

Non-central Chi-Squared Probabilities – Algorithms in R

Martin Mächler
Seminar für Statistik
ETH Zurich

September 24, 2019

Abstract

1 Introduction

2 Non-central chi-squared probabilities: History of R's `pnchisq.c`

The very early versions of R, the oldest still available, the pre-alpha version (before there were version numbers) with source file named `R-unix-src.tar.gz`, dated June 20, 1995, or the oldest still running version 0.16.1, February 11, 1997, already had R functions for the *non*-central chi-squared (called “*Chi-Square*” there) distribution, at the time functions separate from the central chi-squared, the four non-central functions were called `dncchisq()`, `pnchisq()`, etc, note the extra “n”, and had their own separate help page (which was wrongly almost identical to the central chi-squared); e.g. `pnchisq()` with three arguments, already gave the correct result, e.g., for

```
> pnchisq(1,1,1)
```

```
[1] 0.4772499
```

In R version 0.50 “Alpha-4” (September 10, 1997), the help page was correct, and the 4 functions all were shown to have 3 arguments, e.g., `pnchisq(q, df, lambda)`.

The source code `R-<ver>/src/math/pnchisq.c` and then, for 0.62 and newer, using directory name `nmath/` had been practically unchanged from the earliest version up to version 0.61.3 (May 3, 1998). The algorithmic implementation in C was just summing up the Poisson weighted central term `term`, “`while (term >= acc)`” with the constant declaration `double acc = 1.0e-12;`

For R version 0.62 (1998-06-14), on the R level, the **n** versions of function names became deprecated and the noncentrality parameter was changed from `lambda` to `ncp` and added to the “non-n” version of the functions, e.g. R's `pchisq()` became defined as

```
> pchisq
function (q, df, ncp = 0)

  if (missing(ncp))
    .Internal(pchisq(q, df))
  else .Internal(pnchisq(q, df, ncp))
```

and the source file `src/nmath/pnchisq.c` (with timestamp 1998-03-17 04:56 and a size of 2669 bytes) now did implement the algorithm AS 275 by Ding (1992). The NEWS entry (still in the R sources `doc/NEWS.0`) has been

CHANGES IN R VERSION 0.62

NEW FEATURES

.....
.....

o Some of the `t`, `F`, and `chisq` distribution/probability functions now allow a noncentrality parameter ``ncp'`.

But even then, the new `pnchisq.c` with AS 275 contained a comment

Feb.1, 1998 (R 0.62 alpha); M.Maechler: still have INFINITE loop and/or bad precision in some cases.

At the time I had been pretty confident we'd eliminate these cases pretty quickly. In the meant time, many tweaks have been made, to a large extent by myself, and the code of today works accurately in many more cases than in early 1998. On the other hand, the help page has warned for years now that only moderate values of the noncentrality parameter `ncp` where feasible, and still in R 3.6.1 (July 2019), you can find calls to `pchisq()` which lead to an "infinite loop" (at least on 64-bit Linux and Windows), also for small values of `ncp`, e.g.,

```
> pchisq(1.00000012e200, df=1e200, ncp=100)
```

References

Ding, C. G. (1992). Algorithm AS 275: Computing the non-central χ^2 distribution function. *Applied Statistics — Journal of the Royal Statistical Society C*, 41(2):478–482.