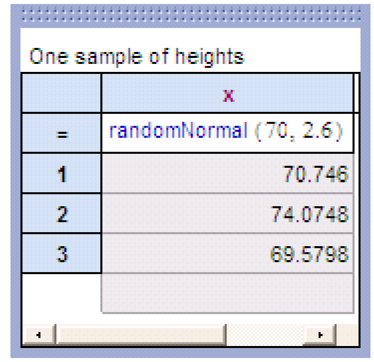
Teaching notes for a Fathom investigation into the sampling distribution of the standardized t statistic for small sample means – adapted from Floyd Bullard's and Julie Grave's original idea for the TI-83

We will simulate the height of an adult male, assuming the population to be Normal with mean 70 inches and standard deviation 2.6 inches.

Simulate the height of one man

- Make a new file
- Make new collection box (Object/New/Collection, or drag one down from the object shelf)
- With that collection box highlighted, drag down a Case Table from the object shelf
- Give the attribute the name x (Click on <New>, type x, press Enter)
- Enter the formula to create the height (Table/Show Formulas; double click the grey formula space; enter randomNormal(70,2.6))
- Collection/New Cases/1



One sample of heights

	x
=	randomNormal (70, 2.6)
1	70.746
2	74.0748
3	69.5798

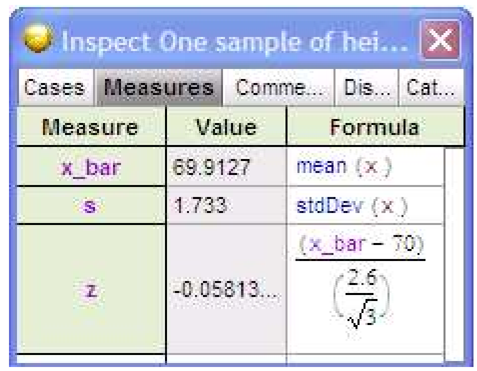
Add 2 more heights to create a small sample of size n=3

- Collection/New Cases/2
- Double click on the name "Collection 1" and change it to "One sample of heights"

Discuss the values you and the class get. Do they look like possible heights of men?

Define the sample statistics $\bar{x}$, s and z (i.e. $z_{\bar{x}}$)

- Double click on the collection box
- Click on the measures tab
- Click <new> and enter x_bar; double click in the formula box and enter mean(x)
- Click <new> and enter s; double click in the formula box and enter stdDev(x)
- Click <new> and enter z; double click in the formula box and enter (x_bar-70)/(2.6/sqrt(3))



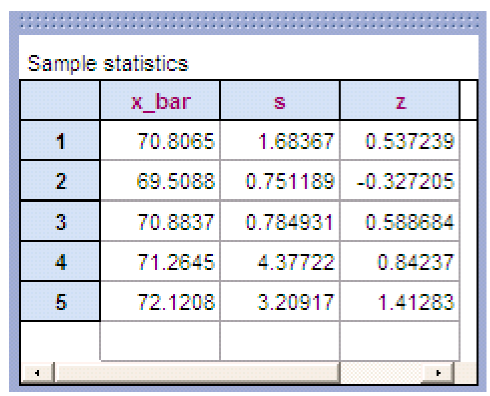
Inspect One sample of hei...

Cases | Measures | Comme... | Dis... | Cat...

Measure	Value	Formula
x_bar	69.9127	mean (x)
s	1.733	stdDev (x)
z	-0.05813...	$\frac{(x\_bar - 70)}{\left(\frac{2.6}{\sqrt{3}}\right)}$

Collect some sample statistics

- Close the Inspection box
- Right click on the collection box and select Collect Measures
- Rename this new collection "Sample statistics"
- With the "Sample statistics" collection box selected, drag down a Case Table from the object shelf. It will display the five sets of statistics ($\bar{x}$, s and z) from five different random samples



Sample statistics

	x_bar	s	z
1	70.8065	1.68367	0.537239
2	69.5088	0.751189	-0.327205
3	70.8837	0.784931	0.588684
4	71.2645	4.37722	0.84237
5	72.1208	3.20917	1.41283

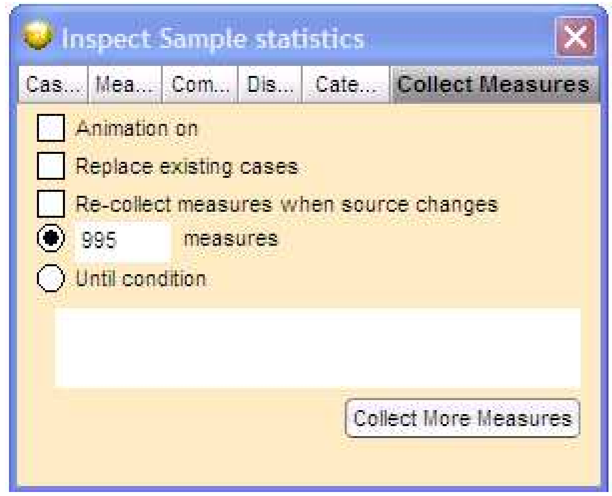
Discuss with the class.

Does everyone understand $z = \frac{\bar{x} - 70}{2.6/\sqrt{3}}$? Go through the formula and review this. Does the class realize that the z values have a standard normal distribution with mean 0 and standard deviation 1? Does anyone have any z values outside the range -3 to 3? That should be quite rare. Write the formula for z on the board and discuss.

(You might also choose to discuss the values of s that you are getting. We know the mean of s should be 2.6. But the distribution of values of s must be skewed right, since the lowest value you could get for s is zero, and on rare occasions s could be much larger than 2.6.)

Collect some more sample statistics

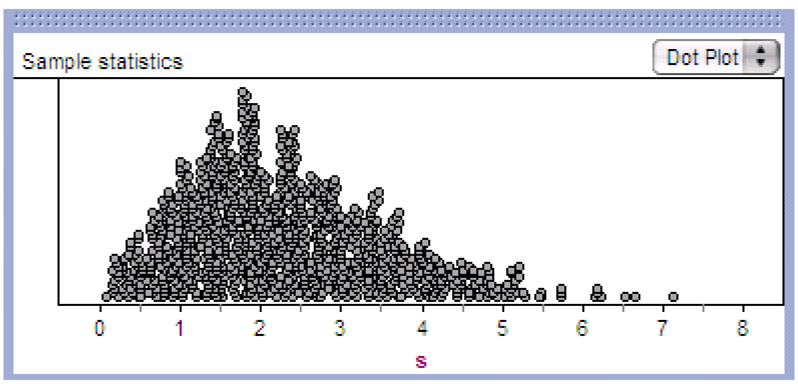
- Double click on the “Sample statistics” collection box
- Click the Collect Measures tab
- Turn off animation
- Collect many more sample statistics (1000 in total works well)

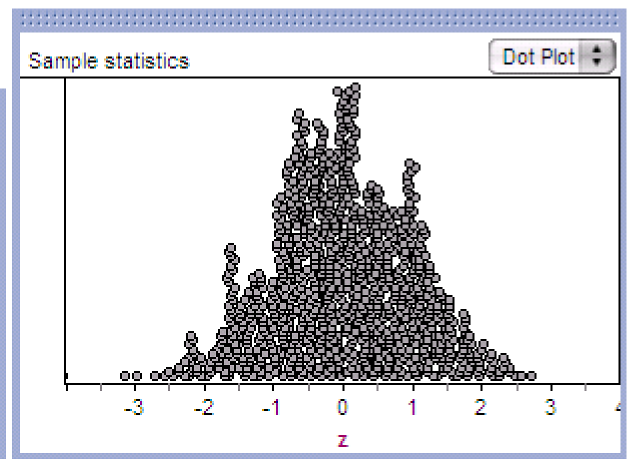


Let’s make a graph of these statistics; this will give us an idea of what their sampling distributions would look like.

- Drag a graph down from the object shelf; grab s and drop it onto the horizontal axis
- Drag another graph down, and grab z and drop it onto the horizontal axis
- Double click on the z graph; that will open the inspection window. Change xLower and xUpper to -4 and 4; and change xAutoRescale to false. (That will keep the window from -4 to 4 when we later create the t distribution graph, so that we can compare the two graphs more readily.)

Discuss these graphs. The graph of s will be skewed right, and the graph of z will be a standard Normal model with mean 0 and standard deviation 1. If you want these graphs to be more smooth, then you should collect more than the 1000 samples I used.





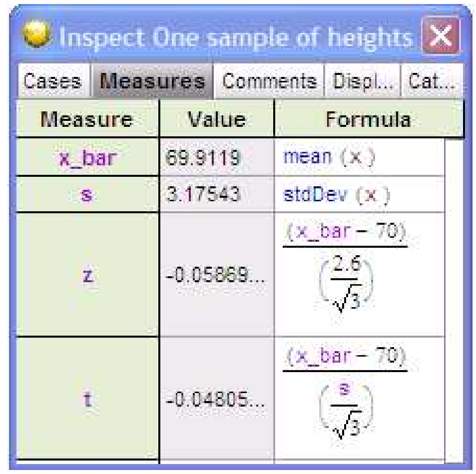
Further discussion: When we are working in the real world, we will usually only have one sample, and we will not know what the population standard deviation is. In this case, it is 2.6 inches, but we wouldn’t know that in most situations.

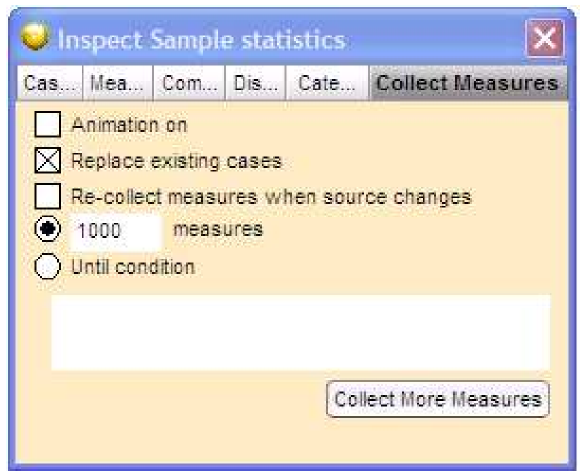
We would have to estimate the population standard deviation by using our sample standard deviation, s.

So, we would actually calculate a different standardized statistic using the formula $t = \dfrac{\bar{x} - 70}{s/\sqrt{3}}$.

Let's put that into Fathom.

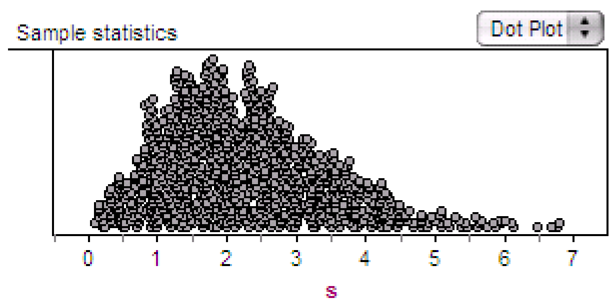
- Double click on the "One sample of heights" collection box.
- Click the Measures tab
- Create another measure called t, and use the formula (x_bar-70)/(stdDev(x)/sqrt(3))



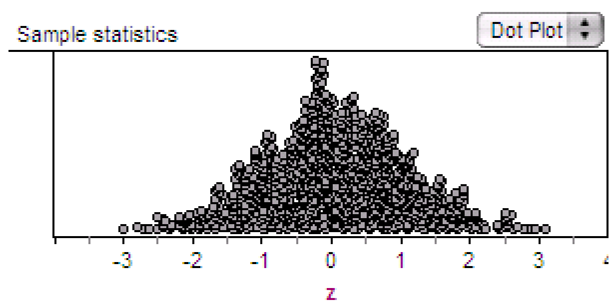


- Double click on the "Sample statistics" collection box
- Click the Collect Measures tab
- Turn on Replace existing cases, and collect 1000 sample statistics
- Right click on the z graph, and choose Duplicate graph; drop the attribute t onto the horizontal axis

Discussion with class: We now have graphs of z and t with exactly the same scale (-4 to 4). The z graph will be pretty much confined to the range -3 to 3, but the t graph will extend beyond -4 and 4. They will both be "bell-shaped" and roughly symmetric. Remember that the true sampling distribution for each is the graph showing all possible sample statistics, so those would be continuous smooth bell-shaped curves.

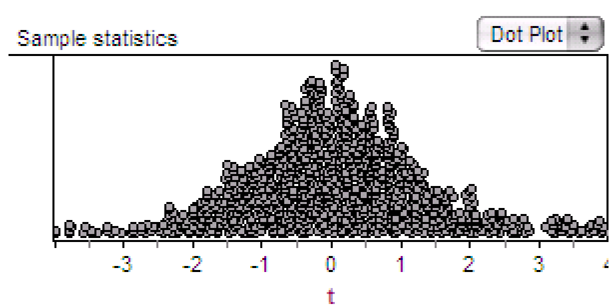


Why is the t distribution wider than the standard normal (z) distribution? The additional random variable s was used to create it. That additional variability results in a standardized variable (t) with a fatter distribution; the tails of the graph are thicker, and it is not uncommon with a small sample of n=3 to get a t value of magnitude 4 or 5 or even higher.



What would cause t to occasionally be very large?

- If our sample of three men has $\bar{x}$ far from 70, then both z and t will be large.
- If our sample of three men has small s, then t will be large.
- If both these things happen, then t will be very large indeed!



If you scroll down the sample statistics case table, you will find some very large values of t.