Statistics with the TI-84 Calculator

Version 2.10 - 2004-01-09 - corrections & additions welcome - Dr. Wm J. Larson - william.larson@ecolint.ch

Table of Contents

Creating a list

Key STAT EDIT Edit and type your list in L1 or L2 etc.

Calculating the Mean, Median, Standard Deviation & Interquartile Range

First enter your data into a list as above.

Key STAT, CALC, 1: 1-Var Stats, L1, Enter. Since L1 is the default for 1-Var Stats, if you entered your data into L1, you need not type L1 again. i.e. keying STAT, CALC, 1: 1-Var Stats, Enter would work.

A set of statistics about L1 will appear. The mean, $\bar{x}$, is at the top of the list. Scrolling down other statistics including Med (the median), Sx (sample standard deviation), σx (the population standard deviation), Q1 (the lower quartile) & Q3 (the upper quartile) will be displayed. The interquartile range = Q3–Q1.

The Names of TI-84 Symbols & Alternative Symbols

TI-84 Symbol	Alternative Symbols	Name	Alternative Name
$\bar{x}$		Mean	Average
Sx	s, s_{n-1}	Sample standard deviation	unbiased estimator of the population standard deviation (IB name)
σx	σ, s_n	Population standard deviation	Sample standard deviation (IB name)
minX	L	Minimum value	the lowest value
Q1		first quartile	
Med	M	Median	
Q3		third quartile	
maxX	H	Maximum value	the highest value

Range

Range = MaxX - MinX

Interquartile Range

Interquartile range = Q3 – Q1.

Using a frequency list

If you are given data points with frequencies for each data point, put the data points in L1 & the frequencies in L2. Then key STAT, CALC, 1: 1-Var Stats, L1, L2.

L1 is the default for the data list, so if there is no frequency list & the data is in L1, you need not type "L1". But there is no default for the frequency list. So if there is a frequency list in L2, you need to type 1-Var Stats L1, L2.

Redisplaying Data

If you cleared the screen (but did not run a new statistics calculation), you can redisplay your data. For example you can redisplay Q1 & Q3 by keying VARS 5:Statistics, PTS & then selecting 7:Q1 or 9:Q3.

stdDev & variance

Be careful stdDev(& variance(which are in LIST MATH and in the CATALOG return Sx (s_{n-1}) and Sx² (s_{n-1}^2) respectively, not σx (s_n) and σx² (s_n^2) as you might suppose.

Be careful to clear the screen

The TI-84 has a tendency to display information from a previous calculation, so when you are making a new calculation, always clear the screen first using CLEAR, CLEAR.

Making Histograms with a TI-84

Enter your data

If your data is just a set of numbers, enter your data into one list, say L1.

If instead your data is a frequency distribution table, enter your data into two lists, say L1 for the values and L2 for the frequencies.

If your data is grouped data, e.g. with class intervals, enter the midpoint of each class interval in L1 and the frequencies in L2.

Set Up Your Plot.

Now key 2nd STAT PLOT and set up your plot.

Choose a plot, say Plot1, by putting the cursor on Plot1 and pressing ENTER.

Turn Plot1 on by putting the cursor on On and pressing ENTER.

Choose to plot a histogram by moving the cursor to the image of a histogram and pressing ENTER.

If you have just a set of numbers in L1, key Xlist: L1 and Freq: 1.

If instead you have the values in L1 and the frequencies in L2, key Xlist: L1 and Freq: L2.

If you want to change Freq from L2 to 1, you must key ALPHA 1.

Set Up Your Window

Key WINDOW.

Set Xmin a little less than your smallest value and Xmax a little more than your biggest value.

Set Xscl to give the size of your class intervals. Xscl can be reset until you are satisfied that your interval size gives a good representation of the data. About 8 to 20 intervals usually give a good representation.

Display Your Histogram

Turn off any other plots and any graphs in Y=.

Now key GRAPH and voila - the histogram!

Display the interval and frequency

To display the interval and frequency use TRACE

Calculating Probabilities for the Normal Distribution

Using ShadeNorm

ShadeNorm will draw the graph and calculate the probability.

Key 2nd DISTR DRAW1: ShadeNorm(lowerbound, upperbound [, μ, σ])

Example

Find P(z < -0.5). (The default vales of μ = 0, σ= 1 are desired, so they need not be entered.)

Key DISTR DRAW 1: ShadeNorm(-100, -.5)

The graph, the lower bound (-100, being 100 standard deviations from the mean, is effectively minus ∞), the upper bound and the P(z<-0.5), i.e. 0.3085 are displayed.

If the graph is not visible, set the Window to:

xmin =	-3
Xmax =	3
Xscl =	1
Ymin =	-.25
Ymax =	.5
Yscl =	.25
Xres =	1

Example

If μ = 55, σ = 10, find P(40 < x < 65).

It's tiresome to reset the window so input this program. It prompts for μ, σ, the lower and upper bounds, sets the widow size and then runs ShadeNorm.

```
PROGRAM: NORMWIND
:Input "MEAN: ",M
:Input "ST DEV: ",S
:-.3/(S*√(2π) ) STO► Ymin
:-3.6*Ymin STO► Ymax
:M-3.5*S STO► Xmin
:M+3.5*S STO► Xmax
:Input "LOWER BND:",L
:Input "UPPER BND:",U
:ShadeNorm(L,U,M,S)
:Stop
```

Using normalcdf

Drawing the cumulative probability distribution graph is very illustrative, but a little time consuming. It's faster just to run normalcdf.

Key 2nd DISTR DISTR (the default)

2: normalcdf(lowerbound, upperbound [, μ, σ])

Example

If μ = 55, σ= 10, find P(x < 65).

Keying DISTR DISTR normalcdf (-E99, 65, 55, 10) will display 0.841.

We usually want the cumulative distribution function (cdf) for the normal distribution. The **p**robability **d**istribution **f**unction (pdf) would be useful to graph the normal curve in Y=, but ShadeNorm already does that.

Significant Digits

Notice that more significant digits are available with the TI-84 than with a normal distribution table in a textbook. However in the real world μ & σ are usually not known with enough accuracy to make this meaningful.

Calculating the Inverse Normal Distribution

Using invNorm

For Φ(a) ≡ P(Z < a) if Φ is known but a is not known, invNorm will calculate a.

Key 2nd DISTR DISTR (the default)

3: invNorm(area, [, μ, σ])

Example

If P(Z < a) = .6, find a. (The default vales of μ = 0, σ= 1 are desired, so they need not be entered.)

Keying DISTR 3: invNorm(.6) will give 0.253347

Example

If x ~ N(100, 5^2) & P(x < a) = .20, find a.

Keying DISTR 3: invNorm(.2, 100, 5) will give 95.8.

Calculating Probabilities for the t-Distribution

Using tcdf

tcdf will calculate the probability, i.e. the **t c**umulative probability **d**istribution **f**unction.

Key 2nd DISTR DISTR (the default)

5: tcdf(lowerbound, upperbound, df)

Example

If df = 20, find P(t < 2).

Keying DISTR DISTR tcdf (-E99, 2, 20) will display 0.9704.

We usually want the cumulative distribution function (cdf) for the normal distribution. The **p**robability **d**istribution **f**unction (pdf) would be useful to graph the normal curve in Y=, but Shade_t already does that.

Inverse t

There is no function corresponding to invNorm for the t-distribution, but you can use the TI-84 equation solver. MATH 0: Solver. Then using tcdf(from 2nd DISTR, key in eqn: 0 = tcdf(-10000, x, 27) - 0.95. Then key ALPHA SOLVE, giving 1.703. We needed a large negative number, so we used -10000. Except for v < 5, -10 would have been large enough.

Or get t from a graph. This method takes longer, but it is more illustrative. Graph 2nd DISTR 5: tcdf(-10000, x, df), where -10000 is the lower bound, df is the degrees of freedom and x is the variable to be graphed. Set Xmin = -4, Xmax = 4, (because the tails beyond t = ± 4 are almost zero, Ymin = -.1, Ymax = 1.1 (because tcdf is the probability that t < x, which must be zero for x = -∞ and one for x = + ∞), Xres = 8 (because the TI-84 calculates t and that's very slow). Now graph the desired probability (e.g. for P(t < t*) = 0.90, graph Y2 = 0.90) and find the intersection of the curves.

Example Find t for n = 10 and p = 0.95. df = n - 1 = 9. Key Y1 = tcdf(-10000, x, 9) Y2 = .95. Use 2nd CALC intersect to find t (x on the screen) = 1.833 at p (y on the screen) = .95, of course.

Calculating Probabilities for the Poisson Distribution (Higher Level only)

Using poissonpdf and poissoncdf

Since the Poisson distribution is discrete, either the cumulative distribution function (cdf) or the probability distribution function (pdf) would be useful. Use the pdf to find the probability that one value is observed (X = Xo) & the cdf to find the probability that one of a range of values is observed (X ≤ Xo).

Key 2nd DISTR DISTR (the default), B: poissonpdf(μ, x) or Key 2nd DISTR DISTR (the default), C: poissoncdf(μ, x)

Example

If μ = 3.75, find P(x = 6).

Keying DISTR DISTR poissonpdf (3.75,6) will display 0.0908.

Example

If μ = 1.4, find P(x ≥ 2) = 1 - P(x ≤ 1)

Keying DISTR DISTR poissoncdf (1.4,1) will display 0.408. P(x ≥ 2) = 1 - 0.408 = 0.592.

Graphing the Poisson distribution

Use poissonpdf to graph the Poisson distribution. For example use Y1 = poissonpdf (6, x). Since the Poisson distribution is discrete, there will only be output for integer values of x. Therefore the grapher has to be set so that integer values of x fall on the pixel elements. In WINDOW set xmin = 0 and xmax to a multiple of 4.7 equal to about 3μ. Trace can be used to read out the values. Since only integer values of x will be traceable, key in 0, 1, 2, ... Unfortunately the values for x = 0 and for p small will be hidden by the axes. If necessary turn off the axes with 2nd FORMAT AxesOff.

Calculating Probabilities for the Binomial Distribution (Higher Level core only)

Using binompdf and binomcdf

Since the Binomial distribution is discrete, either the cumulative distribution function (cdf) or the probability distribution function (pdf) would be useful for calculating probabilities. Use the pdf to find the probability that one value is observed (X = Xo) & the cdf to find the probability that one of a range of values is observed (X ≤ Xo).

Key 2nd DISTR DISTR (the default) 0: binompdf(n, p[, x]) or Key 2nd DISTR DISTR (the default) A: binomcdf(n, p[, x]), where n is the number of trials, p is the probability of a success in one trial and x the desired number of success. If no x is specified, then a list of probabilities will be generated for x equals zero to n.

Example

If n = 6, p = .75, find P(x = 6).

Keying DISTR DISTR pdf (6, .75, 6) will display 0.178.

Example

If n = 6, p = .75, find P(x > 2) = 1 - P(x ≤ 3).

Keying DISTR DISTR binomcdf (6, .75, 3) will display 0.169. P(x > 2) = 1 - 0.169 = 0.831.

Graphing the Binomial distribution

Use binompdf to graph the Binomial distribution. For example use Y1 = binomcdf (6, .75, x). Since the Binomial distribution is discrete, there will only be output for integer values of x. Therefore the grapher has to be set so that integer values of x fall on the pixel elements. In WINDOW set xmin = 0 and xmax to the smallest multiple of 4.7 bigger than n. Trace can be used to read out the values. Since only integer values of x will be traceable, key in 0, 1, 2, ... Unfortunately the values for x = 0 and for p small will be hidden by the axes. If necessary turn off the axes with 2nd FORMAT AxesOff.

Confidence Intervals

Calculating a Z interval

Zinterval can be used to calculate a Confidence Interval. You can enter your entire sample & have the TI-84 calculate □ or you can enter □ directly.

Key STAT TESTS 7: Zinterval. Then if you are entering □ directly select Stats & key ENTER. Then enter σ, □, n & the desired confidence level (as a decimal, not as a % - it's called the "C-Level"), select Calculate & key ENTER.

If you are given the actual sample numbers, i.e. not □, enter them into a list and then you can either calculate □ as described above (key STAT, CALC, 1: 1-Var Stats, L1) & then use Zinterval Stats.

Or you can use Zinterval Data. In Data you must enter σ, n & the desired confidence level as before, but instead of □ you enter the name of the list containing your data, e.g. L1, select Calculate & key ENTER.

Hypothesis Testing

Conducting a Z-Test

Z-Test is used to test a hypothesis. You can enter your entire sample & have the TI-84 calculate □ or you can enter □ directly. Key STAT, select TEST 1: Z-Test. Then if you are entering $\bar{x}$ directly, select Stats & key ENTER. Then enter μ_o, σ, $\bar{x}$, n & the alternative hypothesis. Select Calculate & key ENTER.

If you are using the actual sample numbers, i.e. not $\bar{x}$, enter the data into a list and then use Z-Test Data. In Data you must enter μ_o, σ, n & the alternative hypothesis as before, but instead of $\bar{x}$ you enter the name of the list containing your data, e.g. L1, select Calculate & key ENTER.

Conducting a t-Test (Higher Level only)

t-Test is used to test a hypothesis. It is more realistic than the z test in that s, the standard deviation calculated from the sample, is used, but it requires that the sample be approximately normal or large. For large samples the z & t tests give the same answer.

You can enter your entire sample & have the TI-84 calculate $\bar{x}$ & s or you can enter □ & s directly. Key STAT, select TEST 2: t-Test. Then if you are entering $\bar{x}$ & s directly, select Stats & key ENTER. Then enter μ_o, $\bar{x}$, Sx (i.e. s) n & the alternative hypothesis. Select Calculate & key ENTER.

If you are using the actual sample numbers, i.e. not $\bar{x}$ & s, enter the data into a list and then use Z-Test Data. In Data you must enter μ_o, σ, n & the alternative hypothesis as before, but instead of $\bar{x}$ you enter the name of the list containing your data, e.g. L1, select Calculate & key ENTER.

Conducting a χ^2 Test for Independence i.e. Contingency Tables

χ^2-Test is used to test a hypothesis of independence with a 2-way contingency table.

First enter your data in a matrix. Key MATRIX, select EDIT, select a matrix to fill or edit, key ENTER, change the r × c (number of rows & columns), if necessary & enter your data.

Now key STAT, TESTS, C: χ^2-Test. Key in the name of the matrix containing your data (Observed) and the name of the matrix where you want the expected values placed by keying MATRIX NAMES, selecting the desired matrix name and keying ENTER. Otherwise use matrices A & B which will appear by default as the observed & expected matrices.

Then choose how to display your results: Draw or Calculate. Draw will draw the χ^2 distribution, and report χ^2 (the value of χ^2) & P (the probability of the observed values, if the null hypothesis of independence were true). Calculate will report χ^2, P & df (the number of degrees of freedom). To view the expected value matrix, key MATRIX, EDIT 2:B (assuming you used B, the default). Note that for a χ^2 test df = (r - 1)(c- 1).

Conducting a χ^2 Test for Independence with the Yates Continuity Correction

When the df = 1, i.e. when the observed is a 2 × 2 table, the IB requires that Yates Continuity Correction be applied. $\chi^2 \text{ (corrected)} \equiv \sum \frac{(|Obs - Exp| - 0.5)^2}{Exp}$.

Enter your observed data in a matrix, say [A]. Make sure matrix B is set to be 2 × 2, using Matrix EDIT and keying in 2 × 2. Key STAT, TESTS, C: χ^2-Test [A] [B] ENTER. Unfortunately I have not been able to find a way to get the TI-84 to do the Yates Continuity Correction, so now you have to do it by hand. Copy out the 4 expected values from [B] & do the math. Example suppose the Observed is A = $\begin{bmatrix} 18 & 10 \\ 8 & 14 \end{bmatrix}$. The TI-84 will give the expected B = $\begin{bmatrix} 14.56 & 13.44 \\ 11.44 & 10.56 \end{bmatrix}$. So now by hand do (|18 - 14.56| - 0.5)²/14.56 + (|10 - 13.44| - 0.5)²/13.44 + (|8 - 11.44 - 0.5)²/11.44 + (|14 - 10.56| - 0.5)²/10.56. Luckily it turns out that the numerator of these 4 terms is always the same for a 2 × 2 table, in our example 8.6436. So you only need to calculate 8.6436 × (1/14.56 + 1/13.44 + 1/11.44 + 1/10.56) = 2.81. Now go to the χ^2 table & find that for df = 1 the critical 5% value is 3.841. Since 2.81 < 3.841, we fail to reject (accept) the assumption of independence.

Conducting a χ^2 Goodness of Fit Test (Higher Level only)

A Goodness of Fit Test tests whether the population fits a model, e.g. binomial, Poisson, uniform, normal, etc. The normal, binomial, Poisson, & geometric probability distributions are in 2nd DISTR. There is no χ^2 Goodness of Fit function in the TI-84, but it is easy to calculate. Put the Observed Values in L1 and the Expected Values (the values that you would get if the model you are testing is correct) in L2. In L3, enter the formula (L1 - L2)² /L2. (To enter a formula scroll up to L3, key ENTER & type it in.)

To find the χ^2 test statistic, enter sum(L3). To find p, enter χ^2cdf(sum(L3),E99,df). χ^2cdf is in 2nd DISTR. E99 is a very good approximation to ∞. Note that for a best fit model df = k - m - 1, where k = the number of data categories and m = the number of parameter values estimated on the basis of the sample data.

Regression and Correlation Analysis

Drawing a Scatter Diagram

First enter your data into lists. See above. Then Key 2nd STAT PLOT, choose a Plot, ENTER, Select ON, Type: scatter (the squiggle of dots in the upper left), the names of your x & y lists (E.g. L1 & L2 - note that these are 2nd 1 & 2nd 2). Then Key GRAPH and ZOOM 9: ZoomStat.

Fitting a line

Eleven kinds of regressions for fitting data to a particular type of equation are available. Only 8: LinReg(a+bx) is needed for the IB. Each of them except D accept the following optional parameters Xlistname, Ylistname, freqlist, regeq. regeq is where the fitted regression equation will be stored. The defaults are L1, L2, 1, RegEQ.

If you type the independent variable into L1 & the dependent variable into L2, you can use the defaults, i.e. avoid keying in the list names. It is useful to have the regression equation, so that you can plot it on top of the scatterplot to see if the fit looks good.

You can paste regeq to Y1 by going to Y1 in Y= and then keying VARS 5: Statistics EQ 1: RegEq. Or if you want the Equation saved to Y1 instead of RegEQ, key LinReg(a+bx) L1, L2, Y1. (Or whatever lists & equation you are using.) Note if you are using the defaults (L1, L2, freqlist =1) they are not needed. The commas between L1, L2 & Y1 are required. “Y1” must be keyed as VARS Y-VARS 1: Function Y1. For example key LinReg(ax+b) Y1 (Y1 is in the VARS, Y-VARS, 1:Function menu.)

To get r & r² to appear

To get r & r² to appear in the screen, set the diagnostics on by keying 2nd CATALOG, (x^{-1} - to get to d faster), DiagnosticOn, ENTER, ENTER.

Key STAT, CALC, 8: LinReg(a+bx), ENTER. a, b, r² & r are displayed.

Covariance

Covariance = $\Sigma xy - \bar{x}\,\bar{y}$. Covariance can be calculated from the data displayed by STAT CALC 2: 2-VAR STATS L1, L2. Scrolling down will display Σxy, $\bar{x}$ & $\bar{y}$.

The equations you can fit:

3: MedMed (a sophisticated linear regression which is less sensitive to outliers than LinReg)

4: LinReg(ax+b) (the standard linear regression)

5: QuadReg (a quadratic regression $\{y = ax^2 + bx + c\}$)

6: CubicReg (a cubic regression$\{y = ax^3 + bx^2 + cx + d\}$)

7: QuartReg (a quartic regression $\{y = ax^4 + bx^3 + cx^2 + dx + e\}$)

8: LinReg(a+bx) (a duplication of 4, but useful because some textbooks use one definition of a linear equation, some the other. **<u>The IB uses this one.</u>**)

9: LnReg (a logarithmic regression $\{y = a + b \ln x\}$)

0: ExpReg (an exponential regression $\{y=a\, b^x\}$)

A: PwrReg (a power law regression $\{y=a\, x^b\}$)

B: Logistic (a logistic regression $\{y = c/[1 + a\, e^{-bx}]\}$)

C: SinReg (a sinusoidal regression $\{y = a \sin(bx + c) + d\}$)

Once the data are keyed in, trying several different regressions (not needed for IB) is very quick and easy. The one with the R² closest to 1 is the best fit to the data. (Equations with a number of parameters - a, b, c, etc. - equal to or greater than the number of data points should give r² = 1, but are not normally considered as valid candidate equations. Including too many terms in the regression equation is called “over-fitting” the model.)

The residuals (not needed for IB) are stored in the list RESID & in EQ 1: RegEQ.