ESERCIZIO 5.1

Si seleziona un campione di n=25 osservazioni da una popolazione distribuita secondo una variabile casuale uniforme con media uguale a 80 e deviazione standard uguale a 5.

1) Si descriva la distribuzione di probabilità della media campionaria $\bar{x}$
2) Si calcoli la $P(\bar{x} > 82)$

SVOLGIMENTO

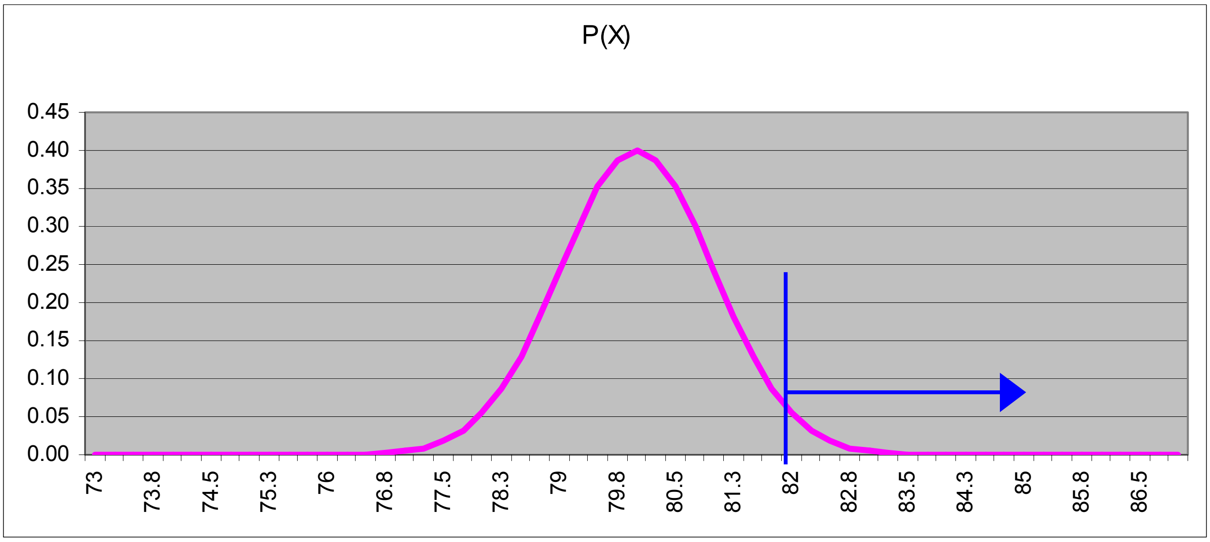
1)

Il teorema del limite centrale, vista la simmetria della distribuzione della popolazione cui si fa riferimento, ci permette di affermare che la distribuzione campionaria di $\bar{x}$ sarà approssimativamente normale. In particolare si ha:

$$\mu_{\bar{x}} = \mu = 80$$

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{5}{\sqrt{25}} = 1$$

2)



$$\bar{x} = 82 \Rightarrow Z = \frac{\bar{x} - \mu_{\bar{x}}}{\sigma_{\bar{x}}} = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}} = \frac{82 - 80}{1} = 2$$

Dalle tavole otteniamo il valore tra la media e il valore Z specifico:

P(0 < Z < z)

	0.00	0.01	0.02	0.03
1.7	0.4554	0.4564	0.4573	0.4582
1.8	0.4641	0.4649	0.4656	0.4664
1.9	0.4713	0.4719	0.4726	0.4732
2.0	0.4772	0.4778	0.4783	[illegible]

Area tra la media e il valore z

Da cui è possibile calcolare la probabilità di interesse come segue:

$$P(\bar{x} > 82) = P(Z > 2) = P(Z > 0) - P(0 < Z < 2) = 0.5 - 0.4772 = 0.0228$$

ESERCIZIO 5.2

Un produttore di batterie per automobili pubblicizza il suo prodotto affermando che la distribuzione relativa alla durata di vita della batteria ha media 54 mesi e deviazione standard 6 mesi.
Un gruppo di tutela dei consumatori decide di verificare quanto affermato dal produttore in questione acquistando 50 batterie e sottoponendole ad opportuni test al fine di determinarne la durata media. Assumendo che quanto affermato dal produttore sia vero,

1) si descriva la distribuzione campionaria della durata media del campione di 50 batterie
2) si calcoli la probabilità che la durata media del campione di 50 batterie sia uguale o inferiore a 52 mesi

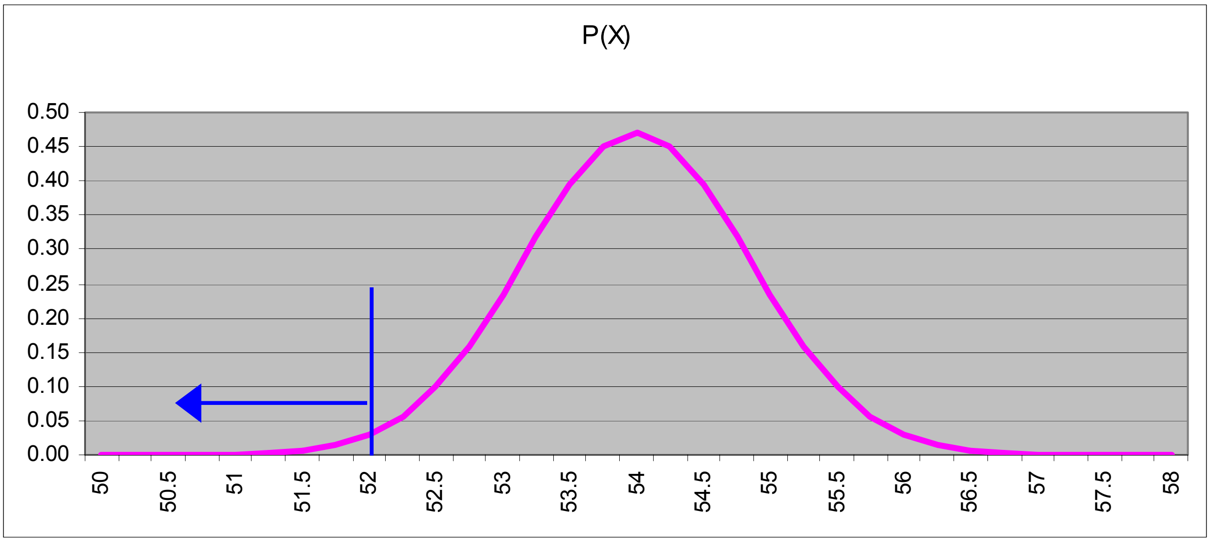
SVOLGIMENTO

1)
Pur non avendo a disposizione informazioni sulla distribuzione di probabilità della v.c. vita delle batterie, si può sfruttare il teorema limite centrale per affermare che la distribuzione campionaria della durata media del campione di 50 batterie si distribuisce approssimativamente come una normale. Inoltre:

$$\mu_{\bar{x}} = \mu = 54 \text{ mesi}$$

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{6}{\sqrt{50}} = 0.85 \text{ mesi}$$

2)



$$\bar{x} = 52 \Rightarrow Z = \frac{\bar{x} - \mu_{\bar{x}}}{\sigma_{\bar{x}}} = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}} = \frac{52 - 54}{0.85} = -2.35$$

Dalle tavole otteniamo il valore tra la media e il valore Z specifico:

P(0 < Z < z)

	0.02	0.03	0.04	0.05
2.0	0.4783	0.4788	0.4793	0.4798
2.1	0.4830	0.4834	0.4838	0.4842
2.2	0.4868	0.4871	0.4875	0.4878
2.3	0.4898	0.4901	0.4904	0.4906

Area tra la media e il valore z

Da cui è possibile calcolare la probabilità di interesse come segue:

$$P(\bar{x} \le 52) = P(Z \le -2.35) = P(Z \ge 2.35) = P(Z \ge 0) - P(0 < Z < 2.235) = 0.5 - 0.4906 = 0.0094$$

ESERCIZIO 5.3

Una grande banca vuole stimare l'ammontare medio di denaro dovuto dai clienti in mora (vale a dire in ritardo nei pagamenti più di due mesi).
La banca decide di selezionare casualmente 100 tra questi clienti, su cui si osserva un ammontare medio di € 2310. Osservazioni fatte negli anni precedenti fanno ritenere plausibile un valore di $\sigma = 89$.
Si costruisca un intervallo di confidenza per l'ammontare medio di denaro dovuto dai clienti in mora

1) al 90%
2) al 95%
3) al 99%
4) Si desidera stimare il parametro μ (ammontare medio di denaro dovuto dai clienti in mora) accettando un errore di 5€. Determinare l'ampiezza campionaria necessaria a tal fine usando un livello di confidenza $\alpha = 0.10$

SVOLGIMENTO

n=100 $\bar{x} = 2310$ $\sigma = 89$

Per risolvere i tre quesiti può essere utile riassumere nella seguente tabella i coefficienti di confidenza con i rispettivi valori per la Z più comunemente utilizzati:

100 (1-α)%	α	α/2	$Z_{\alpha/2}$
90%	0.10	0.050	1.645
95%	0.05	0.025	1.960
99%	0.01	0.005	2.576

$$P\left(-Z_{\frac{\alpha}{2}} \le \frac{\bar{x} - \mu_{\bar{x}}}{\sigma_{\bar{x}}} \le +Z_{\frac{\alpha}{2}}\right) = (1-\alpha)$$

$$P\left(\bar{x} - Z_{\frac{\alpha}{2}}\sigma_{\bar{x}} \le \mu_{\bar{x}} \le \bar{x} + Z_{\frac{\alpha}{2}}\sigma_{\bar{x}}\right) = (1-\alpha)$$

$$P\left(\bar{x} - Z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}} \le \mu_{\bar{x}} \le \bar{x} + Z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\right) = (1-\alpha)$$

Sostituendo i valori dell'esercizio si ha:

1)

$$\alpha = 0.10 \Rightarrow 100 \times (1-\alpha) = 0.90$$

$$P\left(2310 - 1.645\frac{89}{\sqrt{100}} \le \mu_{\bar{x}} \le 2310 + 1.645\frac{89}{\sqrt{100}}\right) = 0.90$$

$$P\left(2310 - 14.64 \le \mu_{\bar{x}} \le 2310 + 14.64\right) = 0.90$$

$$P\left(2295.36 \le \mu_{\bar{x}} \le 2324.64\right) = 0.90$$

2)

$$\alpha = 0.05 \Rightarrow 100 \times (1-\alpha) = 0.95$$

$$P\left(2310 - 1.96\frac{89}{\sqrt{100}} \le \mu_{\bar{x}} \le 2310 + 1.96\frac{89}{\sqrt{100}}\right) = 0.95$$

$$P\left(2310 - 17.44 \le \mu_{\bar{x}} \le 2310 + 17.44\right) = 0.95$$

$P(2292.56 \le \mu_{\bar{x}} \le 2327.44) = 0.95$

3)
$\alpha = 0.01 \Rightarrow 100 \times (1-\alpha) = 0.99$

$$P\left(2310 - 2.576\frac{89}{\sqrt{100}} \le \mu_{\bar{x}} \le 2310 + 2.576\frac{89}{\sqrt{100}}\right) = 0.99$$
$$P(2310 - 22.92 \le \mu_{\bar{x}} \le 2310 + 22.92) = 0.99$$
$$P(2287.08 \le \mu_{\bar{x}} \le 2332.92) = 0.99$$

4)
$$P\left(\bar{x} - Z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}} \le \mu_{\bar{x}} \le \bar{x} + Z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\right) = (1-\alpha)$$
In questo caso si vuole calcolare la numerosità campionaria in grado di assicurare un livello di precisione nella stima fissato. Risolvendo rispetto a n si ha:
$$Z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}} = \varepsilon \Rightarrow Z_{\frac{\alpha}{2}}\sigma = \varepsilon\sqrt{n} \Rightarrow \sqrt{n} = \frac{Z_{\frac{\alpha}{2}}\sigma}{\varepsilon} \Rightarrow n = \frac{Z^2_{\frac{\alpha}{2}}\sigma^2}{\varepsilon^2}$$
Sostituendo i valori, ricordando che per un intervallo al 90% la Z vale 1.96, si ha:
$$n = \frac{Z^2_{\frac{\alpha}{2}}\sigma^2}{\varepsilon^2} = \frac{1.96^2 \times 89^2}{5^2} = 857.22 \approx 858$$

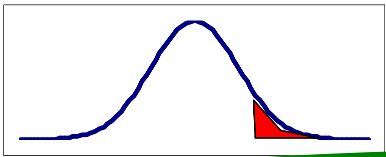
ESERCIZIO 5.4

Si ripeta l'esercizio 5.3 nell'ipotesi di non avere informazioni plausibili sulla varianza della popolazione. Si usi a tal fine il valore di s=89 misurato sul campione di 100 clienti considerato.

SVOLGIMENTO

n=100 $\bar{x}=2310$ $s=89$

In questo caso σ non è noto e invece di usare la distribuzione Z si dovrebbe usare la distribuzione t. Il campione in questione è però di numerosità elevata e quindi non ci sono sensibili variazioni, come si può notare dalle seguenti tabelle che riportano i valori della distribuzione Z e della distribuzione t per i valori di α più comunemente utilizzati:



α

	0.100	0.050	0.025	0.020	0.010	0.005	0.0025	0.001	0.0005
1	3.078	6.314	12.706	15.894	31.821	63.656	127.321	318.289	636.578
2	1.886	2.920	4.303	4.849	6.965	9.925	14.089	22.328	31.600
3	1.638	2.353	3.182	3.482	4.541	5.841	7.453	10.214	12.924
...	...	...	...	...	...	...	...	...	...
100	1.290	1.660	1.984	2.081	2.364	2.626	2.871	3.174	3.390
500	1.283	1.648	1.965	2.059	2.334	2.586	2.820	3.107	3.310
INF	1.282	1.645	1.960	2.054	2.326	2.576	2.807	3.090	3.290

g.d.l

L'ultima riga della tabella riporta i valori della normale e come si vede dalla tabella al crescere dei gradi di libertà i valori della t si avvicinano a quelli della normale.

L'esercizio viene svolto di seguito in maniera da confrontare quanto ottenuto utilizzando l'una o l'altra delle distribuzioni.

$$P\left(-t_{n-1,\frac{\alpha}{2}} \le \frac{\bar{x}-\mu_{\bar{x}}}{s_{\bar{x}}} \le +t_{n-1,\frac{\alpha}{2}}\right)=(1-\alpha)$$

$$P\left(\bar{x}-t_{n-1,\frac{\alpha}{2}}s_{\bar{x}} \le \mu_{\bar{x}} \le \bar{x}+t_{n-1,\frac{\alpha}{2}}s_{\bar{x}}\right)=(1-\alpha)$$

$$P\left(\bar{x}-t_{n-1,\frac{\alpha}{2}}\frac{s}{\sqrt{n}} \le \mu_{\bar{x}} \le \bar{x}+t_{n-1,\frac{\alpha}{2}}\frac{s}{\sqrt{n}}\right)=(1-\alpha)$$

Sostituendo i valori dell'esercizio si ha:

1)

$$\alpha=0.10 \Rightarrow 100\times(1-\alpha)=0.90$$

$$P\left(2310-1.660\frac{89}{\sqrt{100}} \le \mu_{\bar{x}} \le 2310+1.660\frac{89}{\sqrt{100}}\right)=0.90$$

$$P\left(2310-14.78 \le \mu_{\bar{x}} \le 2310+14.78\right)=0.90$$

$$P\left(2295.22 \le \mu_{\bar{x}} \le 2324.78\right)=0.90$$

2)

$$\alpha = 0.05 \Rightarrow 100 \times (1-\alpha) = 0.95$$

$$P\left(2310 - 1.984\frac{89}{\sqrt{100}} \le \mu_{\bar{x}} \le 2310 + 1.984\frac{89}{\sqrt{100}}\right) = 0.95$$

$$P\left(2310 - 17.66 \le \mu_{\bar{x}} \le 2310 + 17.66\right) = 0.95$$

$$P\left(2292.34 \le \mu_{\bar{x}} \le 2327.66\right) = 0.95$$

3)

$$\alpha = 0.01 \Rightarrow 100 \times (1-\alpha) = 0.99$$

$$P\left(2310 - 2.626\frac{89}{\sqrt{100}} \le \mu_{\bar{x}} \le 2310 + 2.626\frac{89}{\sqrt{100}}\right) = 0.99$$

$$P\left(2310 - 23.37 \le \mu_{\bar{x}} \le 2310 + 23.37\right) = 0.99$$

$$P\left(2286.63 \le \mu_{\bar{x}} \le 2333.37\right) = 0.99$$

4)

$$P\left(\bar{x} - t_{n-1,\frac{\alpha}{2}}\frac{s}{\sqrt{n}} \le \mu_{\bar{x}} \le \bar{x} + t_{n-1,\frac{\alpha}{2}}\frac{s}{\sqrt{n}}\right) = (1-\alpha)$$

In questo caso si vuole calcolare la numerosità campionaria in grado di assicurare un livello di precisione nella stima fissato. Risolvendo rispetto a n si ha:

$$t_{n-1,\frac{\alpha}{2}}\frac{s}{\sqrt{n}} = \varepsilon \Rightarrow t_{n-1,\frac{\alpha}{2}}s = \varepsilon\sqrt{n} \Rightarrow \sqrt{n} = \frac{t_{n-1,\frac{\alpha}{2}}s}{\varepsilon} \Rightarrow n = \frac{t^2_{n-1,\frac{\alpha}{2}}s^2}{\varepsilon^2}$$

Sostituendo i valori, ricordando che per un intervallo al 90% la t vale 1.984, si ha:

$$n = \frac{t^2_{\frac{\alpha}{2}}s^2}{\varepsilon^2} = \frac{1.984^2 \times 89^2}{5^2} = 873.33 \approx 874$$

ESERCIZIO 5.5

Una compagnia aerea vuole stimare il numero medio di posti vuoti per volo per l'anno precedente . A tal fine seleziona casualmente 225 voli e ne esamina le registrazioni. Da queste risulta una media campionaria pari a 11.6 posti e uno scarto quadratico medio pari a 4.1 posti.

Si stimi μ, numero medio di posti vuoti per volo durante lo scorso anno, usando:

1) un intervallo di confidenza al 90%
2) un intervallo di confidenza al 95%
3) un intervallo di confidenza al 99%
4) un intervallo di confidenza all'80%

SVOLGIMENTO

n=225 $\quad \bar{x} = 11.6 \quad \sigma = 4.1$

Per risolvere i tre quesiti può essere utile riassumere nella seguente tabella i coefficienti di confidenza con i rispettivi valori per la Z più comunemente utilizzati:

100 (1-α)%	α	α/2	$Z_{\alpha/2}$
90%	0.10	0.050	1.645
95%	0.05	0.025	1.960
99%	0.01	0.005	2.576

$$P\left(-Z_{\frac{\alpha}{2}} \le \frac{\bar{x}-\mu_{\bar{x}}}{\sigma_{\bar{x}}} \le +Z_{\frac{\alpha}{2}}\right) = (1-\alpha)$$

$$P\left(\bar{x}-Z_{\frac{\alpha}{2}}\sigma_{\bar{x}} \le \mu_{\bar{x}} \le \bar{x}+Z_{\frac{\alpha}{2}}\sigma_{\bar{x}}\right) = (1-\alpha)$$

$$P\left(\bar{x}-Z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}} \le \mu_{\bar{x}} \le \bar{x}+Z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\right) = (1-\alpha)$$

Sostituendo i valori dell'esercizio si ha:

1)

$$\alpha = 0.10 \Rightarrow 100 \times (1-\alpha) = 0.90$$

$$P\left(11.6 - 1.645\frac{4.1}{\sqrt{225}} \le \mu_{\bar{x}} \le 11.6 + 1.645\frac{4.1}{\sqrt{225}}\right) = 0.90$$

$$P(11.6 - 0.45 \le \mu_{\bar{x}} \le 11.6 + 0.45) = 0.90$$

$$P(11.15 \le \mu_{\bar{x}} \le 12.05) = 0.90$$

2)

$$\alpha = 0.05 \Rightarrow 100 \times (1-\alpha) = 0.95$$

$$P\left(11.6 - 1.96\frac{4.1}{\sqrt{225}} \le \mu_{\bar{x}} \le 11.6 + 1.96\frac{4.1}{\sqrt{225}}\right) = 0.95$$

$$P(11.6 - 0.54 \le \mu_{\bar{x}} \le 11.6 + 0.54) = 0.95$$

$$P(11.06 \le \mu_{\bar{x}} \le 12.14) = 0.95$$

3)

$\alpha = 0.01 \Rightarrow 100 \times (1-\alpha) = 0.99$

$$P\left(11.6 - 2.576\frac{4.1}{\sqrt{225}} \leq \mu_{\bar{x}} \leq 11.6 + 2.576\frac{4.1}{\sqrt{225}}\right) = 0.99$$

$$P\left(11.6 - 0.70 \leq \mu_{\bar{x}} \leq 11.6 + 0.70\right) = 0.99$$

$$P\left(10.90 \leq \mu_{\bar{x}} \leq 12.30\right) = 0.99$$

4)

$\alpha = 0.20 \Rightarrow 100 \times (1-\alpha) = 0.80$

$$P\left(11.6 - 1.282\frac{4.1}{\sqrt{225}} \leq \mu_{\bar{x}} \leq 11.6 + 1.282\frac{4.1}{\sqrt{225}}\right) = 0.80$$

$$P\left(11.6 - 0.35 \leq \mu_{\bar{x}} \leq 11.6 + 0.35\right) = 0.80$$

$$P\left(11.25 \leq \mu_{\bar{x}} \leq 11.95\right) = 0.80$$

ESERCIZIO 5.6

Un'azienda farmaceutica deve stimare la variazione media della pressione sanguigna nei pazienti che assumono un nuovo tipo di farmaco. Nella prima fase di test del farmaco la normativa consente all'azienda l'utilizzo di soli sei pazienti.
Si supponga che i sei pazienti selezionati a caso abbiano i seguenti valori di variazione nella pressione sanguigna: 1.7, 3.0, 0.8, 3.4, 2.7 e 2.1 punti.

1) Che tipo di ipotesi è necessaria per la costruzione dell'intervallo di confidenza?
2) Si costruisca un intervallo di confidenza al 95% per μ, l'incremento medio nella pressione sanguigna associato con il nuovo farmaco.
3) Verificare come cambia l'intervallo se il campione viene allargato a 8 e a 10 pazienti (nell'ipotesi che i valori di media e scarto rimangano invariati)

SVOLGIMENTO

1)
Per costruire un intervallo di confidenza, è necessario ipotizzare che la pressione sanguigna si distribuisca secondo una legge normale (non è possibile utilizzare il teorema del limite centrale vista la bassa numerosità del campione).

2)
n=6 $\bar{x} = 2.28$ $s = 0.95$

$$P\left(\bar{x} - t_{n-1,\frac{\alpha}{2}} \frac{s}{\sqrt{n}} \leq \mu_{\bar{x}} \leq \bar{x} + t_{n-1,\frac{\alpha}{2}} \frac{s}{\sqrt{n}}\right) = (1-\alpha)$$

Il valore delle tavole per la distribuzione t con a/2=0.025 e 5 g.d.l. è 2.776.
Sostituendo i valori nella formula si ha:

$$P\left(2.28 - 2.776\frac{0.95}{\sqrt{6}} \leq \mu_{\bar{x}} \leq 2.28 + 2.776\frac{0.95}{\sqrt{6}}\right) = 0.95$$

$$P\left(2.28 - 1 \leq \mu_{\bar{x}} \leq 2.28 + 1\right) = 0.95$$

$$P\left(1.29 \leq \mu_{\bar{x}} \leq 3.28\right) = 0.95$$

3)
n=8 $\bar{x} = 2.28$ $s = 0.95$

$$P\left(2.28 - 2.776\frac{0.95}{\sqrt{8}} \leq \mu_{\bar{x}} \leq 2.28 + 2.776\frac{0.95}{\sqrt{8}}\right) = (1-\alpha)$$

$$P\left(2.28 - 0.79 \leq \mu_{\bar{x}} \leq 2.28 + 0.79\right) = 0.95$$

$$P\left(1.49 \leq \mu_{\bar{x}} \leq 3.08\right) = 0.95$$

n=10 $\bar{x} = 2.28$ $s = 0.95$

$$P\left(2.28 - 2.776\frac{0.95}{\sqrt{10}} \leq \mu_{\bar{x}} \leq 2.28 + 2.776\frac{0.95}{\sqrt{10}}\right) = 0.95$$

$$P\left(2.28 - 0.68 \le \mu_{\bar{x}} \le 2.28 + 0.68\right) = 0.95$$
$$P\left(1.60 \le \mu_{\bar{x}} \le 2.96\right) = 0.95$$

ESERCIZIO 5.7

Un fornitore all'ingrosso di gas usa un nuovo macchinario per riempire le bombole che distribuisce. Quando la macchina è ben calibrata i litri medi sono pari a 27, ma fattori non controllabili provocano una variazione nel carico che produce uno scarto quadratico medio pari a 3, secondo una legge normale. Per superare un controllo di certificazione di qualità il produttore desidera stimare la quantità media di carico con un margine di 1 dal suo reale valore usando un intervallo di confidenza al 99%. Quante bombole è necessario analizzare per costruire il campione?

SVOLGIMENTO

$$P\left(\overline{x} - Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \leq \mu_{\overline{x}} \leq \overline{x} + Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}\right) = (1-\alpha)$$

In questo caso si vuole calcolare la numerosità campionaria in grado di assicurare un livello di precisione nella stima fissato. Risolvendo rispetto a n si ha:

$$Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} = \varepsilon \Rightarrow Z_{\frac{\alpha}{2}} \sigma = \varepsilon\sqrt{n} \Rightarrow \sqrt{n} = \frac{Z_{\frac{\alpha}{2}}\sigma}{\varepsilon} \Rightarrow n = \frac{Z_{\frac{\alpha}{2}}^2 \sigma^2}{\varepsilon^2}$$

Sostituendo i valori, ricordando che per un intervallo al 99% la Z vale 2.576, si ha:

$$n = \frac{Z_{\frac{\alpha}{2}}^2 \sigma^2}{\varepsilon^2} = \frac{2.576^2 \times 3^2}{1^2} = 59.71431 \approx 60$$