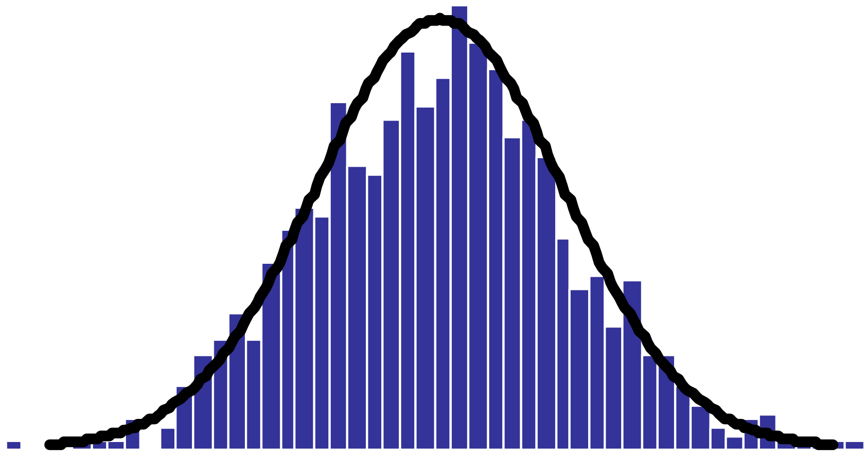
Lecture 3
The Normal Distribution

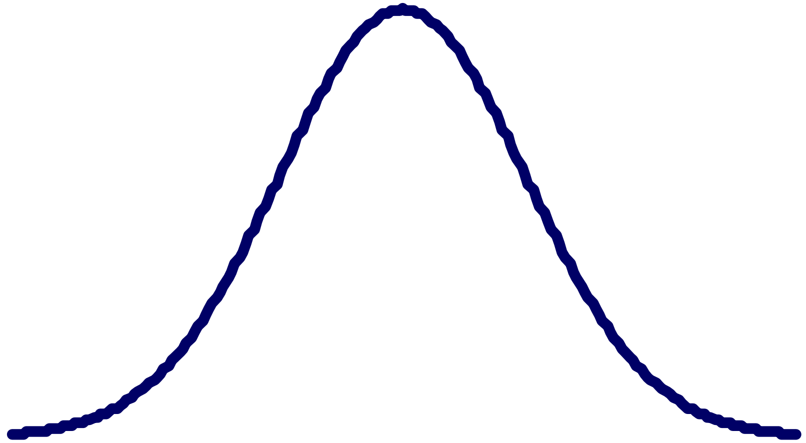
Introduction to probability
The normal distribution
The standard normal distribution
Areas under a normal curve
z-scores and probability
Percentiles for a normal distribution

In the Last Lecture….

- We looked at summarising data.
- The most basic measures summarising ***numerical*** data are measures of ***centre*** and measures of ***spread***.
- The **mean** and **median** are measures of centre. The ***range***, ***standard deviation*** and ***inter-quartile range*** are measures of spread.
- We summarise ***categorical*** data by calculating the ***proportions*** of data in the categories of interest.
- We saw that **population distributions** can take many shapes. We noted that histograms of large samples usually more closely resemble the population from which they are drawn than histograms of small samples.

Population Distributions



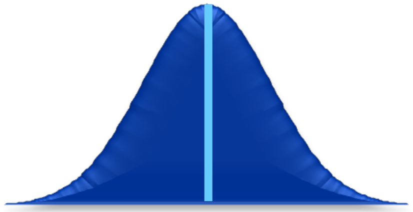


A histogram of a sample

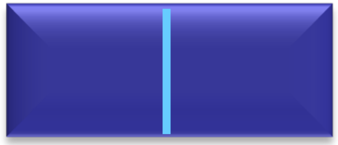
suggests the *shape* of the population from which it is drawn.

Different Population Distributions

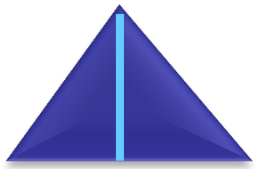
Symmetric



Normal



Uniform



Triangular

Skewed

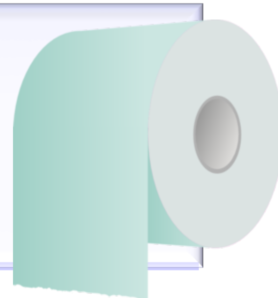
tail left

Skewed Left

tail right

Skewed Right

Review Quiz 1



Toilet paper usage in New Zealand, Western Europe, Britain and North America is respectively, 11, 14, 18 and 23 kg per person per year.

a. Calculate the mean, median, range and standard deviation:

b. Now calculate the mean, median, range and standard deviation if toilet paper usage in each country was to increase by 2kg per person per year.

c. Now calculate the mean, median, range and standard deviation if toilet paper usage in each country was to halve.

Review Quiz 1 continued

Of the statistics shown below, which will change when we:

d. add or subtract a constant to/from each observation in a data set?

e. rescale each observation by multiplying or dividing by a constant?

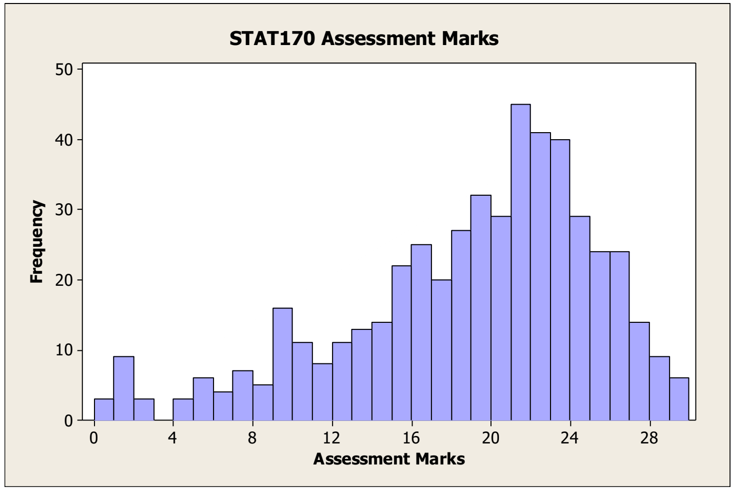
   i. mean
   ii. median
   iii. range
   iv. standard deviation

Review Quiz 2

The following output was obtained from the final assessment marks of 500 randomly selected STAT170 students in 2012:

Descriptive Statistics: Assessment Marks

Variable	N	Mean	StDev	Minimum	Median	Maximum
Assessment Marks	500	18.974	6.356	0.00	20.50	30.00



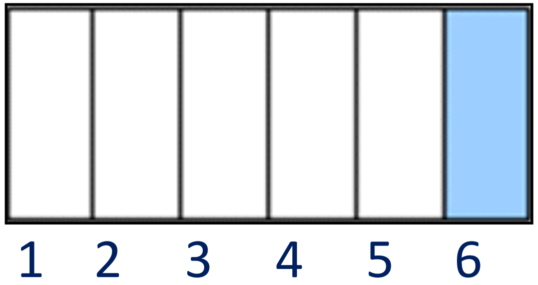
a. Describe the shape of the distribution.
b. How many students did not do any assessments?
c. How many students scored full marks on their assessments?

Introduction to probability

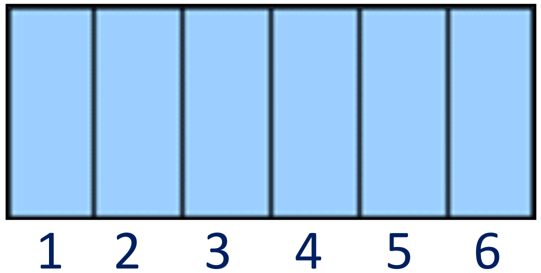
Probability

Example: rolling a die. The probability of rolling a particular number when we roll a die is the same as the proportion of times we would expect to roll that number if we roll the die over and over eg.

The probability of rolling a 6 = 1/6

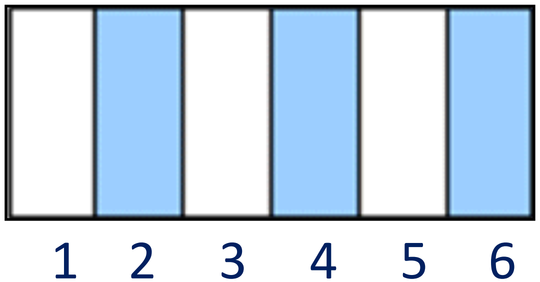


The probability of rolling a 1, 2, 3, 4, 5 or 6 = 1



The probability of rolling an even number

= 1/6 + 1/6 + 1/6 = 1/2



Proportions and Probabilities

- *In the long term*, the *proportion* of throws of a die that are even is expected to be 0.5, ie. 50%.
- Thus the *probability* that *one* throw of a die will be even is also 0.5, ie. 50%
- We can express probabilities and proportions either as decimals or as percentages.
- **Proportions and probabilities *always* lie between 0 and 1, percentages between 0% and 100%.**

Probability and Area

- In the example of rolling a die *we used shaded areas to represent proportions of total 'long term' events*, and hence *probabilities of events*, for a *discrete* distribution.

- The whole distribution represents the results from every throw of a die in the 'long term' (ie. if the die was thrown an infinite number of times). The shaded part of the distribution represents the proportion of observations of interest (eg. the proportion of throws that are expected to result in an even number).

- The idea that shading part of **an *area* under a curve** represents a proportion of all the *values* in that distribution leads us to the concept of *probability*.

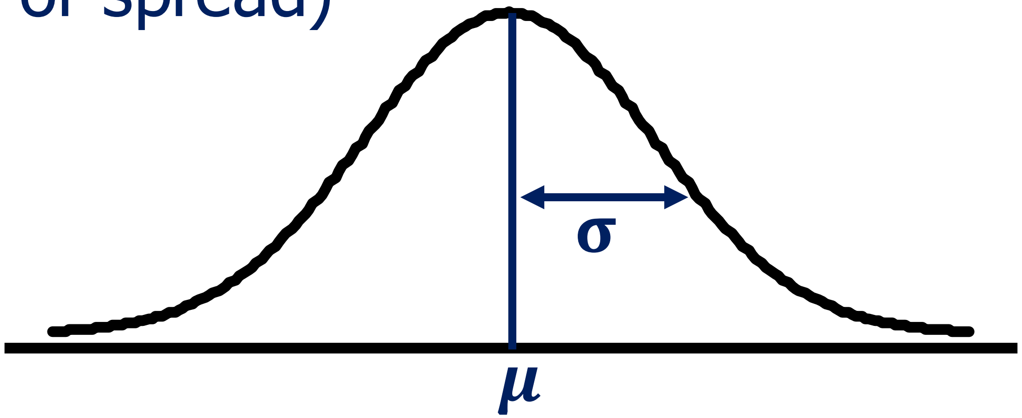
What is Probability?

- A **probability** is a *measure of the expectation that an event will occur*.
- **Probabilities lie between 0 and 1**.
- Now we shall discuss the **normal distribution**, and see how we use areas to represent proportions and probabilities in that *continuous* population distribution.

The normal distribution

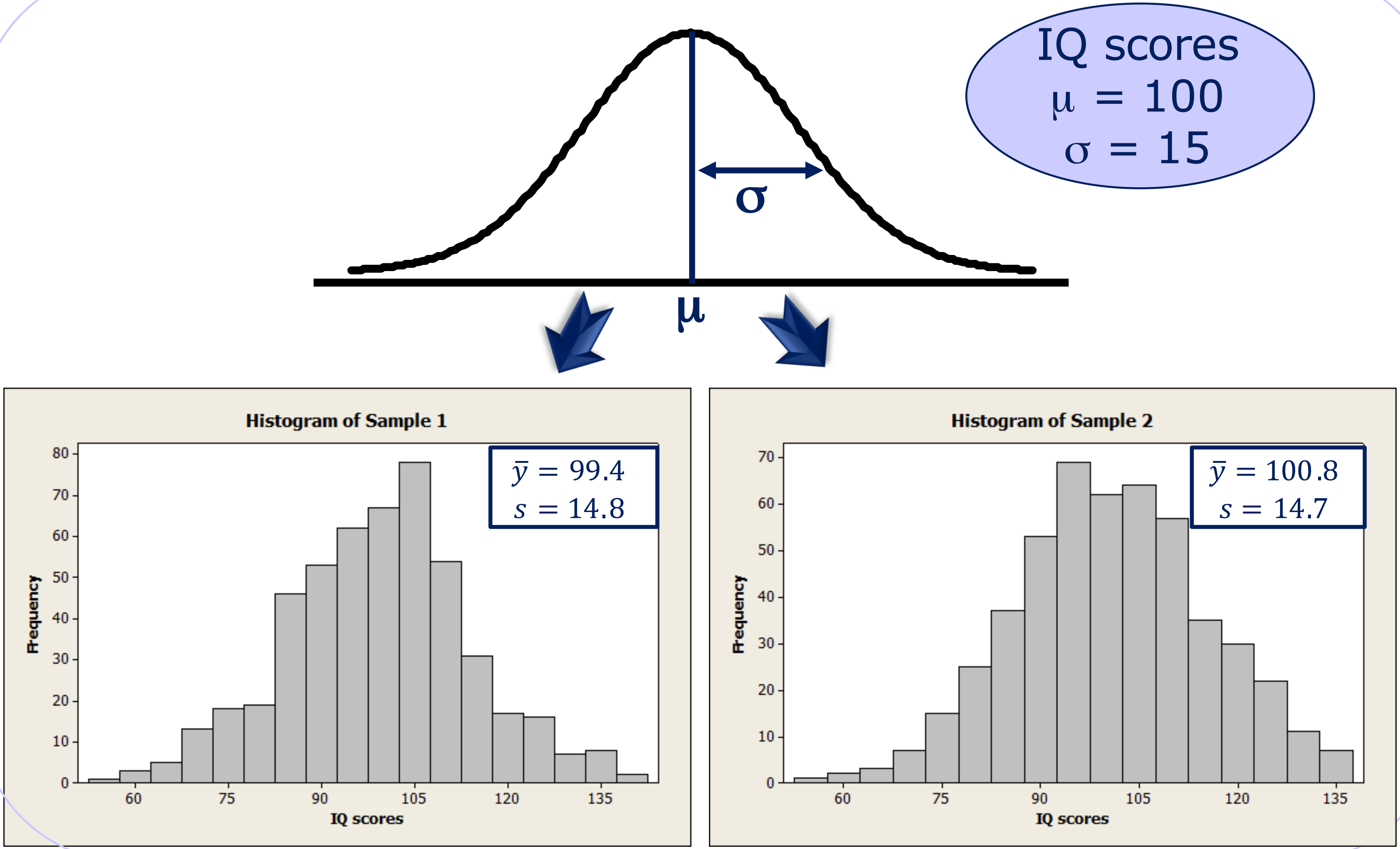
The Normal Distribution

- The normal distribution is a symmetric "bell shaped" distribution with an unlimited range. It is described by two parameters μ (a measure of centre) and σ (a measure of spread)



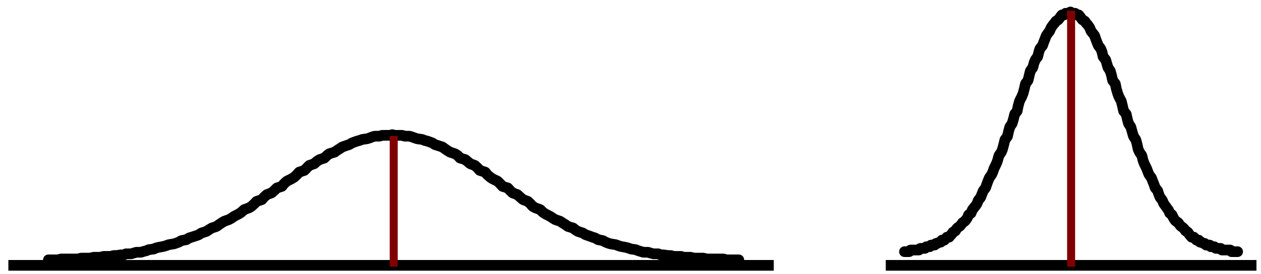
- As you will soon see the normal distribution plays a very important role in statistics.

An Example of a Normal Distribution



The Shape of a Normal Distribution

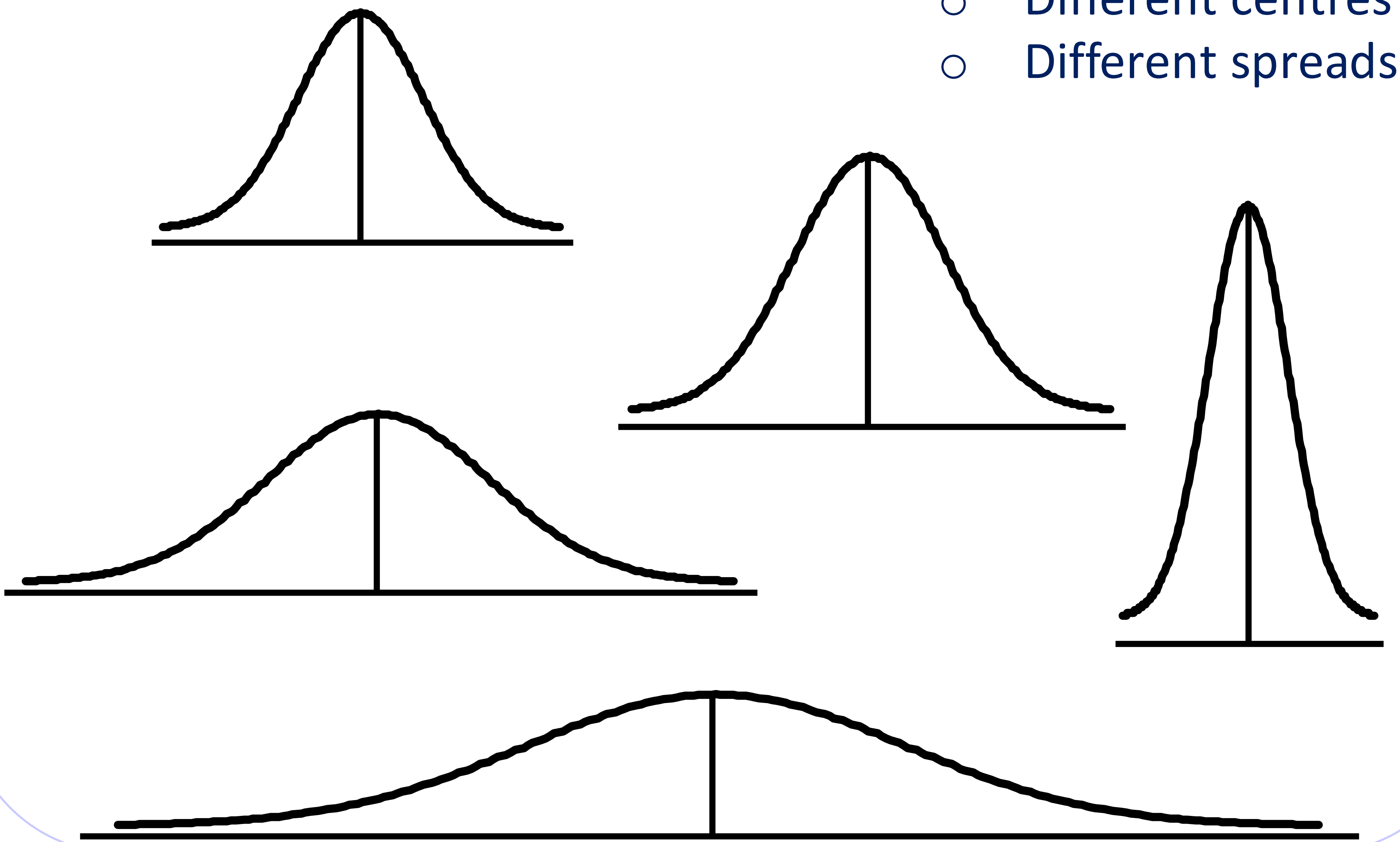
- Normal population distributions may have different means and different standard deviations.



- While every normal population distribution is "bell-shaped", the **height and width** of the "bell" depends on σ, the standard deviation.
- The "bell's" **position along the horizontal axis** depends on μ, its mean.

Examples of Different Normal Population Distributions

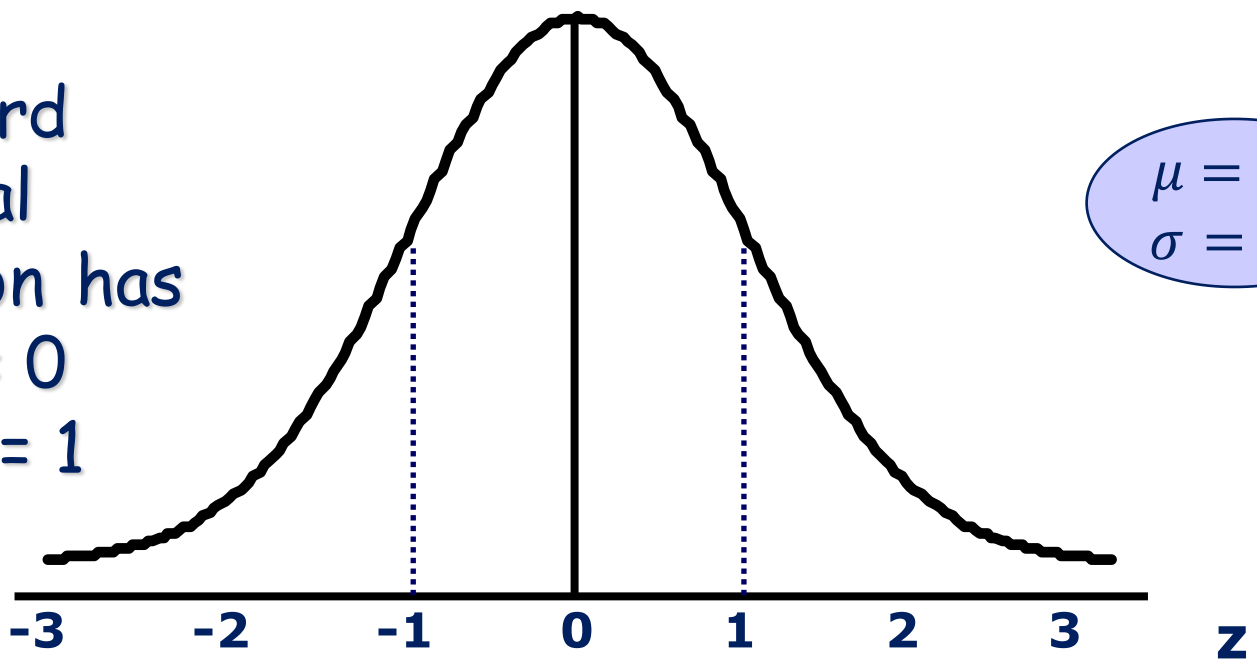
- Different centres
- Different spreads



- The standard normal distribution

The Standard Normal Distribution

The standard normal distribution has mean = 0 and sd = 1



$\mu = 0$
$\sigma = 1$

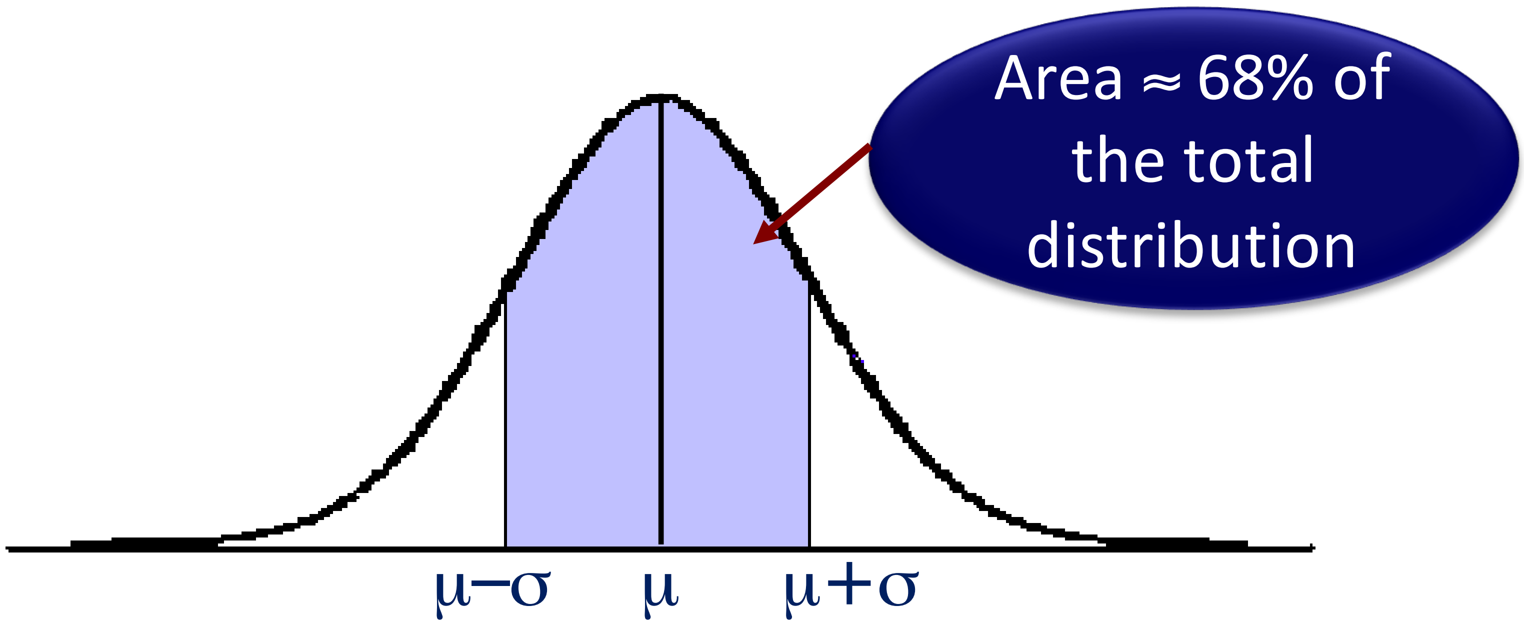
The **steepest parts** of the curve correspond to $\mu - \sigma$ and $\mu + \sigma$, in this case **−1** and **+1**.

- Areas under a normal curve

Area Under a Normal Curve

In any normal distribution, 68.3% of the values lie within one standard deviation of the mean.

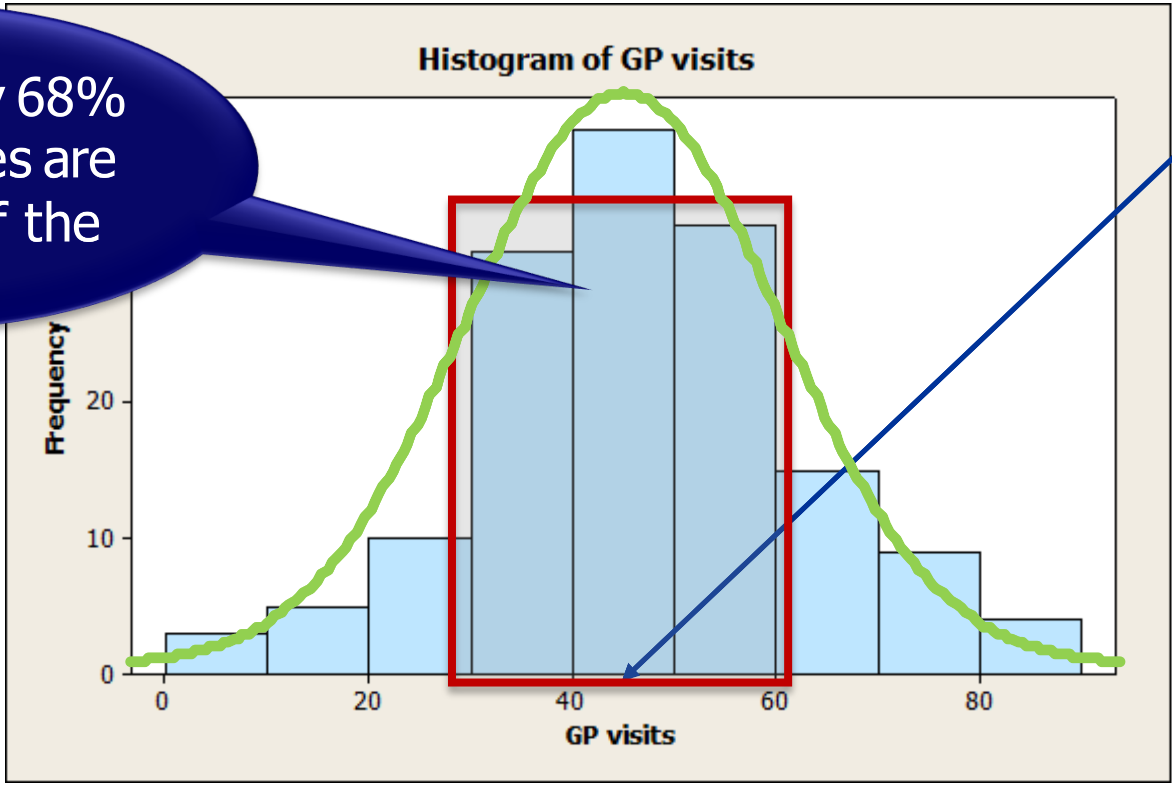
This percentage is represented by the area under the curve between $\mu - \sigma$ and $\mu + \sigma$.



Example: Normal Distribution

Let Y = number of visits to a GP in one year.

Histogram showing the number of visits to a GP in 2004, for a sample of residents of a retirement village.



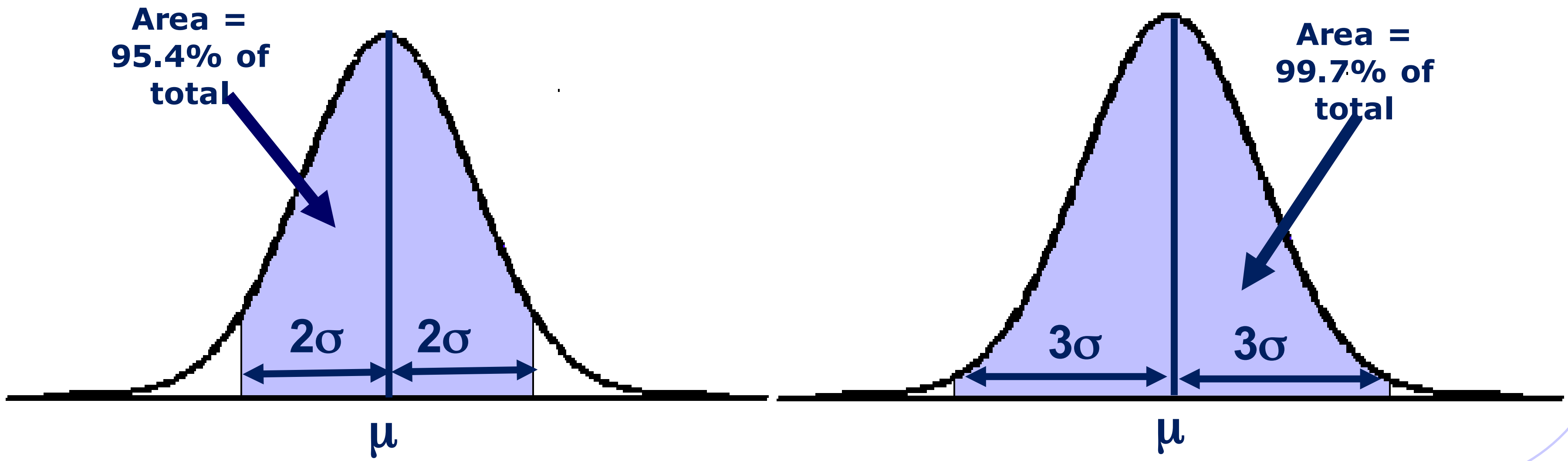
Approximately 68% of these values are within 1 sd of the mean!

$\bar{y} = 45.6$
$s = 16.6$

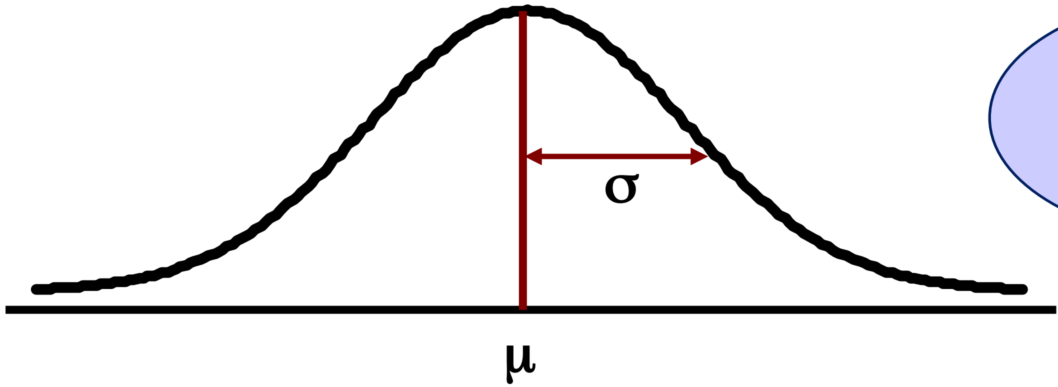
Area Under a Normal Curve

In any normal distribution:

- 95.4% of values lie within two standard deviations of the mean.
- 99.7% of values lie within three standard deviations of the mean.



IQ Scores: A Normal Population Distribution



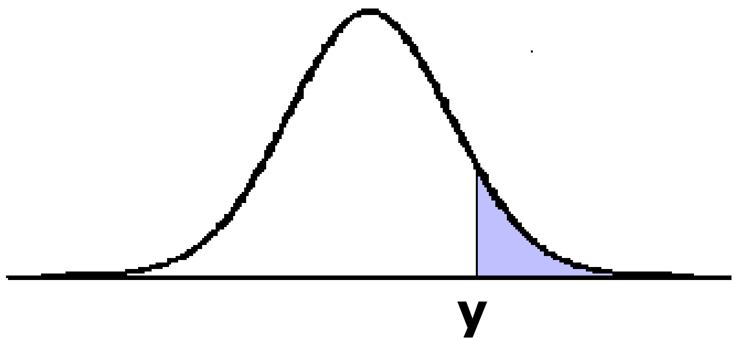
IQ scores
$\mu = 100$
$\sigma = 15$

This means that:

- 68.3% of people have IQs between 85 and 115
- 95.4% of people have IQs between 70 and 130
- **99.7% of people have IQs between 55 and 145**

Very few people will have IQ scores more than 3 standard deviations from the mean.

Area Under a Distribution Curve → Probability



The **shaded area** under the distribution curve above y represents

- the **p**robability that an observation is greater than y

As noted before this is the same as

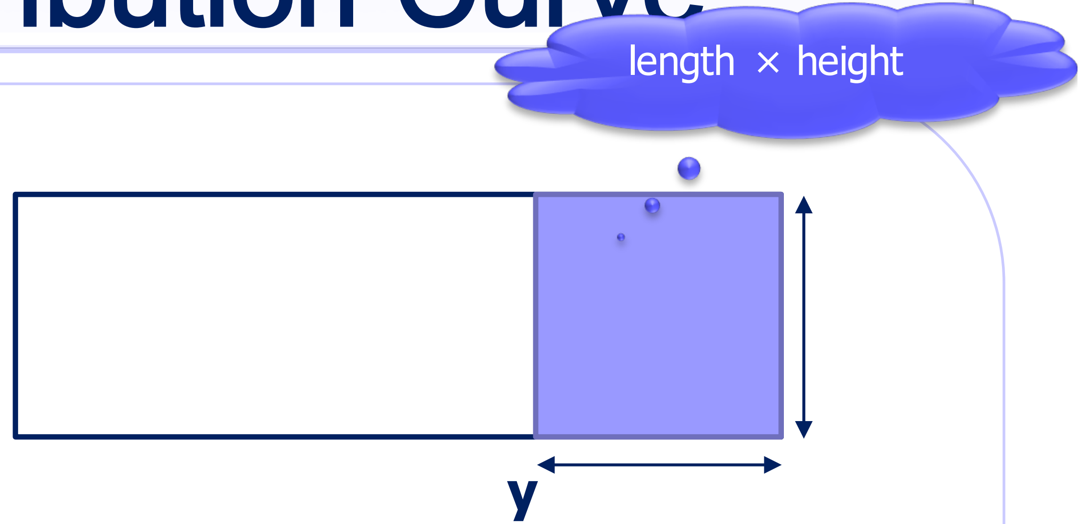
- the **p**roportion of all values greater than y

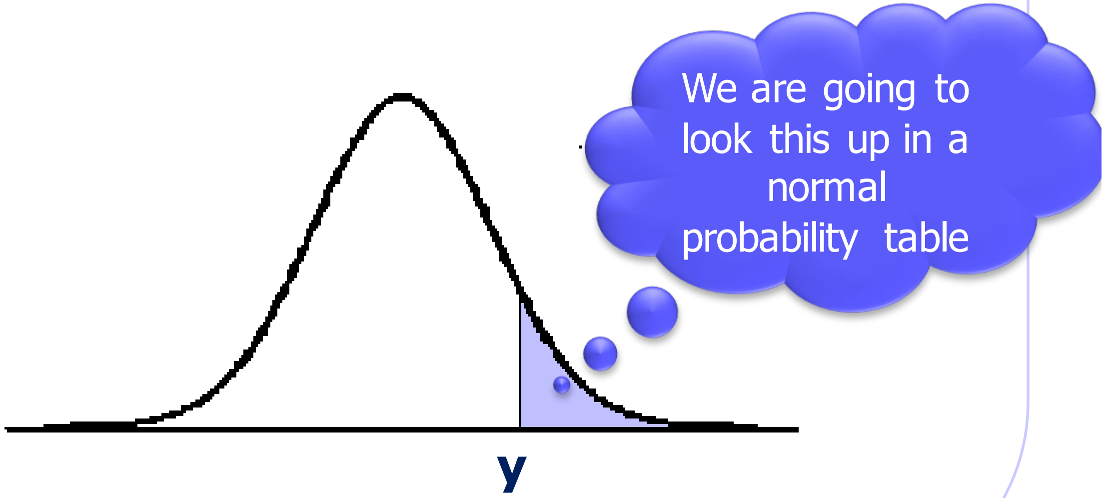
We can write this as a fraction, decimal or percentage – it depends how a question wants you to report the answer.

Quick Quiz: How to Find the Area Under a Distribution Curve

How do you find

- The area above y?





- The area above y?

Using Tables to Find the Area Under a Distribution Curve

- If we know its mean and standard deviation, we can find **areas under *any* normal curve**.

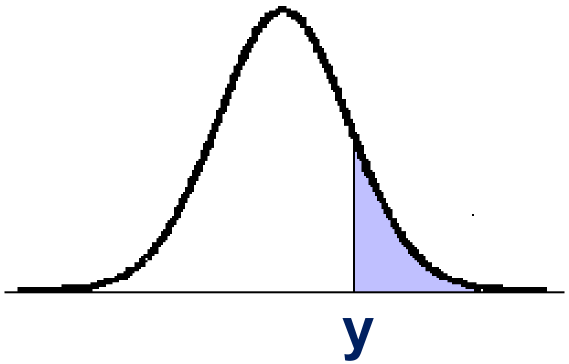
- That is, we can find the ***probability*** that a *randomly selected* value lies above or below any value y, or between any two values y_1 and y_2 !

- We do this by working out how many standard deviations y is from the mean (this is called a z-score), and then looking up the corresponding area in normal tables (z-tables).

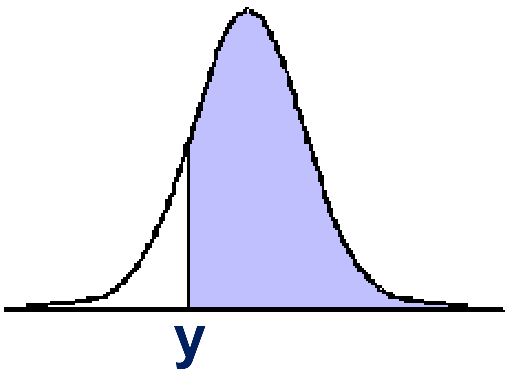
z-scores and probability

z-scores

A z-score is the number of standard deviations that a value, y, is away from its population mean, μ.



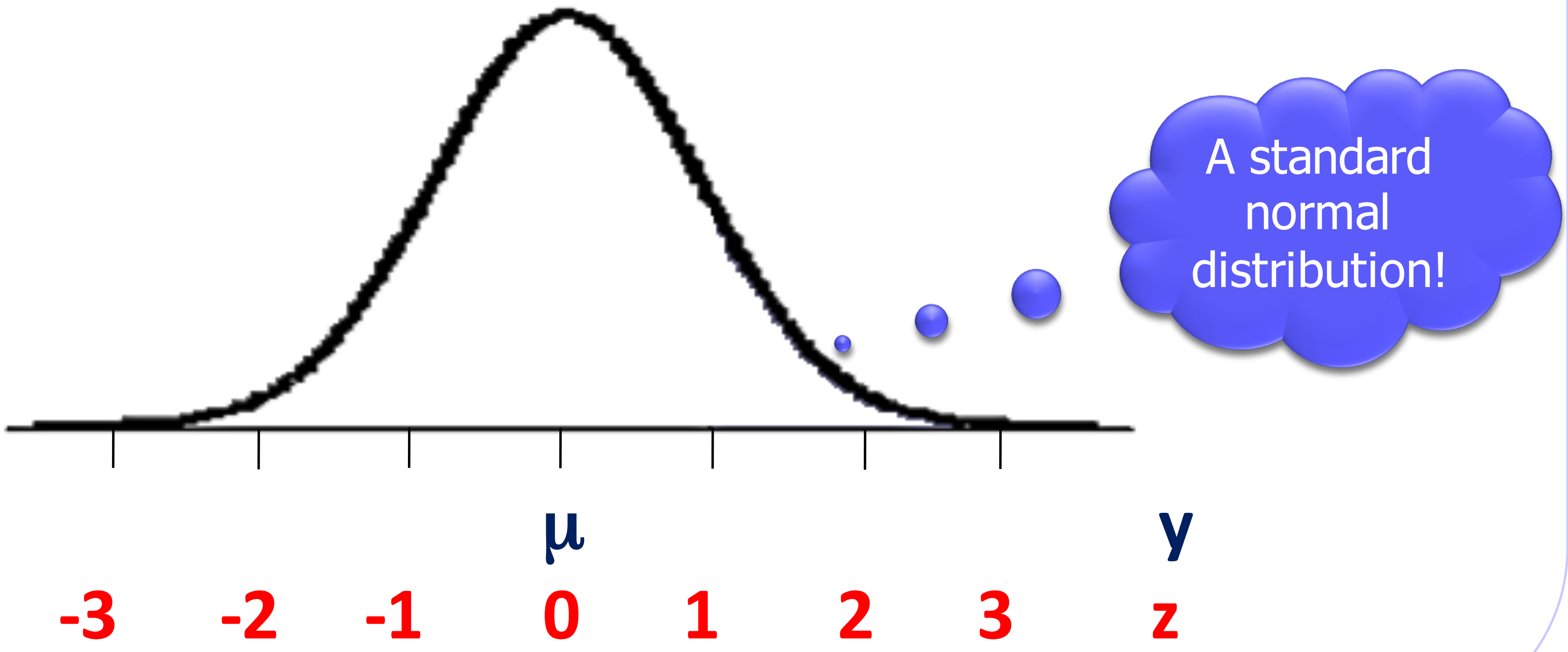
If the y-value is greater than the mean, that is, if $y > \mu$, then the z-score is *positive*.



If the y-value is less than the mean, that is, if $y < \mu$, then the z-score is *negative*.

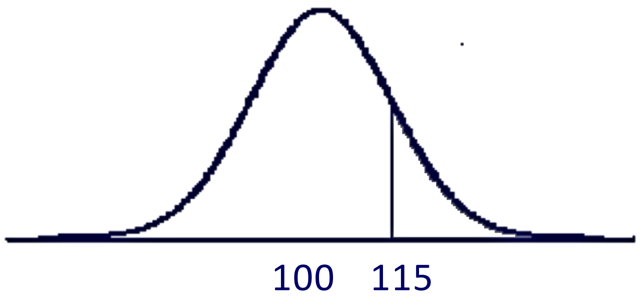
z-scores

A z-score is the number of standard deviations that a value, y, is away from its population mean, μ.



Calculating z-scores

IQ's are normally distributed with mean, μ = 100 and standard deviation, σ = 15

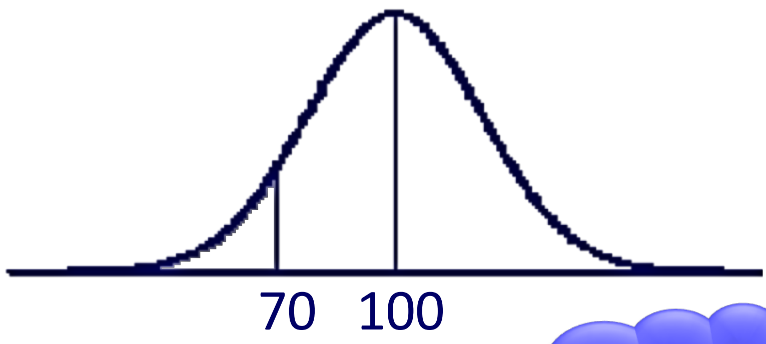


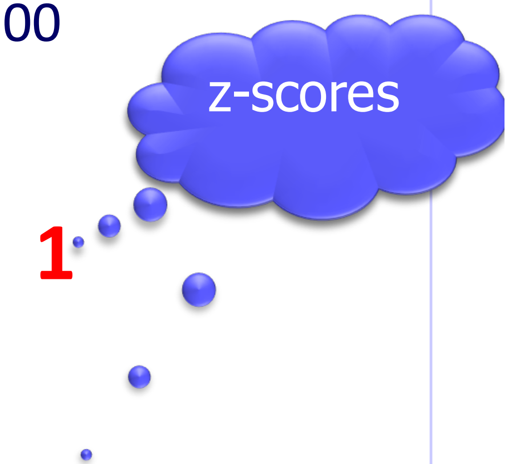
Joe's IQ is 115.

How many sds is Joe's IQ score from the mean? **1**

Calculating z-scores

IQ's are normally distributed with mean, μ = 100 and standard deviation, σ = 15





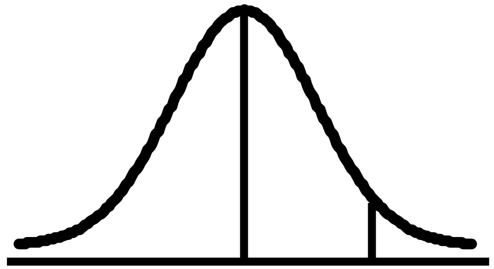
Joe's IQ is 115.
How many sds is Joe's IQ score from the mean? **1**

Fred's IQ is 70.
How many sds is Fred's IQ score from the mean? -2

The z score tells us how many standard deviations the (y) value is from the mean, μ.

Example: Calculating a z-score

IQ scores are normally distributed with μ = 100 and σ = 15.



What z-score represents a person with an IQ of 130?

y = 130 which is 30 points above the mean.

So, 130 is 30/15 = 2 standard deviation above the mean.

Or:

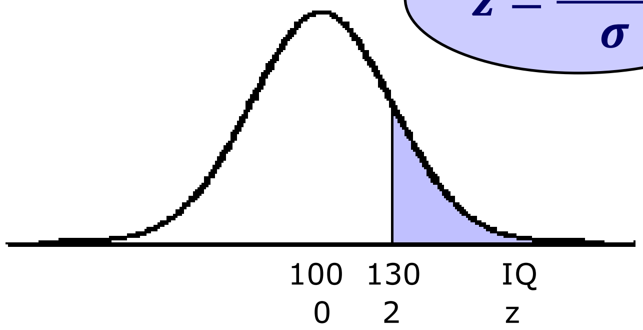
$$z - score = \frac{distance\ between\ y\ and\ the\ mean}{population\ standard\ deviation}$$

$$ie.\ z = \frac{y - \mu}{\sigma} = \frac{130 - 100}{15} = 2$$

Example: Calculating a z-score

What proportion of people do we expect to have IQ scores which are higher than 130?

$$z = \frac{y - \mu}{\sigma}$$



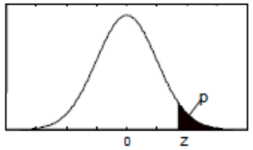
This is equivalent to asking:

What proportion of people have z-scores above 2?

This **proportion** corresponds to the **shaded area** shown above.

We can find this area by looking it up in a **z-table**.

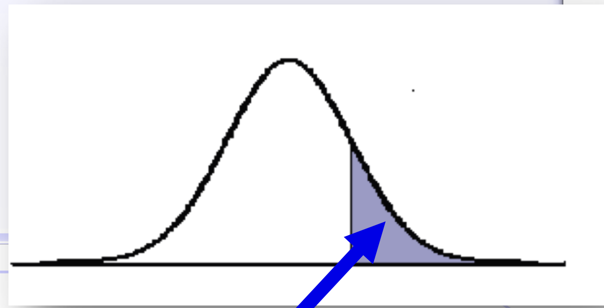
Standard Normal (z) Table



|z|	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641
0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
3.0	.00135	.00131	.00126	.00122	.00118	.00114	.00111	.00107	.00104	.00100
3.1	.00097	.00094	.00090	.00087	.00084	.00082	.00079	.00076	.00074	.00071
3.2	.00069	.00066	.00064	.00062	.00060	.00058	.00056	.00054	.00052	.00050
3.3	.00048	.00047	.00045	.00043	.00042	.00040	.00039	.00038	.00036	.00035
3.4	.00034	.00032	.00031	.00030	.00029	.00028	.00027	.00026	.00025	.00024
3.5	.00023	.00022	.00022	.00021	.00020	.00019	.00019	.00018	.00017	.00017
3.6	.00016	.00015	.00015	.00014	.00014	.00013	.00013	.00012	.00012	.00011
3.7	.00011	.00010	.00010	.00010	.00009	.00009	.00008	.00008	.00008	.00008
3.8	.00007	.00007	.00007	.00006	.00006	.00006	.00006	.00005	.00005	.00005
3.9	.00005	.00005	.00004	.00004	.00004	.00004	.00004	.00004	.00003	.00003
4.0	.000032									

Standard Normal Probabilities: Single tail areas corresponding to z-values for the standardised normal curve

z-table



Z	0.00	0.01	0.02	..	..	..	0.09
0.0							
0.1							
0.2							
…							
…							
4.0							

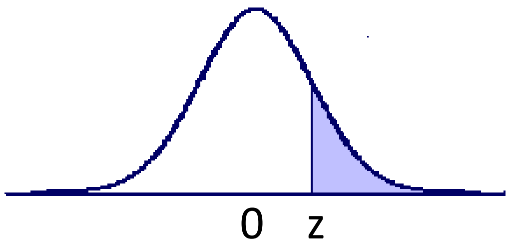
Area in Right hand tail

Using a z-table to Find a Right Hand Tail Area

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
1.2	.1151	.1131	.1112	.1093	.1075	**.1056**	.1038	.1020	.1003	.0985

z score	Area to right
0.50	0.3085
0.95	0.1711
1.00	0.1587
1.08	0.1401
1.25	0.1056

This table gives the area to the right of z-scores specified by the rows (to 1st decimal place) and the columns (2nd decimal place) of the table.



Quiz 4

Use normal tables to find the following areas:

a. area to the right of $z = 1.86$

b. area to the left of $z = 1.86$

c. area to the left of $z = -2.92$

d. area between $z = -1.5$ and $z = 2.14$

e. area between $z = 0.5$ and $z = 1$

Total area under the curve is 1

The normal curve is symmetric.

Solution to Quiz 4a

Solution to Quiz 4b

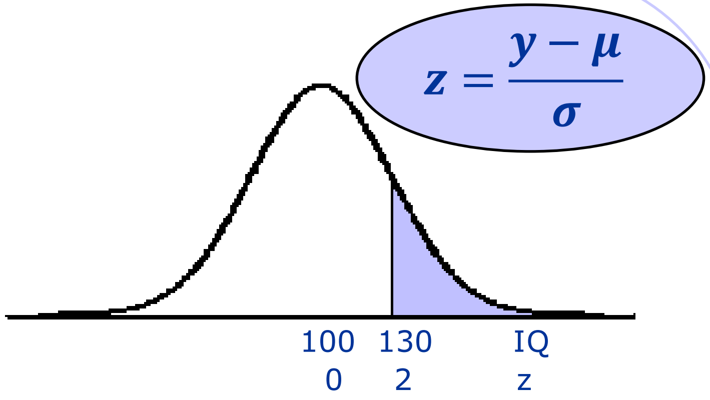
Solution to Quiz 4c

Solution to Quiz 4d

Solution to Quiz 4e

Putting it All Together: Back to the Previous Example:

Previously we calculated a z-score to answer this question:
What proportion of people do we expect to have IQ scores which are higher than 130?

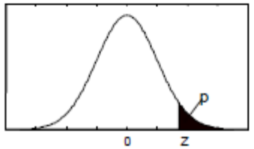


This **proportion** corresponds to the **shaded area** shown above.

We calculated: $$z = \frac{y - \mu}{\sigma} = \frac{130 - 100}{15} = 2$$

Then we found the shaded area in the diagram above by looking up the area in the right hand tail of the **z-table**.

z-table:



|z|	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641
0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
3.0	.00135	.00131	.00126	.00122	.00118	.00114	.00111	.00107	.00104	.00100
3.1	.00097	.00094	.00090	.00087	.00084	.00082	.00079	.00076	.00074	.00071
3.2	.00069	.00066	.00064	.00062	.00060	.00058	.00056	.00054	.00052	.00050
3.3	.00048	.00047	.00045	.00043	.00042	.00040	.00039	.00038	.00036	.00035
3.4	.00034	.00032	.00031	.00030	.00029	.00028	.00027	.00026	.00025	.00024
3.5	.00023	.00022	.00022	.00021	.00020	.00019	.00019	.00018	.00017	.00017
3.6	.00016	.00015	.00015	.00014	.00014	.00013	.00013	.00012	.00012	.00011
3.7	.00011	.00010	.00010	.00010	.00009	.00009	.00008	.00008	.00008	.00008
3.8	.00007	.00007	.00007	.00006	.00006	.00006	.00006	.00005	.00005	.00005
3.9	.00005	.00005	.00004	.00004	.00004	.00004	.00004	.00004	.00003	.00003
4.0	.000032									

Standard Normal Probabilities: Single tail areas corresponding to z-values for the standardised normal curve

So What Proportion of People DO Have IQ Scores Higher than 130???

What ***proportion*** of people do we expect to have IQ scores above 130?

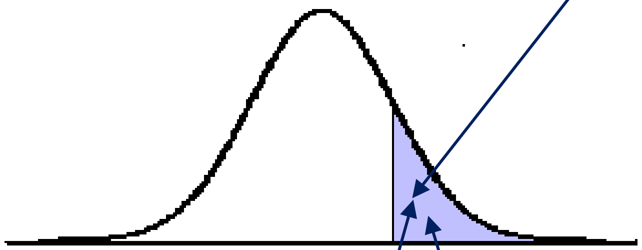
We found that 130 corresponds to a z-score of 2. From the z-table, the **area to the right** of z = 2 is 0.0228.

Therefore, we expect 0.0228, or 2.28% of all people to have IQ scores above 130.

Also, 1 – 0.0228 = 0.9772 = 97.72%. This is the proportion of people expected to have IQ scores less than 130.

So What is the Probability of Having an IQ Score Higher than 130???

So the **area 0.0228** represents:



- the ***proportion*** of all people with IQs above 130
- the ***probability*** that a randomly selected person will have an IQ above 130

NOTE: The probability that a randomly selected person has an IQ less than 70 is also 0.0228. An IQ of 70 would correspond to a z-score of –2 since 70 is 2 standard deviations *below* the mean.

Solving Probability Problems

- Draw a diagram
- Shade in the required area
- Calculate the z score
- Find the corresponding probability using z-table

The Platypus

The platypus is a semiaquatic mammal endemic to eastern Australia, including Tasmania. Together with the four species of echidna, it is one of the five extant species of monotremes which are the only mammals that lay eggs instead of giving birth

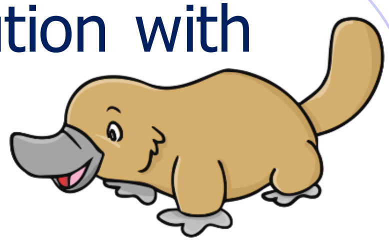
Some platypus statistics:

The body length of the adult male platypus follows a normal distribution with mean of 50cm and standard deviation of 2.45cm. Tail lengths also follow a normal distribution with a mean of 12.5cm and a standard deviation of 0.8cm.

We will find the percentage of adult male platypuses which are expected to have a body length greater than 45cm.

Finding a Probability

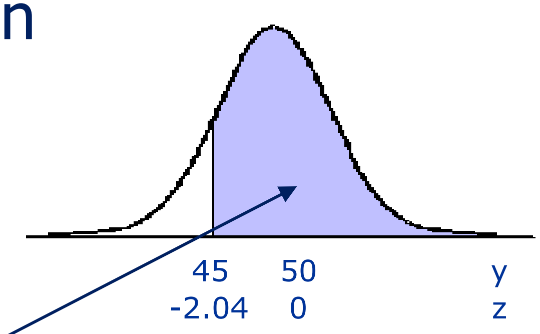
Body lengths of adult males follow a normal distribution with mean of 50cm and standard deviation of 2.45cm.



When y = 45

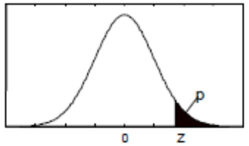
$$z = \frac{\text{distance between y and the population mean}}{\text{population standard deviation}}$$

$$= \frac{y - \mu}{\sigma} = \frac{45 - 50}{2.45} = -2.04$$



Tabulated area = area below z-score of –2.04 = 0.0207
Required area = 1 – 0.0207 = 0.9793
Therefore 97.93% of adult male platypuses are expected to have a body length greater than 45cm.

z-table



|z|	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641
0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
3.0	.00135	.00131	.00126	.00122	.00118	.00114	.00111	.00107	.00104	.00100
3.1	.00097	.00094	.00090	.00087	.00084	.00082	.00079	.00076	.00074	.00071
3.2	.00069	.00066	.00064	.00062	.00060	.00058	.00056	.00054	.00052	.00050
3.3	.00048	.00047	.00045	.00043	.00042	.00040	.00039	.00038	.00036	.00035
3.4	.00034	.00032	.00031	.00030	.00029	.00028	.00027	.00026	.00025	.00024
3.5	.00023	.00022	.00022	.00021	.00020	.00019	.00019	.00018	.00017	.00017
3.6	.00016	.00015	.00015	.00014	.00014	.00013	.00013	.00012	.00012	.00011
3.7	.00011	.00010	.00010	.00010	.00009	.00009	.00008	.00008	.00008	.00008
3.8	.00007	.00007	.00007	.00006	.00006	.00006	.00006	.00005	.00005	.00005
3.9	.00005	.00005	.00004	.00004	.00004	.00004	.00004	.00004	.00003	.00003
4.0	.000032									

Standard Normal Probabilities: Single tail areas corresponding to z-values for the standardised normal curve

Quiz 5

a. The tail of the platypus is mainly made up of a fatty tissue that is used to store energy supplies which the animal can use when there is a shortage of food. The tail also acts as a form of insulation and is used for steering while swimming. What proportion of adult male platypuses are expected to have tails between 10cm and 15cm in length?

tails
$\mu = 12.5$
$\sigma = 0.8$

b. What is the probability of an adult male platypus having a tail which is shorter than 12cm in length?

c. Of 12 male platypuses at Taronga Zoo, how many are expected to have tails which are longer than 12cm in length?

d. Is it unusual for an adult male platypus to have a body length which is longer than 60cm in length?

bodies
$\mu = 50$
$\sigma = 2.45$

Solution to Quiz 5a

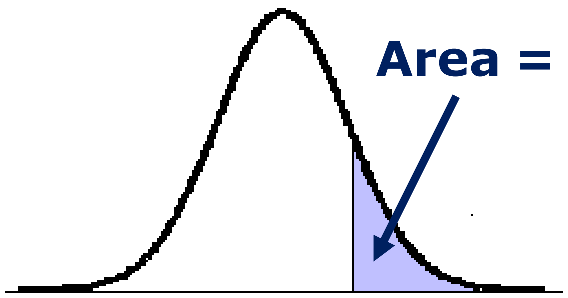
Solution to Quiz 5b

Solution to Quiz 5c

Solution to Quiz 5d

Finding a z-score from an Area

What is the z-score that cuts off a right hand tail area of 0.20?



z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611

The closest tabulated area to 0.2 is 0.2005 (circled).
This tabulated area corresponds to a z-score of 0.84.

Quiz 6

Use your normal probability tables to complete the table on the next slide.

Find the probabilities of obtaining z-scores greater than 0 and greater than 5.

Also find the values of z which cut off probabilities (use right hand tail areas) of 0.0630, 0.0093, 0.1442, 0.00098 and 0.0253 in the right tail.

Quiz 6

z-score	tabulated area
0	
	0.0630
	0.0093
	0.1442
	0.00098
	0.0253
5	

Calculating a y value from a z-score

$$z = \frac{y-\mu}{\sigma} \text{ is equivalent to } y = \mu + \sigma z$$

Find the y value
ie. the raw score

So, given a z-score, we can find the corresponding y-value from any normal distribution.

Example: What is the IQ score of a person who has a z-score corresponding to 2.2?

ie. a person who has an IQ score 2.2 standard deviations above the mean

Let y represent IQ: $y = \mu + \sigma z$

$= 100 + 15 \times 2.2$

$= 133$

Finding Percentiles of a Normal Distribution

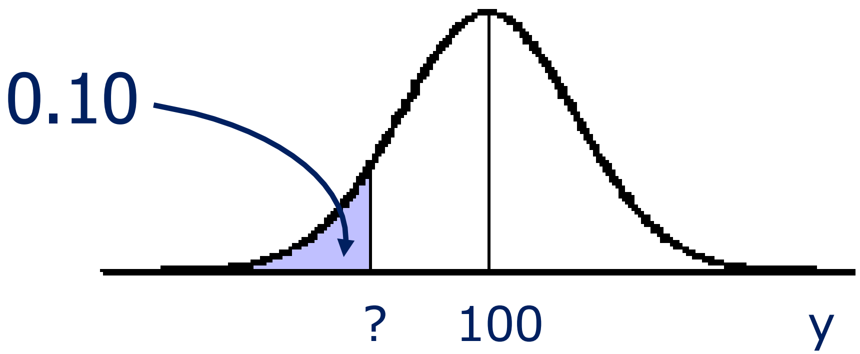
- Recall: the p^{th} percentile is the value such that p% of the values in the distribution are lower than it. **For example, the 50^{th} percentile (the median) has 50% of the values lower than it.**

- We can use the method shown in the last few slides to obtain percentiles from any normal distribution (provided we know μ and σ, of course!)

Example: Finding a Percentile of a Normal Distribution

Find the **IQ score** corresponding to the **10th percentile** of the population. IQ scores follow a normal distribution with a mean of 100 and a standard deviation of 15.

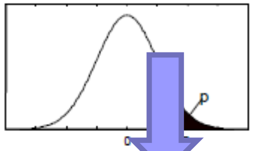
Find the y value ie. the raw score



$\mu = 100$
$\sigma = 15$

Since we are looking for the 10th percentile, 0.1 is the area in the left hand tail, **below** y.

z-table



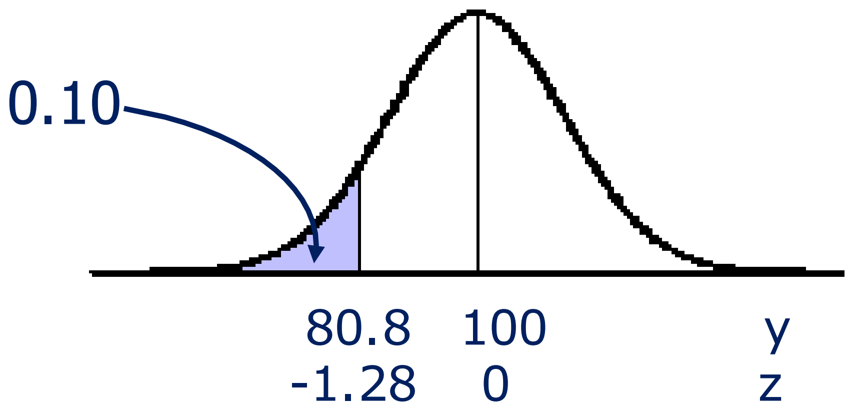
Closest area to 0.1

|z|	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641
0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	[illegible]
1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
3.0	.00135	.00131	.00126	.00122	.00118	.00114	.00111	.00107	.00104	.00100
3.1	.00097	.00094	.00090	.00087	.00084	.00082	.00079	.00076	.00074	.00071
3.2	.00069	.00066	.00064	.00062	.00060	.00058	.00056	.00054	.00052	.00050
3.3	.00048	.00047	.00045	.00043	.00042	.00040	.00039	.00038	.00036	.00035
3.4	.00034	.00032	.00031	.00030	.00029	.00028	.00027	.00026	.00025	.00024
3.5	.00023	.00022	.00022	.00021	.00020	.00019	.00019	.00018	.00017	.00017
3.6	.00016	.00015	.00015	.00014	.00014	.00013	.00013	.00012	.00012	.00011
3.7	.00011	.00010	.00010	.00010	.00009	.00009	.00008	.00008	.00008	.00008
3.8	.00007	.00007	.00007	.00006	.00006	.00006	.00006	.00005	.00005	.00005
3.9	.00005	.00005	.00004	.00004	.00004	.00004	.00004	.00004	.00003	.00003
4.0	.000032									

Standard Normal Probabilities: Single tail areas corresponding to z-values for the standardised normal curve

Example: Finding a Percentile of a Normal Distribution

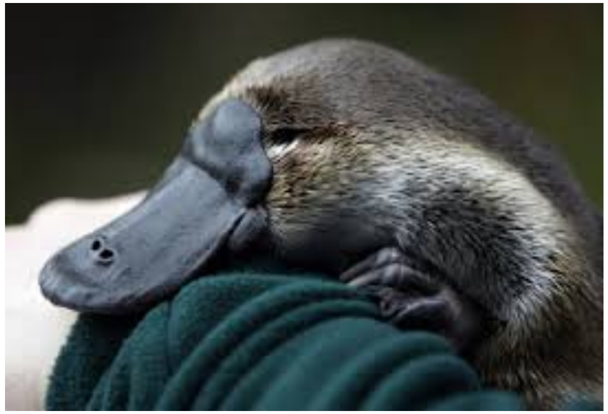
Find the **IQ score** corresponding to the **10^{th} percentile** of the population. IQ scores follow a normal distribution with a mean of 100 and a standard deviation of 15.



z = –1.28 (notice the negative sign since 0.1 is the area in the left hand tail)

$$y = \mu + \sigma z$$
$$= 100 + 15 \times (-1.28)$$
$$= 80.8$$ ie. an IQ of approximately 81.

Quiz 7



Recall the platypus. The platypus' bill is covered with a soft, moist, naked, leathery skin, which, when touched, feels soft and rubbery. The positioning of the bill allows it to breathe while the rest of its body is submerged.

Adult female platypuses have an average bill length of 5.2cm with a standard deviation of 0.25cm. Assuming bill lengths follow a normal distribution, find the length of a bill for an adult female on the 85th percentile of the distribution.

Solution to Quiz 7

Homework Questions

Homework Question 1

A bank audit on 30/6/2012 recorded the balances of a number of school savings accounts. It was found that the balances followed a normal distribution with a mean of \$230 and a standard deviation of \$75:

a. The accounts found to have the lowest 15% of balances are to be offered a saving incentive. What is the highest balance which will ensure the child is offered the incentive?

b. What percentage of school saving accounts were found to have balances between \$200 and \$300?

c. From a randomly selected sample of 500 school savings accounts, how many are expected to have balances over \$230?

Solution to Homework Question 1a

Solution to Homework Question 1b

Solution to Homework Question 1c

Homework Question 2

Where possible give an example of a 5-number data set where:

a. there is no right whisker.

b. the mean is negative.

c. the standard deviation is negative.

d. the range is more than twice the IQR.

e. the range is equal to the IQR.

f. the standard deviation is equal to zero.

Solution to Homework Question 2

Lecture 4 Summary

- Standard normal distributions have a mean equal to 0 and standard deviation equal to 1.

- Other normal distributions can be converted to a standard normal distribution using:

$$z = \frac{y - \mu}{\sigma}$$

- For normal distributions:
  - z scores can be used to find the probability of finding values either greater or smaller than a particular value y.
  - the z-score (and hence y-value) cut-off for a specified area can be found.

Textbook References

Further information on the topics discussed in this lecture can be found in the:

Modern Statistics: An Introduction
by Don McNeil and Jenny Middledorp
(ISBN 9781486007011).

- Chapter 4: Pages 70 – 82