Subject: Re: HISTOGRAM and the Razor's Edge.
Posted by R.G. Stockwell on Thu, 12 Jun 2003 11:33:10 GMT

View Forum Message <> Reply to Message

"Tim Robishaw" <timrobishaw@yahoo.com> wrote in message
news:405594fa.0306112244.6fcd81d6@posting.google.com...
> So I've been reading up on HISTOGRAM and how it is optimized to be
> wicked fast.
> Also, I'm discovering the fancy things one can do with reverse
> indices.
> I've even read JD Smith's exegesis on the topic:
> "HISTOGRAM: The Breathless Horror and Disgust"
>
> So, just when I thought it was safe to start using HISTOGRAM with the
> frequency that Californians use LIKE, I was brought this scary result
> by the guy on the other side of my wall (his name is Tiberius):
>
> IDL> print,
 histogram([-5.50,-5.45,-5.40,-5.35,-5.30,-5.25],min=-5.50,bi nsize=0.05)
>   1   2   0   2   0   1
>
> Wait a minute, this should be a uniform distribution!

No, histogram gives the correct result. Check out what you put into
the histogram function.
IDL> print, [-5.50,-5.45,-5.40,-5.35,-5.30,-5.25],format='(f50.25)'
-5.5000000000000000000000000
-5.4499998092651367000000000
-5.4000000953674316000000000
-5.3499999046325684000000000
-5.3000001907348633000000000
-5.2500000000000000000000000

So, if such razor edge stuff is a concern of yours, preprocess the data
and use integers (i.e. round to integers). This also applies to any
conditional
tests of a float (for instance, for i = 0.1,10,0.001 do begin... etc).

IDL> print, histogram(
 round([-5.50,-5.45,-5.40,-5.35,-5.30,-5.25]/0.05),min=-5.50/ 0.05,binsize=1)
1 1 1 1 1 1

Cheers,
bob