Esercizio

Un'indagine su un campione di $n = 100$ studenti, che hanno sostenuto la prova scritta di matematica, ha prodotto il seguente risultato. Le votazioni in centesimi sono state raggruppate in quattro *classi*.

classe (voto in centesimi)	f_i	f_i/n
20 – 40	10	0.10
40 – 60	20	0.20
60 – 80	50	0.50
80 – 100	20	0.20
	100	1.00

Calcolare media e varianza. Usando l'istogramma delle frequenze o l'ogiva di frequenza, calcolare la mediana.

Le classi sono di uguale ampiezza e contigue.

Esercizio

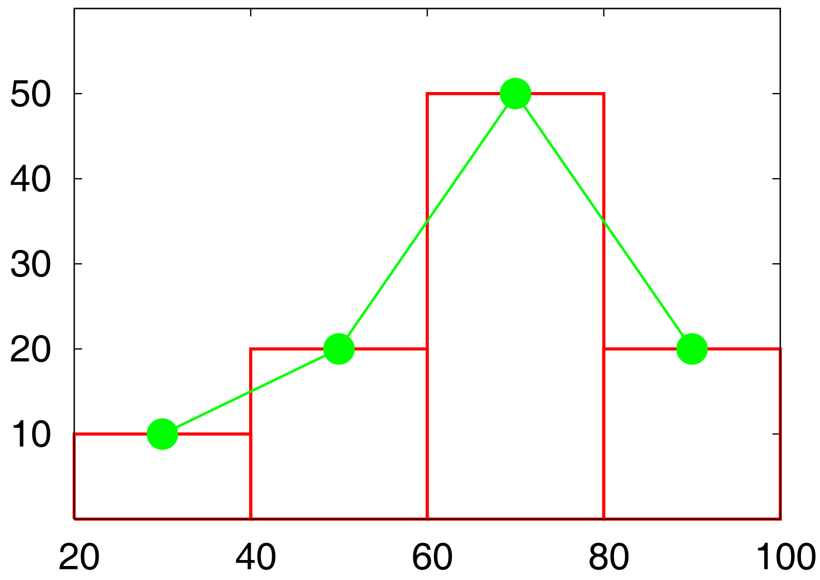
Nell'ipotesi di *distribuzione uniforme*, è naturale associare a ciascuna classe, come *rappresentante*, il valore centrale r_i della classe stessa.

classe	r_i	f_i	F_i
20 – 40	30	10	10
40 – 60	50	20	30
60 – 80	70	50	80
80 – 100	90	20	100

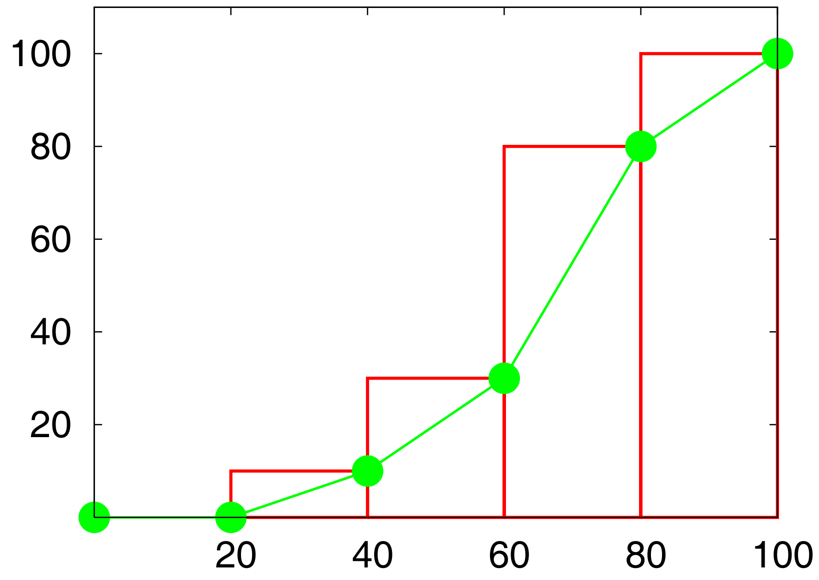
$$\text{media} = \frac{1}{100} \cdot (10 \cdot 30 + 20 \cdot 50 + 50 \cdot 70 + 20 \cdot 90) = 66$$

$$\text{varianza} = \frac{1}{100} \cdot (10 \cdot 36^2 + 20 \cdot 16^2 + 50 \cdot 4^2 + 20 \cdot 24^2) = 304$$

Esercizio



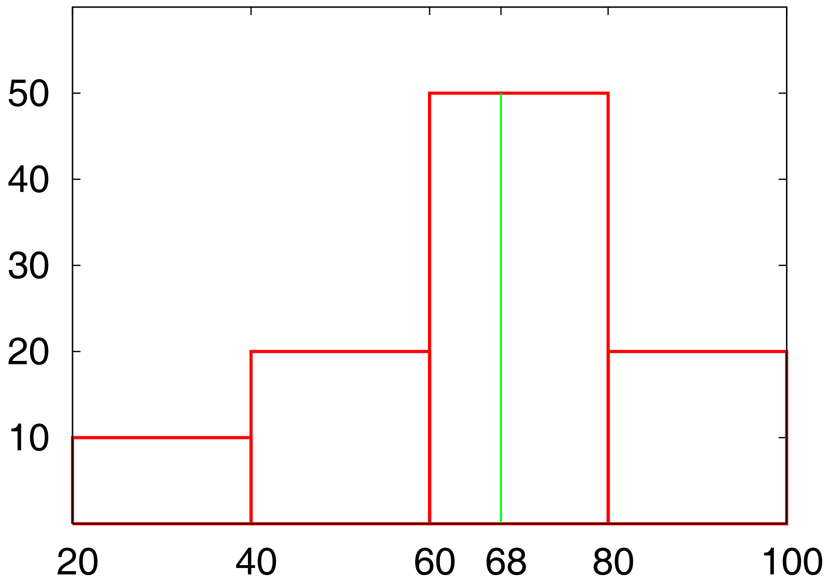
Poligono delle frequenze



Ogiva di frequenza

Esercizio

Calcolo della mediana: con l'istogramma delle frequenze



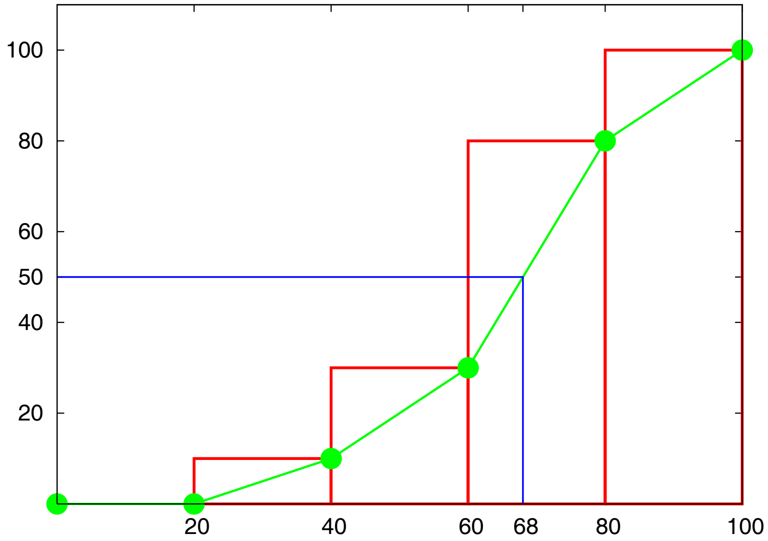
area totale $= 20 \cdot (10 + 20 + 50 + 20) = 2000$

Cerchiamo il valore $x = M_e$ tale che

$$20 \cdot 10 + 20 \cdot 20 + (x - 60) \cdot 50 = 1000 \quad \Rightarrow \quad x = 68 \quad \Rightarrow \quad M_e = 68$$

Esercizio

Calcolo della mediana: con l'ogiva di frequenza

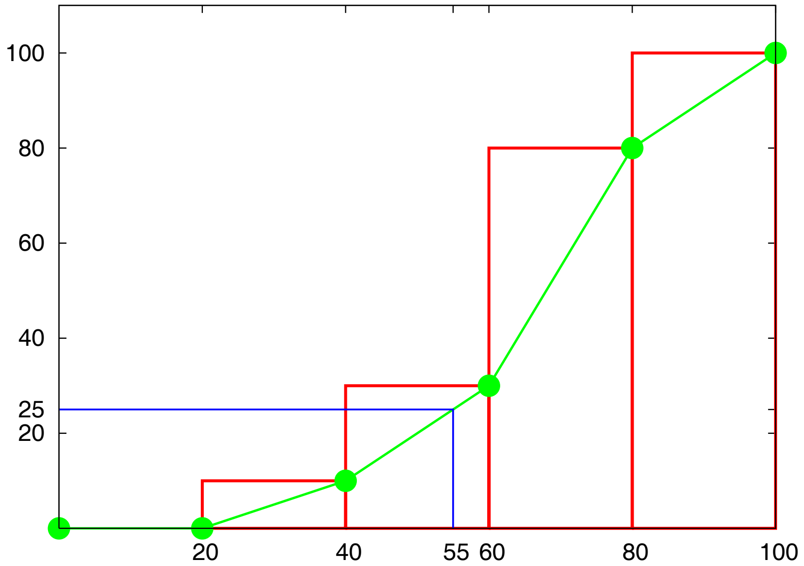


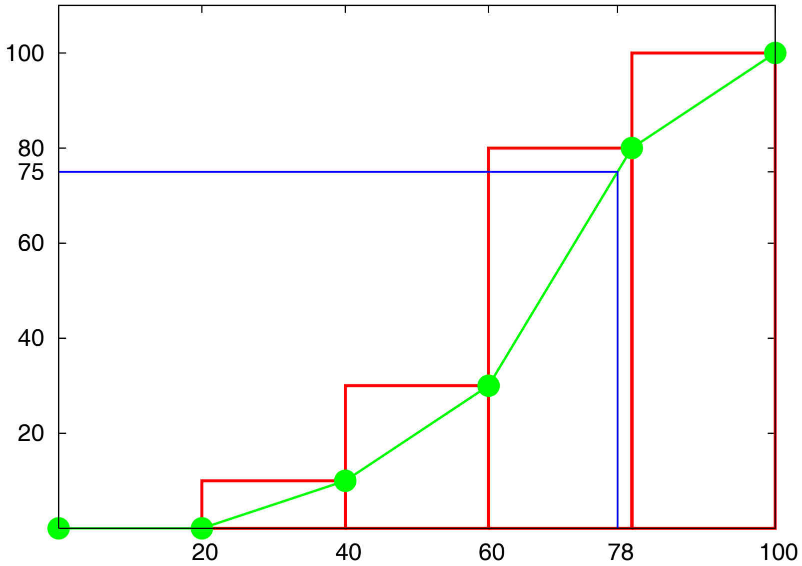
Si considera l'interpolazione lineare sui punti $A = (60, 30)$ e $B = (80, 80)$

$$\begin{cases} y = 50 \\ y - 30 = \frac{5}{2} \cdot (x - 60) \end{cases} \quad \Rightarrow \quad (50 - 30) = \frac{5}{2} \cdot (x - 60) \quad \Rightarrow \quad x = 68 \quad \Rightarrow \quad M_e = 68$$

Esercizio

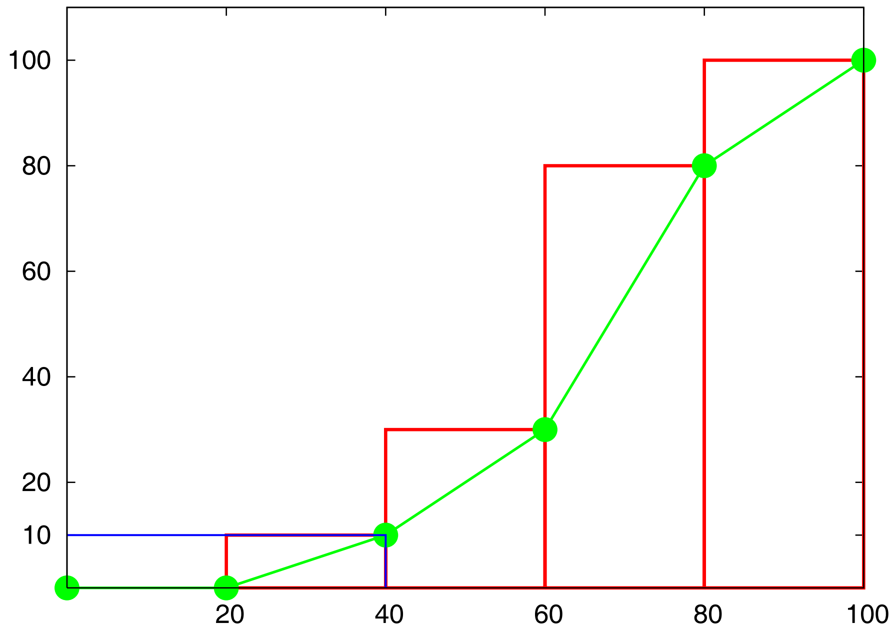
Calcolare i quartili dall'ogiva di frequenza. $q_1 = 55\,,\ q_3 = 78$



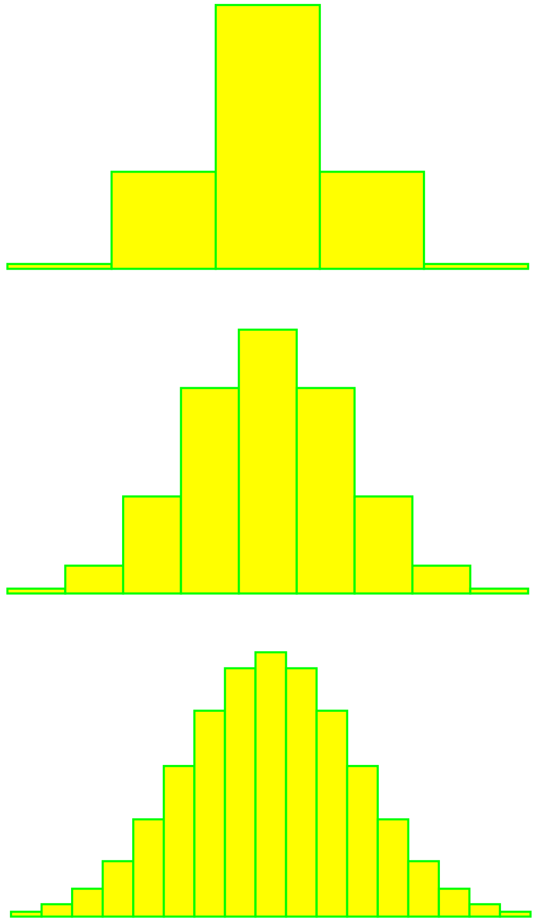


Esercizio

Qual è il voto minimo che bisogna aver preso per non far parte del 10% degli studenti peggiori? risposta: 40



Distribuzione Normale

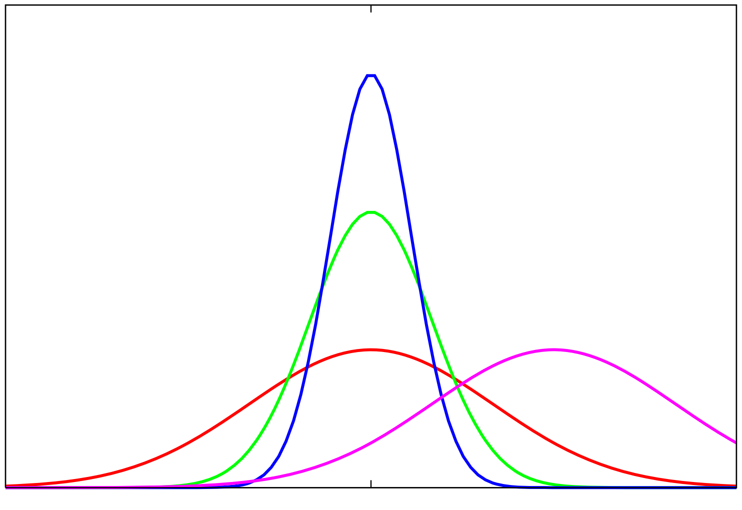


- istogramma delle frequenze di un insieme di misure relative a una grandezza che può variare con continuità
- popolazione molto numerosa, costituita da una quantità praticamente illimitata di individui (popolazione infinita)
- area dell'istogramma uguale a 1 (normalizzata)
- aumentando il numero di intervallini $n = 5, 9, 17, \ldots$ l'istogramma tende a stabilizzarsi intorno a una forma limite: la curva di distribuzione delle frequenze
- nel caso in figura: $y = ae^{-b(x-c)^2}$
  distribuzione normale o gaussiana

Distribuzione Normale

Curve Gaussiane

$$y = ae^{-b(x-c)^2}$$



Se la distribuzione è di tipo gaussiano con

- *media aritmetica* μ
- *deviazione standard* σ

si ha

$$a = \frac{1}{\sigma\sqrt{2\pi}} \quad b = \frac{1}{2\sigma^2} \quad c = \mu$$

La corrispondente curva normale sarà

$$y = \frac{1}{\sigma\sqrt{2\pi}} e^{-\dfrac{(x-\mu)^2}{2\sigma^2}}$$

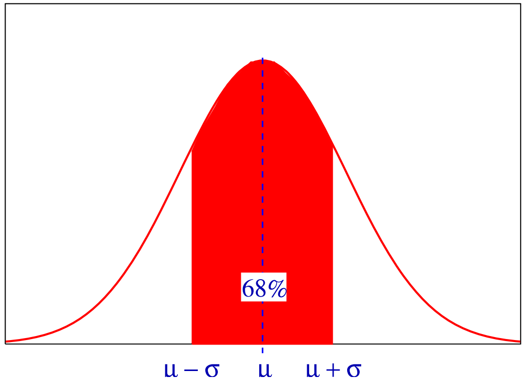
Curva normale standardizzata:

$$y = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \qquad \mu = 0\,, \sigma = 1$$

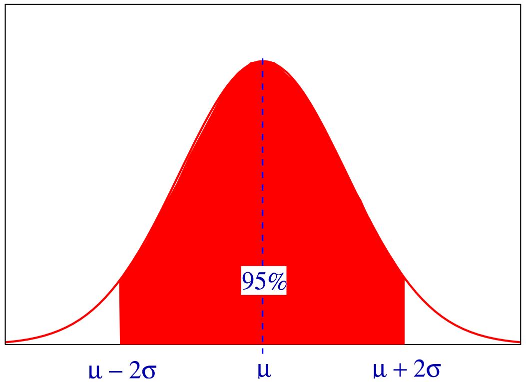
Distribuzione Normale

valori di u	Nell'intervallo $[\mu - u\sigma, \mu + u\sigma]$	Fuori dell'intervallo $[\mu - u\sigma, \mu + u\sigma]$	Nell'intervallo $[\mu + u\sigma, +\infty)$
0	0	1	$0,5$
$0,2$	$0,1586$	$0,8414$	$0,4207$
$0,4$	$0,3108$	$0,6892$	$0,3446$
$0,6$	$0,4514$	$0,5486$	$0,2743$
$0,8$	$0,5762$	$0,4238$	$0,2119$
1	$0,6826$	$0,3174$	$0,1587$
$1,2$	$0,7698$	$0,2302$	$0,1151$
$1,4$	$0,8384$	$0,1616$	$0,0808$
$1,6$	$0,8904$	$0,1096$	$0,0548$
$1,8$	$0,9282$	$0,0718$	$0,0359$
2	$0,9544$	$0,0456$	$0,0228$
$2,2$	$0,9722$	$0,0278$	$0,0139$
$2,4$	$0,9836$	$0,0164$	$0,0082$
$2,6$	$0,9906$	$0,0094$	$0,0047$
$2,8$	$0,9950$	$0,0050$	$0,0025$
3	$0,9974$	$0,0026$	$0,0013$
$3,2$	$0,9986$	$0,0014$	$0,0007$

Fissati due valori x_0, x_1 sull'asse delle ascisse, l'area sottesa dal grafico sull'intervallo $[x_0, x_1]$ rappresenta la porzione di misure che cadono nell'intervallo considerato.



Nell'intervallo $[\mu - \sigma, \mu + \sigma]$ cade circa il 68% delle misure



Nell'intervallo $[\mu - 2\sigma, \mu + 2\sigma]$ cade circa il 95% delle misure

Distribuzione Normale – Esercizi

Esercizio 1. Supponendo che la distribuzione dei pesi degli individui di una popolazione sia gaussiana con media $\mu = 61\,\text{Kg}$ e deviazione standard *(scarto quadratico medio)* $\sigma = 5\,\text{Kg}$,

(a) scrivere l'equazione della gaussiana relativa ai pesi di tale popolazione;

(b) calcolare la percentuale di individui il cui peso è compreso tra 59 Kg e 63 Kg.

Soluzione:

(a) L'equazione della gaussiana è la seguente

$$y = \frac{1}{5\sqrt{2\pi}} \cdot e^{-\frac{(x-61)^2}{50}}$$

(b) L'area sotto il grafico di tale gaussiana nell'intervallo di peso

$$[59\,,\,63] = [61 - 0.4 \cdot 5\,,\, 61 + 0.4 \cdot 5] = [\mu - 0.4\,\sigma\,,\, \mu + 0.4\,\sigma]$$

vale 0.3108. La percentuale cercata è del 31% circa.

Distribuzione Normale – Esercizi

Esercizio 2. Le altezze h di un gruppo di reclute sono distribuite con buona approssimazione secondo una curva gaussiana con media $\mu = 170\,\text{cm}$ e deviazione standard *(scarto quadratico)* $\sigma = 5\,\text{cm}$. Le divise sono disponibili in 5 taglie:

1. per individui di altezza $\leq 161\,\text{cm}$
2. per individui di altezza compresa tra 161 e $167\,\text{cm}$
3. per individui di altezza compresa tra 167 e $173\,\text{cm}$
4. per individui di altezza compresa tra 173 e $179\,\text{cm}$
5. per individui di altezza $> 179\,\text{cm}$.

Stimare il numero delle divise delle varie taglie sapendo che le reclute sono 750.

Soluzione: si tratta di stimare la percentuale di reclute che cade in ciascuna delle quattro differenti classi di altezza:

1. per $h \leq 161 = 170 - 1.8\,\sigma \;\Rightarrow\;$ 3.6% (27 reclute)
2. per $161 < h \leq 167 \;\Rightarrow\; h \in (170 - 1.8\,\sigma\,,\, 170 - 0.6\,\sigma] \;\Rightarrow\;$ 23.8% ($\simeq$ 179 reclute)
3. per $167 < h \leq 173 \;\Rightarrow\; h \in (170 - 0.6\,\sigma\,,\, 170 + 0.6\,\sigma] \;\Rightarrow\;$ 45.1% ($\simeq$ 338 reclute)
4. per $173 < h \leq 179 \;\Rightarrow\; h \in (170 + 0.6\,\sigma\,,\, 170 + 1.8\,\sigma] \;\Rightarrow\;$ 23.8% ($\simeq$ 179 reclute)
5. per $h > 179 = 170 + 1.8\,\sigma \;\Rightarrow\;$ 3.6% (27 reclute)

Tabella Curva Gaussiana

valori di u	Nell'intervallo $[\mu - u\sigma, \mu + u\sigma]$	Fuori dell'intervallo $[\mu - u\sigma, \mu + u\sigma]$	Nell'intervallo $[\mu + u\sigma, +\infty)$
0	0	1	$0,5$
$0,2$	$0,1586$	$0,8414$	$0,4207$
$0,4$	$0,3108$	$0,6892$	$0,3446$
$0,6$	$0,4514$	$0,5486$	$0,2743$
$0,8$	$0,5762$	$0,4238$	$0,2119$
1	$0,6826$	$0,3174$	$0,1587$
$1,2$	$0,7698$	$0,2302$	$0,1151$
$1,4$	$0,8384$	$0,1616$	$0,0808$
$1,6$	$0,8904$	$0,1096$	$0,0548$
$1,8$	$0,9282$	$0,0718$	$0,0359$
2	$0,9544$	$0,0456$	$0,0228$
$2,2$	$0,9722$	$0,0278$	$0,0139$
$2,4$	$0,9836$	$0,0164$	$0,0082$
$2,6$	$0,9906$	$0,0094$	$0,0047$
$2,8$	$0,9950$	$0,0050$	$0,0025$
3	$0,9974$	$0,0026$	$0,0013$
$3,2$	$0,9986$	$0,0014$	$0,0007$

Tabella Curva Gaussiana

aree sottese dalla curva gaussiana sull'intervallo [μ , μ + z σ]

z	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,00	0,0000	0,0040	0,0080	0,0120	0,0160	0,0199	0,0239	0,0279	0,0319	0,0359
0,10	0,0398	0,0438	0,0478	0,0517	0,0557	0,0596	0,0636	0,0675	0,0714	0,0753
0,20	0,0793	0,0832	0,0871	0,0910	0,0948	0,0987	0,1026	0,1064	0,1103	0,1141
0,30	0,1179	0,1217	0,1255	0,1293	0,1331	0,1368	0,1406	0,1443	0,1480	0,1517
0,40	0,1554	0,1591	0,1628	0,1664	0,1700	0,1736	0,1772	0,1808	0,1844	0,1879
0,50	0,1915	0,1950	0,1985	0,2019	0,2054	0,2088	0,2123	0,2157	0,2190	0,2224
0,60	0,2257	0,2291	0,2324	0,2357	0,2389	0,2422	0,2454	0,2486	0,2517	0,2549
0,70	0,2580	0,2611	0,2642	0,2673	0,2704	0,2734	0,2764	0,2794	0,2823	0,2852
0,80	0,2881	0,2910	0,2939	0,2967	0,2995	0,3023	0,3051	0,3078	0,3106	0,3133
0,90	0,3159	0,3186	0,3212	0,3238	0,3264	0,3289	0,3315	0,3340	0,3365	0,3389
1,00	0,3413	0,3438	0,3461	0,3485	0,3508	0,3531	0,3554	0,3577	0,3599	0,3621
1,10	0,3643	0,3665	0,3686	0,3708	0,3729	0,3749	0,3770	0,3790	0,3810	0,3830
1,20	0,3849	0,3869	0,3888	0,3907	0,3925	0,3944	0,3962	0,3980	0,3997	0,4015
1,30	0,4032	0,4049	0,4066	0,4082	0,4099	0,4115	0,4131	0,4147	0,4162	0,4177
1,40	0,4192	0,4207	0,4222	0,4236	0,4251	0,4265	0,4279	0,4292	0,4306	0,4319
1,50	0,4332	0,4345	0,4357	0,4370	0,4382	0,4394	0,4406	0,4418	0,4429	0,4441
1,60	0,4452	0,4463	0,4474	0,4484	0,4495	0,4505	0,4515	0,4525	0,4535	0,4545
1,70	0,4554	0,4564	0,4573	0,4582	0,4591	0,4599	0,4608	0,4616	0,4625	0,4633
1,80	0,4641	0,4649	0,4656	0,4664	0,4671	0,4678	0,4686	0,4693	0,4699	0,4706
1,90	0,4713	0,4719	0,4726	0,4732	0,4738	0,4744	**0,4750**	0,4756	0,4761	0,4767
2,00	0,4772	0,4778	0,4783	0,4788	0,4793	0,4798	0,4803	0,4808	0,4812	0,4817
2,10	0,4821	0,4826	0,4830	0,4834	0,4838	0,4842	0,4846	0,4850	0,4854	0,4857
2,20	0,4861	0,4864	0,4868	0,4871	0,4875	0,4878	0,4881	0,4884	0,4887	0,4890
2,30	0,4893	0,4896	0,4898	0,4901	0,4904	0,4906	0,4909	0,4911	0,4913	0,4916
2,40	0,4918	0,4920	0,4922	0,4925	0,4927	0,4929	0,4931	0,4932	0,4934	0,4936
2,50	0,4938	0,4940	0,4941	0,4943	0,4945	0,4946	0,4948	0,4949	**0,4951**	0,4952
2,60	0,4953	0,4955	0,4956	0,4957	0,4959	0,4960	0,4961	0,4962	0,4963	0,4964
2,70	0,4965	0,4966	0,4967	0,4968	0,4969	0,4970	0,4971	0,4972	0,4973	0,4974
2,80	0,4974	0,4975	0,4976	0,4977	0,4977	0,4978	0,4979	0,4979	0,4980	0,4981
2,90	0,4981	0,4982	0,4982	0,4983	0,4984	0,4984	0,4985	0,4985	0,4986	0,4986
3,00	0,4987	0,4987	0,4987	0,4988	0,4988	0,4989	0,4989	0,4989	0,4990	0,4990