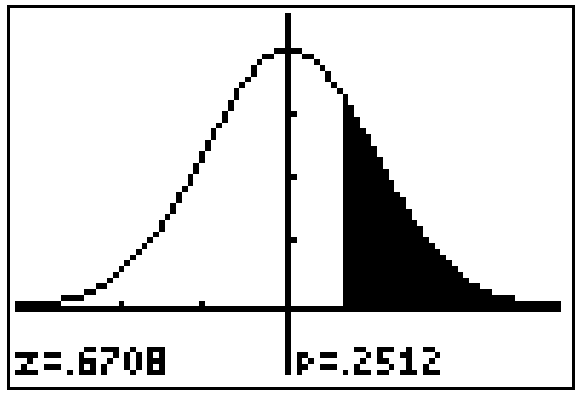
CHAPTER 3



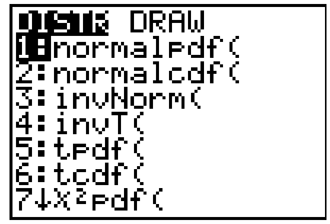
The Normal Distributions

Introduction

In this chapter, we use the TI calculators for calculations involving normal distributions – finding relative frequencies and finding quantiles.

3.1 Density Curves and Normal Distributions

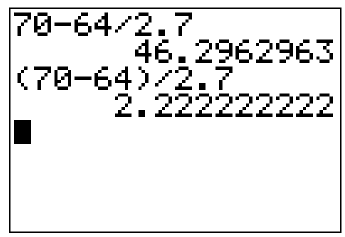
TI calculators have several commands in the `DISTR` menu that can be used for graphing normal distributions, computing normal probabilities, and making inverse normal calculations. In this section, we demonstrate these various functions.



Example 3.4 Standardizing women's heights. The heights of women aged 20 to 29 are approximately Normal with $\mu = 64$ inches and $\sigma = 2.7$ inches. A woman 70 inches tall has what standardized height? The standardized height (or *z*-score) is computed as

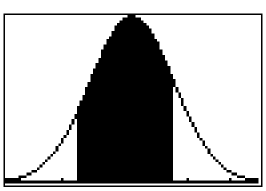
$$z = \frac{height - 64}{2.7}.$$

When using a calculator to perform this calculation, either do the subtraction first, or be sure to use parentheses. Otherwise, errors will result in nonsensical *z*-scores, such as seen at right – there is no way a 70 inch tall woman is more than 46 standard deviations (46*2.7 = 124.2 inches) taller than average! The correct z-score is 2.22 as shown in the bottom computation.

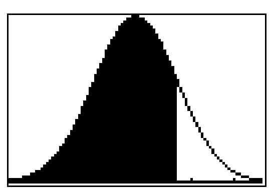


The Normal Distribution and Inverse Normal Commands

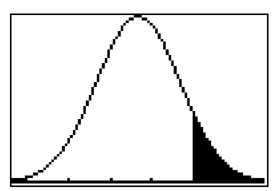
For any $N(\mu,\sigma)$ distribution *X*, we can directly find probabilities with the built-in `normalcdf(` command from the `DISTR` menu. On a TI-89, the command is option 4 on the `Distr` menu. Fill in the boxes as prompted; they will look just like the ones on the `Shade Norm` screen. The command on a TI-83/84 is used as follows. The calculator understands 1E99 as infinity (∞), but practically speaking, any very large number will work.


P(a < *X* < b)

`normalcdf`(a,b,μ,σ)


P(X < k) = P(X ≤ k)

`normalcdf`(–1E 99,k,μ,σ))


P(X > k) = P(X ≥ k)

`normalcdf`(k,1E 99,μ ,σ))

Example 3.5 Who qualifies for college sports? The National Collegiate Athletic Association (NCAA) requires Division II athletes to score at least 820 on the combined mathematics and reading parts of the SAT in order to compete in their first college year. The scores of 1.5 million high school seniors taking the SAT this year are approximately Normal with mean 1026 and standard deviation 209. What percent of high school seniors qualify for Division II college sports?

TI-83/84 Solution. We can simply find the answer using `2:normalcdf(` from the [2nd][VARS] = `DISTR` menu, or use the `DRAW` option on the `DISTR` menu to shade the relevant part of the distribution and give the proportion.

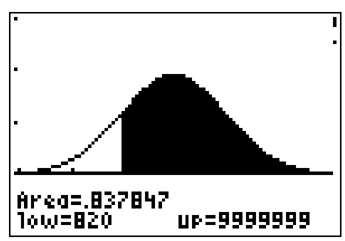
First, simply finding the proportion, the parameters for the `normalcdf` command are the low end of interest, the high end of interest, the mean and the standard deviation. For this example, since we want to find the percentage of high school seniors that score at least 820, the low end is 820. Normal curves extend to infinity (∞) on either end, so the high end is technically $+\infty$. We can for practical purposes use a very large, positive number (a bunch of 9's). The mean was given as 1026 and the standard deviation as 209. We see that about 83.8% of high school seniors would qualify to be Division II athletes.

```
normalcdf(820,99
999999,1026,209)
        .8378465552
```

To graph the Normal distribution and shade the appropriate area, first make certain that all `STAT PLOTS` are turned off. Press [2nd][Y=] and then [4] to turn all plots off. We will need to size the `WINDOW` for this plot. Press [WINDOW]. From the 68-95-99.7 Rule, we know that almost all the area under a Normal curve is within 3 standard deviations of the mean. Three standard deviations here is 3*209 = 627. We add and subtract that from the mean of 1026, so have an `Xmin` of 399 and `Xmax` of 1653. We have set `Xscl` to 209, so we will see a tick mark for each standard deviation above and below the mean. No probability distribution can be negative (we can't see negative frequencies of something!), but we don't want the calculation answers to obscure the graph so we set `Ymin` to -.001. Setting `Ymax` for a plot like this is a little harder. The whole area under the curve is 1, so we know `Ymax` should be less than 1. Play around some to find a good value. Here, I've set `Ymax` to .003. Return to the Home screen and press [2nd][VARS] for the DISTR menu. Press the right arrow key to `DRAW`. Press [ENTER] to select option `1:ShadeNorm`. Enter the parameters just as you did for the `normalcdf` command.

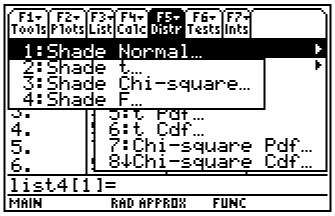
```
WINDOW
 Xmin=399
 Xmax=1653
 Xscl=209
 Ymin=-.001
 Ymax=.003
 Yscl=.001
 Xres=1
```

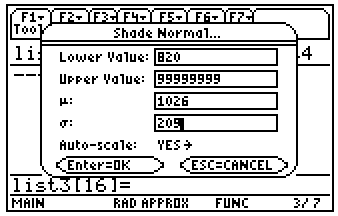
```
ShadeNorm(820,99
99999,1026,209)
```

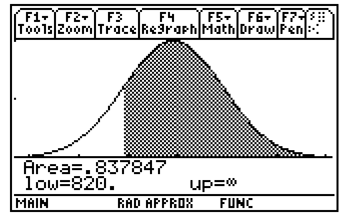


TI-89 Solution. In the Statistics/List Editor, press [F5] (`Distr`). The first option is `Shade`. Press the right arrow to see more options under this. Press [ENTER] to select `Shade Normal`. Enter the lower and upper values of interest then the values of μ

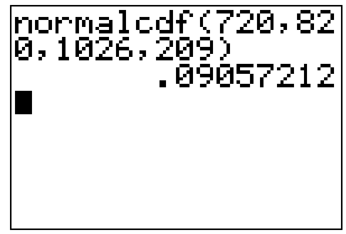
and σ. Select YES to Auto-scale the plot by pressing the right arrow and moving the cursor to highlight YES; this saves having to figure out the window on your own. Note that option 4 on the Distr menu is Normal cdf which asks for the parameters just as Shade Normal does.







Example 3.8 Who qualifies for an athletic scholarship? The NCAA considers a student a "partial qualifier" for Division II athletics if high school grades are satisfactory and the combined SAT score is at least 720. What proportion of all students would be partial qualifiers?



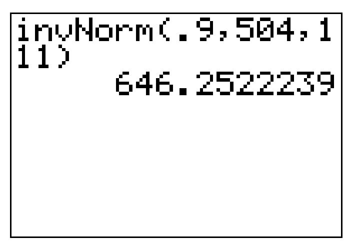
The "partial qualifiers" would have SAT scores between 720 and 820. We see that this is about 9% of all students.

Inverse Normal Calculations

To find the value x for which $P(X \le x)$ equals a desired proportion p (an inverse normal calculation), we use the command invNorm(p, μ, σ). This is option 3 on the TI-83/84 DISTR menu. Note that the proportion p is always to the *left* of the desired value. On a TI-89, press the right arrow to expand the Inverse menu where this is option 1. The following examples demonstrate these commands.

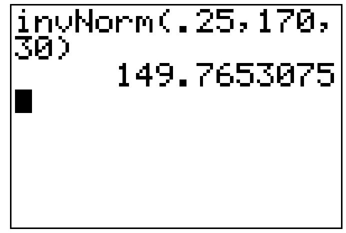
Example 3.9 Find the top 10%. Scores on the SAT reading test in recent years follow approximately the N(504,111) distribution. How high must a student score in order to place in the top 10% of all students taking the SAT?

Solution: If a score marks the beginning of the top 10%, 90% of students score that or less, so p = 90%, or 0.9. We've entered that and the mean and standard deviation as the parameters. An SAT reading score of at least 646.25 (or 650, since scores are reported in multiples of 10) will place a student in the top 10%.



Example 3.11 Find the first quartile. High levels of cholesterol in the blood increase the risk of heart disease. For 14-year-old boys, the distribution of blood cholesterol is approximately Normal with mean $\mu = 170$ milligrams of cholesterol per deciliter of blood

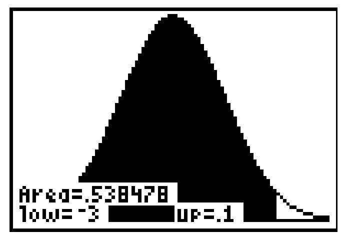
(mg/dl) and standard deviation $\sigma = 80$ mg/dl. What is the first quartile of the distribution of blood cholesterol?



Solution: The first quartile has 25% of the distribution below it. We enter that and the parameters of the distribution to find that the first quartile of blood cholesterol measurements in 14-year-old boys is about 149.8 mg/dl.

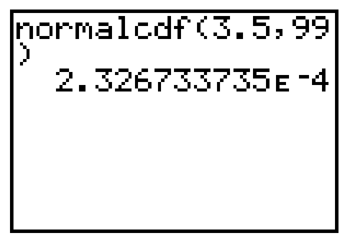
3.2 Common Errors

Why is my curve all black?



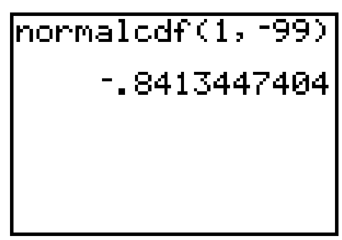
For the standard normal curve, the graph indicates well more than half of the area is of interest between -3 and .1 the message at the bottom says the area is 53.8%. This is a result of having failed to clear the drawing between commands. Press [2nd][PRGM] then [ENTER] to clear the drawing, then reexecute the command.

Probability more than 1



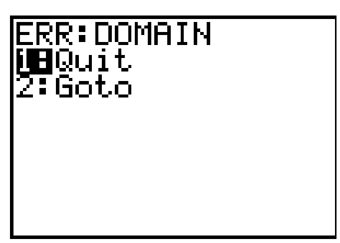
It can't. If the results look like the probability is more than one, check the right side of the result for an exponent. Here it is –4. That means the leading 2 is really in the fourth decimal place, so the probability is 0.0002. The chance a variable is more than 3.5 standard deviations above the mean is about 0.02%.

Negative probability



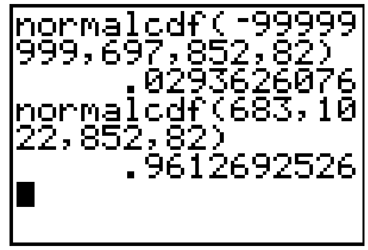
It can't. The low and high ends of the area of interest have been entered in the wrong order. As the calculator does a numerical integration to find the answer, it doesn't care. You should.

Err: Domain

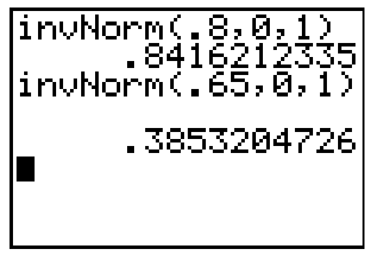


This message comes as a result of having entered the `invNorm` command with parameter 90. (You wanted to find the value that puts you into the top 10% of women's heights, so 90% of the area is to the left of the desired value.) The percentage must be entered as a decimal number. Re-enter the command with parameter .90.

3.3 Selected Exercise Solutions



3.11 To find the percent of years with less than 697 mm of rain, we use `normalcdf` from the `DISTR` ([2nd][VARS]) menu. The parameters for that command are the low value of interest (here, technically $-\infty$, or a very large negative number), the high value of interest (697), the mean (852), and the standard deviation (82). We see that about 2.94% of years will have less than 697 mm of rain. To answer the second part of the question, we use the low value of interest (683) and the high value of interest (1022) and find that about 96.13% or all years will have "normal" rainfall.



3.29 Using software, we can also find these values, using `invNorm` from the `DISTR` ([2nd][VARS]) menu. The parameters for that command are the area to the left of the point of interest, the mean (0 for standard Normal) and the standard deviation (1 for standard Normal). We find that $z = 0.84$ has area 0.8 below it and $z = 0.385$ has area 0.35 above it (0.65 below it).

Continue your practice with these exercises:

3.9 Men's and women's heights.
3.13 Table A.
3.31 Acid rain?
3.33 A milling machine.
3.35 In my Chevrolet.
3.37 The middle half.
3.39 What's your percentile?
3.41 Heights of men and women.
3.43 A surprising calculation.
3.47 Normal is only approximate: ACT scores.
3.49 Are the data normal? Fruit fly thorax lengths.
3.51 Are the data normal? Soil penetrability.
3.53 Where are the quartiles?