Esercitazioni 3-4: Distribuzione di Gauss

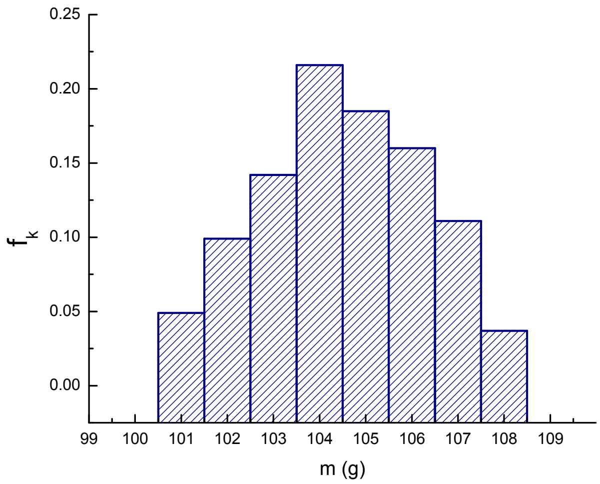
1) ***Uno studente pesa N=162 cubetti di legno, tagliati a mano, e raccoglie i risultati nella seguente tabella:***

massa [g]	***101***	***102***	***103***	***104***	***105***	***106***	***107***	***108***
frequenza	***8***	***16***	***23***	***35***	***30***	***26***	***18***	***6***

a) Rappresentare le misure con un istogramma.
b) Calcolare il valore medio della massa dei cubetti, la deviazione standard e l'errore standard del valore medio.
c) Scrivere la funzione densità di probabilità più adatta a descrivere questo insieme di misure e confrontarla graficamente con l'istogramma delle misure.
d) Se lo studente dovesse pesare altri 100 cubetti dello stesso tipo, quanti se ne dovrebbe aspettare di massa superiore a 106.5 g?

a) Rappresentiamo le misure con un istogramma a intervalli di larghezza Δ=1 g. Calcoliamo per ogni misura m_k ($k = 1...8$) la densità di frequenza relativa $f_k = F_k/\Delta = n_k/(N\Delta)$:

m_k [g]	101	102	103	104	105	106	107	108
n_k	8	16	23	35	30	26	18	6
f_k	0.049	0.099	0.142	0.216	0.185	0.160	0.111	0.037



b) Il valore medio $\overline{m}$ della massa dei cubetti, la deviazione standard σ_m e l'errore standard del valore medio $\sigma_{\overline{m}}$ sono dati da

$$\overline{m} = \frac{\sum_{k=1}^{8} n_k m_k}{N} = 104.50\ \mathrm{g}$$

$$\sigma_m = \sqrt{\frac{\sum_{k=1}^{8} n_k (m_k - \overline{m})^2}{N-1}} = 1.78\ \mathrm{g}$$

$$\sigma_{\overline{m}} = \frac{\sigma_m}{\sqrt{N}} = 0.14\ \mathrm{g}$$

$$\rightarrow \overline{m} = (104.50 \pm 0.14)\ g$$

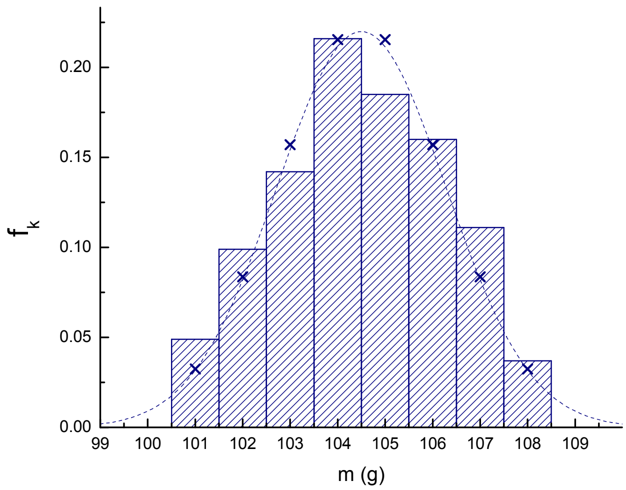
c) Assumendo che le misure siano soggette ad errori casuali e che gli errori sistematici siano trascurabili, la funzione densità di probabilità più adatta a descrivere i dati è la distribuzione di Gauss centrata in $\overline{m}$ con deviazione standard σ_m:

$$G_{\overline{m},\sigma_m}(m) = \frac{1}{\sigma_m\sqrt{2\pi}} e^{-\frac{(m-\overline{m})^2}{2\sigma_m^2}}$$

Per confrontare graficamente la distribuzione con l'istogramma delle misure, calcoliamo il valore della distribuzione di Gauss in corrispondenza dei valori di massa pesati dallo studente:

$$G_{\overline{m},\sigma_m}(m = 101) = \frac{1}{1.78\sqrt{2\pi}} e^{-(101-104.5)^2/(2\cdot 1.78^2)} = 0.0324 = G_{\overline{m},\sigma_m}(m = 108)$$

$$G_{\overline{m},\sigma_m}(m = 102) = 0.0836 = G_{\overline{m},\sigma_m}(m = 107)$$
$$G_{\overline{m},\sigma_m}(m = 103) = 0.1571 = G_{\overline{m},\sigma_m}(m = 106)$$
$$G_{\overline{m},\sigma_m}(m = 104) = 0.2155 = G_{\overline{m},\sigma_m}(m = 105)$$



d) Calcoliamo infine la probabilità di pesare un cubetto e ottenere una massa $m > \widetilde{m}$ =106.5 g. $\widetilde{m}$ dista da $\overline{m}$ per t deviazioni standard, con

$$t = \frac{\widetilde{m} - \overline{m}}{\sigma_m} = \frac{(106.5 - 104.5)\ g}{1.78\ g} = 1.12$$

La probabilità di pesare una massa $m > \widetilde{m} = \overline{m} + t\sigma_m$ è data da

$$p(m > \overline{m} + t\sigma_m) = \frac{1 - p(\overline{m} - t\sigma_m < m < \overline{m} + t\sigma_m)}{2}$$

$p(\overline{m} - t\sigma_m < m < \overline{m} + t\sigma_m)$ è la probabilità per una singola misura di cadere entro t deviazioni standard dal valor medio, e per t=1.12 vale

$$p(\text{entro } t\sigma_m) = \frac{1}{\sqrt{2\pi}} \int_{-t}^{t} e^{-z^2/2} dz \rightarrow p(\text{entro } 1.12\sigma_m) = 0.7373$$

$$\rightarrow p(m > \overline{m} + 1.12\sigma_m) = \frac{1}{2}(1 - 0.7373) = 0.13$$

Pesando 100 cubetti, ci aspettiamo quindi che 13 di questi abbiano massa maggiore di 106.5 g.

2) *Uno studente effettua N=140 misure del periodo di oscillazione di un pendolo semplice ottenendo i seguenti risultati:*

T [s]	***0.17***	***0.18***	***0.19***	***0.20***	***0.21***	***0.22***	***0.23***
frequenza	***5***	***10***	***18***	***32***	***40***	***24***	***11***

a) Calcolare il valor medio del periodo, la deviazione standard e l'errore standard del valore medio.
b) Scrivere la funzione densità di probabilità più adatta a descrivere questo insieme di misure e confrontarla graficamente con l'istogramma delle misure.
c) Se le condizioni di misura non cambiano, che probabilità ha lo studente che misura un' altra volta il periodo del pendolo di trovare un valore superiore a 0.23 s?

a) Il valore medio $\overline{T}$ del periodo di oscillazione, la deviazione standard σ_T e l'errore standard del valore medio $\sigma_{\overline{T}}$ sono dati da

$$\overline{T} = \frac{\sum_{k=1}^{7} n_k T_k}{N} = 0.205\,\mathrm{s}$$

$$\sigma_T = \sqrt{\frac{\sum_{k=1}^{7} n_k (T_k - \overline{T})^2}{N-1}} = 0.015\,\mathrm{s}$$

$$\sigma_{\overline{T}} = \frac{\sigma_T}{\sqrt{N}} = 0.0013\,\mathrm{s}$$

$$\rightarrow \overline{T} = (0.2050 \pm 0.0013)\,\mathrm{s}$$

b) Assumendo che le misure siano soggette ad errori casuali e che gli errori sistematici siano trascurabili, la funzione densità di probabilità più adatta a descrivere i dati è la distribuzione di Gauss centrata in $\overline{T}$ con deviazione standard σ_T:

$$G_{\overline{T},\sigma_T}(T) = \frac{1}{\sigma_T\sqrt{2\pi}} e^{-\frac{(T-\overline{T})^2}{2\sigma_T^2}}$$

Per rappresentare le misure con un istogramma a intervalli e confrontarlo con la distribuzione di Gauss calcoliamo per ogni misura T_k (k=1...7) la densità di frequenza relativa $f_k = F_k/\Delta = n_k/(N\Delta)$, con Δ=0.01 s:

T_k [s]	0.17	0.18	0.19	0.20	0.21	0.22	0.23
n_k	5	10	18	32	40	24	11
f_k	3.571	7.143	12.857	22.857	28.571	17.143	7.857

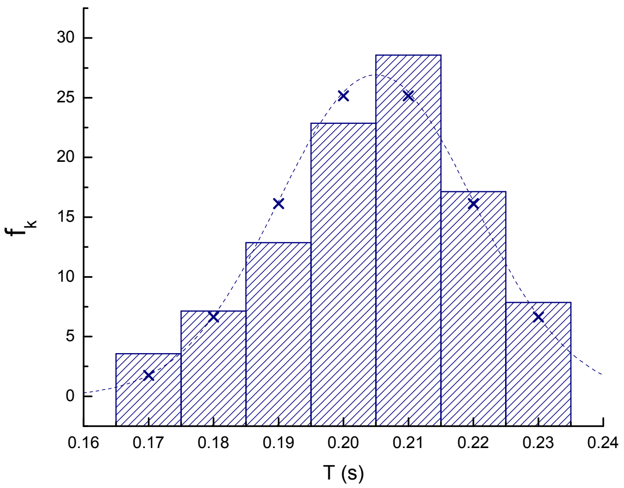
In corrispondenza dei valori di periodo misurati dallo studente la distribuzione di Gauss assume i valori

$$G_{\overline{T},\sigma_T}(T = 0.17) = 1.7481$$

$$G_{\overline{T},\sigma_T}(T = 0.18) = 6.6318 = G_{\overline{T},\sigma_T}(T = 0.23)$$

$$G_{\overline{T},\sigma_T}(T = 0.19) = 16.1314 = G_{\overline{T},\sigma_T}(T = 0.22)$$

$$G_{\overline{T},\sigma_T}(T = 0.20) = 25.1589 = G_{\overline{T},\sigma_T}(T = 0.21)$$



d) Calcoliamo infine la probabilità di misurare un periodo $T > \widetilde{T}$ =0.23 s. $\widetilde{T}$ dista da $\overline{T}$ per t deviazioni standard, con

$$t = \frac{\widetilde{T} - \overline{T}}{\sigma_T} = \frac{(0.23 - 0.205)\,s}{0.015\,s} = 1.7$$

$$p(T > 0.23\,s) = p(T > \overline{T} + 1.7\sigma_T) = \frac{1 - p(\text{entro } 1.7\sigma_T)}{2} = \frac{1}{2}(1 - 0.9109) = 0.045$$

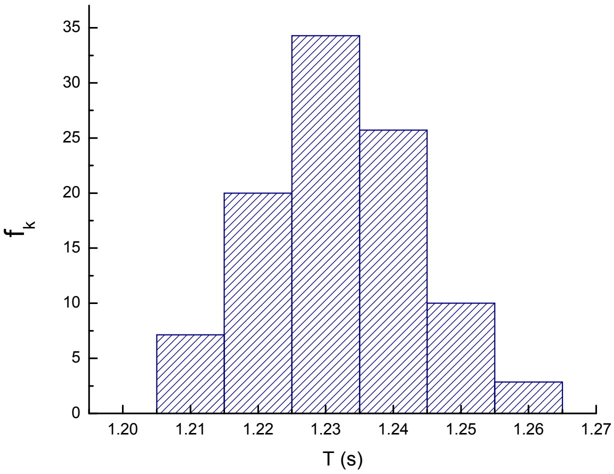
3) ***Uno studente misura per N=70 volte i tempi di caduta da una determinata altezza di una sferetta di acciaio e raccoglie i risultati nella seguente tabella:***

T [s]	***1.21***	***1.22***	***1.23***	***1.24***	***1.25***	***1.26***
frequenza	***5***	***14***	***24***	***18***	***7***	***2***

a) Rappresentare le misure con un istogramma.
b) Calcolare il valore medio del tempo di caduta, la deviazione standard e l'errore standard del valore medio.
c) Nell'ipotesi che le fluttuazioni tra le misure effettuate siano dovute solo ad errori casuali, definire l'intervallo centrato nel valore medio a cui è associato un livello di confidenza dell'85%.

a) Per rappresentare le misure con un istogramma a intervalli calcoliamo per ogni misura T_k (k=1...6) la densità di frequenza relativa $f_k = F_k/\Delta = n_k/(N\Delta)$, con Δ=0.01 s:

T_k[s]	1.21	1.22	1.23	1.24	1.25	1.26
n_k	5	14	24	18	7	2
f_k	7.14	20.00	34.29	25.71	10.00	2.86



b) Il valore medio $\overline{T}$ del tempo di caduta, la deviazione standard σ_T e l'errore standard del valore medio $\sigma_{\overline{T}}$ sono dati da

$$\overline{T} = \frac{\sum_{k=1}^{6} n_k T_k}{N} = 1.232\,\mathrm{s}$$

$$\sigma_T = \sqrt{\frac{\sum_{k=1}^{6} n_k (T_k - \overline{T})^2}{N-1}} = 0.0118\,\mathrm{s}$$

$$\sigma_{\overline{T}} = \frac{\sigma_T}{\sqrt{N}} = 0.0014\,\mathrm{s}$$

$$\rightarrow \overline{T} = (1.2320 \pm 0.0014)\,\mathrm{s}$$

c) In presenza di soli errori casuali, abbiamo l'85% di probabilità che una misura giaccia entro 1.44 deviazioni standard dal valore medio $\overline{T}$:

$$p(\text{entro } t\sigma_T) = 0.85 \text{ per } t = 1.44$$

Pertanto l'intervallo centrato nel valore medio a cui è associato un livello di confidenza dell'85% si estende da $(\overline{T} - 1.44\sigma_T) = 1.2152\,\mathrm{s}$ a $(\overline{T} + 1.44\sigma_T) = 1.2488\,\mathrm{s}$.

4) ***La misura del diametro di 100 rondelle ha fornito il risultato*** $\overline{d} = (6.21 \pm 0.03)\,\mathrm{mm}$***. Supponendo che le fluttuazioni siano dovute solo ad effetti casuali, quante rondelle ci si può aspettare tra le 100 prodotte con diametro*** $d > 6.06\,\mathrm{mm}$***? E con diametro*** $d > 6.72\,\mathrm{mm}$***?***

Dalla deviazione standard della media $\sigma_{\overline{d}}$ ricaviamo la deviazione standard

$$\sigma_d = \sqrt{N}\sigma_{\overline{d}} = 0.3\,\mathrm{mm}$$

$\widetilde{d} = 6.06\,\mathrm{mm}$ dista dal valore medio t deviazioni standard, con

$$t = \frac{\widetilde{d} - \overline{d}}{\sigma_d} = -0.5$$

La probabilità che una rondella abbia diametro $d > \widetilde{d} = 6.06\,\mathrm{mm}$ può essere calcolata come

$$p(d > \overline{d} - 0.5\sigma_d) = 1 - \frac{p(d < \overline{d} - 0.5\sigma_d, d > \overline{d} + 0.5\sigma_d)}{2} = 1 - \frac{1 - p(\text{entro } 0.5\sigma_d)}{2} =$$

$$= 1 - \frac{1 - 0.3829}{2} = 0.69$$

Su 100 rondelle, ci aspettiamo che 69 di queste abbiano diametro maggiore di 6.06 mm.
6.72 mm dista dal valor medio 1.7 deviazioni standard, quindi la probabilità che una rondella abbia diametro d >6.72 mm è

$$p(d > \overline{d} + 1.7\sigma_d) = \frac{p(d < \overline{d} - 1.7\sigma_d, d > \overline{d} + 1.7\sigma_d)}{2} = \frac{1 - p(\text{entro } 1.7\sigma_d)}{2} =$$

$$= \frac{1 - 0.9109}{2} = 0.045$$

Su 100 rondelle, ci aspettiamo che 5 di esse abbiano diametro maggiore di 6.72 mm.

5) ***Per controllare la regolarità del deposito di una vernice su una superficie si misura lo spessore* d *del materiale depositato in 10 diversi punti del substrato, ottenendo i risultati* d*=15, 16, 18, 14, 15, 14, 17, 16, 15, 13* μm.**
a) Calcolare il valore medio dello spessore, la deviazione standard e l'errore standard del valore medio.
b) Supponendo che le differenze di spessore siano di tipo casuale, che frazione di superficie si può prevedere con un deposito inferiore a 13* μm *e che frazione della superficie con un deposito superiore a 15* μm*?

a) Riportiamo in tabella la frequenza assoluta n_k per ciascun valore di spessore d_k (k=1...6) e calcoliamo valor medio, deviazione standard e errore standard del valor medio:

d_k [μm]	13	14	15	16	17	18
n_k	1	2	3	2	1	1

$$\overline{d} = \frac{\sum_{k=1}^{6} n_k d_k}{N} = 15.3\,\mu m$$

$$\sigma_d = \sqrt{\frac{\sum_{k=1}^{6} n_k (d_k - \overline{d})^2}{N-1}} = 1.49\,\mu m$$

$$\sigma_{\overline{d}} = \frac{\sigma_d}{\sqrt{N}} = 0.47\,\mu m$$

$$\rightarrow \overline{d} = (15.3 \pm 0.5)\,\mu m$$

b) Lo spessore 13 μm dista dal valor medio t=-1.54 deviazioni standard; la probabilità che venga misurato uno spessore inferiore a 13 μm è quindi data da

$$p(d < \overline{d} - 1.54\sigma_d) = \frac{1 - p(\text{entro } 1.54\sigma_d)}{2} = \frac{1 - 0.8754}{2} = 0.062$$

Lo spessore 15 μm dista dal valor medio -0.2 deviazioni standard, quindi

$$p(d > \overline{d} - 0.2\sigma_d) = 1 - \frac{1 - p(\text{entro } 0.2\sigma_d)}{2} = 1 - \frac{1 - 0.1585}{2} = 0.58$$

6) ***Uno studente misura la carica dell'elettrone e nota che la sua risposta dista di 2.4 deviazioni standard dal valore atteso. Questa discrepanza è significativa al livello del 5%? E al livello del 2%? O dell'1%?***

Supponiamo che le misure siano governate da una distribuzione normale centrata sul valore atteso x_{atteso} con parametro di larghezza uguale alla deviazione standard σ delle misure dello studente. Sappiamo che

$$t = \frac{|x_{misurato} - x_{atteso}|}{\sigma} = 2.4$$

La probabilità di ottenere una misura che differisce da x_{atteso} per t o più deviazioni standard è

$$p(x < x_{atteso} - t\sigma, x > x_{atteso} + t\sigma) = 1 - p(\text{entro } t\sigma)$$

Dalla tabella dell'integrale normale degli errori, per t=2.4

$$p(x < x_{atteso} - 2.4\sigma, x > x_{atteso} + 2.4\sigma) = 1 - 0.9836 = 0.0164 = 1.64\%$$

La discrepanza tra la misura dello studente e il valore atteso risulta quindi significativa al livello del 5% e al livello del 2%, mentre non risulta significativa al livello dell'1%.

7) ***Uno studente misura g, l'accelerazione di gravità, ripetutamente e ottiene un risultato finale di 9.6 m/s^2 con una deviazione standard $\sigma = 0.1\,m/s^2$. Se la sua misura fosse normalmente distribuita, con centro nel valore accettato 9.8 m/s^2 e con deviazione standard $0.1\,m/s^2$, quale sarebbe la probabilità di ottenere un risultato che differisce dal valore atteso di almeno quanto il suo?***

Data la discrepanza tra valore atteso e valore misurato,

$$t = \frac{|x_{misurato} - x_{atteso}|}{\sigma} = 2$$

$$\rightarrow p(x < x_{atteso} - 2\sigma, x > x_{atteso} + 2\sigma) = 1 - 0.9545 = 0.0455 = 4.6\%$$

Questa discrepanza è inaccettabile al livello del 5%.

8) ***Uno studente misura una grandezza x molte volte e calcola la sua media $\overline{x} = 23$ e la deviazione standard $\sigma_x = 1$. Determinare la frazione di letture attese nell'intervallo a) [22, 24], b) [24, 25], c) [21, 23] nell'ipotesi che le fluttuazioni tra le misure effettuate siano dovute solo ad errori casuali.***

a) L'intervallo [22, 24] corrisponde all'intervallo $[\overline{x} - \sigma_x, \overline{x} + \sigma_x]$: la frazione di misure entro una deviazione standard dal valor medio è il 68.27%.
b) L'intervallo [24, 25] corrisponde all'intervallo $[\overline{x} + \sigma_x, \overline{x} + 2\sigma_x]$: possiamo calcolare $p(\overline{x} + \sigma_x < x < \overline{x} + 2\sigma_x)$ come

$$p(\overline{x} + \sigma_x < x < \overline{x} + 2\sigma_x) = p(x > \overline{x} + \sigma_x) - p(x > \overline{x} + 2\sigma_x)$$

$$\rightarrow p(\overline{x} + \sigma_x < x < \overline{x} + 2\sigma_x) = \frac{1 - p(\text{entro } \sigma_x)}{2} - \frac{1 - p(\text{entro } 2\sigma_x)}{2} =$$

$$= \frac{1-0.6827}{2} - \frac{1-0.9545}{2} = 0.1359$$

L'intervallo [24, 25] contiene quindi il 13.59% delle misure.
c) L'intervallo [21, 23] corrisponde all'intervallo $[\overline{x} - 2\sigma_x, \overline{x}]$, quindi

$$p(\overline{x} - 2\sigma_x < x < \overline{x}) = \frac{p(\text{entro } 2\sigma_x)}{2} = \frac{0.9545}{2} = 0.4773$$

Il 47.73% delle letture giace tra 21 e 23.

9) ***Una macchina confeziona sacchetti di caramelle dal peso medio $\overline{\mu}$=200 g con una deviazione standard σ_μ=10 g; calcolare a) la percentuale di sacchetti che pesano più di 218 g, e b) la percentuale di sacchetti che pesano più di 190 g. c) Se ad un controllo vengono scartati i sacchetti con un peso inferiore a 185 g, su 1000 sacchetti quanti si prevede verranno scartati?***

a) Il peso di 218 g dista dal valor medio t deviazioni standard, con

$$t = \frac{(218 - 200)\text{ g}}{10\text{ g}} = 1.8$$

$$\rightarrow p(\mu > 218\text{ g}) = p(\mu > \overline{\mu} + 1.8\sigma_\mu) = \frac{1 - p(\text{entro } 1.8\sigma_\mu)}{2} = \frac{1-0.9281}{2} = 0.036$$

Il 3.6% dei sacchetti ha peso maggiore di 218 g.
b) La probabilità che il peso sia maggiore di 190 g è

$$p(\mu > 190\text{ g}) = p(\mu > \overline{\mu} - \sigma_\mu) = 1 - p(\mu < \overline{\mu} - \sigma_\mu) = 1 - \frac{1 - p(\text{entro } \sigma_\mu)}{2} = 0.8414$$

L'84% dei sacchetti pesa più di 190 g.
c) La probabilità che il peso sia inferiore a 185 g è

$$p(\mu < 185\text{ g}) = p(\mu < \overline{\mu} - 1.5\sigma_\mu) = \frac{1 - p(\text{entro } 1.5\sigma_\mu)}{2} = \frac{1-0.8664}{2} = 0.067$$

Su 1000 sacchetti si prevede ne vengano scartati 67.

10) ***Il servizio clienti di una compagnia aerea riceve numerose telefonate ogni giorno. Il tempo di attesa in linea di ogni cliente prima di poter parlare con un operatore è distribuito normalmente con media $\overline{T}$=3.1 minuti e deviazione standard σ_T=0.9 minuti.***
a) Quanti utenti rimangono in attesa per più di 4.3 minuti?
b) Qual è il tempo di attesa minimo per il 2.5% dei clienti?

a) Dato

$$t = \frac{(4.3 - 3.1)\text{ minuti}}{0.9\text{ minuti}} = 1.33$$

la probabilità di restare in attesa per più di 4.3 minuti risulta

$$p(T > \overline{T} + 1.33\sigma_T) = \frac{1 - p(\text{entro } 1.33\sigma_T)}{2} = \frac{1-0.8165}{2} = 0.092$$

Il 9.2% dei clienti rimane in attesa più di 4.3 minuti.
b) Vogliamo determinare T_{min} tale per cui la probabilità di restare in attesa per più di T_{min} minuti è 0.025. Se

$$t = \frac{T_{min} - \overline{T}}{\sigma_T}$$

vale che

$$0.025 = p(T > T_{min}) = p(T > \overline{T} + t\sigma_T) = \frac{1 - p(\text{entro } t\sigma_T)}{2}$$

Risolvendo per t dalla tabella dell'integrale normale degli errori,

$$p(\text{entro } t\sigma_T) = 1 - 2 \cdot 0.025 = 0.95 \rightarrow t = 1.96$$

Il 2.5% dei clienti rimane quindi in attesa per più di $T_{min} = (3.1 + 1.96 \cdot 0.9)$ minuti=4.86 minuti.