Subject: Re: Histogram and bin sizes
Posted by Kenneth P. Bowman on Thu, 21 Feb 2008 22:54:19 GMT
View Forum Message <> Reply to Message

In article
<f6219865-59f4-4bf8-8718-67884c9df226@64g2000hsw.googlegroups.com>,
 Conor <cmancone@gmail.com> wrote:

> Arbitrary bin sizes should be pretty easy to program. You just need
> to map your data points appropriately. For instance if you had the
> data set:
>
> x = randomu(seed,100)
>
> and you wanted bins from:
> [0-.1,.1-.3,.3-.35,.35-.8,.8-1]
>
> you might do something like this:
>
> x = randomu(seed,100)
> bins = [[0,.1], [.1,.3], [.3,.35], [.35,.8], [.8,1]]
> newx = fltarr(n_elements(x))
> for i=0,n_elements(bins[0,*])-1 do begin
> w = where(x ge bins[0,i] and x lt bins[1,i], c)
> if c gt 0 then newx[w] = i+.5
> endfor
>
> hist = histogram(newx,binsize=1.0,min=0)
> plothist,newx

This will work, but will be extremely slow because you test every value
in the input array once for every bin.

The VALUE_LOCATE approach will be much faster, particularly for large
numbers of bins, as it does a binary search.

Ken Bowman