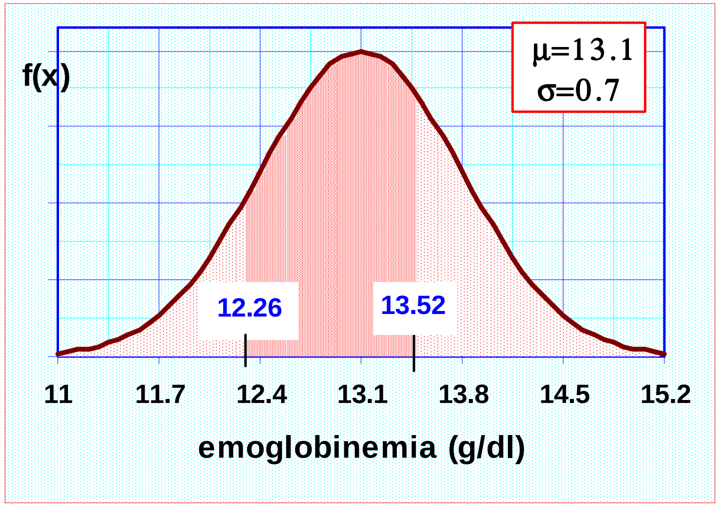
LA DISTRIBUZIONE NORMALE

ESERCITAZIONE

Esempio di applicazione della deviata gaussiana standard

In una popolazione di ragazze di età inclusa tra i 18 e i 25 anni, la concentrazione di emoglobina nel sangue (x) approssima una distribuzione gaussiana con media = 13.1 g/dl e deviazione standard = 0.7 g/dl. In base a queste sole informazioni possiamo calcolare, *ad esempio*, quante ragazze hanno emoglobina inclusa tra 12.26 e 13.52 g/dl.

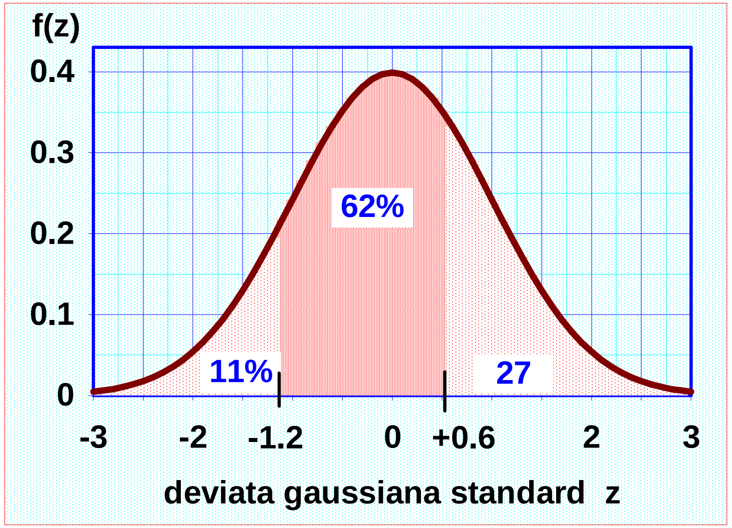
Distribuzione dell'emoglobina in una popolazione di ragazze di età compresa tra i 18 e i 25 anni.



… continua

Infatti: z_1 = (12.26-13.10) / 0.7 = -1.2

z_2 = (13.52-13.10) / 0.7 = +0.6



Nell'11% delle ragazze i valori di Hb sono minori di 12.26 g/dl, e nel 27% sono maggiori di 13.52 g/dl. Quindi il 62% delle ragazze ha valori di Hb compresi tra 12.26 e 13.52 g/dl.

... continua

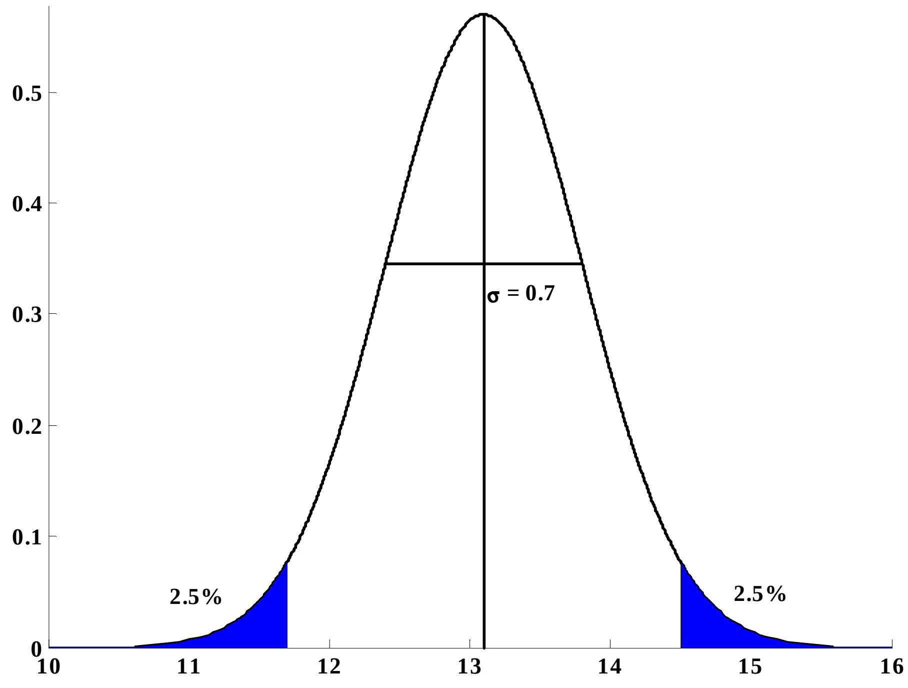
Quali sono i valori che racchiudono il 95% delle osservazioni, che considero come i valori entro cui è compreso il range di normalità?

$$z = \frac{x-\mu}{\sigma} \Rightarrow x = z \cdot \sigma + \mu \quad con\ z_{0.025} = 1.96$$

$$x_{1,2} = \mu \pm z_{0.025} \cdot \sigma$$

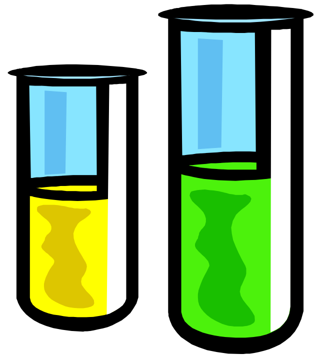
$$x_1 = 13.1 - 1.96 \cdot 0.7 = 11.728$$

$$x_2 = 13.1 + 1.96 \cdot 0.7 = 14.472$$



Esercizio 1

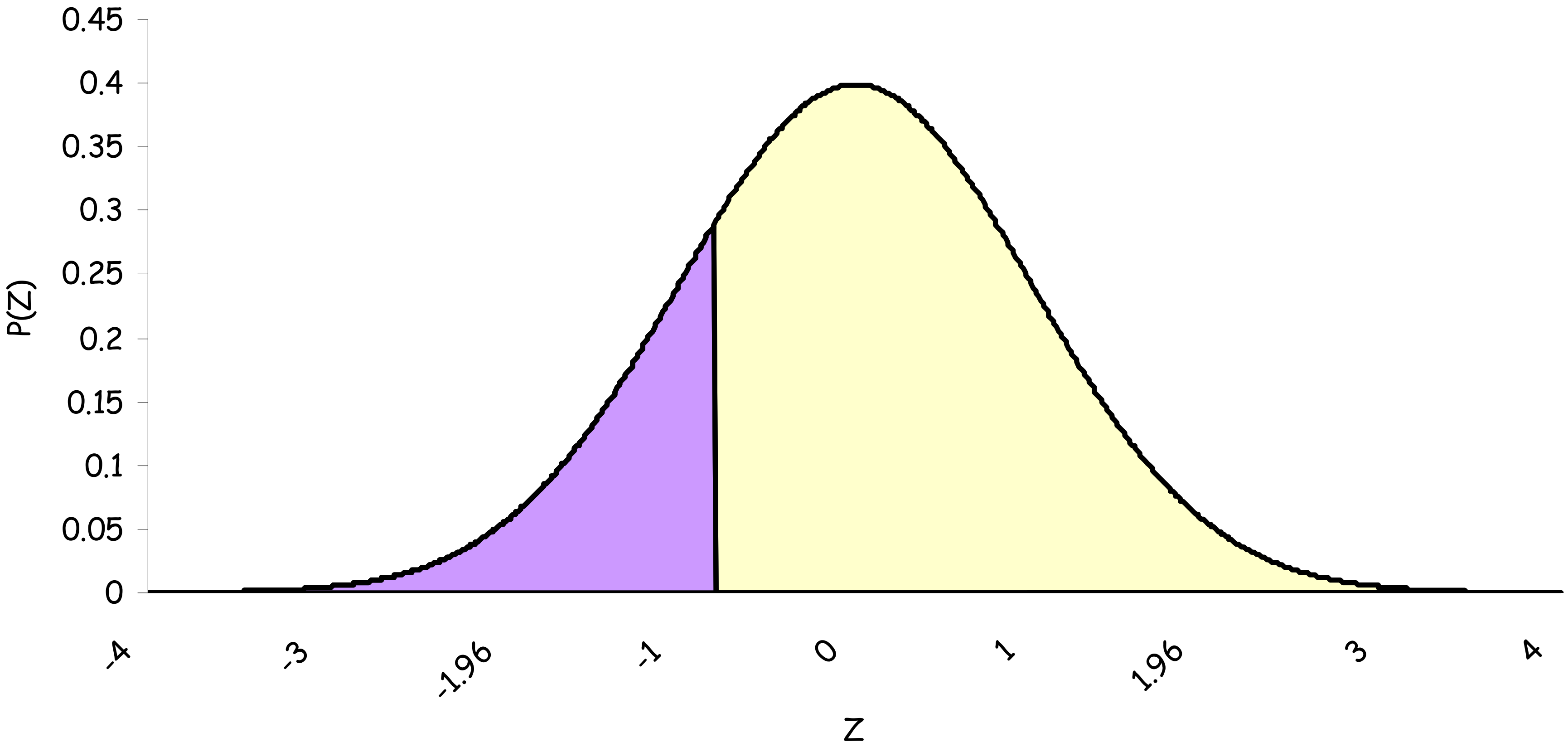
Se si suppone che, nella popolazione degli adulti, il livello di acido urico (mg/100 ml) segua una distribuzione gaussiana con media e d.s. rispettivamente pari a 5.7 e 1 (mg/100ml), si trovi la probabilità che un soggetto scelto a caso da questa popolazione abbia un livello di acido urico:



1. Minore di 4.9 mg/100ml
2. Compreso tra 4.9 e 6.2 mg/100ml
3. Trovare inoltre il valore di acido urico x tale per cui $P(X \geq x)=0.40$

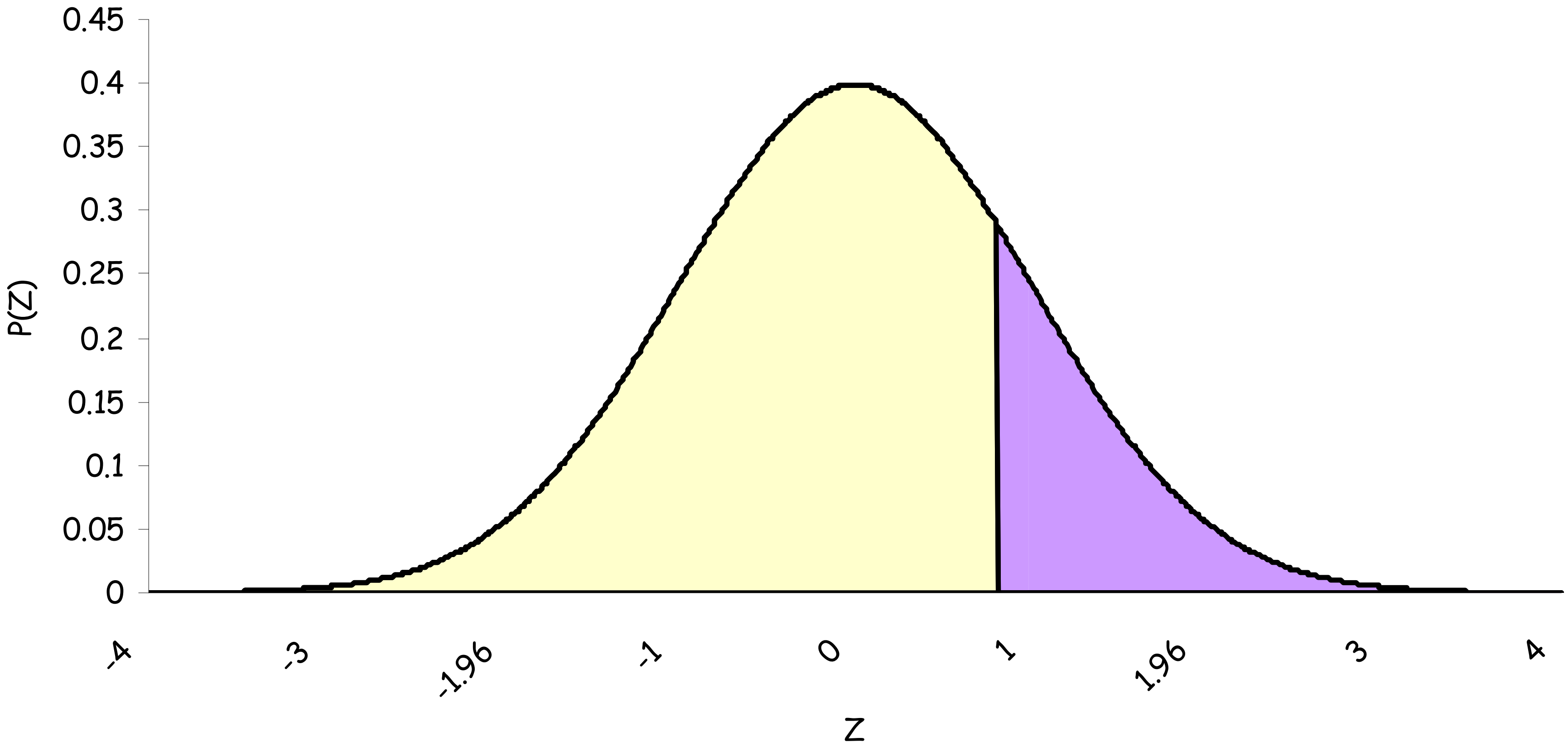
Risposte

1. $P(X<4.9) = P[\underbrace{(X-5.7)/1}_{Z} < (4.9-5.7)/1] =$
   $P(Z<-0.8) = \ldots$



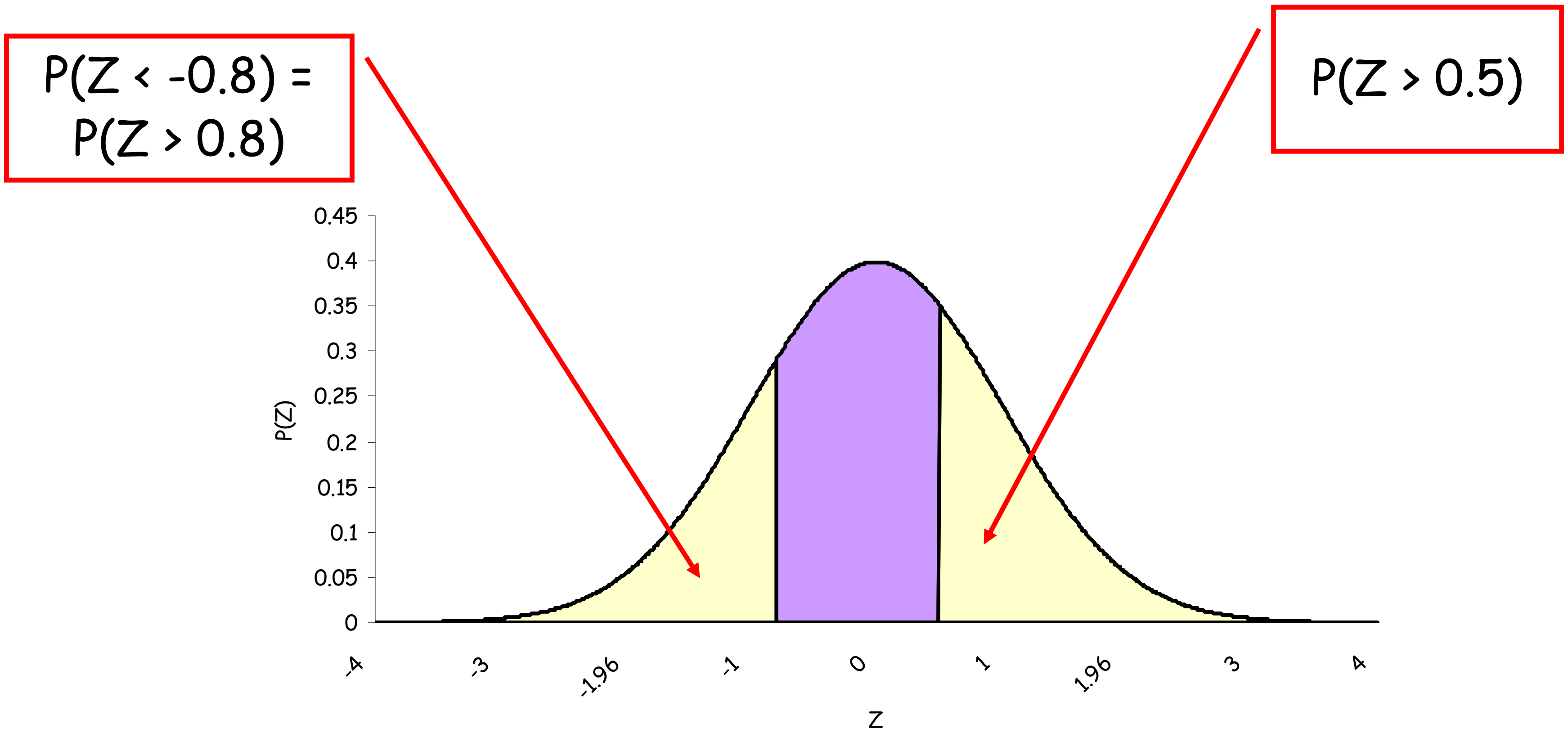
Risposte

1. … = P(Z>0.8) = 0.212



Risposte

2. $P(4.9<X<6.2) = P(-0.8<Z<0.5) = 1-P(Z>0.8)-P(Z>0.5) =$
   $1-0.212-0.308 = 0.479$

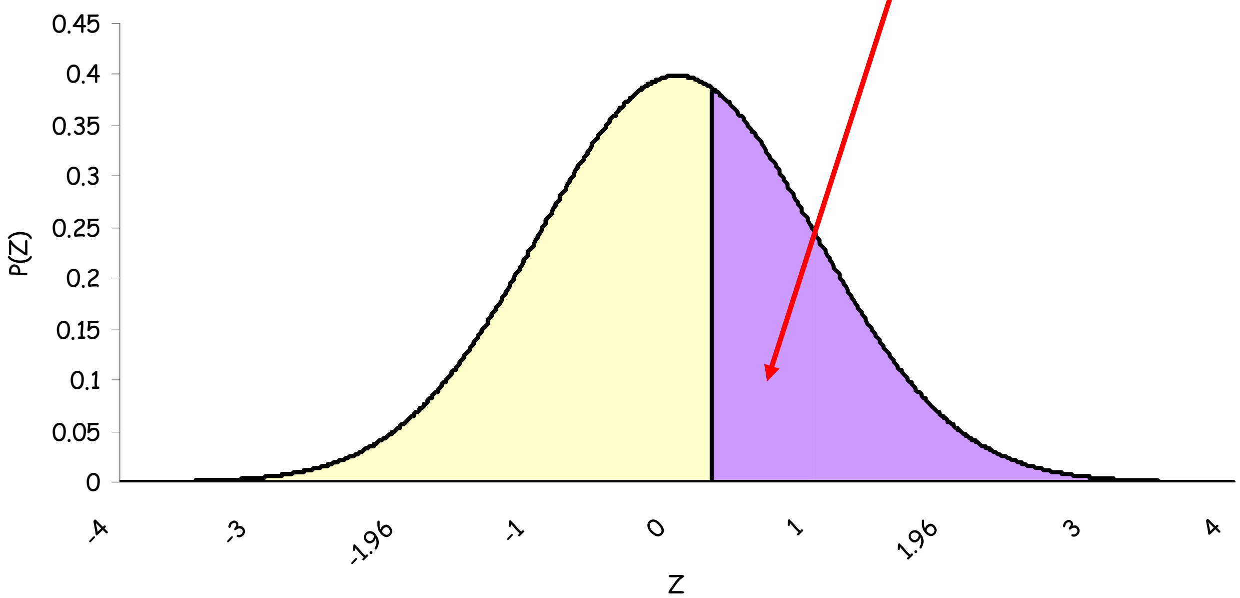


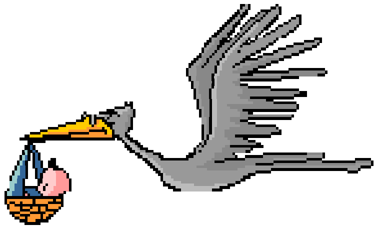
Risposte

3. $P(Z > z_{0.40}) = 0.40 \Rightarrow z_{0.40} = 0.25$

$$\frac{x - 5.7}{1} = 0.25 \Rightarrow x = 5.95$$

$P(Z>z)=0.40$





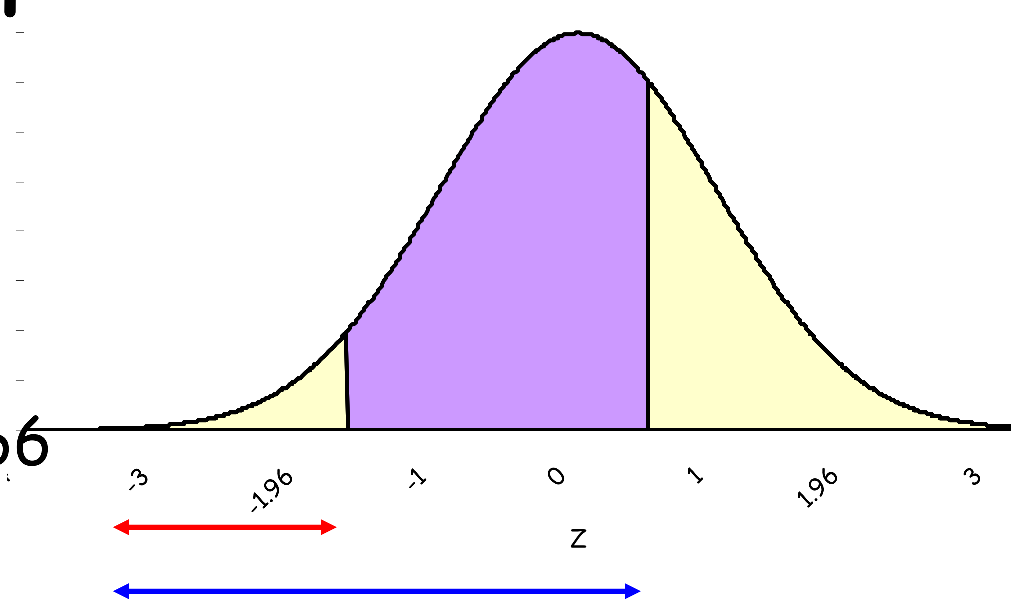
Esercizio 2

Da un'indagine svolta su di un campione di neonati, risulta che la distribuzione dei loro pesi alla nascita è normale con media 3.2 e con σ di 0.6 Kg.

1. Qual è il valore della normale standardizzata (Z) per un peso di 4 Kg?
2. Quale percentuale di neonati presenta un peso alla nascita compreso fra 2.2 e 3.5 Kg?
3. Qual è il peso oltre il quale si trovano il 10% dei valori più elevati?

Risposte

1. $z = \dfrac{x - \mu}{\sigma} = \dfrac{4 - 3.2}{0.6} = 1.33$



2. P(2.2<X<3.5)

   $z_1 = \dfrac{x - \mu}{\sigma} = \dfrac{2.2 - 3.2}{0.6} = -1.66$

   $z_2 = \dfrac{x - \mu}{\sigma} = \dfrac{3.5 - 3.2}{0.6} = 0.5$

   $P(X < 2.2) = P(Z < -1.66) = P(Z > 1.66) = 0.04846$

   $P(X < 3.5) = P(Z < 0.5) = 1 - P(Z > 0.5) = 1 - 0.3085 = 0.6915$

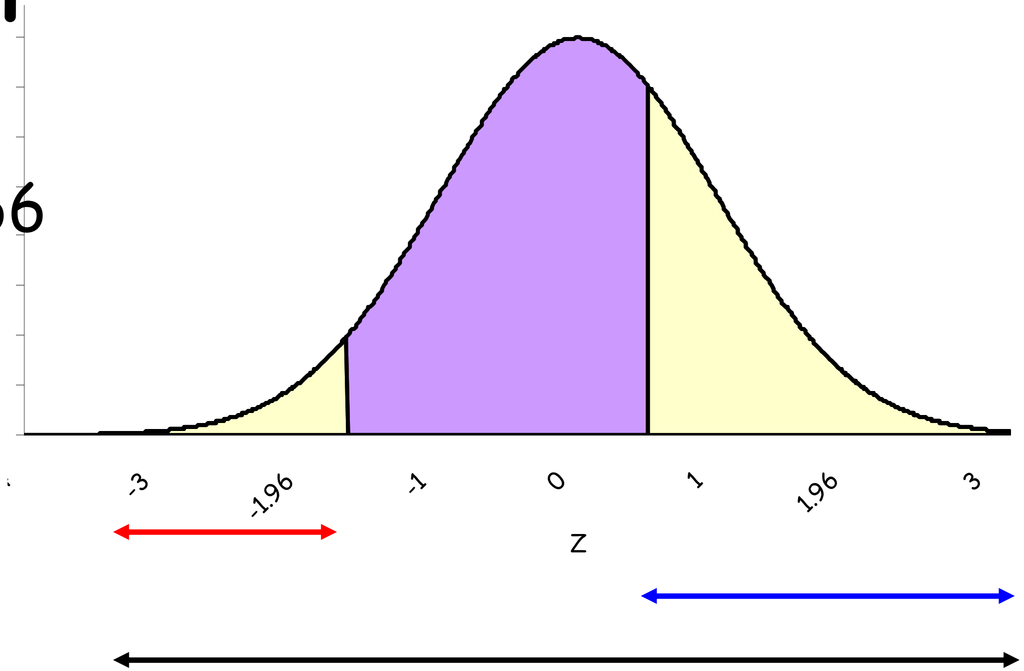
   $P(2.2 < X < 3.5) = P(-1.66 < Z < 0.5) = P(Z < 0.5) - P(Z < -1.66)$
   $= 0.6915 - 0.04846 = 0.64$

Risposte

2. P(2.2<X<3.5)

$$z_1 = \frac{x - \mu}{\sigma} = \frac{2.2 - 3.2}{0.6} = -1.66$$

$$z_2 = \frac{x - \mu}{\sigma} = \frac{3.5 - 3.2}{0.6} = 0.5$$



$P(X < 2.2) = P(Z < -1.66) = P(Z > 1.66) = 0.04846$

$P(X > 3.5) = P(Z > 0.5) = 0.3085$

$P(2.2 < X < 3.5) = P(-1.66 < Z < 0.5) =$

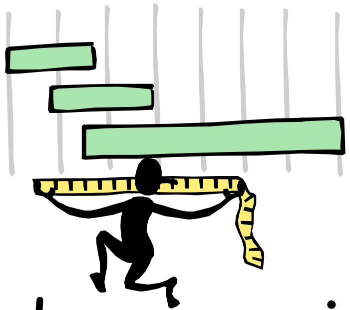
$= 1- P(Z > 0.5)-P(Z < -1.66) = 1 - 0.3085 - 0.04846 = 0.64$

Risposte

3. $P(Z>z) = 10\% = 0.10 \Rightarrow z_{0.10} = 1.28$

$$\frac{x - \mu}{\sigma} = 1.28 \Rightarrow$$

$$\Rightarrow x = 1.28 \times \sigma + \mu = 1.28 \times 0.6 + 3.2 = 3.96$$



Esercizio 3

Ad una visita di controllo, la statura di un bambino di 8 anni è risultata di 130 cm. Da un'indagine condotta su una popolazione di 10000 bambini di pari età (8 anni) è risultato che il 10% ha una statura superiore a 133 cm e il 3% una statura inferiore a 115 cm.

a) Assumete che la distribuzione delle stature sia approssimativamente gaussiana e calcolate media e varianza della statura nella popolazione.

b) Qual è la probabilità di osservare un soggetto con statura inferiore a quella del bambino in esame (130 cm)?

Risposte

a) Poiché:

$P(X > 133) = P(Z > z_1) = 0.1$ e

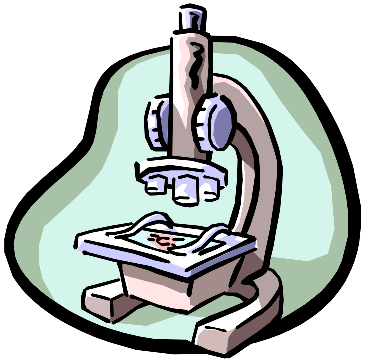
$P(X < 115) = P(Z < z_2) = 0.03$

dalle tavole della distribuzione gaussiana standardizzata si ricava facilmente che $z_1 = 1.28$ e $z_2 = -1.88 \Rightarrow$

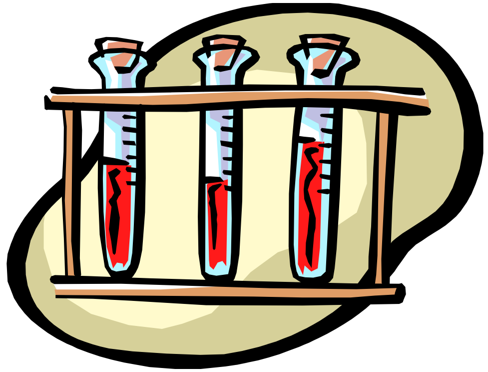
$$\begin{cases} z_1 = \dfrac{x - \mu}{\sigma} = \dfrac{133 - \mu}{\sigma} = 1.28 \\ z_2 = \dfrac{x - \mu}{\sigma} = \dfrac{115 - \mu}{\sigma} = -1.88 \end{cases} \Longrightarrow \begin{cases} \mu = 125.71 \\ \sigma = 5.70 \end{cases}$$

b) $P(X < 130) = P[Z<(130-125.71)/5.7] = P(Z < 0.75) =$

$= 1 - 0.2266 = 0.7734$



Esercizio 4



Dall'esame microscopico dei globuli rossi di un paziente affetto da Plasmodium vivax della malaria, è risultato che la **media e la varianza** delle misure del diametro massimo di un **globulo rosso non infettato** sono rispettivamente **7.6** e **0.9** micron, mentre per un **globulo rosso infettato**, la media e la deviazione standard delle misure del diametro massimo sono rispettivamente **9.6** e **1.0** micron. Assumete che i valori riportati siano uguali ai parametri della popolazione e che il diametro massimo dei globuli rossi, infettati e no, sia **distribuito in modo gaussiano** e calcolate:

Quesiti

a) Quale proporzione di globuli rossi non infettati vi aspettate di trovare con un diametro di misura compreso tra 9.4 e 9.6 micron?

b) Quale proporzione di globuli rossi non infettati vi aspettate di trovare con un diametro di misura compreso tra 7.6 e 9.4 micron?

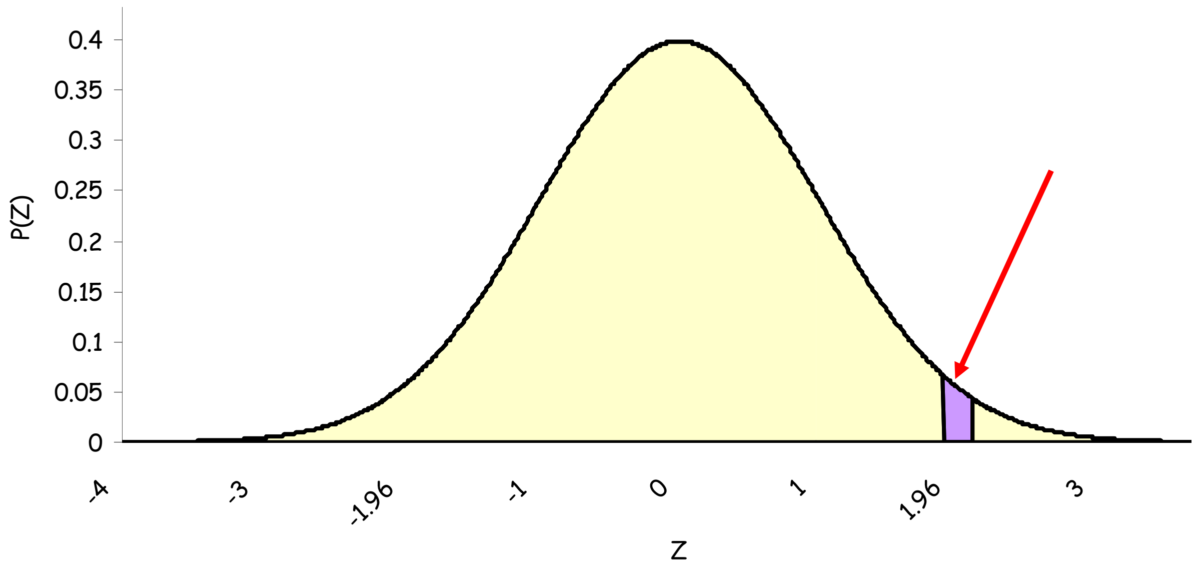
c) Supponete che il 20% dei globuli rossi sia infettato. Quale percentuale di tutti i globuli rossi avrà un diametro superiore a 9.0 micron?

Risposte

a) La proporzione di globuli rossi non infettati con diametro compreso tra 9.4 e 9.6 micron è:

$$P(9.4 < X < 9.6) = P\left[\frac{(9.4-7.6)}{\sqrt{0.9}} < z < \frac{(9.6-7.6)}{\sqrt{0.9}}\right] =$$

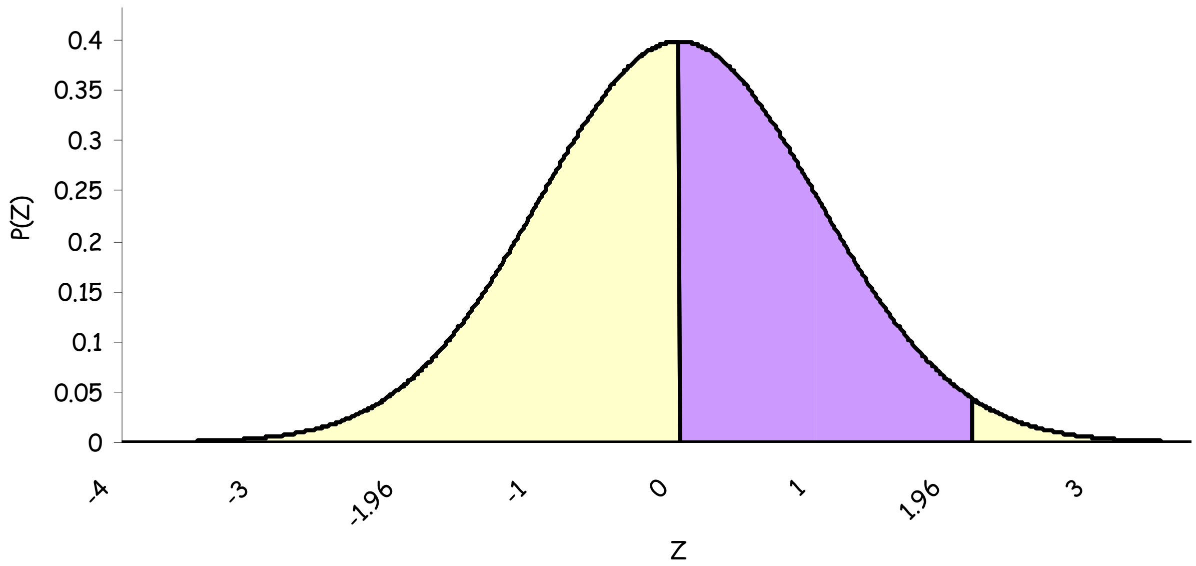
$$= P(1.897 < Z < 2.108) = P(Z > 1.897) - P(Z > 2.109) =$$
$$= 0.028717 - 0.01786 = 0.01086$$



b) La proporzione di globuli rossi non infettati con diametro compreso tra 7.6 e 9.4 micron è:

$$P(9.4 < X < 9.6) = \quad P\left[\frac{(7.6-7.6)}{\sqrt{0.9}} < z < \frac{(9.6-7.6)}{\sqrt{0.9}}\right] =$$

$= P(0 < Z < 1.897) = P(Z > 0) - P(Z > 1.897) =$
$= 0.5 - 0.028717 = 0.471$



c) La proporzione di globuli rossi non infettati (S) con diametro maggiore di 9 micron è:

$$P(X_S > 9) = P\left[Z > \frac{(9-7.6)}{\sqrt{0.9}}\right] = P(Z > 1.4757) = 0.07078$$

La proporzione di globuli rossi infettati (M) con diametro maggiore di 9 micron è:

$$P(X_M > 9) = P\left[Z > \frac{(9-9.6)}{1}\right] = P(Z > -0.6) =$$
$$= 1 - P(Z < -0.6) = 1 - P(Z > 0.6) = 0.72575$$

Se i globuli rossi infettati sono il 20%, la proporzione di tutti i globuli rossi con un diametro maggiore di 9 micron è:

$$P(X > 9) = 0.8 \times 0.07078 + 0.2 \times \mathbf{0.72575} = 0.201774$$

Esercizio 5

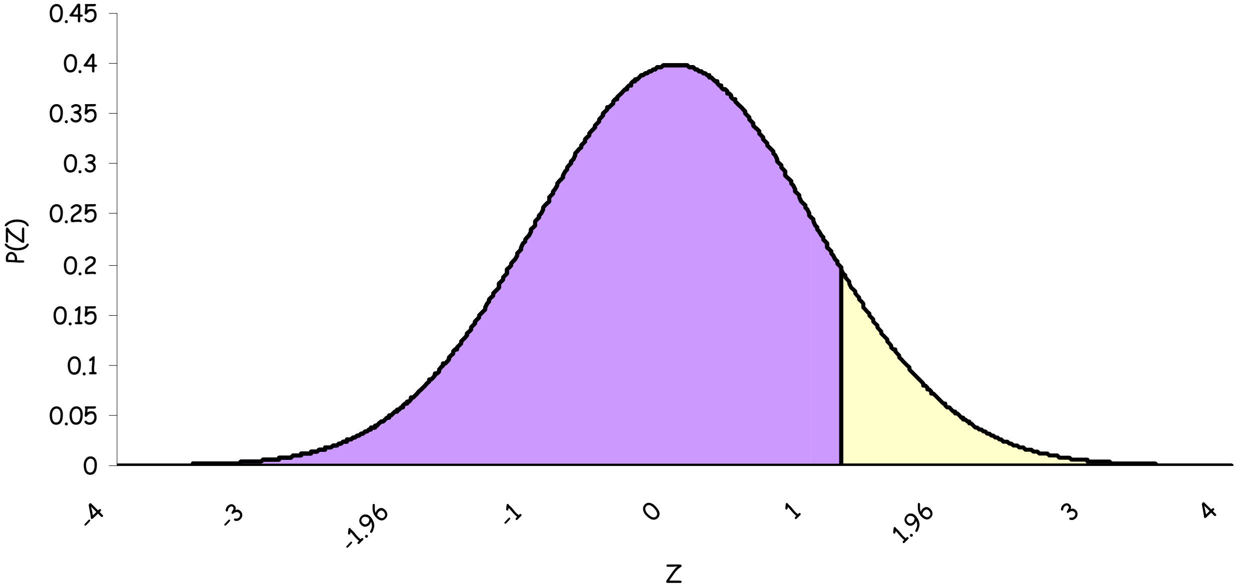
La distribuzione del peso della popolazione di maschi in giovane età è gaussiana con media 67.5 Kg e deviazione standard 2.5 Kg. Se si estraggono a caso 200 soggetti da questa popolazione, quanti avranno:

1. Peso minore di 70.5 Kg?
2. Peso compreso fra 65 e 68 Kg?
3. Qual è il valore di peso superato dal 2.3% della popolazione?

Risposte

1. $P(X < 70.5) = P(Z < 1.2) = 1 - 0.115 = 0.885$

soggetti con peso < 70.5 Kg = 200x0.885 = 177

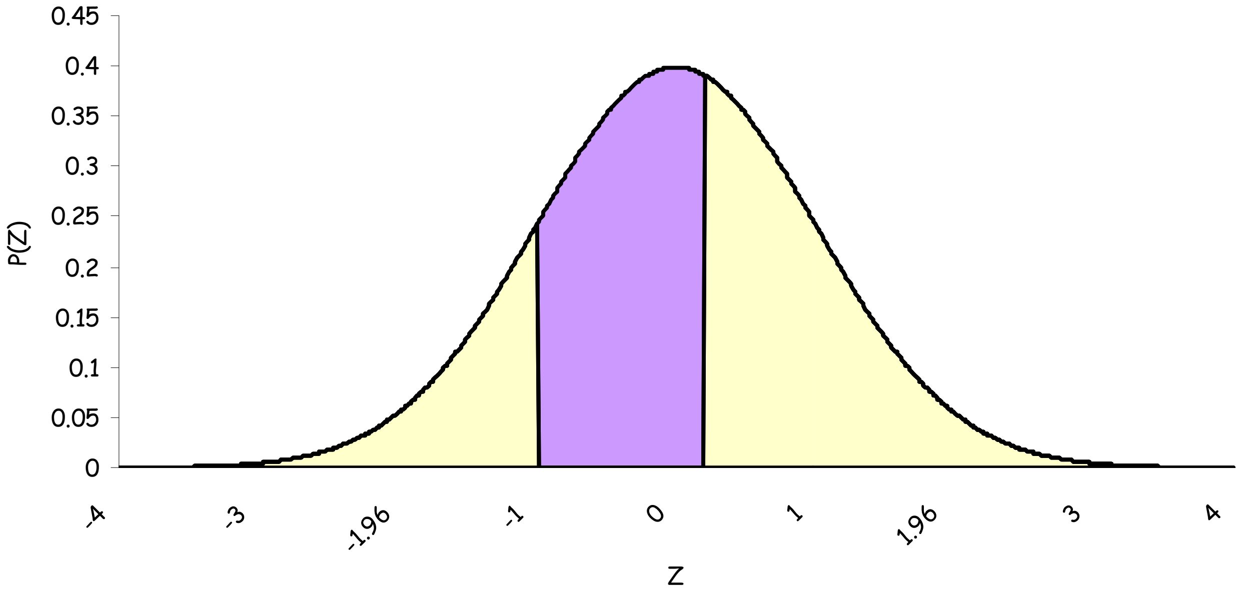


2. $P(65 < X < 68) = P(-1 < Z < 0.2) =$

$= 1 - P(Z > 1) - P(Z > 0.2) = 1 - 0.159 - 0.421 =$

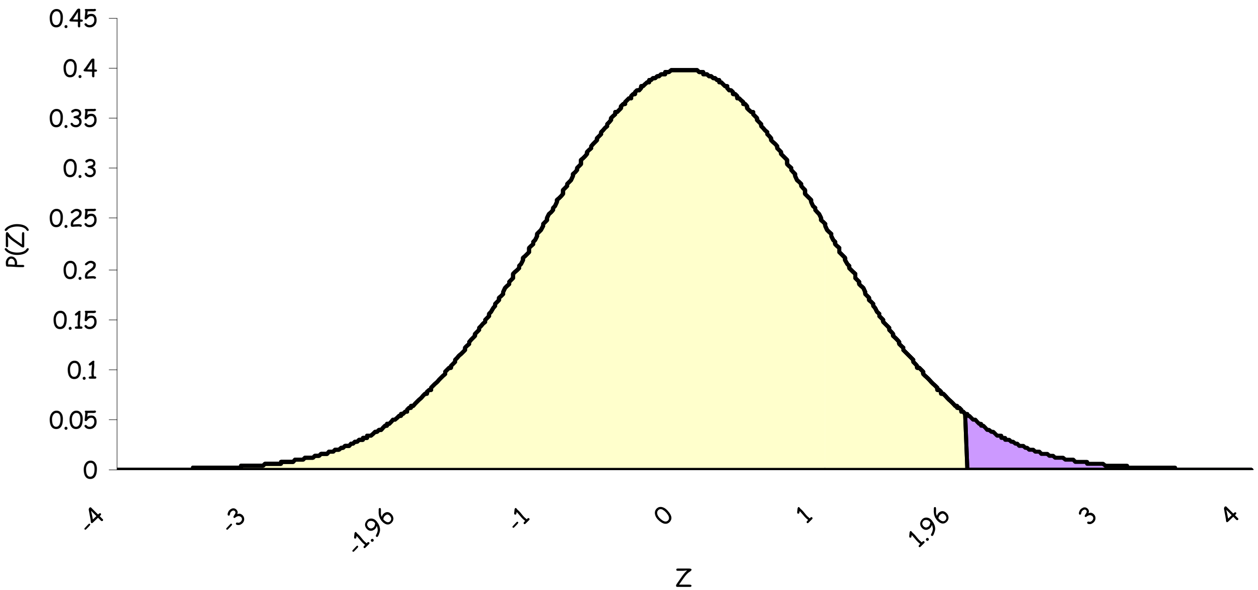
$= 0.42$

\# soggetti con peso compreso tra 65 e 68 Kg
$= 200 \times 0.42 = 84$



3. $P(Z > z_{0.023}) = 0.023 \Rightarrow z_{0.023} = 1.99$

$(x - 67.5) / 2.5 = 1.99 \Rightarrow x = 72.48$



Esercizio 6

I laureati di una certa facoltà hanno una votazione media di 100 con una ds di 4. Supponiamo che la distribuzione dei voti sia normale:

a) Calcolare la percentuale di laureati che ha ottenuto un voto compreso tra 96 e 104

b) Calcolare la percentuale di laureati che ha ottenuto un voto maggiore di 108

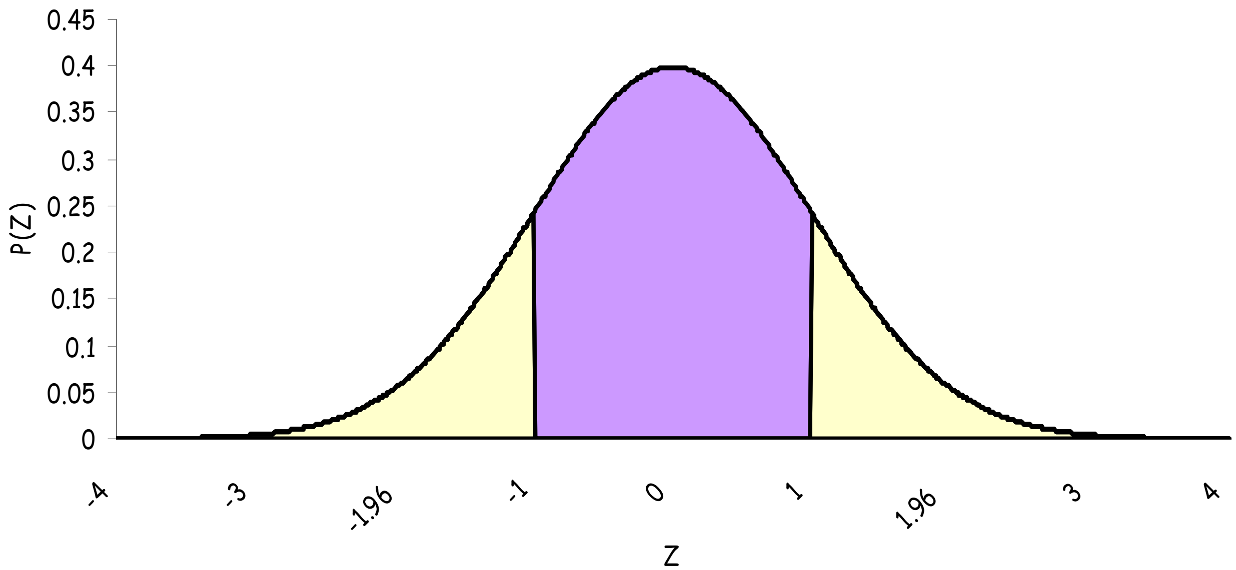
c) Calcolare lo scarto interquartile

Risposte

a) P (96 < L < 104) =

= P [(96-100)/4) < Z < (104-100) /4] = P (-1 < Z < 1) =

= 1-2*0.15866 = 0.6827 = 68.27% dei laureati ha preso un voto compreso tra 96 e 104



Risposte

b) $P(L > 108) =$
$= P(Z > (108-100)/4] = P(Z > 2) = 0.02275 =$
2.275%

c) $P(L < Q_3) = 0.75 =$
$= P[(L - \mu)/\sigma < (Q_3 - \mu)/\sigma] = P(z < z_0) = 0.75$
$z_0 = 0.67 \Rightarrow Q_3 = 0.67 * 4 + 100 = 102.68$

$P(L < Q_1) = 0.25 =$
$= P[(L - \mu)/\sigma < (Q_1 - \mu)/\sigma] = P(z < z_0) = 0.25$
$z_0 = -0.67 \Rightarrow Q_1 = -0.67 * 4 + 100 = 97.32$
$Q_1 - Q_3 = 102.68 - 97.32 = 5.36$

Esercizio 7

Una ditta confeziona scatole di caffé di contenuto medio di 1 Kg, con ds di 6 g. Se la legge impedisce di mettere in commercio col peso dichiarato di 1 Kg confezioni che contengono meno di 985 g, quante confezioni in media, ogni 1000, non potranno essere messe in commercio?

Risposta

$P(X < 985) = P[Z < (985 - 1000)/6] =$

$= P(Z < -2.5) = 0.00621$

confezioni = 0.00621*1000 = 6.21

Esercizio 8

Una macchina produce sbarrette la cui lunghezza è una variabile casuale normale di media μ = 25 cm e ds σ = 0.3 cm

1) Calcolare la probabilità che la lunghezza di una sbarretta differisca dal suo valor medio di almeno 0.5 cm.

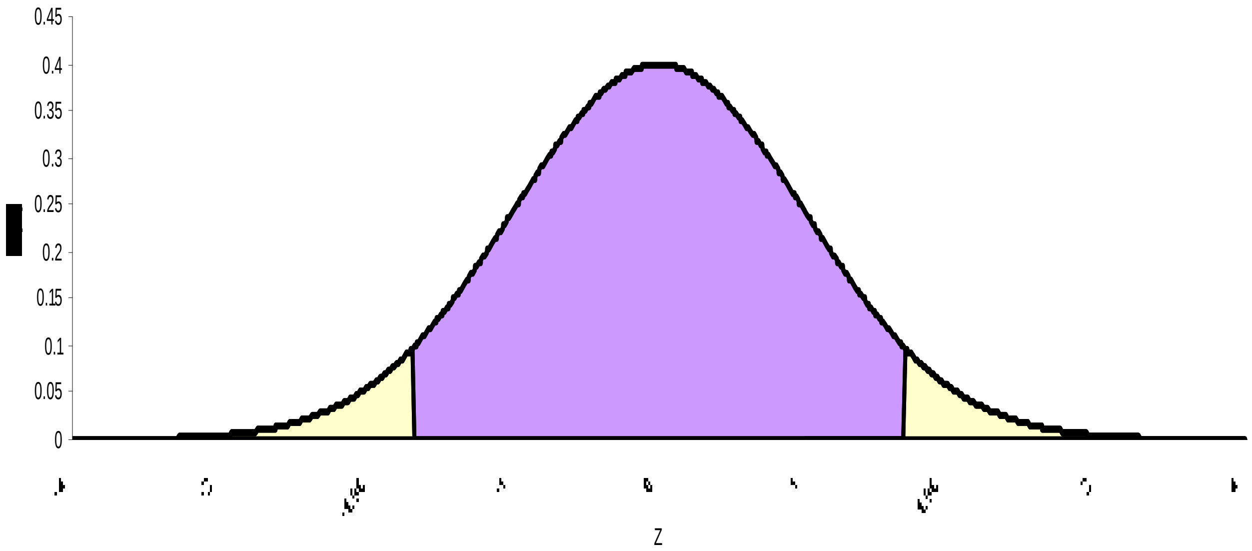
Risposte

1) $P\ (X \leq \mu - 0.5;\ X \geq \mu + 0.5) =$

$= 1- P\ (\mu - 0.5 \leq X \leq \mu + 0.5) =$

$= 1- P(-0.5 / 0.3 \leq Z \leq 0.5 / 0.3) = 1- P(-1.67 \leq Z \leq 1.67)=$

$=1-(1-(2*0.04746)) = 0.09492$



Esercizio 9

La durata delle telefonate urbane segue una distribuzione normale di media μ=10 minuti e σ=3 minuti. Selezionato un campione casuale semplice di 20 telefonate, trovare la distribuzione della media campionaria e la probabilità che la durata media delle telefonate sia compresa tra 9.5 e 10.3 minuti

Risposte

- La distribuzione della media campionaria è:

$$\bar{x} \sim N\left(\mu, \frac{\sigma^2}{n}\right) \Rightarrow \bar{x} \sim N\left(10, \frac{3^2}{20}\right) \Rightarrow \bar{x} \sim N(10{,}0.45)$$

- $P(9.5 < \bar{x} < 10.3) = P(\bar{x} > 9.5) - P(\bar{x} > 10.3) =$

$$= P\left(z > \frac{9.5 - 10}{3/\sqrt{20}}\right) - P\left(z > \frac{10.3 - 10}{3/\sqrt{20}}\right) =$$

$$= P(z > \text{-}0.7454)\text{-}P(z > 0.4472) =$$

$$= 1 - P(z > 0.7454) - P(z > 0.4472) =$$

$$= 1 - 0.22965 - 0.32636 = 0.44399$$

Esercizio 10

Gli occupati di un determinato settore economico vengono pagati con un salario medio di 45 € l'ora e ds pari a 5 €. Assumendo che la distribuzione dei salari possa essere approssimata mediante una distribuzione normale, calcolare:

i) La probabilità che un occupato guadagni meno di 44 € l'ora

Risposte

i) $X \sim N(45, 25) \Rightarrow$

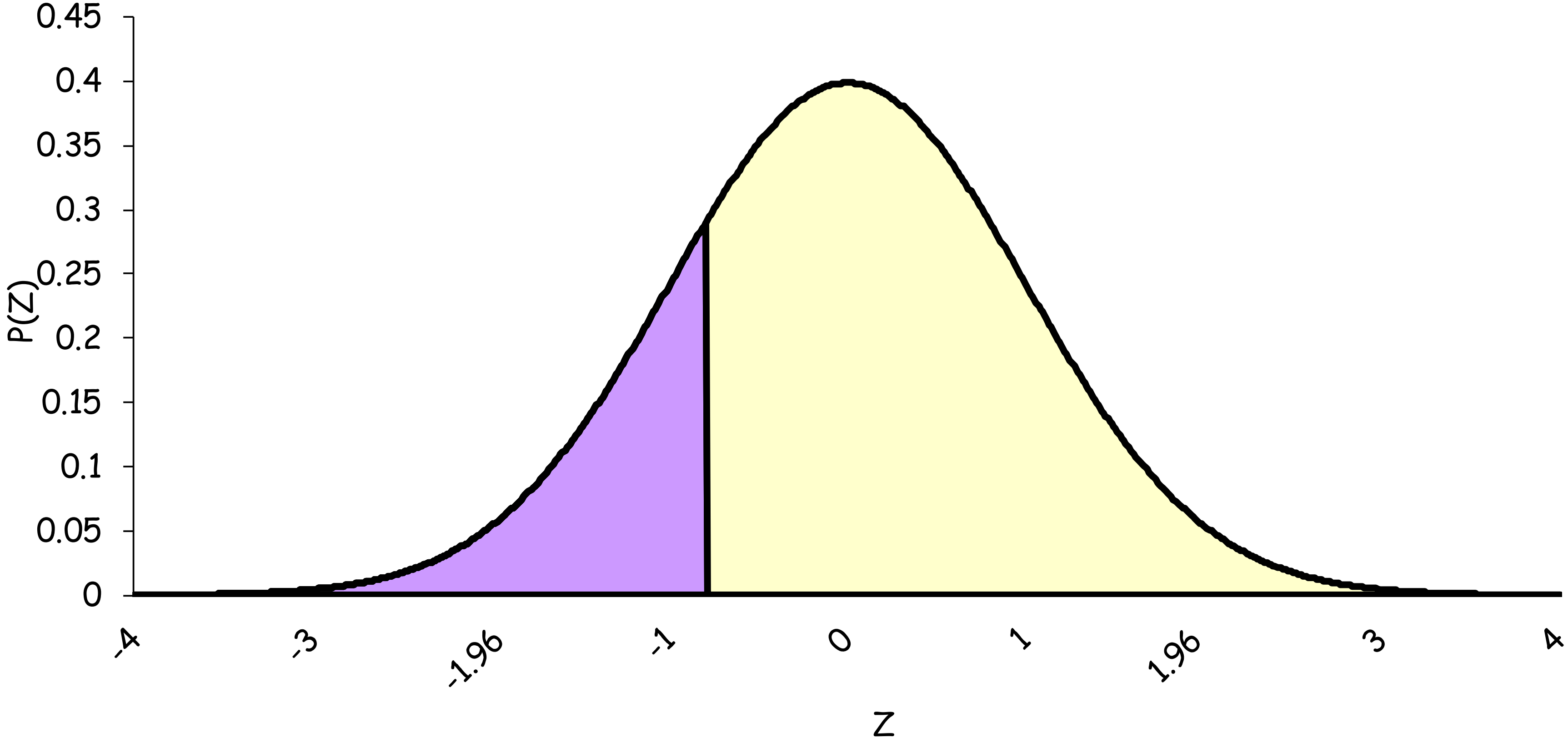
$P(X < 44) = P[Z < (44-45)/5] = P(Z < -0.2) =$

$= P(Z > 0.2) = 0.42074$

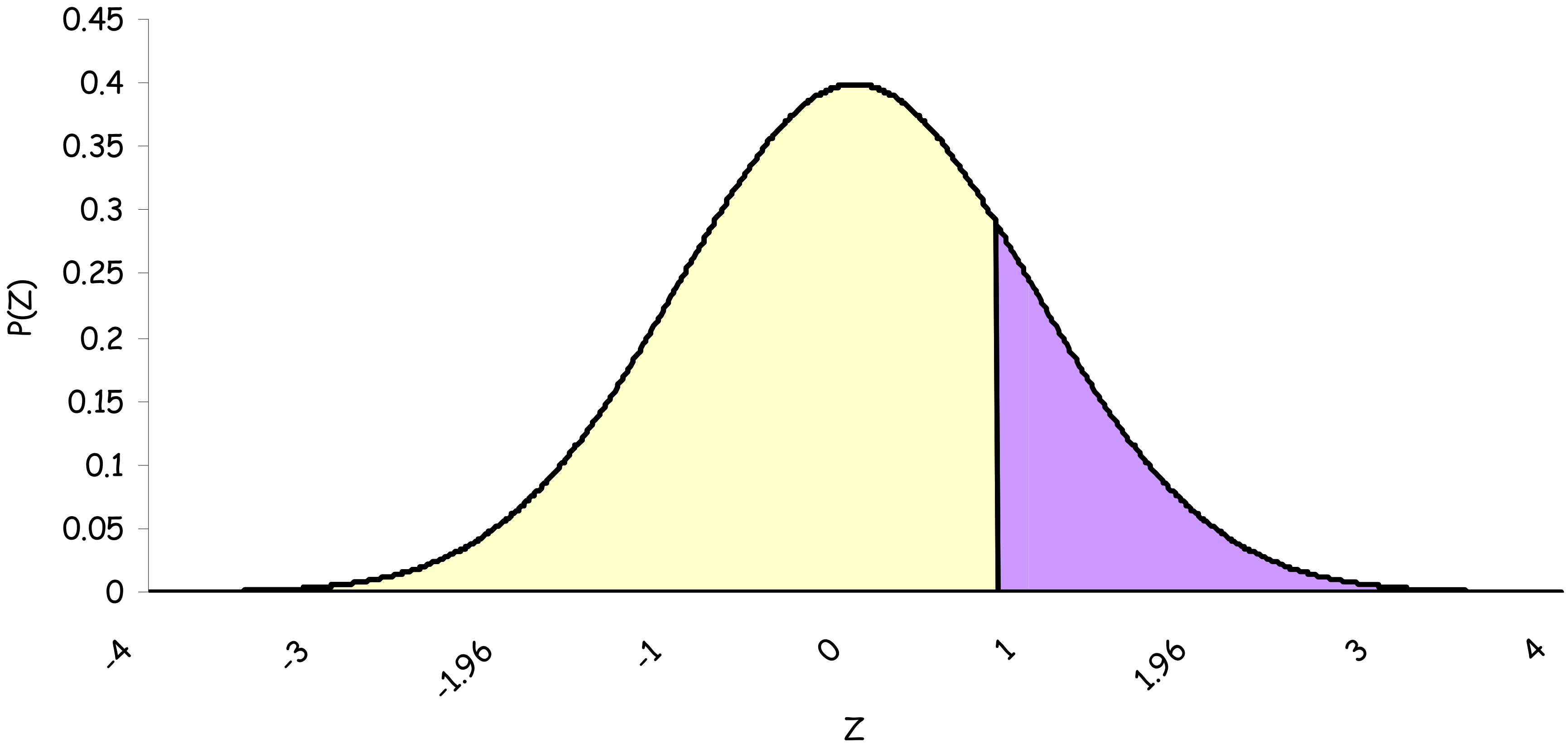
Esercizio 11

In uno studio condotto su bambini di età compresa tra gli 8 e i 15 anni, Eldridge et al. hanno valutato 529 bambini normalmente sviluppati per valutare il tempo trascorso in posizione eretta. I ricercatori hanno trovato che il tempo totale che un bambino trascorre in posizione eretta segue una distribuzione normale con media pari a 5.4 ore e una deviazione standard di 1.3 ore. Assumendo che lo studio si applichi a tutti i bambini con età 8-15 anni,

1. trova la probabilità che un bambino scelto a caso passi meno di 3 ore in posizione eretta in un giorno (24h).
2. In una popolazione di 10000 bambini, quanti te ne aspetti di trovare che passino più di 8.5h in posizione eretta?



1. $P[x<3]=P[\,(x-\mu)/\sigma < (3-\mu)/\sigma\,] = P[z<(3-5.4)/1.3] = P[z<1.85] =0.0322$



2. $P[x>8.5]=P[\ (x-\mu)/\sigma > (8.5-\mu)/\sigma\ ] =$
$P[z>(8.5-5.4)/1.3] = P[z>2.38] =$
$=0.0087$
$N=10000*0.0087= 87$

Es. 12 Malattie cardiovascolari: ipertensione arteriosa

- Media e deviazione standard della pressione arteriosa sistolica per gruppi di età (valori espressi in mmHg)

Età (yr)	Media	SD	Livello limite
1-14	105,0	5,0	115,0
15-44	125,0	10,0	140,0

- Ammettiamo che le persone con valori di PAS superiori al limite dichiarato per gruppo di età siano definite ipertese.
- Qual è la proporzione degli ipertesi nel gruppo di età tra 1 e 14 anni ? Quale nel gruppo di età tra 15 e 44 anni ?

Malattie cardiovascolari: ipertensione arteriosa

- Standardizziamo
- Per x=115 ottengo

$$z = \frac{x-\mu}{\sigma} = \frac{115,0-105,0}{5,0} = 2,00$$

- Per x=140 ottengo

$$z = \frac{x-\mu}{\sigma} = \frac{140,0-125,0}{10,0} = 1,50$$

Malattie cardiovascolari: ipertensione arteriosa

- Dalla tabella 1:
  - **Pr(z>2,00)=0,0228**
  - Il **2,3%** dei bambini tra 1 e 14 anni è iperteso
- Dalla tabella 1:
  - **Pr(z>1,50)=0,0,668**
  - Il **6,7%** dei soggetti tra 15 e 44 anni è iperteso