Let us start chapter 6

Example 1:

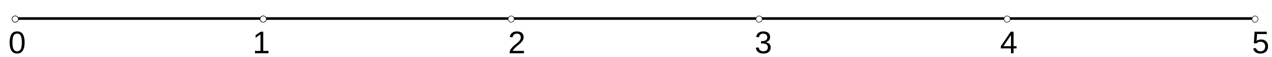
There are 5 employees in a unit.

The random Variable: # of employees that will report for work tomorrow

This is an example of a discrete random variable

Continuous Random Variable:

The temperature outside a certain building is somewhere between 0 and 5 degrees C.



The variable can take on any value on the above number line, **including the whole number values**.

This is an example of a continuous random variable

The probabilities for a continuous random variable are given by the areas under a curve given by a continuous

density function $y = f(x)$.

such that

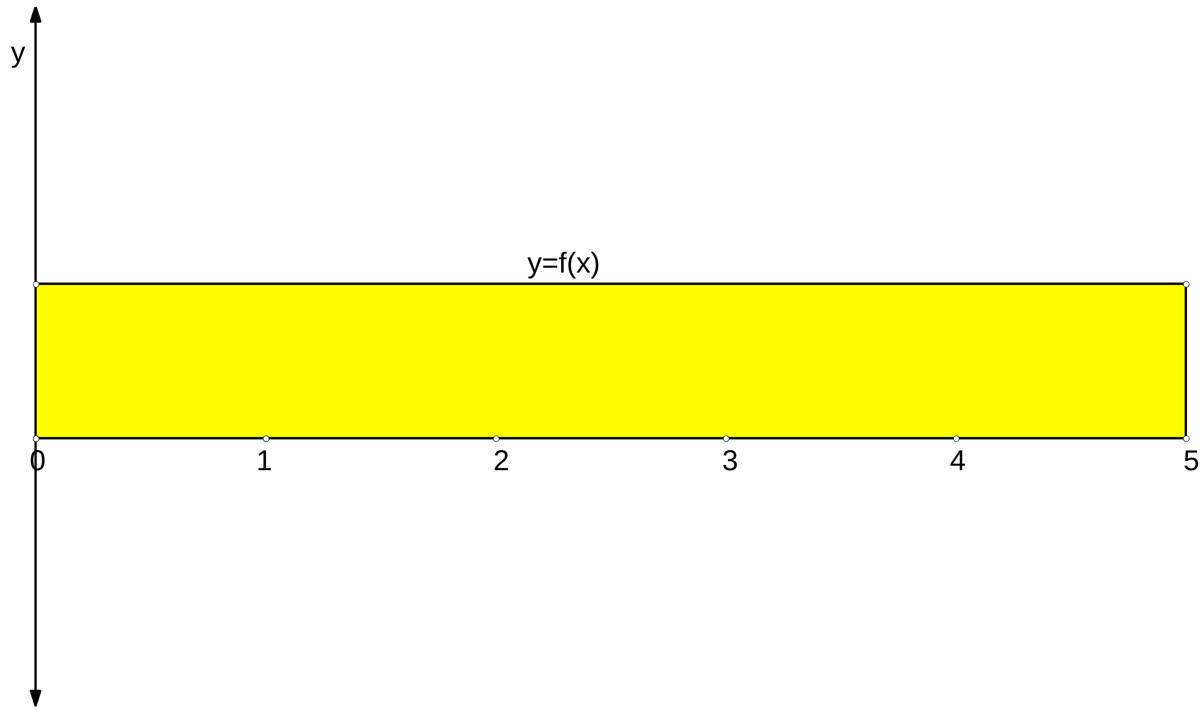
a) $f(x) \geq 0$ **for all x**.
b) **The total area under the curve is 1**.

Example:

A uniform distribution

(**anywhere between the two values uniformly**)

Example



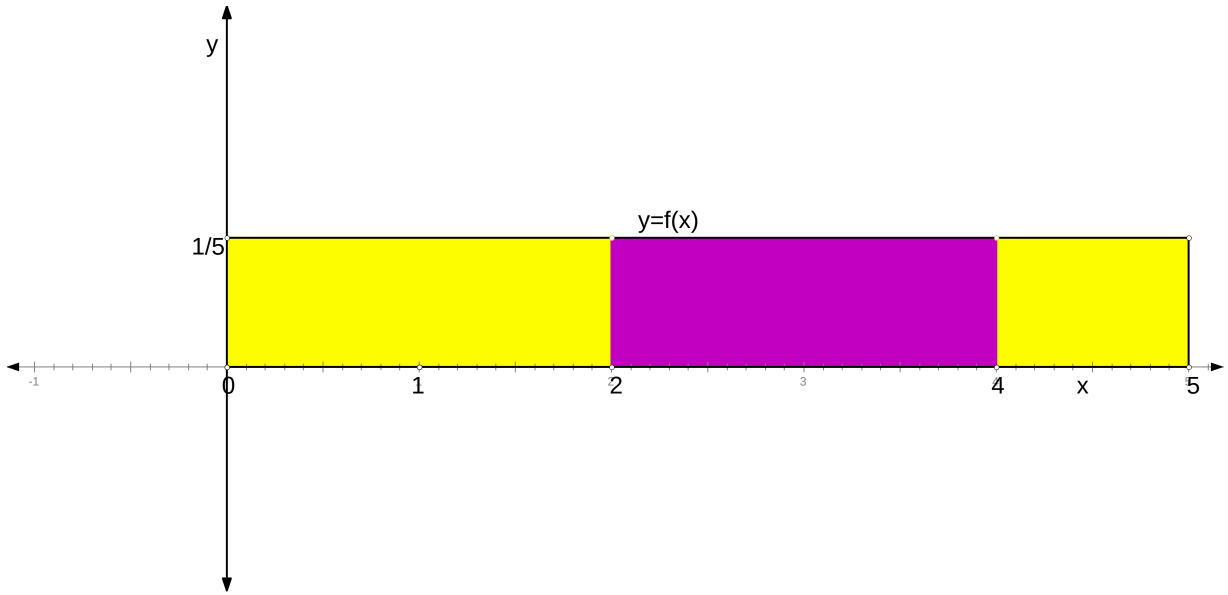
The above shows a uniform distribution on the interval $[0, 5]$

The density curve is a horizontal straight line
In order for the area under the density curve to be 1
the equation of the line has to be $y = \frac{1}{5}$**, that is the density curve is given by** $f(x) = \frac{1}{5}$

The probability of this random variable will take on values between x=a and x=b is given by the area under density curve from x=a to x=b

For example, the area between x=2 and x=4 is

$\mathbf{P}(2 \leq X \leq 4)$
$= 2 \times \frac{1}{5} = 0.4$

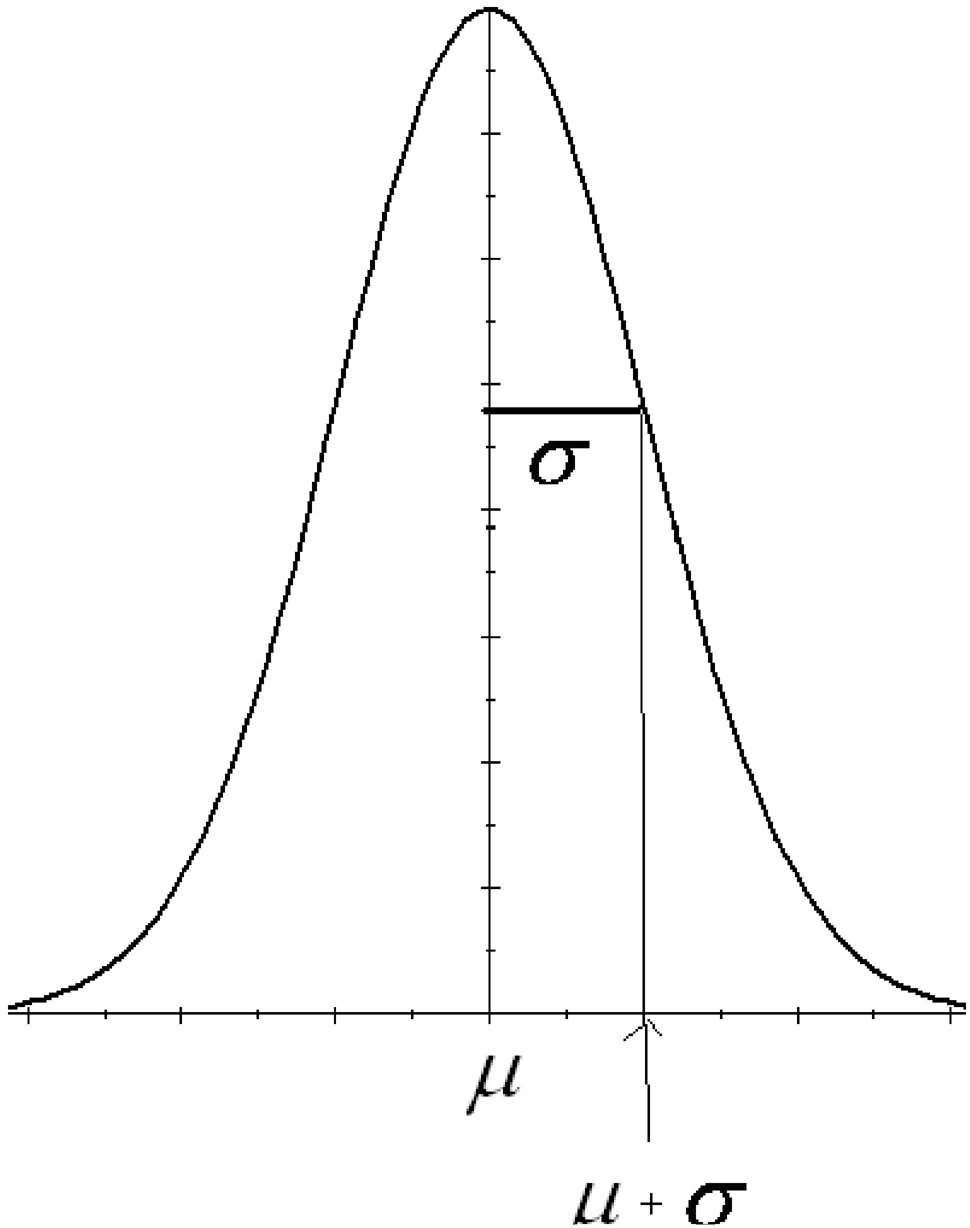


Normal Distribution

The density curve of a normal distribution with mean μ and standard deviation σ is given by

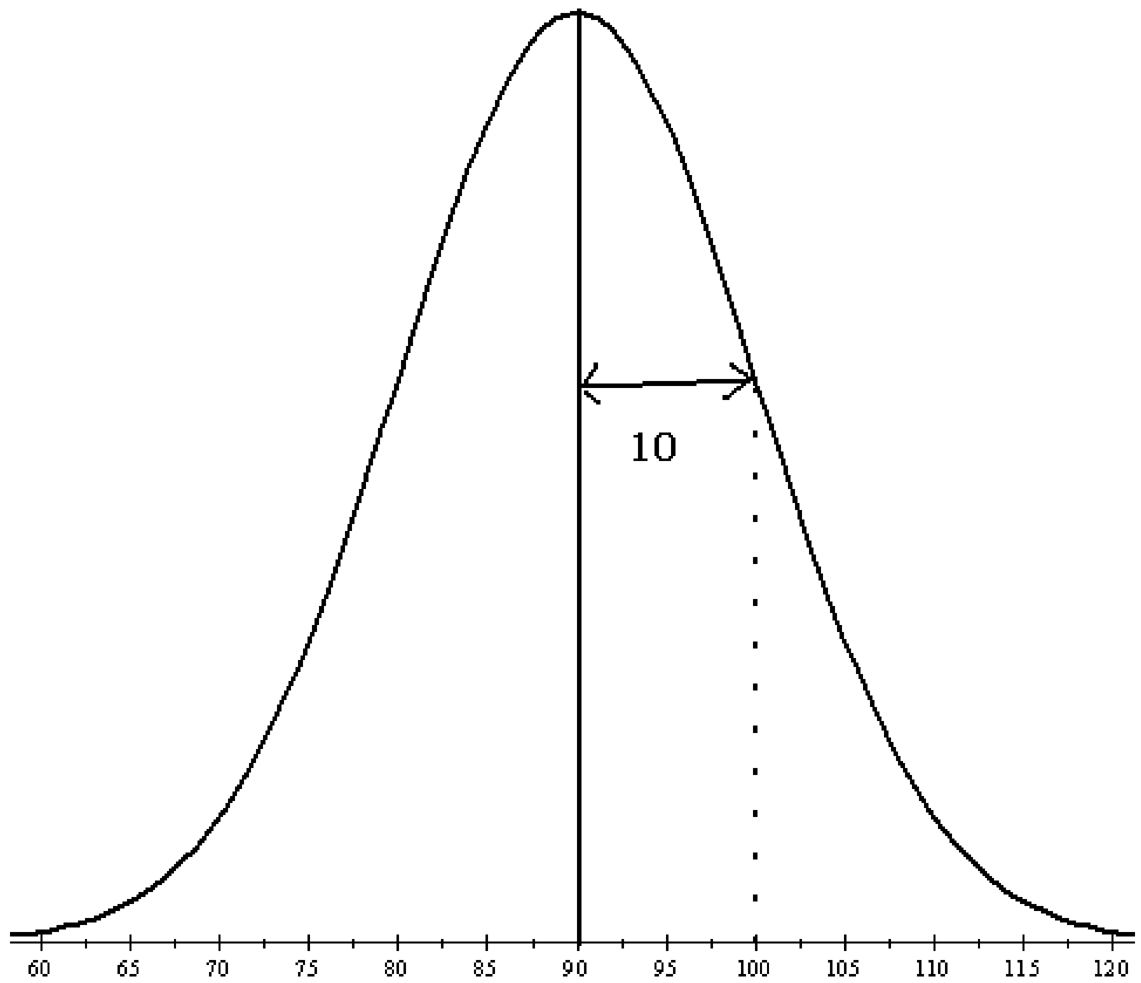
$$\mathbf{f}(\mathbf{x}) = \frac{1}{\sigma\sqrt{2\pi}}\mathbf{e}^{-((x-\mu)^2/(2\times\sigma^2))}$$

Its graph is a bell shaped curved determined by the parameters μ and σ as shown below


σ
μ
μ + σ

Example:

Given that the time that the students take to finish a certain standardized test shows a normal distribution with mean 90 minutes and standard deviation 10 minutes.

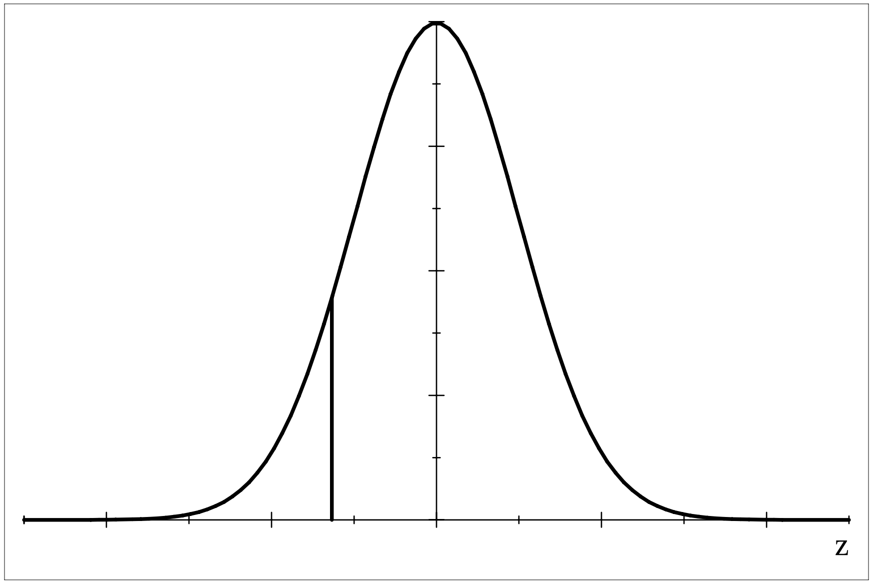


Terminology:

A normal distribution with the mean 0 and standard deviation 1 is called the standard normal distribution. **Standard normal variable is denoted by z**.

$$\mathbf{s(z)} = \frac{1}{\sqrt{2\pi}} \mathbf{e}^{-z^2/2}$$

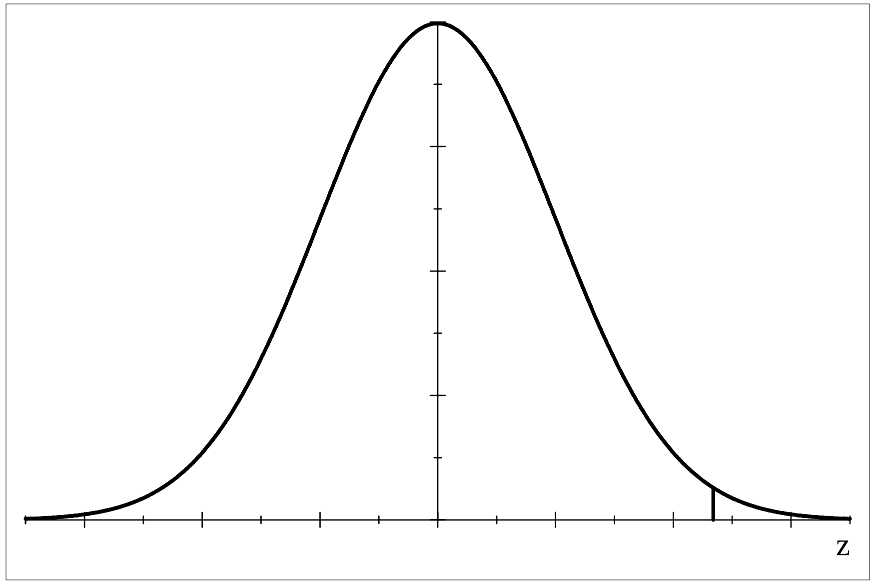
$s(z)$

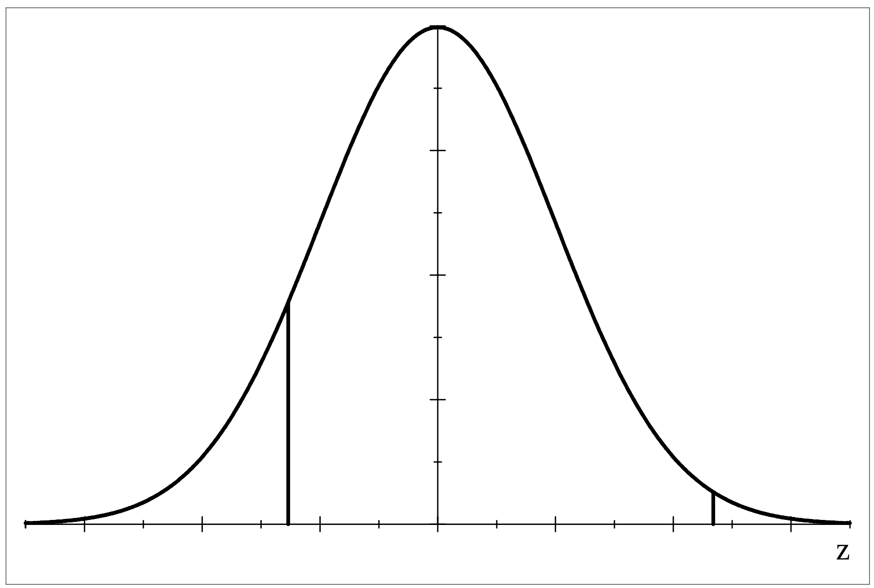


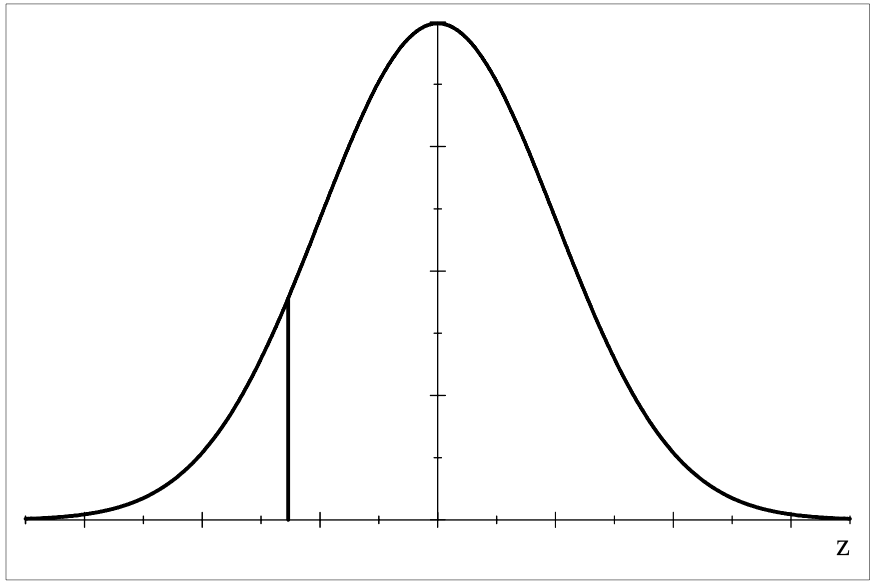
s

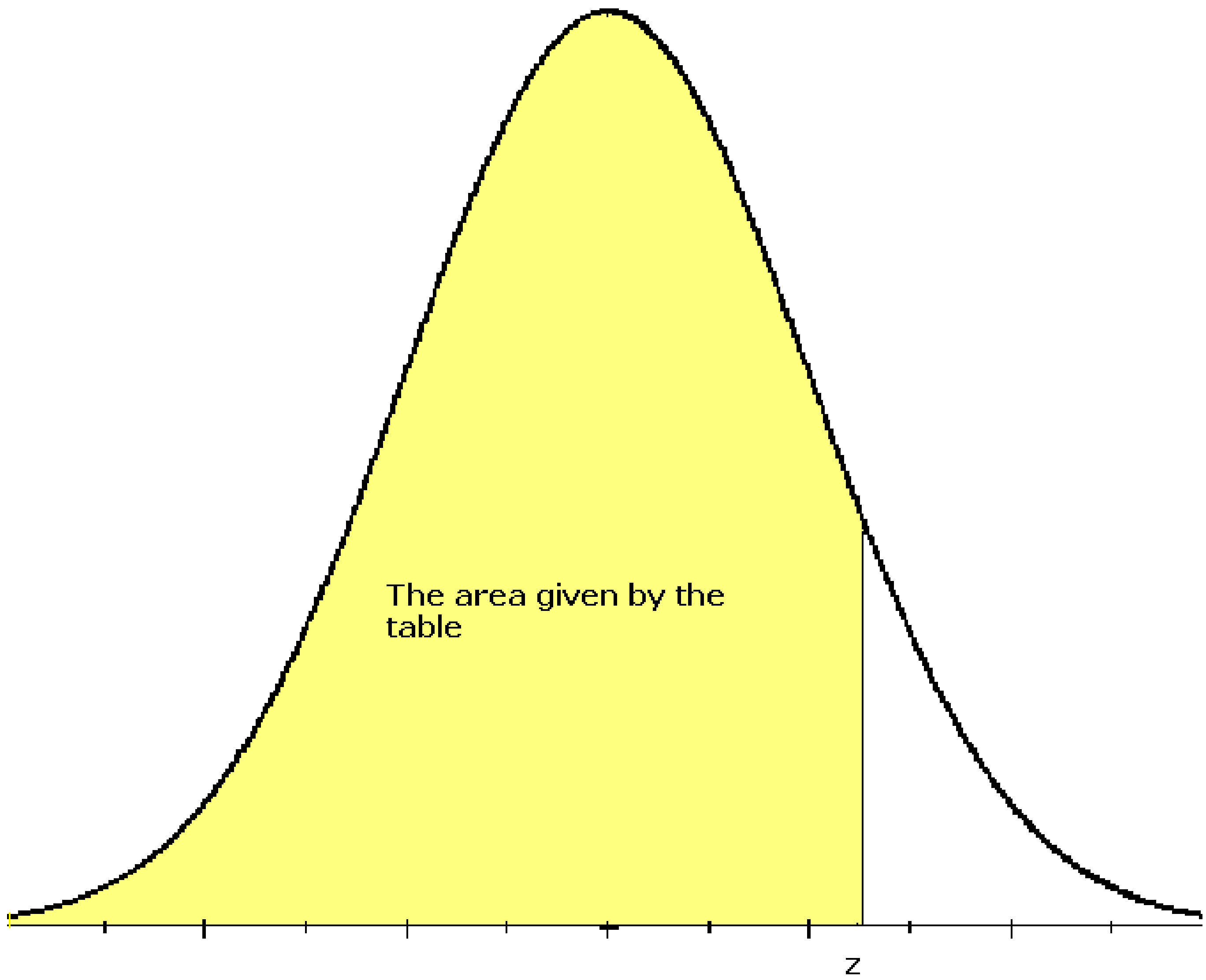
Total area under this curve is 1.

We are going to use a standard normal curve areas table that is attached at the end of this lesson. The table will give us the cummulative areas for a given value of z.


z


z


z

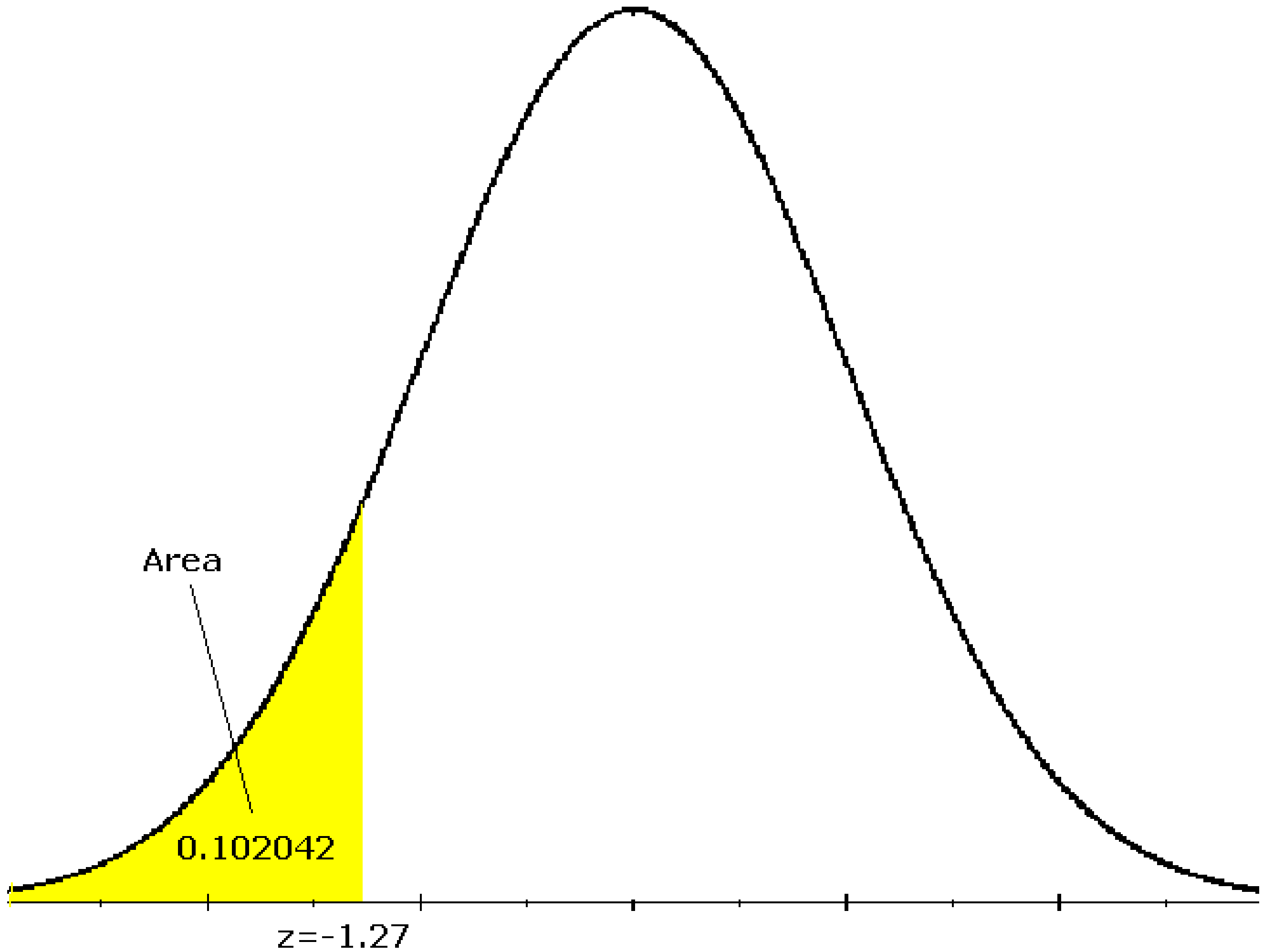

The area given by the table
z

For Example, if we would like to see the area under the standard normal curve to the left of $z = -1.27$

We can go right of the row containing -1.2 and look at the entry corresponding to the column of 0.07

	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
-3	0.00135	0.001306	0.001264	0.001223	0.001183	0.001144	0.001107	0.00107	0.001035	0.001001
-2.9	0.001866	0.001807	0.00175	0.001695	0.001641	0.001589	0.001538	0.001489	0.001441	0.001395
-2.8	0.002555	0.002477	0.002401	0.002327	0.002256	0.002186	0.002118	0.002052	0.001988	0.001926
-2.7	0.003467	0.003364	0.003264	0.003167	0.003072	0.00298	0.00289	0.002803	0.002718	0.002635
-2.6	0.004661	0.004527	0.004396	0.004269	0.004145	0.004025	0.003907	0.003793	0.003681	0.003573
-2.5	0.00621	0.006037	0.005868	0.005703	0.005543	0.005386	0.005234	0.005085	0.00494	0.004799
-2.4	0.008198	0.007976	0.00776	0.007549	0.007344	0.007143	0.006947	0.006756	0.006569	0.006387
-2.3	0.010724	0.010444	0.01017	0.009903	0.009642	0.009387	0.009137	0.008894	0.008656	0.008424
-2.2	0.013903	0.013553	0.013209	0.012874	0.012545	0.012224	0.011911	0.011604	0.011304	0.011011
-2.1	0.017864	0.017429	0.017003	0.016586	0.016177	0.015778	0.015386	0.015003	0.014629	0.014262
-2	0.02275	0.022216	0.021692	0.021178	0.020675	0.020182	0.019699	0.019226	0.018763	0.018309
-1.9	0.028717	0.028067	0.027429	0.026803	0.02619	0.025588	0.024998	0.024419	0.023852	0.023295
-1.8	0.03593	0.035148	0.03438	0.033625	0.032884	0.032157	0.031443	0.030742	0.030054	0.029379
-1.7	0.044565	0.043633	0.042716	0.041815	0.04093	0.040059	0.039204	0.038364	0.037538	0.036727
-1.6	0.054799	0.053699	0.052616	0.051551	0.050503	0.049471	0.048457	0.04746	0.046479	0.045514
-1.5	0.066807	0.065522	0.064255	0.063008	0.06178	0.060571	0.05938	0.058208	0.057053	0.055917
-1.4	0.080757	0.07927	0.077804	0.076359	0.074934	0.073529	0.072145	0.070781	0.069437	0.068112
-1.3	0.0968	0.095098	0.093418	0.091759	0.090123	0.088508	0.086915	0.085343	0.083793	0.082264
-1.2	0.11507	0.113139	0.111232	0.109349	0.107488	0.10565	0.103835	0.102042	0.100273	0.098525
-1.1	0.135666	0.1335	0.131357	0.129238	0.127143	0.125072	0.123024	0.121	0.119	0.117023
-1	0.158655	0.156248	0.153864	0.151505	0.14917	0.146859	0.144572	0.14231	0.140071	0.137857
-0.9	0.18406	0.181411	0.178786	0.176186	0.173609	0.171056	0.168528	0.166023	0.163543	0.161087
-0.8	0.211855	0.20897	0.206108	0.203269	0.200454	0.197663	0.194895	0.19215	0.18943	0.186733
-0.7	0.241964	0.238852	0.235762	0.232695	0.22965	0.226627	0.223627	0.22065	0.217695	0.214764
-0.6	0.274253	0.270931	0.267629	0.264347	0.261086	0.257846	0.254627	0.251429	0.248252	0.245097
-0.5	0.308538	0.305026	0.301532	0.298056	0.294599	0.29116	0.28774	0.284339	0.280957	0.277595
-0.4	0.344578	0.340903	0.337243	0.333598	0.329969	0.326355	0.322758	0.319178	0.315614	0.312067
-0.3	0.382089	0.37828	0.374484	0.3707	0.366928	0.363169	0.359424	0.355691	0.351973	0.348268
-0.2	0.42074	0.416834	0.412936	0.409046	0.405165	0.401294	0.397432	0.39358	0.389739	0.385908
-0.1	0.460172	0.456205	0.452242	0.448283	0.44433	0.440382	0.436441	0.432505	0.428576	0.424655
0	0.5	0.496011	0.492022	0.488034	0.484047	0.480061	0.476078	0.472097	0.468119	0.464144

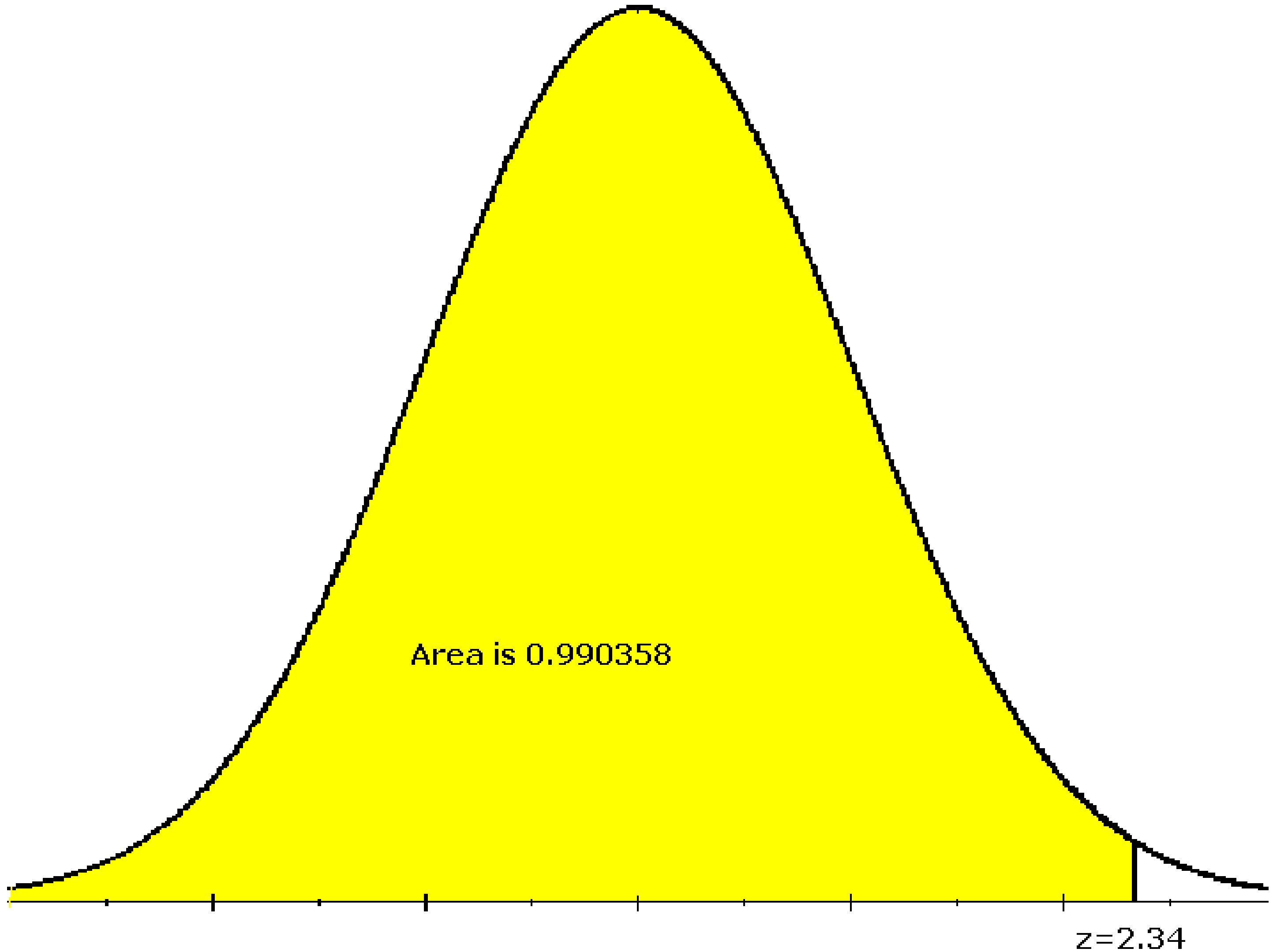
To get the desired answer



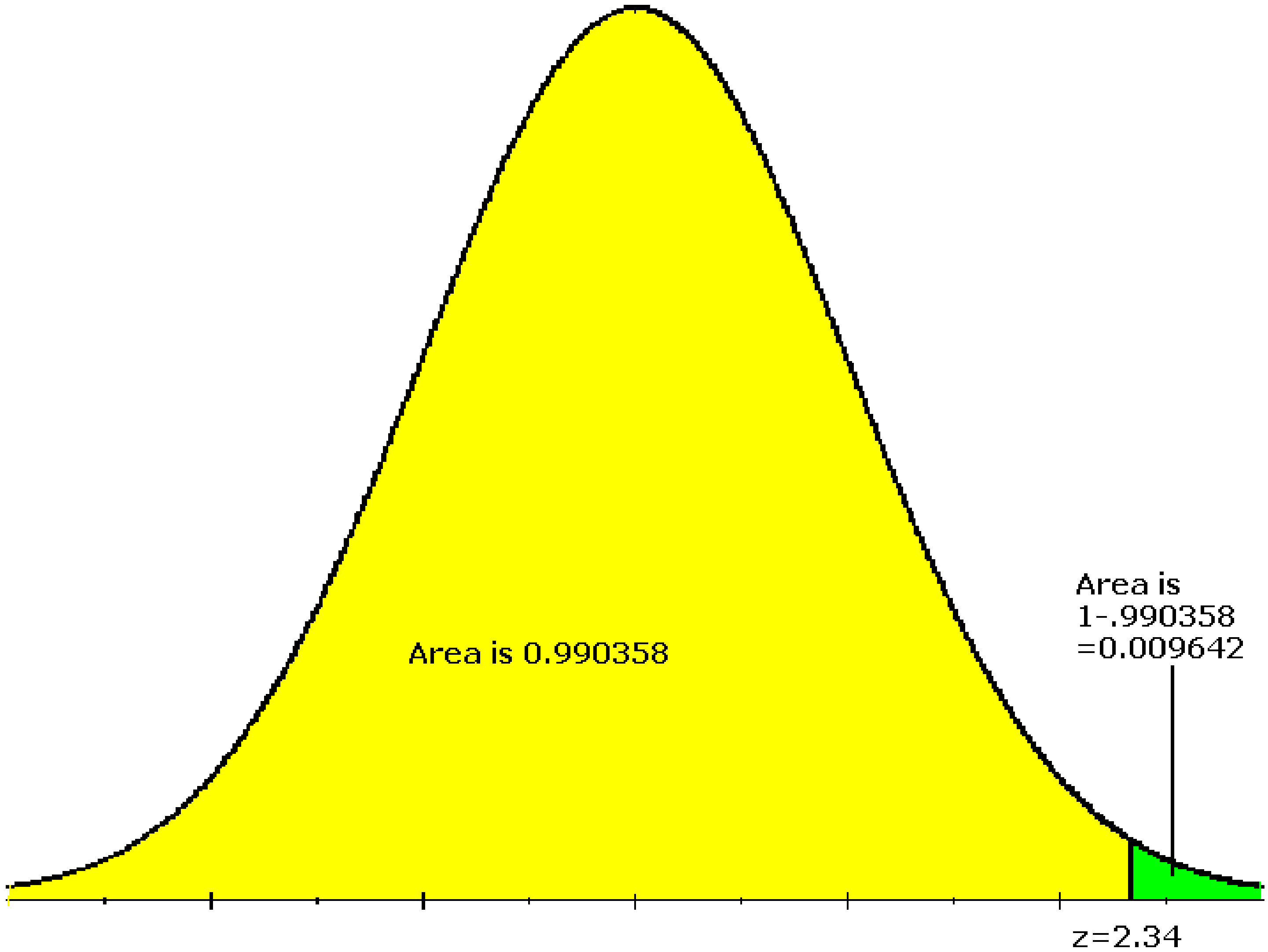
If we would like to get the area under the standard normal curve to the left of $z = 2.34$

We can go right of the row containing 2.3 and look at the entry corresponding to the column of 0.04

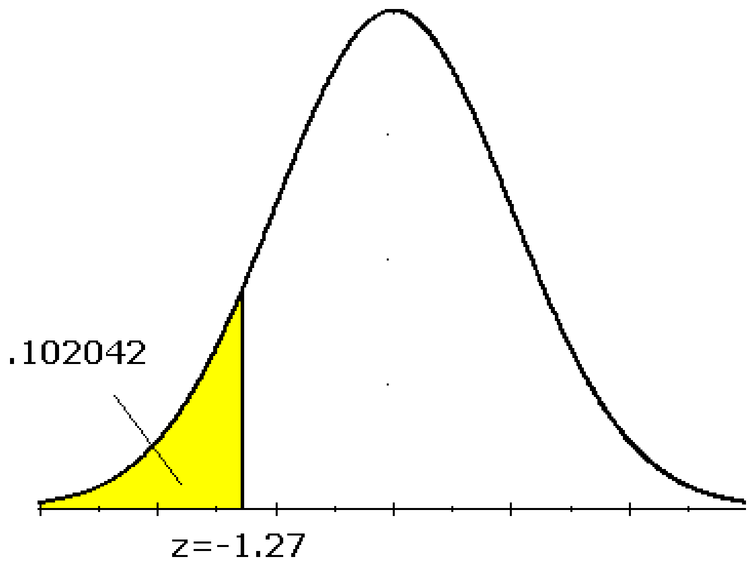
	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0	0.5	0.503989	0.507978	0.511966	0.515953	0.519939	0.523922	0.527903	0.531881	0.535856
0.1	0.539828	0.543795	0.547758	0.551717	0.55567	0.559618	0.563559	0.567495	0.571424	0.575345
0.2	0.57926	0.583166	0.587064	0.590954	0.594835	0.598706	0.602568	0.60642	0.610261	0.614092
0.3	0.617911	0.62172	0.625516	0.6293	0.633072	0.636831	0.640576	0.644309	0.648027	0.651732
0.4	0.655422	0.659097	0.662757	0.666402	0.670031	0.673645	0.677242	0.680822	0.684386	0.687933
0.5	0.691462	0.694974	0.698468	0.701944	0.705401	0.70884	0.71226	0.715661	0.719043	0.722405
0.6	0.725747	0.729069	0.732371	0.735653	0.738914	0.742154	0.745373	0.748571	0.751748	0.754903
0.7	0.758036	0.761148	0.764238	0.767305	0.77035	0.773373	0.776373	0.77935	0.782305	0.785236
0.8	0.788145	0.79103	0.793892	0.796731	0.799546	0.802337	0.805105	0.80785	0.81057	0.813267
0.9	0.81594	0.818589	0.821214	0.823814	0.826391	0.828944	0.831472	0.833977	0.836457	0.838913
1	0.841345	0.843752	0.846136	0.848495	0.85083	0.853141	0.855428	0.85769	0.859929	0.862143
1.1	0.864334	0.8665	0.868643	0.870762	0.872857	0.874928	0.876976	0.879	0.881	0.882977
1.2	0.88493	0.886861	0.888768	0.890651	0.892512	0.89435	0.896165	0.897958	0.899727	0.901475
1.3	0.9032	0.904902	0.906582	0.908241	0.909877	0.911492	0.913085	0.914657	0.916207	0.917736
1.4	0.919243	0.92073	0.922196	0.923641	0.925066	0.926471	0.927855	0.929219	0.930563	0.931888
1.5	0.933193	0.934478	0.935745	0.936992	0.93822	0.939429	0.94062	0.941792	0.942947	0.944083
1.6	0.945201	0.946301	0.947384	0.948449	0.949497	0.950529	0.951543	0.95254	0.953521	0.954486
1.7	0.955435	0.956367	0.957284	0.958185	0.95907	0.959941	0.960796	0.961636	0.962462	0.963273
1.8	0.96407	0.964852	0.96562	0.966375	0.967116	0.967843	0.968557	0.969258	0.969946	0.970621
1.9	0.971283	0.971933	0.972571	0.973197	0.97381	0.974412	0.975002	0.975581	0.976148	0.976705
2	0.97725	0.977784	0.978308	0.978822	0.979325	0.979818	0.980301	0.980774	0.981237	0.981691
2.1	0.982136	0.982571	0.982997	0.983414	0.983823	0.984222	0.984614	0.984997	0.985371	0.985738
2.2	0.986097	0.986447	0.986791	0.987126	0.987455	0.987776	0.988089	0.988396	0.988696	0.988989
2.3	0.989276	0.989556	0.98983	0.990097	0.990358	0.990613	0.990863	0.991106	0.991344	0.991576
2.4	0.991802	0.992024	0.99224	0.992451	0.992656	0.992857	0.993053	0.993244	0.993431	0.993613
2.5	0.99379	0.993963	0.994132	0.994297	0.994457	0.994614	0.994766	0.994915	0.99506	0.995201
2.6	0.995339	0.995473	0.995604	0.995731	0.995855	0.995975	0.996093	0.996207	0.996319	0.996427
2.7	0.996533	0.996636	0.996736	0.996833	0.996928	0.99702	0.99711	0.997197	0.997282	0.997365
2.8	0.997445	0.997523	0.997599	0.997673	0.997744	0.997814	0.997882	0.997948	0.998012	0.998074
2.9	0.998134	0.998193	0.99825	0.998305	0.998359	0.998411	0.998462	0.998511	0.998559	0.998605
3	0.99865	0.998694	0.998736	0.998777	0.998817	0.998856	0.998893	0.99893	0.998965	0.998999

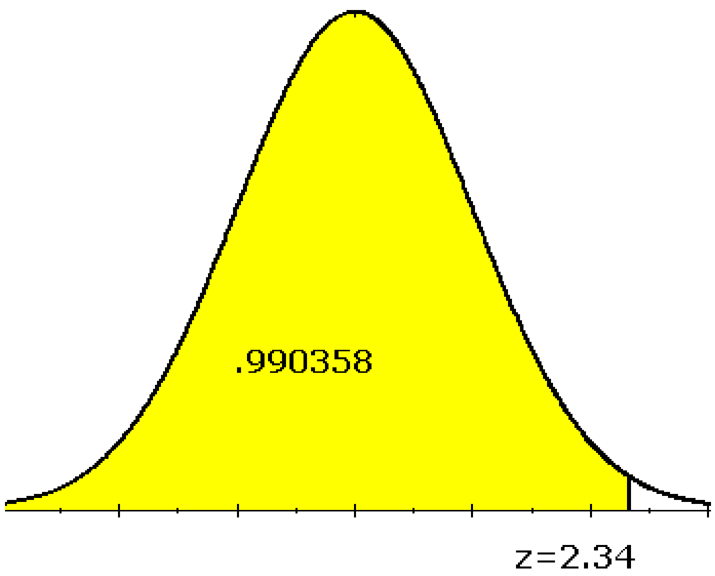


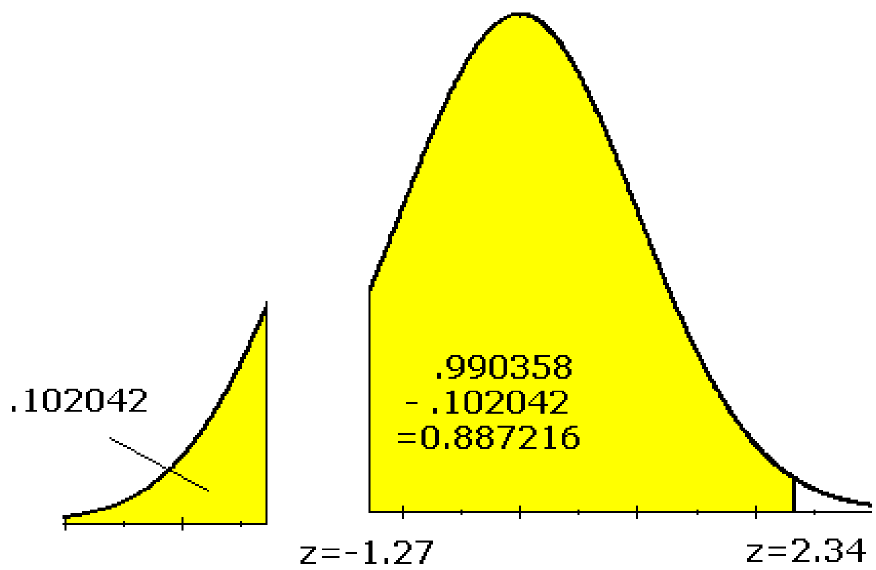
The area under the z-curve to the right of $z = 2.34$ **is** $1 - .990358 = 0.009642$



Now let us find the area under the normal curve between $z = -1.27$ and $z = 2.34$


.102042
z=-1.27


.990358
z=2.34

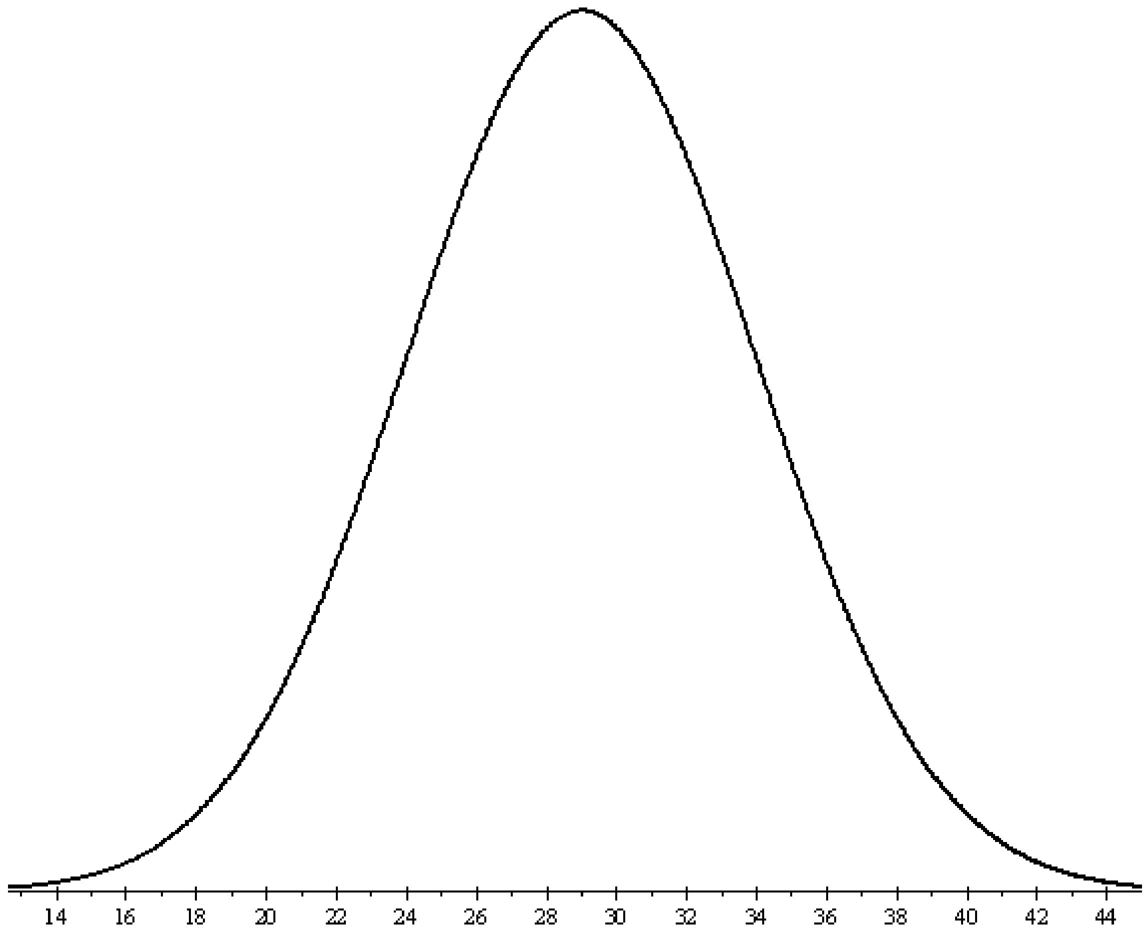

.102042
.990358
- .102042
=0.887216
z=-1.27
z=2.34

We can standardize a variable that follows a normal distribution to srandard normal distribution by using the conversion rule

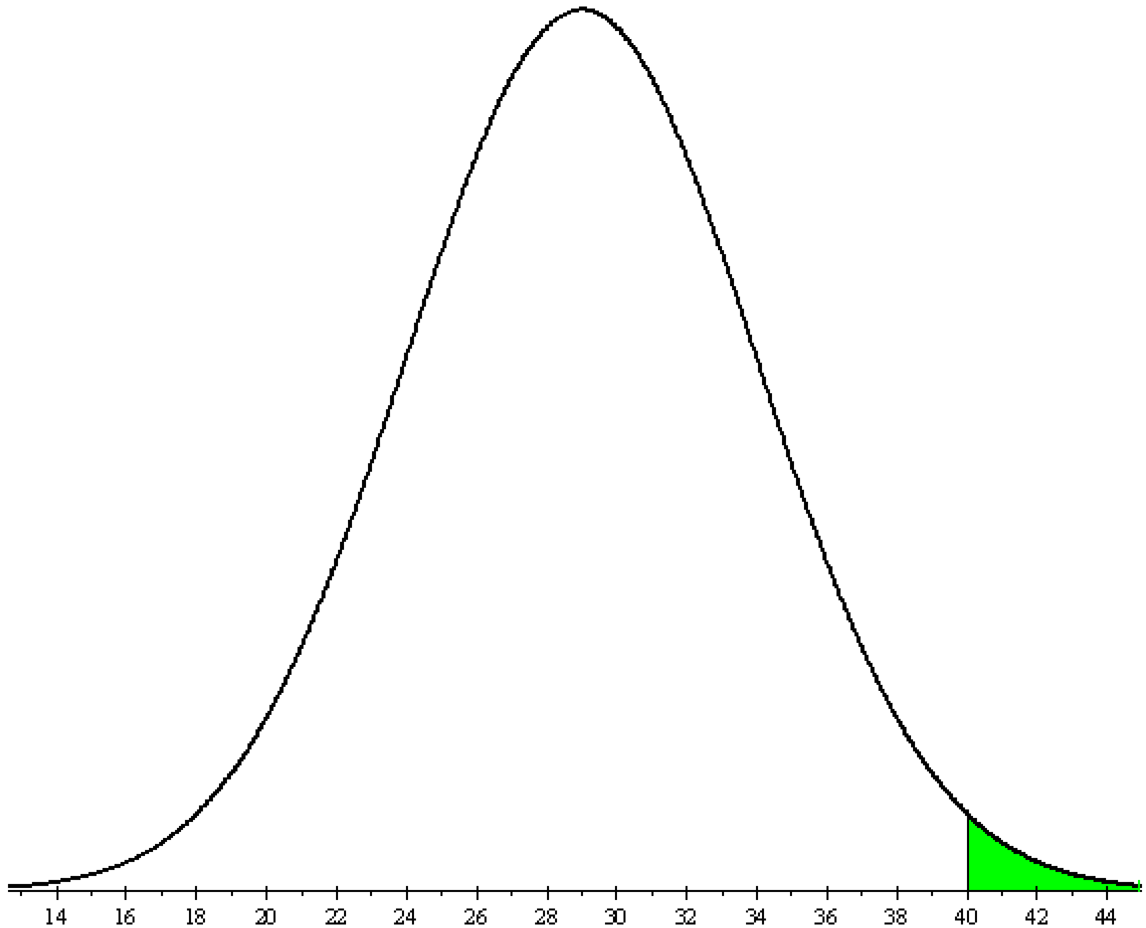
$$\boxed{z = \frac{x - \mu}{\sigma}}$$

Let us use this conversion rule to answer questions about any normal distribution
Example:

Scores on a certain performance test for document analysists show approximately a normal distribution with mean 29 points and standard deviation 6 points.



a) To find the percentage of analysisists who scored more than 40 points.



First, **we shall find the standard score for** $x = 40$

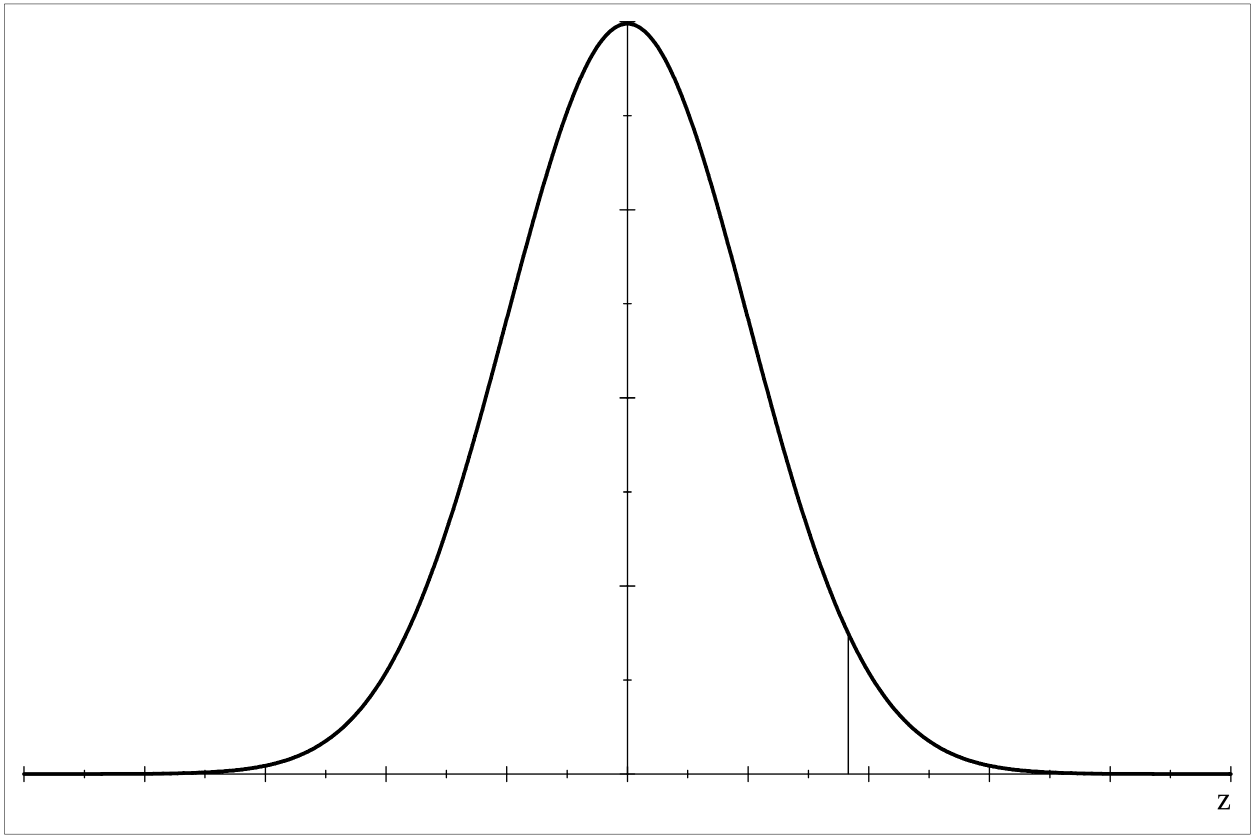
$$\mathbf{z} = \frac{40-29}{6} = \mathbf{1.83}$$

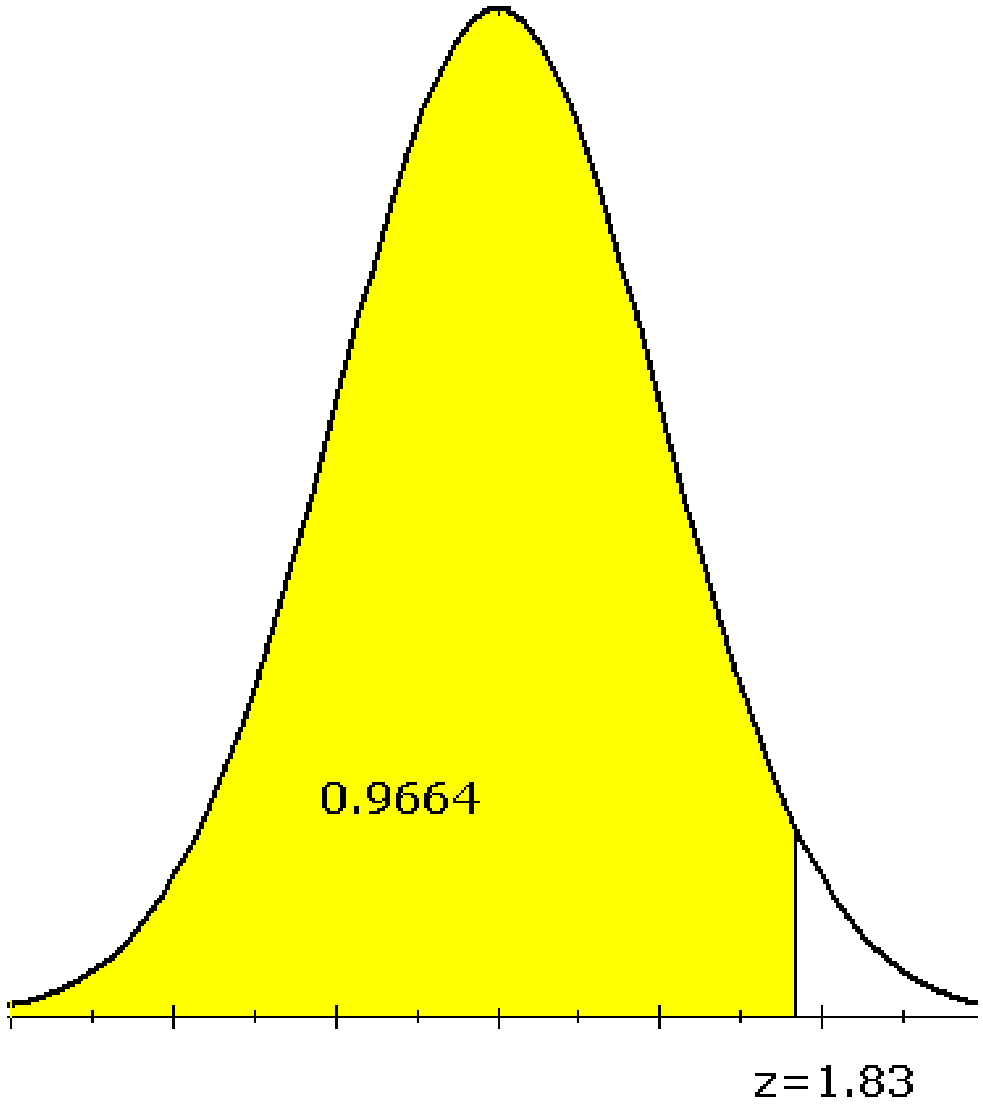
The proportion of scores above 40 can be obtained by finding the area under the standard normal curve to the right of z=1.83

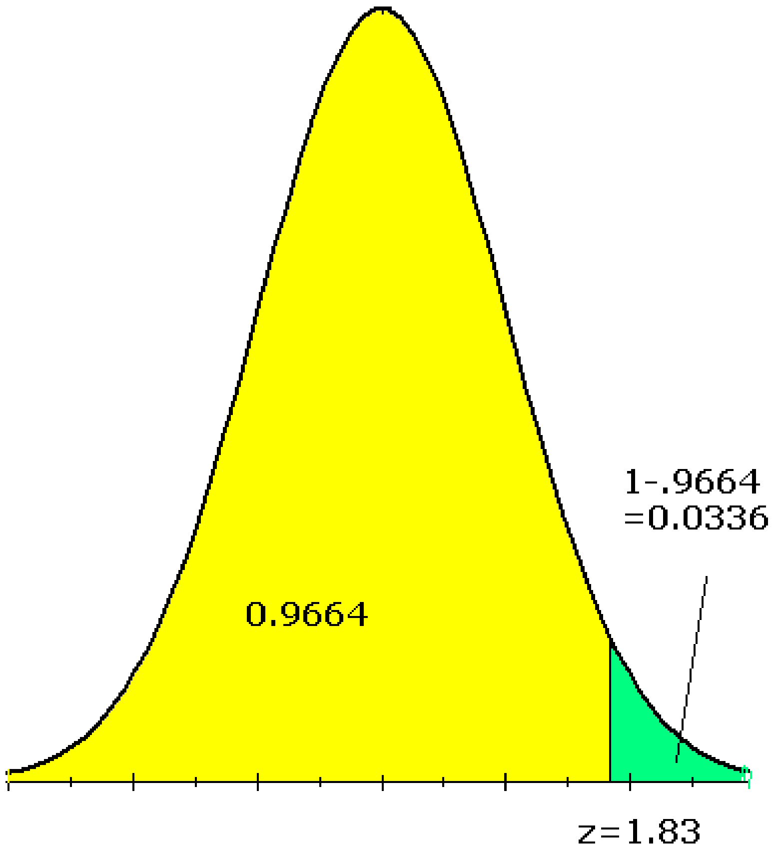
First, we shall find the area under the standard normal curve to the left of z=1.83 from the z-table

$$.1.83 = 1.80 + .03$$

$$\begin{array}{ccc} & & .03 \\ & & \downarrow \\ 1.80 & \rightarrow & .9664 \end{array}$$


z


0.9664
z=1.83



b) **Find the percentage of analysists who scored between 20 and 40 points**.

For $x = 20$,

$\mathbf{z} = \frac{20-29}{6} = \mathbf{-1.5}$

$$\begin{array}{ccc} & \mathbf{.00} & & \mathbf{.09} \\ & \downarrow & \downarrow & \\ \mathbf{-1.5} & 0.0668 & & \end{array}$$

For $x = 40$
we already saw that
$\mathbf{z} = \frac{40-29}{6} = \mathbf{1.83}$

and that

$$\begin{array}{ccc} & & \mathbf{.03} \\ & & \downarrow \\ \mathbf{1.80} & \rightarrow & \mathbf{.9664} \end{array}$$

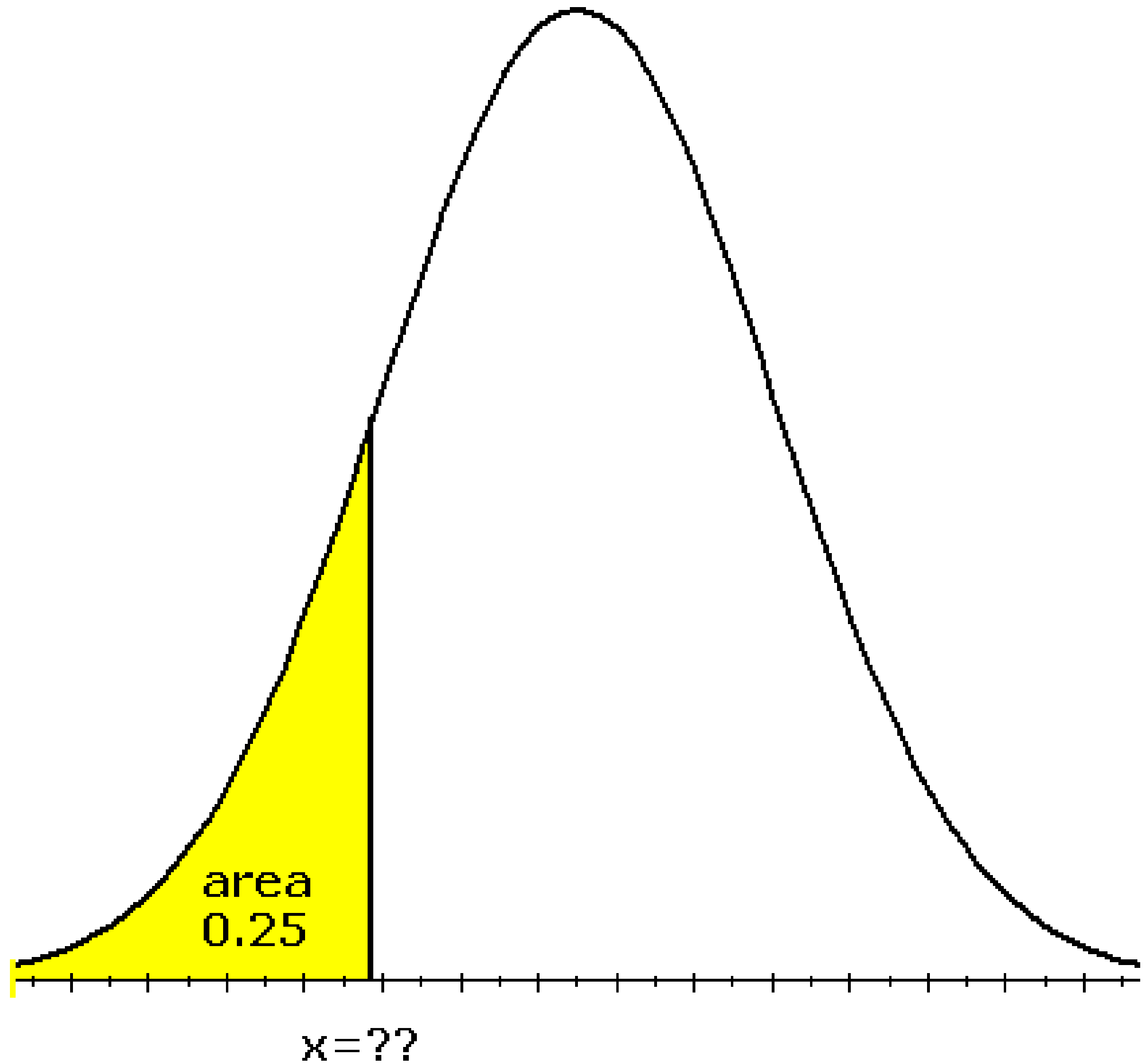
The proportion of scores between 20 **and** 40 **is the same as the area under the standard normal curve from** $z = -1.50$ **and** $z = 1.83$

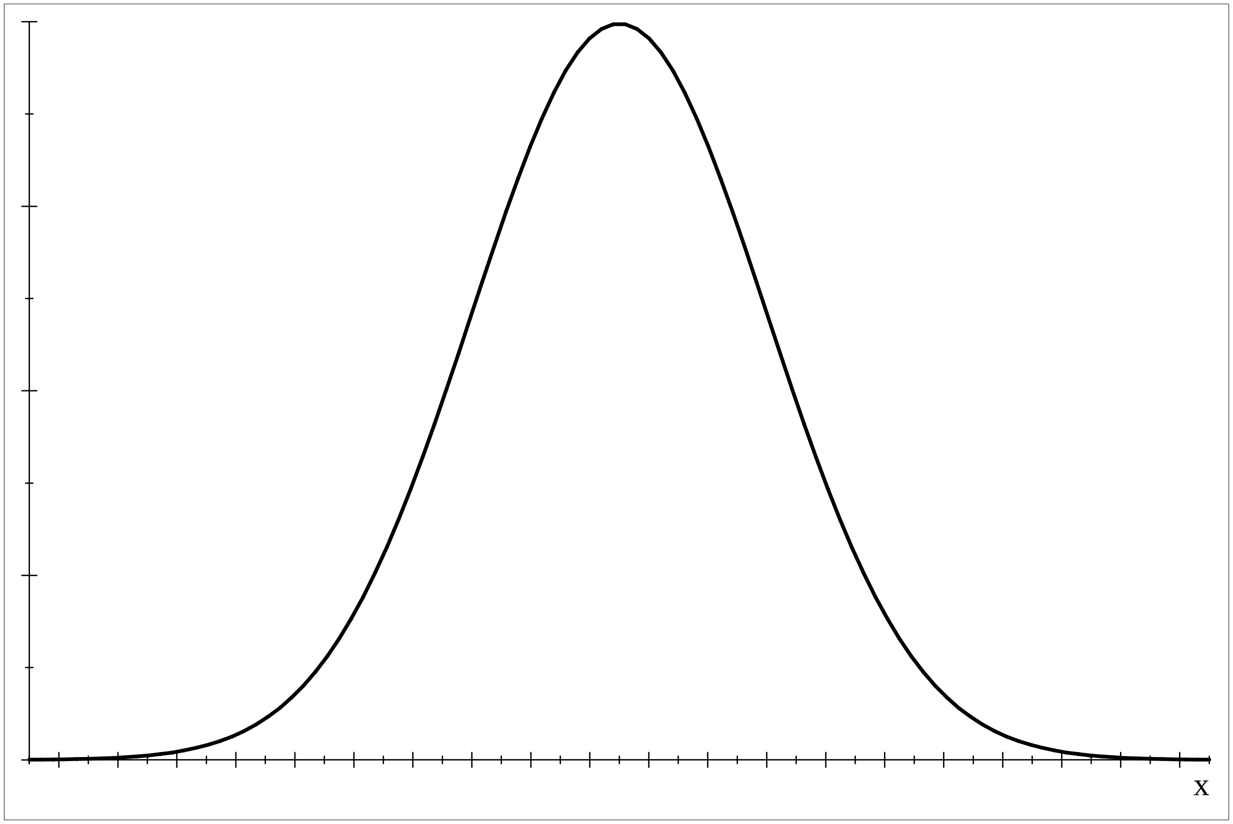
is $.9664 - 0.668 = 0.2984$
about 89.**96**% **of the scores are between 20 and 40**.

c) **Find the first quartile of this distribution**.

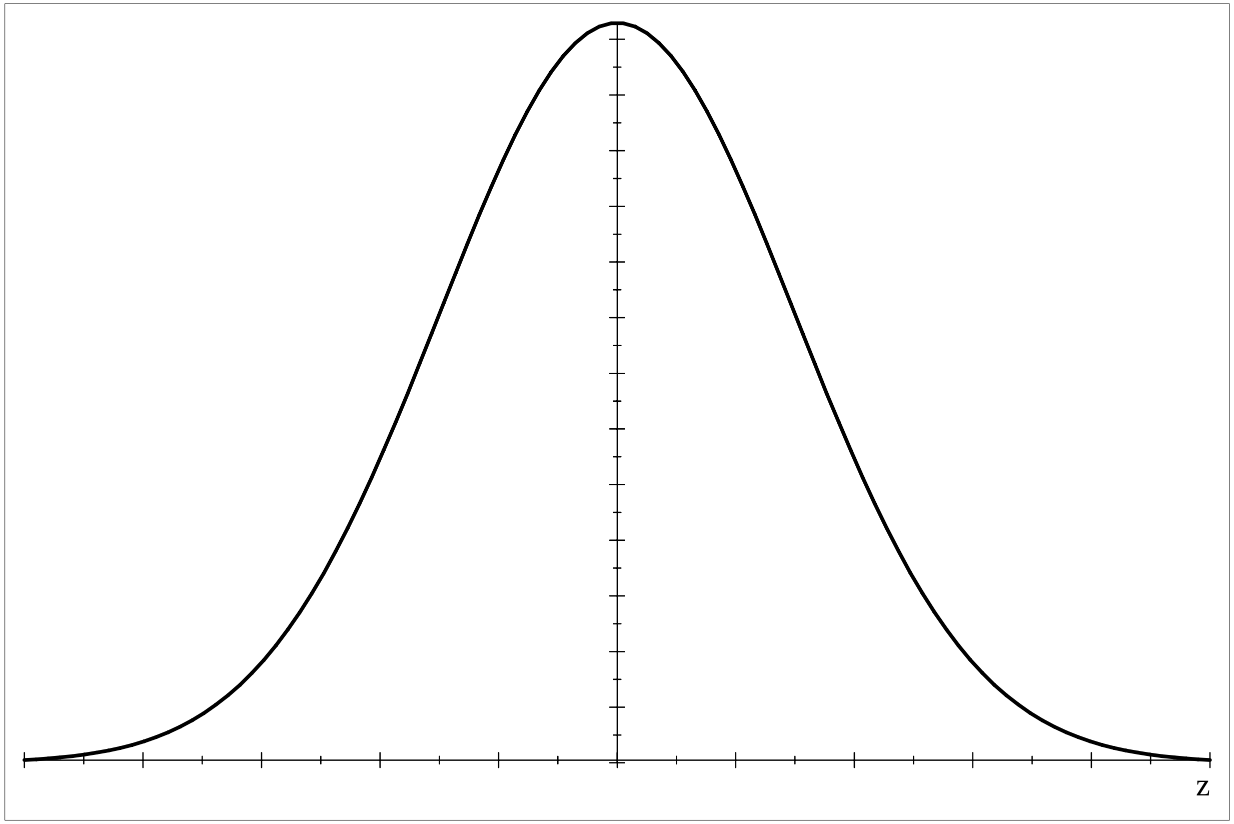
Who is the first quartile,
the number such that 25% **of the data is below this number**.

That is x such that the area to the left of x is approximately 0.25.


area
0.25
x=??



If we are using the z-table, first we have to find z such that the area below this value of z is 0.25.



Look for the area value of 0.25 in the table.

Since, we could not find 0.25 exactly in the area portion, we are using the value .2486 that is closest to 0.25

$$\begin{array}{ccc} & & .07 \\ & & \uparrow \\ -0.6 & \leftarrow & .2486 \end{array}$$

$\mathbf{z = -0.67}$

This gives us $z = -.67$ **as the first quartile for the standard normal distribution**

Now, we shall find the first Quartile for our normal distribution of interest that is shall solve the following equation for x

$$\frac{x-29}{6} = -.67$$

$\rightarrow$

$\mathbf{x - 29 = -.67 \times 6}$

$\mathbf{x = 29 -.67 \times 6}$

x$= 24.98$

is the first quartile of this distribution.

It is just a coincidence that the first quartile is approximately 25

d) You would like to give a special raise to the analysits who are in the top 2%. At least what score must one have to get this special raise?

First let us find the value of z such that the area under the standard normal curve to the left of this value is

$\mathbf{1 - .02} = 0.98$

Look in the area portion

$$\begin{array}{ccc} & & \mathbf{.05} \\ & & \uparrow \\ \mathbf{2.0} & \leftarrow & \mathbf{.9798} \end{array}$$

z=2.05

To find x solve,

$$\frac{x-29}{6} = \mathbf{2.05} \quad \rightarrow \quad \mathbf{x - 29 = 2.05 \times 6} \quad \rightarrow \quad \mathbf{x = 29 + 2.05 \times 6 = 41.3}$$

Must have at least 41.3 points to be in the top 2%

Example 2:

The processing time for an application at a certain office shows approximately a normal distribution with mean of 20.7 days and standard deviation of 2.9 days. Answer the following questions:

a) What percentage of applications are processed within 21 days?

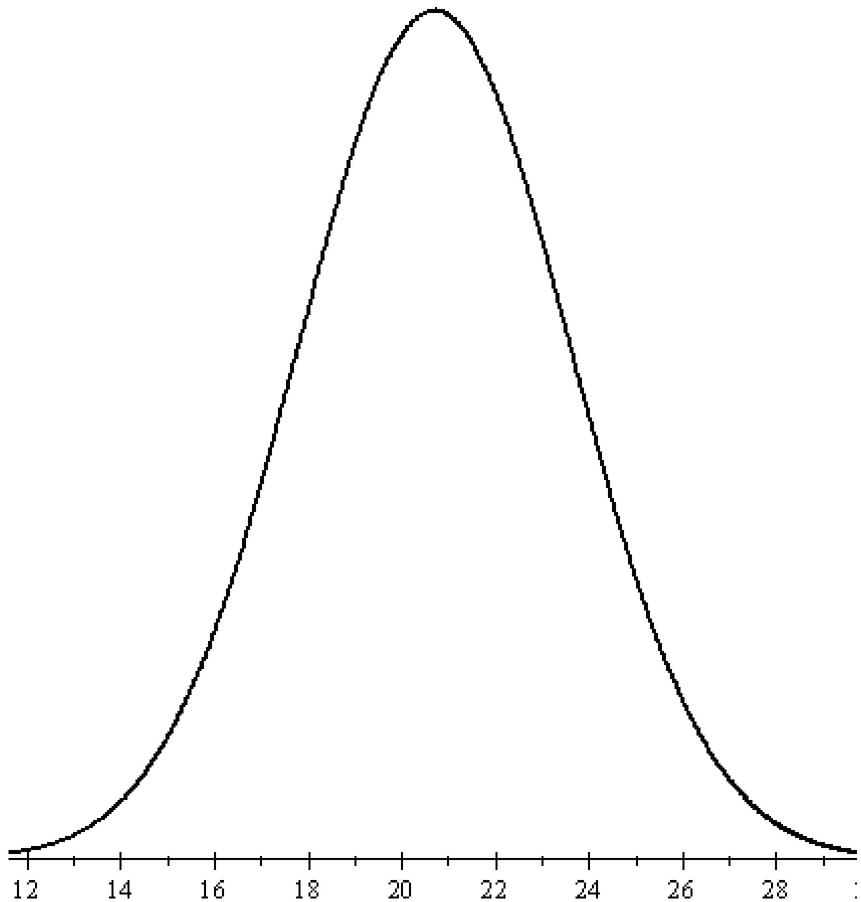
b) What percentage of applications take from 14 to 28 days?

c) What percentage of applications take more than 30 days?

d) How fast does an application takes to be processed to within the fastest 1%.

e) A lawyer's office has sent five such applications to this office. Calculate the probability that at least one of these will be processed within 18 days.

(state the assumptions that you made in your calculations.)



a) What percentage of applications are processed within 21 days?

Standardize

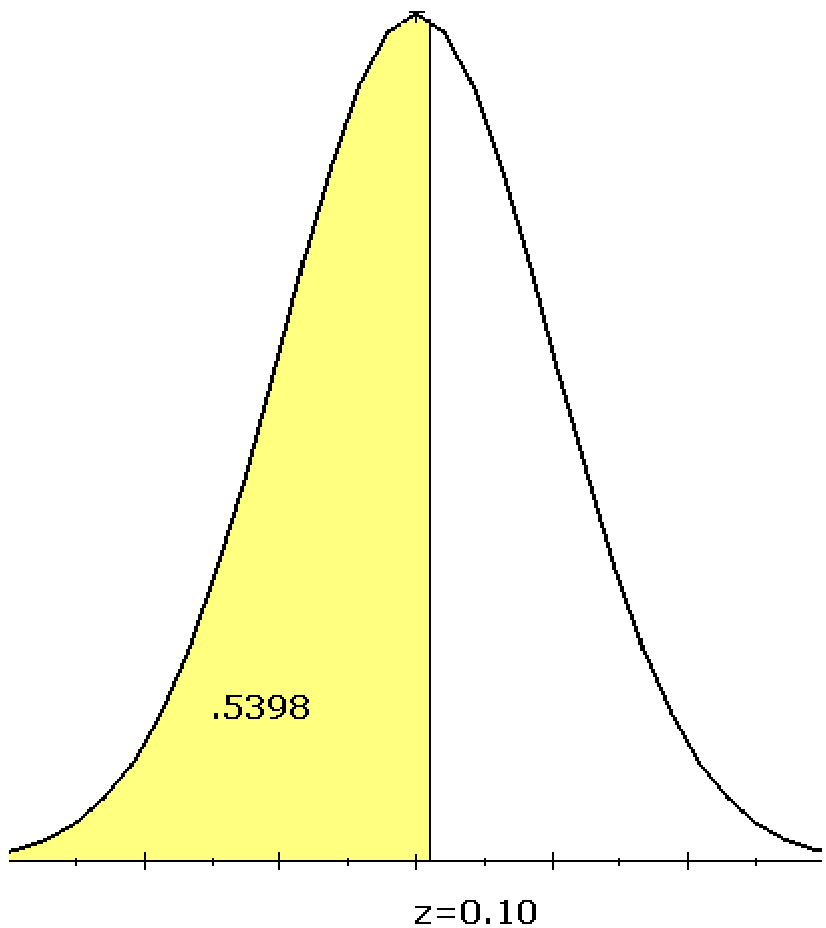
$\mathbf{z} = \frac{21-20.7}{2.9} = \mathbf{0.1034482759 \cong 0.10}$

From the table the area to the left of $z = .10$ **is**

.00 .09

↓

.1 .5398



The area to the left of $z = 0.1$ **is** $.0.5398$

53.98% are processed within 21 days.

b) What percentage of applications take from 14 to 28 days?

Between
14 to 28

For $x = 14: \quad z = \dfrac{14-20.7}{2.9} = -2.310344828$
For $x = 28: \quad z = \dfrac{28-20.7}{2.9} = 2.517241379$

To find the area under the z-curve from $z = -2.31$ **to** $z = 2.52$

Area between $z = -2.31$ **and** $z = 0$ **is**

$$\begin{array}{lll} & \mathbf{.01} & \qquad\qquad\qquad \mathbf{.09} \\ & \downarrow & \\ \mathbf{-2.3} \rightarrow & \mathbf{.0104} & \end{array}$$

Area between to the left of $z = 2.52$

$$\begin{array}{ll} & \mathbf{.02} \\ & \downarrow \\ \mathbf{2.5} \rightarrow & \mathbf{.9941} \end{array}$$

The area between $z = -2.31$ **to** $z = 2.52$

is $.9941 - .0104 = 0.9837$
.

98.37%

c) What percentage of applications take more than 30 days?
z$=\frac{30-20.7}{2.9} = 3.206896552 \cong 3.21$

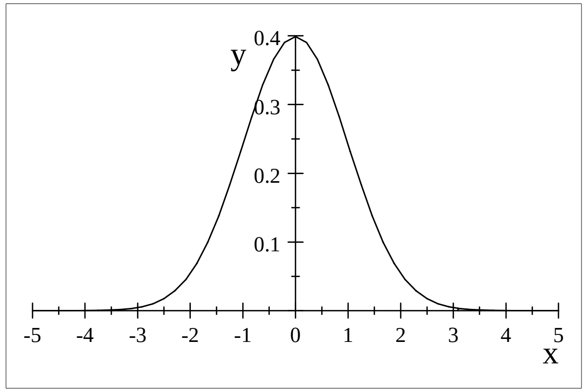
Off the limits on the chart

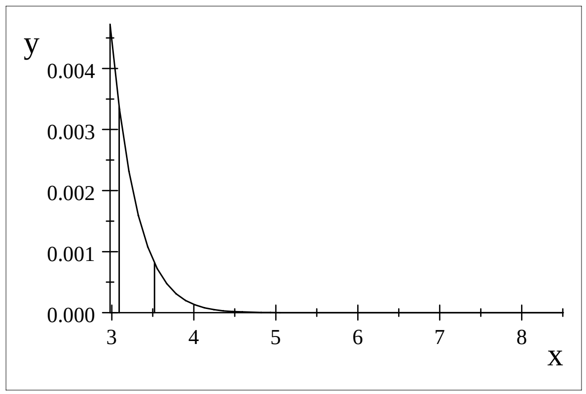
The largest z **value on the chart is** $z = 3.09$

	$\mathbf{.09}$
$\mathbf{3.0}$	$\mathbf{.9990}$

The area to the right of $z = 3.09$ **is**
$\mathbf{1 - .9990 = 0.001}$

$s(z)$





the area to the right of $z = 3.52$
is approximately 0
actually
less than $.001$

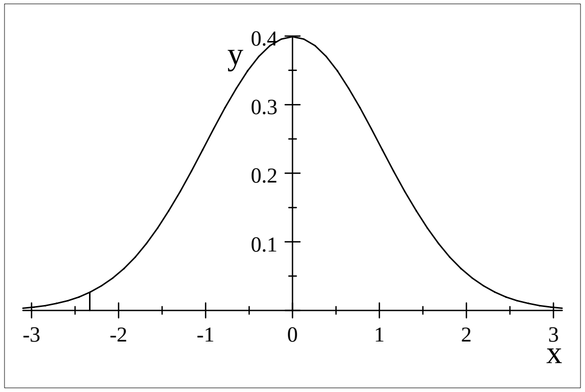
d) **How fast does an application takes to be processed to within the fastest 1**%.

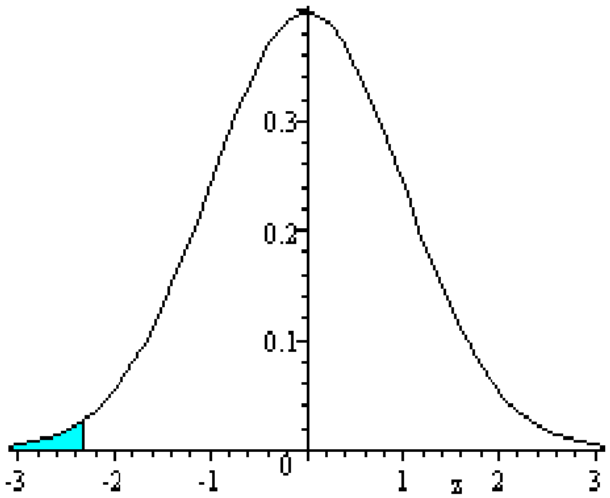
To find X **such that only 1**% **of the area is to the left of this value of** X

We are doing a reverse calculation here

We take 0.01 in the table

$s(z)$





the blue area is 0.01

the value closest to 0.01 is 0.0099

$$\begin{array}{ccc} & & \mathbf{.03} \\ & & \uparrow \\ \mathbf{-2.3} & \leftarrow & \mathbf{.0099} \end{array}$$

our value of z **is**

$\mathbf{z = -2.33}$

We still have to find X

$\frac{X-20.7}{2.9} = -2.33$, **Solution is**: $\{X = 13.943\}$

e) A lawyer's office has sent five such applications to this office. Calculate the probability that at least one of these will be processed within 18 days.

Assumption1: We are assuming that the same trend holds for the future applications.

First:
We find the probability that any of them will be processed within 18 days

$$\mathbf{z} = \frac{18-20.7}{2.9} = -\mathbf{0.9310344828} \cong -\mathbf{0.93}$$

$$\begin{array}{ccc} & & \mathbf{.03} \\ & & \downarrow \\ \mathbf{-.9} & \rightarrow & \mathbf{.1762} \end{array}$$

A

The probability that any application will take less than 18 days is
$\mathbf{0.1762}$

assuming that
The probability that any ONE application will take less than 18 days is
$\mathbf{0.1762}$

Assuming independence,

we have a binomial distribution with

$\mathbf{n} = \mathbf{5}$

$\mathbf{p} = .\mathbf{1762}$

to find the probability

for

1 or more

a better way is:

$\mathbf{1} - \mathbf{P}(\mathbf{0})$

$= 1 - C(5,0)(0.1762)^{0}(1 - .1762)^{5}$

$= 1 - (1 - .1762)^{5}$

$= \mathbf{0.620589856}$

f)

A lawyer's office has sent FIFTEEN such applications to this office. **Calculate the probability that at least one of these will be processed within 18 days**.

$\mathbf{1} - (1 - .1782)^{15} = \mathbf{0.947338642}$

Another widely used continuous distribution

is exponential distribution