Re: Continuity correction.

Source: http://sci.tech−archive.net/Archive/sci.stat.math/2007−10/msg00245.html

- *From*: "David A. Heiser" <dah_box1@xxxxxxxxxxxxx>

- *Date*: Wed, 24 Oct 2007 11:13:42 −0700

"a.riva@UCL" <anto.ilmago@xxxxxxxxx> wrote in message
news:1193216668.258398.172480@xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx

> On 23 Oct, 19:55, "David A. Heiser" <dah_b...@xxxxxxxxxxxxx> wrote:
>
> > "a.riva@UCL" <anto.ilm...@xxxxxxxxx> wrote in message
> >
> > news:1193155810.098400.152250@xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
> >
> > > On 23 Oct, 17:03, Ken Butler <but...@xxxxxxxxxxxxxxxx>
> > > wrote:
> > >
> > > > On Tue, 23 Oct 2007 15:44:31 −0000,
> > > > "a.riva@UCL"
> > >
> > > > <anto.ilm...@xxxxxxxxx> wrote:
> > >
> > > > > So, this means that if I want
> > > > > to calculate the probability
> > > > > associated
> > > > > to X>=n (with X positive
> > > > > this corresponds to the right
> > > > > tail of the
> > > > > distribution) I should
> > > > > convert everything to the
> > > > > calculation of
> > > > > P(Z>=n−0.5) so that I
> > > > > include n. And using
> > > > > NORMSDIST(Abs(z)), the
> > > > > cumulative probability left
> > > > > to n−0.5 is
> > > > > NORMSDIST(Abs(z(n−0.5))),
> > > > > and

> the probability associated
> with the right tail is 1−
> NORMSDIST(Abs(z(n−0.5))).
> Is this right?

I don't claim to be an expert in
NORMSDIST, but it looks right to me.
−−
Ken Butler, Lecturer (Statistics)
University of Toronto at Scarborough
butler (at)
utsc.utoronto.cahttp://www.utsc.utoronto.ca/~butler

Great!!!

Thanks a lot!!! Unfortunately Excel is really limited for statistical
analyses, but it's generally the software that we use to create all
our databases, and sometimes it's really quicker to try to implement
it instead of moving all the data in other softwares...

Cheers,

Antonio.

++
Actually for Excel 2003 and 2007, NORMSDIST is accurate from z values of −37.5 upward. The lower limit is from the lower limit of the IEEE−754 double. There is a region of around z from −3 to −8 where the accuracy drops
to 6 significant decimal digits. Outside of this the accuracy is in the region of 14 decimal digits.

NORMSDIST has been pretty well tested. See my URL on this (search for "Excel
Faults")

What is your rational for saying that Excel is limited here?

David Heiser− Hide quoted text −

− Show quoted text −

Hi David.

Sorry for my comment about Excel. I didn't explain accurately what I mean.
I was not actually expressing doubts about the accuracy of the functions that Excel already contains. I have to say that I find them really useful and definitely precise and accurate. What I exactly meant is that some statistical tests or functions or graphs are not included in Excel. For example, non parametric tests, GLMs, boxplots, ranking function which takes ties into consideration. If one wants to perform them, they have to be "built" combining the existing functions. This is absolutely fine, but not always straightforward :−)
It's true that Excel wasn't born as a statistical analysis software, but it's more and more used also for this purpose.

++
Yes you are right here. Excel is very limited on statistical functions/routines.

From about 1998 on, it was very popular to write new introductory stat

textbooks that included use of Minitab, Excel, et. al, to solve problems. Publishers loved this, because they could include a cdrom with all the problems sets, and their own add−ins to do things that the software could not do. They now had a reason to increase textbook costs, and create a market to sell their add−in software. All the hidden stat−teacher−authors now had a reason to revise/rewrite their texts for new editions, making the used texts no longer useable. Great marketing ploy.

The text books were a hidden means to sell software. This is still continuing today.

MIcrosoft viewed Excel as a business tool. The stat functions and routines were just incidental add−ins that were in the original Graymatter software that Microsoft bought, since the focus was on the financial functions it contained. The stat is all 1960's stuff and has never been updated as to context.

The wide use of Microsoft Office, essentially has created an international body of users, that rely on Excel and the functions to solve their problems.

There are hidden issues regarding "competition", "copy−rights", "algorithm patents" and "market dominance" (especially in Europe), that make the issue of expanding Excel's stat capabilities (in competition to all the other

commercial stat software and Excell add−ns) an uncertain area to proceed in. Microsoft would rather list available commercial add−ins. They see no marketing advantage to add to Excel's stat capabilities.

So there you have it, a 1960's view of statistics, locked into all future time.

If business pushed Microsoft to add to the stat capabilites, Microsoft probably would. However, that is not what business' want. Just spend some time on the Excel related lists.

Like everything else in our world, it is only a matter of dollars.

David Heiser
++=

> Just this.
>
> By the way, could you explain to me what is the logic behind the fact that NORMSDIST returns the cumulative probability "left−to−right" whilst TDIST returns the cumulative probability "right−to−left"?
>
> Thanks, and I hope I explained myself better now :−)
>
> Antonio.

++
No I can't. It all comes from historic usage, of negative being on the left and positive being on the right. The normal is viewed as a stand−alone concept. The t distribution only as a test, not a visual concept of the distribution. The F distribution shares this too. The chi−square is like the normal, a stand−alone distribution.

David Heiser
++

.