The Distribution of BOLD Susceptibility effects in the Brain is Non-Gaussian.

Stephen José Hanson
Department of Psychology
Rutgers University (Newark)

Benjamin Martin Bly
Department of Psychology
Rutgers University (Newark) &
Department of Radiology
UMDNJ-NJMS

Abstract

A key assumption underlying fMRI analysis in the General Linear Model is that the underlying distribution of BOLD Susceptibility is gaussian. Analysis of several common data sets and experimental paradigms shows that the underlying distribution for the BOLD signal is NON-Gaussian. Further identification shows that the distribution is most likely GAMMA and implications for hemodynamic modeling are discussed as well as recommendations concerning inferential testing in "heavy-tailed" environments.

Introduction

Neuronal activities in the brain are known to engage metabolic processes that alter hemoglobin oxygen consumption. Although the exact mechanism relating neural activity and metabolic oversupply of oxygen is still unexplained, this purported coupling forms the basis for the transient signal that is know as the Blood Oxygenation Level Dependent (BOLD) response. This generic hemodynamic response is at the heart of virtually all neuroimaging involving functional Magnetic Resonance Imaging (fMRI). Since the early 1980s when the BOLD signal was first proposed as a measure of neural activity[1] there has been an explosive growth of statistical methods and analysis procedures, all focused on detecting this signal in the background of noise. In the most common paradigm—the General Linear Model (GLM), various tests have been used (**F, t, r, z** scores) to detect condition-dependent variation in the BOLD response in an assumed background of Gaussian noise. In fact, the assumptions underlying the GLM are much stronger: with random variables $\boldsymbol{Y_{1,\ldots}Y_n}$ independently normally distributed with a common variance, σ^2 , and with means given by $E(Y)=X\beta$. With these assumptions the general linear hypothesis can be expressed as

$$H_0: \boldsymbol{A}b=0 \qquad (1)$$

In the Maximum Likelihood framework, the model being tested for say a simple ANOVA case would be:

$$m_{ij} = m + a_I + b_j + e_{ij}, \quad i=1,\ldots rows, j=1, \ldots columns. \qquad (2)$$

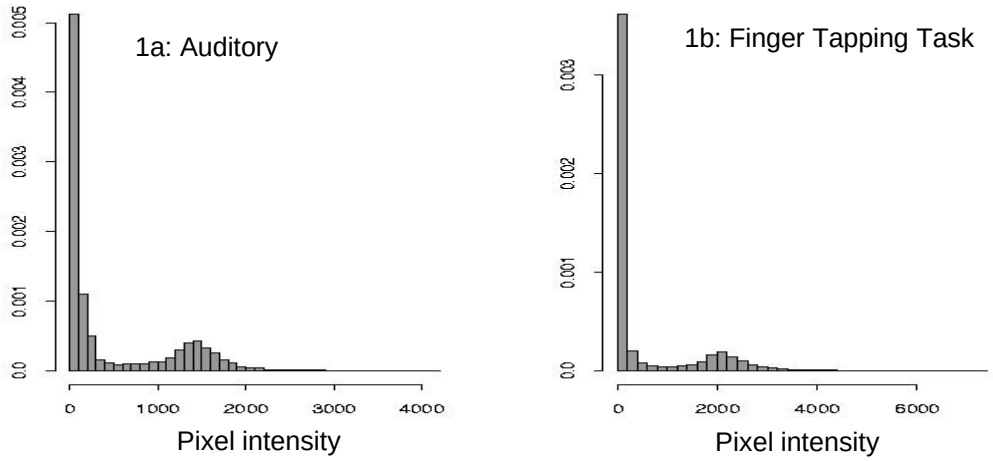
In this case we simply ask for

$$H_0: b_1 = \ldots b_c=0 \qquad (3)$$

Where the error in the model, represented by the ε_{ij} are also assumed to be normally distributed with the initial random variables. Tests of such Gaussian error properties are usually performed with residuals of the model fit to the random variables of interest. In the MLE case, *both* the random variables $\boldsymbol{Y_{1,\ldots}Y_n}$ and the error $\boldsymbol{(e_{ij})}$ must be Gaussian. This appears to be a natural assumption for large data sets and recent experimental work[2] indicates that power estimates can be reasonable even for small samples (<20). Many of the statistical assumptions concerning the estimators and their efficiencies are embodied in software that has been widely used, such as SPM (Statistical Parametric Mapping). However, given the standard GLM assumptions, if the underlying distribution of BOLD susceptibility effects in the brain is *not* Gaussian, then statistics based on this assumption may, at minimum lead to incorrect characterization of ROIs or, in overly conservative hypothesis testing, to higher rates of false negatives. Unfortunately, a natural corrective to, say, *heavier tails* would be to increase the statistical threshold in order to detect the signal of interest or to smooth with larger and larger kernels increasingly flirting with a high rate of false positives.

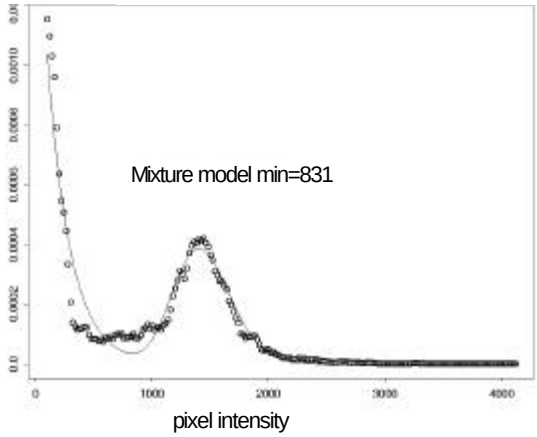
At present, we are unaware of any analysis that has attempted an identification of the underlying distribution of BOLD susceptibility effects in the brain, or shown that the actual distribution is homogeneous, unimodal or even continuous. In this paper we identify the distribution, showing it is non-gaussian, and nonhomogeneous. We further provide a maximum likelihood approach to partitioning an early noise component from the BOLD signal. We use the residue to identify the actual form of the distribution of the BOLD response during standard boxcar presentation of auditory stimuli, response-evoked motor activity and an event-related perceptual task.

Fig. 1: Frequency histograms of BOLD Susceptibility

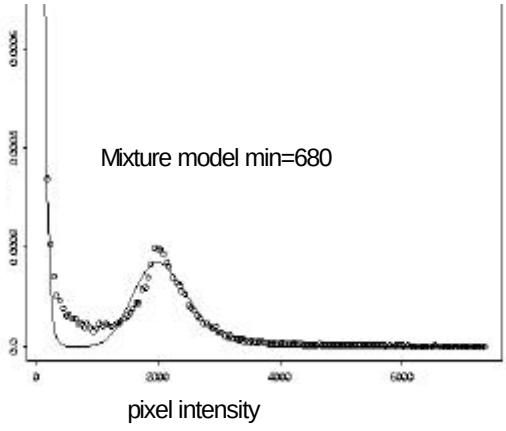


Examples of the whole-image MR signal-intensity distribution for pixels measured using gradient-echo pulse sequences designed to detect BOLD effects is shown in Figures 1a and 1b. The data in Figure 1a were collected from a GE 1.5T (Echospeed Horizon 5.X) MR scanner with a single subject, listening to a 1 kHz tone with a TR of 4s and a TE of 60ms, while the data in Figure 1b were collected from a subject finger tapping at a fixed rate with identical measurement parameters Two other data sets used in this analysis included another auditory monitoring task conducted in a different MRI scanner (SIEMENS 1.5T Vision magnetom, TR 4s, TE 64ms), and a perceptual/memory task involving judgement of event boundaries in a continuous viewing task. Note that the distributions shown are bimodal with a low-intensity first peak near signal-intensity values of 40-50 followed by a rapid decline, leading to a minimum near pixel values of 500-700. A second peak occurs near signal-intensity values of 1500-1600. A potentially serious problem for determining the identity of the density is the inflation of a hypothetical underlying distribution associated with the second peak (where the signal is actually is embedded) in its early tail due to overlap with a hypothetical distribution associated with the first peak. It is possible to “eyeball” a cutoff near an intensity of 600 and simply discard the initial peak. As long as the fall-off of the first distribution is fast enough, the contamination of its tail into the early portions of the second distribution will be insignificant. Although generally not specified, analysis packages such as SPM (Statistical Parametric Mapping) employ such a procedure. A cutoff is determined, either by fixing a value below which no pixel values will be considered, or by determining the minimum between the two peaks and then discarding the residue below the cutoff or the derived minimum. With the now-normative procedures for chosing effects of interest in neuroimaging paradigms, in which pixel intensities well below 90% in the cumulative distribution function are of little interest, (as in SPM), ignoring properties of this whole distribution may be equivalent to less naïve ways of factoring these two distributions, although given what follows, not necessarily more desirable.

Fig. 2a: Mixture model for Auditory Task (GE)



2b: Mixture model to Auditory (SIEMENS)

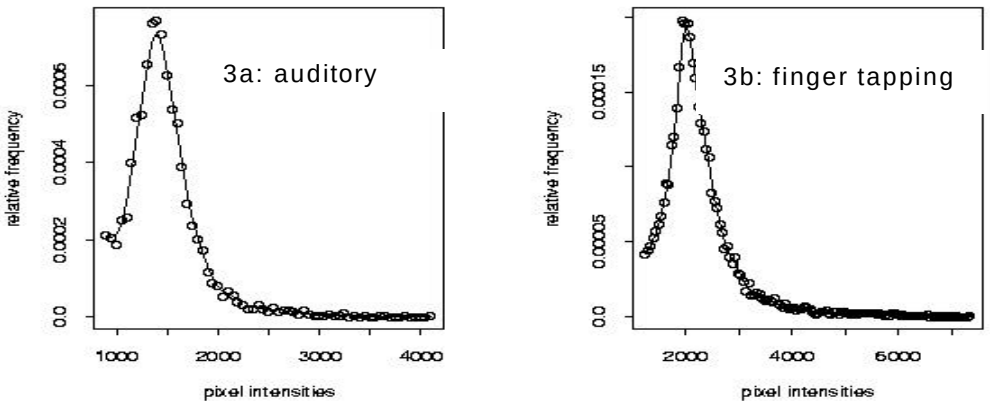


In fact, if we believe the effects of interest could involve early portions of the second distribution, it is probably worth a more informed approach to factoring the two distributions by an independent mixture distribution [$k_1(x) + k_2(x)$], for example, in which the effects of the first kernel can be completely subtracted from the total mixture and hence, assuming a valid estimate, completely controlled. In any case, for our present analysis, it is crucial to approach this problem by use of a mixture analysis in order to produce bi-directional tail estimates for the second distribution that reflect minimal influence by the upper-tail of the first distribution. Quantal-Quantal (QQ) plots are commonly used for distribution identification[3] however, this approach depends critically on the tail properties that are often a "signature" of a particular distribution. Maximum likelihood estimates of best linear fits to the QQ- plots provides a comparative method of examining several common densities and their relation to the recovered distributional response we aim to describe. In order to examine the second distribution it is important, therefore, to take account of possible contamination of its lower tail. We suppose that the observed distribution can be approximated by a mixture of an exponential and normal kernel of the following form:

$$P(X=x)= \alpha\ (e^{-\beta X}) + 1-\alpha\ (e^{-(\mu-X)\hat{}2/\sigma})$$

Irrespective of the actual distribution of the second peak (note that the use of a gaussian kernel in this procedure does *not* imply that the recovered distribution will be gaussian), the mixture distribution will allow a normative estimate of the two bumps and allow us to factor one from the other by simple subtraction. The use of the gaussian allows us to grossly "capture" the second bump while estimating the rate of falloff of the first bump. Also if the second bump is truly Gaussian then the MLE will be optimal. Once the second distribution is recovered QQ-plots can be used to identify it. The mixture model described above was fit with an iterative Maximum Likelihood method to the histogram in Figure 1a for the BOLD response to an auditory stimulus, measured with a GE Echospeed Horizon MR scanner. The resultant maximum likelihood fit is shown in Figure 2a, using the following estimated parameters: α, β, μ, σ. The minimum of this estimated exact function was then found using a minimum search method producing a value of 831 (pixel intensity). Rather than truncate at this minimum, the estimated exponential component was subtracted from the data histogram, leaving the estimated rightmost distribution, uncontaminated by the early peak or upper-tail of the leftmost distribution.
Figure 2b shows another typical mixture fit to the other auditory listening task with data coming from a Siemens Vision magnetom. The minimum in this case was smaller at 680. This method estimates a reduced peak for the leftmost part of the data in both examples, indicating the presence of the lower-tail of the rightmost distribution. The properties of the recovered tail are critical for the next phase of the analysis. The recovered distributions associated with the rightmost peak for both examples are shown in Figure 3a-b, with a nonparametric fit through the points. Notice the skew now apparent in the recovered rightmost distribution.

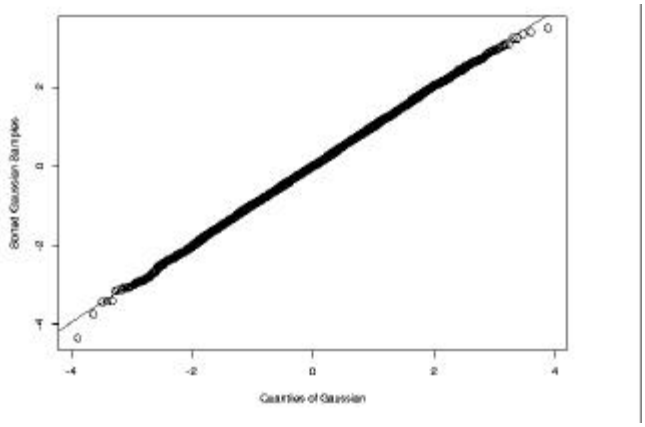
Fig. 3 Nonparametric (Parzen) Fit to recovered BOLD Distributions



In the final phase of the analysis, we apply a QQ analysis to the recovered distribution. This method plots the cumulative density of the distribution against candidate cumulative distributions or random samples in order to best identify the source distribution of the data. Consider the example shown in Figure 4, where 10000 random samples of a Gaussian density are plotted against the standard Gaussian coordinates, similar to a probit plot. If the data are linear in these coordinates, then the data are considered not significantly different from a Gaussian source distribution. Note that the extremes of the distribution are most sensitive and will show small variations even for ground truth distributions. In general Quantal-Quantal (QQ-plots) are used toassess whether data have a particular distribution, or whether two data sets have the same distribution. If the distributions are the same, then the plot will be approximately a straight line. The key sensitivity or signature of the underlying distribution is to be found in the “tails”. These extreme points have more variability than points toward the center. A plot with a "U" shape means that one distribution is skewed relative to the other. An "S" shape implies that one distribution has longer tails than the other. If, in a QQ-plot a data set compared with a Gaussian candidate is either bent down on the left and bent up on the right then this means that the data have heavier tails than the Gaussian. Consider again in Figure 4 where we show the QQ-plot for the Gaussian source against a Gaussian standard which in this case is, not surprisingly, linear.

However, note the extreme points, which even in the case of a known source distribution can deviate from the best linear unbiased estimators in either tail. This is where the greatest

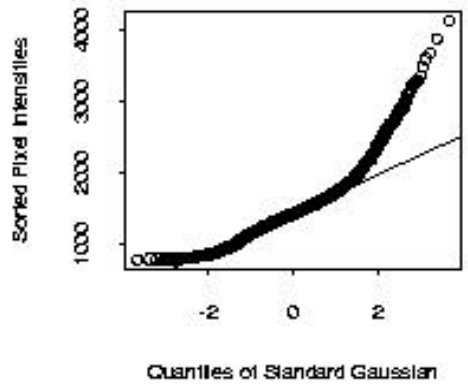
Fig. 4: QQplot for 10k Random Gaussian Samples



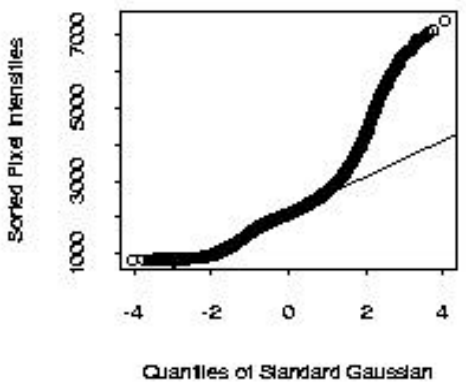
sensitivity of the QQ-plot is useful for identification of the source distribution; excursions from linearity in the tails are in fact diagnostic. If we apply the same analysis to the BOLD susceptibility, we find a significant deviation both at the tails and towards the central tendency of the distribution. Note in Figure 5a, where we have plotted the recovered BOLD values from the mixture model shown in Figure 3. Clearly, the Gaussian standard ignores significant aspects of the tail properties of the bold susceptibility which appear to be both more skewed towards larger intensity values and with slightly heavier tails. The next three plots (Figure 5b-d) confirm this finding in three disparate behavioral contexts including finger tapping, another auditory tone detection and a perceptual/memory task. The auditory detection task also involves a different magnet/vendor than the other three experiments.

Fig. 5: Quantal-Quantal Plots for BOLD vs GAUSSIAN

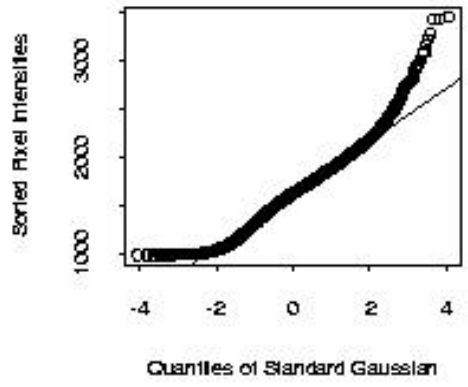
5a: Auditory Listening Task (GE)



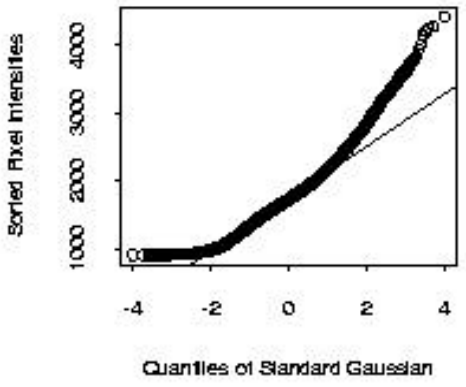
5b: Finger Tapping task



5c: Auditory Listening Task (SIEMENS)



5d: Event Perception task



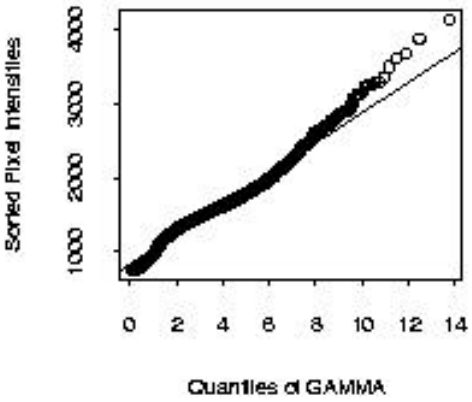
Each case has its own serious individual deviations from Gaussian, some cases possessing tail deviations while others show both tail and center peak excursions from the linear Gaussian baseline. Clearly then, in describing BOLD susceptibility, other distributional possibilities than Gaussian should be considered.

Identification of the Bold Susceptibility Distribution.

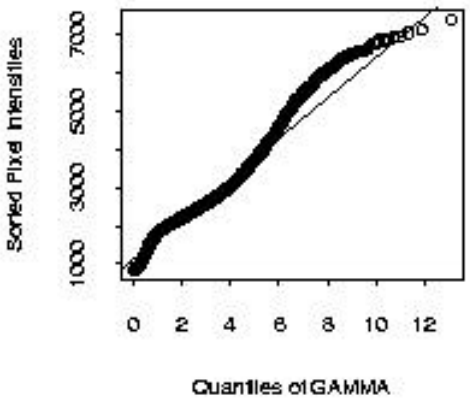
It is possible to submit the BOLD susceptibility data to simple transformations that would lead to linearity or near linearity in the QQ-plots. Three common distributional hypotheses were examined including the Gamma, Log-Normal and Weibel (fixing power exponents at 4 different values). By far the most linearizing case was the Gamma density. Shown in Figure 6a-d are QQ-Plots of the BOLD susceptibility and a Gamma standard. In each case it is clear there is significant linearity and tail deviation well within sampling error (compare for example to the Gaussian cases in Figure 4).

Fig. 6: Quantal-Quantal Plots for BOLD vs GAMMA

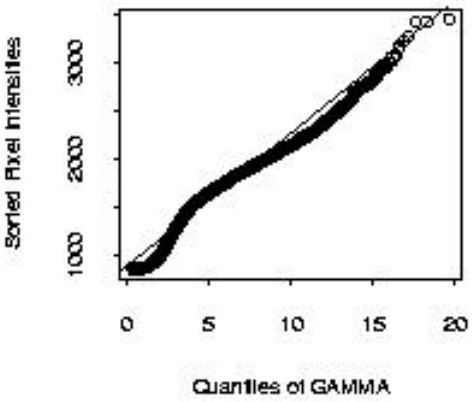
6a: Auditory Task (GE)



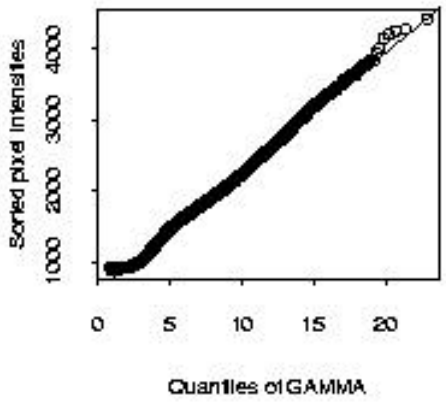
6b: Finger Tapping Task



6c: Auditory Task (SIEMENS)



6d: Event Perception Task



Neither Log-Normal nor Weibel cases proved to be within 15% of the goodness of fit of the Gamma. In Figure 7 we provide a ML estimate of λ and K, the shape and scale parameters of the Gamma for the each case shown as a QQ-plot above. In these cases, we see in the following table that λ, often interpreted as an average rate of an underlying exponential process is near .02 (interpretable as 2% exponential diffusion from any given voxel to another voxel) while K, often

associated with the **number** of underlying exponential processes associated with the stochastic process varies between 20 and 35 putative exponential processes.

Fig. 7a:

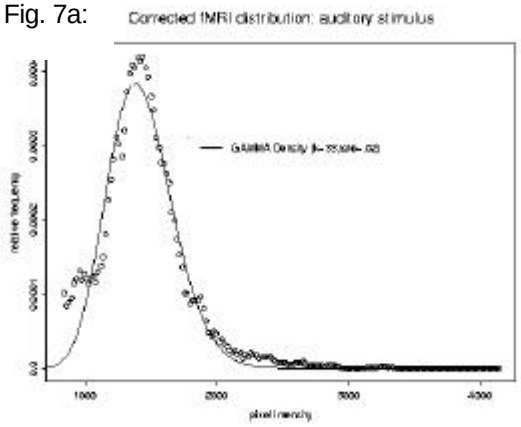


Fig 7b:

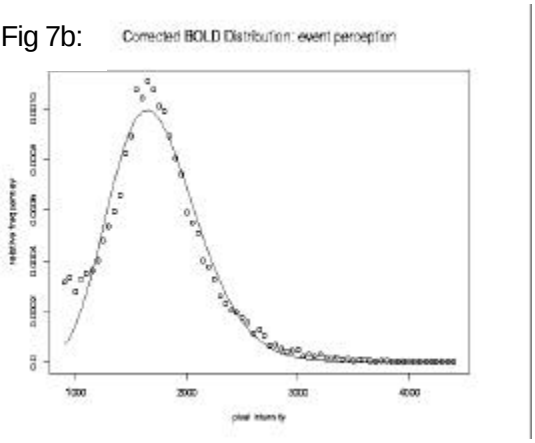


TABLE 1. Parameters for the Gamma Distribution

Parameter/ Data Set	Auditory (GE)	Auditory (SIEMENS)	Finger Tapping	Event Perception
λ	.02	.03	.01	.01
K	30	35	23	19

A Theoretical Rationale for Gamma density.

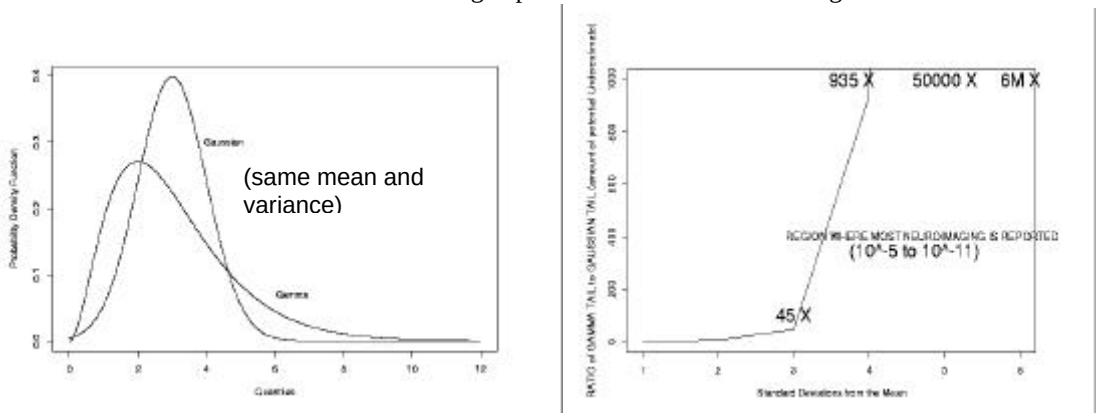
The Gamma density arises under many common stochastic assumptions concerning the sum of K independent exponential processes that possess the same or similar rate parameter. Assuming the hemodynamic response can be compartmentalized across the brain into K independent exponential sinks ("drains") then the distributional outcome in any discrete or continuous model of blood flow would have a strong Gamma signature. This might be so if the distribution of the vasculature serving the cortex could be approximated with a relatively small number of independent sources and sinks[5]. If each voxel corresponds to an average fixed number of sources that are exponentially decaying ("draining") at rate λ then the voxel intensity distribution will tend towards Gamma. The similarity of the distribution across a range of tasks, scanners, individual subjects, and putative numbers of independent sources suggests that the sensitivity of this analysis is rather coarse. Hence, the present identification is likely to be little-affected by the particulars of data collection, but to reflect underlying properties of the cerebrovascular dynamics as measured by BOLD imaging.

Conclusions: Implications for statistical analysis and interpretation BOLD susceptibility

Statistical methods employed in the various data-analysis packages used by researchers in neuroimaging (SPM, AFNI, etc.) generally assume that distributional properties of the BOLD response are Gaussian[4]. Clearly, this assumption violates the present distributional identification, and test-statistics such as **F, t, r or z** scores will have different power in the context of the Gamma environment that is suggested by the present work. One of the ways researchers have adapted to this mismatch between data and model is by using extreme test-statistic threshold values in an effort to avoid Type-I errors. This strategy often makes so-called statistical significance estimates virtually meaningless and has the potential for even more grievous statistical outcomes. For example, the use of unduly extreme significance thresholds can produce high false-negative rates in statistical brain maps. Note in Figure 8a for example, where we show

a Gamma density and a Gaussian density with the same mean and variance, that the Gamma density shows a heavier tail as indicated in our analysis above with actual BOLD data.

Fig 8: GAMMA and GAUSSIAN compared with same mean and variance (left).
Ratio of two distributions showing exponential increase in tail weight.



In Figure 8b, we show the ratio of this Gamma tail to the Gaussian tail which indicates the potential for **underdetection** in the BOLD data. Note, for example, that at 4-6 standard deviations from the mean, the difference in sensitivity is a factor of 50,000 to $6x10^6$. Such a difference could easily account for the difference between conventional significance thresholds (e.g. p<.01) and extreme significance thresholds that are often used in neuroimaging research (typically combined with independence corrections as severe as 10^{-5}, or even 10^{-11}). Note that Figure 8b suggests that in a hypothetical brain volume **at the same ostensible threshold for an appropriate test-statistic,** a Gamma distributional analysis[6] could resemble a standard analysis using a significance threshold as much as two orders of magnitude higher. Analysis methods presently used in the neuroimaging field may therefore lead to serious misinterpretation and a bias towards false identification of focal or modular areas engaged by complex functions that depend on far broader collections of brain regions. Relationships between areas where functional disassociations might be found would also be subject to misinterpretation if one area were merely a false negative in the presence of another area with detected activity. In any case, the gamma or associated heavy tailed environments could form the basis for a rational and more sensitive analysis of neuroimaging data.

References

1. Ogawa, S. and Lee, T.M. "Magnetic resonance imaging of blood vessels at high fields: in vivo and in vitro measurements and image simulation." Magn Reson Med. 16(1): 9-18, 1990.
2. Desmond, J."Statistical Methods: Estimating Sample Size and using Random Effects Models. Talk at Cognitive Neuroscience Society, San Francisco, 2000.
3. Chambers, J. M., Cleveland, W. W., Kleiner, B. and Tukey, P. A. (1983), Graphical Methods or Data Analysis, Wadsworth, Belmont, California.
4. In fact, many researchers have tried to show that due to little autocorrelational structure in the BOLD time series that Gaussian violations would be unlikely. Zarahn, E., Aguirre, G. K. & D'Esposito, M. (1997), Empircal Analysis of BOLD fmri Statistics, Neuroimage, 5, 179-197.
5. R.M Henkelman, Xumei Huang, Qs Xiang, G.J. Stanisz, S. D. Swanson, M.J. Bronskill. (1993) Quantitative interprestation of Magnetization Transfer. Magn. Reson. Med., 29: 759-766.

Alder, R., Feldman, R and Taqqu, M. S.(1998). A Practical Guide to Heavy Tails: Statistical Techniques and Applications.

Acknowledgement: We would like to thank McDonnell Foundation for support of this research and support of the RUMBA laboratories at Rutgers University.

Filename:	BOLDpost.doc
Directory:	C:\WINDOWS\Desktop
Template:	C:\Program Files\Microsoft Office\Templates\Normal.dot
Title:	The Distribution of fMRI Suceptibility effects is GAMMA
Subject:	
Author:	stephen j. hanson
Keywords:	
Comments:	
Creation Date:	05/02/01 7:09 AM
Change Number:	2
Last Saved On:	05/02/01 7:09 AM
Last Saved By:	stephen j. hanson
Total Editing Time:	0 Minutes
Last Printed On:	05/02/01 7:09 AM

As of Last Complete Printing

Number of Pages:	8
Number of Words:	3,623 (approx.)
Number of Characters:	20,656 (approx.)