USING THE NORMAL DISTRIBUTION

Mr C's IB Standard Notes

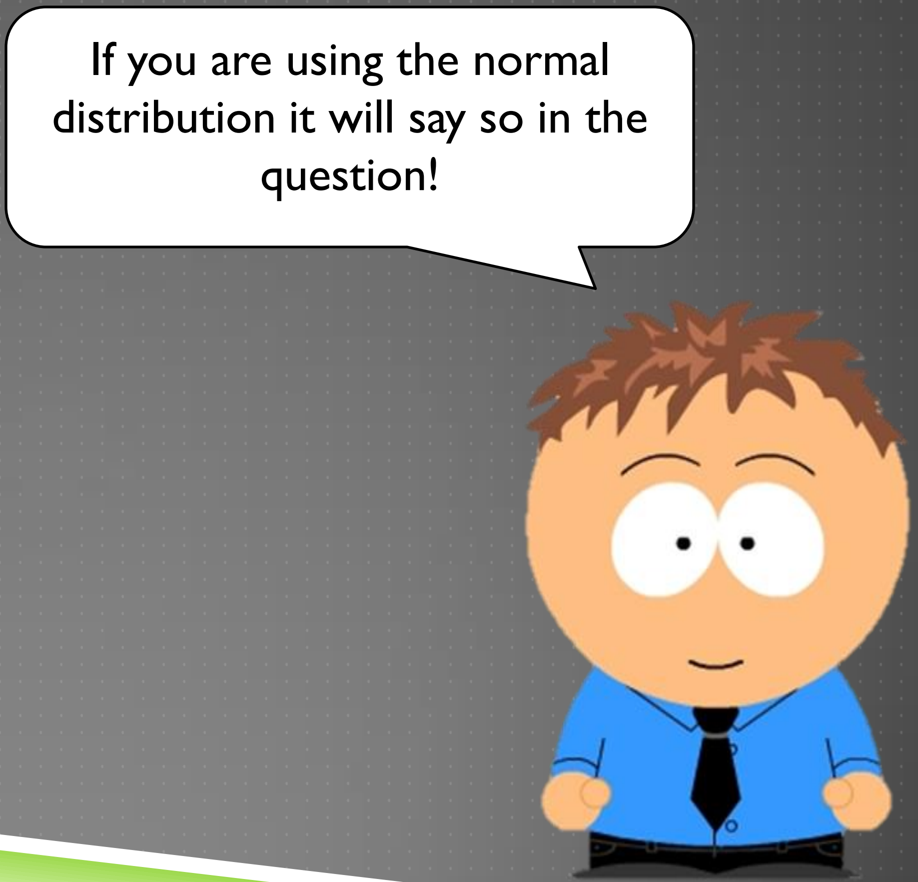
In this PDF you can find the following:

1. General Tips
2. Example Questions
3. Worked Solutions
4. Summary
5. Bonus Questions

GENERAL TIPS

Does the question involve the normal distribution?

E.g. The time taken for students to complete a test is **normally distributed** with a mean of 32 minutes and standard deviation of 6 minutes



GENERAL TIPS

<u>Always state the distribution</u>

E.g. The time taken for students to complete a test is **normally distributed** with a mean of 32 minutes and standard deviation of 6 minutes

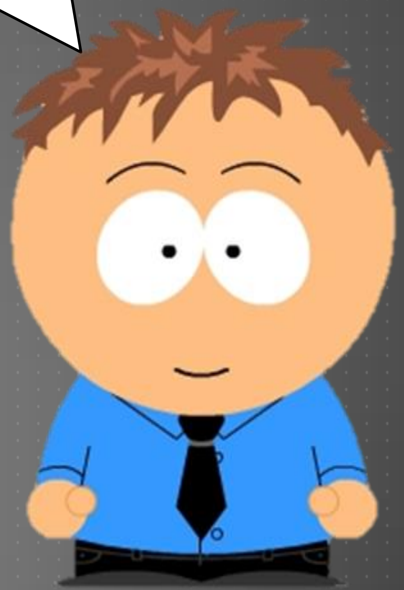
$$X \sim N(32, 6^2)$$

You get a mark if you use the information from the question and write down the distribution

$$X \sim N(\mu, \sigma^2)$$

mean

variance



GENERAL TIPS

<u>Always state the distribution</u>

E.g. Weights of rabbits are asssumed normally distributed with a mean of 2.6kg.

$$X \sim N(2.6, \sigma^2)$$

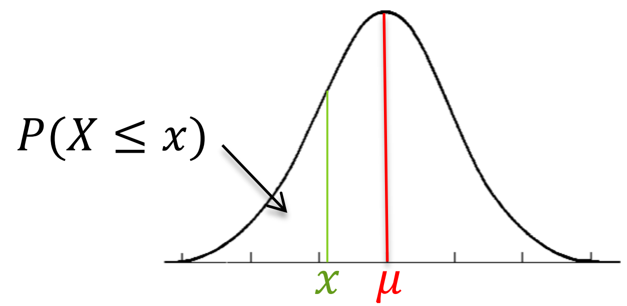
Even if you don't know one of the values, still state the distribution

GENERAL TIPS

Things you might need to work out

- The **mean** σ
- The **standard deviation** σ
- The **variance**
  (remember that's just standard deviation squared)
- **Areas**
  $P(x \text{ is less than some value})$
  $P(x \text{ is greater than some value})$

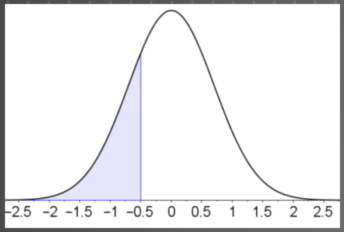
Drawing a picture will make it clearer what needs to be calculated

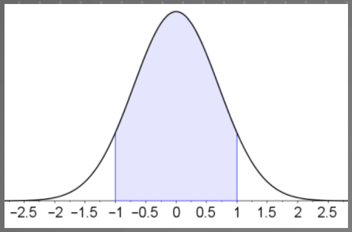


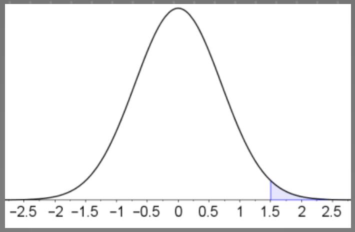
GENERAL TIPS

<u>Using the GDC - NormalCDF</u>

- Found in 2nd -> Vars menu
- Needs **mean** and **standard deviation** (not variance!)
- Needs upper and lower bounds.







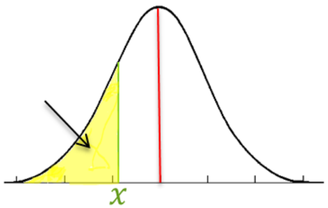
lower = -9999
upper = 0.5

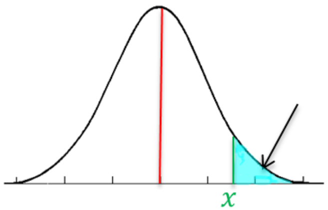
lower = -1
upper = 1

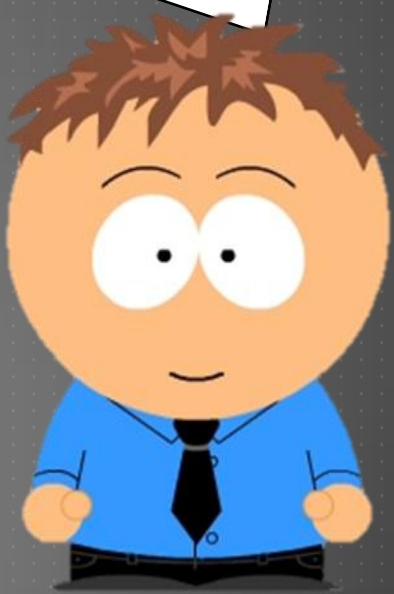
lower = 1.5
upper = 9999

You will use NormalCDF when trying to calculate probabilities below or above a certain value (x).

This corresponds to the areas on the curve





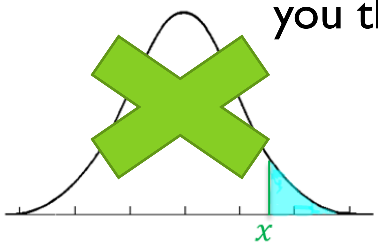


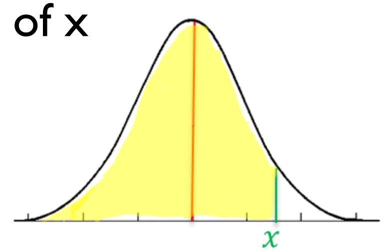
GENERAL TIPS

<u>Using the GDC - InvNorm</u>

- Found in 2^{nd} -> Vars menu
- Needs **area** (probability), **mean** and **standard deviation**

If you know what the probability/area already is, the inverse normal function tells you the value of x



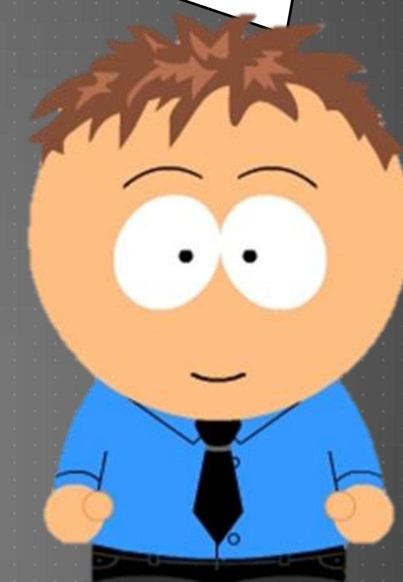


It always uses the area to the left of the value of x (i.e. $P(X \leq x)$)

Note. It makes no difference if the inequality is strict (≤) or not (<). The data is continuous, and so:

$$P(X \leq 10) = P(X < 10)$$

This is different if you use the binomial distribution, where: $P(X \leq 10) = P(X < 11)$



EXAMPLES

In each of the below examples the variable that is assumed normally distributed.

IQ tests have a mean of 100 and standard deviation of 20

What is the probability a randomly selected person has an IQ above 120?

1.

What score must you achieve to be in the top 5%?

2.

A supermarket needs to perform quality checks on some of their produce. A sample of 150 apples is collected.

The mean weight is 125g, and 30% weigh more than 140g. Estimate the standard deviation.

3.

The standard deviation is expected to be 8g. What should the mean have been?

4.

SOLUTIONS

IQ tests have a mean of 100 and standard deviation of 20

1.

What is the probability a randomly selected person has an IQ above 120?

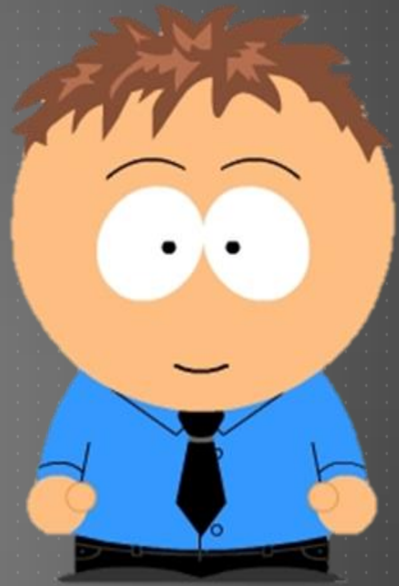
SOLUTIONS

IQ tests have a mean of 100 and standard deviation of 20

1.

What is the probability a randomly selected person has an IQ above 120?

$X \sim N(100, 20^2)$



SOLUTIONS

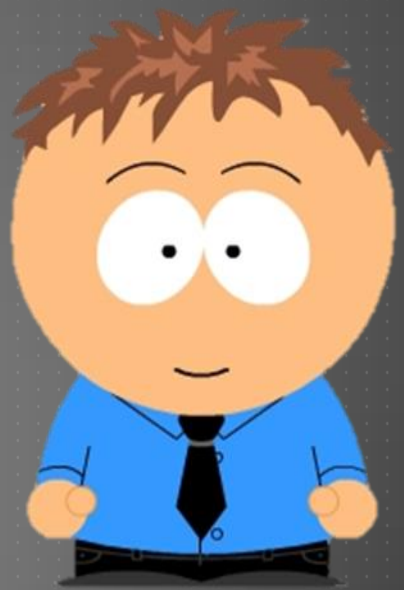
IQ tests have a mean of 100 and standard deviation of 20

1.

What is the probability a randomly selected person has an IQ above 120?

$X \sim N(100, 20^2)$

Draw a picture to show what you are trying to calculate



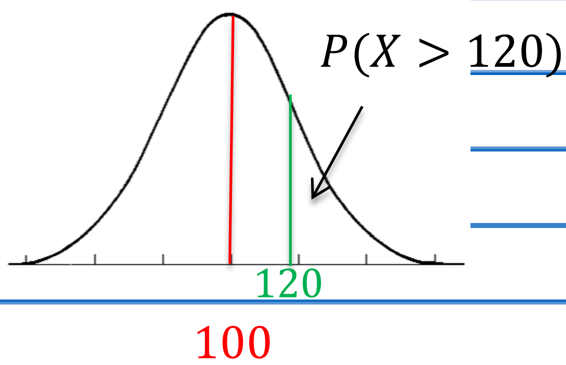
SOLUTIONS

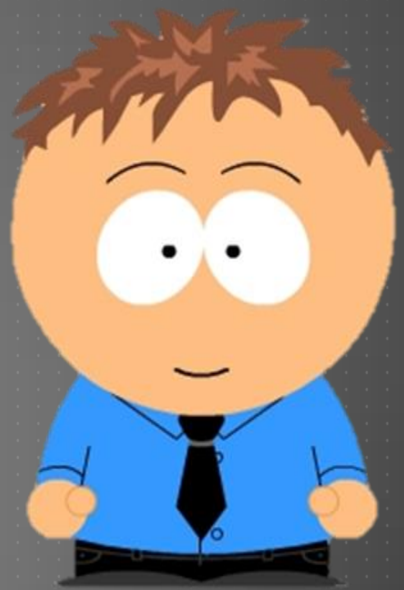
IQ tests have a mean of 100 and standard deviation of 20

1.

What is the probability a randomly selected person has an IQ above 120?

$X \sim N(100, 20^2)$





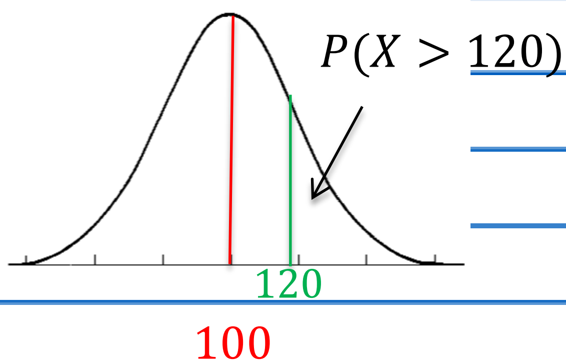
SOLUTIONS

IQ tests have a mean of 100 and standard deviation of 20

1.

What is the probability a randomly selected person has an IQ above 120?

$X \sim N(100, 20^2)$



Use GDC to calculate area

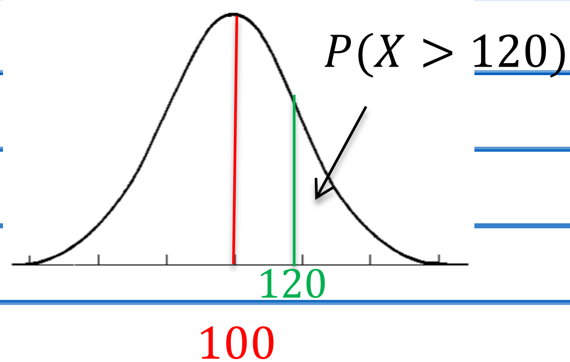
SOLUTIONS

IQ tests have a mean of 100 and standard deviation of 20

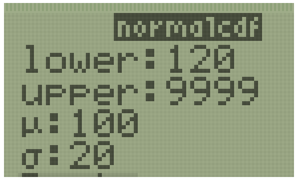
1.

What is the probability a randomly selected person has an IQ above 120?

$X \sim N(100, 20^2)$



Use GDC to calculate area



SOLUTIONS

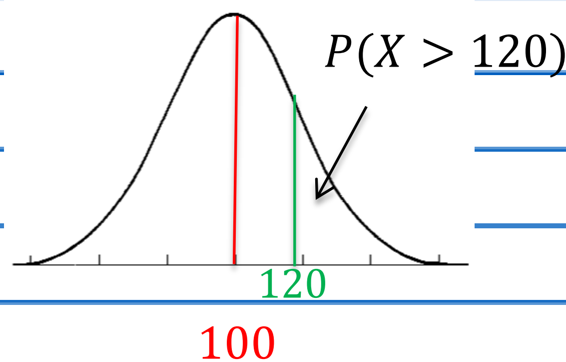
IQ tests have a mean of 100 and standard deviation of 20

1.

What is the probability a randomly selected person has an IQ above 120?

$X \sim N(100, 20^2)$

Using GDC
$P(X > 100) = 0.15$



Use GDC to calculate area

```
normalcdf(120,9999,100,20)
    .1586552596
```

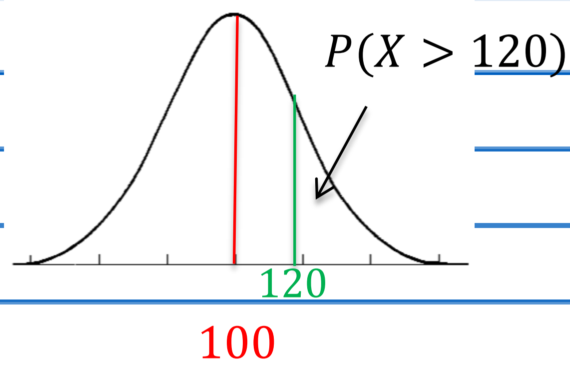
SOLUTIONS

IQ tests have a mean of 100 and standard deviation of 20

1.

What is the probability a randomly selected person has an IQ above 120?

$X \sim N(100, 20^2)$

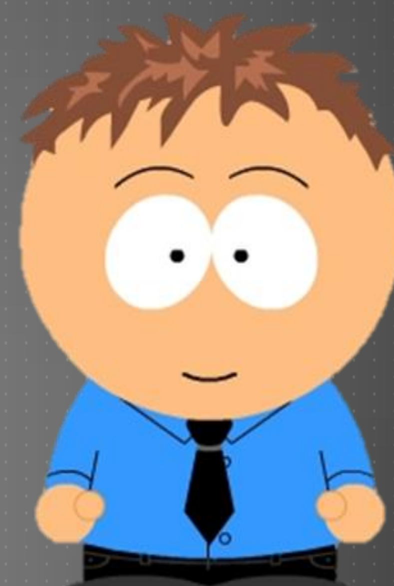


Using GDC

$P(X > 100) = 0.15$

Well done if you got that!

Let’s try the next question



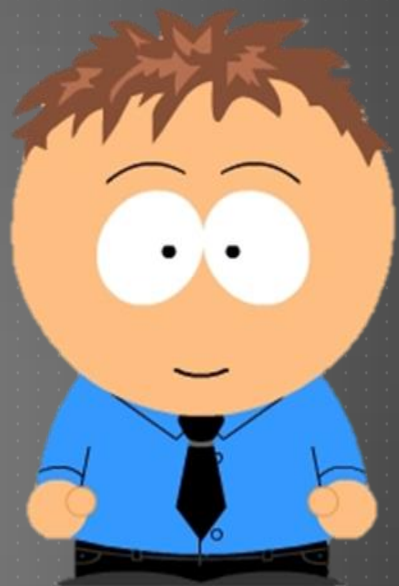
SOLUTIONS

IQ tests have a mean of 100 and standard deviation of 20

2.

What score must you achieve to be in the top 5%?

This question is a little bit different, we have the probability and need to find a value



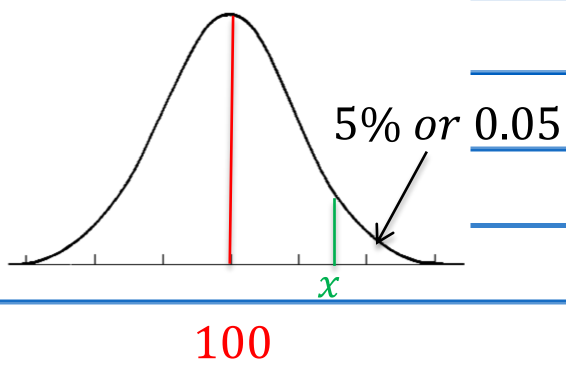
SOLUTIONS

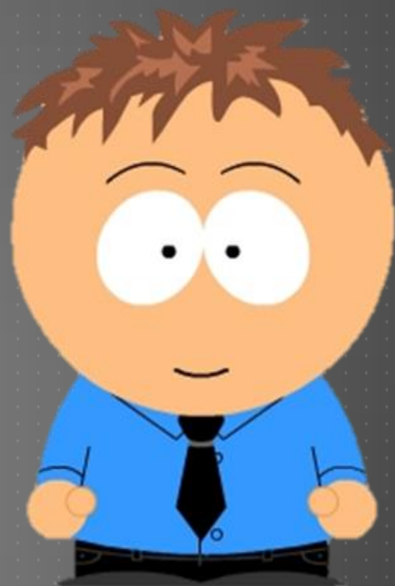
IQ tests have a mean of 100 and standard deviation of 20

2.

What score must you achieve to be in the top 5%?

$X \sim N(100, 20^2)$





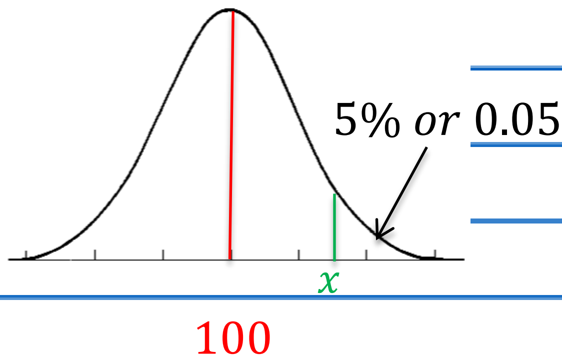
SOLUTIONS

IQ tests have a mean of 100 and standard deviation of 20

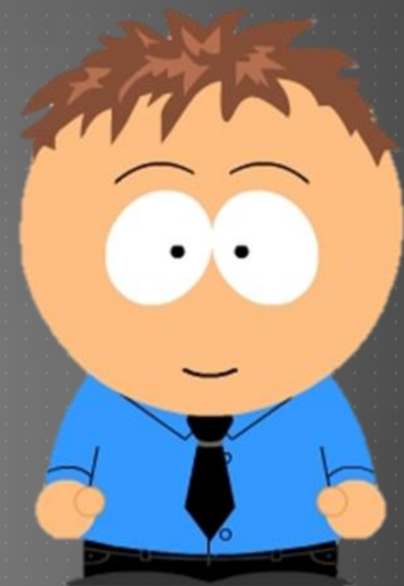
2.

What score must you achieve to be in the top 5%?

$X \sim N(100, 20^2)$



We want to use GDC, however we must remember to use area to the left



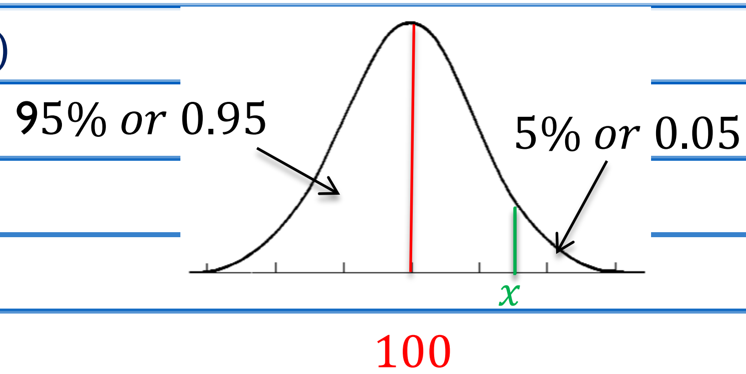
SOLUTIONS

IQ tests have a mean of 100 and standard deviation of 20

2.

What score must you achieve to be in the top 5%?

$X \sim N(100, 20^2)$



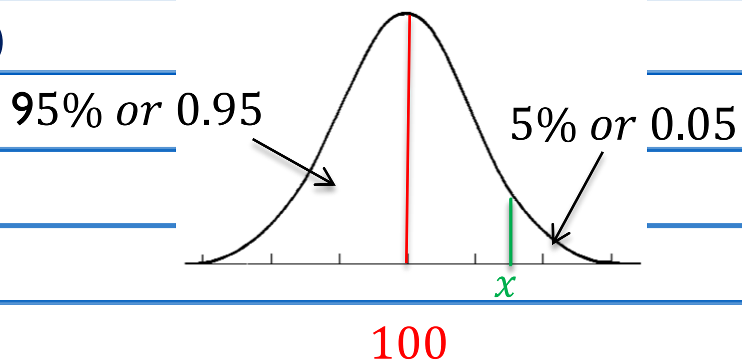
SOLUTIONS

IQ tests have a mean of 100 and standard deviation of 20

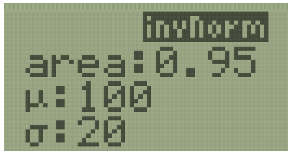
2.

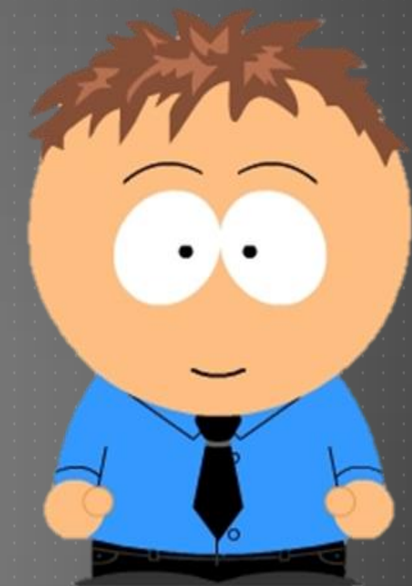
What score must you achieve to be in the top 5%?

$X \sim N(100, 20^2)$



We have area, so use InvNorm





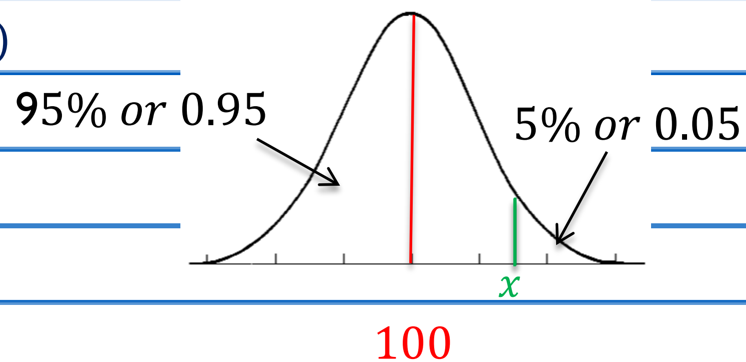
SOLUTIONS

IQ tests have a mean of 100 and standard deviation of 20

2.

What score must you achieve to be in the top 5%?

$X \sim N(100, 20^2)$

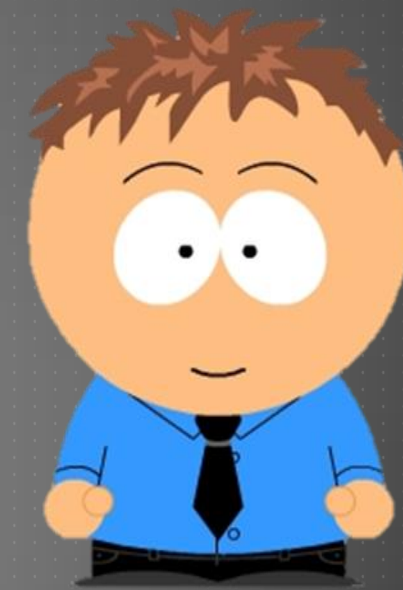


Using GDC

$x = 133$

We have area, so use InvNorm

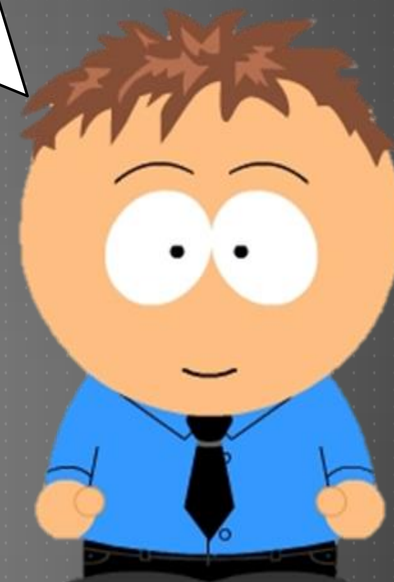
```
invNorm(0.95,100,20)
           132.897072
```



I hope those two examples made sense. In the next two we'll introduce the **standard normal distribution**, which is useful if mean or standard deviation is unknown.

$$z = \frac{x-\mu}{\sigma} \quad \text{(in formula book)}$$

Before answering questions 3 and 4 we'll take a quick look at standardizing, and why it's so useful.



STANDARD DISTRIBUTION

$X \sim N(\mu, \sigma^2)$

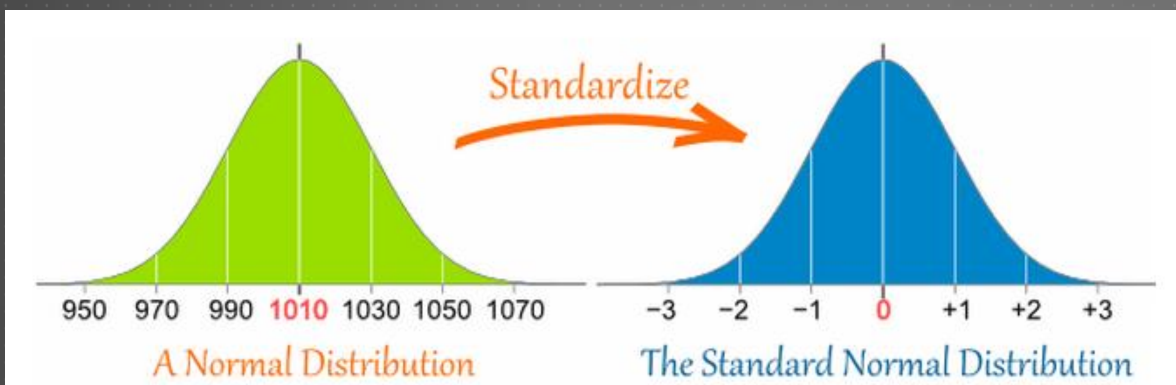
- Subtract the mean
- Divide by the standard deviation
- Call new distribution Z

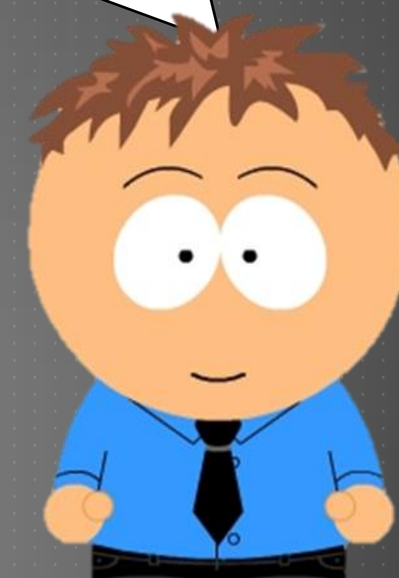
$Z \sim N(0, 1^2)$

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

By doing this we can convert from a distribution with any mean and standard deviation, to one where $\mu = 0$ and $\sigma = 1$

This is perfect if we didn't know one of the values to begin with





EXAMPLES

In each of the below examples the variable that is assumed normally distributed.

IQ tests have a mean of 100 and standard deviation of 20

What is the probability a randomly selected person has an IQ above 120?

1.

What score must you achieve to be in the top 5%

2.

A supermarket needs to perform quality checks on some of their produce. A sample of 150 apples is collected.

The mean weight is 125g, and 30% weigh more than 140g. Estimate the standard deviation.

3.

The standard deviation is expected to be 8g. What should the mean have been?

4.

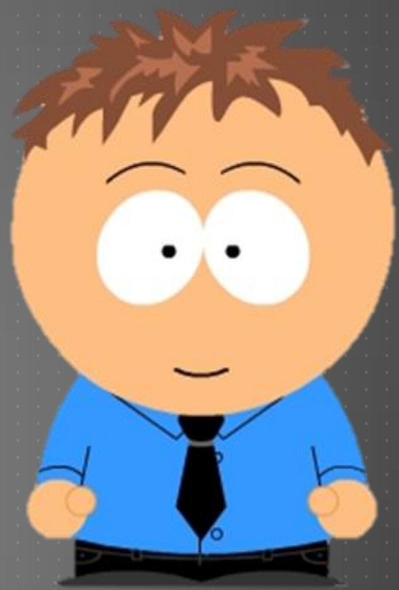
SOLUTIONS

A supermarket needs to perform quality checks on some of their produce. A sample of 150 apples is collected.

3.

The mean weight is 125g, and 30% weigh more than 140g. Estimate the standard deviation.

As usual we start by writing the distribution



SOLUTIONS

A supermarket needs to perform quality checks on some of their produce. A sample of 150 apples is collected.

3.

The mean weight is 125g, and 30% weigh more than 140g. Estimate the standard deviation.

$X \sim N(125, \sigma^2)$

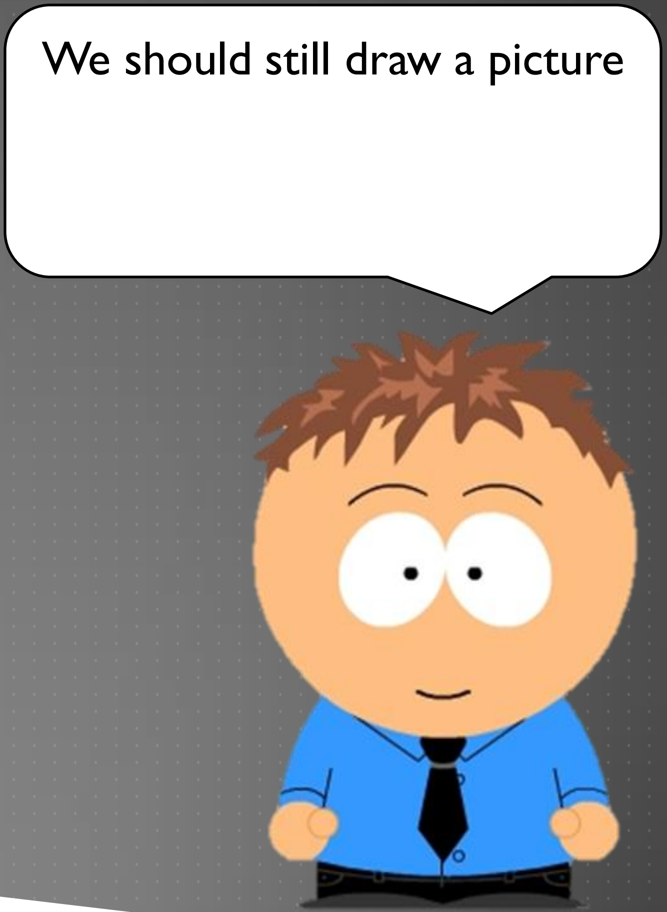
SOLUTIONS

3.

A supermarket needs to perform quality checks on some of their produce. A sample of 150 apples is collected.

The mean weight is 125g, and 30% weigh more than 140g. Estimate the standard deviation.

$X \sim N(125, \sigma^2)$



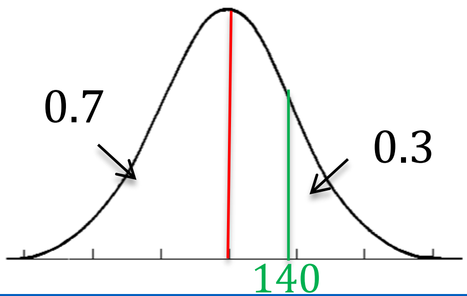
SOLUTIONS

3.

A supermarket needs to perform quality checks on some of their produce. A sample of 150 apples is collected.

The mean weight is 125g, and 30% weigh more than 140g. Estimate the standard deviation.

$X \sim N(125, \sigma^2)$



$\mu = 125 \quad \sigma = ?$

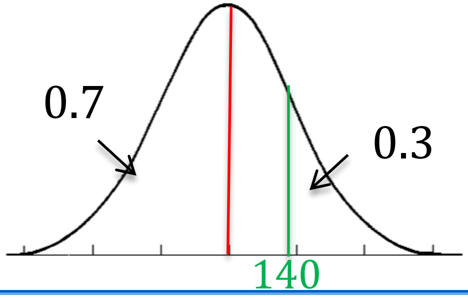
SOLUTIONS

3.

A supermarket needs to perform quality checks on some of their produce. A sample of 150 apples is collected.

The mean weight is 125g, and 30% weigh more than 140g. Estimate the standard deviation.

$X \sim N(125, \sigma^2)$



$\mu = 125 \quad \sigma = ?$

Since standard deviation is unknown we must standardise.

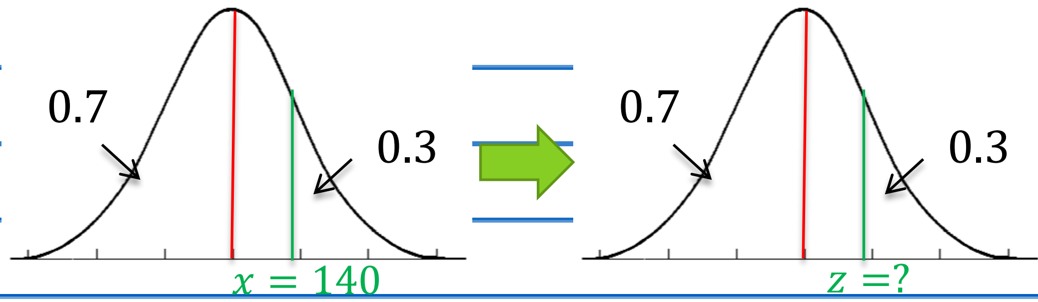
SOLUTIONS

3.

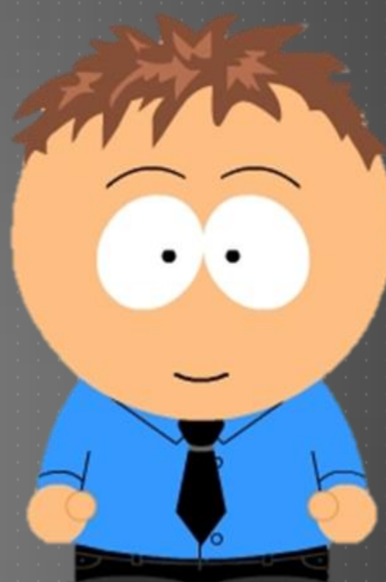
A supermarket needs to perform quality checks on some of their produce. A sample of 150 apples is collected.

The mean weight is 125g, and 30% weigh more than 140g. Estimate the standard deviation.

$X \sim N(125, \sigma^2)$



$\mu = 125 \quad \sigma = ?$ $\qquad \mu = 0 \quad \sigma = 1$



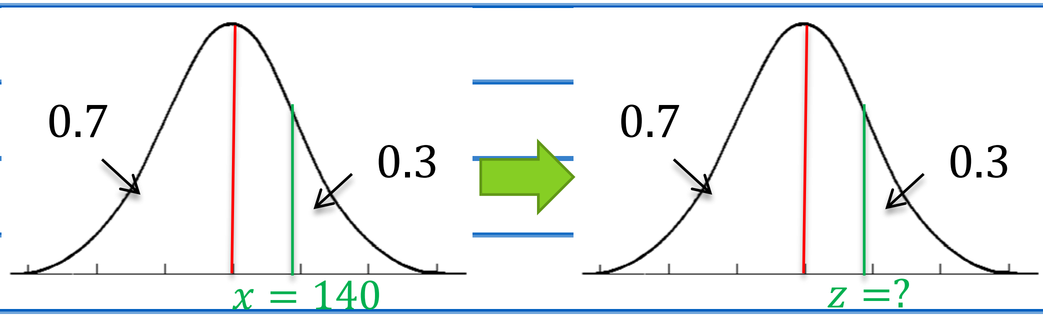
SOLUTIONS

3.

A supermarket needs to perform quality checks on some of their produce. A sample of 150 apples is collected.

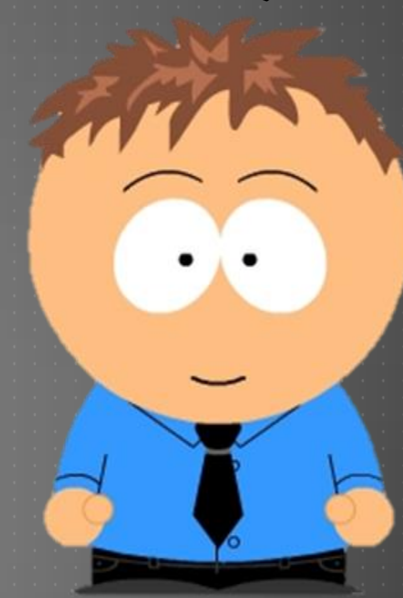
The mean weight is 125g, and 30% weigh more than 140g. Estimate the standard deviation.

$X \sim N(125, \sigma^2)$



$\mu = 125 \quad \sigma = ?$ $\qquad \mu = 0 \quad \sigma = 1$

We now have to calculate the new value of z.
This is just like question 2
(area known, value unknown)



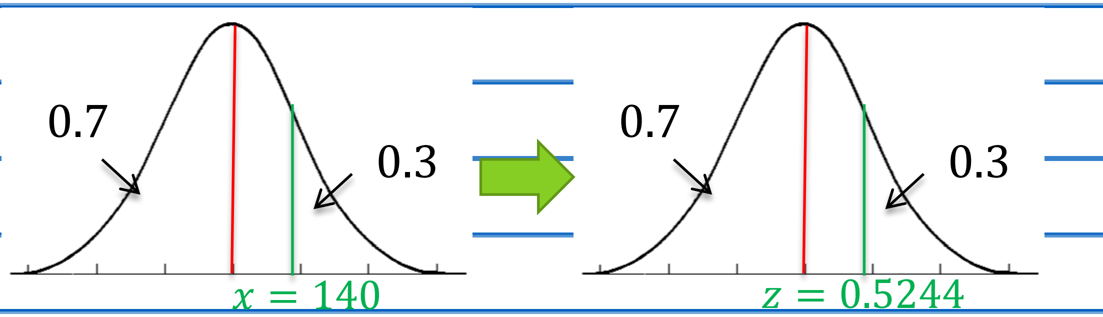
SOLUTIONS

A supermarket needs to perform quality checks on some of their produce. A sample of 150 apples is collected.

3.

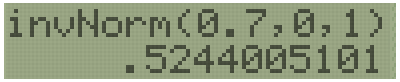
The mean weight is 125g, and 30% weigh more than 140g. Estimate the standard deviation.

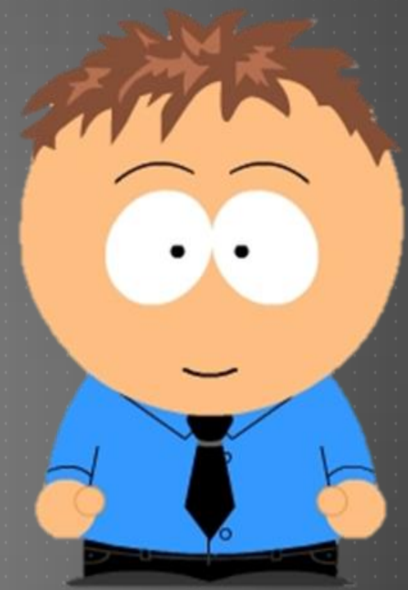
$X \sim N(125, \sigma^2)$



$\mu = 125 \quad \sigma = ?$ $\qquad \mu = 0 \quad \sigma = 1$

Use GDC





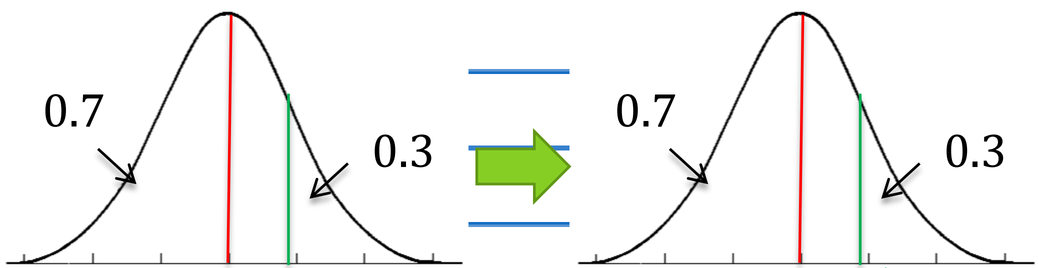
SOLUTIONS

3.

A supermarket needs to perform quality checks on some of their produce. A sample of 150 apples is collected.

The mean weight is 125g, and 30% weigh more than 140g. Estimate the standard deviation.

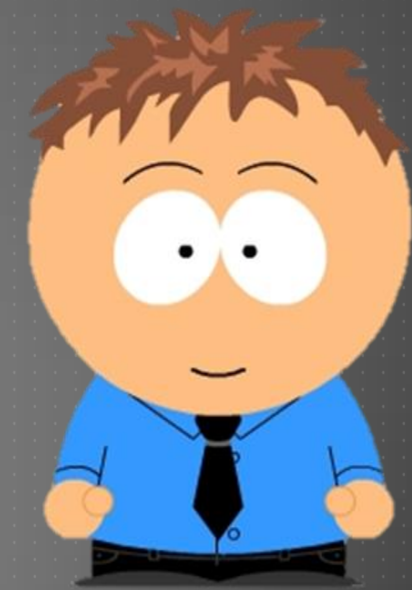
$X \sim N(125, \sigma^2)$



$x = 140$ $\qquad$ $z = 0.5244$

$\mu = 125 \quad \sigma = ?$ $\qquad$ $\mu = 0 \quad \sigma = 1$

The final step is to convert back to find the original standard deviation



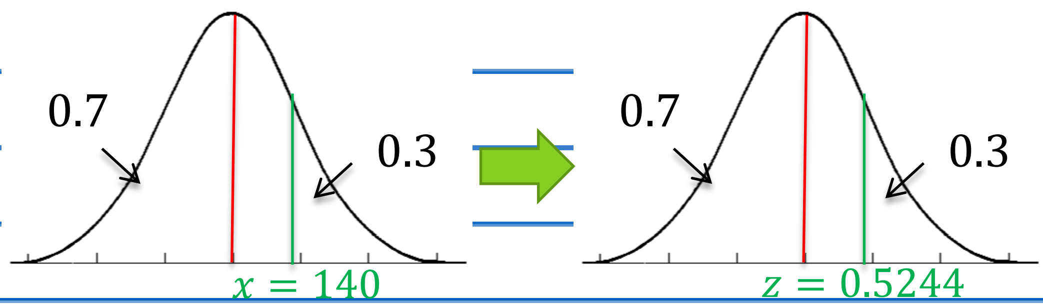
SOLUTIONS

3.

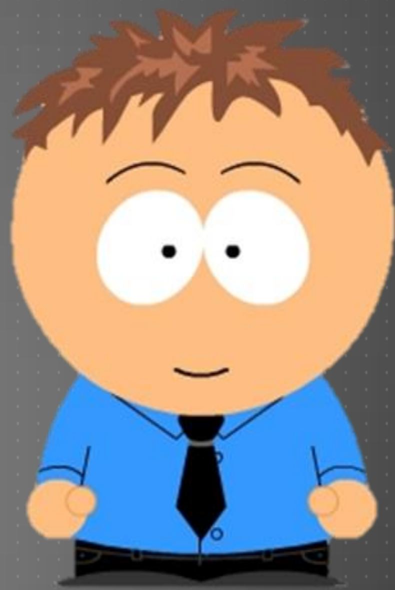
A supermarket needs to perform quality checks on some of their produce. A sample of 150 apples is collected.

The mean weight is 125g, and 30% weigh more than 140g. Estimate the standard deviation.

$X \sim N(125, \sigma^2)$



$\mu = 125 \quad \sigma = ?$ $\qquad$ $\mu = 0 \quad \sigma = 1$



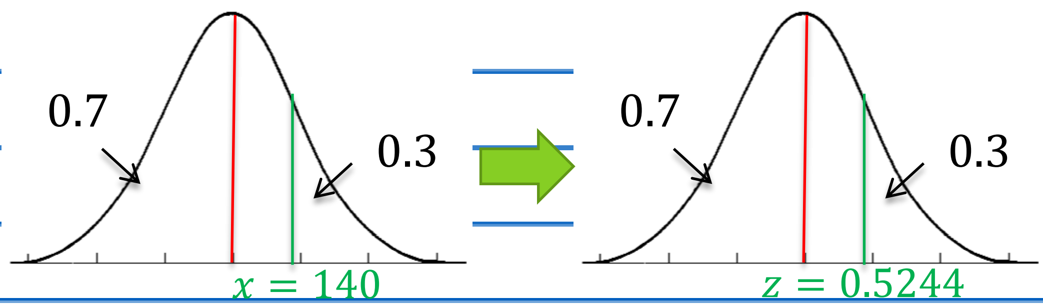
SOLUTIONS

A supermarket needs to perform quality checks on some of their produce. A sample of 150 apples is collected.

3.

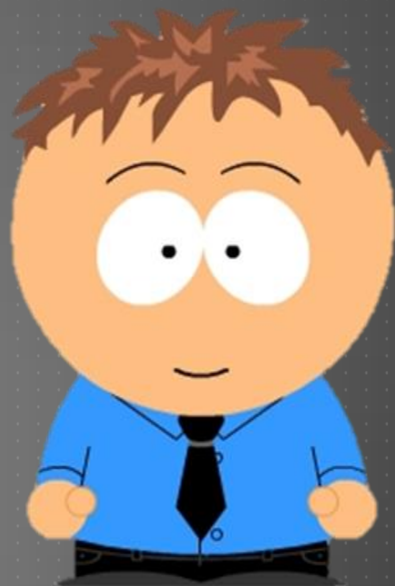
The mean weight is 125g, and 30% weigh more than 140g. Estimate the standard deviation.

$X \sim N(125, \sigma^2)$



$\mu = 125 \quad \sigma = ?$ $\qquad$ $\mu = 0 \quad \sigma = 1$

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



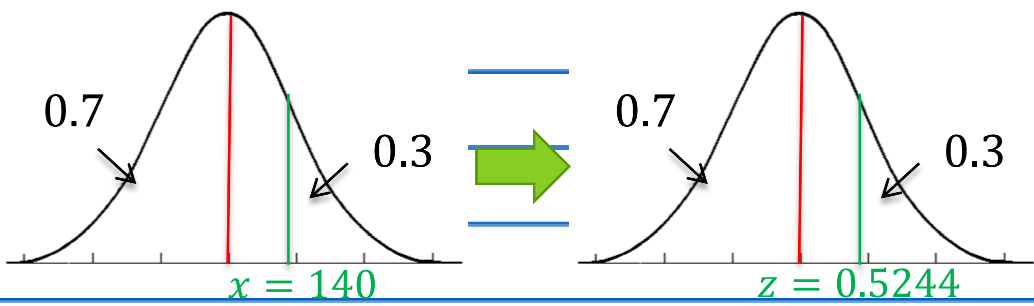
SOLUTIONS

3.

A supermarket needs to perform quality checks on some of their produce. A sample of 150 apples is collected.

The mean weight is 125g, and 30% weigh more than 140g. Estimate the standard deviation.

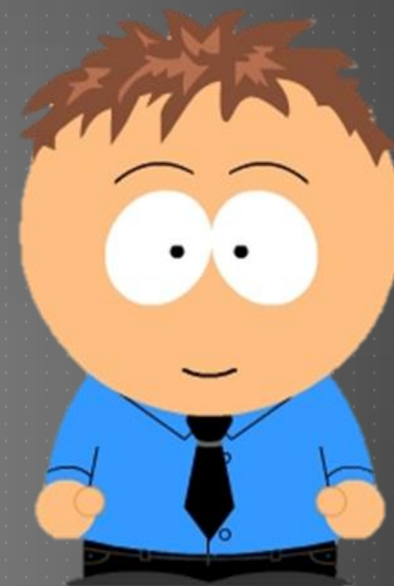
$X \sim N(125, \sigma^2)$



$\mu = 125 \quad \sigma = ?$ $\qquad \mu = 0 \quad \sigma = 1$

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

Substitute in values and rearrange to find σ



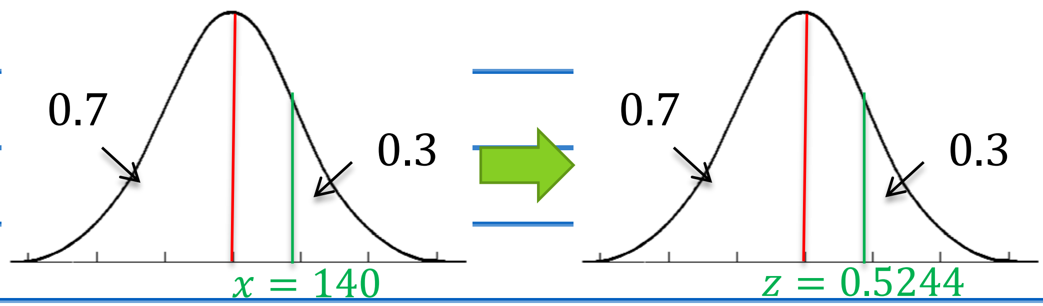
SOLUTIONS

3.

A supermarket needs to perform quality checks on some of their produce. A sample of 150 apples is collected.

The mean weight is 125g, and 30% weigh more than 140g. Estimate the standard deviation.

$X \sim N(125, \sigma^2)$

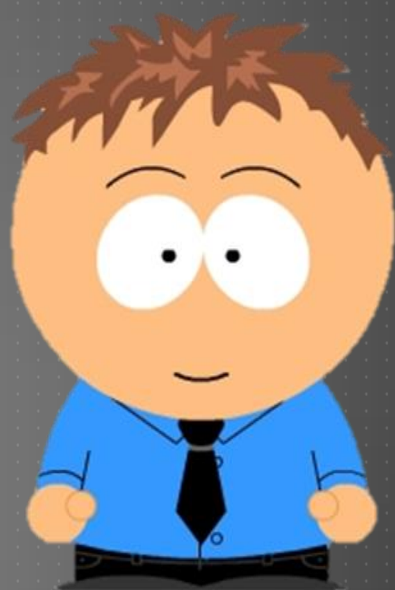


$\mu = 125 \quad \sigma = ?$ $\qquad \mu = 0 \quad \sigma = 1$

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

$$0.5244 = \frac{140 - 125}{\sigma}$$

$$\sigma = \frac{140 - 125}{0.5244} = \mathbf{28.6}$$



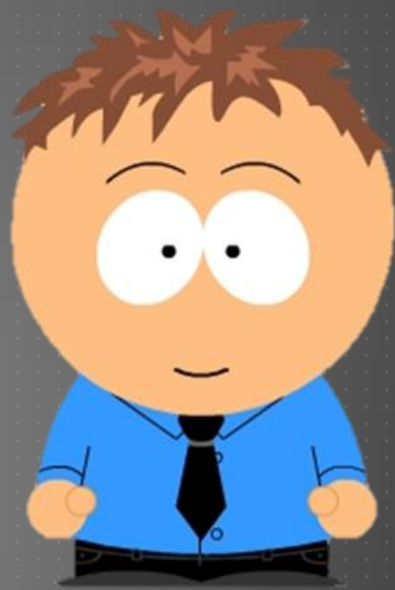
SOLUTIONS

4.

A supermarket needs to perform quality checks on some of their produce. A sample of 150 apples is collected.

The standard deviation is expected to be 8g. What should the mean have been?

This time we have the mean and need to work out the standard deviation.



SOLUTIONS

4.

A supermarket needs to perform quality checks on some of their produce. A sample of 150 apples is collected.

The standard deviation is expected to be 8g. What should the mean have been?

This will use the same method as the previous question. Try have a go yourself first

SOLUTIONS

A supermarket needs to perform quality checks on some of their produce. A sample of 150 apples is collected.

4.

The standard deviation is expected to be 8g. What should the mean have been?

$X \sim N(\mu, 8^2)$

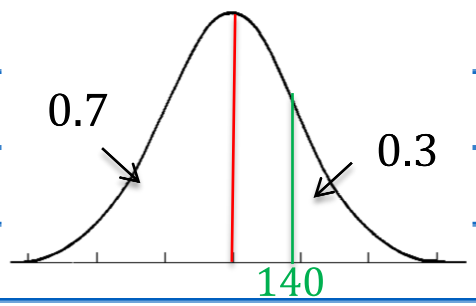
SOLUTIONS

4.

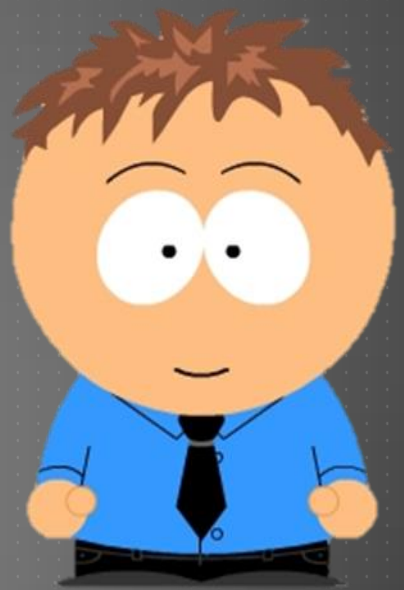
A supermarket needs to perform quality checks on some of their produce. A sample of 150 apples is collected.

The standard deviation is expected to be 8g. What should the mean have been?

$X \sim N(\mu, 8^2)$



$\mu = ?$ $\sigma = 8$



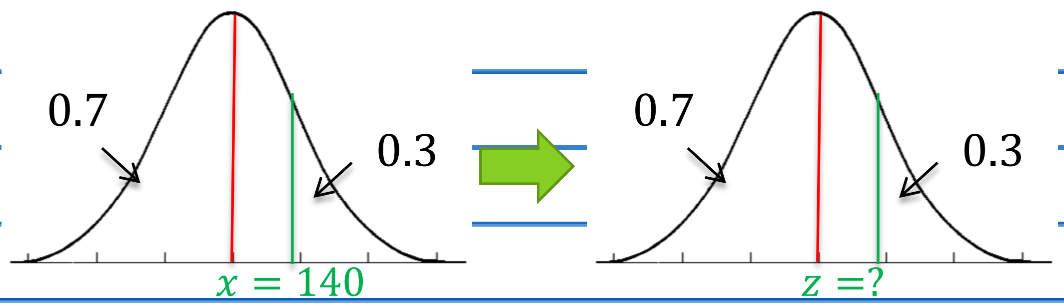
SOLUTIONS

4.

A supermarket needs to perform quality checks on some of their produce. A sample of 150 apples is collected.

The standard deviation is expected to be 8g. What should the mean have been?

$X \sim N(\mu, 8^2)$ $\qquad$ $X \sim N(\mu, 8^2)$



$\mu = ?$ $\quad$ $\sigma = 8$ $\qquad$ $\mu = 0$ $\quad$ $\sigma = 1$

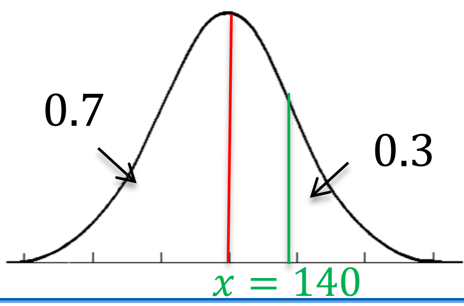
SOLUTIONS

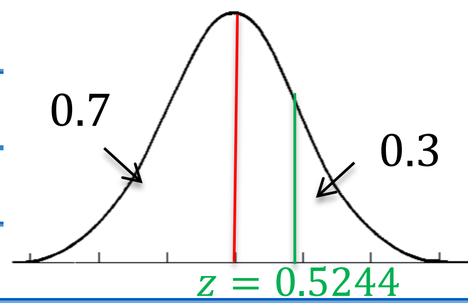
4.

A supermarket needs to perform quality checks on some of their produce. A sample of 150 apples is collected.

The standard deviation is expected to be 8g. What should the mean have been?

$X \sim N(\mu, 8^2)$ $\quad\quad$ $X \sim N(\mu, 8^2)$





$\mu = ?$ $\quad\quad$ $\sigma = 8$ $\quad\quad\quad\quad$ $\mu = 0$ $\quad\quad$ $\sigma = 1$

Using GDC

$z = 0.5244$

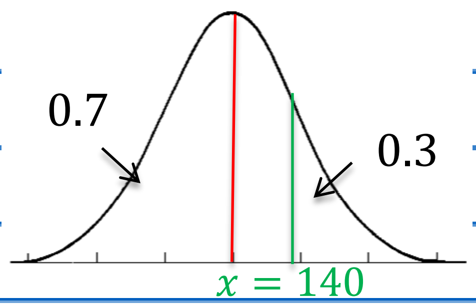
SOLUTIONS

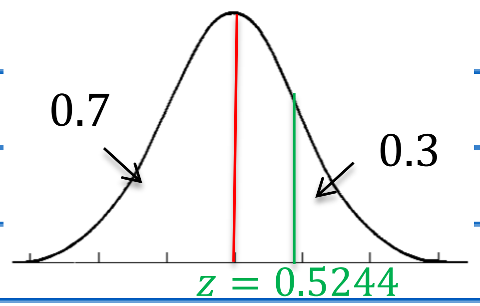
4.

A supermarket needs to perform quality checks on some of their produce. A sample of 150 apples is collected.

The standard deviation is expected to be 8g. What should the mean have been?

$X \sim N(\mu, 8^2)$	$X \sim N(\mu, 8^2)$



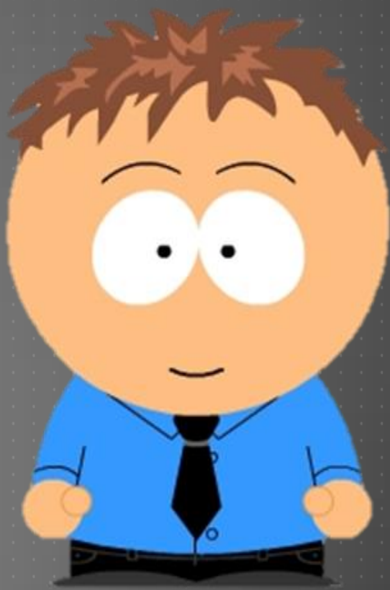


$\mu = ?$ $\sigma = 8$ $\mu = 0$ $\sigma = 1$

Using GDC

$z = 0.5244$ $\sigma z = x - \mu$

$z = \dfrac{x - \mu}{\sigma}$ $\mu = x - \sigma z$



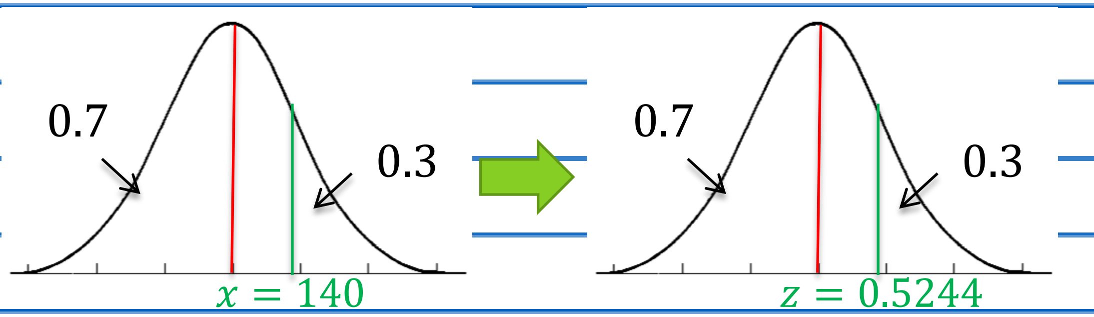
SOLUTIONS

4.

A supermarket needs to perform quality checks on some of their produce. A sample of 150 apples is collected.

The standard deviation is expected to be 8g. What should the mean have been?

$X \sim N(\mu, 8^2)$ $\quad$ $X \sim N(\mu, 8^2)$

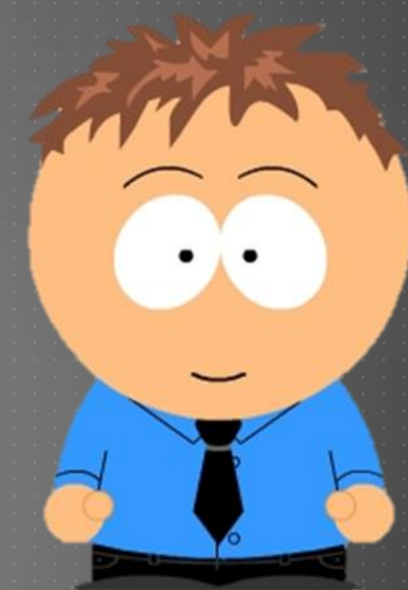


$\mu = ?$ $\quad \sigma = 8$ $\qquad \mu = 0$ $\quad \sigma = 1$

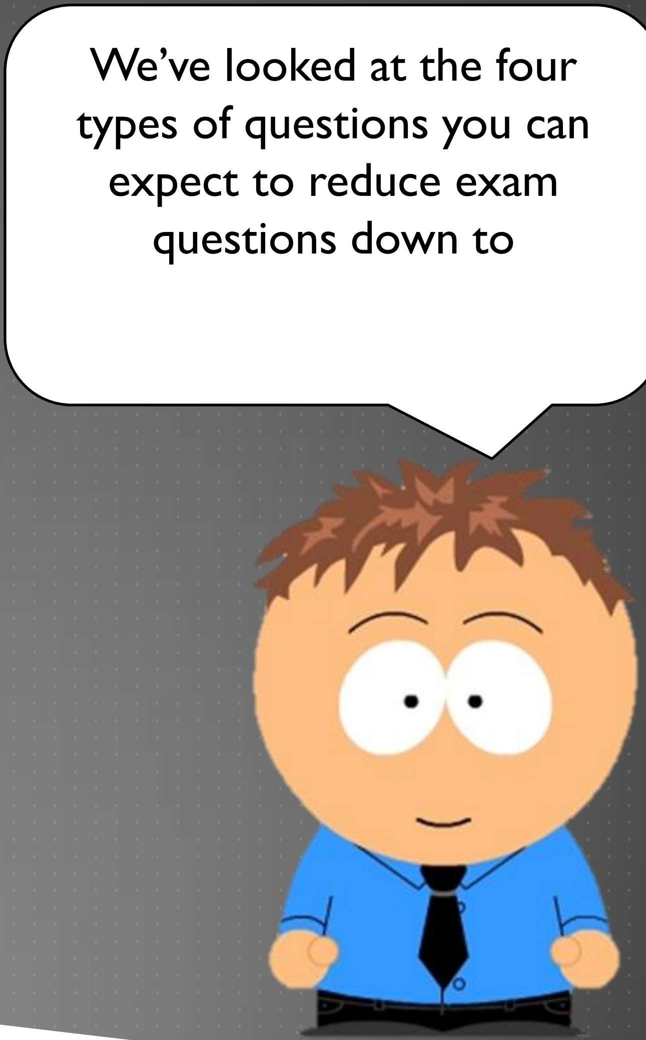
Using GDC

$z = 0.5244$ $\qquad \sigma z = x - \mu$ $\qquad \mu = 140 - 8 \times 0.5244$

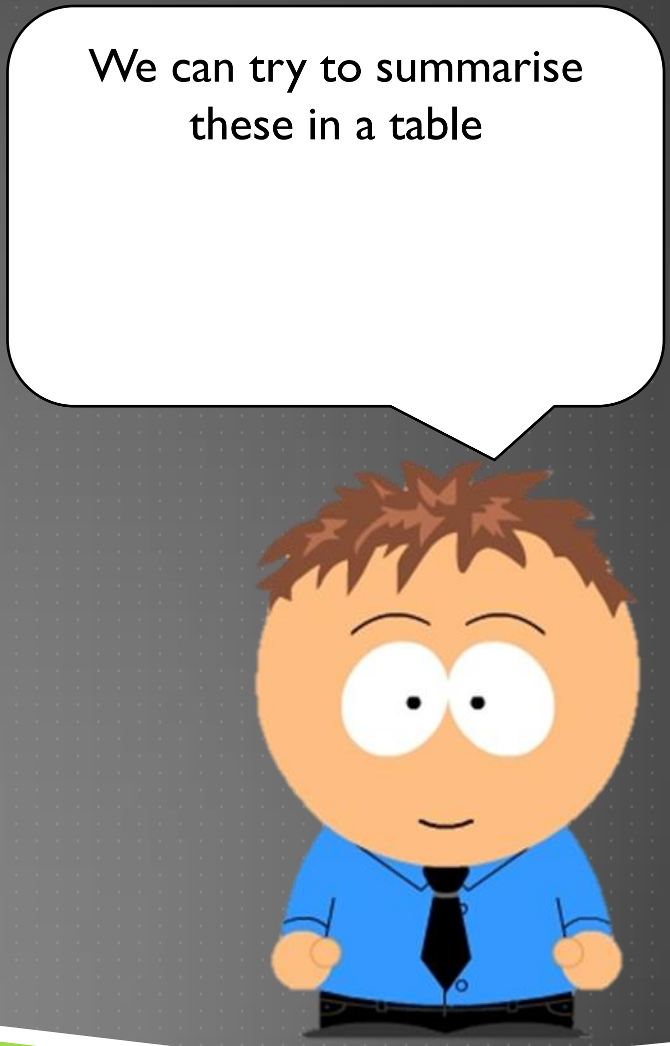
$z = \dfrac{x - \mu}{\sigma}$ $\qquad \mu = x - \sigma z$ $\qquad = \mathbf{136g}$



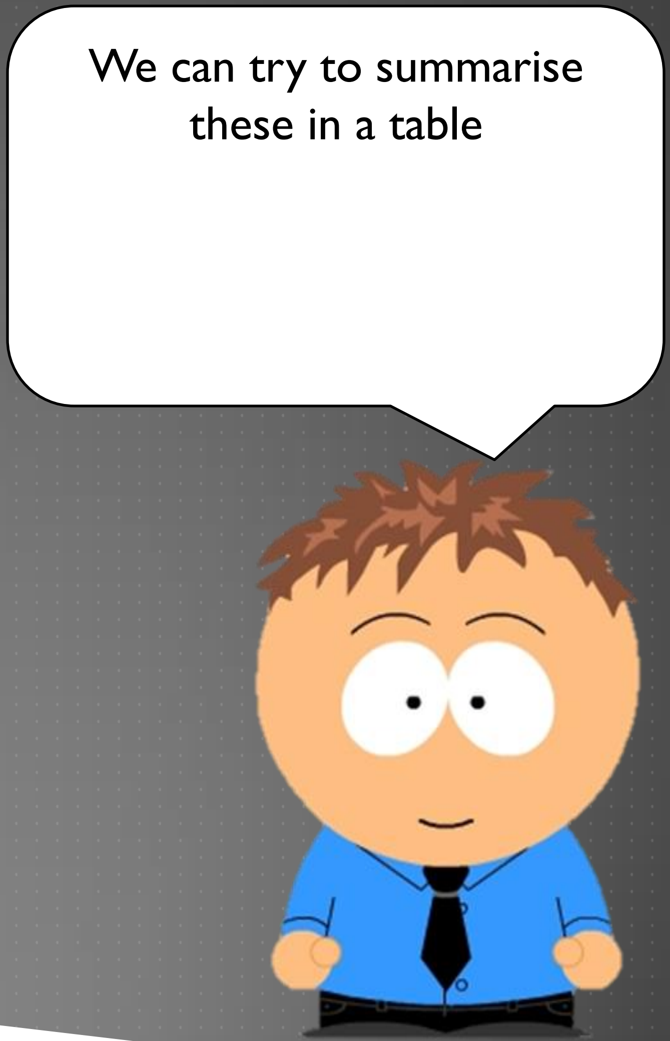
SUMMARY


We've looked at the four types of questions you can expect to reduce exam questions down to

SUMMARY

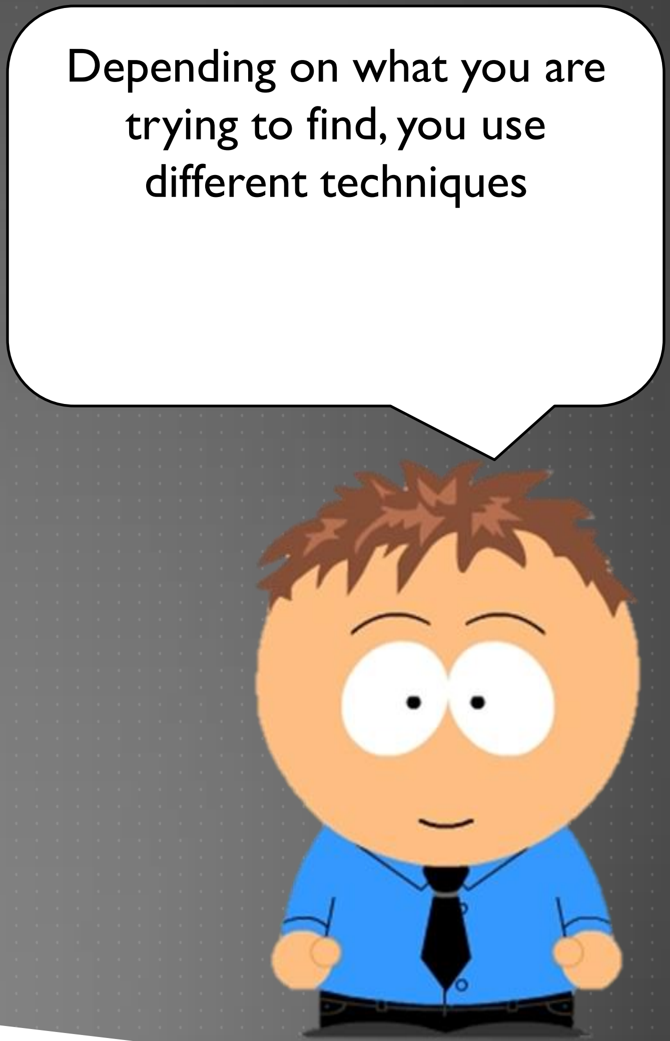


SUMMARY



SUMMARY

Mean	Standard Deviation	Area / Probability	Value of x	Use
			✗	
		✗		
	✗			
✗				

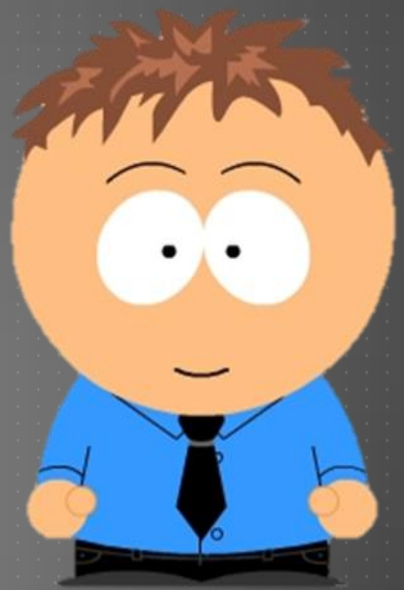


SUMMARY

Mean	Standard Deviation	Area / Probability	Value of x	Use
			X	InvNorm
		X		
	X			
X				

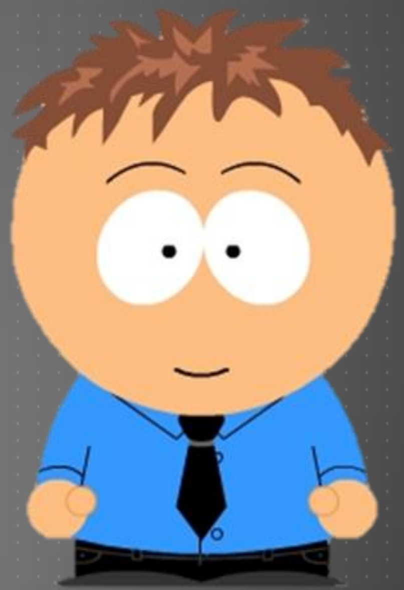
SUMMARY

Mean	Standard Deviation	Area / Probability	Value of x	Use
			X	InvNorm
		X		NormalCDF
	X			
X				



SUMMARY

Mean	Standard Deviation	Area / Probability	Value of x	Use
			✖	InvNorm
		✖		NormalCDF
	✖			Standard Normal + InvNorm
✖				



SUMMARY

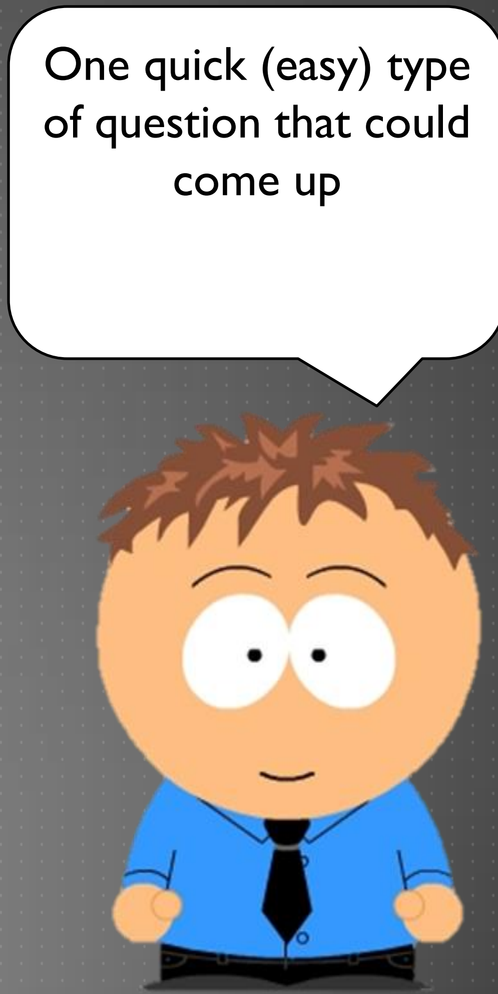
Mean	Standard Deviation	Area / Probability	Value of x	Use
			✗	InvNorm
		✗		NormalCDF
	✗			Standard Normal + InvNorm
✗				

BONUS QUESTION 1
HOW MANY STANDARD DEVIATIONS?

A battery life follows a normal distribution with mean 16 hours and standard deviation 5 hours.

A particular battery has life of 10.2 hours

How many standard deviations below the mean is this?



BONUS QUESTION 1
HOW MANY STANDARD DEVIATIONS?

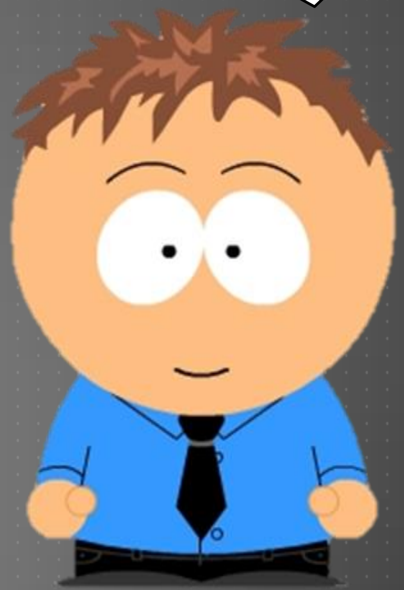
A battery life follows a normal distribution with mean 16 hours and standard deviation 5 hours.

A particular battery has life of 10.2 hours

How many standard deviations below the mean is this?

This is what 'z' scores actually calculate.

z is the number of standard deviations from the mean



BONUS QUESTION 1
HOW MANY STANDARD DEVIATIONS?

A battery life follows a normal distribution with mean 16 hours and standard deviation 5 hours.

A particular battery has life of 10.2 hours

How many standard deviations below the mean is this?

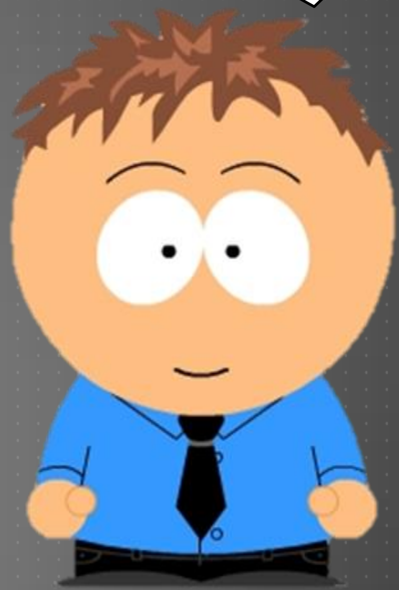
$$z = \frac{x - \mu}{\sigma} \qquad z = \frac{10.2 - 16}{5}$$

$$z = -1.16$$

$$\mathbf{1.16\ standard\ deviations}$$

This is what 'z' scores actually calculate.

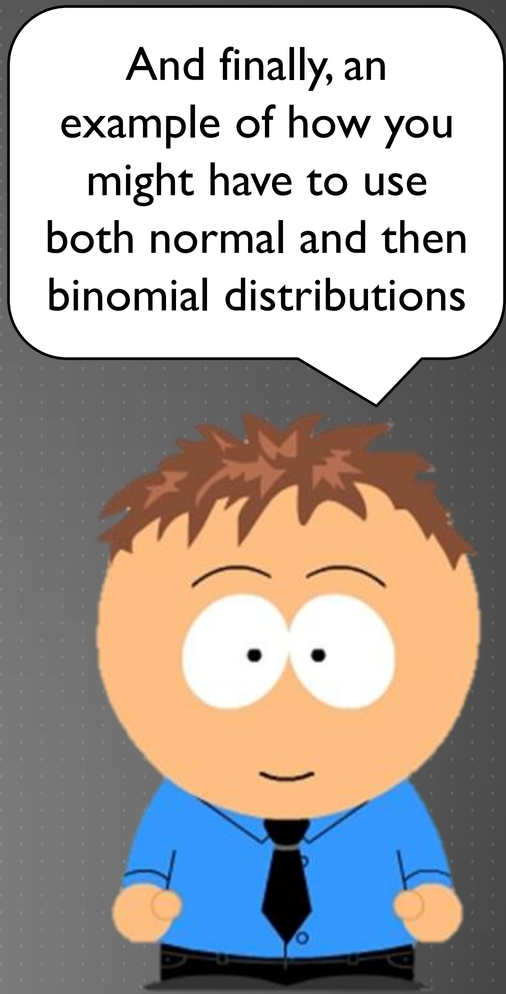
z is the number of standard deviations from the mean



BONUS QUESTION 2
USING THE BINOMIAL DISTRIBUTION ALSO

A battery life follows a normal distribution with mean 16 hours and standard deviation 5 hours.

A sample of 20 laptops are taken. What is the probability that 5 have a battery life between 11 and 13 hours?

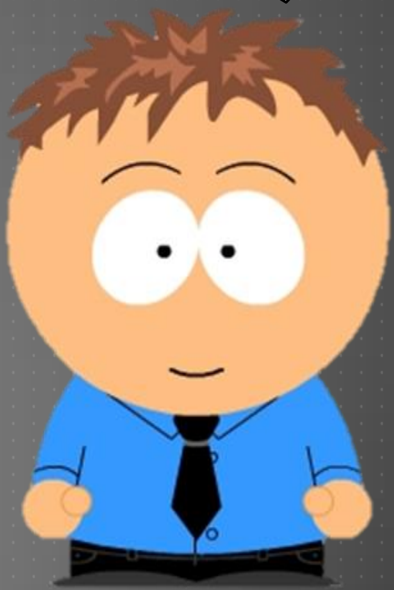


BONUS QUESTION 2
USING THE BINOMIAL DISTRIBUTION ALSO

A battery life follows a normal distribution with mean 16 hours and standard deviation 5 hours.

A sample of 20 laptops are taken. What is the probability that 5 have a battery life between 11 and 13 hours?

Work out the probability that a random laptop has battery life between 11 and 13 hours



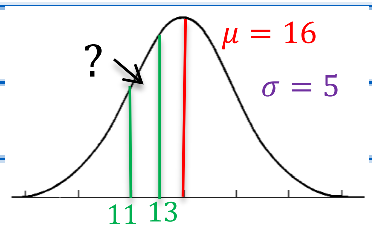
BONUS QUESTION 2
USING THE BINOMIAL DISTRIBUTION ALSO

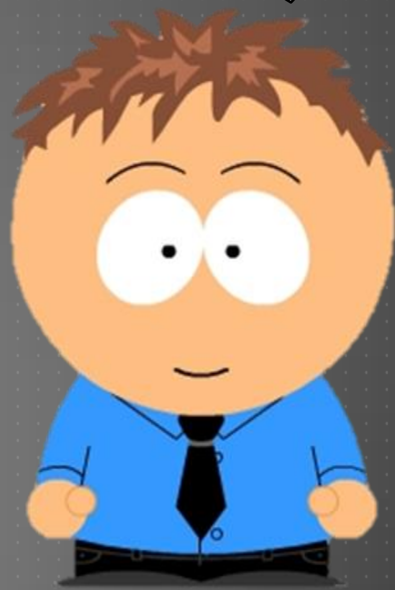
A battery life follows a normal distribution with mean 16 hours and standard deviation 5 hours.

A sample of 20 laptops are taken. What is the probability that 5 have a battery life between 11 and 13 hours?

Work out the probability that a random laptop has battery life between 11 and 13 hours

$X \sim N(16, 5^2)$

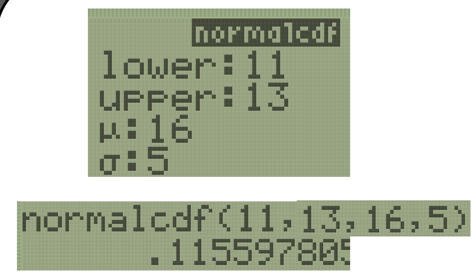




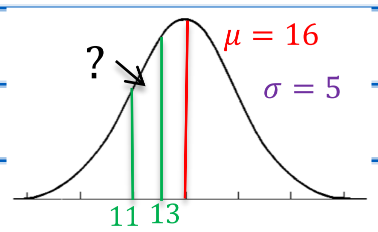
BONUS QUESTION 2
USING THE BINOMIAL DISTRIBUTION ALSO

A battery life follows a normal distribution with mean 16 hours and standard deviation 5 hours.

A sample of 20 laptops are taken. What is the probability that 5 have a battery life between 11 and 13 hours?



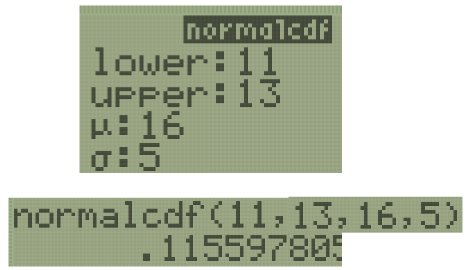
$X \sim N(16, 5^2)$



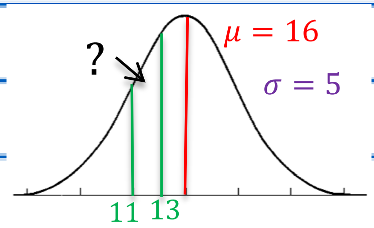
BONUS QUESTION 2
USING THE BINOMIAL DISTRIBUTION ALSO

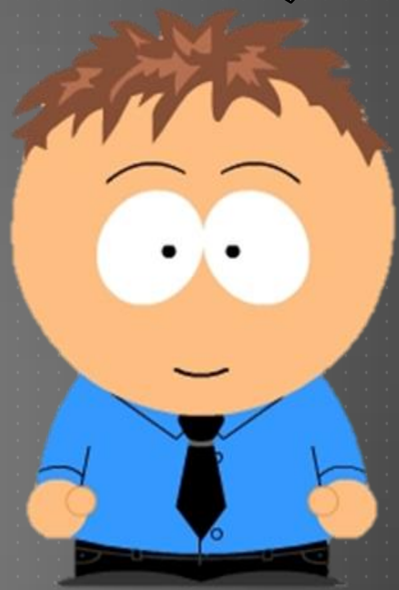
A battery life follows a normal distribution with mean 16 hours and standard deviation 5 hours.

A sample of 20 laptops are taken. What is the probability that 5 have a battery life between 11 and 13 hours?



$X \sim N(16, 5^2)$ $\qquad$ $P(11 \le x \le 13) = 0.1156$





BONUS QUESTION 2
USING THE BINOMIAL DISTRIBUTION ALSO

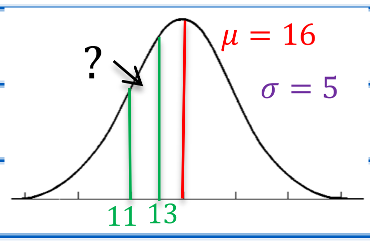
A battery life follows a normal distribution with mean 16 hours and standard deviation 5 hours.

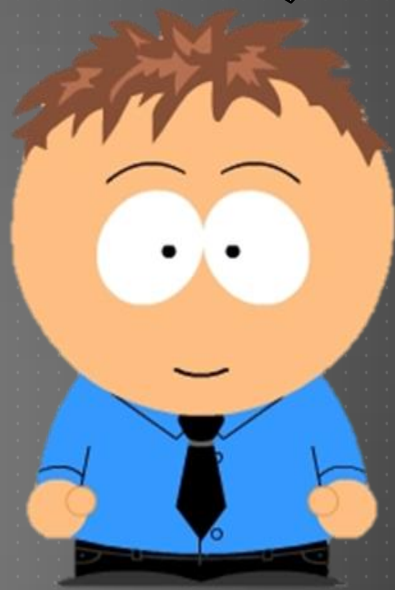
A sample of 20 laptops are taken. What is the probability that 5 have a battery life between 11 and 13 hours?

Now you can use the probability for a binomial distribution

$$X \sim B(n, p)$$

$X \sim N(16, 5^2)$ $\qquad$ $P(11 \leq x \leq 13) = 0.1156$



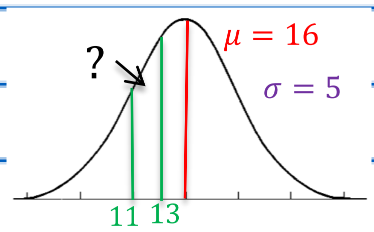


BONUS QUESTION 2
USING THE BINOMIAL DISTRIBUTION ALSO

A battery life follows a normal distribution with mean 16 hours and standard deviation 5 hours.

A sample of 20 laptops are taken. What is the probability that 5 have a battery life between 11 and 13 hours?

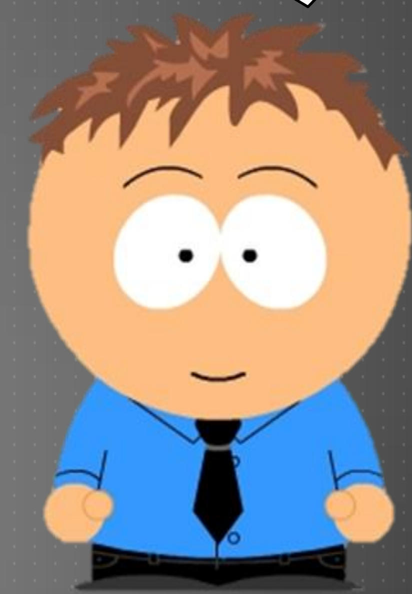
$X \sim N(16, 5^2)$ $\quad P(11 \leq x \leq 13) = 0.1156$



$Y \sim B(20, 0.1156)$

Now you can use the probability for a binomial distribution

$X \sim B(n, p)$

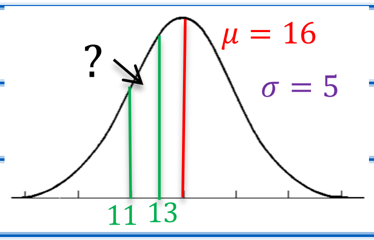


<u>BONUS QUESTION 2</u>
USING THE BINOMIAL DISTRIBUTION ALSO

A battery life follows a normal distribution with mean 16 hours and standard deviation 5 hours.

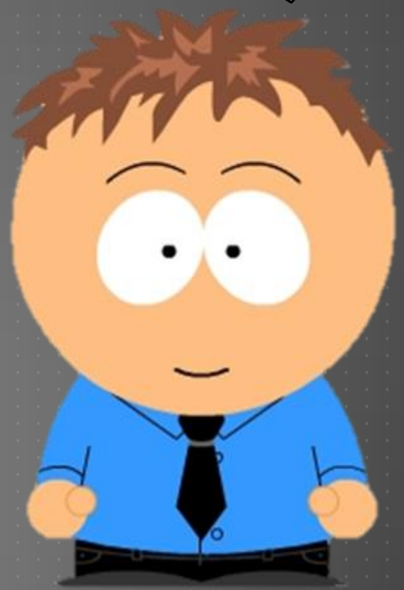
A sample of 20 laptops are taken. What is the probability that 5 have a battery life between 11 and 13 hours?

$X \sim N(16, 5^2)$ $\quad P(11 \leq x \leq 13) = 0.1156$



$Y \sim B(20, 0.1156)$

To find the probability **exactly** 5 satisfy the condition, use binomPDF

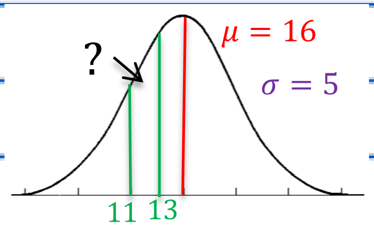


BONUS QUESTION 2
USING THE BINOMIAL DISTRIBUTION ALSO

A battery life follows a normal distribution with mean 16 hours and standard deviation 5 hours.

A sample of 20 laptops are taken. What is the probability that 5 have a battery life between 11 and 13 hours?

$X \sim N(16, 5^2)$ $P(11 \leq x \leq 13) = 0.1156$



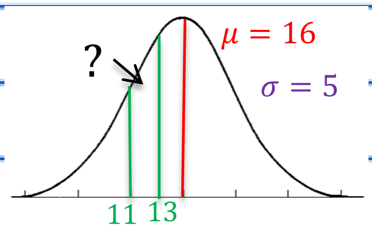
$Y \sim B(20, 0.1156)$

BONUS QUESTION 2
USING THE BINOMIAL DISTRIBUTION ALSO

A battery life follows a normal distribution with mean 16 hours and standard deviation 5 hours.

A sample of 20 laptops are taken. What is the probability that 5 have a battery life between 11 and 13 hours?

$X \sim N(16, 5^2)$ $\quad$ $P(11 \leq x \leq 13) = 0.1156$



$Y \sim B(20, 0.1156)$

$P(Y = 5) = \mathbf{0.0507}$