

Prerequisiti:

- Primi elementi di probabilità e statistica.
- Nozioni di calcolo combinatorio.
- Rappresentazione di punti, rette e curve in un piano cartesiano.
- Concetto di distribuzione di probabilità

#### OBIETTIVI DI APPRENDIMENTO

Una volta completata l'unità, gli allievi devono essere in grado di:

- *riconoscere, rappresentare e utilizzare la distribuzione normale, servendosi eventualmente di uno strumento di calcolo automatico*
- *riconoscere, rappresentare e utilizzare la distribuzione di Poisson, servendosi eventualmente di uno strumento di calcolo automatico*

Questa unità interessa il Liceo Scientifico, compresa l'opzione scienze applicate, ed il Liceo delle Scienze Umane, opzione economico-sociale. Interessa inoltre gli Istituti Tecnici e gli Istituti Professionali, ma con modalità diverse che saranno specificate a tempo debito. Tecnici e Professionali ne affronteranno lo studio nel 2° biennio, i Licei nella 5<sup>a</sup> classe.

**80.1 Distribuzione normale.**

**80.2 Distribuzione di Poisson.**

*Verifiche.*

**Una breve sintesi  
per domande e risposte.**

**Distribuzione normale.  
Distribuzione di Poisson**

**Unità 80**

## 80.1 DISTRIBUZIONE NORMALE

**80.1.1** Un esempio particolarmente importante di distribuzione continua è costituito dalla **distribuzione normale** (o **distribuzione di Gauss**<sup>(1)</sup>). Si tratta di una distribuzione la cui funzione di densità di probabilità  $p(x)$  è tale che:

$$[1] \quad p(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}},$$

dove i parametri  $\mu$  e  $\sigma$  sono la media e la deviazione standard della distribuzione.

Ad onor del vero per poter affermare che si tratti di una funzione di probabilità bisogna accertarsi che siano soddisfatte le due condizioni note. Ora, che sia  $p(x) > 0$  è facile constatarlo. Che sia  $\int_{-\infty}^{+\infty} p(x) dx = 1$  è invece molto complicato dimostrarlo teoricamente, anche se possibile: lo assumiamo per vero. In realtà, mediante un idoneo software matematico, fissati i valori di  $\mu$  e  $\sigma$  (per esempio:  $\mu=0$ ,  $\sigma=1$ ) si trova che:

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{x^2}{2}} dx = 1.$$

La funzione  $X$  tale che:

$$P[X = x] = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

è detta **variabile aleatoria normale** di parametri  $\mu$  e  $\sigma$ . È indicata con la scrittura  **$N(\mu, \sigma^2)$** .

Il suo grafico, in un piano riferito ad un sistema di assi cartesiani ( $Oxy$ ), si chiama **curva normale della probabilità** (o **gaussiana** o anche **curva a campana** per la sua forma caratteristica).

In figura 1, dove in ordinate sono rappresentate le probabilità, è disegnata la gaussiana di parametri  $\mu=12$  e  $\sigma=2,5$ .

Si può stabilire che la curva normale della probabilità (ogni curva, non solo quella dell'esempio) è simmetrica rispetto alla retta  $x=\mu$ , la quale divide pertanto l'area sotto il grafico della curva in due parti uguali, ciascuna di area 0,5.

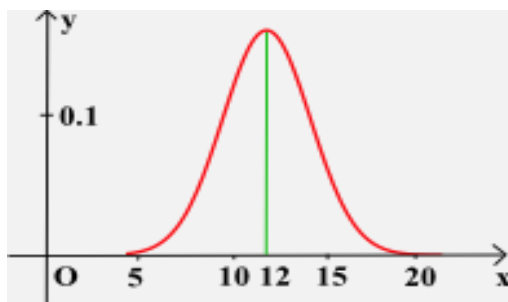


FIG. 1

Al fine di valutare “visivamente” il ruolo dei parametri  $\mu$  e  $\sigma$  visualizza sullo schermo di un computer le gaussiane corrispondenti ai seguenti valori:

- a)  $\mu=0$ ,  $\sigma^2=1$ ; b)  $\mu=0$ ,  $\sigma^2=2$ ; c)  $\mu=2$ ,  $\sigma^2=1$ ; d)  $\mu=2$ ,  $\sigma^2=2$ .

<sup>1</sup> **Gauss**, Carl Friedrich; matematico tedesco, 1777–1855.

**80.1.2** La distribuzione normale ha la caratteristica di approssimare, sotto particolari condizioni, la distribuzione binomiale. Nella formula [1], con riferimento alla distribuzione binomiale  $B(n,p)$ ,  $x$  prende il posto di  $k$  e  $p(x)$  quello di  $p_k$ ; inoltre  $\mu=np$  e  $\sigma=\sqrt{npq}$ , essendo  $q=1-p$ , mentre  $\pi$  ed  $e$  sono due celebri numeri trascendenti ( $\pi\approx 3,14159$ ;  $e\approx 2,71828$ ). Per la precisione, circa la possibilità di utilizzare la distribuzione normale in sostituzione della distribuzione binomiale, valgono le seguenti condizioni:

- l'approssimazione è ottima ogni volta che  $n\geq 10$  e  $|p-q|<0,04$ ;
- l'approssimazione è accettabile quando  $n\geq 30$ ,  $np>5$  ed  $nq>5$ .

S'intende che  $n$  è il numero delle prove effettuate,  $p$  la probabilità che l'evento si verifichi in una prova e  $q=1-p$  la probabilità contraria.

A dire il vero, oggi giorno questa caratteristica della distribuzione normale, come approssimazione della distribuzione binomiale, è irrilevante dal momento che esistono appositi software matematici che consentono di calcolare direttamente i valori della variabile aleatoria binomiale. Cosa che d'altro canto hai avuto modo di constatare nella precedente unità. Ragion per cui i pochi esempi che andiamo a fornire hanno lo scopo di dare solo e semplicemente un'idea di quanto andiamo dicendo e non già di calcolare approssimazioni più o meno buone della distribuzione binomiale. Sono insomma una curiosità storica, se si vuole, giacché danno l'idea di come si operasse quando il calcolo automatico non c'era ancora.

Vediamo allora qualche situazione in cui la distribuzione normale approssima la distribuzione binomiale.

- **ESERCIZIO 1.** Calcolare la probabilità che, lanciando 20 volte una moneta "testa-croce", esca 8 volte "testa", nell'ipotesi che i due eventi "esce testa" ed "esce croce" siano equiprobabili.

**RISOLUZIONE.** Osserviamo anzitutto che in ogni lancio la probabilità dell'evento "esce testa" è  $p=0,5$  e così pure  $q=0,5$ . Volendo applicare la [1] per risolvere la questione, dobbiamo preventivamente calcolare  $\mu$  e  $\sigma$ . Ora si ha:  $\mu=np=20\times 0,5=10$ ,  $\sigma=\sqrt{npq}=\sqrt{20\times 0,5\times 0,5}\approx 2,2361$ . Per cui, con uno strumento di calcolo automatico:

$$P[X=8] = \frac{1}{2,2361 \sqrt{2\pi}} e^{-\frac{(8-10)^2}{2\times 2,2361^2}} \approx 0,1196.$$

Se avessimo applicato la formula della distribuzione binomiale, avremmo ottenuto, sempre con uno strumento di calcolo automatico:

$$P[B=8] = \binom{20}{8} \times 0,5^8 \times 0,5^{12} = \frac{20!}{8!12!} \times 0,5^{20} \approx 0,1201.$$

Affermando quindi che la probabilità cercata, che di fatto è 0,1201, sia invece 0,1196, si commette un errore relativo:

$$\varepsilon_r = \frac{0,1201 - 0,1196}{0,1201} \approx 0,5\%$$

certamente trascurabile.

- **ESERCIZIO 2.** Calcolare la probabilità che, estraendo a caso per 10 volte una carta da un mazzo da bridge, esca 3 volte una carta di cuori.

**RISOLUZIONE.** Osserviamo anzitutto che la probabilità dell'evento "esce una carta di cuori" è  $p=0,25$  e che quella dell'evento contrario è ovviamente  $q=0,75$ .

Il valore esatto della probabilità, ottenuto ricorrendo alla distribuzione binomiale e calcolato con l'aiuto di

uno strumento di calcolo automatico, è il seguente:

$$P[B=3] = \binom{10}{3} \times 0,25^2 \times 0,75^7 \approx 0,250.$$

Invece, applicando la formula approssimata [1], dopo aver calcolato che  $\mu=2,5$  e  $\sigma \approx 1,3693$ , si ottiene (sempre con l'aiuto di un calcolatore):

$$P[X=3] = \frac{1}{1,3693 \sqrt{2\pi}} e^{-\frac{(3-2,5)^2}{2 \times 1,3693^2}} \approx 0,272.$$

Affermare quindi che la probabilità cercata sia 0,272 (invece di 0,250) significa commettere un errore relativo:

$$\varepsilon_r = \frac{0,272-0,250}{0,250} \approx 8,8\%$$

Non è un errore enorme, ma non è neppure trascurabile.

- Nel primo dei due esempi esaminati, essendo  $n=20$  e  $|p-q|=0$ , è soddisfatta una condizione affinché si possa ricorrere alla distribuzione normale come approssimazione della distribuzione binomiale: si è ottenuta, in effetti, un'approssimazione eccellente.
- Nel secondo dei due esempi, essendo  $n=10$ ,  $|p-q|=0,5$ ,  $np=10 \times 0,25=2,5$ , non è soddisfatta alcuna delle condizioni per cui si possa ricorrere alla distribuzione normale come approssimazione della distribuzione binomiale e, in realtà, non si è ottenuta un'approssimazione soddisfacente.

- **ESERCIZIO 3.** Da un'urna contenente i 90 numeri della tombola si fanno 100 estrazioni casuali, rimettendo ogni volta il numero estratto nell'urna. Calcolare la probabilità che per 15 volte esca un multiplo di 6. **RISOLUZIONE.** Osserviamo anzitutto che la probabilità dell'evento "esce un multiplo di 6" è  $p = \frac{15}{90} = \frac{1}{6}$  e che quella dell'evento contrario è ovviamente  $q = 1 - p = \frac{5}{6}$ .

Il valore esatto della probabilità, calcolato ricorrendo alla distribuzione binomiale, è il seguente:

$$P[B=15] = \binom{100}{15} \times \left(\frac{1}{6}\right)^{15} \times \left(\frac{5}{6}\right)^{85} \approx 0,100.$$

Invece, applicando la formula approssimata [1], dopo aver calcolato che  $\mu=100 \times \frac{1}{6} \approx 16,6667$  e  $\sigma = \sqrt{100 \times \frac{1}{6} \times \frac{5}{6}} \approx 3,7268$ , si ottiene:

$$P[X=15] = \frac{1}{3,7268 \sqrt{2\pi}} e^{-\frac{(15-16,6667)^2}{2 \times 3,7268^2}} \approx 0,096.$$

In questa circostanza, è vero che si ha  $|q-p| = \left|\frac{1}{6} - \frac{5}{6}\right| \approx 0,7$  ma risulta  $n > 30$  ed inoltre  $np > 5$  ed  $nq > 5$ . Quindi l'approssimazione, pur non essendo eccellente, dovrebbe comunque essere accettabile. In effetti, prendendo come buona la seconda valutazione della probabilità si commette il seguente errore relativo:

$$\varepsilon_r = \frac{0,100-0,096}{0,100} = 4\%.$$

**80.1.3** Occupiamoci adesso del caso in cui si debba calcolare la probabilità che una variabile normale assuma un valore compreso in un determinato intervallo, limitato o illimitato che sia. Oltre che con un calcolo diretto (sempre con l'uso di un software) si può fare ricorso ad un'apposita tabella che fornisce alcune probabilità e, sulla scorta di queste, calcolare quella che di volta in volta interessa. Ma procediamo con ordine, incominciando con la costruzione di siffatta tabella. Per tale scopo è preferibile ser-

virsi non della formula [1], ma di una sua trasformata.

A tal riguardo operiamo la trasformazione di equazioni:

$$X = \frac{x - \mu}{\sigma}, \quad Y = \sigma y$$

dove abbiamo posto  $y$  al posto di  $p(x)$ . La [5] diventa allora:

$$Y = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

la quale, scritta nelle solite variabili  $x, y$  – al posto rispettivamente di  $X, Y$  – diventa:

$$y = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}.$$

È chiamata **curva normale standardizzata** (o **standard**) e la distribuzione di probabilità relativa è detta **distribuzione normale standardizzata** (o **standard**). Ha media 0 e deviazione standard 1. Per disegnarne il grafico è preferibile scegliere un sistema di riferimento non monometrico (Fig. 2).

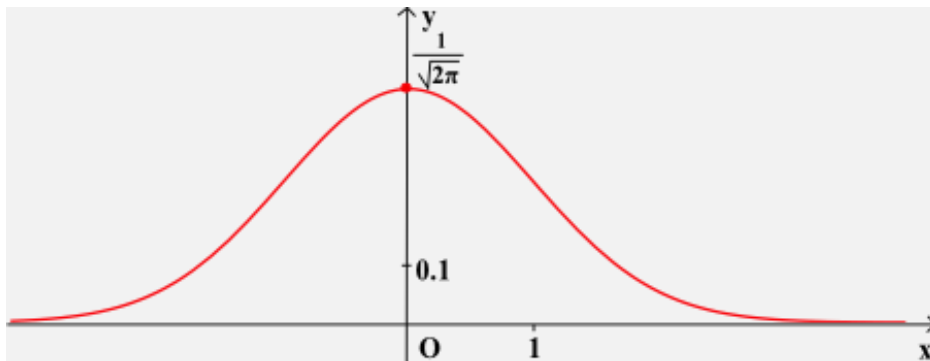


FIG. 2

Facciamo notare che la nota simmetria del grafico si ha questa volta rispetto all'asse  $y$ .

La probabilità  $P[0 \leq N \leq k]$  che la variabile normale  $N(0,1)$  assuma un valore compreso fra 0 e  $k$  è uguale all'area della regione finita  $A$  di piano (Fig. 3) delimitata dalla gaussiana, dall'asse  $x$ , dall'asse  $y$  e dalla retta di equazione  $x=k$ , e perciò è dipendente soltanto da  $k$ . Quest'area, come sai, può essere calcolata con un metodo di integrazione numerica (per esempio: metodo dei trapezi).

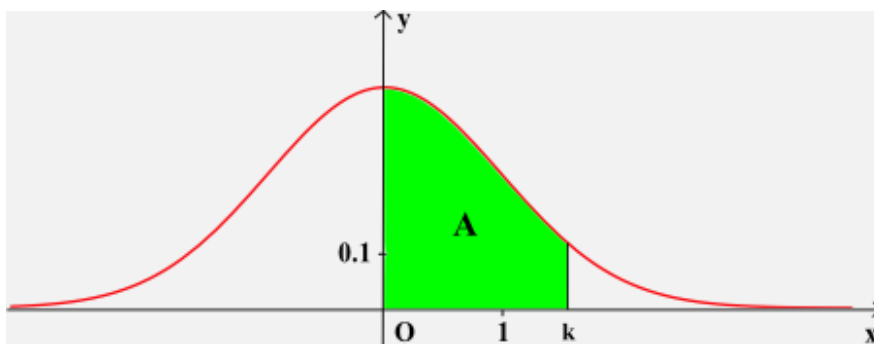


FIG. 3

Ebbene, si può costruire una volta per tutte una tabella (Tab. 1) che fornisce la probabilità che la variabile normale standardizzata  $N$  acquisti un valore qualsiasi compreso fra 0 e  $k$ .

**Tab. 1 - Tabella della distribuzione normale standardizzata**

<b>k</b>	<b>P[0≤N≤k]</b>	<b>k</b>	<b>P[0≤N≤k]</b>	<b>k</b>	<b>P[0≤N≤k]</b>	<b>k</b>	<b>P[0≤N≤k]</b>
<b>0,00</b>	000000	<b>0,1</b>	0,03983	<b>1,1</b>	0,36433	<b>2,1</b>	0,48214
<b>0,01</b>	0,00399	<b>0,2</b>	0,07926	<b>1,2</b>	0,38493	<b>2,2</b>	0,48610
<b>0,02</b>	0,00798	<b>0,3</b>	0,11791	<b>1,3</b>	0,40320	<b>2,3</b>	0,48928
<b>0,03</b>	0,01197	<b>0,4</b>	0,15542	<b>1,4</b>	0,41924	<b>2,4</b>	0,49180
<b>0,04</b>	0,01595	<b>0,5</b>	0,19146	<b>1,5</b>	0,43319	<b>2,5</b>	0,49379
<b>0,05</b>	0,01994	<b>0,6</b>	0,22575	<b>1,6</b>	0,44520	<b>2,6</b>	0,49534
<b>0,06</b>	0,02392	<b>0,7</b>	0,25804	<b>1,7</b>	0,45543	<b>2,7</b>	0,49653
<b>0,07</b>	0,02790	<b>0,8</b>	0,28814	<b>1,8</b>	0,46407	<b>2,8</b>	0,49744
<b>0,08</b>	0,03188	<b>0,9</b>	0,31594	<b>1,9</b>	0,47128	<b>2,9</b>	0,49813
<b>0,09</b>	0,03586	<b>1,0</b>	0,34134	<b>2,0</b>	0,47725	<b>3,0</b>	0,49864
						<b>3,5</b>	0,49977
						<b>4,0</b>	0,49997

**Nota Bene.** Se a ciascuno dei valori  $P[0 \leq N \leq k]$  della precedente tabella si somma 0,5 si ottengono i valori della funzione di ripartizione della variabile normale standardizzata:

$$F(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-t^2/2} dt$$

ovviamente per  $x \geq 0$ . Se invece  $x < 0$  i valori di  $F(x)$  si ottengono per simmetria rispetto al valore 0,5. La figura 4, che rappresenta tale funzione di ripartizione chiarisce quanto stiamo dicendo.

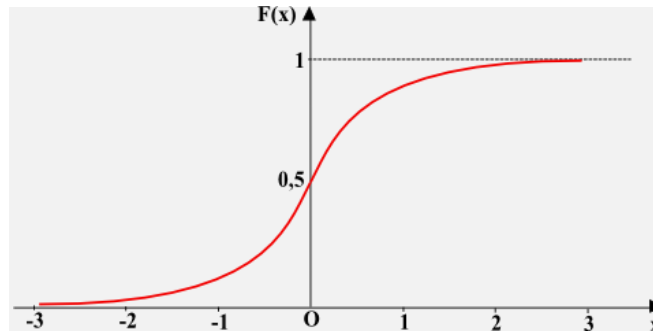


FIG. 4

Allora, tenendo presente la tabella 1 ed il fatto che l'asse  $y$ , che adesso è l'asse di simmetria e perciò divide la suddetta regione in due parti congruenti (ciascuna di area 0,5) si ha per esempio che:

- $P[0,2 \leq N \leq 1,4] = P[0 \leq N \leq 1,4] - P[0 \leq N \leq 0,2] = 0,41924 - 0,07926 = 0,3399$ ;
- $P[-1,2 \leq N \leq 1,2] = 2 P[0 \leq N \leq 1,2] = 2 \times 0,38493 = 0,7698$ ;
- $P[-0,8 \leq N \leq 0,6] = P[-0,8 \leq N \leq 0] + P[0 \leq N \leq 0,6] = P[0 \leq N \leq 0,8] + P[0 \leq N \leq 0,6] = 0,28814 + 0,22575 = 0,5138$ ;
- $P[N \leq 1,4] = 0,5 + P[0 \leq N \leq 1,4] = 0,5 + 0,41924 = 0,91924$ ;
- $P[N \geq 0,2] = 0,5 - P[0 \leq N \leq 0,2] = 0,5 - 0,07926 = 0,42074$ .

Proponiamo a te, per esercizio, di calcolare le seguenti probabilità, utilizzando la tabella della distribuzione normale standardizzata:

$$P[0,3 \leq N \leq 1,7], \quad P[-0,5 \leq N \leq 0,9], \quad P[N \leq 2,4], \quad P[N \leq -1,5], \quad P[N \geq 1,2], \quad P[N \geq -0,2].$$

**80.1.4** Se  $k$  ha altre cifre decimali oltre la prima (quando  $k > 0,1$ ) od oltre la seconda (quando  $0 < k < 0,1$ ), il calcolo di  $P[0 \leq N \leq k]$ , sempre che ci si voglia avvalere della tabella 1, richiede il ricorso alla cosiddetta *interpolazione lineare*.

Un esempio basta per capire come si fa. Si voglia allora calcolare  $P[0 \leq N \leq 1,44]$ .

Tra le ascisse 1,4 ed 1,5 l'andamento delle aree sotto la curva normale standardizzata, calcolate entro i limiti 0 e  $k$ , si può supporre lineare; rappresentabile perciò col segmento di retta AB (Fig. 5), dove A ha come coordinate  $x_A=1,4$  ed  $y_A=P[0 \leq N \leq 1,4]$  e B ha come coordinate  $x_B=1,5$  ed  $y_B=P[0 \leq N \leq 1,5]$ .

Il valore cercato  $P[0 \leq N \leq 1,44]$  è allora l'ordinata  $y_P$  del punto P della retta AB di ascissa  $x_P=1,44$ .

Pertanto, tenendo presente l'equazione della retta passante per due punti, risulta:

$$\frac{y_P - y_A}{y_B - y_A} = \frac{x_P - x_A}{x_B - x_A};$$

da qui, dopo aver sostituito i rispettivi valori numerici e dopo qualche calcolo, segue:  $y_P \approx 0,4248$ .

In definitiva:  $P[0 \leq N \leq 1,44] \approx 42,48\%$ .

In realtà non serve ricorrere alla tabella 1 poiché si può impiegare direttamente lo stesso software matematico utilizzato per compilare la tabella medesima. Si trova in tal caso:  $P[0 \leq N \leq 1,44] \approx 42,50\%$ .

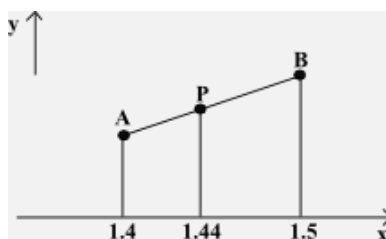


FIG. 5

Ti proponiamo, per esercizio, di calcolare le seguenti probabilità:

$$P[0,32 \leq N \leq 1,7], \quad P[-0,54 \leq N \leq 0,69], \quad P[N \leq 1,472], \quad P[N \leq -1,25], \quad P[N \geq 1,12], \quad P[N \geq -0,222].$$

**80.1.5** La tabella 1 permette pure di giungere al calcolo della probabilità:

$$P[\alpha \leq X \leq \beta]$$

riguardante una distribuzione normale di media  $\mu$  e deviazione standard  $\sigma$  qualsiasi.

La teoria permette, infatti, di dimostrare che si ha:

$$[2] \quad P[\alpha \leq X \leq \beta] = P\left[\frac{\alpha - \mu}{\sigma} \leq N \leq \frac{\beta - \mu}{\sigma}\right].$$

Per esempio, se  $\mu=12$  e  $\sigma=2,5$  ed inoltre  $\alpha=10$  e  $\beta=13$ , si ha:

$$\begin{aligned} P[10 \leq X \leq 13] &= P\left[\frac{10-12}{2,5} \leq N \leq \frac{13-12}{2,5}\right] = P[-0,8 \leq N \leq 0,4] = \\ &= P[-0,8 \leq N \leq 0] + P[0 \leq N \leq 0,4] = P[0 \leq N \leq 0,8] + P[0 \leq N \leq 0,4] = 0,28814 + 0,15542 = 0,4435. \end{aligned}$$

• **ESERCIZIO.** L'altezza media di un gruppo di persone è 175 cm con una deviazione standard di 14 cm. Ammesso che la distribuzione delle altezze delle persone sia quella normale, trovare la probabilità che una persona, scelta a caso nel gruppo considerato, abbia un'altezza:

- 1) compresa fra 170 cm e 180 cm;
- 2) minore di 170 cm.

**RISOLUZIONE.** Si ha a che fare con una variabile aleatoria normale  $X$  di parametri  $\mu=175$  e  $\sigma=14$ .

Nel 1° caso la probabilità cercata è allora:

$$P[170 \leq X \leq 180] = P\left[\frac{170-175}{14} \leq N \leq \frac{180-175}{14}\right] = P[-0,35714 \leq N \leq 0,35714] = 2 P[0 \leq N \leq 0,35714].$$

Mediante il solito software matematico, se non si vuole ricorrere alla tabella 1 e ad un processo d'interpolazione, si trova:  $P[0 \leq N \leq 0,35714] \approx 0,1395$ . Pertanto:

$$P[170 \leq X \leq 180] = 2 \times 0,1395 \approx 27,90\%.$$

Per rispondere alla seconda domanda bisogna ricordare che la curva normale della probabilità è simmetrica rispetto alla retta  $x = \mu = 175$  e che quella retta divide l'area sotto la curva in due parti, aventi ciascuna area 0,5. Pertanto si ha:  $P[X < 170] = 0,5 - P[170 \leq X \leq 175]$ . D'altra parte, sempre per ragioni di simmetria:  $P[170 \leq X \leq 175] = P[175 \leq X \leq 180]$ , perciò:

$$P[170 \leq X \leq 175] = \frac{1}{2} P[170 \leq X \leq 180] \approx \frac{1}{2} 0,2790 = 0,1395.$$

Di conseguenza:

$$P[X < 170] \approx 0,5 - 0,1395 = 36,05\%.$$

**ESERCIZIO.** Considerata la variabile aleatoria normale  $X$ , di media  $\mu$  e deviazione standard  $\sigma$ , calcola la probabilità che  $X$  assuma un valore compreso tra:

a)  $\mu - \sigma$  e  $\mu + \sigma$ , b)  $\mu - 2\sigma$  e  $\mu + 2\sigma$ , c)  $\mu - 3\sigma$  e  $\mu + 3\sigma$ .

**80.1.6** La distribuzione normale svolge un ruolo molto importante in campo statistico, come avremo modo di constatare nella prossima unità. Ma già adesso vogliamo far vedere come la tabella 1 sia utile in situazioni interessanti come quella che si presenta nella risoluzione del seguente problema.

• **PROBLEMA.** La prova di un concorso consiste in una batteria di quesiti a scelta multipla, con 4 alternative di cui una sola corretta. In seguito alla revisione degli elaborati, i punteggi ottenuti risultano distribuiti normalmente con una media uguale a 108 ed una deviazione standard pari a 9,6. Qual è il punteggio minimo per risultare vincitore di un posto, se i posti messi a concorso costituiscono il 7% del numero dei concorrenti?

**RISOLUZIONE.** La distribuzione di probabilità riguardante il nostro problema è espressa dalla formula [1], dove adesso  $\mu = 108$  e  $\sigma = 9,6$  (Fig. 6).

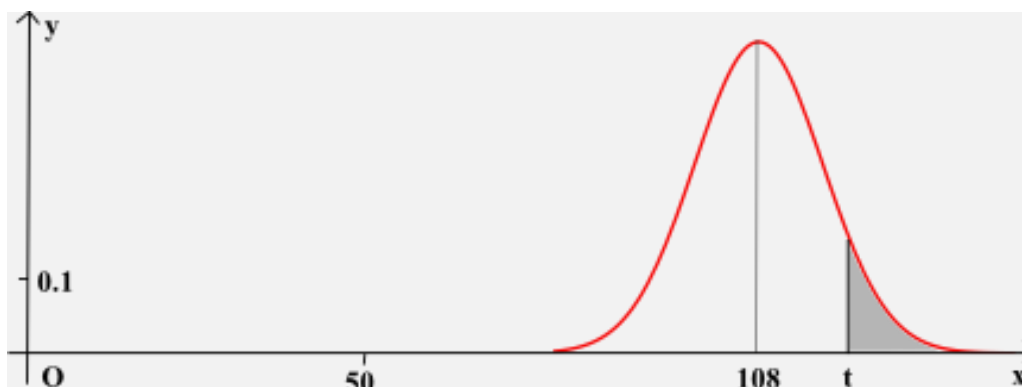


FIG. 6

Il problema consiste nel trovare per quali punteggi  $t$ , con  $t > \mu$ , l'area  $A(t)$  tratteggiata in figura non supera il 7% di tutta l'area sotto la curva, che notoriamente è 1; cioè per quale  $t$  risulta  $A(t) \leq 0,07$ .

Ora, posto  $X = N(\mu, \sigma^2)$ , si ha evidentemente:  $A(t) = 0,5 - P[\mu \leq X \leq t]$ . D'altronde:

$$P[\mu \leq X \leq t] = P\left[\frac{\mu - \mu}{\sigma} \leq N \leq \frac{t - \mu}{\sigma}\right] = P\left[0 \leq N \leq \frac{t - 108}{9,6}\right].$$

Quindi deve risultare:

$$0,5 - P\left[0 \leq N \leq \frac{t-108}{9,6}\right] \leq 0,07 \quad \text{ossia: } P\left[0 \leq N \leq \frac{t-108}{9,6}\right] \geq 0,43.$$

Cercando il numero 0,43 nella tabella 1, colonna  $P[0 \leq N \leq k]$ , si nota che esso non c'è; ma c'è 0,41924 che costituisce il più grande valore che non supera 0,43. Poiché a quel valore di  $P[0 \leq N \leq k]$  corrisponde il valore 1,4 di  $k$ , il valore  $t$ , con buona approssimazione, deve essere tale che:

$$\frac{t-108}{9,6} > 1,4 \quad \text{ossia: } t-108 > 9,6 \times 1,4 \quad \text{e perciò: } t > 121,4.$$

È questo il punteggio minimo richiesto.

## 80.2 DISTRIBUZIONE DI POISSON <sup>(2)</sup>

**80.2.1** La **distribuzione di Poisson** <sup>(3)</sup> (detta anche **poissoniana**) è un esempio di *distribuzione discreta e infinita*. Essa è la distribuzione generata dalla formula seguente, detta a sua volta **formula di Poisson**:

$$[3] \quad p_k = \frac{\mu^k}{k!} e^{-\mu}$$

e fornisce la probabilità  $p_k$  che in infinite prove, ripetute nelle stesse condizioni, l'evento si presenti  $k$  volte. Questa formula, quando  $k$  varia da 0 ad  $\infty$ , genera il seguente sistema di numeri:

$$\frac{\mu^0}{0!} e^{-\mu}, \quad \frac{\mu^1}{1!} e^{-\mu}, \quad \frac{\mu^2}{2!} e^{-\mu}, \quad \dots, \quad \frac{\mu^k}{k!} e^{-\mu}, \quad \dots;$$

i quali costituiscono effettivamente una distribuzione di probabilità poiché la loro somma è uguale ad 1. Per dimostrarlo basta tener presente lo sviluppo di  $e^x$  in serie di potenze di  $x$ . Si ha, precisamente:

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!},$$

ragione per cui, dopo aver sostituito  $x$  con  $\mu$ :

$$\sum_{k=0}^{\infty} \frac{\mu^k}{k!} e^{-\mu} = e^{-\mu} \sum_{k=0}^{\infty} \frac{\mu^k}{k!} = e^{-\mu} \cdot e^{\mu} = 1.$$

Una variabile capace di assumere i valori 0, 1, 2, ...,  $k$ , ... con le rispettive probabilità date dalla [3], si chiama **variabile aleatoria di Poisson** di parametro  $\mu$ . È indicata con la scrittura:

$$\mathbf{P}(\mu) \text{ o anche, se c'è luogo ad equivoci, con } \mathbf{Poisson}(\mu).$$

Onde avere un'idea più concreta di tale distribuzione, abbiamo rappresentato (Fig. 7) la funzione:

$$y = \frac{\mu^x}{x!} e^{-\mu}$$

con  $x \geq 0$ , per  $\mu=8$ , evidenziando i punti della curva aventi ascisse intere da 1 a 15.

<sup>2</sup> Questo paragrafo, comprese le verifiche, riguarda i Licei e, per quanto concerne i Tecnici, esclusivamente il settore Tecnologico, indirizzo Elettronica ed Elettrotecnica. È opzionale sia per gli altri indirizzi del settore Tecnologico sia per il settore Economico sia infine per gli Istituti Professionali.

<sup>3</sup> **Poisson**, Simeon Denis; matematico francese, 1781-1840.

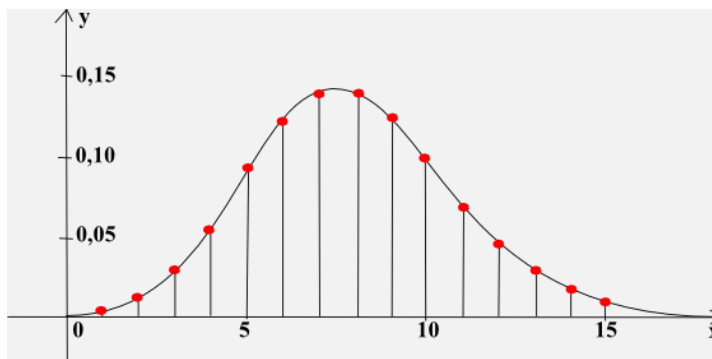


FIG. 7

A te proponiamo invece di visualizzare sullo schermo di un computer le distribuzioni di Poisson corrispondenti ai seguenti valori del parametro  $\mu$ :

a)  $\mu=1$ , b)  $\mu=3$ , c)  $\mu=6$ , d)  $\mu=10$ .

La teoria mostra che la media  $M(P)$  e la deviazione standard  $\sigma(P)$  della variabile di Poisson sono date dalle formule seguenti:

$$M(P) = \mu, \quad \sigma(P) = \sqrt{\mu}.$$

Cosicché il parametro  $\mu$  della variabile di Poisson costituisce nel medesimo tempo la media e la varianza di tale variabile.

**80.2.2** Anche la distribuzione di Poisson può essere utilizzata come approssimazione della distribuzione binomiale. Ciò si ha precisamente quando il numero  $n$  delle prove è molto grande (in pratica quando  $n \geq 30$ ) e la probabilità  $p$  di ogni evento elementare è molto piccola (in pratica quando  $np < 5$ ). Per questo motivo la distribuzione di Poisson è nota anche come *legge degli eventi rari*. Nell'approssimazione si assume  $\mu = np$ . Ma, come nel caso della distribuzione normale e per gli stessi motivi addotti, quest'aspetto della distribuzione di Poisson è oggi irrilevante.

Al giorno d'oggi, in effetti, l'aspetto più interessante è fornito da una singolare interpretazione della formula di Poisson. Senza soffermarci sulla spiegazione del fatto, diciamo che:

**La formula [3] esprime la probabilità che un certo evento si presenti esattamente  $k$  volte in un dato "dominio" sapendo che si verifica mediamente  $\mu$  volte in quel dominio.**

Il dominio di cui si parla può essere un insieme di oggetti (alle condizioni descritte sopra) ma anche un intervallo di tempo, una lunghezza, una superficie, un volume.

A maggior chiarimento vediamo alcuni esercizi, nella risoluzione dei quali ci serviremo di un software matematico.

- Nei primi due esempi il dominio è un intervallo di tempo.

ESERCIZIO 1. Si sa che in una certa città avvengono mediamente 7 incidenti automobilistici al mese. Calcolare la probabilità che capitino 4 incidenti nel mese di gennaio.

RISOLUZIONE. Basta applicare la formula di Poisson, dove  $\mu=7$ ,  $k=4$ . Per cui:

$$P[k=4] = \frac{7^4}{4!} \approx 9,12\%.$$

ESERCIZIO 2. Al centralino di un'azienda di medie proporzioni arrivano mediamente 2 telefonate al minuto. Qual è la probabilità che in mezz'ora arrivino a quel centralino:

- a) esattamente 60 telefonate? b) non più di 60 telefonate?

RISOLUZIONE. Il modello di riferimento è quello di Poisson, vale a dire la formula [3], dove  $\mu=2 \times 30=60$ . Precisamente si ha:

$$\text{a) } P[k=60] = \frac{60^{60}}{60!} e^{-60} \approx 0,0514; \quad \text{b) } P[0 \leq k \leq 60] = \sum_{k=0}^{60} \frac{60^k}{k!} = 0,5343.$$

- Nel terzo esempio il dominio è una superficie.

ESERCIZIO 3. Contro una data superficie sono lanciati  $n$  proiettili. Supposto che la superficie sia ripartita in  $S$  unità, calcolare la probabilità che in una unità prescelta cadano  $k$  proiettili.

RISOLUZIONE. Basta applicare la formula di Poisson, dove:

$$\mu = \frac{n}{S}.$$

- Nel quarto esempio il dominio è un insieme di oggetti.

ESERCIZIO 4. I controlli eseguiti hanno rivelato che mediamente lo 0,4% delle bottigliette di coca-cola indicano, sotto il tappo, la vincita di un premio. Il gestore di un bar ha acquistato una partita di 500 bottigliette. Qual è la probabilità che esattamente lo 0,4% di esse contenga il premio?

RISOLUZIONE. Il modello di riferimento esatto è la distribuzione binomiale ed il relativo calcolo è lasciato a te. Ti forniamo solo il risultato, che è all'incirca 27,12%.

Siccome tuttavia ci troviamo in presenza di una situazione in cui sono soddisfatte le condizioni per ricorrere alla distribuzione di Poisson come approssimazione della distribuzione binomiale ( $n=500 > 30$ ,  $np=500 \times 0,004=2 < 5$ ), possiamo ricorrere proprio alla formula [3], nella quale adesso  $\mu=2$ , che è per l'appunto lo 0,4% di 500. Si trova:  $P[k=2] \approx 27,07\%$ .

Quale errore si commette assumendo come valore della probabilità cercata quello calcolato con la formula di Poisson al posto della distribuzione binomiale?

- Anche in questo quinto esempio il dominio è un insieme di oggetti, che questa volta sono intervalli di tempo.

ESERCIZIO 5. Una sostanza radioattiva emette in modo irregolare particelle che possono essere conteggiate da un contatore Geiger <sup>(4)</sup>. Il numero di tali particelle, conteggiato ad intervalli regolari di 15 secondi, è riportato nella tabella sottostante.

Numero particelle	0	1	2	3	4	5	6	7	8	9	10	11
Numero intervalli	1	8	20	32	36	34	25	16	14	8	2	1

Una volta trovata la distribuzione delle frequenze relative del numero degli intervalli in cui sono emesse  $k$  particelle, una rappresentazione grafica di questa distribuzione mostra intuitivamente che essa può essere assimilata con buona approssimazione ad una distribuzione di Poisson. Trovare la distribuzione delle frequenze assolute derivante da questa distribuzione di Poisson.

RISOLUZIONE. Lasciamo a chi legge la risoluzione della prima parte dell'esercizio e occupiamoci della seconda parte. Si tratta di trovare il numero  $N_k$  di intervalli durante i quali sono emesse  $k$  particelle, ipotizzando che la probabilità che sia

$$p_k = \frac{\mu^k}{k!} e^{-\mu}$$

la probabilità di avere un intervallo durante il quale siano emesse  $k$  particelle, essendo  $\mu$  la media della distribuzione. Allora, se  $N$  è il numero totale degli intervalli in cui sono state conteggiate le particelle emesse, risulta evidentemente:

<sup>4</sup> Geiger, Hans Wilhelm, fisico tedesco, 1882-1945

$$N_k = N \cdot p_k.$$

Basta pertanto conoscere  $N$  e  $p_k$ .

Riguardo ad  $N$ , si ha semplicemente:

$$N = 1 + 8 + 20 + 32 + 36 + 34 + 25 + 16 + 14 + 8 + 2 + 1 = 197.$$

Riguardo a  $p_k$ , bisogna prima calcolare  $\mu$ . Cosa abbastanza semplice. Si ha infatti:

$$\mu = \frac{1 \cdot 0 + 8 \cdot 1 + 20 \cdot 2 + 32 \cdot 3 + 36 \cdot 4 + 34 \cdot 5 + 25 \cdot 6 + 16 \cdot 7 + 14 \cdot 8 + 8 \cdot 9 + 2 \cdot 10 + 1 \cdot 11}{1 + 8 + 20 + 32 + 36 + 34 + 25 + 16 + 14 + 8 + 2 + 1} \approx 4,746.$$

Utilizzando infine la formula seguente:

$$N_k = 197 \cdot \frac{4,746^k}{k!} e^{-4,746},$$

magari con l'ausilio di uno strumento di calcolo automatico, si ottiene la distribuzione delle frequenze assolute degli intervalli in cui sono emesse  $k$  particelle, dedotta dalla distribuzione di Poisson. Questa distribuzione è rappresentata dalla tabella sottostante, la quale – come si può constatare – non è molto diversa da quella reale, ottenuta sperimentalmente.

k=Numero particelle	0	1	2	3	4	5	6	7	8	9	10	11	≥12
N <sub>k</sub> =Numero intervalli	2	8	19	30	36	34	27	18	11	6	3	1	2

## VERIFICHE

(Si suggerisce l'uso di un idoneo software matematico)

### Distribuzione normale (nn. 1-9).

- Ad un concorso a 240 posti partecipano 3860 concorrenti. Calcolare la probabilità che, prendendo a caso  $n$  concorrenti,  $X$  di essi risultino vincitori, sapendo che la distribuzione di probabilità è normale e che: **a)**  $n=50$ ,  $X=4$ ; **b)**  $n=200$ ,  $X=12$ ; **c)**  $n=1000$ ,  $X=60$ ; **d)**  $n=2000$ ,  $X=130$ .  
[R.  $p_a \approx 20,39\%$ ;  $p_b \approx 11,59\%$ ;  $p_c \approx 5,02\%$ ;  $p_d \approx 3,22\%$ ]
- Una variabile aleatoria è distribuita normalmente con media  $\mu=80$  e deviazione standard  $\sigma=3,2$ . Calcolare la probabilità che assuma un valore:  
**a)** compreso fra 80 e 90;                      **b)** compreso fra 70 e 90;                      **c)** minore di 80;  
**d)** minore di 83;                                      **e)** maggiore di 85;                                      **f)** maggiore di 75;  
**g)** minore di 75 o maggiore di 82;                      **h)** compreso fra  $\mu-\sigma$  e  $\mu+\sigma$ ;  
**i)** compreso fra  $\mu-2\sigma$  e  $\mu+2\sigma$ ;                      **l)** compreso fra  $\mu-3\sigma$  e  $\mu+3\sigma$ .  
[R.  $p_a \approx 49,9\%$ ; ...;  $p_d \approx 82,5\%$ ; ...;  $p_g \approx 32,5\%$ ; ...]
- Il peso medio di un gruppo di persone è 72,5 kg con una deviazione standard di 3,4 kg. Ammesso che la distribuzione dei pesi delle persone sia normale, calcolare la probabilità che una persona, scelta a caso nel gruppo, abbia un peso:  
**a)** compreso fra 62 kg e 82 kg;                      **b)** minore di 72,5 kg;  
**c)** maggiore di 68 kg;                                      **d)** minore di 65 kg o maggiore di 85 kg.  
[R. ...;  $p_c \approx 90,7\%$ ;  $p_d \approx 1,4\%$ ]
- È stato controllato che gli impiegati di una certa città lavorano in media 38 ore settimanali con una deviazione standard di 2,2 ore settimanali. Calcolare quale probabilità si ha di trovare uno tra quegli impiegati che settimanalmente lavori:  
**a)** più di 36 ore;                                      **b)** meno di 36 ore;

- c) almeno 37 ore; d) meno di 34 o più di 42 ore,  
supponendo che la distribuzione di probabilità sia normale.

[R.  $p_a \approx 81,8\%$ ; ...;  $p_d \approx 6,9\%$ ]

5. Il grano prodotto annualmente da un'azienda agricola è stato in media, negli ultimi anni, di 80 q con una deviazione standard di 7,4 q. Calcolare quale probabilità si ha di ottenere quest'anno una quantità di grano: a) superiore a 80 q; b) compresa fra 75 e 85 q; c) inferiore a 70 q.  
[R. ...;  $p_c \approx 8,7\%$ ]
6. Una variabile aleatoria  $X$  è distribuita normalmente, con media 75,6 e deviazione standard 8,2. Dopo aver disegnato in un piano cartesiano ortogonale (Oxy) l'andamento della curva della probabilità associata alla  $X$ , calcolare per quale valore di  $t$  l'area della regione piana situata nel semipiano  $x \geq t$ , al di sotto della curva e al di sopra dell'asse  $x$ , è pari al  $k\%$  dell'area complessiva della regione situata al di sotto della curva e al di sopra dell'asse  $x$ , essendo: a)  $k=12$ ; b)  $k=50$ ; c)  $k=62$ .  
[R. a)  $t \approx 85,24$ ; ...]
7. I punteggi ottenuti in una prova di selezione dai concorrenti sono distribuiti normalmente, con media 124 e deviazione standard 11,5. Fra i 1788 concorrenti sono selezionati soltanto i migliori 350. Qual è il punteggio minimo (intero) conseguito dai concorrenti selezionati?  
[R. 134]
8. I punteggi ottenuti in una prova di selezione dai concorrenti sono distribuiti normalmente, con media 64 e deviazione standard 8. Fra i 478 concorrenti che hanno sostenuto la prova sono stati selezionati i migliori 25. Qual è il massimo punteggio, approssimato per difetto al primo decimale, che può aver ottenuto il primo dei concorrenti esclusi?  
[R. 76,1]
9. I tempi ottenuti dai 150 atleti che partecipano ad una gara podistica sono distribuiti normalmente, con media 2,15 h e deviazione standard 0,75 h. Qual è il tempo minimo che garantisce di classificarsi nei primi 15?  
[R.  $\approx 2,26$  h]

#### Distribuzione di Poisson (nn. 10-18).

10. Le statistiche mostrano che in una scatola, contenente 1000 pezzi di ricambio, si trovano mediamente 9 pezzi difettosi.
- Ricorrendo alla distribuzione binomiale calcolare la probabilità di trovare in uno scatolone contenete 5000 pezzi di ricambio:
    - esattamente 45 pezzi difettosi; B) non meno di 45 pezzi difettosi.
  - Spiegare se sono soddisfatte le condizioni per ricorrere alla distribuzione di Poisson.
    - In entrambi i casi precedenti eseguire il calcolo ricorrendo alla distribuzione di Poisson.
    - In entrambi i casi, quale errore si commette se si utilizza la distribuzione di Poisson al posto della distribuzione binomiale?
11. Le statistiche mostrano che nel periodo di punta, per un certo casello autostradale passano mediamente 8 veicoli ogni minuto. Qual è la probabilità che, in un minuto preso a caso nel periodo di punta, per quel casello passino: a) al più 4 veicoli? b) almeno 4 veicoli?  
[R.  $\approx 9,96\%$ ;  $\approx 95,76\%$ ]
12. Le statistiche mostrano che una famiglia riceve, in media, 5 telefonate al giorno. Qual è la probabilità che, in un giorno preso a caso, quella famiglia: a) riceva 10 telefonate? b) ne riceva almeno 2?  
[R.  $\approx 1,8\%$ ;  $\approx 96,0\%$ ]
13. Le statistiche mostrano che una famiglia ha, in media, due figli. Qual è la probabilità che una famiglia, scelta a caso: a) non abbia figli? b) ne abbia 1? c) ne abbia 2? d) ne abbia 3? e) ne abbia più di 3?  
[R. a)  $\approx 13,5\%$ ; b)  $\approx 27,1\%$ ; c)  $\approx 27,1\%$ ; d)  $\approx 18,0\%$ ; e)  $\approx 14,3\%$ ]

14. Al dottor Cesare, medico di famiglia, sono state affidate le persone di una piccola comunità. È stato rilevato il numero dei pazienti di quella comunità che nell’arco di un anno hanno richiesto le sue cure ed i dati sono riportati nella tabella sottostante.

Numero visite	0	1	2	3	4	5	6	7	8	9	10	11	12
Numero pazienti	2	5	18	34	46	64	57	54	37	27	16	7	5

Una volta trovata la distribuzione delle frequenze relative del numero dei pazienti che richiedono k visite, una rappresentazione grafica di questa distribuzione mostra intuitivamente che essa può essere assimilata con buona approssimazione ad una distribuzione di Poisson. Trovare la distribuzione delle frequenze assolute derivante da questa distribuzione di Poisson.

[R. I risultati sono sintetizzati nella tabella sottostante]

Numero visite	0	1	2	3	4	5	6	7	8	9	10	11	12	≥13
Numero pazienti	1	6	17	35	51	61	60	50	37	25	14	8	4	3

15. Su una scacchiera vengono fatti cadere a caso 100 monetine (le cose sono fatte in modo che tutte le monetine rimangano sulla scacchiera). Calcolare su quanti riquadri della scacchiera cadono k monetine, con  $k=0,1,2,3,4,5$ . (Si ricorda che i riquadri di una scacchiera sono 64).

[R. I risultati sono sintetizzati nella tabella sottostante]

Numero di monete per riquadro	0	1	2	3	4	5
Numero di riquadri	13	21	16	9	3	1

16. Una certa zona di guerra è stata suddivisa in 100 riquadri pressoché della stessa area. È previsto che su quella zona verranno lanciate 75 bombe. Dopo il lancio una bomba è caduta in un dato riquadro. Calcolare la probabilità che in quello stesso riquadro: a) cada al più un’altra bomba; b) non cadano altre bombe.

[R. ≈83% ; ≈47% ]

17. Con riferimento all’esercizio precedente, si supponga che ognuno dei 100 riquadri sia ripartito a sua volta in 16 riquadri più piccoli.

- Calcolare la probabilità che una bomba (e una soltanto) cada in uno dei riquadri “piccoli”.
- Supposto di sapere che in un determinato riquadro “grande” siano cadute 2 bombe, calcolare la probabilità che una di esse sia caduta in uno dei riquadri “piccoli” in cui è stato suddiviso quello grande.

[R. ≈4,47% ; ... ]

18. Le statistiche mostrano che una persona passa, in media, 4 ore al giorno davanti al televisore.

1. Qual è la probabilità che una persona, scelta a caso, passi davanti al televisore:

A) 4 ore al giorno? B) più di 4 ore al giorno? C) meno di 4 ore al giorno?

2. Qual è la probabilità che una persona, scelta a caso, passi davanti al televisore:

A) esattamente 100 ore in un mese di 30 giorni? B) non più di 100 ore in un mese di 30 giorni?

[R. 1A) ≈19,5% ; 1B) ≈37,1% ; 1C) ... ; 2A) ≈0,68% ; 2B) ≈3,47% ]

## UNA BREVE SINTESI PER DOMANDE E RISPOSTE

### DOMANDE.

1. Sotto quali condizioni una distribuzione binomiale può essere approssimata da una distribuzione normale?

2. Quali sono le proprietà principali della curva normale della probabilità?
3. Indicata con  $X$  una generica variabile aleatoria normale, di media aritmetica  $\mu$  e deviazione standard  $\sigma$ :  
a) qual è la formula che permette di calcolare  $P[a \leq X \leq b]$  per mezzo della distribuzione normale standardizzata? b) quanto vale la probabilità che la variabile  $X$  acquisti un valore compreso fra  $\mu - k\sigma$  e  $\mu + k\sigma$ , con  $k$  numero reale positivo assegnato?
4. Le altezze delle persone di una certa comunità sono distribuite normalmente con media 172 cm e deviazione standard 9 cm. Quali sarebbero la media e la deviazione standard se tutte le persone fossero più alte di 2 cm?
5. Se un evento casuale si presenta mediamente  $m$  volte ogni ora, quanto vale la probabilità che si presenti esattamente  $n$  volte in un'ora stabilita?
6. Posto che contro una superficie, ripartita in  $s$  unità vengano lanciati  $n$  proiettili, quanto vale la probabilità che in una prescelta unità cadano  $k$  proiettili?
7. Sai indicare una variabile aleatoria che abbia media 0 e deviazione standard 1? Sai indicare una variabile aleatoria la cui media, a parte l'unità di misura eventuale, è numericamente uguale alla sua varianza?

**RISPOSTE.**

1. La distribuzione binomiale  $B(n,p)$  può esser approssimata dalla distribuzione normale  $N(\mu, \sigma^2)$ , dove  $\mu = np$  e  $\sigma = \sqrt{np(1-p)}$  in due casi:  
a) ogni volta che  $n \geq 10$  e  $|p-q| < 0,04$  nel qual caso l'approssimazione è eccellente;  
b) quando  $n \geq 30$ ,  $np > 5$  ed  $nq > 5$ , nel qual caso l'approssimazione è accettabile.
2. La curva normale della probabilità:  
- è simmetrica rispetto alla retta  $x = \mu$ ;  
- è collocata interamente nel semipiano  $y > 0$ ;  
- l'area sotto il suo grafico è uguale ad 1.
3. Si ha:  

$$P[a \leq X \leq b] = P\left[\frac{a-\mu}{\sigma} \leq N \leq \frac{b-\mu}{\sigma}\right]$$
mentre la probabilità che la variabile  $X$  acquisti un valore compreso fra  $\mu - k\sigma$  e  $\mu + k\sigma$  è:  

$$P[\mu - k\sigma \leq X \leq \mu + k\sigma] = 2 P[0 \leq N \leq k].$$
4. La media aumenterebbe essa stessa di 2 cm, diventando 174 cm, mentre la deviazione standard resterebbe immutata.
5. La probabilità cercata  $p_n$  è data dalla formula di Poisson:  

$$p_n = \frac{m^n}{n!} e^{-m}.$$
6. La probabilità cercata  $p_k$  è data dalla formula di Poisson:  

$$p_k = \frac{m^k}{k!} e^{-m}, \text{ dove } m = \frac{n}{S}.$$
7. Ha media 0 e deviazione standard 1 la variabile aleatoria normale standard. La variabile aleatoria di Poisson è invece tale che la sua media è uguale alla sua varianza.