

# Hypergeometric Functions for Statistical Computing

V. E. McHale

June 1, 2022

## Abstract

We perform familiar statistical computations transparently with hypergeometric functions. The program was suggested by Shaw, we offer some commentary.

## 1 Introduction

The hypergeometric function is given by

$${}_pF_q(a_1, \dots, a_p; b_1, \dots, b_q; z) = \sum_{n=0}^{\infty} \frac{(a_1)_n \cdots (a_p)_n z^n}{(b_1)_n \cdots (b_q)_n n!}$$

where  $(a)_n = a(a+1)\cdots(a+n-1)$ ,  $(a)_0 = 1$ .

It is implemented in J as the hypergeometric conjunction `H.`,  
(a1...ap H. b1...bq) z =  ${}_pF_q(a_1, \dots, a_p; b_1, \dots, b_q; z)$ .

## 2 Statistics

### 2.1 Erf

We have [9]:

$$\begin{aligned} \operatorname{erf}(z) &= \frac{2}{\sqrt{\pi}} \int_0^z e^{-t^2} dt \\ &= \frac{2z}{\sqrt{\pi}} {}_1F_1\left(\frac{1}{2}; \frac{3}{2}; -z^2\right) \\ &= \frac{2ze^{-z^2}}{\sqrt{\pi}} {}_1F_1\left(1; \frac{3}{2}; z^2\right) \end{aligned}$$

by expansion of power series [6, p. 408][5].

The former has convergence problems, as pointed out by Shaw [7]. We can see this experimentally, viz.

```

erfbad =: {{ 2p_0.5 * y * (0.5 H. 1.5) (-y^2) }}
erfbad 7
_814.656

```

The latter:

```

erf =: {{ 2p_0.5 * y * (^(-y^2)) * (1 H. 1.5) (y^2) }}

```

## 2.2 Normal CDF

The above allows us to define the cdf for the standard normal  $N(0, 1)$  [3], viz.

```

cdf =: {{ -: >: erf (y%:2) }}

```

This is equivalent to the approach was used by Marsaglia in 2004 [5], independently of Shaw [7].

## 2.3 t-Distribution CDF

For  $\nu$  degrees of freedom we have [1]:

$$\frac{1}{2} + x \frac{\Gamma(\frac{1}{2}(\nu + 1))}{\sqrt{\pi\nu}\Gamma(\frac{\nu}{2})} {}_2F_1\left(\frac{1}{2}, \frac{1}{2}(\nu + 1); \frac{3}{2}; -\frac{x^2}{\nu}\right)$$

for  $|x| < \sqrt{\nu}$  ( ${}_2F_1$  converges if and only if  $|z| < 1$ ).

Hui [4] gives the J dyad:

```

cdf =: 4 : 0
    assert. (%:x)>|y
    0.5 + y * (!:-x-1) * ((0.5,-:1+x) H. 1.5 (-*:y)%x) % (%:o.x) * !<:-:x
)

```

(note that !n is  $\Gamma(n + 1)$ )

## 2.4 $\chi^2$ -distribution

The CDF for  $r$  degrees of freedom is given by [8]

$$\frac{\gamma(\frac{r}{2}, \frac{x}{2})}{\Gamma(\frac{r}{2})}$$

where  $\gamma$  is the incomplete gamma function [10]

$$\begin{aligned} \gamma(a, x) &= \int_0^x t^{a-1} e^{-t} dt \\ &= a^{-1} x^a e^{-x} {}_1F_1(a; 1 + a; -x) \end{aligned}$$

We have the following elegant J implementation [2]

```

gamma =: !@:<:
incgam =: {{ (1 H. (1+x) % x&(( * ^ ) * (^-)^)) y }} % gamma@[
chisqcdf =: incgam&-:

```

### 3 Conclusion

Implementations in J are fairly transparent; we do not need to translate mathematical notions to loops.

### References

- [1] D. E. Amos. Representations of the central and non-central  $t$  distributions. *Biometrika*, 51(3 and 4), 1964. 2
- [2] Roger Hui. Chi squared cdf. [https://code.jsoftware.com/wiki/Essays/Chi\\_Squared\\_CDF](https://code.jsoftware.com/wiki/Essays/Chi_Squared_CDF). 2
- [3] Roger Hui. Normal cdf. [https://code.jsoftware.com/wiki/Essays/Normal\\_CDF](https://code.jsoftware.com/wiki/Essays/Normal_CDF). 2
- [4] Roger Hui. t-distribution cdf. [https://code.jsoftware.com/wiki/Essays/t-Distribution\\_CDF](https://code.jsoftware.com/wiki/Essays/t-Distribution_CDF). 2
- [5] George Marsaglia. Evaluating the normal distribution. *Journal of Statistical Software*, 11, 2004. 1, 2
- [6] Keith B. Oldham, Jan Myland, and Jerome Spanier. *An Atlas of Mathematical Functions*. Springer, second edition, 2009. 1
- [7] Ewart Shaw. Hypergeometric functions and cdfs in j. *Vector*, 18(4), 2002. 1, 2
- [8] Eric W. Weinstein. Chi-squared distribution. <https://mathworld.wolfram.com/Chi-SquaredDistribution.html>. 2
- [9] Eric W. Weinstein. Erf. <https://mathworld.wolfram.com/Erf.html>. 1
- [10] Eric W Weinstein. Incomplete gamma function. <https://mathworld.wolfram.com/IncompleteGammaFunction.html>. 2