Re: (hyper)sensitivity of goodness−of−fit tests

Source: http://sci.tech−archive.net/Archive/sci.stat.math/2006−11/msg00521.html

- *From*: "Old Mac User" <chendrixstats@xxxxxxxxx>
- *Date*: 21 Nov 2006 14:43:52 −0800

RF... Thanks for your post.

This is the point I've tried to make, but it doesn't seem to be getting through the fog.

Regardless of what the OP intends to do with this and all the discussions, there is plainly a lack of fit. It's not that much work to fit these data to an alternative model that will behave properly in both tails and in the middle as well. <sigh>

Have a great Thanksgiving!!

OMU

Reef Fish wrote:

> David Jones wrote:
>
> > Richard Ulrich wrote:
> >
> > > On 20 Nov 2006 07:15:59 −0800, "Old Mac User"
> > > <chendrixstats@xxxxxxxxx> wrote:
> > >
> > > > Why would you go to so much trouble to
> > > > numb the test so that it
> >
> > fails
> >
> > > > to show significance?
> > >
> > > − The obvious next step seems to me to be to *quantify* the
> > > amount of deviation of the fit. The Ns are huge, so the tests
> > > are big, but does it *matter* to the OP? What is the purpose
> > > of
> > > the fit?

The data you posted indicate a strong
departure from an exponential
distribution. Just a simple plot of the data
reveals this, even
without applying a Chi−sq test to verify it.
IMHO this departure

is

bad enough to disallow using an exponential
approximation. Why not
consider another distribution (gamma,
perhaps) and move on with it?
It would be much faster to do that than to
waste most of the data
just to numb the Chi−sq test. OMU

I agree with OMU, "fit something else" −− *if* the sample
size was
wisely chosen to detect a bad 'fit' that matters. But if the

sample

is big because data just happens to be there... what matters
next
is the OP's purpose, and how much the fit (and non−fit)
matters.
Does it matter that the observed 'tail' is much fatter than
predicted?

There are trends in the deviations of the fit. The best clue
about
the nature of a distribution is often in the question, "How was
it
generated?" Does a reason suggest itself? For the purpose of
the
OP, unmentioned so far, it *might* be enough to describe the
fit,
and describe the deviations.

[rest of post included without additional comments.]

morfyssster@xxxxxxxxx wrote:

Thank you very much for
your responses.

What if I took random
subsets of the observed data,
and conducted
the goodness−of−fit tests
using these smaller subsets
and then

used

the average of the p−values
corresponding to these tests?
Would
such an approach be valid?

On Nov 17, 4:16 am, "Reef
Fish"
<large_nassua_grou...@xxxxxxxxx>
wrote:

morfyss...@xxxxxxxxx
wrote:

I
have
a
large
amount
of
empirical
data
consisting
of

interarrival

times
that
I
believe
are
exponentially
distributed.
Looking
at
the

quantile−quantile
plot
between
the
empirical
and
theoretical/fitted
distribution,
I
see
an
almost
perfect
linear
relationship.

While agreeing what what others have posted in response, I would like to point out:

(i) the Q−Q plot should not be looked at for "an almost perfect linear relationship", but for departures from a 1:1 line;

Good point to emphasize. In infinitessimal departure from a perfect linear fit, but with residuals
−−−−−−−−−−−−−−+++++++++−−−−−−−−−−−−−−−−−−−−
pattern is a significant departure. That is in fact the kind of departure the eye can easily detect where as the analytic tests will NOT.

Any systematic SMALL departures are equally BAD, e.g.
+++++++++−−−−−−−−−−−−−−−−−−−−−−++++++++++++++

or
++++++++++−−−−−−−−−−−−−−−−−−−−−−−+++++++++++++++++−−−−−−−−−−−−−−−−−−−−−−−−−

of the kind of TOO FEW or TOO MANY runs.

−− Reef Fish Bob.

(ii) there are other graphical procedures more or less specifically designed to look for departures from an exponential distribution .. see "log−survivor plots" and/or "log−survival plots" for example;

(iii) there is an extensive literature on survival analysis / reliability / inter−arrival times that contain various well−understood alternatives to the exponential. As Richard has pointed out, context is important, and it may be that other similar applications have already homed in on a suitable alternative for the case here.

(iv) other explanations of apparent departures from the null hypothesis in large samples arise from some other aspect of the null hypothesis not holding: for example the data may not consist of statistically independent values, or the data may not arise from a fixed distribution, for example if there are seasonal/time−within−day effects that are not being modelled: again context is important here.

David Jones

.