Maths for Brain Imaging: Lecture 3

W.D. Penny
Wellcome Department of Imaging Neuroscience,
University College, London WC1N 3BG.

October 18, 2006

1 Inference for random fields

A random field is a set of random variables defined at every point in space. To find out if our z-score is 'significant' we need to find out the probability of getting a score that size (or greater) in the abscence of signal. In the absence of signal, we have just error fields. In brain imaging the error fields are spatially correlated and can be described by stochastic processes over space.

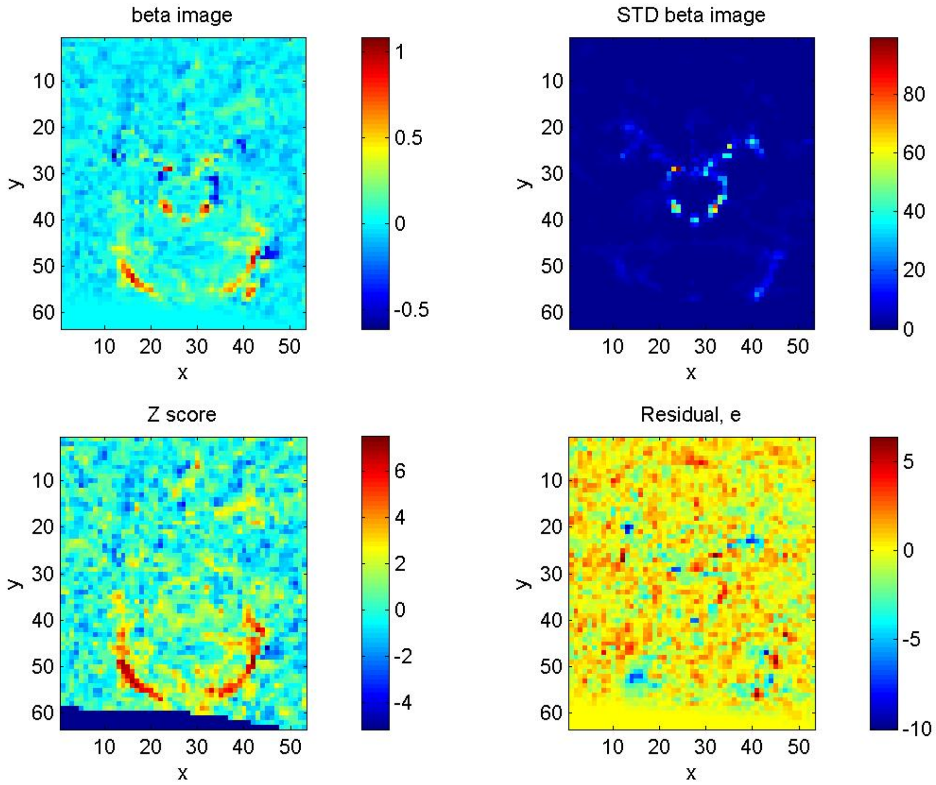


Figure 1: Face data: U1 effect.

1.1 Family Wise Error

We wish to find the probability, under the null hypothesis, that the maximum statistic over the field is larger than some threshold u. That is

$$p(U_{max} > u|H_0) \tag{1}$$

This is the probability of a Family Wise Error (FWE). An FWE is a false positive anywhere in the image.

2 Gaussian processes

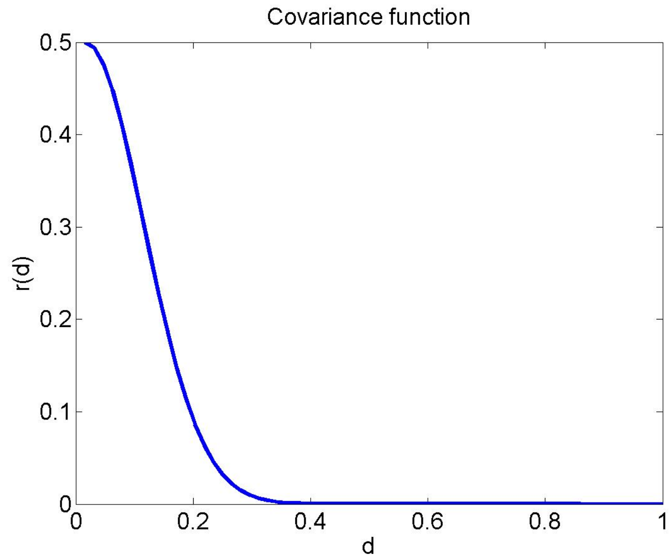
A stochastic process $x(v)$ is a Gaussian process if for any N samples the joint distribution $p(x(v_1), x(v_2), ..., x(v_N))$ is a multivariate Gaussian.

A Gaussian random field/process has a Gaussian distribution at every point and at every collection of points.

Gaussian processes (GPs) are therefore defined by a mean function $m(v)$ and a covariance function

$$r(u, v) = E\left([x(u) - m(u)]^T [x(v) - m(v)]\right) \qquad (2)$$

A Gaussian process is stationary if $r(u, v) = r(u - v)$. We can then write the covariance function as $r(d)$ where $d = r - v$. In what follows we will assume the mean function to be zero at all points. GPs and their properties are then defined solely by their covariance function.



2.1 Example 1

A Gaussian covariance function is given by

$$r(d) = \sigma^2 \exp\left(-\frac{d^2}{2s^2}\right) \tag{3}$$

with power $\sigma^2 = 0.5$ and smoothness $s^2 = 0.1^2$

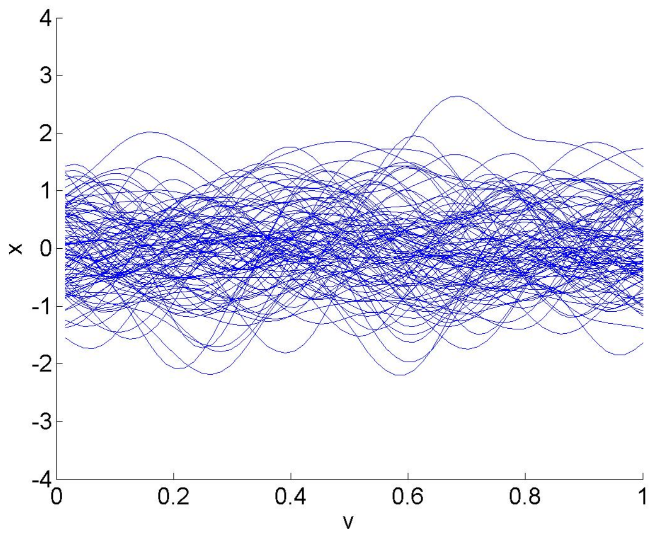


Figure 2: 100 realisations of a Gaussian process with previous Gaussian covariance function. See also NETLAB demo.

2.2 Power and roughness

For stationary processes, the distribution of the max statistic is determined solely by the power and roughness.

The power or variance of a stationary zero-mean Gaussian process is given by the covariance function at lag 0

$$E(|x(v)|^2) = r_x(0) \tag{4}$$

Given any Gaussian process we can create a new one by taking derivatives eg. $y = x'(v) = dx(v)/dv$. Using Fourier methods (see eg. page 325 in [3]) or making use of symmetry properties of the covariance function (see eg. page 314 in [3]) it can be shown that the covariance function of y is given by

$$r_y(v) = -r_x''(v) \tag{5}$$

The power of the stochastic process y is

$$E(|y(v)|^2) = r_y(0) \tag{6}$$

Combining this with the result above shows that variance of the slope is given by

$$E(|\frac{dx(v)}{dv}|^2) = -r_x^{''}(0) \tag{7}$$

The ‘roughness’, λ is then given by the following ratio

$$\lambda^{1/2} = \frac{-r_x^{''}(0)}{r(0)} \tag{8}$$

For unit power fields we have $\lambda^{1/2} = -r_x^{''}(0)$. The ‘smoothness’ is defined as the inverse of the roughness.

For the results that follow, the covariance function can be chosen arbitrarily. However, some results are simplified if the covariance function has a particular form. For example, the covariance function could itself be Gaussian.

2.3 Gaussian covariance function

A Gaussian covariance function is given by

$$r(d) = \sigma^2 \exp\left(-\frac{d^2}{2s^2}\right) \tag{9}$$

At distance 0, $r(0) = \sigma^2$. Spatial derivatives are then given by

$$r'(d) = -\frac{d}{s^2} r(d) \tag{10}$$

Hence, $r'(0) = 0$. The second derivative is given by

$$\begin{aligned} r''(d) &= -\frac{d}{s^2} \times -\frac{d}{s^2} r(d) - \frac{1}{s^2} r(d) \\ &= \left(\frac{d^2 - s^2}{s^4}\right) r(d) \end{aligned} \tag{11}$$

Re-arranging shows that the roughness is given by,

$$\begin{aligned} \lambda^{1/2} &= -\frac{r''(0)}{r(0)} \\ &= \frac{1}{s^2} \end{aligned} \tag{12}$$

The roughness of a GP with a Gaussian CF is therefore $1/s^2$. The smoothness is then the square of the length scale s.

3 Crossings of one-dimensional processes

In a stationary 1-dimensional zero-mean Gaussian field the expected number of crossings, N_c, in the interval $[0, 1]$ of the level u is (page 606, [3])

$$E(N_c) = p_x(u)E(|x'(t)|) \tag{13}$$

That is the density at u multiplied by the expected slope. The density is the usual Gaussian

$$p_x(u) = \frac{1}{(2\pi)^{1/2}\sigma} \exp\left(-\frac{u^2}{2\sigma^2}\right) \tag{14}$$

and it can be shown that

$$E(|x'(t)|)^2 = \frac{-2r^{''}(0)}{\pi} \tag{15}$$

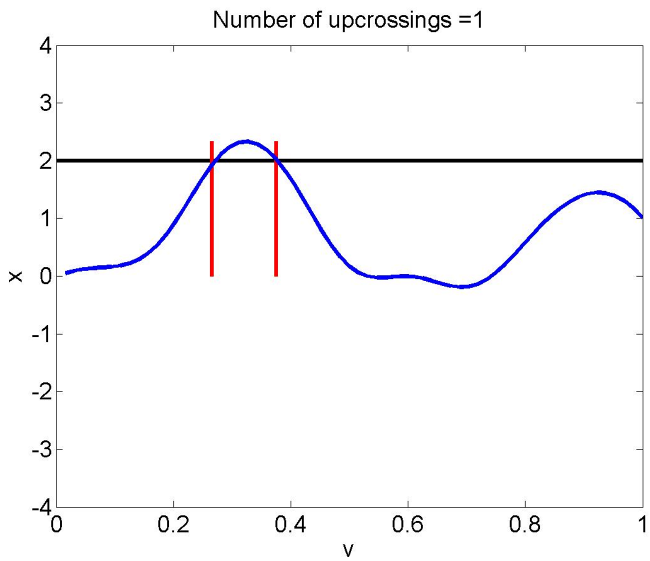


Figure 3: Crossings of a 1D field

The crossing density is therefore

$$E(N_c) = \frac{\lambda^{1/2}}{\pi\sigma} \exp\left(-\frac{u^2}{2\sigma^2}\right) \tag{16}$$

The expected number of upcrossings, N_u, is therefore half that (see also page 67, [1])

$$E(N_u) = \frac{\lambda^{1/2}}{2\pi\sigma} \exp\left(-\frac{u^2}{2\sigma^2}\right) \tag{17}$$

where $\sigma^2 = E(|x(v)|^2) = r(0)$ is the power and $\lambda^{1/2} = -r^{''}(0)$ is the roughness. So, the greater the roughness the more upcrossings we expect. At high thresholds, u, $E(c)$ is the probability that the maximum of the process is larger than u.

4 Multi-dimensional processes

We assume standard Gaussian variates at each location (ie. $\sigma = r(0) = 1$). We first define an *excursion set* as

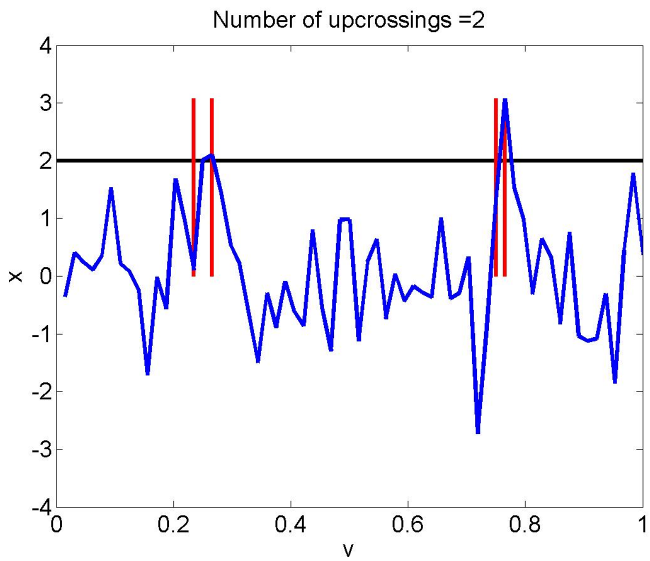


Figure 4: Crossings of a rougher higher power 1D field

the set of voxels where the statistical field exceeds a fixed threshold u.

4.1 Euler characteristic

See eg. [4]. The Euler characteristic, c, counts the number of disconnected components minus the number of 'holes' plus the number of 'hollows'. For high thresholds u the holes and hollows disappear and c counts the number of local maxima.

For large x the Euler characteristic, c, approaches the number of local maxima. Raising the threshold further either the global maxima is above threshold or it is not. So the *expected value* of c is then the probability that the global maximum exceeds the threshold u.

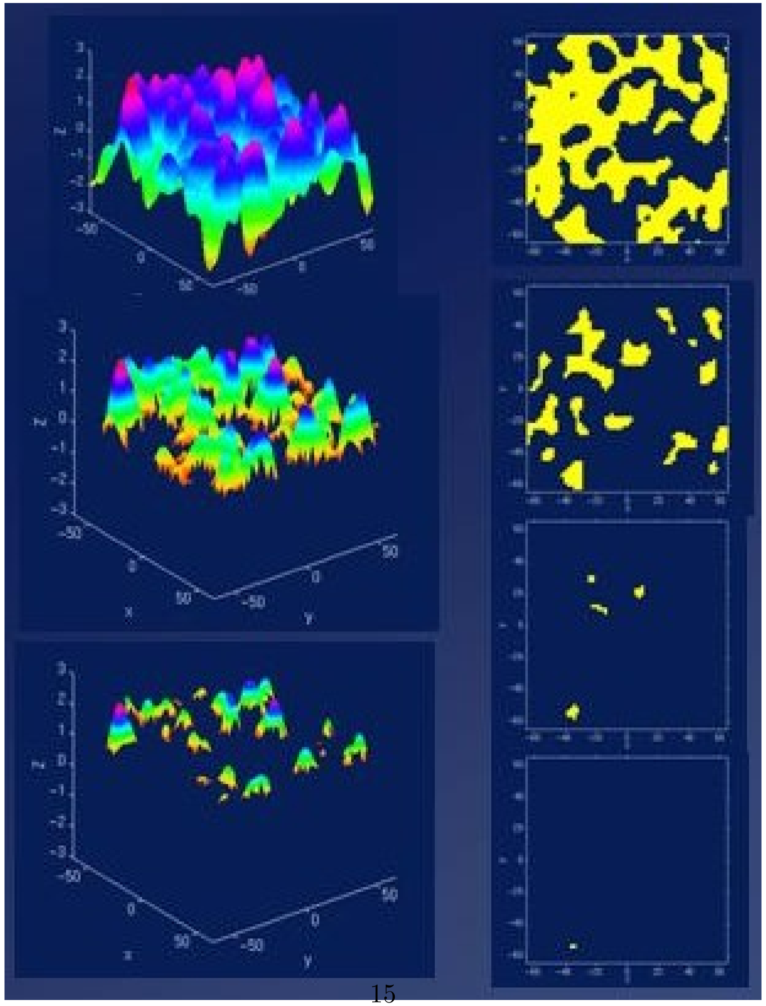


Figure 5: Thresholding a 2D field

4.2 Expected Euler characteristic

The expected value of c for an N-dimensional stationary Gaussian process is given by (page 111 [1])

$$E[c] = V|\Lambda|^{1/2}(2\pi)^{-(N+1)/2}b(N,u)\exp\left(-\frac{u^2}{2}\right) \tag{18}$$

where

$$b(N,u) = \sum_{j=0}^{(N-1)/2}(-1)^j\frac{(2j)!}{j!2^j}u^{N-1-2j} \tag{19}$$

This general result rests on a theorem from differential topology known as Morse's theorem. Results for dimensions $N < 3$ can be derived without this.

For $N = 1$ we have $b(N,u) = 1$

$$E[c] = V\frac{\lambda^{1/2}}{2\pi}\exp\left(-\frac{u^2}{2}\right) \tag{20}$$

which is the same result as earlier for the expected number of upcrossings (assuming $V = 1$, $\sigma = 1$).

For $N = 2$ (eg. brain slice) we have $b(N, u) = u$ and

$$E[c] = V|\Lambda|^{1/2}(2\pi)^{-3/2}u \exp\left(-\frac{u^2}{2}\right) \tag{21}$$

For $N = 3$ (eg. brain volume) we have $b(N, u) = u^2 - 1$ and

$$E[c] = V|\Lambda|^{1/2}(2\pi)^{-2}(u^2 - 1) \exp\left(-\frac{u^2}{2}\right) \tag{22}$$

4.3 Gaussian smoothing

One can create a Gaussian process by convolving IID Gaussian noise with a Gaussian kernel (ie. a Gaussian with covariance matrix Λ^{-1}. For a 3D field, if the principal axes of Λ coincide with the x, y and z directions then the off-diagonal elements of Λ are zero. If f_x, f_y and f_z are the Full Width at Half Maximums (FWHMs) in the x, y and z directions then the roughness is given by [7]

$$|\Lambda|^{1/2} = (f_x f_y f_z)^{-1}(4 \ln 2)^{3/2} \tag{23}$$

If we then define the number of resels as

$$R = \frac{V}{f_x f_y f_z} \tag{24}$$

then for volumetric data we can write (see eg. [])

$$E[c] = R(4\ln 2)^{3/2}(2\pi)^{-2}(u^2 - 1)\exp\left(-\frac{u^2}{2}\right) \tag{25}$$

The above formula only applies to stationary Gaussian fields with Gaussian CFs (these can be created by smoothing IID data with a Gaussian kernel). But because roughness is a property at zero lag, in practice the above formula works well if the covariance function at zero lag is similar to that of a Gaussian CF. It does'nt matter what the tails of the CF look like. So the result can be used for non-Gaussian covariance functions as long as the above holds [6].

4.4 Slice data

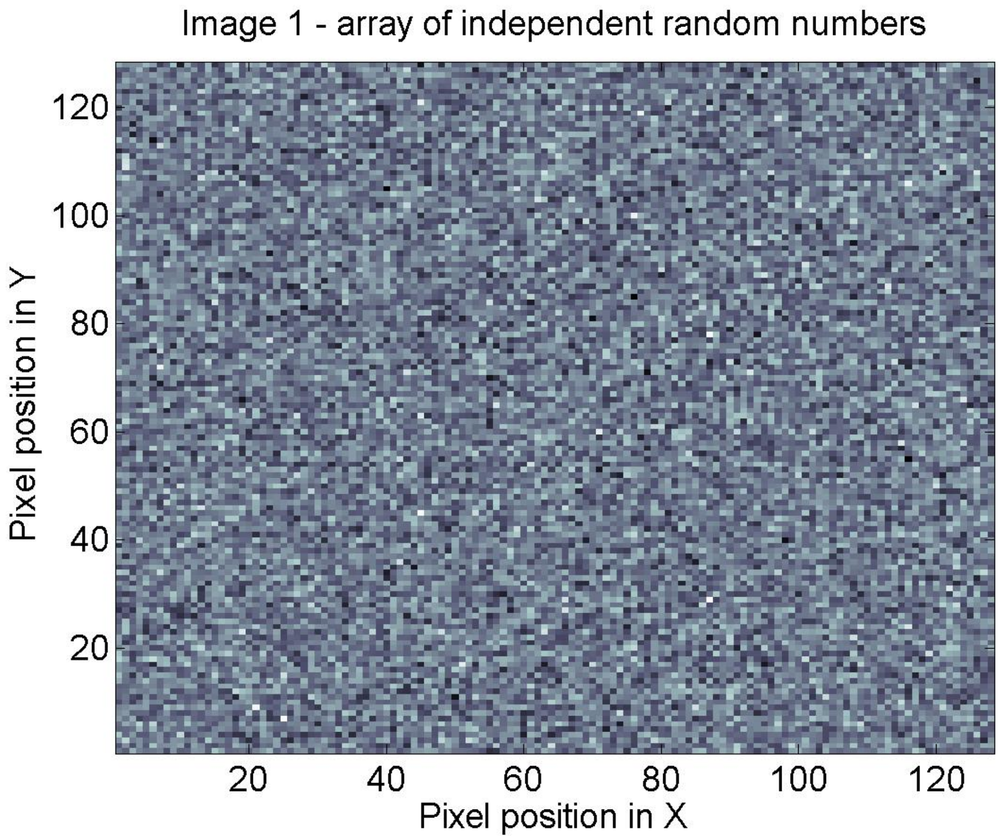


Image 1 - array of independent random numbers
Pixel position in Y
120
100
80
60
40
20
20
40
60
80
100
120
Pixel position in X

Figure 6:

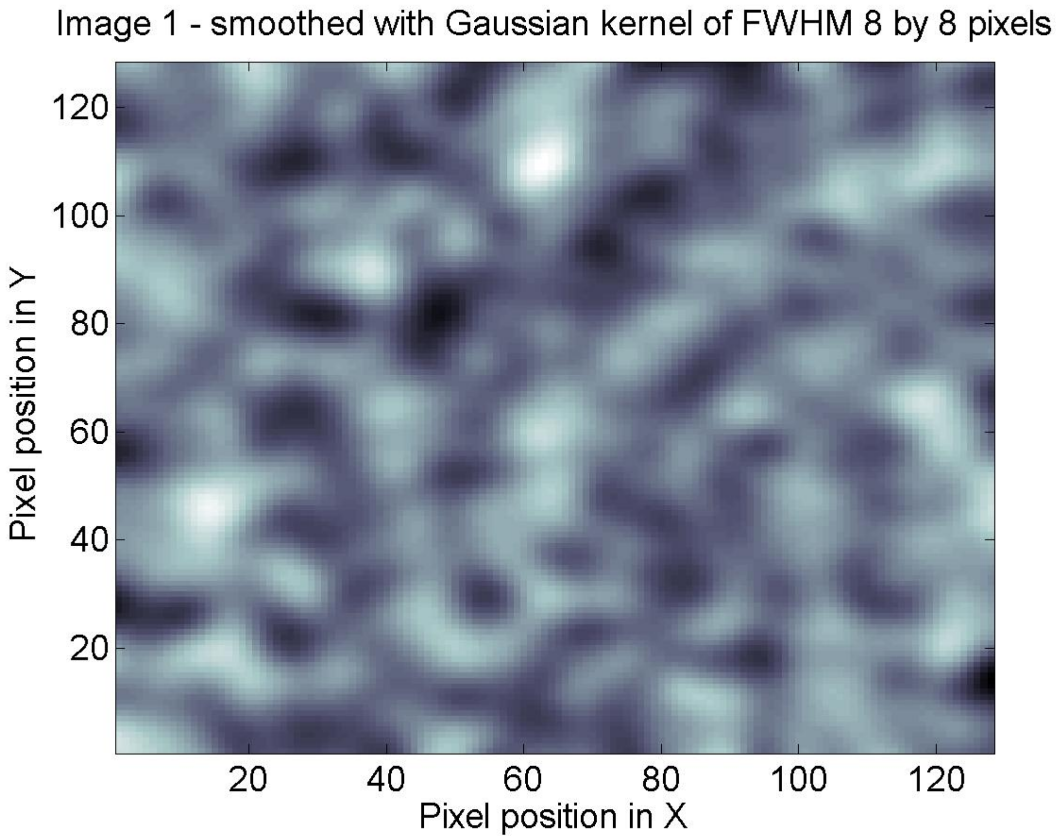


Image 1 - smoothed with Gaussian kernel of FWHM 8 by 8 pixels
120
100
80
60
40
20
Pixel position in Y
20
40
60
80
100
120
Pixel position in X

Figure 7:

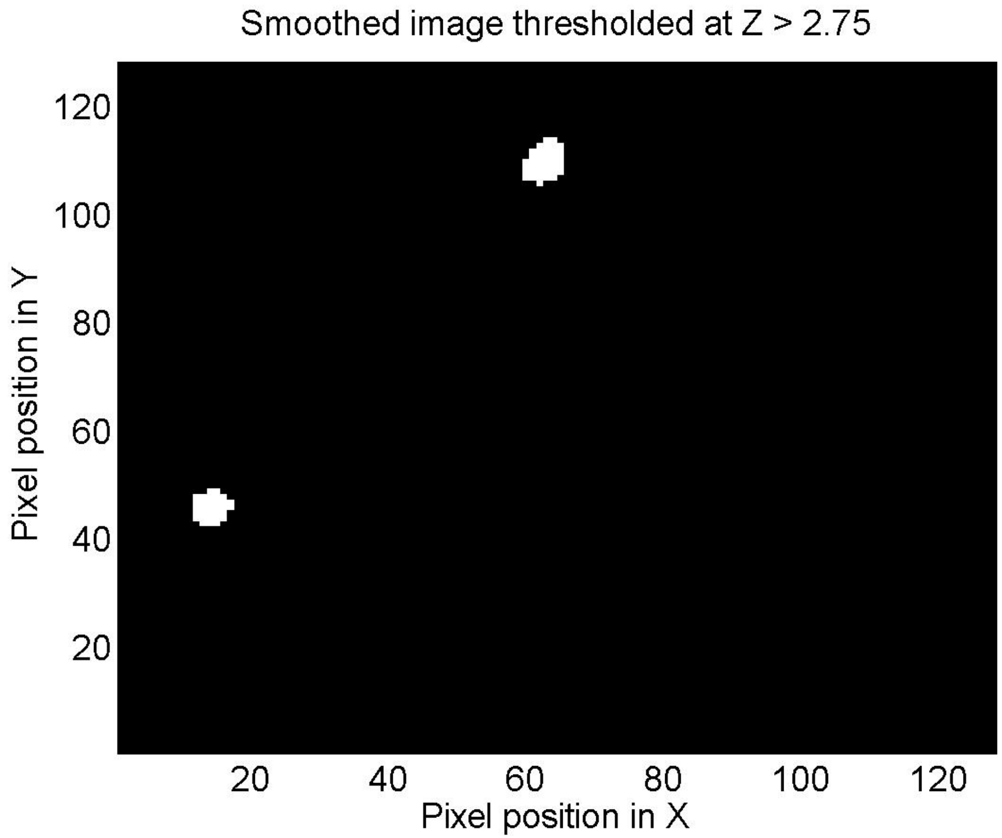

Smoothed image thresholded at Z > 2.75
Pixel position in Y
Pixel position in X
120
100
80
60
40
20
20
40
60
80
100
120

Figure 8:

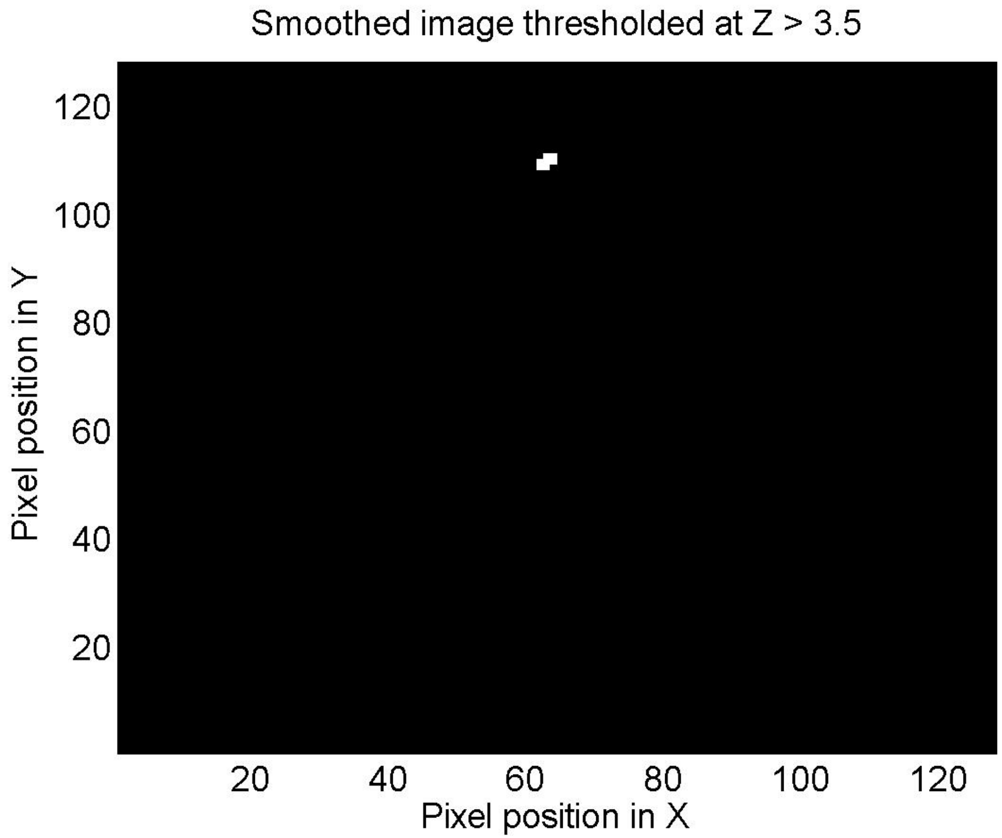

Smoothed image thresholded at Z > 3.5
Pixel position in Y
120
100
80
60
40
20
20
40
60
80
100
120
Pixel position in X

Figure 9:

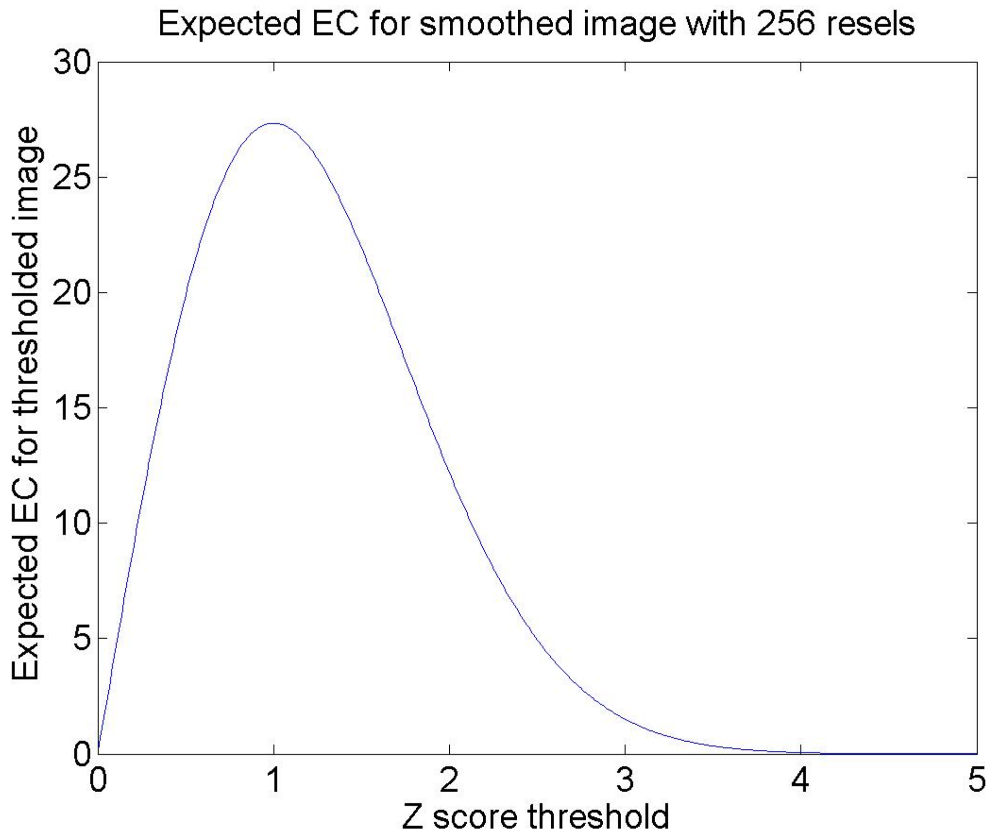

Expected EC for smoothed image with 256 resels
Expected EC for thresholded image
Z score threshold
30
25
20
15
10
5
0
0
1
2
3
4
5

Figure 10:

5 Further issues

5.1 Estimating roughness

Roughness can be estimated using numerical derivatives of the residuals in each of the x, y and z directions. These are then averaged over different residual images (SPM uses 64). These are stored as a Resels Per Voxel (RPV) image, then averaged over voxels [7]. See eg. face data.

5.2 Discretisation

The application of continuous theory to data sampled at discrete points requires that voxel size be eg. 3 times as small as the smoothness of the field. The theory in [7] also requires the search region to be considerably larger than the smoothness (see below).

5.3 Non-Gaussian processes

The results have been extended to t, χ^2 and F random fields [6]. This extension also provides accurate approximations for small search volumes (see Small Volume Correction (SVC) button in SPM). In this work Worsley derives the 'unified formula'

$$E(c) = \sum_{N=1}^{3} R_N(V) p_N(u) \tag{26}$$

where N is the dimension of the field, V is the search volume, $R_N(V)$ is the number of resels in dimension V, and $p_N(V)$ is the EC density for threshold u. The above equation can be solved for u to find the appropriate threshold.

5.4 Inferences about extent

For a given level u, one can work out the probability that the extent of an activation is greater than k. This is known as a cluster-level inference [2].

5.5 Nonstationary fields

The assumption of stationarity is reasonable for PET or smoothed fMRI data. But functional data projected onto unfolded or flattened cortical surfaces or anatomical data such as deformation vectors are highly non-isotropic. Such data can be dealt with by warping voxel coordinates so the effective FWHM is constant [5]. The method has a minor impact on height inferences but a major impact on extent inferences. It is therefore most useful for eg. cluster-level inference for VBM.

References

[1] R.J. Adler. *The geometry of random fields.* John Wiley, New York., 1981.

[2] K.J. Friston, K.J. Worsley R.S.J. Frackowiak, J.C. Mazziotta, and A.C. Evans. Assessing the significance of focal activations using their spatial extent. *Human Brain Mapping*, 1:214–220, 1994.

[3] A. Papoulis. *Probability, Random Variables, and Stochastic Processes.* McGraw-Hill, Singapore, 3 edition, 1991.

[4] K. J. Worsley. The geometry of random images. *Chance*, 9(1):27–40, 1996.

[5] K. J. Worsley, M. Andermann, T. Koulis, D. MacDonald, and A. C. Evans. Detecting changes in nonisotropic images. *Human Brain Mapping*, 8(2):98–101, 1999.

[6] Keith J. Worsley, S. Marrett, P. Neelin, A.C. Vandal, K.J. Friston, and A.C. Evans. A unified statistical approach for determining significant signals in images of cerebral activation. *Human Brain Mapping*, 4:58–73, 1996.

[7] K.J. Worsley, A.C. Evans, S. Marrett, and P. Neelin. A three dimensional statistical analysis for CBF activation studies in human brain. *Journal of Cerebral Blood Flow and Metabolism*, 12:900–918, 1992.