime di coefficiente esponenziali e rapporti di probabilità. Specificare un valore maggiore o uguale a 50 e inferiore a 100.

## Funzioni Aggiuntive di comando CSORDINAL

Il linguaggio della sintassi dei comandi consente inoltre di:

- Specificare test personalizzati di effetti contro una combinazione lineare di effetti o un valore (utilizzando il comando CUSTOM ).
- Fissare valori di altre variabili del modello a valori diversi dai loro mezzi quando si calcolano i rapporti di probabilità cumulativi per fattori e covariate (utilizzando il comando ODDSRATIOS ).
- Utilizzare valori non etichettati come categorie di riferimento personalizzate per i fattori quando vengono richiesti rapporti di odds (utilizzando il comando ODDSRATIOS ).
- Specificare un valore di tolleranza per la verifica della singolarità (utilizzando il comando CRITERIA ).
- Produrre una tabella di funzioni generali stimabili (utilizzando il sottocomando PRINT ).
- Salvare più di 25 variabili di probabilità (utilizzando il comando SAVE ).

Vedere *Command Syntax Reference* per informazioni dettagliate sulla sintassi.

## Regressione di Cox di campioni complessi

La procedura Complex Samples Cox Regression esegue l'analisi di sopravvivenza per campioni estratti con metodi di campionamento complessi. Facoltativamente, è possibile richiedere analisi per una sottopopolazione.

**Esempi.** Un'agenzia governativa incaricata dell'applicazione della legge è preoccupata per i tassi di recidiva nella loro area di giurisdizione. Una delle misure di recidiva è il tempo fino al secondo arresto per i delinquenti. L'agenzia vorrebbe modellare il tempo di riarresto utilizzando la regressione Cox, ma teme che l'ipotesi di rischio proporzionale non sia valida per le varie categorie di età.

I ricercatori medici indagano i tempi di sopravvivenza per i pazienti che esitano un ictus post - ischemico. C'è il potenziale per più casi per soggetto, dal momento che le storie del paziente cambiano come si fa notare l'insorgenza di eventi non morte significativi e i tempi di questi eventi registrati. Il campione viene anche lasciato troncato nel senso che i tempi di sopravvivenza osservati sono "gonfiati" dalla durata della riabilitazione, perché mentre l'insorto del rischio inizia al momento dell'ictus ischemico, solo i pazienti che sopravvivono oltre il programma di riabilitazione sono nel campione.

Campioni complessi Cox Considerazioni sui dati di regressione

**Ora di sopravvivenza.** La procedura applica la regressione Cox all'analisi dei tempi di sopravvivenza, ovvero il tempo che intercorre prima del verificarsi di un evento. Ci sono due modi per specificare il tempo di sopravvivenza, a seconda dell'ora di inizio dell'intervallo:

- **Time=0.** Comunemente si avranno informazioni complete sull'inizio dell'intervallo per ogni soggetto e avranno semplicemente una variabile contenente i tempi di fine (o creare una singola variabile con i tempi di fine dalle variabili Data & Ora; vedi sotto).
- **Varia per soggetto.** Questo è opportuno quando si ha **sinistra - troncamento**, chiamato anche **voce ritardata**; ad esempio, se si stanno analizzando i tempi di sopravvivenza per i pazienti che stanno uscendo da un programma di riabilitazione post - ictus, si potrebbe considerare che la loro insorga di rischio inizia al momento dell'ictus. Tuttavia, se il tuo campione include solo i pazienti che sono sopravvissuti al programma di riabilitazione, allora il tuo campione viene lasciato troncato nel senso che i tempi di sopravvivenza osservati sono "gonfiati" dalla durata della riabilitazione. Si può rendere conto di questo specificando il tempo in cui si è uscita la riabilitazione come momento di ingresso nello studio.

**Data & Variabili.** Le variabili Data & Ora non possono essere utilizzate per definire direttamente l'inizio e la fine dell'intervallo; se si hanno delle variabili Data & Ora, è opportuno utilizzarle per creare variabili contenenti i tempi di sopravvivenza. Se non vi è alcun tronco di sinistra, semplicemente creare una variabile contenente i tempi di fine in base alla differenza tra la data di entrata nello studio e la data di osservazione. Se c'è il tronco di sinistra, creare una variabile contenente orari di inizio, in base alla differenza tra la data di inizio dello studio e la data di ingresso, e una variabile contenente i tempi di fine, in base alla differenza tra la data di inizio dello studio e la data di osservazione.

**Stato evento.** È necessaria una variabile che registra se il soggetto ha vissuto l'evento di interesse entro l'intervallo. I soggetti per i quali l'evento non si è verificato sono giustamente censurati.

**L'identificativo oggetto.** È possibile incorporare facilmente i predittori piecewise, a tempo determinato, suddividendo le osservazioni per un singolo soggetto su più casi. Ad esempio, se si stanno analizzando i tempi di sopravvivenza per i pazienti post - ictus, le variabili che rappresentano la loro storia medica dovrebbero essere utili come predittori. Nel corso del tempo possono sperimentare importanti eventi medici che alterano la loro storia clinica. La tabella seguente mostra come strutturare un tale dataset: *ID paziente* è l'identificativo del soggetto, *Ora di fine* definisce gli intervalli osservati, *Stato* registra importanti eventi medici e *Storia precedente dell'infarto* e *Storia precedente dell'emorragia* sono predittori piecewise - costanti.

Tabella 1. Struttura dati per incorporare i predittori a tempo determinato piecewise				
<b>ID paziente</b>	<b>Ora di fine</b>	<b>Stato</b>	<b>Storia precedente dell'attacco cardiaco</b>	<b>Storia precedente dell'emorragia</b>
1	5	Attacco Cardiaco	N	N

Tabella 1. Struttura dati per incorporare i predittori a tempo determinato piecewise (Continua)

ID paziente	Ora di fine	Stato	Storia precedente dell'attacco cardiaco	Storia precedente dell'emorragia
1	7	Emorragia	Sì	N
1	8	Inattivo	Sì	Sì
2	24	Inattivo	N	N
3	8	Attacco Cardiaco	N	N
3	15	Inattivo	Sì	N

**Ipotesi.** I casi nel file dati rappresentano un campione da un design complesso che dovrebbe essere analizzato in base alle specifiche nel file selezionato nella finestra di dialogo [Complex Campioni Plan](#).

In genere, i modelli di regressione Cox assumono rischi proporzionali, ossia il rapporto dei rischi tra un caso e l'altro non dovrebbe variare nel tempo. Se questa assunzione non tiene, potrebbe essere necessario aggiungere dei predittori dipendenti dal tempo al modello.

**Analisi di Kaplan-Meier.** Se non si seleziona alcun preditore (o non si immette alcun preditore selezionato nel modello) e si sceglie il metodo limite del prodotto per il calcolo della curva di sopravvivenza al basale sulla scheda Opzioni, la procedura esegue un tipo di analisi di sopravvivenza Kaplan - Meier.

Per ottenere campioni complessi Cox Regressione

1. Dai menu, scegliere:

**Analizzare > campioni complessi > Cox Regressione...**

2. Selezionare un file del piano. Facoltativamente, selezionare un file di probabilità congiunta personalizzato.
3. Fare clic su **Continua**.
4. Specificare il tempo di sopravvivenza selezionando i tempi di entrata e di uscita dallo studio.
5. Selezionare una variabile di stato dell'evento.
6. Fare clic su Definisci evento e definire almeno un valore evento.

Facoltativamente, è possibile selezionare un identificativo di soggetto.

## Definire l'evento

Specificare i valori che indicano un evento terminale si è verificato.

- **Valore individuale (s).** Specificare uno o più valori inserendoli nella griglia o selezionandoli da un elenco di valori con etichette di valore definite.
- **Intervallo di valori.** Specificare una gamma di valori inserendo i valori minimi e massimi o selezionando i valori da un elenco con etichette di valore definite.

## Predittori

Scheda Predittori consente di specificare i fattori e le covariate utilizzate per costruire effetti di modello.

**Fattori.** I fattori sono predittori categoriali; possono essere numerici o stringa.

**Covariati.** Le covariate sono predittori di scala; devono essere numeriche.

**Predicatori tempo - dipendenti.** Ci sono certe situazioni in cui l'assunzione dei rischi proporzionali non regge. Ovvero i rapporti di pericolo cambiano nel tempo; i valori di uno (o più) dei tuoi predittori sono diversi in momenti diversi. In questi casi, è necessario specificare predittori dipendenti dal tempo. Per ulteriori informazioni, consultare l'argomento [“Definire il predittore dipendente dal tempo”](#) a pagina 33. I predittori dipendenti dal tempo possono essere selezionati come fattori o covariate.

## Definire il predittore dipendente dal tempo

La Define Time - Dipendente Predictor finestra di dialogo consente di creare un predittore che dipende dalla variabile temporale integrata,  $T_.$ . È possibile utilizzare questa variabile per definire covariate a tempo determinato in due modi generali:

- Se si desidera stimare un modello di regressione Cox esteso che ammette rischi non proporzionali, è possibile farlo definendo il predittore dipendente dal tempo come una funzione della variabile temporale  $T_$  e della covariata in questione. Un esempio comune sarebbe il semplice prodotto della variabile del tempo e il predittore, ma possono essere specificate anche funzioni più complesse.
- Alcune variabili possono avere valori diversi in periodi di tempo diversi ma non sono sistematicamente correlati al tempo. In questi casi, è necessario definire un **predicatore a tempo determinato**, che può essere fatto utilizzando espressioni logiche. Le espressioni logiche assumono il valore 1 se vero e 0 se false. Utilizzando una serie di espressioni logiche, è possibile creare il proprio predicatore a tempo determinato da una serie di misurazioni. Ad esempio, se si ha la pressione arteriosa misurata una volta alla settimana per le quattro settimane del tuo studio (identificato come  $BP1$  a  $BP4$ ), è possibile definire il tuo predittore dipendente dal tempo come  $(T_ < 1) * BP1 + (T_ > = 1 \& T_ < 2) * BP2 + (T_ > = 2 \& T_ < 3) * BP3 + (T_ > = 3 \& T_ < 4) * BP4$ . Si noti che esattamente uno dei termini tra parentesi sarà pari a 1 per ogni caso dato e il resto sarà uguale a tutti gli 0. In altre parole, questa funzione significa che se il tempo è inferiore a una settimana, utilizzare  $BP1$ ; se è superiore a una settimana ma inferiore a due settimane, utilizzare  $BP2$  e così via.

*Nota:* Se il tuo predittore segmentato, tempo - dipendente è costante all'interno dei segmenti, come nell'esempio di pressione arteriosa indicato sopra, può essere più facile per te specificare il predicatore piecewise, dipendente dal tempo, suddividendolo in più casi. Consultare la discussione su Oggetti identificativi in [“Regressione di Cox di campioni complessi”](#) a pagina 31 per ulteriori informazioni.

Nella finestra di dialogo Definizioni Tempo - Dipendente Dipendente è possibile utilizzare i controlli della funzione - building per costruire l'espressione per la covariata dipendente dal tempo, oppure è possibile inserirla direttamente nell'area di testo dell'espressione numerica. Notare che le costanti di stringa devono essere racchiuse tra virgolette o apostrofi e le costanti numeriche devono essere digitate in formato americano, con il punto come delimitatore decimale. Alla variabile risultante viene dato il nome che si specifica e deve essere incluso come fattore o covariato nella scheda Predittori.

## Sottogruppi

**Baseline Strati.** Per ogni valore di questa variabile viene calcolata una funzione di pericolo di baseline separata e di sopravvivenza, mentre si stima un unico insieme di coefficienti di modello in strati.

**Variabile di sottopopolazione.** Specificare una variabile per definire una sottopopolazione. L'analisi viene eseguita solo per la categoria selezionata della variabile sottopopolazione.

## Modello

**Specificare effetti del modello.** Per impostazione predefinita, la procedura costruisce un modello di effetti principali utilizzando i fattori e le covariate specificate nella finestra di dialogo principale. In alternativa, è possibile costruire un modello personalizzato che include effetti di interazione e termini nidificati.

Termini Non Nidificati

Per i fattori e le covariate selezionati:

**Interazione.** Crea il termine di interazione ad alto livello per tutte le variabili selezionate.

**Effetti principali.** Consente di creare un termine di effetti principali per ciascuna variabile selezionata.

**Tutti - 2 vie** Consente di creare tutte le possibili interazioni a due vie delle variabili selezionate.

**Tutti - 3 vie** Consente di creare tutte le possibili interazioni a tre vie delle variabili selezionate.

**Tutti - 4 vie** Consente di creare tutte le possibili interazioni a quattro vie delle variabili selezionate.

**Tutti - 5 vie** Consente di creare tutte le possibili interazioni a cinque vie delle variabili selezionate.

## Termini Nidificati

Questa procedura consente di costruire termini nidificati per il modello. I termini nidificati sono utili per modellare l'effetto di un fattore o di una covariata i cui valori non interagiscono con i livelli di un altro fattore. Per esempio, una catena di supermercati può seguire le abitudini di spesa dei propri clienti in più negozi. Poiché ogni cliente frequenta un solo negozio, l'effetto *Cliente* può definirsi **nidificato** all'interno dell'effetto *Negozi*.

È inoltre possibile includere effetti di interazione, ad esempio termini polinomiali che interessano la stessa covariata, o aggiungere più livelli di nidificazione al termine nidificato.

**Limitazione.** I termini nidificati sono sottoposti alle seguenti restrizioni:

- Tutti i fattori compresi in un'interazione devono essere univoci. Di conseguenza, se  $A$  è un fattore, non è possibile specificare  $A^*A$ .
- Tutti i fattori compresi in un effetto nidificato devono essere univoci. Di conseguenza, se  $A$  è un fattore, non è possibile specificare  $A(A)$ .
- Nessun effetto può essere nidificato all'interno di una covariata. Di conseguenza, se  $A$  è un fattore e  $X$  è una covariata, non è possibile specificare  $A(X)$ .

## Statistica

**Informazioni di progettazione di esempio.** Visualizza le informazioni di riepilogo sul campione, incluso il conteggio non ponderato e la dimensione della popolazione.

**Riepilogo eventi e censuramenti.** Visualizza le informazioni di riepilogo sul numero e la percentuale di casi censurati.

**Rischio impostato nei tempi di evento.** Visualizza il numero di eventi e numero a rischio per ogni ora di evento in ogni strato di base.

**Parametri.** Questo gruppo consente di controllare la visualizzazione delle statistiche relative ai parametri del modello.

- **Stima.** Visualizza le stime dei coefficienti.
- **stima esponenziata.** Visualizza la base del logaritmo naturale alzato al potere delle stime dei coefficienti. Mentre la stima ha delle belle proprietà per i test statistici, la stima esponenziale, o  $\exp(B)$ , è più facile da interpretare.
- **Errore standard.** Visualizza l'errore standard per ogni stima del coefficiente.
- **Intervallo di confidenza.** Visualizza un intervallo di confidenza per ogni stima di coefficiente. Il livello di confidenza per l'intervallo è impostato nella finestra di dialogo Opzioni.
- **T-Test.** Visualizza un test  $t$  di ogni stima di coefficiente. L'ipotesi nulla per ogni prova è che il valore del coefficiente sia di 0.
- **Le covarianze delle stime dei parametri.** Visualizza una stima della matrice di covarianza per i coefficienti di modello.
- **Correlazioni delle stime dei parametri.** Visualizza una stima della matrice di correlazione per i coefficienti di modello.
- **Effetto di progettazione.** Il rapporto della varianza della stima alla varianza ottenuta ipotizzando che il campione sia un semplice campione casuale. È una misura dell'effetto derivato dalla specificazione di un disegno complesso, dove valori più piccoli indicano effetti maggiori.
- **radice di piazza dell'effetto design.** È una misura dell'effetto derivato dalla specificazione di un disegno complesso, dove valori più piccoli indicano effetti maggiori.

**Ipotesi modello.** Questo gruppo consente di produrre una prova dell'assunzione di rischi proporzionali. Il test confronta il modello montato su un modello alternativo che include predittori dipendenti dal tempo  $x^*_TF$  per ogni predicatore  $x$ , dove  $_TF$  è la funzione di tempo specificata.

- **Funzione tempo.** Specifica il modulo di  $_TF$  per il modello alternativo. Per la funzione **identity**,  $_TF=T_$ . Per la funzione **log**,  $_TF=\log(T_)$ . Per **Kaplan - Meier**,  $_TF=1-S_{KM}(T_)$ , dove  $S_{KM}(\cdot)$  è la stima Kaplan - Meier della funzione di sopravvivenza. Per **rank**,  $_TF$  è il rango - ordine di  $T_$  tra i tempi di fine osservati.
- **Le stime dei parametri per il modello alternativo.** Visualizza il preventivo, l'errore standard e l'intervallo di confidenza per ogni parametro nel modello alternativo.
- **matrice di Covarianza per modello alternativo.** Visualizza la matrice delle covarianze stimate tra parametri nel modello alternativo.

**Sopravvivenza baseline e funzioni di pericolo cumulativo.** Visualizza la funzione di sopravvivenza baseline e la funzione di rischio cumulativo di base insieme ai relativi errori standard.

*Nota:* se nel modello sono inclusi i predittori dipendenti dal tempo definiti nella scheda Predittori, questa opzione non è disponibile.

## Grafici

La scheda Plotti consente di richiedere complotti della funzione di pericolo, funzione di sopravvivenza, log - minus - log della funzione di sopravvivenza e uno meno la funzione di sopravvivenza. È inoltre possibile scegliere di tracciare intervalli di confidenza lungo le funzioni specificate; il livello di confidenza è impostato sulla scheda Opzioni.

**Modelli di previsione.** È possibile specificare un modello di valori predittori da utilizzare per i complotti richiesti e il file di sopravvivenza esportato nella scheda Export. Si noti che queste opzioni non sono disponibili se i predittori dipendenti dal tempo definiti sulla scheda Predittori sono inclusi nel modello.

- **Fattori di trama a.** Per impostazione predefinita, ogni fattore viene valutato al suo livello più alto. Inserire o selezionare un livello diverso se desiderato. In alternativa, è possibile scegliere di tracciare linee separate per ogni livello di un singolo fattore selezionando la casella di spunta per quel fattore.
- **Trama Covariati at.** Ogni covariata viene valutata alla sua media. Inserire o selezionare un valore diverso se si desidera.

## Test di ipotesi

**Statistica del test.** Questo gruppo consente di selezionare il tipo di statistica utilizzata per le ipotesi di prova. È possibile scegliere tra  $F$ , rettificato  $F$ , chi - quadrato e il chi quadrato rettificato.

**Gradi di campionamento della libertà.** Questo gruppo ti dà il controllo sui gradi di progettazione del campionamento di libertà utilizzati per calcolare i valori  $p$  per tutte le statistiche di prova. Se in base al disegno di campionamento, il valore è la differenza tra il numero di unità di campionamento primarie e il numero di strati nella prima fase di campionamento. In alternativa, è possibile impostare un grado di libertà personalizzato specificando un intero positivo.

**Regolazione per Confronti multipli.** Quando si eseguono test di ipotesi con più contrasti, il livello di significatività generale può essere adattato in base ai livelli di significatività dei contrasti inclusi. Questo gruppo consente di scegliere il metodo di regolazione.

- **Differenza meno significativa.** Questo metodo non controlla la probabilità generale di rifiuto delle ipotesi che alcuni contrasti lineari siano diversi dai valori di ipotesi null.
- **Sidak sequenziale.** Una procedura di Sidak con scarti sequenzialmente decrescenti, molto meno conservativa in termini di rifiuto di singole ipotesi, ma che mantiene lo stesso livello di significatività globale.
- **Bonferroni sequenziale.** Una procedura di Bonferroni con scarti sequenzialmente decrescenti, molto meno conservativa in termini di rifiuto di singole ipotesi, ma che mantiene lo stesso livello di significatività globale.
- **Sidak.** Questo metodo fornisce limiti più stretti rispetto all'approccio di Bonferroni.
- **Bonferroni.** Questo metodo corregge il livello di significatività osservato tenendo conto del fatto che si stanno verificando contrasti multipli.

## Salva

**Salva Variabili.** Questo gruppo consente di salvare le variabili relative al modello al dataset attivo per un ulteriore utilizzo nella diagnostica e la segnalazione dei risultati. Da notare che nessuno di questi è disponibile quando i predittori dipendenti dal tempo sono inclusi nel modello.

- **Funzione Sopravvivenza.** Salva la probabilità di sopravvivenza (il valore della funzione di sopravvivenza) ai valori temporali e predittivi osservati per ogni caso.
- **Abbassa l'intervallo di confidenza per la funzione di sopravvivenza.** Salva il limite inferiore dell'intervallo di confidenza per la funzione di sopravvivenza ai valori temporali e predittivi osservati per ogni caso.
- **Alto limite di intervallo di confidenza per la funzione di sopravvivenza.** Salva il limite superiore dell'intervallo di confidenza per la funzione di sopravvivenza ai valori temporali e predittivi osservati per ogni caso.
- **Funzione di rischio cumulativo.** Salva il pericolo cumulativo, o meglio si tratta di  $\ln$  (sopravvivenza), ai valori temporali e predittivi osservati per ogni caso.
- **Abbassa l'intervallo di confidenza per la funzione di pericolo cumulativo.** Salva il limite inferiore dell'intervallo di confidenza per la funzione di pericolo cumulativo nei valori di tempo e predittori osservati per ogni caso.
- **Intervallo di confidenza superiore per la funzione di pericolo cumulativo.** Salva il limite superiore dell'intervallo di confidenza per la funzione di pericolo cumulativo nei valori temporali e predittivi osservati per ogni caso.
- **Valore previsto di predittore lineare.** Salva la combinazione lineare di predittori corretti per i coefficienti di regressione dei tempi di regressione. Il predittore lineare è il rapporto tra la funzione di pericolo e il pericolo di base. Sotto il modello di rischio proporzionale, questo valore è costante nel tempo.
- **residua Schoenfeld.** Per ogni caso non censurato e ogni parametro non ridondante nel modello, il residuo Schoenfeld è la differenza tra il valore osservato del predittore associato al parametro del modello e il valore previsto del predittore per i casi nel rischio impostato al momento dell'evento osservato. I residui di Schoenfeld possono essere utilizzati per aiutare a valutare l'assunzione dei rischi proporzionali; ad esempio, per un predittore  $x$ , i tracciati dei residui di Schoenfeld per il predittore dipendente dal tempo  $x^* \ln(T_)$  contro il tempo dovrebbero mostrare una linea orizzontale a 0 se i pericoli proporzionali trattiene. Una variabile separata viene salvata per ogni parametro non ridondante nel modello. I residui di Schoenfeld vengono calcolati solo per casi non censurati.
- **residuo di Martingale.** Per ogni caso, il residuo martingale è la differenza tra la censura osservata (0 se censurata, 1 se non) e l'aspettativa di un evento durante il tempo di osservazione.
- **Devianza residuo.** I residui di devianza sono residui di martingale "regolati" per apparire più simmetrici circa 0. I lotti di residui di devianza contro i predittori non dovrebbero rivelare schemi.
- **Cox -Residuo Snell.** Per ogni caso, il residuo di Cox -Snell è l'aspettativa di un evento durante il tempo di osservazione, ovvero la censura osservata meno il residuo della martingala.
- **Punteggio residuo.** Per ogni caso e ogni parametro non ridondante nel modello, il punteggio residuo è il contributo del caso al primo derivato della pseudo verosimigliante. Una variabile separata viene salvata per ogni parametro non ridondante nel modello.
- **residuo DFBeta.** Per ogni caso e ogni parametro non ridondante nel modello, il residuo DFBeta approssimerà la variazione del valore della stima del parametro quando il caso viene rimosso dal modello. I casi con residui di DFBeta relativamente grandi possono esercitare un'influenza indebita sull'analisi. Una variabile separata viene salvata per ogni parametro non ridondante nel modello.
- **Residui aggregati.** Quando più casi rappresentano un singolo soggetto, il residuo aggregato per un soggetto è semplicemente la somma dei residui di caso corrispondenti su tutti i casi appartenenti allo stesso soggetto. Per il residuo di Schoenfeld, la versione aggregata è la stessa della versione non aggregata perché il residuo di Schoenfeld è definito solo per i casi non censurati. Questi residui sono disponibili solo quando viene specificato un identificativo di soggetto sulla scheda Tempo e Evento.

**Nomi delle variabili salvate.** La generazione automatica dei nomi garantisce che tu mantenga tutto il tuo lavoro. I nomi personalizzati consentono di eliminare/sostituire i risultati di precedenti esecuzioni senza dover prima eliminare le variabili salvate nell'Editor dei dati.

## Esportazione

**Esporta modello come dati IBM SPSS Statistica .** Scrive un dataset in formato IBM SPSS Statistica contenente la correlazione dei parametri o la matrice di covarianza con le stime dei parametri, gli errori standard, i valori di significatività e i gradi di libertà. L'ordine delle variabili nel file di matrice è il seguente.

- **rowtype\_**. I valori di Takes (e le etichette di valore), COV (Covarianze), CORR (Correlazioni), EST (stime dei parametri), SE (errori standard), SIG (Livelli di Significanza) e DF (Sampling design gradi di libertà). Esiste un caso separato con tipo di riga COV (o CORR) per ogni parametro modello, più un caso separato per ciascuno degli altri tipi di riga.
- **varname\_**. Valori di Takes P1, P2, ..., corrispondente ad un elenco ordinato di tutti i parametri del modello, per i tipi di righe COV o CORR, con etichette di valore corrispondenti alle stringhe di parametro mostrate nella tabella delle stime dei parametri. Le celle sono vuote per altri tipi di righe.
- **P1, P2, ...** Queste variabili corrispondono ad un elenco ordinato di tutti i parametri del modello, con etichette variabili corrispondenti alle stringhe di parametri mostrate nella tabella delle stime dei parametri e assumono valori in base al tipo di riga. Per i parametri ridondanti, tutte le covarianze sono impostate a zero; le correlazioni sono impostate sul valore mancante di sistema; tutte le stime dei parametri sono impostate a zero; e tutti gli errori standard, i livelli di significatività e i gradi residui di libertà sono impostati sul valore mancante di sistema.

*Nota:* questo file non è immediatamente utilizzabile per ulteriori analisi in altre procedure che leggono un file di matrice a meno che tali procedure non accettino tutti i tipi di riga esportati qui.

**Esporta funzione di sopravvivenza come dati IBM SPSS Statistica .** Scrive un dataset in formato IBM SPSS Statistica contenente la funzione di sopravvivenza; errore standard della funzione di sopravvivenza; limiti superiori e inferiori dell'intervallo di confidenza della funzione di sopravvivenza; e la funzione di rischio cumulativo per ogni guasto o momento dell'evento, valutato alla base e nei modelli predittori specificati sulla scheda Plot. L'ordine delle variabili nel file di matrice è il seguente.

- **Variabile di strati baseline.** Vengono prodotte tabelle di sopravvivenza separate per ogni valore della variabile strati.
- **Variabile tempo di sopravvivenza.** Il tempo dell'evento; viene creato un caso separato per ogni ora di evento univoco.
- **Sur\_0, LCL\_Sur\_0, UCL\_Sur\_0.** Funzione di sopravvivenza al basale e limiti superiori e inferiori del suo intervallo di confidenza.
- **Sur\_R, LCL\_Sur\_R, UCL\_Sur\_R.** Funzione di sopravvivenza valutata allo schema "reference" (vedere la tabella dei valori del pattern nell'output) e i limiti superiori e inferiori del suo intervallo di confidenza.
- **Sur\_#. #, LCL\_Sur\_#. #, UCL\_Sur\_#. #, ...** Funzione di sopravvivenza valutata in ognuno dei modelli di predittore specificati sulla scheda Grado e i limiti superiori e inferiori dei relativi intervalli di confidenza. Vedere la tabella dei valori di modello nell'output per abbinare i modelli con il numero #. #.
- **Haz\_0, LCL\_Haz\_0, UCL\_Haz\_0.** Funzione di pericolo cumulativo di base e limiti superiori e inferiori del suo intervallo di confidenza.
- **Haz\_R, LCL\_Haz\_R, UCL\_Haz\_R.** Funzione di pericolo cumulativa valutata allo schema "reference" (vedere la tabella dei valori di modello nell'output) e i limiti superiori e inferiori del suo intervallo di confidenza.
- **Haz\_#. #, LCL\_Haz\_#. #, UCL\_Haz\_#. #, ...** Funzione di pericolo cumulativa valutata in ciascuno dei modelli di predittore specificati sulla scheda Grado e i limiti superiori e inferiori dei relativi intervalli di confidenza. Vedere la tabella dei valori di modello nell'output per abbinare i modelli con il numero #. #.

**Modello di esportazione come XML.** Saves tutte le informazioni necessarie per prevedere la funzione di sopravvivenza, incluse le stime dei parametri e la funzione di sopravvivenza al basale, in formato XML (PMML). È possibile utilizzare questo file modello per applicare le informazioni del modello ad altri file di dati per il calcolo del punteggio.

## Opzioni

**Stima.** Questi controlli specificano criteri di stima dei coefficienti di regressione.

- **Numero massimo di iterazioni.** Numero massimo di iterazioni eseguite dall'algoritmo. Specificare un intero non negativo.
- **Massimo numero di dimezzamenti.** Ad ogni iterazione, la dimensione di passo viene ridotta di un fattore di 0,5 finché la verosimiglianza log non aumenta o non si raggiunge il numero massimo di dimezzamenti. Specificare un intero positivo.
- **Limitare le iterazioni in base alla modifica delle stime dei parametri.** Quando è selezionato, l'algoritmo si arresta dopo un'iterazione nella quale la variazione assoluta o relativa nelle stime dei parametri è minore del valore specificato, che deve essere positivo.
- **Limitare le iterazioni in base al cambio di logaritmo.** Quando è selezionato, l'algoritmo si arresta dopo un'iterazione nella quale la variazione assoluta o relativa nella funzione di verosimiglianza log è minore del valore specificato, che deve essere positivo.
- **Visualizza cronologia di iterazione.** Visualizza la cronologia delle iterazioni per le stime dei parametri e la pseudo logaritmo della verosimiglianza e stampa l'ultima valutazione della modifica delle stime dei parametri e della pseudo logaritmo. La tabella della cronologia delle iterazioni stampa ogni  $n$  iterazioni che iniziano con la  $0$ -esima iterazione (le stime iniziali), dove  $n$  è il valore dell'incremento. Se è richiesta la cronologia delle iterazioni, l'ultima iterazione viene sempre visualizzata indipendentemente da  $n$ .
- **Tie di rottura metodo per la stima dei parametri.** Quando ci sono i tempi di insuccesso osservati, uno di questi metodi viene utilizzato per spezzare i legami. Il metodo Efron è più computazionalmente costoso.

**Funzioni di sopravvivenza.** Questi controlli specificano criteri per le elaborazioni che coinvolgono la funzione di sopravvivenza.

- **Metodo per stimare le funzioni di sopravvivenza baseline.** Il metodo **Breslow** (o Nelson - Aalan o empirico) stima il pericolo cumulativo di base da una funzione di passo non decrescente con passaggi ai tempi di insuccesso osservati, calcola quindi la sopravvivenza al basale da parte della relazione sopravviv= $\exp$  (rischio cumulativo). Il metodo **Efron** è più computazionalmente costoso e si riduce al metodo Breslow quando non ci sono legami. Il metodo **limite prodotto** stima la sopravvivenza al basale grazie ad una funzione continua non in continuo aumento; quando non ci sono predittori nel modello, questo metodo riduce la stima Kaplan - Meier.
- **Intervalli di confidenza delle funzioni di sopravvivenza.** L'intervallo di confidenza può essere calcolato in tre modi: in unità originali, tramite una trasformazione log o una trasformazione log - minus - log. Solo la trasformazione log - minus - log garantisce che i limiti dell'intervallo di confidenza risiederanno tra il 0 e il 1, ma la trasformazione del log generalmente sembra eseguire "al meglio".

**Valori utente mancanti.** Tutte le variabili devono avere valori validi per un caso da inserire nell'analisi. Questi controlli consentono di decidere se i valori mancanti dell'utente sono trattati come validi tra i modelli categoriali (inclusi i fattori, le variabili evento, strati e sottopopolazione) e le variabili di progettazione del campionamento.

**Intervallo di confidenza (%).** Questo è il livello di intervallo di confidenza utilizzato per le stime del coefficiente, le stime del coefficiente esponenziato, le stime della funzione di sopravvivenza e le stime di funzione di pericolo cumulativo. Specificare un valore maggiore o uguale a 0, e inferiore a 100.

## Funzioni Aggiuntive Comando CSCOXREG

Il linguaggio di comando consente anche di:

- Eseguire test di ipotesi personalizzati (utilizzando il comando CUSTOM e /PRINT LMATRIX).
- Specifica di tolleranza (utilizzando /CRITERIA SINGULAR).
- Tabella delle funzioni generali stimabili (utilizzare /PRINT GEF).
- Più modelli predittori (utilizzando più sottocomandi PATTERN ).

- Numero massimo di variabili salvate quando viene specificato un rootname (utilizzando il comando SAVE). La finestra di dialogo onora il valore predefinito CSCOXREG di 25 variabili.

Vedere *Command Syntax Reference* per informazioni dettagliate sulla sintassi.



## Informazioni particolari

---

Queste informazioni sono state sviluppate per prodotti e servizi offerti negli Stati Uniti. Questo materiale potrebbe essere disponibile da IBM in altre lingue. Tuttavia, può essere necessario possedere una copia del prodotto o versione del prodotto in tale lingua per potervi accedere.

IBM può non offrire i prodotti, i servizi o le funzioni presentati in questo documento in altri paesi. Consultare il proprio rappresentante locale IBM per informazioni sui prodotti ed i servizi attualmente disponibili nella propria zona. Qualsiasi riferimento ad un prodotto, programma o servizio IBM non implica o intende dichiarare che solo quel prodotto, programma o servizio IBM può essere utilizzato. Qualsiasi prodotto funzionalmente equivalente al prodotto, programma o servizio che non violi alcun diritto di proprietà intellettuale IBM può essere utilizzato. È comunque responsabilità dell'utente valutare e verificare la possibilità di utilizzare altri prodotti, programmi o servizi non IBM.

IBM può avere applicazioni di brevetti o brevetti in corso relativi all'argomento descritto in questo documento. La fornitura del presente documento non concede alcuna licenza a tali brevetti. È possibile inviare per iscritto richieste di licenze a:

*IBM Director of Licensing  
IBM Corporation  
North Castle Drive, MD-NC119  
Armonk, NY 10504-1785  
it*

Per richieste di licenze relative ad informazioni double-byte (DBCS), contattare il Dipartimento di Proprietà Intellettuale IBM nel proprio paese o inviare richieste per iscritto a:

*Intellectual Property Licensing  
Legal and Intellectual Property Law  
IBM Japan Ltd.  
19-21, Nihonbashi-Hakozakicho, Chuo-ku  
Tokyo 103-8510, Japan*

IBM (INTERNATIONAL BUSINESS MACHINES CORPORATION) FORNISCE LA PRESENTE PUBBLICAZIONE "NELLO STATO IN CUI SI TROVA" SENZA GARANZIE DI ALCUN TIPO, ESPRESSE O IMPLICITE, IVI INCLUSE, A TITOLO DI ESEMPIO, GARANZIE IMPLICITE DI NON VIOLAZIONE, DI COMMERCIALIBITÀ E DI IDONEITÀ PER UNO SCOPO PARTICOLARE. Alcune giurisdizioni non consentono la rinuncia ad alcune garanzie espresse o implicite in determinate transazioni, pertanto, la presente dichiarazione può non essere applicabile.

Queste informazioni potrebbero contenere imprecisioni tecniche o errori tipografici. Le modifiche alle presenti informazioni vengono effettuate periodicamente; tali modifiche saranno incorporate nelle nuove pubblicazioni della pubblicazione. IBM si riserva il diritto di apportare miglioramenti e/o modifiche al prodotto o al programma descritto nel manuale in qualsiasi momento e senza preavviso.

I riferimenti in queste informazioni a siti Web non IBM vengono forniti solo per comodità e non implicano in alcun modo l'approvazione di tali siti web. I materiali presenti in tali siti web non sono parte dei materiali per questo prodotto IBM e l'utilizzo di tali siti è a proprio rischio.

IBM può utilizzare o distribuire qualsiasi informazione fornita in qualsiasi modo ritenga appropriato senza incorrere in alcun obbligo verso l'utente.

Coloro che detengono la licenza su questo programma e desiderano avere informazioni su di esso allo scopo di consentire: (i) uno scambio di informazioni tra programmi indipendenti ed altri (compreso questo) e (ii) l'utilizzo reciproco di tali informazioni, dovrebbe rivolgersi a:

*IBM Director of Licensing  
IBM Corporation  
North Castle Drive, MD-NC119*

*Armonk, NY 10504-1785*

*it*

Tali informazioni possono essere disponibili, in base ad appropriate clausole e condizioni, includendo in alcuni casi, il pagamento di una tassa.

Il programma concesso in licenza descritto nel presente documento e tutto il materiale concesso in licenza disponibile sono forniti da IBM in base alle clausole dell'Accordo per Clienti IBM (IBM Customer Agreement), dell'IBM IPLA (IBM International Program License Agreement) o qualsiasi altro accordo equivalente tra le parti.

Gli esempi client e dati delle prestazioni citati sono presentati esclusivamente a scopo illustrativo. Gli effettivi risultati delle prestazioni possono variare in base alle configurazioni e alle condizioni operative specifiche.

Le informazioni relative a prodotti non IBM sono ottenute dai fornitori di quei prodotti, dagli annunci pubblicati i da altre fonti disponibili al pubblico. IBM non ha testato quei prodotti e non può confermarne la precisione della prestazione, la compatibilità o qualsiasi altro reclamo relativo ai prodotti non IBM. Le domande sulle funzionalità dei prodotti non IBM devono essere indirizzate ai fornitori di tali prodotti.

Le dichiarazioni relative all'orientamento o alle intenzioni future di IBM sono soggette a modifica o a ritiro senza preavviso e rappresentano solo mete e obiettivi.

Questa pubblicazione contiene esempi di dati e prospetti utilizzati quotidianamente nelle operazioni aziendali. Pertanto, per maggiore completezza, gli esempi includono nomi di persone, società, marchi e prodotti. Tutti i nomi contenuti nel manuale sono fittizi e ogni riferimento a persone o aziende reali è puramente casuale.

#### LICENZA DI COPYRIGHT:

Queste informazioni contengono programmi applicativi di esempio in linguaggio sorgente, che illustrano tecniche di programmazione su varie piattaforme operative. È possibile copiare, modificare e distribuire questi programmi di esempio sotto qualsiasi forma senza alcun pagamento a IBM, allo scopo di sviluppare, utilizzare, commercializzare o distribuire i programmi applicativi in conformità alle API (application programming interface) a seconda della piattaforma operativa per cui i programmi di esempio sono stati scritti. Questi esempi non sono stati testati approfonditamente tenendo conto di tutte le condizioni possibili. IBM, quindi, non può garantire o sottintendere l'affidabilità, l'utilità o il funzionamento di questi programmi. I programmi di esempio sono forniti "NELLO STATO IN CUI SI TROVANO", senza alcun tipo di garanzia. IBM non intende essere responsabile per alcun danno derivante dall'uso dei programmi di esempio.

Ogni copia o qualsiasi parte di questi programmi di esempio o qualsiasi lavoro derivato, devono contenere le seguenti informazioni relative alle leggi sul diritto d'autore:

© Copyright IBM Corp. 2021. Le porzioni di questo codice derivano da IBM Corp. Programmi di esempio.

© Copyright IBM Corp. 1989 - 2021. Tutti i diritti riservati.

## Marchi

---

IBM, il logo IBM e ibm.com sono marchi o marchi registrati di International Business Machines Corp., registrati in molte giurisdizioni in tutto il mondo. Altri nomi di prodotti e servizi possono essere marchi di IBM o di altre società. Un elenco corrente dei marchi IBM è disponibile sul web in "Copyright and trademark information" all'indirizzo [www.ibm.com/legal/copytrade.shtml](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.

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 della The Open Group negli Stati Uniti e/o negli altri paesi.

Java e tutti i marchi e i logo basati su Java sono marchi o marchi registrati di Oracle e/o associate.



# Indice analitico

## B

Bonferroni

- in Campioni Complessi [19, 24, 28](#)
- in Campioni complessi Cox Regressione [35](#)

## C

Calcolo del punteggio di Fisher

- in Complex Campioni Ordinale Regressione [30](#)

campionamento

- disegno complesso [3](#)

Campionamento casuale semplice

- in Procedura guidata di campionamento [4](#)

campionamento complesso

- pianificazione dell'analisi [8](#)

- piano di campionamento [3](#)

campionamento sequenziale

- in Procedura guidata di campionamento [4](#)

campionamento sistematico

- in Procedura guidata di campionamento [4](#)

Campionamento SPA

- in Procedura guidata di campionamento [4](#)

Campioni complessi

- opzioni [13, 14, 16, 17](#)

- Test sull'ipotesi [19, 24, 28](#)

- valori mancanti [12, 16](#)

categoria di riferimento

- in Complex Campioni General Linear Model [20](#)

- in Complex Campioni Logistic Regression [22](#)

categorie previste

- in Complex Campioni Logistic Regression [24](#)

- in Complex Campioni Ordinale Regressione [29](#)

chi-quadrato

- in Campioni Complessi [19, 24, 28](#)

- in Campioni complessi Cox Regressione [35](#)

Chi-quadrato adattato

- in Campioni Complessi [19, 24, 28](#)

- in Campioni complessi Cox Regressione [35](#)

cluster

- in Analisi Preparazione guidata [8](#)

- in Procedura guidata di campionamento [3](#)

coefficiente di rischio

- in Campioni Complessi Crosstabs [15](#)

coefficiente di variazione (COV)

- in Campioni Complessi Crosstabs [15](#)

- in Complex Campioni Descrittivi [13](#)

- in Complex Campioni Ratio [17](#)

- in Frequenze Campioni Frequenze [12](#)

conteggio non ponderato

- in Campioni Complessi Crosstabs [15](#)

- in Complex Campioni Descrittivi [13](#)

- in Complex Campioni Ratio [17](#)

- in Frequenze Campioni Frequenze [12](#)

contrasti

- in Complex Campioni General Linear Model [20](#)

contrasti di deviazione

contrasti di deviazione (*Continua*)

- in Complex Campioni General Linear Model [20](#)

contrasti di differenza

- in Complex Campioni General Linear Model [20](#)

contrasti di Helmert

- in Complex Campioni General Linear Model [20](#)

contrasti polinomiali

- in Complex Campioni General Linear Model [20](#)

contrasti ripetuti

- in Complex Campioni General Linear Model [20](#)

contrasti semplici

- in Complex Campioni General Linear Model [20](#)

convergenza delle probabilità

- in Complex Campioni Logistic Regression [25](#)

- in Complex Campioni Ordinale Regressione [30](#)

convergenza parametri

- in Complex Campioni Logistic Regression [25](#)

- in Complex Campioni Ordinale Regressione [30](#)

Correlazioni delle stime dei parametri

- in Complex Campioni General Linear Model [19](#)

- in Complex Campioni Logistic Regression [23](#)

- in Complex Campioni Ordinale Regressione [27](#)

correzione di Bonferroni sequenziale

- in Campioni Complessi [19, 24, 28](#)

- in Campioni complessi Cox Regressione [35](#)

correzione Sidak sequenziale

- in Campioni Complessi [19, 24, 28](#)

- in Campioni complessi Cox Regressione [35](#)

Covarianze delle stime dei parametri

- in Complex Campioni General Linear Model [19](#)

- in Complex Campioni Logistic Regression [23](#)

- in Complex Campioni Ordinale Regressione [27](#)

cronologia delle iterazioni

- in Complex Campioni Logistic Regression [25](#)

- in Complex Campioni Ordinale Regressione [30](#)

## D

Descrittive campioni complessi

- Statistiche [13](#)

- valori mancanti [14](#)

Differenza di rischio

- in Campioni Complessi Crosstabs [15](#)

differenza meno significativa (LSD)

- in Campioni Complessi [19, 24, 28](#)

- in Campioni complessi Cox Regressione [35](#)

dimensione campione

- in Procedura guidata di campionamento [5, 6](#)

dimensione di popolamento

- in Campioni Complessi Crosstabs [15](#)

- in Complex Campioni Descrittivi [13](#)

- in Complex Campioni Ratio [17](#)

- in Frequenze Campioni Frequenze [12](#)

- in Procedura guidata di campionamento [6](#)

dimezzamenti

- in Complex Campioni Logistic Regression [25](#)

- in Complex Campioni Ordinale Regressione [30](#)

## E

effetto della progettazione  
in Campioni complessi Cox Regressione 34  
in Campioni Complessi Crosstabs 15  
in Complex Campioni Descrittivi 13  
in Complex Campioni General Linear Model 19  
in Complex Campioni Logistic Regressione 23  
in Complex Campioni Ordinale Regressione 27  
in Complex Campioni Ratio 17  
in Frequenze Campioni Frequenze 12  
errore standard  
in Campioni Complessi Crosstabs 15  
in Complex Campioni Descrittivi 13  
in Complex Campioni General Linear Model 19  
in Complex Campioni Logistic Regressione 23  
in Complex Campioni Ordinale Regressione 27  
in Complex Campioni Ratio 17  
in Frequenze Campioni Frequenze 12

## F

file del piano 2  
Frequenze campioni complessi  
Statistiche 12

## G

gradi di libertà  
in Campioni Complessi 19, 24, 28  
in Campioni complessi Cox Regressione 35

## I

Informazioni sulla progettazione del campione  
in Campioni complessi Cox Regressione 34  
interruzione del rapporto di lavoro  
in Complex Campioni Logistic Regressione 25  
in Complex Campioni Ordinale Regressione 30  
intervalli di confidenza  
in Campioni Complessi Crosstabs 15  
in Complex Campioni Descrittivi 13  
in Complex Campioni General Linear Model 19, 21  
in Complex Campioni Logistic Regressione 23  
in Complex Campioni Ordinale Regressione 27  
in Complex Campioni Ratio 17  
in Frequenze Campioni Frequenze 12  
iterazioni  
in Complex Campioni Logistic Regressione 25  
in Complex Campioni Ordinale Regressione 30

## L

livello di confidenza  
in Complex Campioni Logistic Regressione 25  
in Complex Campioni Ordinale Regressione 30

## M

media  
in Complex Campioni Descrittivi 13  
medie marginali stimate  
in Complex Campioni General Linear Model 20

metodo di campionamento  
in Procedura guidata di campionamento 4  
Metodo di campionamento di Brewer  
in Procedura guidata di campionamento 4  
Metodo di campionamento di Murthy  
in Procedura guidata di campionamento 4  
Metodo di campionamento di Sampford  
in Procedura guidata di campionamento 4  
Metodo di stima di Breslow  
in Campioni complessi Cox Regressione 38  
Metodo di stima Efron  
in Campioni complessi Cox Regressione 38  
Metodo Newton - Raphson  
in Complex Campioni Ordinale Regressione 30  
Misura della dimensione  
in Procedura guidata di campionamento 4  
Modello lineare generalizzato campioni complessi  
funzioni aggiuntive del comando 21  
medie stimate 20  
modello 18  
opzioni 21  
Salva variabili 20  
Statistiche 19

## P

percentuali di colonna  
in Campioni Complessi Crosstabs 15  
percentuali di riga  
in Campioni Complessi Crosstabs 15  
percentuali di tabella  
in Campioni Complessi Crosstabs 15  
in Frequenze Campioni Frequenze 12  
pesi campione  
in Analisi Preparazione guidata 8  
in Procedura guidata di campionamento 6  
pesi campione di input  
in Procedura guidata di campionamento 3  
pianificazione dell'analisi 8  
piano di campionamento 3  
predittore dipendente dal tempo  
in Campioni complessi Cox Regressione 33  
probabilità cumulativa  
in Complex Campioni Ordinale Regressione 29  
Probabilità di inclusione  
in Procedura guidata di campionamento 6  
probabilità di risposta  
in Complex Campioni Ordinale Regressione 27  
probabilità prevista  
in Complex Campioni Logistic Regressione 24  
in Complex Campioni Ordinale Regressione 29  
Proporzione del campione  
in Procedura guidata di campionamento 6  
pseudo R2 statistiche  
in Complex Campioni Logistic Regressione 23  
in Complex Campioni Ordinale Regressione 27

## R

R2 statistica  
in Complex Campioni General Linear Model 19  
Radice quadrata dell'effetto del disegno  
in Campioni complessi Cox Regressione 34

Radice quadrata dell'effetto del disegno (*Continua*)  
in Campioni Complessi Crosstabs 15  
in Complex Campioni Descrittivi 13  
in Complex Campioni General Linear Model 19  
in Complex Campioni Logistic Regression 23  
in Complex Campioni Ordinale Regressione 27  
in Complex Campioni Ratio 17  
in Frequenze Campioni Frequenze 12

Rapporti campioni complessi  
Statistiche 17  
valori mancanti 17

Rapporti odds  
in Campioni Complessi Crosstabs 15  
in Complex Campioni Logistic Regression 24  
in Complex Campioni Ordinale Regressione 29

Regressione di Cox di campioni complessi  
Analisi Kaplan - Meier 31  
Definisci evento 32  
Esportazione del modello 37  
grafici 35  
modello 33  
opzioni 38  
predittore dipendente dal tempo 33  
predittori 32  
Salva variabili 36  
Sottogruppi 33  
Statistiche 34  
Test sull'ipotesi 35  
variabili di data e ora 31

Regressione logistica campioni complessi  
categoria di riferimento 22  
funzioni aggiuntive del comando 25  
modello 22  
opzioni 25  
Rapporti odds 24  
Salva variabili 24  
Statistiche 23

Regressione ordinale campioni complessi  
modello 27  
opzioni 30  
probabilità di risposta 27  
Rapporti odds 29  
Salva variabili 29  
Statistiche 27

residui  
in Campioni Complessi Crosstabs 15  
in Complex Campioni General Linear Model 20

Residui adattati  
in Campioni Complessi Crosstabs 15

residui aggregati  
in Campioni complessi Cox Regressione 36

Residui di Cox-Snell  
in Campioni complessi Cox Regressione 36

Residui di devianza  
in Campioni complessi Cox Regressione 36

residui di martingale  
in Campioni complessi Cox Regressione 36

residui di punteggio  
in Campioni complessi Cox Regressione 36

Residui parziali di Schoenfeld  
in Campioni complessi Cox Regressione 36

## S

Sidak corretta  
in Campioni Complessi 19, 24, 28  
in Campioni complessi Cox Regressione 35

somma  
in Complex Campioni Descrittivi 13

sottopopolazione  
in Campioni complessi Cox Regressione 33

statistica F  
in Campioni Complessi 19, 24, 28  
in Campioni complessi Cox Regressione 35

statistica F regolata  
in Campioni Complessi 19, 24, 28  
in Campioni complessi Cox Regressione 35

stima del campionamento  
in Analisi Preparazione guidata 9

stime del parametro  
in Campioni complessi Cox Regressione 34  
in Complex Campioni General Linear Model 19  
in Complex Campioni Logistic Regression 23  
in Complex Campioni Ordinale Regressione 27

Strati di baseline  
in Campioni complessi Cox Regressione 33

stratificazione  
in Analisi Preparazione guidata 8  
in Procedura guidata di campionamento 3

## T

tabelle di classificazione  
in Complex Campioni Logistic Regression 23  
in Complex Campioni Ordinale Regressione 27

Tabelle di contingenza campioni complessi  
Statistiche 15

Test dei pericoli proporzionali  
in Campioni complessi Cox Regressione 34

test delle linee parallele  
in Complex Campioni Ordinale Regressione 27

test t  
in Complex Campioni General Linear Model 19  
in Complex Campioni Logistic Regression 23  
in Complex Campioni Ordinale Regressione 27

## V

valori attesi  
in Campioni Complessi Crosstabs 15  
in Complex Campioni General Linear Model 20

Valori cumulativi  
in Frequenze Campioni Frequenze 12

valori mancanti  
in Campioni Complessi 12, 16  
in Complex Campioni Descrittivi 14  
in Complex Campioni General Linear Model 21  
in Complex Campioni Logistic Regression 25  
in Complex Campioni Ordinale Regressione 30  
in Complex Campioni Ratio 17





**IBM.**<sup>®</sup>