8. STATISTICA INDUTTIVA

8.1 Regressione lineare

In molto esperimenti al ricercatore interessa l'analisi delle variazioni di due o più variabili per evidenziare le eventuali relazioni esistenti tra di loro e predire valori interessanti delle variabili non note sperimentalmente.
Noi ci occupiamo solo di esperimenti con due variabili.
La relazione che viene presa in esame è la dipendenza di una variabile rispetto all'altra

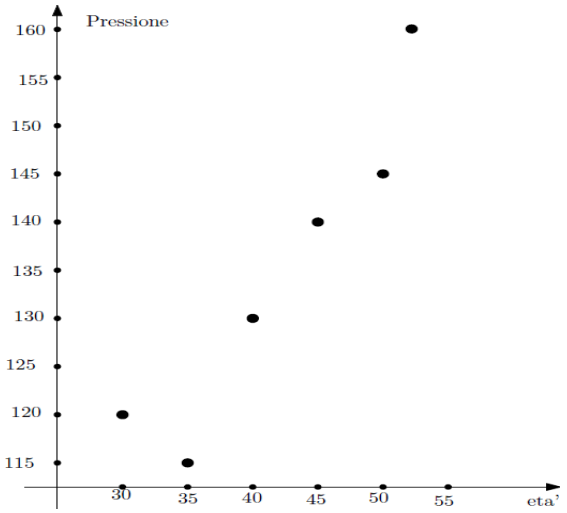
Il rapporto di dipendenza in matematica si indica con *funzione,* in statistica lo indicheremo con la parola *regressione.*
Si indica come indipendente (X), una variabile per cui i livelli possono essere fissati o possono essere fissati sperimentalmente (es. le dosi di una sostanza), oppure possono essere semplicemente rilevati (es. la temperatura).
Si indica come variabile dipendente (Y), una variabile la cui variazione si assume essere la risposta alle variazioni della variabile indipendente.

Esempio Nella seguente tabella sono riportate le età e le pressioni arteriose massime (in $mmHg$ millimetri di mercurio) di 6 soggetti maschili.

ETÀ	30	35	40	45	50	52
PRESSIONE	120	115	130	140	145	160

OBIETTIVO: cercare se esiste una dipendenza del valore della pressione massima dall'età maschile.

Per capire questa dipendenza la prima cosa da fare è rappresentare i dati nel *diagramma di dispersione*: un grafico cartesiano in cui mettiamo in ascissa la variabile indipendente (in questo caso l'età) e in ordinata la variabile dipendente (in questo caso la pressione).
I dati sperimentali sono quindi rappresentati da punti nel grafico di dispersione.
Noi saremo interessati a quelle variabili che dipendono linearmente l'una dall'altra. (graficamente parlando saremo interessati a quei grafici a dispersione in cui i punti sembrano allinearsi su una retta).
Nel nostro esempio:



I dati sembrano effettivamente disporsi su una retta. Esistono diversi tipi di regressione:
uno lineare è quello di regressione basato sulla retta dei minimi quadrati (dal grafico sembra che questo possa essere un buon modello per rappresentare la dipendenza dei dati)

Vedere il grafico non basta però a capire se effettivamente tra i dati esiste una relazione lineare (l'effetto visivo potrebbe confonderci), dobbiamo osservare matematicamente se esiste questa dipendenza.
Per far questo introduciamo il *coefficiente di correlazione* r, che ci permette di capire se esiste o meno una buona relazione lineare tra le variabili.

$$r = \frac{\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \overline{x})^2 \sum_{i=1}^{n}(y_i - \overline{y})^2}} \quad \text{con } \overline{x}, \overline{y} \text{ medie dei dati } x_i, y_i$$

Il numeratore può essere espresso più semplicemente :

$$\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y}) = \sum_{i=1}^{n}(x_i y_i - x_i \overline{y} - y_i \overline{x} + \overline{x}\,\overline{y}) = \sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \overline{y} - \sum_{i=1}^{n} y_i \overline{x} + \sum_{i=1}^{n} \overline{x}\,\overline{y}$$

$$= \sum_{i=1}^{n} x_i y_i - n\overline{x}\,\overline{y} - n\overline{x}\,\overline{y} + n\overline{x}\,\overline{y} = \sum_{i=1}^{n} x_i y_i - n\overline{x}\,\overline{y}$$

Perché:

$$\frac{\sum_{i=1}^{n} x_i}{n} = \overline{x} \quad e \quad \frac{\sum_{i=1}^{n} y_i}{n} = \overline{y}$$

Dunque:

$$r = \frac{\sum_{i=1}^{n} x_i y_i - n\overline{x}\,\overline{y}}{\sqrt{\sum_{i=1}^{n}(x_i - \overline{x})^2 \sum_{i=1}^{n}(y_i - \overline{y})^2}}$$

Il coefficiente di regressione:

- È sempre $-1 \leq r \leq 1$,
- Se r è vicino a 1 o -1 c'è una buona correlazione lineare,
- Se $r = \pm 1$ la correlazione è perfetta
- Se $r < 0$ la corrispondenza è inversa, all'aumentare della variabile x diminuisce la variabile y
- Se $r = 0$ o si avvicina a 0, significa che i dati non sono in relazione lineare, in questo caso non ha senso approssimare i dati con la retta dei minimi quadrati.

Per i calcoli si può utilizzare la seguente tabella, nel nostro esempio:

$$\overline{x} = \frac{\sum_{i=1}^{n} x_i}{n} = 42 \qquad \overline{y} = \frac{\sum_{i=1}^{n} y_i}{n} = 135$$

x_i	$x_i - \overline{x}$	$(x_i - \overline{x})^2$	y_i	$y_i - \overline{y}$	$(y_i - \overline{y})^2$	$x_i y_i$
30	-12	144	120	-15	225	3600
35	-7	49	115	-20	400	4025
40	2	4	135	-5	25	5200
45	3	9	140	5	25	6300
50	8	64	145	10	100	7250
52	10	100	160	25	625	8320
		TOT: 370			TOT: 1400	TOT: 34695

Quindi abbiamo che $r = \frac{(34695 - 6 \cdot 42 \cdot 135)}{\sqrt{370 \cdot 1400}} = \frac{675}{719.722} \cong 0.938$

Notiamo che r si avvicina ad 1 ne segue che esiste una buona correlazione lineare tra le variabili.

I dati possono essere rappresentati dalla retta dei minimi quadrati: $y = mx + q$ dove:

$$m = \frac{\sum_{i=1}^{n} x_i y_i - n\overline{x}\,\overline{y}}{\sqrt{\sum_{i=1}^{n} (x_i - \overline{x})^2}} \quad \text{e} \quad q = \overline{y} - m\overline{x}$$

Nel nostro caso: $m = 1.824$ e $q = 58.38$

$$y = 1.82x + 58.38$$

La retta ci permette di fare delle previsioni, cioè

- Dire, presumibilmente, qual è la pressione massima di un maschio di 47 anni:

$$y = 1.82 \cdot 47 + 58.38 = 143.92$$

- Dire quale presunta età può avere un uomo con pressione massima pari a $150\ mmHg$

$$x = \frac{y - 58.38}{1.82} = 50.34$$

8.2 Distribuzioni di probabilità

VARIABILI ALEATORIE DISCRETE

Nell'ambito del calcolo delle probabilità si introduce un nuovo concetto di "variabile", quello di variabile aleatoria, che è fondamentale in svariate applicazioni.

Intuitivamente una variabile aleatoria è una variabile i cui valori sono numeri reali determinati dall'esito di un esperimento aleatorio. Per esempio sono varabili aleatorie:

- T che rappresenta il numero di testa uscite nel lancio di 100 monete,
- H che rappresenta l'altezza di una persona scelta a caso nella popolazione,
- N che rappresenta in numero di autovetture che giungono in un giorno ad un casello autostradale.

Una variabile aleatoria si può anche interpretare come una funzione, per esempio se consideriamo la variabile aleatoria X "somma dei numeri ottenuti nel lancio dei due dadi", essa si può interpretare come la funzione che associa ad ogni possibile esito del lancio la somma dei due numeri ottenuti.

Questa osservazione consente di dare una definizione più formale di variabile aleatoria:

Definizione Si chiama *variabile aleatoria*, una funzione che associa ad ogni possibile esito di un esperimento aleatorio un numero reale.

Se Ω è lo spazio campione di un esperimento e X è una variabile aleatoria legata all'esperimento, allora $X: \Omega \longrightarrow \mathbb{R}$

Esempio

Consideriamo l'esperimento che consiste nel lancio di tre monete eque, indichiamo con X il numero di "testa" che si ottiene. Rappresentiamo la variabile aleatoria X.

X rappresenta il numero di teste uscite lanciando tre monete, quindi X può assumere solo i valori $\{0,1,2,3\}$, lo spazio campione $\Omega = \{(CCC), (CCT), (CTC), (CTT), (TCC), (TCT), (TTC), (TTT)\}$, possiamo rappresentare X come segue:

$$\begin{array}{rcl} X: \quad \Omega & \longrightarrow & \{0,1,2,3\} \\ (CCC) & \longmapsto & 0 \\ (TCC), (CTC), (CCT) & \longmapsto & 1 \\ (TTC), (CTT), (TCT) & \longmapsto & 2 \\ (TTT) & \longmapsto & 3 \end{array}$$

Una variabile aleatoria X, che assume un numero finito o numerabile di valori si dice *discreta*.

L'evento: "X assume il valore x_i" (che rappresenta un evento essendo la controimmagine di x_i e quindi un sottoinsieme di Ω), si rappresenta con il simbolo $X = x_i$.

Possiamo anche calcolare la probabilità di $X = x_i$, essa è uguale alla somma delle probabilità degli eventi elementari la cui immagine tramite X è uguale ad x_i.

Esempio: Considerando l'esempio precedente, osserviamo che:

- La probabilità che $X = 0$ è uguale alla probabilità dell'evento elementare CCC, quindi vale $\frac{1}{8}$.
- La probabilità che $X = 1$ è uguale alla somma delle probabilità degli eventi elementari TCC, CTC, CCT, quindi vale $\frac{3}{8}$.
- La probabilità che $X = 2$ è uguale alla somma delle probabilità degli eventi elementari TTC, CTT, TCT, quindi vale $\frac{3}{8}$.

- La probabilità che $X = 3$ è uguale alla probabilità dell'evento elementare TTT, quindi vale $\frac{1}{8}$.

Formalmente scriviamo:

$$p(X = 0) = \frac{1}{8}, \quad p(X = 1) = \frac{3}{8}, \quad p(X = 2) = \frac{3}{8}, \quad p(X = 3) = \frac{1}{8}$$

Supponendo ancora che X sia una variabile aleatoria discreta che assume i valori $x_1, x_2, \dots x_n$, possiamo associare a ciascuno degli eventi, $X = x_1, \dots, X = x_n$, la rispettiva probabilità, si definisce così una nuova funzione:

Definizione: Sia X una variabile aleatoria discreta che assume i valori $x_1, x_2, \dots, x_n$ rispettive probabilità $p_1, p_2, \dots, p_n$; si chiama *distribuzione di probabilità (o densità),* della variabile aleatoria X, la funzione che associa a ciascun x_i la rispettiva probabilità p_i.

La distribuzione di probabilità di una variabile aleatoria discreta si rappresenta tramite una tabella:

x_i	x_1	x_2	…	x_n
$p(X = x_i)$	p_1	p_2	…	p_n

Poiché gli eventi $X = x_i$, sono disgiunti si ha che $p_1 + p_2 + \cdots + p_n = 1$.

Esempio Riprendendo l'esempio precedente si ha che

x_i	0	1	2	3
$p(X = x_i)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

C'è una certa similitudine tra il concetto di distribuzione di probabilità e quello di distribuzione di frequenze relative in statistica, la similitudine continua anche nella definizione di concetti di *media, varianza, scarto quadratico medio.*

Definizioni: Sia X una variabile aleatoria discreta che assume i valori $x_1, x_2, \dots, x_n$, con probabilità rispettive $p_1, p_2, \dots, p_n$.

a) Si chiama *media (o valore atteso),* della variabile aleatoria X, e si indica con $E(X)$, o con la lettera μ, il numero:

$$\mu = E(X) = x_1 p_1 + \cdots + x_n p_n$$

b) Si definisce *varianza* di X, e si indica con il simbolo $V(X)$ o con il simbolo σ^2, il numero così definito

$$\sigma^2 = V(X) = \big(x_1 - E(X)\big)^2 \cdot p_1 + \cdots + \big(x_n - E(X)\big)^2 \cdot p_n$$

c) Si definisce *scarto quadratico medio (o deviazione standard)* di X, e si indica con il simbolo σ, la radice quadrata della varianza, $\sigma = \sqrt{\sigma^2}$.

Esempio Calcola media, varianza, scarto quadratico medio dell'esempio precedentemente studiato. (quante teste lanciando tre monete).

VARIABILI ALEATORIE CONTINUE

Esistono parecchi fenomeni reali per la cui descrizione le variabili aleatorie discrete non sono adatte. Per esempio è necessaria una variabile aleatoria *continua* (ovvero una variabile che può assumere tutti i valori reali in un dato intervallo), per descrivere il tempo di vita di un apparecchio soggetto a guasti casuali, o per descrivere il gioco del lancio delle freccette. Inoltre facendo ad esempio riferimento al lancio delle freccette, è inutile chiedersi la probabilità che la freccetta colpisca esattamente un punto sul disco, ma ha più senso cercare di capire la probabilità che la freccetta colpisca una zona ben definita del disco. Si è dunque interessati alla probabilità non di eventi rappresentati da singoli punti ma di eventi rappresentati da intervalli. Proprio per questo una variabile aleatoria X è definita assegnando una funzione che permette di calcolare la probabilità che X assumi valori in un qualsivoglia intervallo.

Definizione Una variabile aleatoria continua X, viene definita assegnando una funzione f, detta *densità (di probabilità)* di X, che soddisfi le seguenti proprietà:

- $f(x) \geq 0$ per ogni $x \in \mathbb{R}$,
- $\int_{-\infty}^{+\infty} f(x)dx = 1$

La probabilità che X assuma valori in un determinato intervallo I è data dall'integrale della sua densità sull'intervallo I.

Osservazioni:

a) Se l'intervallo $I = [a, b]$ si riduce ad un punto, cioè se $a = b$, risulta che:

$$p(X \in I) = \int_a^b f(x)dx \qquad e \qquad p(X = a) = \int_a^a f(x)dx = 0$$

Perciò se X è una variabile aleatoria continua, la probabilità che essa assuma un qualsivoglia valore reale prefissato è sempre nulla.

b) Data la densità di probabilità f di una variabile aleatoria continua X, il valore $f(a)$ da essa assunto quando $x = a$, NON ha (come accada nel caso discreto) il significato di probabilità dell'evento $X = a$, infatti questa probabilità è sempre uguale a zero, mentre il valore assunto da f in a, in generale è un numero positivo.

Diamo ora alcune definizioni che riprendiamo dal caso discreto:

Definizioni: Data una variabile aleatoria continua X, di densità f,

- si dice *media* o *valore atteso* di X e si indica con il simbolo $E(X)$ (o con la lettera μ), il numero se esiste, così definito

$$\mu = E(X) = \int_{-\infty}^{+\infty} xf(x)dx$$

- si dice *varianza* di Xe si indica con il simbolo $V(X)$, o con σ^2, il numero, se esiste, così definito

$$\sigma^2 = V(X) = \int_{-\infty}^{+\infty} (x-\mu)^2 f(x)dx$$

- si definisce *deviazione standard* di X (e si indica con il simbolo σ) la radice quadrata della varianza

$$\sigma = \sqrt{V(x)}$$

Definizione Sia una variabile aleatoria continua X, avente come densità la funzione f, si chiama *funzione di ripartizione* di X la funzione che, per ogni $x \in \mathbb{R}$, è così definita:

$$F(x) = p(X \leq x) = \int_{-\infty}^{x} f(t)dt$$

Esistono diverse funzioni notevoli di densità di probabilità :

- uniforme,
- binomiale
- Poisson
- Gauss,

Concentreremo la nostra attenzione in modo particolare sull'ultima.

8.3 Distribuzione normale (o di Gauss)

La funzione di densità di questa particolare distribuzione è data da:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \qquad \forall\, x \in \mathbb{R}$$

Con μ e σ rispettivamente valor medio e deviazione standard di una variabile aleatoria continua (v.a.c.) X.
Cerchiamo di capire l'andamento di $f(x)$ e andiamo a studiarla.

- Dominio $\mathcal{D} = \mathbb{R}$
- Segno $f(x) > 0 \quad \forall\ x \in \mathbb{R},$ è simmetrica rispetto alla retta $x = \mu$.
- Limiti agli estremi del dominio:

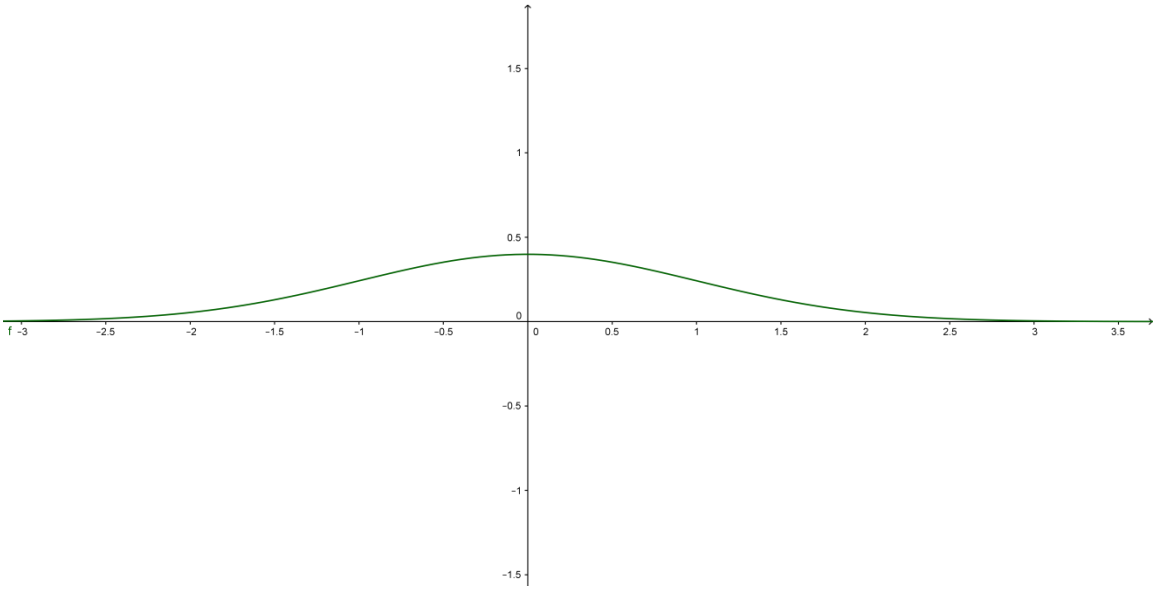
  $$\lim_{x \mapsto \pm\infty} f(x) = \lim_{t \mapsto \pm\infty} e^{-t^2} = 0$$

  La retta $y = 0$ è asintoto orizzontale.
- Crescenza e decrescenza: $f'(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \cdot \left(-\frac{2}{2\sigma^2}(x-\mu)\right) \geq 0 \iff x \leq \mu$

  Si ha che $x = \mu$ è punto di massimo, e il massimo vale $f(\mu) = \frac{1}{\sqrt{2\pi}\sigma}$
- Concavità e convessità: $f''(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \cdot \left(-\frac{1}{\sigma^2} + \frac{(x-\mu)^2}{\sigma^4}\right) \geq 0$

  Se e solo se $\frac{(x-\mu)^2}{\sigma^2} - 1 \geq 0 \iff (x-\mu)^2 \geq \sigma^2 \iff x \leq -\sigma + \mu \ \wedge \ x \geq \sigma + \mu$

  $x = \mu - \sigma$ e $x = \mu + \sigma$ sono punti di flesso. Il flesso vale: $f(\mu + \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}}$.



Esempi di caratteristiche che si distribuiscono normalmente.

i. Errori di misurazione di una grandezza fisica
ii. Peso e altezza di una popolazione omogenea.
iii. Dimensioni di oggetti prodotti in serie.

La probabilità $p(X \leq x) = F(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ si calcola utilizzando varie tabelle dove è riportato il valore della DISTRIBUZIONI NORMALE STANDARD.

Sia X una v.a.c. con media μ e scarto quadratico σ, allora diremo che $X \in \mathcal{N}(\mu, \sigma)$.

Definizione Sia X una v.a.c. con $X \in \mathcal{N}(0,1)$, funzione di densità $f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$, la funzione di partizione $F(x) = p(X \leq x) = \int_{-\infty}^{x} f(t)dt$ associata ad X è detta *distribuzione normale standard.*

$f(x)$ è una funzione pari e solitamente si indica con $\varphi(x) = p(X \leq x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt$.

REGOLE GENERALE:

i. $p(X \leq x) = \varphi(x)$
ii. $p(X > x) = 1 - \varphi(x)$
iii. $p(X \leq -x) = \varphi(-x) = 1 - \varphi(x)$
iv. $p(-x \leq X \leq x) = \varphi(x) - \varphi(-x) = \varphi(x) - \big(1 - \varphi(x)\big) = 2\varphi(x) - 1$
$= 1 - 2\varphi(-x) = 1 - 2\big(1 - \varphi(x)\big) = 1 - 2 + 2\varphi(x) = 2\varphi(x) - 1$

Dalle tavole si ottiene che :

- $p(X \leq 0) = \frac{1}{2}$
- $p(-1 \leq X \leq 1) = 2\varphi(1) - 1 \cong 0.6827$
- $p(-2 \leq X \leq 2) \cong 0.9545$
- $p(-3 \leq X \leq 3) \cong 0.9973$

Come facciamo però se X non è standard? Dobbiamo trovare un modo per renderla standard!!!

STANDARIZZAZIONE DI UNA VARIABILE ALEATORIA CONTINUA

Se X è una variabile aleatoria continua con $X \in \mathcal{N}(\mu, \sigma)$, allora $Z = \frac{X-\mu}{\sigma}$ è una v.a.c. con $Z \in \mathcal{N}(0,1)$ che si chiama *standarizzazione di* X.

Dunque:

$$p(X \leq x) = p\left(\frac{X-\mu}{\sigma} \leq \frac{x-\mu}{\sigma}\right) = p\left(Z \leq \frac{x-\mu}{\sigma}\right) = \varphi\left(\frac{x-\mu}{\sigma}\right)$$

Quindi qualsiasi sia la v.a.c. che considero, grazie alla standarizzazione riesco a leggere la probabilità che mi interessa.

Da quanto detto precedentemente

- $p(\mu - \sigma \leq X \leq \mu + \sigma) = p(-1 \leq Z \leq 1) \cong 0.6827$
- $p(\mu - 2\sigma \leq X \leq \mu + 2\sigma) = p(-2 \leq Z \leq 2) \cong 0.9545$
- $p(\mu - 3\sigma \leq X \leq \mu + 3\sigma) = p(-3 \leq Z \leq 3) \cong 0.9973$

Esercizio (uso della tabella)

Il livello di glucosio a digiuno in un gruppo di pazienti non diabetici si distribuisce normalmente con $\mu = 98\, mg/100\, ml$ e $\sigma = 8\, mg/100ml$.

Determinare:

a) $p(X < 82\, mg/100\, ml)$
b) $p(90 \leq X \leq 106\, mg/100ml)$
c) $p(X > 116\, mg/100\, ml)$

SVOLGIMENTO

$X \in \mathcal{N}(98,8)$ utilizzando la sua standardizzazione Z otteniamo:

a) $p(X < 82mg/100\, ml) = p\left(Z < \frac{82-98}{8}\right) = p(Z < -2) = \varphi(-2) = 1 - \varphi(2) = 1 - 0.9772 = 0.0228 = 2{,}28\,\%$
b) $p(90 \leq X \leq 106) = p\left(\frac{90-98}{8} \leq Z \leq \frac{106-98}{8}\right) = p(-1 \leq Z \leq 1) \cong 68{,}26\%$
c) $p(X > 116) = p\left(Z > \frac{116-98}{8}\right) = p(Z > 2{,}25) = 1 - p(Z < 2{,}25) = 1 - \varphi(2{,}25) = 1 - 0{,}9878$

$\cong 1{,}22\%$.

Teorema (del limite centrale)
Siano $X_1, X_2, \dots, X_n$ variabili aleatorie indipendenti, aventi tutte la stessa distribuzione di probabilità, (quindi stessa media μ e stessa varianza σ^2). Allora

$$\lim_{n\mapsto+\infty} p\left(\frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}} \leq x\right) = p(\mathcal{N}(0,1) \leq x)$$

Qual è il suo significato? In sostanza possiamo approssimare le media di $X_1, X_2, \dots, X_n$ con una variabile normale avente la media comune μ e la varianza $\frac{\sigma^2}{n}$.
Questo teorema ci porta a formulare delle stime attraverso gli intervalli di confidenza.

INTERVALLI DI CONFIDENZA.
Sia X una caratteristica di una popolazione P che vogliamo studiare. Supponiamo che X sia normalmente distribuita con media μ e varianza σ^2, di cui non siamo sempre a conoscenza. Estraiamo un campione casuale di n elementi di P, siano $x_1, x_2, \dots, x_n$ i valori di X per gli elementi del campione.
Possiamo pensare di avere $X_1, X_2, \dots, X_n$ v.a.c. del tipo $\mathcal{N}(\mu, \sigma^2)$ e che $x_1, x_2, \dots, x_n$ siano i valori assunti da $X_1, X_2, \dots, X_n$ dopo aver svolto l'esperimento.
Sia $\overline{X} = \frac{X_1+X_2+\cdots+X_n}{n}$ la media campionaria.
Per il teorema del limite centrale $\frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}} \xrightarrow[n\mapsto+\infty]{} \mathcal{N}(0,1)$,
ossia per $n \gg 0$ si ha che $\overline{X} \sim \mathcal{N}(\mu, \frac{\sigma^2}{n})$
Possiamo usare queste informazioni per ottenere stime di tipo probabilistico per la media μ:
Scegliamo una probabilità $P \in (0,1)$, possiamo costruire un intervallo (μ_1, μ_2), detto intervallo di confidenza tale che $p\big(\mu \in (\mu_1, \mu_2)\big) = P$.

Esempio Sia X una v.a.c con $\mathcal{N}(\mu, \sigma^2)$ con $\sigma^2 = 9$. Da un campione casuale di 5 elementi della popolazione P si ottiene $\overline{x} = 61$.
Determinare intervallo (μ_1, μ_2) tale che la $p\big(\mu \in (\mu_1, \mu_2)\big) = 95\%$
Lo scopo è quindi quello di determinare un intervallo di confidenza al 95% per μ.
Sto quindi cercando Z una v.a.c. del tipo $\mathcal{N}(0,1)$ tale che $p(-z \leq Z \leq z) = 0{,}95$. Determinato Z, finalmente posso trovare la media μ sfruttando la relazione $Z = \frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}}$.
Ora:
$p(-z \leq Z \leq z) = 0.95$ ma $p(-z \leq Z \leq z) = 2p(Z < z) - 1$ quindi

$$2p(Z < z) - 1 = 0{,}95 \;\Rightarrow p(Z < z) = \frac{0{,}95+1}{2} = 0{,}9750$$

Cerco $0{,}9750$ nella tabella e vedo che corrisponde a $z = 1{,}96$.
Otteniamo dunque $-1{,}96 < z < 1{,}96$ ch3e per il teorema del limite centrale

$$-1{,}96 < \frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}} < 1{,}96$$

Sviluppando i conti:

$$-1{,}96 \cdot \frac{\sigma}{\sqrt{n}} < \overline{X} - \mu < 1{,}96 \cdot \frac{\sigma}{\sqrt{n}}$$

$$-1{,}96 \cdot \frac{\sigma}{\sqrt{n}} - \overline{X} < -\mu < 1{,}96 \cdot \frac{\sigma}{\sqrt{n}} - \overline{X}$$

$$\overline{X} - 1{,}96 \cdot \frac{\sigma}{\sqrt{n}} < \mu < \overline{X} + 1{,}96 \cdot \frac{\sigma}{\sqrt{n}}$$

Sostituiamo con i valori di cui siamo a conoscenza e otteniamo:

$$61 - 1{,}96 \cdot \frac{3}{\sqrt{5}} < \mu < 61 + 1{,}96 \cdot \frac{3}{\sqrt{5}}$$

$$58{,}37 < \mu < 63{,}63$$

Si ha che al 95% $\mu \in (58{,}37; 63{,}63)$.

Osservazione: all'aumentare di P aumenta l'ampiezza dell'intervallo $(\mu_1;\ \mu_2)$. Infatti considerando $p(\mu \in (\mu_1, \mu_2)) = 0{,}98$ si ha che:

$$p(-z \leq Z \leq z) = 0.98 \Rightarrow 2p(Z < z) - 1 = 0{,}98 \ \ \Rightarrow p(Z < z) = \frac{0{,}98 + 1}{2} = 0{,}9898$$

$z = 2{,}32$ (valore cercato nella tabella) $\quad \overline{X} - 2{,}32 \cdot \frac{\sigma}{\sqrt{n}} < \mu < \overline{X} + 2{,}32 \cdot \frac{\sigma}{\sqrt{n}}$

$$61 - 2{,}32 \cdot \frac{3}{\sqrt{5}} < \mu < 61 + 2{,}32 \cdot \frac{3}{\sqrt{5}}$$

$$57{,}89 < \mu < 64{,}11$$

Quindi $\mu \in (57{,}89; 64{,}11)$ al 98%.