

Re: How to identify flat (even) distributions?

Source: <http://sci.tech-archive.net/Archive/sci.stat.math/2008-12/msg00058.html>

- *From:* illywhacker <illywacker@xxxxxxxx>
 - *Date:* Thu, 11 Dec 2008 03:25:08 -0800 (PST)
-

On Dec 11, 11:57 am, Steve555 <foursh...@xxxxxxxxxxxxxxxx> wrote:

maybe the chi-squared one will be the fastest to compute.

They all involve summing over all ten score frequencies: it is hard to see that one can avoid this!

Either way, my hunch is that it would be useful, when devising a music recommendation system, to eliminate – or give a low weighting to – the scores of these people.

It is usually better to decide exactly what you are trying to achieve before trying to achieve it. Many questions of this nature are simply linked to a failure to define the goal precisely. Once the goal is defined, whether these people are 'useful' or not should be a question of calculation based on hypothesized models, not hunches. However difficult you may think it is to model human behaviour, you have no choice if you wish to pursue this type of application. You may as well do it explicitly, thereby making your assumptions explicit, for otherwise you will in any case be doing it implicitly, and your assumptions will be hidden and hence un-analysable. Your explicit models may seem ludicrously simplistic, but this is the nature of the application. Your implicit models will also be ludicrously simplistic, but you would not be forced to face up to it and admit it.

illywhacker;

.