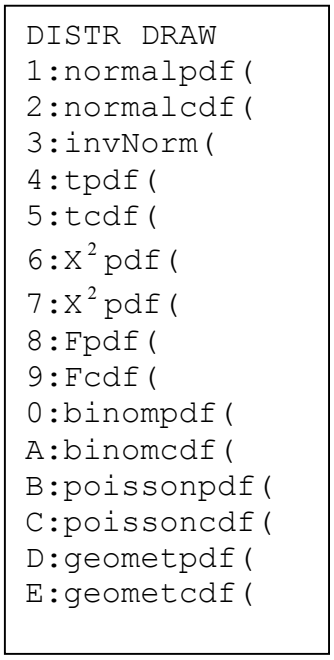
Chapter 7
Random Variables and Probability Distributions

7.0 Introduction

The TI-83's built-in capability to handle the common probability distributions in elementary statistics is an incredible time-saver. It doesn't take very long to flip to the back of the book and look at the tables, but sometimes it can be a pain when a desk is small, or cluttered, and there's a lot of flipping. The TI-83 will easily calculate values for the commonly performed table look-ups and does so through the use of algorithms programmed into the calculator. The functions that perform in the place of table look-ups can be seen by pressing the following:

2^{nd} > Distr

These functions appear for your inspection as:



There are a couple things we need to make note of here. First of all, notice that suspiciously unmatched left parenthesis? It certainly offends our mathematical sense of decency but as we have previously mentioned the unmatched left parenthesis is a signal to you, the calculator user, that some further information is required in order to actually evaluate the function you are choosing. Your tasks when using one of these functions are (a) to understand what values are being asked for, and in what order, and then to supply the right parenthesis. (In some cases the calculator will forgive you if you don't supply

that right parenthesis, but we know that you, being a mathematical purist, will supply that right parenthesis.) The second thing to notice is that almost all these functions are "pdf" functions and "cdf" functions. They look very similar when you are choosing them in the menu, and it is possible you might choose the incorrect one if you are not sufficiently attentive.

Before we actually perform some calculations with the standard normal distribution on the TI-83 there is an unfortunate misnomer that should be pointed out. To the mathematical statisticians among you, "pdf" and "cdf" have specific connotations – "probability density function" and "cumulative distribution function," or just "distribution function." The TI-83 is not quite completely on board with this notation. For continuous probability functions, the TI-83 "cdf" functions calculate the probability that the value of the variable is *between* two numbers.

7.1 Probabilities associated with normal distributions

Example 7.25: Children's heights.

Problem #1: $P(a < x < b)$

In poor countries the growth of children can be an important indicator of general levels of nutrition and health. Data suggests that a reasonable model for the probability distribution of the continuous numerical variable x = height of a randomly selected five-year-old child is a normal distribution with a mean of μ = 100 cm and standard deviation σ = 6 cm. What proportion of the heights is between 94 cm and 112 cm?

As you may recall, in order to answer this question, we must find

$$P(94 < x < 112)$$

Because the normal curve we are investigating is not the standard normal curve – and thus does not match the Normal Curve Tables – we have to do a little preliminary work. Translating to an equivalent problem for the standard normal distribution,

$$a^* = \frac{a-\mu}{\sigma} = \frac{94-100}{6} = -1.00$$
$$b^* = \frac{b-\mu}{\sigma} = \frac{112-100}{6} = 2.00$$

Then

$$P(94 < x < 112) = P(-1.00 < z < 2.00)$$
$$= (z \text{ curve area to the left of } 2.00) - (z \text{ curve area to the left of } -1.00)$$
$$= .9772 - .1587$$
$$= .8185$$

The TI-83 can take the place of this table lookup and save a great deal of time. The function we will use to find the area between 0.9772 and 0.1587 for the standard normal distribution is the "normalcdf" function. Here's the sequence:

2nd > DISTR > `normalcdf(`

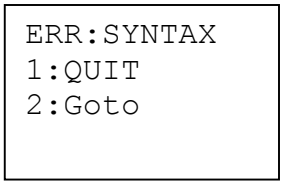
Now we are a bit worried. We know -- because of that left parenthesis – that we are supposed to supply some information, and that the order we enter it is important. A brief referral to the TI-83 manual tells us the order:

normalcdf (lower bound, upper bound [, μ, σ])

On the TI-83 you must supply the lower bound and upper bound with the appropriate values for *z*, and the mean and standard deviation. The square brackets, [ ] tell us the mean and standard deviation are options. In the case of the standard normal distribution, we don't need to fill in those options – we will return to them later. At this point, continue the keystroke entry we began earlier:

2nd > DISTR > `normalcdf(-1,2)` > ENTER

and the calculator should respond with lighting speed: .8185946784. It is also possible it responded with lightning speed in a stern and stentorian manner:



If you see this message, you have most likely pressed the "subtraction" key next to the 6 on the TI-83, rather than the "(–)" key under the 3 on the –83. This second option is the one you want.

If you are reasonably comfortable with the calculator and its workings, and don't mind working with those added parameters, you can actually save even more time on this problem. Remember, our model for the heights of these prehistoric 5-year olds is a

normal distribution with a mean of μ = 100 cm and standard deviation σ = 6 cm. Try out this sequence:

2nd > DISTR > `normalcdf(94,112,100,6)` > ENTER

Did you get the same answer, $P(94 < x < 112) = .8185946784$?

Do you think that is just an accident?

No, probably not – it is the TI performing the calculations with the original values, without your needing to transform them to *z* values. Now there is a time saver!

Problem #2: $P(a < x)$

The next calculation in this exercise: how do we find the probability that a randomly chosen child will be taller than 110 cm? Here is where that problem with the "cdf" mentioned above kicks in. We must recast this problem so that it fits the syntax of the cdf function of the TI. This is not difficult, but it's a bit of a trick. We will conceptually translate

$P(a < x)$ into $P(a < x < b)$, and think of *b* as a "very large" number.

Just how large a very large number is will depend on the particular normal distribution, but our knowledge of the normal distribution suggests that 10 standard deviations above the mean should be "large enough" so that we haven't missed too much of the probability in the right tail. So let's try this:

2nd > DISTR > `normalcdf(110,100+10*6,100,6)` > ENTER

We are telling the calculator to find the probability of getting a value…

1. `110,100+10*6` → "between 110 and 6 standard deviations above the mean,
2. `100,6` → "in a distribution with a mean of 100 and standard deviation of 6.

And the TI delivers once again! $P(110 < x) = 0.0477903304$. Of course, our mathematical sensibility still may be offended by all these decimals, so we suggest rounding to something reasonable. Say, $P(110 < x) = 0.0478$.

7.2 Conclusion

As can be seen, these probability distribution calculations can be terrific time savers. Just be careful to choose the right function, "pdf" or "cdf", and if you don't want to standardize the distribution to the standard normal, be sure to use those optional parameters carefully.