Subject: histogram and binsize problems
Posted by Chad Bender on Thu, 27 Feb 2003 16:52:33 GMT
View Forum Message <> Reply to Message

Hi-

I'm encountering some confusion using the histogram function when specifying the min, max, and binsize keywords. I looked at JD Smith's tutorial. But I couldn't find the answer to my question there, as all I really want is a plain old histogram, not some fancy array manipulation.

First:
IDL> print, !version
{ x86 linux unix 5.4 Sep 25 2000 32 32}

I have an floating point array containing 500 elements, with values restricted to a grid ranging from 0.000 to 0.150 with increment 0.005. What I want is to plot a histogram of this array, with one bin representing each element on my grid (ie. [0.000,0.005,0.010 ... 0.150]).

The first thing I tried was:
IDL> hist=histogram(data,min=0.0,max=0.15,binsize=0.005)

This returns a 31 element vector, which seemed correct as I have 31 possible grid values. So I plot it using:
IDL> plot, findgen(31)*0.005, hist, psym=10

Unfortunately, the plot has several bins EQ 0, where there should be nonzero values. For example, the plot I'm currently looking at has a value of 0 for the bin at 0.065, but I know that 'data' has values there.
IDL> help, where(data EQ 0.065)
<Expression> LONG = Array[43]

So from this I assume that the bins that IDL setup are structured something like this: bin0--[0.000,0.00499], bin1--[0.005,0.00999], etc, and I'm experiencing a floating point error or something where the elements with value 0.065 got included in the bin for 0.06 or 0.07.

To solve this I try to center the grid points in the bin range by doing something like this:
IDL> hist=histogram(data,min=-0.0025,max=0.1525,binsize=0.005)
IDL> help, hist
HIST LONG = Array[32]

Plotting hist shows that every bin that should have data in it does. But now I have 32 bins, instead of 31. This prompted me to realize that in the first example (running from 0 to 0.15), should have resulted in only 30 bins if IDL is making each bin of equal size, and not creating 'half

bins' or something on the ends. So my question is, how does IDL determine the range of each bin, and also the number of bins. And more importantly (for the immediate short term), how can I get a histogram containing 1 bin for every point on my grid, with each bin centered on its corresponding point.

Thanks
Chad

--
################################
Chad Bender
Dept of Physics and Astronomy
SUNY Stony Brook
cbender@mail.astro.sunysb.edu
