
Subject: Re: histogram, how to trasfer from linear bins to logarithmic bin?

Posted by [penteado](#) on Mon, 28 Sep 2009 02:48:38 GMT

[View Forum Message](#) <> [Reply to Message](#)

On Sep 27, 4:30 pm, Gavin Ge <jiancha...@gmail.com> wrote:

> HI, I have a question that now I get a set of datas, which are
> linearly binned into the histogram, but how to transfer it into a
> logarithmic bins? Thanks.

Either do the histogram of the log of your variable, then use the
exponential of the result, or see

http://www.dfanning.com/code_tips/valuelocate.html

Subject: Re: histogram, how to trasfer from linear bins to logarithmic bin?

Posted by [David Fanning](#) on Mon, 28 Sep 2009 22:54:03 GMT

[View Forum Message](#) <> [Reply to Message](#)

Gavin Ge writes:

> HI, I have a question that now I get a set of datas, which are
> linearly binned into the histogram, but how to transfer it into a
> logarithmic bins?

Here is one way to create a histogram with logarithmic
bins:

http://www.dfanning.com/code_tips/logbins.html

Cheers,

David

--

David Fanning, Ph.D.

Coyote's Guide to IDL Programming (www.dfanning.com)

Sepore ma de ni thui. ("Perhaps thou speakest truth.")

Subject: Re: histogram, how to trasfer from linear bins to logarithmic bin?

Posted by [Brian Larsen](#) on Tue, 29 Sep 2009 14:31:00 GMT

[View Forum Message](#) <> [Reply to Message](#)

Along the lines that David is saying here, when all the value_locate
stuff was hot on the group I wrote up a few routines that *might* be
useful to you.

One of them "log_bins" is like findgen for log spaced bins over a given range. I use this constantly for binning particle energy spectra.

http://people.bu.edu/balarsen/IDLdoc/log_bins.html

Another is "histo_bins" this creates a histogram with use specified bins, it is exactly what David did in his logbins page just bundled up for easier (for me) use.

http://people.bu.edu/balarsen/IDLdoc/histo_bins.html

To recreate David's example:

```
IDL> v1 = Randomu(-3L, 1000) * 100
IDL> v2 = Randomu(-5L, 1000) * 100
IDL> ratio = v1 / v2
IDL> bins=log_bins(0.01, 1000, 6)
IDL> print, log_bins(0.01, 1000, 6)
    0.0100000    0.100000    1.00000    10.0000    100.000
1000.00
IDL> hist=histo_bins(ratio, bins)
IDL> plot, bins, hist, /xlog, psym=10
```

For me this is easier than the labels and xyouts stuff.

Cheers,

Brian

Brian Larsen
Boston University
Center for Space Physics

Subject: Re: histogram, how to trasfer from linear bins to logarithmic bin?

Posted by [Brian Larsen](#) on Tue, 29 Sep 2009 14:35:16 GMT

[View Forum Message](#) <> [Reply to Message](#)

To reply to my own post, I note that there is a 1/2 bin issue that I haven't bothered to fix in plotting from histo_bins. I normally have 100+ bins so I can't tell anyway. If someone uses this and fixes it let me know.

Cheers,

Brian

Brian Larsen
Boston University
Center for Space Physics

Subject: Re: histogram, how to trasfer from linear bins to logarithmic bin?

Posted by [JDS](#) on Tue, 29 Sep 2009 21:33:05 GMT

[View Forum Message](#) <> [Reply to Message](#)

On Sep 28, 6:54 pm, David Fanning <n...@dfanning.com> wrote:

> Gavin Ge writes:

>> HI, I have a question that now I get a set of datas, which are
>> linearly binned into the histogram, but how to transfer it into a
>> logarithmic bins?

>

> Here is one way to create a histogram with logarithmic

> bins:

>

> http://www.dfanning.com/code_tips/logbins.html

This would better be described as a way to create a histogram using any arbitrary bins of your own devising; pretty cool indeed, since you can design those bins in whatever way is useful. It does require you to sort your array beforehand, and so in this case would be less efficient than just taking the histogram of the log of your data.

JD

Subject: Re: histogram, how to trasfer from linear bins to logarithmic bin?

Posted by [David Fanning](#) on Tue, 29 Sep 2009 21:51:20 GMT

[View Forum Message](#) <> [Reply to Message](#)

JDS writes:

> This would better be described as a way to create a histogram using
> any arbitrary bins of your own devising; pretty cool indeed, since you
> can design those bins in whatever way is useful. It does require you
> to sort your array beforehand, and so in this case would be less
> efficient than just taking the histogram of the log of your data.

Actually, as I realized a couple of weeks ago

in Australia when I was teaching this example, the array does NOT need to be sorted in this example. The cutoff vector needs to be monotonically increasing, but the array you are partitioning does not need to be.

I've been using this method (without sorting) to process quite a lot of data recently, and I am **extremely** pleased with how darn fast it is!

I'll update the web page soon, but Thursday I'm heading to my oldest son's wedding and a weekend of partying. Even Value_Locate will be pretty far from my thoughts, I imagine. I'd leave Coyote in charge and have him to it, but, well, you know...

Cheers,

David

--

David Fanning, Ph.D.
Coyote's Guide to IDL Programming (www.dfanning.com)
Sepore ma de ni thui. ("Perhaps thou speakest truth.")

Subject: Re: histogram, how to trasfer from linear bins to logarithmic bin?

Posted by [JDS](#) on Wed, 30 Sep 2009 18:58:00 GMT

[View Forum Message](#) <> [Reply to Message](#)

On Sep 29, 5:51 pm, David Fanning <n...@dfanning.com> wrote:

> JDS writes:

>> This would better be described as a way to create a histogram using
>> any arbitrary bins of your own devising; pretty cool indeed, since you
>> can design those bins in whatever way is useful. It does require you
>> to sort your array beforehand, and so in this case would be less
>> efficient than just taking the histogram of the log of your data.

>

> Actually, as I realized a couple of weeks ago
> in Australia when I was teaching this example,
> the array does NOT need to be sorted in this
> example. The cutoff vector needs to be monotonically
> increasing, but the array you are partitioning does
> not need to be.

>

> I've been using this method (without sorting) to
> process quite a lot of data recently, and I am
> **extremely** pleased with how darn fast it is!

Very good point. The sort is over the bin vector, which can be (and usually is) much shorter than the data vector. And you will likely setup your bin boundary vector sorted to begin with. That said, for me HISTOGRAM(ALOG10) is still faster than HISTOGRAM(VALUE_LOCATE) (see below). You'll also note some "sky is falling" razors-edge differences between bins if you look closely.

JD

```
+++++
```

```
n=100000000L
a=randomu(sd,n)*1.e8
```

```
t=systime(1)
h=histogram(a,alog10(a),BINSIZE=1)
print,'Hist(log))      ',systime(1)-t
```

```
t=systime(1)
mn=min(a,MAX=mx)
mn=alog10(mn) & mx=alog10(mx)
nbin=ceil(mx-mn)
bin=10.^(mn + findgen(nbin))
h2=histogram(value_locate(bin,a))
print,'Hist(value_locate)',systime(1)-t
```

```
print,h,h2
END
```

```
Hist(log))      1.9417701
Hist(value_locate) 3.7843559
      7      60      646      6203      63126
629722
      6286765  62867178  30146293
      7      60      647      6202      63126
629724
      6286771  62867205  30146258
```

Subject: Re: histogram, how to trasfer from linear bins to logarithmic bin?
Posted by [David Fanning](#) on Wed, 30 Sep 2009 19:25:58 GMT
[View Forum Message](#) <> [Reply to Message](#)

JDS writes:

> Very good point. The sort is over the bin vector, which can be (and

> usually is) much shorter than the data vector. And you will likely
> setup your bin boundary vector sorted to begin with. That said, for
> me HISTOGRAM(ALOG10) is still faster than HISTOGRAM(VALUE_LOCATE) (see
> below). You'll also note some "sky is falling" razors-edge
> differences between bins if you look closely.

Yes, I can believe this for logarithmic bins.
I should have been more careful to point out that
the partitioning examples I was raving about are
actually using irregularly sized bins, something
that can only be done in the IDL Way (as far as
I know) with Value_Locate.

Cheers,

David

--

David Fanning, Ph.D.
Fanning Software Consulting, Inc.
Coyote's Guide to IDL Programming: <http://www.dfanning.com/>
Sepore ma de ni thui. ("Perhaps thou speakest truth.")

Subject: Re: histogram, how to trasfer from linear bins to logarithmic bin?
Posted by [David Fanning](#) on Wed, 30 Sep 2009 19:29:28 GMT
[View Forum Message](#) <> [Reply to Message](#)

JD writes:

> Very good point. The sort is over the bin vector, which can be (and
> usually is) much shorter than the data vector. And you will likely
> setup your bin boundary vector sorted to begin with. That said, for
> me HISTOGRAM(ALOG10) is still faster than HISTOGRAM(VALUE_LOCATE) (see
> below). You'll also note some "sky is falling" razors-edge
> differences between bins if you look closely.
>
> Hist(log)) 1.9417701
> Hist(value_locate) 3.7843559

By the way, when I ran your example on my (aging) Windows
machine, I got these results:

Hist(log)) 6.6090002
Hist(value_locate) 5.1719999

Hard to say what *that* means. :-)

Cheers,

David

--

David Fanning, Ph.D.

Fanning Software Consulting, Inc.

Coyote's Guide to IDL Programming: <http://www.dfanning.com/>

Sepore ma de ni thui. ("Perhaps thou speakest truth.")

Subject: Re: histogram, how to trasfer from linear bins to logarithmic bin?

Posted by [Jeremy Bailin](#) on Thu, 01 Oct 2009 18:53:31 GMT

[View Forum Message](#) <> [Reply to Message](#)

On Sep 30, 3:29 pm, David Fanning <n...@dfanning.com> wrote:

> JD writes:

>> Very good point. The sort is over the bin vector, which can be (and
>> usually is) much shorter than the data vector. And you will likely
>> setup your bin boundary vector sorted to begin with. That said, for
>> me HISTOGRAM(ALOG10) is still faster than HISTOGRAM(VALUE_LOCATE) (see
>> below). You'll also note some "sky is falling" razors-edge
>> differences between bins if you look closely.

>

>> Hist(log)) 1.9417701

>> Hist(value_locate) 3.7843559

>

> By the way, when I ran your example on my (aging) Windows
> machine, I got these results:

>

> Hist(log)) 6.6090002

> Hist(value_locate) 5.1719999

>

> Hard to say what *that* means. :-)

>

> Cheers,

>

> David

>

> --

> David Fanning, Ph.D.

> Fanning Software Consulting, Inc.

> Coyote's Guide to IDL Programming: <http://www.dfanning.com/>

> Sepore ma de ni thui. ("Perhaps thou speakest truth.")

It should also depend on the number of bins... for M bins and an N

element data vector, the hist(log) version should work in $O(N)$ while the hist(value_locate) version should work in $O(N \log M + M)$ (the final +M is for setting up the bin cutoffs, but in most cases $N \gg M$ and it doesn't matter). So it depends how well-optimized log is vs. the log of the number of bins... which I can imagine could vary between architectures and C libraries!

-Jeremy.

Subject: Re: histogram, how to transfer from linear bins to logarithmic bin?

Posted by [Jeremy Bailin](#) on Thu, 01 Oct 2009 18:59:51 GMT

[View Forum Message](#) <> [Reply to Message](#)

On Sep 29, 5:33 pm, JDS <jdtsmith.nos...@yahoo.com> wrote:

> On Sep 28, 6:54 pm, David Fanning <n...@dfanning.com> wrote:

>

>> Gavin Ge writes:

>>> HI, I have a question that now I get a set of datas, which are

>>> linearly binned into the histogram, but how to transfer it into a

>>> logarithmic bins?

>

>> Here is one way to create a histogram with logarithmic

>> bins:

>

>> http://www.dfanning.com/code_tips/logbins.html

>

> This would better be described as a way to create a histogram using

> any arbitrary bins of your own devising; pretty cool indeed, since you

> can design those bins in whatever way is useful. It does require you

> to sort your array beforehand, and so in this case would be less

> efficient than just taking the histogram of the log of your data.

>

> JD

As I mentioned in my value_locate essay, value_locate combined with histogram can be very powerful! I'm sure you've already thought of these, but two really good uses are:

(1) using reverse_indices to find out which elements fall into the arbitrary bins (and all associated tricks based on reverse_indices)

(2) using histogram on very sparse data (which is how I solved the histogramming-the-surface-of-the-sphere issue we were discussing)

-Jeremy.
