3.**Durante il picco massimo di un'epidemia di influenza il 15% della popolazione è a casa ammalato: a) qual è la probabilità che in una classe di 20 alunni ce ne siano più di due assenti per l'influenza? b) descrivere le operazioni da compiere per verificare che, se l'intera scuola ha 500 alunni, la probabilità che ce ne siano più di 50 influenzati è maggiore del 99%.**

Soluzione

a) Con opportune ipotesi aggiuntive si può accettare che la variabile aleatoria X= <<numero di alunni malati>> segua una distribuzione binomiale con **n=20** $\boldsymbol{p = 0,15 \quad e\ q = 1 - p = 0,85}$.

Gli eventi che costituiscono le <<prove ripetute>> devono essere indipendenti e la probabilità che un alunno sia malato deve avere sempre valore uguale a p

La probabilità che un evento di probabilità p si verifichi k volte in n prove è

$$\boldsymbol{P(k) = C_{n,k}\, p^k q^{n-k}}$$

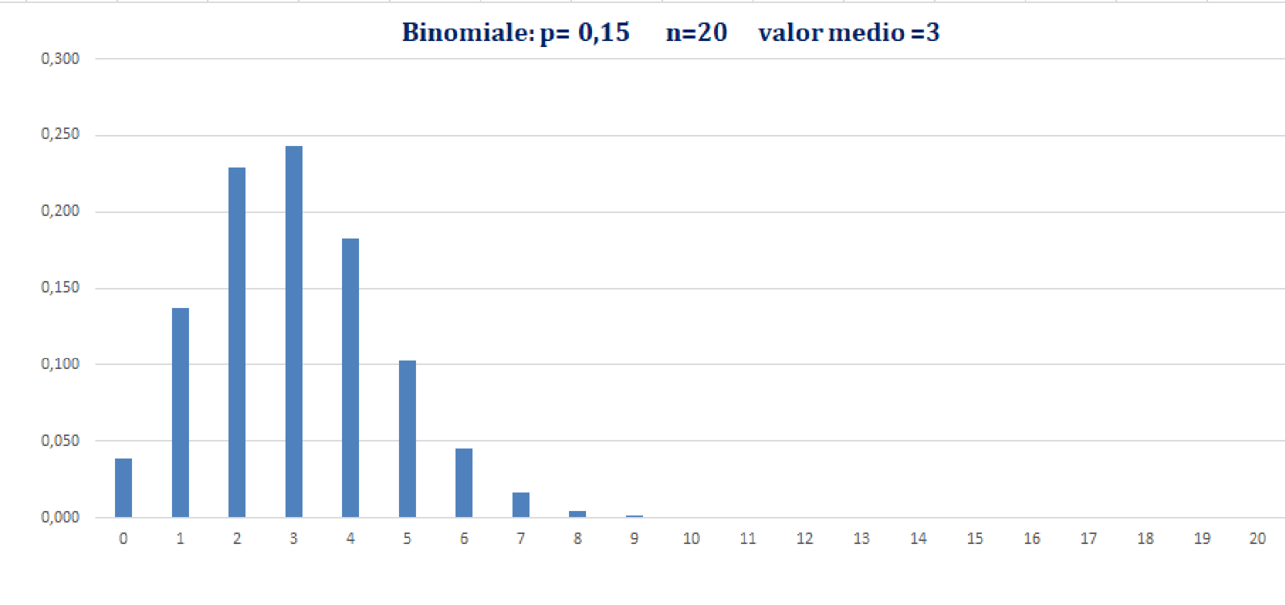
Il valor medio della distribuzione è uguale a $\boldsymbol{np}$.

Nel nostro caso la distribuzione è quella della tabella e della figura sottostanti.

La probabilità che nella classe di 20 alunni ce ne siano più di due assenti per l'influenza è uguale a

$\boldsymbol{P(X > 2) = 1 - [P(0) + P(1) + P(2)] \approx 1 - 0,405}$ **=0,595** $\boldsymbol{\approx 60\%}$

	n=20	p=(0,15)	E(X)=3
	Binomiale	Binomiale cumulata	
k	P(X=k)	P(X<=x)	P(X>2)
0	0,039	0,039	**0,595**
1	0,137	0,176	
2	0,229	0,405	
3	0,243	0,648	
4	0,182	0,830	
5	0,103	0,933	
6	0,045	0,978	
7	0,016	0,994	
8	0,005	0,999	
9	0,001	1,000	
10	0,000	1,000	
11	0,000	1,000	
12	0,000	1,000	
13	0,000	1,000	
14	0,000	1,000	
15	0,000	1,000	
16	0,000	1,000	
17	0,000	1,000	
18	0,000	1,000	
19	0,000	1,000	
20	0,000	1,000	



b)Se si considerano 500 alunni, dato l'elevato numero di unità statistiche, la distribuzione binomiale può essere approssimata con una distribuzione normale la quale è una distribuzione continua.

La funzione di distribuzione normale $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$

e il suo grafico è la curva di Gauss , simmetrica rispetto alla retta $x = \mu$ e dalla caratteristica forma a campana

Nel nostro caso ha $\mu = np = 500 \cdot 0{,}15 = 75 \quad e \quad \sigma = \sqrt{npq} = \sqrt{500 \cdot 0{,}15 \cdot 0{,}85} \approx 7{,}98$

Se X è un valore della variabile aleatoria, la probabilità che X appartenga all'intervallo $[a; b]$ è

$$P[a \leq X \leq b] = \int_a^b f(x)dx$$

$Geometricamente$ è $l'area\ sottesa\ dalla\ curva\ sull'intervallo [a; b]$
In particolare

$$\int_{-\infty}^{\mu} f(x)dx = \int_{\mu}^{+\infty} f(x)dx = 0{,}5 \qquad \int_{-\infty}^{+\infty} f(x)dx = 1$$

Nel nostro caso interessa

$$P[X > 50] = \int_{50}^{+\infty} f(x)dx)$$

Poiché manca un'espressione elementare per la primitiva di $f(x)$, il valore dell'integrale è fornito da adeguati strumenti di calcolo oppure da tabelle.

Solitamente si fa riferimento alla distribuzione normale standardizzata ,

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

risultato di una trasformazione affine che porta nell'origine il valor medio della distribuzione e riduce il valore di σ al valore unitario.

I valori di X diventano $Z = \frac{X-\mu}{\sigma}$ e compariranno sull'asse x , espressi i unità di deviazione standard

mentre le ordinate sono moltiplicate per σ.

La trasformazione conserva le aree.

La funzione di ripartizione è Φ(z)= $\boldsymbol{P(Z \le z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt}$ **.**

Spesso si prendono in considerazione , per Φ(z,) alcuni valori di riferimento, da considerare noti.

$\boldsymbol{\Phi(0) = 0,5}$ **Φ(1,65)=0,95 Φ(1,96) =0,975 Φ(2,58) =0,995 .**

Una proprietà importante è inoltre, **uguaglianza delle aree delle due "code",** per cui

$$\boldsymbol{\Phi(-z) = 1 - \Phi(z)}$$

Il quesito chiede di verificare se la probabilità che ci siano più di 50 alunni influenzati è maggiore del 99% ,quindi si deve valutare il valore complementare della funzione di ripartizione

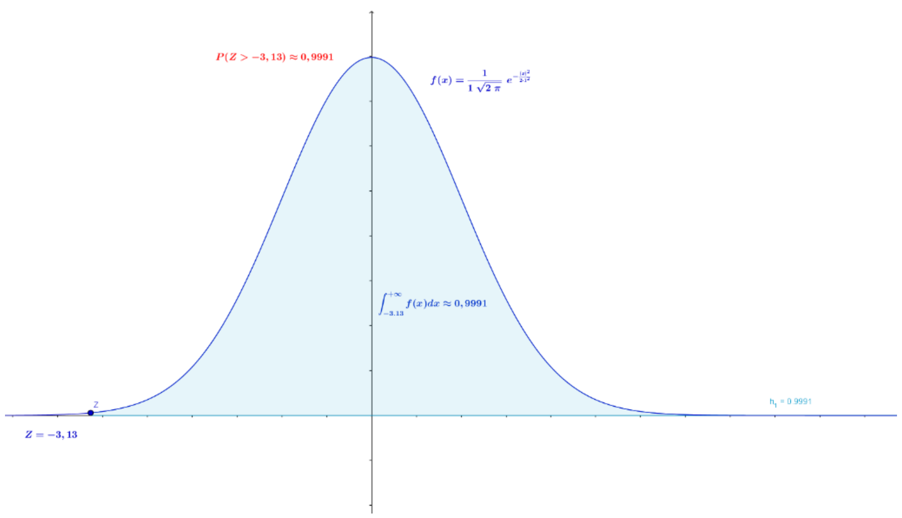
$$\boldsymbol{P(Z > z) = 1 - \Phi(z)}$$

Il primo passo è l'operazione di standardizzazione

$$z = \frac{x-\mu}{\sigma} = \frac{50-75}{7{,}98} \approx -3{,}13$$

In seguito, possiamo seguire uno dei procedimenti seguenti:

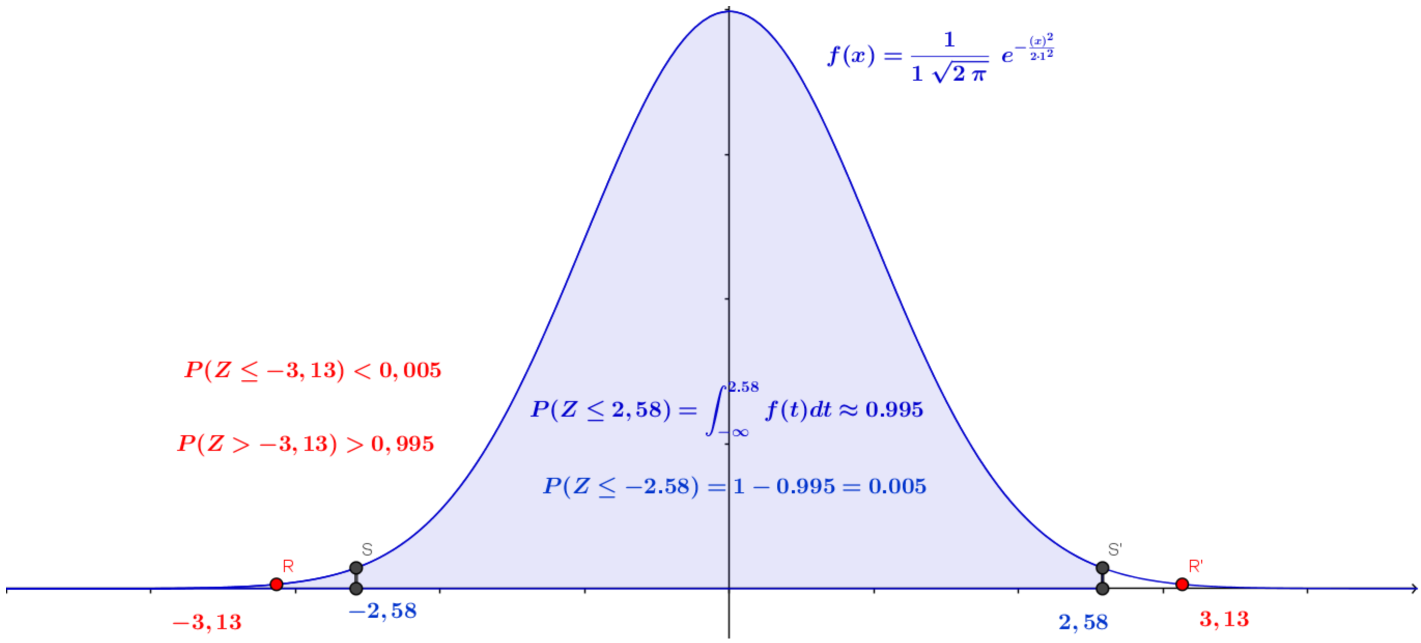
1. Con l'aiuto di un software o di strumenti di calcolo adeguati ,si calcola immediatamente l'area sottesa dalla curva nell'intervallo $[-3{,}13; +\infty[$, ovvero il valore di $\int_{-3{,}13}^{+\infty} f(t)dt$



Il valore trovato è 0,9991 , quindi maggiore di 0.99

2. Ricordando che Φ(2,58) =0,995 e quindi Φ (-2,58) =1-0,995 =0,005 osserviamo che:

per la monotonia della funzione di ripartizione
Φ(-3,13)< Φ(-2,58) pertanto
Φ (-3,13) < 0,005→ $\boldsymbol{P(Z > -3,13) = 1 - \Phi(-3,13) > 1 - 0,005}$ **=0,995→**
$\boldsymbol{P(Z > -3,13) > 99\%}$



3. Si legge il valore di **Φ (3,31) =0,99913** su una tabella della funzione di ripartizione della distribuzione normale standardizzata.

Essendo Φ (-3,31) = 1 - Φ (3,31) sarà

1- Φ (-3,31) = $P(Z > -3,31)$ **= Φ (3,31) =0,99913**

Tavole della funzione di ripartizione della variabile Normale Standardizzata:

$$\Phi(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2} dt$$

	Seconda cifra decimale di z									
z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.50000	0.50399	0.50798	0.51197	0.51595	0.51994	0.52392	0.52790	0.53188	0.53586
0.1	0.53983	0.54380	0.54776	0.55172	0.55567	0.55962	0.56356	0.56749	0.57142	0.57535
0.2	0.57926	0.58317	0.58706	0.59095	0.59483	0.59871	0.60257	0.60642	0.61026	0.61409
0.3	0.61791	0.62172	0.62552	0.62930	0.63307	0.63683	0.64058	0.64431	0.64803	0.65173
0.4	0.65542	0.65910	0.66276	0.66640	0.67003	0.67364	0.67724	0.68082	0.68439	0.68793
0.5	0.69146	0.69497	0.69847	0.70194	0.70540	0.70884	0.71226	0.71566	0.71904	0.72240
0.6	0.72575	0.72907	0.73237	0.73565	0.73891	0.74215	0.74537	0.74857	0.75175	0.75490
0.7	0.75804	0.76115	0.76424	0.76730	0.77035	0.77337	0.77637	0.77935	0.78230	0.78524
0.8	0.78814	0.79103	0.79389	0.79673	0.79955	0.80234	0.80511	0.80785	0.81057	0.81327
0.9	0.81594	0.81859	0.82121	0.82381	0.82639	0.82894	0.83147	0.83398	0.83646	0.83891
1.0	0.84134	0.84375	0.84614	0.84849	0.85083	0.85314	0.85543	0.85769	0.85993	0.86214
1.1	0.86433	0.86650	0.86864	0.87076	0.87286	0.87493	0.87698	0.87900	0.88100	0.88298
1.2	0.88493	0.88686	0.88877	0.89065	0.89251	0.89435	0.89617	0.89796	0.89973	0.90147
1.3	0.90320	0.90490	0.90658	0.90824	0.90988	0.91149	0.91308	0.91466	0.91621	0.91774
1.4	0.91924	0.92073	0.92220	0.92364	0.92507	0.92647	0.92785	0.92922	0.93056	0.93189
1.5	0.93319	0.93448	0.93574	0.93699	0.93822	0.93943	0.94062	0.94179	0.94295	0.94408
1.6	0.94520	0.94630	0.94738	0.94845	0.94950	0.95053	0.95154	0.95254	0.95352	0.95449
1.7	0.95543	0.95637	0.95728	0.95818	0.95907	0.95994	0.96080	0.96164	0.96246	0.96327
1.8	0.96407	0.96485	0.96562	0.96638	0.96712	0.96784	0.96856	0.96926	0.96995	0.97062
1.9	0.97128	0.97193	0.97257	0.97320	0.97381	0.97441	0.97500	0.97558	0.97615	0.97670
2.0	0.97725	0.97778	0.97831	0.97882	0.97932	0.97982	0.98030	0.98077	0.98124	0.98169
2.1	0.98214	0.98257	0.98300	0.98341	0.98382	0.98422	0.98461	0.98500	0.98537	0.98574
2.2	0.98610	0.98645	0.98679	0.98713	0.98745	0.98778	0.98809	0.98840	0.98870	0.98899
2.3	0.98928	0.98956	0.98983	0.99010	0.99036	0.99061	0.99086	0.99111	0.99134	0.99158
2.4	0.99180	0.99202	0.99224	0.99245	0.99266	0.99286	0.99305	0.99324	0.99343	0.99361
2.5	0.99379	0.99396	0.99413	0.99430	0.99446	0.99461	0.99477	0.99492	0.99506	0.99520
2.6	0.99534	0.99547	0.99560	0.99573	0.99585	0.99598	0.99609	0.99621	0.99632	0.99643
2.7	0.99653	0.99664	0.99674	0.99683	0.99693	0.99702	0.99711	0.99720	0.99728	0.99736
2.8	0.99744	0.99752	0.99760	0.99767	0.99774	0.99781	0.99788	0.99795	0.99801	0.99807
2.9	0.99813	0.99819	0.99825	0.99831	0.99836	0.99841	0.99846	0.99851	0.99856	0.99861
3.0	0.99865	0.99869	0.99874	0.99878	0.99882	0.99886	0.99889	0.99893	0.99896	0.99900
3.1	0.99903	0.99906	0.99910	0.99913	0.99916	0.99918	0.99921	0.99924	0.99926	0.99929
3.2	0.99931	0.99934	0.99936	0.99938	0.99940	0.99942	0.99944	0.99946	0.99948	0.99950
3.3	0.99952	0.99953	0.99955	0.99957	0.99958	0.99960	0.99961	0.99962	0.99964	0.99965
3.4	0.99966	0.99968	0.99969	0.99970	0.99971	0.99972	0.99973	0.99974	0.99975	0.99976
3.5	0.99977	0.99978	0.99978	0.99979	0.99980	0.99981	0.99981	0.99982	0.99983	0.99983
3.6	0.99984	0.99985	0.99985	0.99986	0.99986	0.99987	0.99987	0.99988	0.99988	0.99989
3.7	0.99989	0.99990	0.99990	0.99990	0.99991	0.99991	0.99992	0.99992	0.99992	0.99992
3.8	0.99993	0.99993	0.99993	0.99994	0.99994	0.99994	0.99994	0.99995	0.99995	0.99995
3.9	0.99995	0.99995	0.99996	0.99996	0.99996	0.99996	0.99996	0.99996	0.99997	0.99997

4. Se si dispone di una tavola che fornisce il valore dell'integrale $\int_0^z f(t)\,dt$ si sfrutta l'interpretazione geometrica per arrivare al risultato desiderato

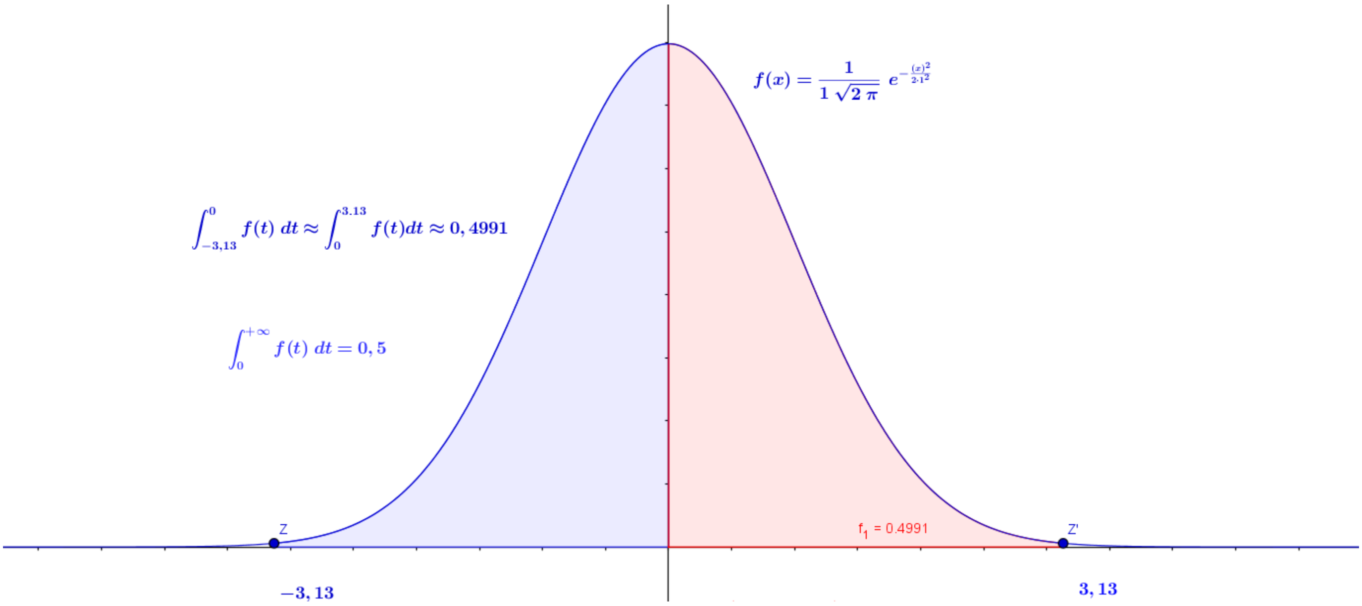
La parte annerita rappresenta l'area sottostante la distribuzione normale standardizzata dalla media aritmetica a z.

z	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	00000	00399	00792	01197	01595	01994	02392	02790	03188	03586
0,1	03983	04380	04776	05172	05567	05962	06356	06749	07142	07535
0,2	07926	08317	08706	09095	09483	09871	10257	10642	11026	11409
0,3	11791	12172	12552	12930	13307	13683	14058	14431	14803	15173
0,4	15542	15910	16276	16640	17003	17364	17724	18082	18439	18793
0,5	19146	19497	19847	20194	20540	20884	21226	21566	21904	22240
0,6	22575	22907	23237	23565	23891	24215	24537	24857	25175	25490
0,7	25804	26115	26424	26730	27035	27337	27637	27935	28230	28524
0,8	28814	29103	29389	29673	29955	30234	30511	30785	31057	31327
0,9	31594	31859	32121	32381	32639	32';94	33147	33398	33646	33891
1,0	34134	34375	34614	34849	35083	35314	35543	35769	35993	36214
1,1	36433	36650	36864	37076	37286	37493	37698	37900	38100	38298
1,2	38493	38686	38877	39065	39251	39435	39617	39796	39973	40147
1,3	40320	40490	40658	40824	40988	41149	41309	41466	41621	41774
1,4	41924	42073	42220	42364	42507	42647	42786	42922	43056	43189
1,5	43319	43448	43574	43699	43822	43943	44062	44179	44295	44408
1,6	44520	44630	44738	44845	44950	45053	45154	45254	45352	45449
1,7	45543	45637	45728	45818	45907	45994	46080	46164	46246	46327
1,8	46407	46485	46562	46637	46712	46784	46856	46926	46995	47062
1.9	47128	47193	47257	47320	47381	47441	47500	47558	47615	47670
2,0	47725	47778	47831	47882	47932	47982	48030	48077	48124	48169
2,1	48214	48257	48300	48341	48382	48422	48461	48500	48537	48574
2,2	48610	48645	48679	48713	48745	48778	48809	48840	48870	48899
2,3	48928	48956	48983	49010	49036	49061	49086	49111	49134	49158
2,4	49180	49202	49224	49245	49266	49286	49305	49324	49343	49361
2,5	49379	49396	49413	49430	49446	49461	49477	49492	49506	49520
2,6	49534	49547	49560	49573	49585	49598	49609	49621	49632	49643
2,7	49653	49664	49674	49683	49693	49702	49711	49720	49728	49736
2,8	49745	49752	49760	49767	49774	49781	49788	49795	49801	49807
2,9	49813	49819	49825	49831	49836	49841	49846	49851	49856	49861
3,0	49865	49869	49874	49878	49882	49886	49889	49893	49897	49900
3,1	49903	49906	49910	49913	49916	49918	49921	49924	49926	49929
3,2	49931	49934	49936	49938	49940	49942	49944	49946	49948	49950
3,3	49952	49953	49955	49957	49958	49960	49961	49962	49964	49965
3,4	49966	49968	49969	49970	49971	49972	49973	49974	49975	49976
3,5	49977	49978	49978	49979	49980	49981	49981	49982	49983	49983
3,6	49984	49985	49985	49986	49986	49987	49987	49988	49988	49989
3,7	49989	49990	49990	49990	49991	49991	49991	49992	49992	49992
3,8	49993	49993	49993	49994	49994	49994	49994	49995	49995	49995
3,9	49995	49995	49995	49996	49996	49996	49996	49996	49997	49997

La tavola fornisce il risultato $\int_0^{3,13} f(t)dt = 0{,}49913$ che è uguale . per simmetria al valore di

$$\int_{-3,13}^{0} f(t)dt$$

al quale va aggiunto il valore di $\int_0^{+\infty} f(t)dt = 0{,}5$



Pertanto si ritrova il risultato del primo metodo $\int_{-3,13}^{+\infty} f(t)dt \approx 0.9991$

La probabilità che tra 500 alunni ce ne siano più di 50 influenzati è maggiore del 99%.