

Come costruire una distribuzione di frequenze per caratteri quantitativi continui

Consideriamo i dati contenuti nel primo foglio di lavoro (quello denominato **dati**) del file **esempio2.xls**. I dati si riferiscono al carattere *ricavo per ogni fornitura* richiesta alla Società XXXX dai clienti (indicati in colonna 1) nelle diverse settimane (colonna 2) dell'anno 2000. Le modalità rilevate del carattere in esame sono contenute in colonna 3. L'unità di misura utilizzata è l'Euro (la virgola è il separatore delle migliaia).

Vogliamo ricavare la distribuzione di frequenza del carattere ricavo per ogni fornitura, che è quantitativo continuo. Sappiamo che è necessario suddividere in classi l'intervallo nel quale sono state rilevate le modalità del carattere. Consideriamo inizialmente il caso di classi tutte della stessa ampiezza.

Ricaviamo il valore della modalità minima e massima rilevata, cioè il minimo ed il massimo ricavo per cliente:

- nelle caselle **D4** e **D5** scrivi il titolo **min=** e **max=**
- nella casella **E4** scrivi la formula **=min(C4:C102)** e premi **invio**; nella casella **E5** scrivi la formula **=max(C4:C102)** e premi **invio**

Il minimo ed il massimo dovrebbero risultare 1,764 e 58,328.

Adesso sappiamo che le modalità assunte dal carattere ricavo per cliente assume valori diciamo nell'intervallo che va da 0 a 60000 Euro. Il numero di classi in cui suddividere tale intervallo dipende dal fine dell'analisi. Supponiamo di volerlo suddividere in 12 classi equiampie, ciascuna di ampiezza 5000 Euro. Quindi possiamo creare la prima colonna della distribuzione di frequenza, ponendo in ciascuna casella il valore che separa le classi in cui vogliamo raggruppare i dati.

- nella casella **E7** scrivi il titolo **lim sup classi di ricavi**
- nella caselle **E8** scrivi 5000
- nella caselle **E9** scrivi 10000
- seleziona le caselle **E8** ed **E9** e poni il mouse nell'angolo in basso a destra della casella fino a che non compare una croce scura. Adesso premi il tasto sinistro del mouse e trascina in basso il rettangolo fino a coprire anche la casella **E19**.

Sarà comparso l'elenco 5,000, 10,000, ..., 60,000. Per far assumere ai numeri il formato richiesto seleziona le caselle d'interesse, poi, nel menu principale, **formato, celle, numero** ed, in questo caso, **valuta** e poni **simbolo** pari a **nessuno** agendo sulla freccia rivolta verso il basso (il simbolo per l'Euro non è ancora previsto).

Adesso bisogna creare in corrispondenza di ciascuna classe di modalità la corrispondente frequenza.

- nella casella **F7** scrivi il titolo **Frequenze assolute**
- seleziona col mouse le caselle **F8:F19**

- scrivi la formula **=frequenza(C4:C102;E8:E19)**, poi tieni premuti contemporaneamente i tasti \uparrow (il tasto delle maiuscole) e **Ctrl** mentre premi il tasto **invio**.

Adesso la casella F8 contiene il numero di valori tra 0 e 5000, la casella F9 il numero di valori tra 5000 e 10000, ..., la casella F19 il numero di valori tra 55000 e 60000.

Per controllare con non siano stati commessi errori, somma le frequenze

- nella casella **E20** scrivi **totale**
- seleziona la casella **F20** e successivamente il simbolo Σ nel menu in alto (è come se avessi digitato la formula **=somma(F8:F19)**) e premi **invio**.

Il risultato deve essere 99.

Per calcolare le frequenze relative

- nella casella **G7** scrivi il titolo **Frequenze relative**
- seleziona la casella **G8**
- scrivi la formula **=F8/\$F\$20** e premi **invio**. Comparirà il valore della frequenza relativa della classe di modalità (0, 5000].
- per effettuare automaticamente la stessa operazione (cioè senza riscrivere la formula) nelle caselle sottostanti, puoi tagliare ed incollare la formula. Per farlo seleziona la casella **G8** e poni il mouse nell'angolo in basso a destra della casella fino a che non compare una croce scura. Adesso premi il tasto sinistro del mouse e trascina in basso il rettangolo fino a coprire anche la casella **G19**.

Come risultato hai ottenuto le frequenze relative. Attenzione: se compaiono tutti **0** non è detto che tu abbia commesso errori, può semplicemente voler dire che il formato selezionato per i numeri non prevede decimali. Per rimediare, seleziona le caselle da **G8** a **G19**. Nel menu principale seleziona **formato**, poi **celle** ed infine **numero** nella sotto-finestra **numero**. In **posizioni decimali** scrivi **2**.

A questo punto, per controllare con non siano stati commessi errori, somma le frequenze: seleziona la casella **G20** e successivamente il simbolo Σ nel menu in alto (è come se avessi digitato la formula **=somma(G8:G19)**) e premi **invio**. Il risultato deve essere 1. Se questo non accade può voler dire che nel menu principale in **strumenti/opzioni/calcolo** non è selezionato **automatico**. Puoi selezionarlo e ripetere l'operazione oppure, se è selezionato **manuale**, basta selezionare le caselle dalla **G8** alla **G20** e premere il tasto **F9**.

Per calcolare le frequenze percentuali devi operare esattamente come per le frequenze relative, scrivendo però nella casella **H7** il titolo **%** e nella casella **H8** la formula **=(F8/\$F\$20)*100**.

Come disegnare un istogramma

Adesso vogliamo disegnare l'istogramma della distribuzione dei valori delle forniture. Teniamo presenti due punti fondamentali:

- dal momento che il carattere è continuo (e le classi sono contigue), nel grafico non ci deve essere spazio tra una colonna e l'altra
- ogni colonna corrisponderà ad una classe di valori ed excel posiziona al centro della base di ciascuna barra il valore usato per indicare la classe. Finora abbiamo usato il limite superiore di ogni classe: dovremo calcolarci il valore centrale.

Calcoliamo i valori centrali di ciascuna classe:

- seleziona l'indicativo **F** di colonna. Nel menu principale seleziona **inserisci** e **colonne**. Le frequenze assolute si sono spostate nella colonna **G**. Nella colonna **F** possiamo inserire i valori centrali.
- nella casella **F7** scrivi il titolo **valori centrali classi**
- nella casella **F8** scrivi $=E8/2$ e premi **invio** (il limite inferiore della classe è 0)
- nella casella **F9** scrivi $=(E8+E9)/2$ e premi **invio**
- copia la formula nelle caselle in basso, fino alla casella **F19**

Adesso disegniamo l'istogramma delle frequenze assolute contenute nelle caselle dalla **G8** alla **G19** usando i punti centrali (caselle dalla **F8** alla **F19**) come indicatori delle classi:

- seleziona le caselle da **G7** a **G19** (incluso il titolo) che contengono le frequenze
- nel menu principale seleziona **inserisci** e poi **grafico**
- Passo 1
 - tra i **tipi standard** seleziona **istogramma** e a destra la prima delle scelte disponibili
 - seleziona **Avanti**.
- Passo 2
 - in **intervallo dati** assicurati che sia visualizzato l'elenco delle caselle in cui sono contenute le frequenze e che in **serie in** sia selezionato **colonne**
 - seleziona **serie** per visualizzare la finestra nascosta. Per aggiungere le **etichette asse categorie (X)** seleziona la freccia rossa sulla destra. Ora seleziona le caselle dalla **F8** alla **F19** e premi **invio**

- seleziona **Avanti**.
- Passo 3
 - in **titolo del grafico** scrivi **distribuzione dei ricavi**, in **asse delle categorie (X)** scrivi **ricavi (in Euro)** ed in **asse dei valori (Y)** scrivi **n. commissioni**
 - nella pagina **assi** lascia selezionato **automatico** dopo **asse delle categorie (X)**
 - nella pagina **griglia** deseleziona tutte le tipologie di griglia
 - nella pagina **legenda** deseleziona **mostra legenda**
 - in **etichette di dati** assicurati che sia selezionato **assente**
 - in **tabella dati** assicurati che non sia selezionato **mostra tabella dati**
 - Seleziona **Avanti**.
- Passo 4
 - in **posizione grafico** deve essere selezionato **come oggetto in** ed accanto deve esserci scritto il nome del foglio su cui stai lavorando
 - seleziona **fine**

Ci sono ancora un po' di cose da aggiustare:

- le colonne non sono unite (e questo, come sappiamo è sbagliato)
- sull'asse delle ascisse, l'associazione tra le classi e i corrispondenti identificativi non è chiara.

Ora:

- seleziona un punto nell'area centrale del grafico (che non sia su di una colonna) e nel menu principale **formato** ed **area del tracciato**. Seleziona **assente** sotto **bordo** e sotto **area**, e poi **OK**
- seleziona una delle colonne del grafico e poi, nel menu principale, seleziona **formato**, **serie dei dati selezionati**, **opzioni** e, ad esempio agendo sulla freccia rivolta verso il basso, fai diventare **0** la **distanza tra le barre** e poi **OK**
- seleziona l'asse delle ascisse e poi nel menu principale **formato**, **asse selezionato**, **allineamento** e scrivi **90** in corrispondenza dei gradi.

Se vuoi puoi modificare la dimensione del grafico e la sua posizione sul foglio di lavoro.

Come costruire una distribuzione di frequenza e come disegnare un istogramma nel caso di classi di differente ampiezza

Riprendiamo l'esempio precedente e supponiamo di voler suddividere l'intervallo da 0 a 60000 Euro, su cui la variabile ricavi assume valori, nelle quattro classi $[0, 5000]$, $[5000, 10000]$, $[10000, 30000]$, $[30000, 60000]$. Possiamo creare la prima colonna della distribuzione di frequenza, ponendo in ciascuna casella il valore che separa le classi in cui vogliamo raggruppare i dati.

- nella casella **E7** scrivi il titolo **lim sup classi di ricavi**
- nella caselle **E8** scrivi 5000
- nella caselle **E9** scrivi 10000
- nella caselle **E10** scrivi 30000
- nella caselle **E11** scrivi 60000

Abbiamo scritto l'elenco 5,000, 10,000, ..., 60,000. (per scegliere il corretto formato dei numeri vedi la spiegazione fornita in precedenza).

Adesso crea la colonna delle frequenze:

- nella casella **F7** scrivi il titolo **Frequenze assolute**
- seleziona col mouse le caselle **F8:F11**
- scrivi la formula **=frequenza(C4:C102;E8:E11)**, poi tieni premuti contemporaneamente i tasti \uparrow (il tasto delle maiuscole) e **Ctrl** mentre premi il tasto **invio**.

Adesso la casella F8 contiene il numero di valori tra 0 e 5000, la casella F9 il numero di valori tra 5000 e 10000, la casella F10 il numero di valori tra 10000 e 30000 e la casella F11 il numero di valori tra 10000 e 60000.

Somma le frequenze. Sappiamo che il risultato deve essere 99.

- nella casella **E12** scrivi **totale**
- seleziona la casella **F12** e successivamente il simbolo Σ nel menu in alto e premi **invio**.

Per calcolare le frequenze relative e le frequenze percentuali si opera esattamente come già detto in precedenza:

- nella casella **G7** scrivi il titolo **Frequenze relative**
- seleziona la casella **G8**, scrivi la formula **=F8/\$F\$12** e premi **invio**. Comparirà il valore della frequenza relativa della classe di modalità $(0, 5000]$.

- seleziona di nuovo la casella **G8** e copia la formula nelle caselle sottostanti fino alla **G11**.

Come risultato hai ottenuto le frequenze relative.

Ora somma le frequenze: seleziona la casella **G12** e successivamente il simbolo Σ nel menu in alto e premi **invio**. Il risultato deve essere 1.

Per le frequenze percentuali opera esattamente come sopra scrivendo però nella casella **H7** il titolo % e nella casella **H8** la formula $=(F8/\$F\$12)*100$.

Adesso dobbiamo disegnare l'istogramma della distribuzione. Consideriamo separatamente il caso di classi di differente ampiezza perché Excel disegna gli istogrammi come se fossero diagrammi a barre, cioè non tiene conto delle differenti ampiezze delle classi. La soluzione consiste nell'utilizzare per disegnare il grafico le densità di frequenza delle varie classi (le altezze dei rettangoli corrispondenti alle varie classi di diversa ampiezza). Il risultato grafico non sarà ottimale, dal momento che rappresenta comunque un compromesso.

Calcoliamo i valori centrali di ciascuna classe:

- seleziona l'indicativo **F** di colonna. Nel menu principale seleziona **inserisci** e **colonne**. Le frequenze assolute si sono spostate nella colonna **G**. Nella colonna **F** possiamo inserire i valori centrali.
- nella casella **F7** scrivi il titolo **valori centrali classi**
- nella casella **F8** scrivi $=E8/2$ e premi **invio** (il limite inferiore della classe è 0)
- nella casella **F9** scrivi $=(E8+E9)/2$ e premi **invio**
- copia la formula nelle caselle in basso, fino alla casella **F11**

Calcoliamo le densità di frequenza (frequenza/ampiezza) delle classi. Qui lo facciamo usando le frequenze assolute, ma ricordiamo che per i confronti è necessario usare le frequenze relative.

- seleziona l'indicativo **G** di colonna. Nel menu principale seleziona **inserisci** e **colonne**. Le frequenze assolute si sono spostate nella colonna **H**. Nella colonna **G** possiamo inserire i valori delle densità di frequenza.
- nella casella **G7** scrivi il titolo **densità di frequenza**
- nella casella **G8** scrivi $=H8/E8$ e premi **invio** (il limite inferiore della classe è 0)

- nella casella **G9** scrivi $=H9/(E9-E8)$ e premi **invio**
- copia la formula nelle caselle in basso, fino alla casella **G11**

Per disegnare l'istogramma:

- seleziona le caselle da **G8** a **G11** che contengono le frequenze
- nel menu principale seleziona **inserisci** e poi **grafico**
- Passo 1
 - tra i **tipi standard** seleziona **istogramma** e a destra la prima delle scelte disponibili.
 - seleziona **Avanti**.
- Passo 2
 - in **intervallo dati** assicurati che sia visualizzato l'elenco delle caselle in cui sono contenute le frequenze e che in **serie in** sia selezionato **colonne**
 - seleziona **serie** per visualizzare la finestra nascosta. Per aggiungere le **etichette asse categorie (X)** seleziona la freccia rossa sulla destra. Ora seleziona le caselle dalla **F8** alla **F11** e premi **invio**
 - seleziona **Avanti**.
- Passo 3
 - in **titolo del grafico** scrivi **distribuzione dei ricavi**, in **asse delle categorie (X)** scrivi **ricavi (in Euro)** ed in **asse dei valori (Y)** scrivi **densità di frequenza**
 - nella pagina **assi** lascia selezionato **automatico** dopo **asse delle categorie (X)**
 - nella pagina **griglia** deseleziona tutte le tipologie di griglia
 - nella pagina **legenda** deseleziona **mostra legenda**
 - in **etichette di dati** assicurati che sia selezionato **assente**
 - in **tabella dati** assicurati che non sia selezionato **mostra tabella dati**
 - seleziona **Avanti**.
- Passo 4
 - in **posizione grafico** deve essere selezionato **come oggetto in** ed accanto deve esserci scritto il nome del foglio su cui stai lavorando
 - seleziona **fine**

Adesso per unire le colonne e per cambiare la posizione dei valori centrali, opera come spiegato in precedenza nel caso di classi equiampie.