NIfTI-1 Statistical Distributions:
Descriptions and Sample C Functions

Robert W Cox, PhD
Scientific and Statistical Computing Core
National Institute of Mental Health
National Institutes of Health
Department of Health and Human Services

[Draft of March 7, 2004]

Introduction

The NIfTI-1 image format (http://nifti.nimh.nih.gov/dfwg) contains a field `intent_code` which signals the interpretation of the data values in the file. A number of codes are defined that correspond to standard statistical distributions. My goal in this document is to specify what these codes refer to, and also to document the user-level sample C functions provided at the NIfTI website.

The table below lists the distributions, their NIfTI-1 codes, and the number of parameters each distribution has.

Code	Macro Name	Distribution Name	Parameters
2	`NIFTI_INTENT_CORREL`	correlation statistic	1
3	`NIFTI_INTENT_TTEST`	t statistic (central)	1
4	`NIFTI_INTENT_FTEST`	F statistic (central)	2
5	`NIFTI_INTENT_ZSCORE`	$N(0,1)$ statistic	0
6	`NIFTI_INTENT_CHISQ`	chi-squared (central)	1
7	`NIFTI_INTENT_BETA`	Beta variable (central)	2
8	`NIFTI_INTENT_BINOM`	Binomial variable	2
9	`NIFTI_INTENT_GAMMA`	Gamma distribution	2
10	`NIFTI_INTENT_POISSON`	Poisson distribution	1
11	`NIFTI_INTENT_NORMAL`	$N(\mu,\sigma^2)$ normal	2
12	`NIFTI_INTENT_FTEST_NONC`	noncentral F statistic	3
13	`NIFTI_INTENT_CHISQ_NONC`	noncentral chi-squared	2
14	`NIFTI_INTENT_LOGISTIC`	Logistic distribution	2
15	`NIFTI_INTENT_LAPLACE`	Laplace distribution	2
16	`NIFTI_INTENT_UNIFORM`	Uniform distribution	2
17	`NIFTI_INTENT_TTEST_NONC`	noncentral t statistic	2
18	`NIFTI_INTENT_WEIBULL`	Weibull distribution	3
19	`NIFTI_INTENT_CHI`	chi statistic (central)	1
20	`NIFTI_INTENT_INVGAUSS`	inverse Gaussian variable	2
21	`NIFTI_INTENT_EXTVAL`	Extreme value type I	2
22	`NIFTI_INTENT_PVAL`	p-value	0

Distribution codes 2...10 are from the AFNI package.

The following 3 books are repeatedly referred to in the sections below:

[U] Univariate Discrete Distributions, by NL Johnson, S Kotz, AW Kemp.

[C1] Continuous Univariate Distributions, vol. 1, by NL Johnson, S Kotz, N Balakrishnan.

[C2] Continuous Univariate Distributions, vol. 2, NL Johnson, S Kotz, N Balakrishnan.

The chapter numbers in these books continue across volumes; that is, [U] runs from Chapter 1 to 11, [C1] from Chapter 12 to 21, and [C2] from Chapter 22 to 33.

These books contain a wealth of information about many distribution functions. I will not attempt to duplicate their coverage or to provide any coherent introduction to statistical theory. My primary goal here is simply to define precisely to which distributions the above table refers.

For each distributional parameter, its usual mathematical symbol is used in the discussion below. The NIfTI-1 header field name is also given (e.g., `intent_p1`). Please see the `nifti1.h` header file for details about how the distributional parameters are stored in a NIfTI-1 header.

2) `NIFTI_INTENT_CORREL` = correlation statistic [C2, chap 32]

This random variable is often denoted by the variable R, and takes on values in the interval $[-1, 1]$. The distribution has a single positive parameter, here called ν [`intent_p1`] (equal to [C2's] $n - 2$). Properties:

- $R/\sqrt{1 - R^2}$ is t-distributed with ν degrees of freedom
- $(R + 1)/2$ is Beta-distributed with parameters $a = b = \nu/2$.
- The density of R is proportional to $(1 - R^2)^{\nu/2-1}$.

This distribution is the special case of the **null** distribution of a general correlation coefficient; that is, $E[R] = 0$. [C2, chap 32] deals with the more complex case when the expected value of R is not 0.

The correlation statistic arises in these FMRI papers (among others):

> PA Bandettini, A Jesmanowicz, EC Wong, JS Hyde, Processing strategies for time-course data sets in functional MRI of the human brain. *Magn. Reson. Med.* **30**: 161-173 (1993).
>
> RW Cox, A Jesmanowicz, and JS Hyde, Real-time functional magnetic resonance imaging. *Magn. Reson. Med.* **33**: 230-236 (1995).

Given an M-dimensional voxel time series vector $\mathbf{x}$ fitted to an $L + 1$ dimensional model $\mathbf{x} = \alpha\mathbf{r} + \sum_{k=1}^{L} \gamma_k \mathbf{s}_k + \boldsymbol{\xi}$, where $\mathbf{r}$ is the given reference time series model being fitted,

$\{\mathbf{s}_k : k = 1 \ldots L\}$ is the given set of baseline vectors, $\{\alpha, \gamma_1, \ldots, \gamma_L\}$ are unknowns to be estimated, and $\boldsymbol{\xi} \sim N(0, \sigma^2\mathbf{I})$ (i.e., Gaussian white noise), then under the null hypothesis ($\alpha = 0$), the correlation coefficient R of $\mathbf{x}$ with $\mathbf{r}$, after removing the baseline model with linear least squares from the data, has the null correlation distribution with parameter $\nu = M - L - 1$; that is, one degree of freedom is lost for every regressor.

3) `NIFTI_INTENT_TTEST` = t statistic (central) [C2, chap 28]

This distribution takes on values in the unbounded interval $[0, \infty)$. The t-distribution has a single positive parameter ν [`intent_p1`], usually referred to as the "degrees of freedom". The mean of t_ν is 0. As $\nu \to \infty$, the distribution of t_ν approaches the standard normal distribution $N(0, 1)$. The t-statistic often arises when testing if a single regression coefficient is significantly different from zero. The density is proportional to $[1 + x^2/\nu]^{-(\nu+1)/2}$. Also, t_ν^2 is distributed like a central F-variable with $(1, \nu)$ degrees of freedom.

4) `NIFTI_INTENT_FTEST` = F statistic (central) [C2, chap 27]

This distribution takes on values in $[0, \infty)$. The F-distribution has two positive parameters, ν_1 [`intent_p1`] and ν_2 [`intent_p2`], usually referred to as the "numerator degrees of freedom" and the "denominator degrees of freedom", respectively. The density is proportional to $x^{\nu_1/2-1} / [1 + \nu_1 x/\nu_2]^{(\nu_1+\nu_2)/2}$. The F-distribution often arises when testing if a multi-parameter regression model makes any significant contribution to fitting the data.

5) `NIFTI_INTENT_ZSCORE` = $N(0, 1)$ statistic [C1, chap 13]

The most famous of distributions, it takes on values in $(-\infty, \infty)$. This particular statistical code takes no parameters, and has the density function

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \; .$$

The z-score often arises as a transformation of some other statistic, in order to bring that value into a standard form. In such cases, it serves much the same psychological purpose as a p-value.

6) `NIFTI_INTENT_CHISQ` = chi-squared (central) [C1, chap 18]

This distribution takes on values in $[0, \infty)$. The χ^2 distribution has one positive parameter ν [`intent_p1`], usually referred to as the "degrees of freedom". If $\{\xi_k : k = 1 \ldots L\}$ is a set of $N(0, 1)$ random variables, then $X = \sum_{k=1}^{L} \xi_k^2$ has this distribution with L degrees of freedom. The χ_ν^2 distribution is a special case of the Gamma distribution with shape parameter $\nu/2$ and scale parameter 2.

7) NIFTI_INTENT_BETA = Beta variable (central) [C2, chap 25]

This distribution takes on values in $[0,1]$. It has 2 positive parameters, usually called a [intent_p1] and b [intent_p2] (although [C2] uses the symbols p and q). Its density is proportional to $x^{a-1}(1-x)^{b-1}$. If χ^2_ν and χ^2_μ are two independent chi-squared variables, then the ratio $\chi^2_\nu/(\chi^2_\nu+\chi^2_\mu)$ has a Beta distribution with $a=\nu/2$ and $b=\mu/2$.

8) NIFTI_INTENT_BINOM = Binomial variable [U, chap 3]

This is a discrete distribution with two parameters: n=number of trials [intent_p1], and p=probability of "success" per trial [intent_p2], with $n>0$ and $p\in(0,1)$. The distribution takes on values $\{0,\ 1,\ 2,\ \ldots,\ n\}$; the probability that value x occurs is

$$P_x = \binom{n}{x} p^x (1-p)^{n-x} \ .$$

If n independent experiments are carried out, each of which separately has probability p of success, then P_x is the probability that x of them are successful. One might use this in FMRI for some sort of group statistics.

9) NIFTI_INTENT_GAMMA = Gamma distribution [C1, chap 17]

This distribution takes on values in $[0,\infty)$. It has two positive parameters; the "shape" α [intent_p1] and the "scale" β [intent_p2]. The density is

$$f(x) = \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-x/\beta} \ .$$

This function is often used for non-probabilistic purposes in FMRI time series model fitting; it should be called a "gamma variate" or "gamma density" function, but a few benighted folk call this a "gamma function".

10) NIFTI_INTENT_POISSON = Poisson distribution [U, chap 4]

This is a discrete distribution with one positive parameter, the mean θ [intent_p1]. It takes on all nonpositive integral values $\{0,\ 1,\ 2,\ \ldots\}$. The probability that x occurs is

$$P_x = \frac{1}{x!} e^{-\theta} \theta^x \ .$$

I have no idea what use this distribution might have in FMRI data analysis. But I'm also sure someone will think of one, someday.

11) NIFTI_INTENT_NORMAL = $N(\mu,\sigma^2)$ normal [C1, chap 13]

This distribution takes values in $(-\infty,\infty)$. It takes two parameters: μ ("location" or "mean") [intent_p1], and σ ("scale" or "standard deviation") [intent_p2]. Its density function is

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/(2\sigma^2)} \ .$$

This distribution arises disgustingly often.

12) NIFTI_INTENT_FTEST_NONC = noncentral F statistic [C2, chap 30]

This distribution takes on values in $[0, \infty)$. It has three positive parameters: ν_1 ("numerator degrees of freedom") [intent_p1], ν_2 ("denominator degrees of freedom") [intent_p2], and λ ("noncentrality parameter") [intent_p3]. When $\lambda = 0$, the distribution is the same as the central F-distribution. The noncentral F density function is too complex to give here. This statistic arises in power analysis of regression models (including ANOVA).

If $\chi'^2_{\nu_1}(\lambda)$ is a noncentral chi-squared variable and $\chi^2_{\nu_2}$ is an independent central chi-squared variable, then the ratio $[\chi'^2_{\nu_1}(\lambda)/\nu_1][\chi^2_{\nu_2}/\nu_2]^{-1}$ has the noncentral F distribution.

13) NIFTI_INTENT_CHISQ_NONC = noncentral chi-squared [C2, chap 29]

This distribution takes on values in $[0, \infty)$. It has two positive parameters: ν ("degrees of freedom") [intent_p1], and λ ("noncentrality parameter") [intent_p2]. When $\lambda = 0$, the distribution is the same as the central chi-squared distribution. If $\{\xi_k : k = 1 \dots L\}$ is a set of $N(0,1)$ random variables, and $\{c_k : k = 1 \dots L\}$ is a set of constants, then $X = \sum_{k=1}^{L} (\xi_k + c_k)^2$ has this distribution with $\nu = L$ and $\lambda = \sum_{k=1}^{L} c_k^2$. The noncentral chi-squared density function is too complex to give here (unless you like Bessel functions).

14) NIFTI_INTENT_LOGISTIC = Logistic distribution [C2, chap 23]

This distribution takes values in $(-\infty, \infty)$. It takes two parameters: μ ("location") [intent_p1], and σ ("scale") [intent_p2]. Its density function is

$$f(x) = \frac{1}{4\sigma} \mathrm{sech}^2 \left(\frac{x - \mu}{2\sigma} \right) .$$

This distribution is often used as a long-tailed replacement for the Normal distribution, since its density decays proportionally as $\exp[-|x - \mu|/\sigma]$ rather than as $\exp[-|x - \mu|^2/(2\sigma^2)]$ as $|x| \to \infty$.

15) NIFTI_INTENT_LAPLACE = Laplace distribution [C2, chap 24]

This distribution takes values in $(-\infty, \infty)$. It takes two parameters: μ ("location") [intent_p1], and σ ("scale") [intent_p2]. Its density function is

$$f(x) = \frac{1}{2\sigma} e^{-|x-\mu|/\sigma} .$$

This distribution is also sometimes used as a long-tailed replacement for the Normal distribution. It has the interesting property that the maximum likelihood estimator for μ, given a sample of i.i.d. Laplace variables, is the sample median. This property is analogous to the Normal distribution's property that the maximum likelihood estimator of the location is the sample mean.

16) `NIFTI_INTENT_UNIFORM` **= Uniform distribution [C2, chap 26]**

This distribution takes two parameters, a [`intent_p1`], and b [`intent_p2`], with $a < b$ required. Values are in the interval $[a, b]$, and are all equally likely.

17) `NIFTI_INTENT_TTEST_NONC` **= noncentral t statistic [C2, chap 31]**

This distribution takes on values in $(-\infty, \infty)$. It has two parameters: ν ("degrees of freedom") [`intent_p1`], which must be positive, and δ ("noncentrality parameter") [`intent_p2`], which is not restricted. When $\delta = 0$, the distribution is the same as the standard t-distribution. If z has a $N(0, 1)$ distribution, χ^2_ν has a central chi-squared distribution, and δ is a constant, then the ratio $t = (z + \delta)/\sqrt{\chi^2_\nu/\nu}$ has a noncentral t-distribution. Also, t^2 follows a noncentral F-distribution with parameters (ν_1, ν_2, λ) given by $(1, \nu, \delta^2)$. The noncentral t density function is far far too complex to give here.

18) `NIFTI_INTENT_WEIBULL` **= Weibull distribution [C1, chap 21]**

This distribution depends on three parameters: ξ_0 ("location") [`intent_p1`], α ("scale") [`intent_p2`], and c ("power") [`intent_p3`]. This distribution takes values in $[\xi_0, \infty)$. Its cumulative distribution function is

$$F(x) = 1 - \exp\left[-\{(x - \xi_0)/\alpha\}^c\right] \ ,$$

and so its density is proportional to

$$[(x - \xi_0)/\alpha]^{c-1} \exp\left[-\{(x - \xi_0)/\alpha\}^c\right] \ .$$

19) `NIFTI_INTENT_CHI` **= chi statistic (central) [C1, chap 18]**

This distribution takes values in $[0, \infty)$. It has one positive parameter ν ("degrees of freedom") [`intent_p1`]. This distribution is just the distribution of the positive square root of a central chi-squared variable with the same degrees of freedom. Special cases: $\nu = 1$ is the half-normal distribution (i.e., the distribution of $|z|$ when $z \sim N(0, 1)$); $\nu = 2$ is the Rayleigh distribution; and $\nu = 3$ is the Maxwell-Boltzmann distribution. The density is proportional to $e^{-x^2/2} x^{\nu-1}$.

20) `NIFTI_INTENT_INVGAUSS` **= inverse Gaussian variable [C1, chap 15]**

This distribution takes values in $[0, \infty)$. It has two positive parameters: μ ("mean") [`intent_p1`], and λ [`intent_p2`]. The density function is

$$f(x) = \left[\frac{\lambda}{2\pi x^3}\right]^{1/2} \exp\left\{-\frac{\lambda}{2\mu}\left(\frac{x}{\mu} - 2 + \frac{\mu}{x}\right)\right\} \ .$$

The mean is μ and the variance is μ^3/λ. As $\lambda \to \infty$, the distribution becomes "more normal".

21) `NIFTI_INTENT_EXTVAL` = Extreme value type I [C2, chap 22]

This distribution takes values in $(-\infty, \infty)$. It has two parameters: ξ ("location") [`intent_p1`], and $\theta > 0$ ("scale") [`intent_p2`]. The cumulative distribution function is $F(x) = \exp\left[-e^{-(x-\xi)/\theta}\right]$. The density is left as an exercise in differentiation.

22) `NIFTI_INTENT_PVAL` = *p*-value

There isn't much to say here. By p-value is usually meant the tail mass of some probability distribution. Or maybe the double-sided tail mass. In any case, the values stored in a NIfTI-1 dataset when `intent_code=NIFTI_INTENT_PVAL` should be in the interval $[0, 1]$ for this to make any sense. This "distribution" takes no parameters.

Sample C Functions

The file `nifti_stats.c` contains a set of functions to simplify using the above distributions. Most of this file comprises utility functions, not suitable for direct calling by a user-level program. Most of these utility functions in `nifti_stats.c` are from the the cdflib functions by Brown and Lovato; the noncentral t cdf is by Krishnamoorthy; a few distribution and inverse functions are by myself.

The user-callable function names all start with "`nifti_`". Each function has the same five input arguments:

```
( double val , int code , double p1 , double p2 , double p3 )
```

and each function returns a `double`. The input `code` determines which distribution is used. The inputs `p1`, `p2`, and `p3` are the parameters of the distribution (corresponding to `intent_p1`, etc., as described for each distribution above); depending on `code`, none or all of these parameters may be used.

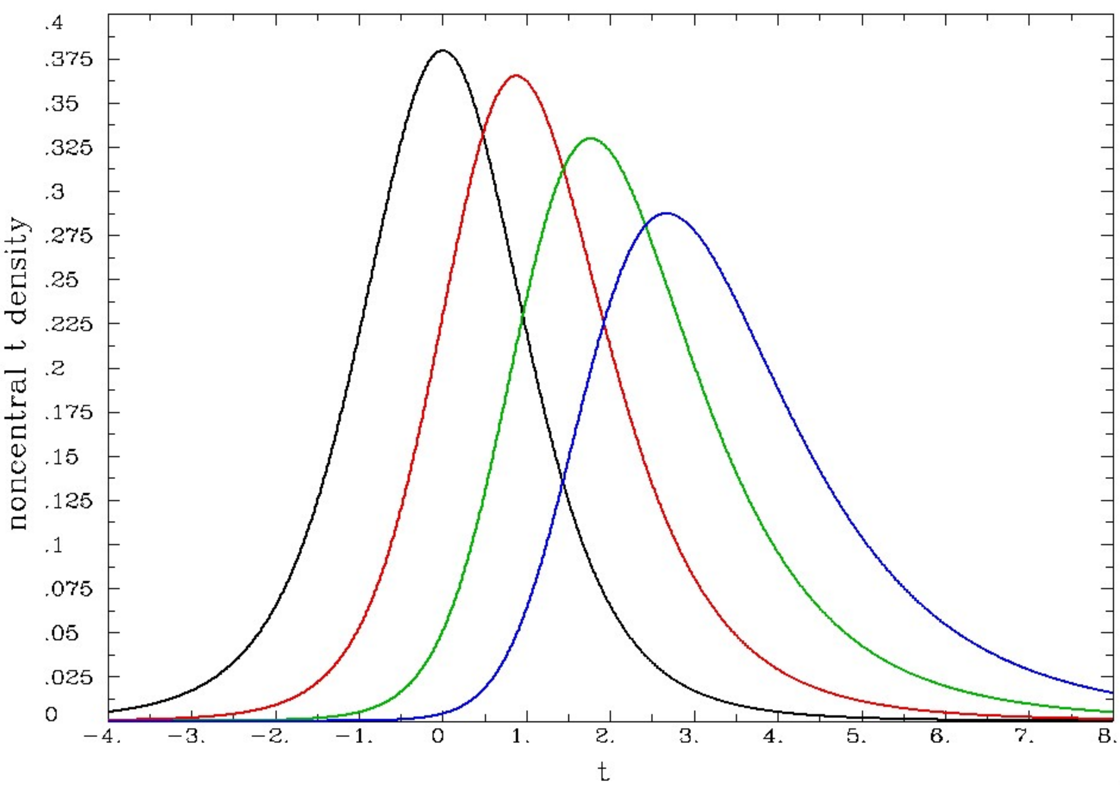
- `nifti_stat2cdf`: given the statistic `val` as input, it returns the cumulative distribution function (cdf)
- `nifti_stat2rcdf`: given the statistic `val` as input, it returns the reversed cdf (1−cdf)
- `nifti_cdf2stat`: given the cdf as input `val`, it returns the corresponding statistic
- `nifti_rcdf2stat`: given 1−cdf as input `val`, it returns the corresponding statistic
- `nifti_stat2zscore`: given the statistic `val` as input, it returns the value of z such that the cdf of $N(0, 1)$ at z has the same value as the cdf of the `code` distribution at `val`; this function is useful for computing z-scores that correspond to "two-sided" tails
- `nifti_stat2hzscore`: given the statistic `val` as input, it returns the value of z such that the cdf of $N(0, 1)$ at z has the value $\frac{1}{2}(1 + \text{cdf})$ for the cdf of the `code` distribution at `val`; in other words, this computes the point corresponding to `val`'s cdf on the half-normal distribution; this function is useful for computing z-scores that correspond to "one-sided" tails

Some attempt has been made to make these functions robust; however, if extreme values are input, inaccurate results may be returned, particularly when very far out in the tails of a distribution.

The file `nifti_stats.c` also contains a sample main program that will read values, codes, and parameters from the command line and print out distributional results from the above functions. This serves as a tool to see how to call these functions, as a way of testing them, and as a way to make graphs (combined with some graphing program). Running the `nifti_stats` program with no command line arguments will result in a brief usage printout.

As an example, the commands below were used to create a plot of the noncentral t density:

```
./nifti_stats -d -4:8:.01 TTEST_NONC 5 0 > t0
./nifti_stats -d -4:8:.01 TTEST_NONC 5 1 > t1
./nifti_stats -d -4:8:.01 TTEST_NONC 5 2 > t2
./nifti_stats -d -4:8:.01 TTEST_NONC 5 3 > t3
1dplot -one -dx .01 -xzero -4 -xlabel 't' -ylabel 'noncentral t density' t?
rm -f t?
```



Each curve has $\nu = 5$; the noncentrality parameter δ increases from 0 to 3, as the curves shift rightwards. Program `1dplot` is included in the AFNI package.