Appendices

A: Co-ordinate Systems

A:1 Referring to voxels by position

***Real co-ordinates, x*∈ Ξ**

Ξ is the subset of $\Re^3$ which is imaged. Ξ is partitioned into K voxels, $\varsigma = \{V_k\}_1^K$ $V_k \subset \Xi$, $k = 1,\ldots,K$; $V_k \cap V_{k'} = \phi$ for $k \neq k'$; and $\bigcup_{k=1}^{K} V_k = \Xi$.

Sometimes it is convenient to refer to a voxel in an image using Cartesian co-ordinates. For an image $\boldsymbol{Y} = (Y_1,\ldots,Y_K)$, abusing the notation somewhat, let $Y_{(\boldsymbol{x})}$ be the value of the voxel containing point $\boldsymbol{x}$:

$$Y_{(\boldsymbol{x})} = \begin{cases} \sum_{k=1}^{K} Y_k \{\boldsymbol{x} \in V_k\} & \text{for } \boldsymbol{x} \in \Xi \\ 0 & \text{for } \boldsymbol{x} \in \overline{\Xi} \end{cases}$$

where $\overline{\Xi} = \Re^3 \setminus \Xi$ the compliment of Ξ in $\Re^3$, and a logical expression in brackets "{•}" takes the value one if the argument is true and zero otherwise, as advocated by Knuth (1992).

Voxel Co-ordinates

If the voxels are identical in shape and size, and are regularly arranged, then the *D*-dimensional image is conveniently stored in a *D*-dimensional array. In this case, voxels are most conveniently referred to by their array indices. Usually the image is stored in an array such that increases in the row, column, and plane indices correspond to increases in the X, Y, and Z directions respectively.

This co-ordinate system is referred to as *voxel co-ordinates*, since the position of each voxel is specified by the displacement in each axial dimension measured in whole numbers of voxels from a given origin. For the standard orientation of the co-ordinate axes in PET, the most left-posterior-lower voxel in the image space is (1,1,1) (fig.92).

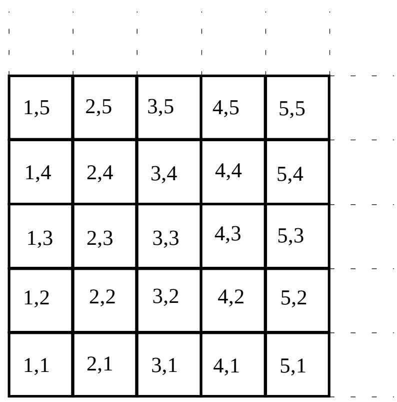

Figure 92
Voxel co-ordinates of voxels (pixels) at the left-lower of a two dimensional image.

Suppose that Ξ is cuboid, oriented parallel to the axes, and partitioned into $K = Xdim \times Ydim \times Zdim$ cuboid voxels. Here $Xdim$ is the width (X-dimension) of Ξ, measured in voxels. If the centre of the left-posterior-lower voxel of Ξ is at (x_o,y_o,z_o), and voxels are of uniform dimensions $h_x \times h_y \times h_z$ then the voxel co-ordinates $\boldsymbol{x}^v = (x^v,y^v,z^v)$ are related to the real co-ordinates $\boldsymbol{x}$ by:

$$\boldsymbol{x} = (x,y,z) = (x^v h_x + x_o,\ y^v h_y + y_o,\ z^v h_z + z_o)$$
$$\quad \text{for } \boldsymbol{x}^v \in [-0.5, Xdim+0.5] \times [-0.5, Ydim+0.5] \times [-0.5, Zdim+0.5]$$
$$\boldsymbol{x}^v = (\ (x-x_o)/h_X,\ (y-y_o)/h_Y,\ (z-z_o)/h_Z\ )$$
$$\quad \text{for } \boldsymbol{x} \in \Xi$$

Denote voxel co-ordinates using brackets thus: $\boldsymbol{Y}_{[\boldsymbol{x}^v]} = \boldsymbol{Y}_{(\boldsymbol{x})}$, where $\boldsymbol{x}$ and $\boldsymbol{x}^v$ are related as above.

A:2 Tri-linear interpolation

Recall that the reconstructed images $\hat{\lambda}$ are estimates of λ, itself a step function approximating the continuous function $\lambda(\boldsymbol{x})$. Thus, an image $\boldsymbol{Y}$ derived from $\hat{\lambda}$ can be regarded as a step function approximating an underlying continuous function $Y(\boldsymbol{x})$. In many cases it is necessary to obtain estimates of $Y(\boldsymbol{x})$ for arbitrary locations, for example when re-sampling an image after a change of co-ordinate axes. In these situations some form of interpolation of the image vector $\boldsymbol{Y}$ is desirable.

If we view the image $\boldsymbol{Y}$ as approximating $Y(\boldsymbol{x})$ at the centres of the voxels, then for locations $\boldsymbol{x}$ other than the voxel centres, an estimate $\tilde{Y}(\boldsymbol{x})$ of $Y(\boldsymbol{x})$ can be obtained by interpolating between the values at neighbouring voxel locations. Tri-linear interpolation is usually employed in PET image analysis. This is the simplest form of interpolation for three-dimensional data. The interpolated value at a given point is a linear combination of the values of the eight neighbouring voxels whose centres define the cuboid containing the point.

Let $\boldsymbol{x}^v = (x^v, y^v, z^v)$ be $\boldsymbol{x}$ in voxel co-ordinates, and let $\boldsymbol{t} = (t_X,\ t_Y,\ t_Z) = \boldsymbol{x}^v - (\lfloor x^v \rfloor, \lfloor y^v \rfloor, \lfloor z^v \rfloor)$ for $\lfloor \bullet \rfloor$ the *floor* function, rounding the argument towards minus infinity. Then, for $\lceil \bullet \rceil$ the *ceil* function, rounding towards plus infinity:

$$
\begin{aligned}
\tilde{Y}(\boldsymbol{x}) = \\
& (1-t_X)(1-t_Y)(1-t_Z)\ \boldsymbol{Y}_{[(\lfloor x^v \rfloor, \lfloor y^v \rfloor, \lfloor z^v \rfloor)]} && + \quad (1-t_X)(1-t_Y)(t_Z)\ \boldsymbol{Y}_{[(\lfloor x^v \rfloor, \lfloor y^v \rfloor, \lceil z^v \rceil)]} \\
+\ & (1-t_X)(t_Y)(1-t_Z)\ \boldsymbol{Y}_{[(\lfloor x^v \rfloor, \lceil y^v \rceil, \lfloor z^v \rfloor)]} && + \quad (1-t_X)(t_Y)(t_Z)\ \boldsymbol{Y}_{[(\lfloor x^v \rfloor, \lceil y^v \rceil, \lceil z^v \rceil)]} \\
+\ & (t_X)(1-t_Y)(1-t_Z)\ \boldsymbol{Y}_{[(\lceil x^v \rceil, \lfloor y^v \rfloor, \lfloor z^v \rfloor)]} && + \quad (t_X)(1-t_Y)(t_Z)\ \boldsymbol{Y}_{[(\lceil x^v \rceil, \lfloor y^v \rfloor, \lceil z^v \rceil)]} \\
+\ & (t_X)(t_Y)(1-t_Z)\ \boldsymbol{Y}_{[(\lceil x^v \rceil, \lceil y^v \rceil, \lfloor z^v \rfloor)]} && + \quad (t_X)(t_Y)(t_Z)\ \boldsymbol{Y}_{[(\lceil x^v \rceil, \lceil y^v \rceil, \lceil z^v \rceil)]}
\end{aligned}
$$

Clearly interpolation of $\boldsymbol{Y}$ introduces some smoothing.

B: Smoothing Convolution

B:1 Smoothing convolution

For a continuous function $Y(\boldsymbol{x})$, $\boldsymbol{x}\in\Re^D$, smoothing is achieved by convolving the function with a filter kernel $\mathrm{f}(\boldsymbol{x})$, to obtain a new function $Y\otimes\mathrm{f}(\boldsymbol{x})$:

$$X\otimes\mathrm{f}(\boldsymbol{x}) = \int\mathrm{f}(\boldsymbol{r})\ Y(\boldsymbol{x}+\boldsymbol{r})\mathrm{d}\boldsymbol{r} \tag{76}$$

Here integration is over the whole range of $\boldsymbol{x}$. The filter kernel, $\mathrm{f}(\boldsymbol{x})$, satisfies:

$$\int\mathrm{f}(\boldsymbol{x})\mathrm{d}\boldsymbol{x} = 1$$

The kernel is a continuous function, usually with a single local maximum at the origin, and with value decreasing as $\boldsymbol{x}$ becomes distant from $\mathbf{0}$.

B:2 Moving average filter

The discrete analogue of convolution is that of a moving average filter. Although technically incorrect, smoothing of images is frequently described as "convolution with a kernel". Suppose our discretisation of the image space $\Xi\subset\Re^D$ is of K voxels $\varsigma=\{\mathrm{V}_k:\ k = 1,\ldots,K\}$, and let $\boldsymbol{x}_k$ be the centre of voxel k. Then for an image of voxel values, $\boldsymbol{Y} = \{Y_1,\ldots,Y_K\}$, the smoothed version of this is $\boldsymbol{Y}^{\mathrm{S}} = \{Y^{\mathrm{S}}_1,\ldots,Y^{\mathrm{S}}_K\}$, given by:

$$Y^{\mathrm{S}}_k = \sum_{k'=1}^{K}\mathrm{f}(\boldsymbol{x}_{k'} - \boldsymbol{x}_k)\ Y_{k'} \tag{77}$$

For a particular voxel, the smoothed image is obtained by positioning the filter kernel on the centre of the voxel, evaluating it on the lattice of points corresponding to the centres of the voxels to obtain the weightings for the voxels, and then summing the weighted voxel values.

Regular discretisation, constant weights, moving average

If the voxels are identical in size and shape, and are regularly arranged, then the set of weights ($\mathrm{f}(\boldsymbol{x}_{k'} - \boldsymbol{x}_k)$) for any voxel k will be identical (ignoring boundary effects). In this case the weights can be computed in advance, giving an image of the filter kernel. The weights can then be explicitly normalised to sum to unity, and the smoothing is a simple moving average.

B:3 Edge effects & boundary truncation smoothing

For voxels k close to the edge of the image space, the filter kernel when located at these voxels will have positive values outside the image space. I.e. $\mathrm{f}(\boldsymbol{x}-\boldsymbol{x}_k)$ is non negative for some $\boldsymbol{x}\in\overline{\Xi} = \Re^D\setminus\Xi$. For such a voxel k, the sum of the weights $\mathrm{f}(\boldsymbol{x}_{k'} - \boldsymbol{x}_k)$ over all voxels k' is less than one. The remaining "weight" of the smoothing kernel corresponds to locations outside the image space for which there are no voxels. The values for edge voxels in the smoothed image are generally reduced towards zero, an *edge effect*. The effect is as if a boundary of voxels with zero value were placed round the edge of the image. This scenario is *zero-boundary smoothing.*

Truncated smoothing

If the image space adequately contains the volume of interest, in our case the voxels corresponding to the brain, then the zero-boundary edge effect is not of any consequence. As can be seen from the raw images (ch.1) the brain fits just inside the image space in the X and Y directions, but is truncated in the Z direction due to the limited axial length of the tomograph.

The effect can be avoided by using *boundary-truncation* smoothing (eqn.78). Here the weights of the filter kernel are normalised at each voxel. The effect is as if the filter kernel is truncated when it reaches the edge of the image.

$$Y_k^S = \frac{1}{\sum_{k'=1}^{K} \mathrm{f}(\boldsymbol{x}_k - \boldsymbol{x}_{k'})} \sum_{k'=1}^{K} \mathrm{f}(\boldsymbol{x}_k - \boldsymbol{x}_{k'})\, Y_{k'} \qquad (78)$$

B:4 Gaussian kernels

The filter kernel used is almost universally Gaussian, by which we mean that it is the probability density function (PDF) of a D-variate normal distribution with zero mean and variance-covariance matrix $\boldsymbol{\Sigma}$ (eqn.79).

$$\mathrm{f}(\boldsymbol{x}) = \frac{1}{(2\pi)^{D/2}\sqrt{|\boldsymbol{\Sigma}|}} \exp\left(-\tfrac{1}{2}\,\boldsymbol{x}\boldsymbol{\Sigma}^{-1}\boldsymbol{x}^{\mathrm{T}}\right) \qquad (79)$$

The filters usually used are orthogonal, with variance-covariance matrices with zero off-diagonal elements. The filter is then completely specified by the D-tuple containing the variances, and is ellipsoidal in shape, with axes parallel to the image axes. The convolution integral factorises into D one dimensional component integrals, which simplifies and speeds up computation.

Relationship of FWHM to variance-covariance matrix

As with image resolution, the shape of the kernel is expressed in FWHM.[56] This is related to the variance for a Gaussian PDF as follows: A univariate Gaussian PDF with variance σ^2 has maxima $1/(\sigma\sqrt{2\pi})$ at $x = 0$. The FWHM l is then the width of the PDF at half this height, $\mathrm{f}(l/2) = 1/(2\sigma\sqrt{2\pi})$, so $l = \sigma\sqrt{8\ln(2)}$. This is extended to D-dimensional orthogonal kernels in the obvious way. If a spherical orthogonal kernel with $\boldsymbol{\Sigma} = \sigma^2\,\mathbf{I}_D$ is used then it is common to just quote $\sigma\sqrt{8\ln(2)}$ as the FWHM. Some authors prefer to specify FWHM in terms of voxels.

Some common filters

Commonly used three-dimensional filters are 10mm×10mm×12mm, and 20mm×20mm×24mm, with variance-covariance matrices of

$$\boldsymbol{\Sigma} = \begin{pmatrix} 10^2 & 0 & 0 \\ 0 & 10^2 & 0 \\ 0 & 0 & 12^2 \end{pmatrix} \frac{1}{8\ln(2)} \text{ and } \boldsymbol{\Sigma} = \begin{pmatrix} 20^2 & 0 & 0 \\ 0 & 20^2 & 0 \\ 0 & 0 & 24^2 \end{pmatrix} \frac{1}{8\ln(2)} \text{ respectively.}$$

[56]Recall that the Full Width at Half Maximum (FWHM) is the width of the (point spread) function at half its maximum.

C: Some Results For Smoothing Convolution

The following results for smoothing convolution of random fields are useful.

C:1 Smoothing convolution: commutative for even kernels

Let $f_1(\boldsymbol{x})$ and $f_2(\boldsymbol{x})$ be any two even functions of $\boldsymbol{x} \in \Re^D$. Then $f_1 \otimes f_2 = f_2 \otimes f_1$, that is, smoothing convolution is commutative for even functions.

Proof:

$$f_1 \otimes f_2(\boldsymbol{u}) = \int_{\Re^D} f_2(\boldsymbol{v})\, f_1(\boldsymbol{u}+\boldsymbol{v}) d\boldsymbol{v}$$

$$= \int_{\Re^D} f_2(\boldsymbol{w}-\boldsymbol{u})\, f_1(\boldsymbol{w}) d\boldsymbol{w} \qquad \text{under change of variables } \boldsymbol{w} = \boldsymbol{u}+\boldsymbol{v} \ (\dagger)$$

$$= \int_{\Re^D} f_2(\boldsymbol{u}-\boldsymbol{w})\, f_1(\boldsymbol{w}) d\boldsymbol{w} \qquad \text{by even property of } f_2$$

$$= \int_{\Re^D} f_2(\boldsymbol{u}+\boldsymbol{v})\, f_1(-\boldsymbol{v}) d\boldsymbol{v} \qquad \text{by change of variables } \boldsymbol{v} = -\boldsymbol{w} \ (\dagger)$$

$$= \int_{\Re^D} f_2(\boldsymbol{u}+\boldsymbol{v})\, f_1(\boldsymbol{v}) d\boldsymbol{v} \qquad \text{by even property of } f_1$$

$$= f_2 \otimes f_1(\boldsymbol{u})$$

So $f_1 \otimes f_2 = f_2 \otimes f_1$, since convolution is commutative for even kernels. The Jacobeans for the changes of variables (†), are $\det(-\mathbf{I}_D)$, the determinant of the negative of the $D \times D$ identity matrix, which has absolute value 1.

C:2 Double smoothing convolution: Associative for even kernels

Let $X(\boldsymbol{x})$ be any function, $\boldsymbol{x} \in \Re^D$, and let $f_1(\boldsymbol{x})$ and $f_2(\boldsymbol{x})$ be two even filter kernels. Then $(X \otimes f_1) \otimes f_2 = X \otimes (f_1 \otimes f_2)$, that is, smoothing convolution is associative.

Proof:

$$(X \otimes f_1) \otimes f_2 = \int_{\Re^D} f_2(\boldsymbol{v})\, X \otimes f_1(\boldsymbol{x}+\boldsymbol{v}) d\boldsymbol{v} = \int_{\Re^D} f_2(\boldsymbol{v}) \int_{\Re^D} f_1(\boldsymbol{u})\, X(\boldsymbol{x}+\boldsymbol{u}+\boldsymbol{v}) d\boldsymbol{u} d\boldsymbol{v}$$

$$= \int_{\Re^D} f_2(\boldsymbol{v}) \int_{\Re^D} f_1(\boldsymbol{w}-\boldsymbol{v})\, X(\boldsymbol{x}+\boldsymbol{w}) d\boldsymbol{w} d\boldsymbol{v} \qquad \text{on change of variables } \boldsymbol{w} = \boldsymbol{u}+\boldsymbol{v}$$

$$= \int_{\Re^D} \int_{\Re^D} f_2(\boldsymbol{v})\, f_1(\boldsymbol{w}-\boldsymbol{v})\, d\boldsymbol{v}\, X(\boldsymbol{x}+\boldsymbol{w}) d\boldsymbol{w}$$

$$= \int_{\Re^D} \int_{\Re^D} f_2(-\boldsymbol{u})\, f_1(\boldsymbol{w}+\boldsymbol{u})\, d\boldsymbol{u}\, X(\boldsymbol{x}+\boldsymbol{w}) d\boldsymbol{w} \qquad \text{on change of variables } \boldsymbol{u} = -\boldsymbol{v}$$

$$= \int_{\Re^D} \int_{\Re^D} f_2(\boldsymbol{u})\, f_1(\boldsymbol{w}+\boldsymbol{u})\, d\boldsymbol{u}\, X(\boldsymbol{x}+\boldsymbol{w})d\boldsymbol{w} \qquad \text{by even property of } f_2$$

$$= \int_{\Re^D} f_1 \otimes f_2(\boldsymbol{w})\, X(\boldsymbol{x}+\boldsymbol{w})d\boldsymbol{w} = X\otimes(f_1\otimes f_2)$$

C:3 Double smoothing convolution: Order unimportant

Let $X(\boldsymbol{x})$ be any function, $\boldsymbol{x}\in\Re^D$, and let $f_1(\boldsymbol{x})$ and $f_2(\boldsymbol{x})$ be two filter kernels. Then $(X\otimes f_1)\otimes f_2 = (X\otimes f_2)\otimes f_1$, that is the order of smoothing is unimportant.

Proof:

For even filter kernels f_1 and f_2 this result follows as a corollary of the previous two results. However, it holds for general f_1 & f_2:

$$(X\otimes f_1)\otimes f_2 = \int_{\Re^D} f_2(\boldsymbol{v})\, X\otimes f_1(\boldsymbol{x}+\boldsymbol{v})d\boldsymbol{v} = \int_{\Re^D} f_2(\boldsymbol{v}) \int_{\Re^D} f_1(\boldsymbol{u})\, X(\boldsymbol{x}+\boldsymbol{u}+\boldsymbol{v})d\boldsymbol{u}d\boldsymbol{v}$$

$$= \int_{\Re^D} f_1(\boldsymbol{u}) \int_{\Re^D} f_2(\boldsymbol{v})\, X(\boldsymbol{x}+\boldsymbol{u}+\boldsymbol{v})d\boldsymbol{v}d\boldsymbol{u} = \int_{\Re^D} f_1(\boldsymbol{u})\, X\otimes f_2(\boldsymbol{x}+\boldsymbol{u})d\boldsymbol{u} = (X\otimes f_2)\otimes f_1$$

C:4 Combining Gaussian kernels: Double smoothing

Let $X(\boldsymbol{x})$ be any function, $\boldsymbol{x}\in\Re^D$, and let $f_1(\boldsymbol{x})$ and $f_2(\boldsymbol{x})$ be two Gaussian filter kernels with variance-covariance matrices $\boldsymbol{\Sigma}_1$ and $\boldsymbol{\Sigma}_2$ respectively. Then $(X\otimes f_1)\otimes f_2 = (X\otimes f_2)\otimes f_1 = X\otimes(f_1\otimes f_2) = X\otimes(f_2\otimes f_1) = X\otimes f$ where $f(\boldsymbol{x})$ is a Gaussian filter kernel with variance-covariance matrix $\boldsymbol{\Sigma}=\boldsymbol{\Sigma}_1+\boldsymbol{\Sigma}_2$.

Proof:

Since the kernels are even functions of $\boldsymbol{x}\in\Re^D$, the associativity and commutativity properties give $(X\otimes f_1)\otimes f_2 = (X\otimes f_2)\otimes f_1 = X\otimes(f_1\otimes f_2) = X\otimes(f_2\otimes f_1)$. It remains to prove that $f_1\otimes f_2$ has the required form. This can be done using Fourier transforms, or directly as follows:

$$f_1\otimes f_2(\boldsymbol{u}) = \frac{1}{(2\pi)^D\sqrt{|\boldsymbol{\Sigma}_1||\boldsymbol{\Sigma}_2|}} \int_{\Re^D} \exp\left(-\tfrac{1}{2}\left[\boldsymbol{v}^{\mathrm{T}}\boldsymbol{\Sigma}_2^{-1}\boldsymbol{v} + (\boldsymbol{u}+\boldsymbol{v})^{\mathrm{T}}\boldsymbol{\Sigma}_1^{-1}(\boldsymbol{u}+\boldsymbol{v})\right]\right)d\boldsymbol{v}$$

$$= \frac{1}{(2\pi)^D\sqrt{|\boldsymbol{\Sigma}_1||\boldsymbol{\Sigma}_2|}} \int_{\Re^D} \exp\left(-\tfrac{1}{2}\left[\boldsymbol{v}^{\mathrm{T}}\left(\boldsymbol{\Sigma}_1^{-1}+\boldsymbol{\Sigma}_2^{-1}\right)\boldsymbol{v} + \boldsymbol{u}^{\mathrm{T}}\boldsymbol{\Sigma}_1^{-1}\boldsymbol{u} + 2\boldsymbol{v}^{\mathrm{T}}\boldsymbol{\Sigma}_1^{-1}\boldsymbol{u}\right]\right)d\boldsymbol{v}$$

using $\boldsymbol{u}^{\mathrm{T}}\boldsymbol{\Sigma}_1^{-1}\boldsymbol{v} = \left(\boldsymbol{u}^{\mathrm{T}}\boldsymbol{\Sigma}_1^{-1}\boldsymbol{v}\right)^{\mathrm{T}} = \boldsymbol{v}^{\mathrm{T}}\boldsymbol{\Sigma}_1^{-1}\boldsymbol{u}$ (since $\boldsymbol{\Sigma}_1^{-1}$ is symmetric)

$$= \frac{1}{(2\pi)^D \sqrt{|\Sigma_1||\Sigma_2|}} \times$$

$$\int_{\Re^D} \exp\left(-\tfrac{1}{2} \begin{bmatrix} \left(\boldsymbol{v} + \left(\Sigma_1^{-1} + \Sigma_2^{-1}\right)^{-1} \Sigma_1^{-1} \boldsymbol{u} \right)^{\mathrm{T}} \left(\Sigma_1^{-1} + \Sigma_2^{-1}\right) \left(\boldsymbol{v} + \left(\Sigma_1^{-1} + \Sigma_2^{-1}\right)^{-1} \Sigma_1^{-1} \boldsymbol{u} \right) \\ - \boldsymbol{u}^{\mathrm{T}} \Sigma_1^{-1} \left(\Sigma_1^{-1} + \Sigma_2^{-1}\right)^{-1} \Sigma_1^{-1} \boldsymbol{u} + \boldsymbol{u}^{\mathrm{T}} \Sigma_1^{-1} \boldsymbol{u} \end{bmatrix} \right) d\boldsymbol{v}$$

$$= \frac{1}{(2\pi)^D \sqrt{|\Sigma_1||\Sigma_2|}} \times$$

$$\int_{\Re^D} \exp\left(-\tfrac{1}{2} \left[\left(\boldsymbol{v} + \left(\Sigma_1^{-1} + \Sigma_2^{-1}\right)^{-1} \Sigma_1^{-1} \boldsymbol{u} \right)^{\mathrm{T}} \left(\Sigma_2^{-1} (\Sigma_2 + \Sigma_1) \Sigma_1^{-1} \right) \left(\boldsymbol{v} + \left(\Sigma_1^{-1} + \Sigma_2^{-1}\right)^{-1} \Sigma_1^{-1} \boldsymbol{u} \right) \right] \right) d\boldsymbol{v}$$

$$\times \exp\left(-\tfrac{1}{2} \left[\boldsymbol{u}^{\mathrm{T}} \Sigma_1^{-1} \left(\Sigma_1^{-1} + \Sigma_2^{-1}\right)^{-1} \left(\Sigma_1^{-1} + \Sigma_2^{-1}\right) \boldsymbol{u} - \boldsymbol{u}^{\mathrm{T}} \Sigma_1^{-1} \left(\Sigma_1^{-1} + \Sigma_2^{-1}\right)^{-1} \Sigma_1^{-1} \boldsymbol{u} \right] \right)$$

$$= \frac{1}{(2\pi)^D \sqrt{|\Sigma_1||\Sigma_2|}} \times (2\pi)^{D/2} \sqrt{\left| \left(\Sigma_2^{-1} (\Sigma_2 + \Sigma_1) \Sigma_1^{-1} \right)^{-1} \right|}$$

$$\times \exp\left(-\tfrac{1}{2} \left[\boldsymbol{u}^{\mathrm{T}} \Sigma_1^{-1} \left(\Sigma_1^{-1} + \Sigma_2^{-1}\right)^{-1} \Sigma_2^{-1} \boldsymbol{u} \right] \right)$$

$$= \frac{1}{(2\pi)^{D/2} \sqrt{|\Sigma_1||\Sigma_2|}} \times \sqrt{\left| \Sigma_1 (\Sigma_2 + \Sigma_1)^{-1} \Sigma_2 \right|} \times \exp\left(-\tfrac{1}{2} \left[\boldsymbol{u}^{\mathrm{T}} \left(\Sigma_2 \left(\Sigma_1^{-1} + \Sigma_2^{-1}\right) \Sigma_1 \right)^{-1} \boldsymbol{u} \right] \right)$$

$$= \frac{1}{(2\pi)^{D/2} \sqrt{|\Sigma_1||\Sigma_2|}} \times \sqrt{\left| (\Sigma_2 + \Sigma_1)^{-1} \right|} \sqrt{|\Sigma_1||\Sigma_2|} \times \exp\left(-\tfrac{1}{2} \left[\boldsymbol{u}^{\mathrm{T}} (\Sigma_2 + \Sigma_1)^{-1} \boldsymbol{u} \right] \right)$$

$$= \frac{1}{(2\pi)^{D/2} \sqrt{|\Sigma_1 + \Sigma_2|}} \exp\left(-\tfrac{1}{2} \left[\boldsymbol{u}^{\mathrm{T}} (\Sigma_1 + \Sigma_2)^{-1} \boldsymbol{u} \right] \right)$$

C:5 Covariance function of smoothed white noise processes

Let $Z(\boldsymbol{x})$, $\boldsymbol{x}\in\Re^D$, be the field formed by convolving a weakly stationary continuous white noise random field[57] $X(\boldsymbol{x})$ of variance σ^2 ($\mathrm{Var}[X(\boldsymbol{x})] = \sigma^2\ \forall\boldsymbol{x}$) with a kernel $\mathrm{f}(\boldsymbol{x})$, $Z(\boldsymbol{x}) = X\otimes\mathrm{f}(\boldsymbol{x})$. Clearly $Z(\boldsymbol{x})$ is a strictly stationary continuous random field. Moreover, the covariance function is $\mathrm{C}(\boldsymbol{h}) = \sigma^2\ \mathrm{f}\otimes\mathrm{f}(\boldsymbol{h})$

Proof:
This result is easily proved directly from the convolution integral (eqn.76):

$$\begin{aligned}
\mathrm{C}(\boldsymbol{h}) &= \mathrm{Cov}[Z(\boldsymbol{x}),Z(\boldsymbol{x}+\boldsymbol{h})]\\
&= \mathrm{Cov}\left[\int_{\Re^D}\mathrm{f}(\boldsymbol{r}')X(\boldsymbol{x}+\boldsymbol{r}')d\boldsymbol{r}', \int_{\Re^D}\mathrm{f}(\boldsymbol{r})X(\boldsymbol{x}+\boldsymbol{h}+\boldsymbol{r})d\boldsymbol{r}\right]\\
&= \int_{\Re^D}\mathrm{f}(\boldsymbol{r})\,\mathrm{Cov}\left[\int_{\Re^D}\mathrm{f}(\boldsymbol{r}')X(\boldsymbol{x}+\boldsymbol{r}')d\boldsymbol{r}', X(\boldsymbol{x}+\boldsymbol{h}+\boldsymbol{r})\right]d\boldsymbol{r}\\
&= \int_{\Re^D}\mathrm{f}(\boldsymbol{r})\int_{\Re^D}\mathrm{f}(\boldsymbol{r}')\,\mathrm{Cov}\left[X(\boldsymbol{x}+\boldsymbol{r}'), X(\boldsymbol{x}+\boldsymbol{h}+\boldsymbol{r})\right]d\boldsymbol{r}'d\boldsymbol{r}\\
&= \int_{\Re^D}\mathrm{f}(\boldsymbol{r})\,\mathrm{f}(\boldsymbol{r}+\boldsymbol{h})\sigma^2 d\boldsymbol{r}
\end{aligned}$$

(since $\mathrm{Cov}[X(\boldsymbol{x}+\boldsymbol{r}'),X(\boldsymbol{x}+\boldsymbol{h}+\boldsymbol{r})]=\sigma^2$ if $\boldsymbol{r}'=\boldsymbol{h}+\boldsymbol{r}$ and is zero otherwise)

$$= \sigma^2\ \mathrm{f}\otimes\mathrm{f}(\boldsymbol{h})$$

Corollary:

The field $Z(\boldsymbol{x})$ formed by convolving a white noise field of variance σ^2 with a Gaussian kernel $\mathrm{f}(\boldsymbol{x})$ with variance-covariance matrix $\boldsymbol{\Sigma}$, $\mathrm{f}(\boldsymbol{x}) = \exp(-\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\boldsymbol{x}/2) / \sqrt{(2\pi)^D\,|\boldsymbol{\Sigma}|}$, is a strictly stationary continuous field with covariance function:

$$\begin{aligned}
\mathrm{C}(\boldsymbol{h}) &= \sigma^2\ \mathrm{f}\otimes\mathrm{f}(\boldsymbol{h})\\
&= \sigma^2 \exp(-\boldsymbol{h}^{\mathrm{T}}(2\boldsymbol{\Sigma})^{-1}\boldsymbol{h}/2) / \sqrt{(2\pi)^D\,|2\boldsymbol{\Sigma}|}
\end{aligned}$$

(a Gaussian kernel with variance-covariance $2\boldsymbol{\Sigma}$, by result 4)

$$= \frac{\sigma^2}{2^D\pi^{D/2}\sqrt{|\boldsymbol{\Sigma}|}}\exp(-\boldsymbol{h}^{\mathrm{T}}(\boldsymbol{\Sigma})^{-1}\boldsymbol{h}/4)$$

So, for $\boldsymbol{h} = \boldsymbol{0}$, $\mathrm{Var}[Z(\boldsymbol{x})] = \dfrac{\sigma^2}{2^D\pi^{D/2}\sqrt{|\boldsymbol{\Sigma}|}}$

Corollary:

A strictly stationary continuous random field with zero mean, variance σ^2, and auto-correlation function $\mathrm{R}(\boldsymbol{h}) = \exp(-\boldsymbol{h}^{\mathrm{T}}(\boldsymbol{\Sigma})^{-1}\boldsymbol{h}/4)$, can be obtained by convolving a white noise random field of variance $\sigma^2 2^D\pi^{D/2}\sqrt{|\boldsymbol{\Sigma}|}$ with a Gaussian kernel of variance-covariance matrix $\boldsymbol{\Sigma}$, $\mathrm{f}(\boldsymbol{x}) = \exp(-\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\boldsymbol{x}/2) / \sqrt{(2\pi)^D\,|\boldsymbol{\Sigma}|}$.

---

[57]A random process $X(t)$ is white noise if $\mathrm{E}[X(t)] = 0$, and if $X(t_1)$ and $X(t_2)$ are independent, for all points t, t_1 & t_2 in the parameter space.

C:6 Smoothness of smoothed Gaussian white noise fields

The field obtained by convolving a continuous white noise Gaussian random field (defined on $\Re^D$) with a kernel $f(\boldsymbol{x})$, is itself a strictly stationary continuous Gaussian random field with zero mean. If the variance of the white noise process is chosen such that the resulting field has unit variance, then Adler (1981) shows that the variance-covariance matrix of partial derivatives is:

$$\boldsymbol{\Lambda} = \frac{\displaystyle\int_{\Re^D} \frac{\partial f(\boldsymbol{x})}{\partial \boldsymbol{x}} \frac{\partial f(\boldsymbol{x})}{\partial \boldsymbol{x}^{\mathrm{T}}} d\boldsymbol{x}}{\displaystyle\int_{\Re^D} f^2(\boldsymbol{x}) d\boldsymbol{x}}$$

C:7 Smoothness of (Gaussian) smoothed Gaussian white noise

The strictly stationary continuous standard Gaussian (zero mean, unit variance) random field formed by convolving a white noise Gaussian random field with a Gaussian kernel, $f(\boldsymbol{x}) = \exp(-\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\boldsymbol{x}/2) / \sqrt{(2\pi)^D |\boldsymbol{\Sigma}|}$, has variance-covariance matrix of partial derivatives $\boldsymbol{\Lambda} = \boldsymbol{\Sigma}^{-1}/2 = (2\boldsymbol{\Sigma})^{-1}$. By result 5, the variance of the white noise field must be $2^D\pi^{D/2}\sqrt{|\boldsymbol{\Sigma}|}$ for the smoothed field to have unit variance.

Proof:

This follows from the previous result by direct integration.

Corollary:

A strictly stationary continuous Gaussian random field with zero mean, unit variance, variance-covariance matrix of partial derivatives $\boldsymbol{\Lambda}$, and Gaussian auto-correlation function can be obtained by convolving a white noise Gaussian random field of variance $(2\pi)^{D/2} / \sqrt{|\boldsymbol{\Lambda}|}$ with a Gaussian kernel with variance-covariance matrix $\boldsymbol{\Sigma} = (2\boldsymbol{\Lambda})^{-1}$. This observation provides the framework for simulating Gaussian random fields.

C:8 Secondary smoothing

Consider a strictly stationary continuous standard Gaussian (zero mean, unit variance) random field $Y(\boldsymbol{x})$, $\boldsymbol{x} \in \Re^D$, with Gaussian auto-correlation function and variance-covariance matrix of partial derivatives $\boldsymbol{\Lambda}_Y$. Let $f(\boldsymbol{x})$ be a Gaussian kernel with variance-covariance matrix $\boldsymbol{\Sigma}$. Let $Z = c^{-1/2} \times Y \otimes f$, where the constant c is chosen such that $\mathrm{Var}[Z(\boldsymbol{x})] = 1\ \forall \boldsymbol{x}$. Then Z is also a strictly stationary standard Gaussian random field, with variance-covariance matrix of partial derivatives $\boldsymbol{\Lambda}_Z = (2\boldsymbol{\Sigma} + \boldsymbol{\Lambda}_Y^{-1})^{-1}$. Furthermore, $c = 1/\sqrt{\left|2\boldsymbol{\Lambda}_Y\boldsymbol{\Sigma} + \mathbf{I}_D\right|}$.

If $\boldsymbol{\Sigma}_Y = (2\boldsymbol{\Lambda}_Y)^{-1}$, then $\boldsymbol{\Lambda}_Z = (2\boldsymbol{\Sigma} + 2\boldsymbol{\Sigma}_Y)^{-1}$, $c = \dfrac{\sqrt{\left|\boldsymbol{\Sigma}_Y\right|}}{\sqrt{\left|\boldsymbol{\Sigma} + \boldsymbol{\Sigma}_Y\right|}}$, and $\boldsymbol{\Sigma}_Z = (2\boldsymbol{\Lambda}_Z)^{-1}$

$= \boldsymbol{\Sigma} + \boldsymbol{\Sigma}_Y$

Proof:

Since $Y(\boldsymbol{x})$ can be generated by convolving a white noise Gaussian random field $X(\boldsymbol{x})$, of variance $(2\pi)^{D/2} / \sqrt{|\Lambda_Y|}$, with Gaussian kernel $f_Y(\boldsymbol{x})$ with variance-covariance matrix $\Sigma_Y = (2\Lambda_Y)^{-1}$ (corollary to result 7). Result 4 then gives that $\sqrt{c}\times Z(\boldsymbol{x}) = Y\otimes f$ is equivalent to a field obtained by convolving a white noise random field, of variance $(2\pi)^{D/2} / \sqrt{|\Lambda_Y|}$, with a Gaussian kernel $f_Z = f_Y\otimes f$, with variance-covariance matrix $\Sigma + \Sigma_Y$. Then, by result 7, $Z(\boldsymbol{x})$ has variance-covariance matrix of partial derivatives $\Lambda_Z = (2\Sigma + 2\Sigma_Y)^{-1} = (2\Sigma + \Lambda_Y{}^{-1})^{-1}$. It remains to identify the constant c. Since $Z(\boldsymbol{x})$ is of unit variance, $\sqrt{c}\times Z(\boldsymbol{x})$ has variance c, but, regarding $\sqrt{c}\times Z(\boldsymbol{x})$ as $X\otimes(f_Y\otimes f)$, result 5 gives its variance as:

$$c = \frac{\sqrt{|\Sigma_Y|}}{\sqrt{|\Sigma + \Sigma_Y|}} = \frac{\sqrt{|(2\Lambda_Y)^{-1}|}}{\sqrt{|\Sigma + (2\Lambda_Y)^{-1}|}} = \frac{1}{\sqrt{|2\Lambda_Y||\Sigma + (2\Lambda_Y)^{-1}|}} = \frac{1}{\sqrt{|2\Lambda_Y\Sigma + I_D|}}$$

C:9 Effect of scaling on smoothness

Let $Z(\boldsymbol{x})$, $\boldsymbol{x}\in\Re^D$, be a strictly stationary continuous random field with variance-covariance matrix of partial derivatives Λ_Z. Then the field $Y(\boldsymbol{x}) = c\times Y(\boldsymbol{x})$, for constant c, has variance-covariance matrix of partial derivatives $\Lambda_Y = c^2\,\Lambda_Z$.

Proof:

Trivial. Using the definition of Λ_Z, and the chain rule for differentiation…

$$\Lambda = \begin{pmatrix} \operatorname{var}\left[\partial Y/\partial x_1\right] & \operatorname{cov}\left[\partial Y/\partial x_1, \partial Y/\partial x_2\right] & \cdots \\ \operatorname{cov}\left[\partial Y/\partial x_1, \partial Y/\partial x_2\right] & \operatorname{var}\left[\partial Y/\partial x_2\right] & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix} \text{Where } \boldsymbol{x} = (x_1, x_2, \ldots).$$

$$= \begin{pmatrix} \operatorname{var}\left[\frac{\partial Y}{\partial X}\frac{\partial X}{\partial x_1}\right] & \operatorname{cov}\left[\frac{\partial Y}{\partial X}\frac{\partial X}{\partial x_1}, \frac{\partial Y}{\partial X}\frac{\partial X}{\partial x_2}\right] & \cdots \\ \operatorname{cov}\left[\frac{\partial Y}{\partial X}\frac{\partial X}{\partial x_1}, \frac{\partial Y}{\partial X}\frac{\partial X}{\partial x_2}\right] & \operatorname{var}\left[\frac{\partial Y}{\partial X}\frac{\partial X}{\partial x_2}\right] & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix} \ldots$$

$$= \left(\partial Y/\partial X\right)^2 \begin{pmatrix} \operatorname{var}\left[\partial X/\partial x_1\right] & \operatorname{cov}\left[\partial Y/\partial x_1, \partial X/\partial x_2\right] & \cdots \\ \operatorname{cov}\left[\partial X/\partial x_1, \partial X/\partial x_2\right] & \operatorname{var}\left[\partial X/\partial x_2\right] & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix} = c^2\,\Lambda$$

D: Expected Euler Characteristics

D:1 The χ^2-field

The expected Euler characteristic $\chi(A_u(U,\Psi))$ of the excursion set $A_u(U,\Psi)$ of a homogeneous (strictly stationary) Chi-squared random field $U(\boldsymbol{x})$ with n degrees of freedom, defined on $\boldsymbol{x}\in\Psi$, a compact, convex subset of $\Re^D$ (with boundary of zero Lebesgue measure), for a threshold u is, for $D \geq 2$ and under mild conditions on the component fields:

$$\mathrm{E}[\chi(A_u(U,\Psi))] = \frac{\lambda(\Psi)|\Lambda|^{1/2} u^{\frac{n-D}{2}} \exp\left(-\frac{u}{2}\right)}{(2\pi)^{\frac{D}{2}} 2^{\frac{n-1}{2}} \Gamma\left(\frac{n}{2}\right)} \mathrm{P}_{D,n}(u)$$

where $\mathrm{P}_{D,n}(u)$ is a polynomial of degree D -1 in u with integer coefficients:

$$\mathrm{P}_{D,n}(u) = \sum_{j=0}^{\frac{D-1}{2}} \sum_{k=0}^{D-1-2j} \binom{n-1}{D-1-2j-k} \frac{(-1)^{D-1+j+k}(D-1)!}{2^j\, j!\, k!} u^{j+k}$$

Terms with factorials of negatives in the denominator are taken as zero.
(Worsley, 1994, Theorem 3.5)

D:2 The *F*-field

For $F(\boldsymbol{x})$ an F-field with n,m degrees of freedom, the expected Euler characteristic of the excursion set (over Ψ, a compact, convex subset of $\Re^D$, with boundary of zero Lebesgue measure) above a threshold u is, for $m + n > D \geq 2$:

$$\mathrm{E}[\chi(A_u(F,\Psi))] = \frac{\lambda(\Psi)|\Lambda|^{1/2} \Gamma\left(\frac{1}{2}(m+n-D)\right)}{(2\pi)^{D/2} 2^{\frac{D-2}{2}} \Gamma(m/2)\Gamma(n/2)} \left(\frac{nu}{m}\right)^{\frac{n-D}{2}} \left(1+\frac{nu}{m}\right)^{-\frac{m+n-2}{2}} \mathrm{K}_{D,m,n}(u)$$

where $\mathrm{K}_{D,m,n}(u)$ is a polynomial of degree D -1 in nu/m with integer coefficients:

$$\mathrm{K}_{D,m,n}(u) = (-1)^{D-1}(D-1)! \sum_{j=0}^{\frac{D-1}{2}} \frac{\Gamma\left(\frac{1}{2}(m+n+D)+j\right)}{\Gamma\left(\frac{1}{2}(m+n-D)\right) j!} \times \sum_{k=0}^{D-1-2j} \binom{m-1}{k} \binom{n-1}{D-1-2j-k} (-1)^{j+k} \left(\frac{nu}{m}\right)^{j+k}$$

(Worsley, 1994, Theorem 4.6)

D:3 The *t*-field

For $T(\boldsymbol{x})$ a t-field with n degrees of freedom, the expected Euler characteristic of the excursion set (over Ψ, a compact, convex subset of $\Re^D$, with boundary of zero Lebesgue measure) above a threshold u is, for $n \geq D \geq 2$:

$$\mathrm{E}[\chi(A_u(T,\Psi))] = \frac{\lambda(\Psi)|\Lambda|^{1/2}}{(2\pi)^{\frac{D+1}{2}}}\left(1+\frac{u^2}{n}\right)^{-\frac{n-1}{2}} Q_{D,n}(u)$$

where $\mathrm{Q}_{D,n}(u)$ is a polynomial of degree D -1 in u:

$$\mathrm{Q}_{D,n}(u) = \sum_{j=0}^{\frac{D-1}{2}} \frac{(-1)^j (D-1)!}{2^j\, j!(D-1-2j)!} \frac{\Gamma\left(\frac{1}{2}(n+1)\right)}{\Gamma\left(\frac{1}{2}(n+2-D+2j)\right)\left(\frac{1}{2}n\right)^{\frac{D-1-2j}{2}}} u^{D-1-2j}$$

(Worsley, 1994, Theorem 5.3)

E: "Transform" Functions

Suppose t is drawn from a distribution with Cumulative Distribution Function $F(x)$, then an equivalent standard Gaussian variate z is one with equal extremum probability:

$$\Phi(z) = F(t) \quad \Leftrightarrow \quad z = \Phi^{-1}(F(t))$$

(The distribution function method for functions of random variables.) Since the normal distribution is continuous, $\Phi(z)$ is strictly monotonic increasing. Thus Φ^{-1} exists and a unique z is specified. $\Phi^{-1}(F(\bullet))$ is thus a function "transforming" a random variable from one distribution to a standard Gaussian distribution, and has become known (in PET) as a *transform function*.

Transform function for Students' t-distribution

The computing environments used by most PET centres to analyse images do not have built in statistical distribution functions (PDFs, CDFs, inverse CDFs), so they must be explicitly coded. Since the evaluation of these functions is becoming a lost art in these days of comprehensive tables and sophisticated statistics packages, we review the computations for Students t-distribution.

For the t-distribution with df degrees of freedom, the Cumulative Distribution Function $F_T(\bullet)$ can be expressed in terms of the incomplete Beta function, $I_x(a,b)$, as follows:

$$F_T(t) = \frac{1}{\sqrt{df}\,\beta\left(df/2, 1/2\right)} \int_{-\infty}^{t} \left(1 + \frac{u^2}{dx}\right)^{-\frac{df+1}{2}} = \begin{cases} 1 - \frac{1}{2} \times I_{\frac{df}{df+t^2}}\left(\frac{df}{2}, \frac{1}{2}\right) & t \geq 0 \\ \frac{1}{2} \times I_{\frac{df}{df+t^2}}\left(\frac{df}{2}, \frac{1}{2}\right) & t \leq 0 \end{cases}$$

Here $\beta(a,b)$ is Beta function:

$$\beta(a,b) = \frac{\Gamma(a)\,\Gamma(b)}{\Gamma(a+b)} = \int_0^1 u^{a-1}\,(1-u)^{b-1}\,\mathrm{d}u \qquad a,b > 0$$

and the incomplete Beta function is:

$$I_x(a,b) = \frac{\beta_x(a,b)}{\beta(a,b)} = \frac{1}{\beta(a,b)} \int_0^x u^{a-1}\,(1-u)^{b-1}\,\mathrm{d}u \qquad a,b > 0$$

As is well known, the CDF of the normal distribution is related to the error function:

$$\Phi(x) = \frac{1}{2} + \frac{1}{2}\,\mathrm{erf}\left(\frac{x}{\sqrt{2}}\right) \qquad \mathrm{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-u^2}\,\mathrm{d}u$$

$$\Leftrightarrow \quad \Phi^{-1}(p) = \sqrt{2}\;\mathrm{erf}^{-1}(2p - 1) \qquad 0 \leq p \leq 1$$

The incomplete Beta function and the inverse of the error function are supplied in many engineering and imaging packages. Alternatively, various published solutions for their approximation exist. Thus the "transform" function for Students t-distribution can be painlessly coded. See "Numerical Recipes" (Press *et. al.*) for algorithms and relationships for other distributions.

F: Ordering Theorem

Theorem

Consider two sets of real numbers: $X=\{x_i\}_{i=1}^{n}$, $Y=\{y_j\}_{j=1}^{n}$, such that $x_i \le y_i$ $\forall$ $i = 1,\dots,n$. Order X and Y from largest to smallest, with ties broken arbitrarily, giving ordered sets $X=\{x_{(i)}\}_{i=1}^{n}$, $Y=\{y_{(i)}\}_{i=1}^{n}$, where $x_{(i)} \ge x_{(j)}$ & $y_{(i)} \ge y_{(j)}$ $\forall$ $1 \le i \le j \le n$.

Then, $x_{(i)} \le y_{(i)}$ $\forall$ $i = 1,\dots,n$.

Proof:

Suppose $x_{(k)} = x_{i_k}$, $y_{(k)} = y_{j_k}$

and $S_{X,k} = \{i_l : l = 1,\dots,k\}$, the k largest x_i, and
$S_{Y,k} = \{j_l : l = 1,\dots,k\}$, the k largest y_j.

$k = 1$

$y_{(1)} = y_{j_1}$	$\ge y_{i_1}$	by definition of maximum
	$\ge x_{i_1}$	by hypothesis
	$= x_{(1)}$	by definition of x_{i_1}

$k > 1$

If $i_k \notin S_{Y,k\text{-}1}$ then $y_{(k)}$	$\ge y_{j_k}$	$(i_k \notin S_{Y,k\text{-}1} \Rightarrow y_{j_k}$ not in k-1 largest $y)$
	$\ge x_{i_k}$	by hypothesis
	$= x_{(k)}$	by definition of x_{i_k}

otherwise, $i_k \in S_{Y,k\text{-}1}$, and $\exists$ i' such that $i' \in S_{X,k\text{-}1}$ but $i' \notin S_{Y,k\text{-}1}$ †

then $y_{(k)}$	$\ge y_{j'}$	because $i' \notin S_{Y,k\text{-}1}$
	$\ge x_{i'}$	by hypothesis
	$\ge x_{(k)}$	because $i' \in S_{X,k\text{-}1}$

† $S_{X,k\text{-}1}$ has k-1 members, but i_k is not one of them, by definition. However, $S_{Y,k\text{-}1}$ has k-1 members, one of which is i_k. Therefore $S_{X,k\text{-}1}$ must contain an element not contained in $S_{Y,k\text{-}1}$.

G: Smoothness of *t*-Fields

Let $X_1(\boldsymbol{x}),\ldots,X_n(\boldsymbol{x})$ $\boldsymbol{x}\in\Psi\subset\Re^D$ be independent, identically distributed, strictly stationary Gaussian random fields with zero mean and variance σ^2. Suppose that the variance-covariance matrix of partial derivatives of the field is $\boldsymbol{\Lambda}$.

Consider the *t*-field $T(\boldsymbol{x})$ with n-1 degrees of freedom formed as the one-sample *t*-statistic of $\{X_1(\boldsymbol{x}),\ldots,X_n(\boldsymbol{x})\}$ at each point $\boldsymbol{x}\in\Psi$:

$$T(\boldsymbol{x}) = \frac{M(\boldsymbol{x})}{S(\boldsymbol{x})/\sqrt{n}}$$

$$\text{where}\quad M(\boldsymbol{x}) = \frac{1}{n}\sum_{i=1}^{n} X_i(\boldsymbol{x})$$

$$\text{and}\quad S(\boldsymbol{x})^2 = \frac{1}{n-1}\sum_{i=1}^{n} (X_i(\boldsymbol{x}) - M(\boldsymbol{x}))^2$$

Let $\boldsymbol{\Lambda}_T$ be the variance-covariance matrix of partial derivatives of the *t*-field. This can be related to $\boldsymbol{\Lambda}$ using a simplification of the argument used in the appendix of Worsley *et al*. (1992). (With thanks to Dr. Worsley for pointing this out.) The argument is as follows:

Since $X_1(\boldsymbol{x}),\ldots,X_n(\boldsymbol{x})$ and their partial derivatives are all independent Gaussian random variables with zero expectations (Adler, 1981, p.31), conditioning on $X_1(\boldsymbol{x}),\ldots,X_n(\boldsymbol{x})$, we have

$$\begin{aligned}
\boldsymbol{\Lambda}_T = \mathrm{Var}\left(\frac{\partial T}{\partial \boldsymbol{x}}\right) &= \mathrm{Var}\left(\sum_{i=1}^{n}\left(\frac{\partial T}{\partial X_i}\right)\left(\frac{\partial X_i}{\partial \boldsymbol{x}}\right)\right) \\
&= \mathrm{E}\left[\sum_{i=1}^{n}\left(\frac{\partial T}{\partial X_i}\right)^2 \mathrm{Var}\left(\frac{\partial X_i}{\partial \boldsymbol{x}}\right)\right] \\
&= \mathrm{E}\left[\sum_{i=1}^{n}\left(\frac{\partial T}{\partial X_i}\right)^2\right]\boldsymbol{\Lambda}\sigma^2
\end{aligned}$$

From the definition of $T(\boldsymbol{x})$ we have

$$\frac{\partial T}{\partial X_i} = \frac{\partial}{\partial X_i}\frac{M}{S / n^{1/2}} = \frac{1}{n^{1/2}\, S} - \frac{n^{1/2}\, M(X_i - M)}{(n-1)\, S^3}$$

and so

$$\sum_{i=1}^{n}\left(\frac{\partial T}{\partial X_i}\right) = \frac{1}{S^2} + \frac{nM^2}{(n-1)\, S^4}$$

Hence

$$\Lambda_T = \mathrm{E}\left[\frac{1}{S^2} + \frac{nM^2}{(n-1)\, S^4}\right]\Lambda\sigma^2$$

$$= \lambda_n \Lambda \quad \text{say.}$$

This can be further simplified using the fact that $U = [nM^2 + (n-1)S^2]/\sigma^2$ has a χ^2 distribution with n degrees of freedom, independent of T, and that $\mathrm{E}[1/U^2] = 1/(n\text{-}2)$, giving

$$\lambda_n = \mathrm{E}\left[\frac{(T^2 + n - 1)^2}{(n-1)\,(n-2)}\right]$$

Integrating over the density of T gives:

$$\lambda_n = \int_{-\infty}^{+\infty} \frac{(t^2 + n - 1)^2}{(n-1)\,(n-2)}\, \mathrm{f}_T(t)\, \mathrm{d}t \;=\; 2\int_{0}^{+\infty} \frac{(t^2 + n - 1)^2}{(n-1)\,(n-2)}\, \mathrm{f}_T(t)\, \mathrm{d}t$$

where $\mathrm{f}_T(\bullet)$ is the PDF of a Student's t-distribution with n-1 degrees of freedom. The integral is finite only for $n \geq 4$ ($df \geq 3$).

Values of λ_n for $n = 4,\ldots,199$ are given to 4dp in the following table, computed using an adaptive recursive Newton-Cotes eight panel rule. Note that λ_n tends to 1 from above as n tends to infinity.

λ_n	+0	+1	+2	+3	+4	+5	+6	+7	+8	+9
0	–	∞	∞	∞	20.9819	4.3698	2.6638	2.0859	1.8015	1.6342
10	1.5243	1.4467	1.3891	1.3446	1.3092	1.2804	1.2565	1.2363	1.2191	1.2042
20	1.1912	1.1798	1.1696	1.1605	1.1524	1.1450	1.1383	1.1323	1.1267	1.1215
30	1.1168	1.1124	1.1083	1.1045	1.1010	1.0977	1.0946	1.0918	1.0890	1.0865
40	1.0840	1.0818	1.0796	1.0775	1.0756	1.0737	1.0720	1.0703	1.0687	1.0671
50	1.0657	1.0643	1.0629	1.0616	1.0604	1.0592	1.0580	1.0569	1.0559	1.0549
60	1.0539	1.0529	1.0520	1.0511	1.0503	1.0494	1.0486	1.0479	1.0471	1.0464
70	1.0457	1.0450	1.0443	1.0437	1.0431	1.0424	1.0419	1.0413	1.0407	1.0402
80	1.0396	1.0391	1.0386	1.0381	1.0376	1.0372	1.0367	1.0363	1.0358	1.0354
90	1.0350	1.0346	1.0342	1.0338	1.0335	1.0331	1.0327	1.0324	1.0320	1.0317
100	1.0314	1.0310	1.0307	1.0304	1.0301	1.0298	1.0295	1.0292	1.0289	1.0287
110	1.0284	1.0281	1.0279	1.0276	1.0274	1.0271	1.0269	1.0266	1.0264	1.0262
120	1.0259	1.0257	1.0255	1.0253	1.0251	1.0249	1.0247	1.0245	1.0243	1.0241
130	1.0239	1.0237	1.0235	1.0233	1.0231	1.0230	1.0228	1.0226	1.0224	1.0223
140	1.0221	1.0219	1.0218	1.0216	1.0215	1.0213	1.0212	1.0210	1.0209	1.0207
150	1.0206	1.0205	1.0203	1.0202	1.0200	1.0199	1.0198	1.0196	1.0195	1.0194
160	1.0193	1.0191	1.0190	1.0189	1.0188	1.0187	1.0186	1.0184	1.0183	1.0182
170	1.0181	1.0180	1.0179	1.0178	1.0177	1.0176	1.0175	1.0174	1.0173	1.0172
180	1.0171	1.0170	1.0169	1.0168	1.0167	1.0166	1.0165	1.0164	1.0163	1.0162
190	1.0162	1.0161	1.0160	1.0159	1.0158	1.0157	1.0157	1.0156	1.0155	1.0154

H: Poline's Bivariate Approach

Poline & Mazoyer (1994a) address the problem of intense focal activations being missed by suprathreshold cluster size tests with low thresholds, by including the mean voxel value of the suprathreshold cluster into the testing procedure.

Methodology

The bivariate parameter space $P = \mathbf{Z}^{+}\times\Re^{+} = \{\{0,1,\ldots\}\times[0,\infty)\}$ for the size (in voxels) and magnitude of a suprathreshold cluster is partitioned into rejection and acceptance regions by an "iso-cumulative" curve. Let $C(s,m)$ be the number of clusters with size $> s$ and mean voxel value $> m$ in an single statistic image, with expected value $E[C(s,m)]$. Let $P_{inf}(s,m) = \{(s',m')\in P: E[C(s',m')] \leq E[C(s,m)]\} \subseteq P$ for $(s,m)\in P$. The boundary of this region $\{(s',m')\in P: E[C(s',m')] = E[C(s,m)\}$ is an *iso-cumulative* curve, so called since for each point (s',m') on the curve, the expected number of clusters with size $> s'$ and mean voxel value $> m'$ is constant. Let $I(s,m)$ be the number of suprathreshold clusters in an image with attributes $(s',m')\in P_{inf}(s,m)$. The rejection region for a level α test is then $P_{inf}(s_\alpha, m_\alpha)$, where (s_α, m_α) are chosen such that $\Pr(I(s_\alpha,m_\alpha) \geq 1) = \alpha$ under H_W. The rejection region is unique, although (s_α, m_α) are not.

Considering the testing of a suprathreshold cluster with attributes (s,m), if $\Pr(I(s,m)\geq 1) \leq \alpha$ then $P_{inf}(s,m)\subseteq P_{inf}(s_\alpha, m_\alpha)$, since the iso-cumulative curves do not cross. Since $(s,m)\in P_{inf}(s,m)$, (s,m) is in the rejection region and the omnibus null hypothesis for the suprathreshold cluster of voxels is rejected.

Since direct estimation of the rejection region from simulated statistic images is difficult, Poline & Mazoyer (1994) assumed a Poisson distribution for $I(s,m)$, suggested by the law of rare events. Then, $\Pr(I(s,m)\geq 1) = 1-e^{-E[I(s,m)]}$, and only $E[I(s,m)]$ needs to be estimated to apply the test.

It remains to estimate $E[I(s,m)]$. For a given image, an empirical iso-cumulative curve can be computed as the boundary of $\hat{P}_{inf}(s,m) = \{(s',m')\in P: C(s',m') \leq C(s,m)\}$, a step function, passing through (s,m). This gives an estimate, $\hat{I}(s,m)$, of $I(s,m)$ as the number of objects in image with attributes $(s',m')\in \hat{P}_{inf}(s,m)$. Computing the mean of $\hat{I}(s,m)$ over many simulated images gives an estimate of $E[I(s,m)]$.

Comments

The description given above differs slightly to that of Poline & Mazoyer (1994a), which was not as rigorous in its definition of the iso-cumulative curves. A summary of the approach appears in Poline & Mazoyer (1994b).

A simpler approach would be to approximate the iso-cumulative curves with some function f, as $f(s,m) = c$. The function f then constitutes a statistic describing a suprathreshold portion of the statistic image. The null distribution of f_{max}, the largest f-value for clusters in a single image is then easily simulated, and the appropriate quantile for a level α test based on the f-values of clusters estimated. This avoids the necessity of the Poisson assumption and the lengthy computation of $\hat{I}(s,m)$ for each object from the simulated data.

Hierarchical decomposition

In addition to the bivariate approach for suprathreshold clusters, Poline & Mazoyer (1994a) propose a hierarchical decomposition of the image into *objects* whose size and mean amplitude are analysed, thus avoiding having to choose a threshold for cluster identification. Essentially the local maxima are iteratively "cut off" to form the objects. Fig.93 illustrates the objects after hierarchical decomposition of a simple 1D image. (For rigorous definitions and further explanation the reader is referred to Poline & Mazoyer (1994a) and the references there.) This represents an interesting direction. However, the hierarchical nature of the decomposition leads to objects whose attributes are not independent. The effect of this is probably negligible.

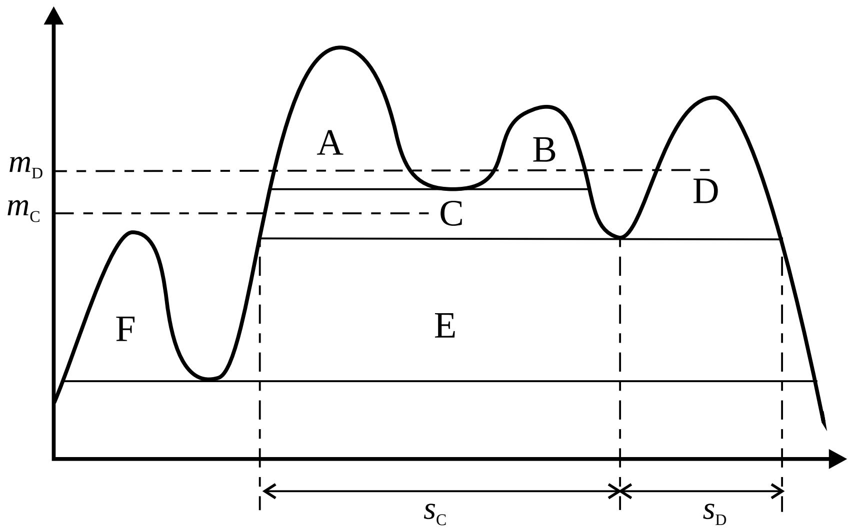

Figure 93
Hierarchical decomposition of a continuous one-dimensional "image".
The means and sizes for "objects" C & D are shown.

References

Adler RJ, Hasofer AM (1976) “Level Crossings for Random Fields” *Ann. Prob.* **4**:1–12

Adler RJ (1976) “Excursions Above a Fixed Level by *n*-Dimensional Random Fields” *Journal of Applied Probability* **13**:276–289

Adler RJ (1981) *The Geometry of Random Fields* Wiley, New York.

Bailey DL, Jones T, Spinks TJ, Gillardi M-C, Townsend DW (1991) “Noise equivalent count measurements in a neuro-PET scanner with retractable septa” *IEEE Transactions on Medical Imaging* **10**:256–260

Bailey DL, Jones T, Watson JDG, Schnorr L, Frackowiak RSJ (1993) “Activation studies in 3D PET: evaluation of true signal gain” In *Quantification of Brain Function: Tracer Kinetics and Image Analysis in PET* K Uemura *et al.* (eds). Excerpta Medica International Congress Series (1030), Elsevier Science Publishers. Pages 341–347

Besag JE (1974) “Spatial Interaction and the Statistical Analysis of Lattice Systems” *Journal of the Royal Statistical Society* **B36**:192-236

Besag J, Green PJ (1993) “Spatial Statistics and Bayesian Computation” *Journal of the Royal Statistical Society* **B55**:25–37

Blair RC, Karniski W (1994) “Distribution-Free Statistical Analyses of Surface and Volumetric Maps” In *Functional Neuroimaging: Technical Foundations.* Eds. RW Thatcher *et al.* Academic Press, San Diego, California.

Brownell GL, Budinger TF, Lauterbur PC, McGeer PL (1982) “Positron Emission Tomography and Nuclear Magnetic Resonance Imaging” *Science* **215**:619–626

Cherry SR, Dahlbom M, Hoffman EJ (1991) “3D PET using a conventional multislice tomograph without septa” *Journal of Computer Assisted Tomography* **15**:655–668

Cherry SR, Woods RP, Hoffman EJ, Mazziotta JC (1993) “Improved detection of focal cerebral blood flow changes using three-dimensional positron emission tomography” *Journal of Cerebral Blood Flow and Metabolism* **13**:630–638

Clark C, Carson R (1993) “Analysis of Covariance in Statistical Parametric Mapping” Letter to the Editor, *Journal of Cerebral Blood Flow and Metabolism* **13**:1038

D’Agostino RB (1971) “An Omnibus Test of Normality for Moderate and Large Size Samples” *Biometrika* **58**:341–348

D’Agostino RB, Tietjen GL (1971) “Simulation of Probability Points of b_2 for Small Samples” *Biometrika* **58**:669–672

Defrise M, Townsend DW, Clack R (1989) “Three-Dimensional Image Reconstruction from Complete Projections” *Physics in Medicine and Biology* **34**:573–587

Defrise M, Townsend D, Geissbuhler A (1990) "Implementation of Three-Dimensional Image Reconstruction for Multi-Ring Positron Tomographs" *Physics in Medicine and Biology* **35**:1361–1372

Dempster AP, Laird NM, Rubin DB (1977) "Maximum Likelihood from Incomplete Data via the EM Algorithm" *Journal of the Royal Statistical Society* **B39**:1–38

Derin H, Elliott H (1987) "Modelling and Segmentation of Noisy and Textured Images Using Gibbs Random Fields" *IEEE Transactions on Pattern Analysis and Machine Intelligence* **9**:39-55

Draper NR, Smith H (1981) *Applied Regression Analysis* (2nd ed.) John Wiley & Sons, New York.

Dubes RC, Jain AK (1989) "Random Field Models in Image Analysis" *Journal of Applied Statistics* **16**:131-164

Dubes RC, Jain AK (1993) "Random Field Models in Image Analysis" *Advances in Applied Statistics,* p121–154. Supplement to *Journal of Applied Statistics* **21**(1/2)

Eberl S, Kanno I, Fulton RR, Iida H, Chew PPL, Lee KS, Hutton BF (1993) "Automatic 3D Spatial Alignment for Correcting Interstudy Patient Motion in Serial PET Studies" In *Quantification of Brain Function: Tracer Kinetics and Image Analysis in PET* K Uemura *et al.* (eds). Excerpta Medica International Congress Series (1030), Elsevier Science Publishers. p419–426

Edgington ES (1964) "Randomization Tests" *The Journal of Psychology* **57**:445–449

Edgington ES (1969a) *Statistical Inference: The Distribution Free Approach* McGraw-Hill series in Psychology, New York.

Edgington ES (1969b) "Approximate Randomization Tests" *The Journal of Psychology* **72**:143–149

Edgington ES (1980) *Randomization Tests*. Marcel Dekker, New York & Basel.

English RJ, Brown SE (1986) *Single Photon Emission Computed Tomography.* Society of Nuclear Medicine.

Filliben JJ (1975) "The Probability Plot Correlation Coefficient Test for Normality" *Technometrics* **17**:111

Fishbach GD (1993) "Mind and Brain" in *Readings from Scientific American Magazine* Scientific American, Inc.

Fisher RA (1935) *The Design of Experiments*. Edinburgh: Oliver & Boyd.

Ford I (1985) Discussion of "A Smoothed EM Approach to Indirect Estimation Problems, with Particular Reference to Stereology and Emission Tomography" by BW Silverman *et al. Journal of the Royal Statistical Society* **B52**:309

Ford I, McColl JH, McCormack AG, McCrory SJ (1991) "Statistical Issues in the Analysis of Neuroimages" *Journal of Cerebral Blood Flow and Metabolism* **11**:A89–A95

Ford I (1994) "Some Non-Ontological and Functionally Unconnected Views on Current Issues in the Analysis of PET datasets" Invited paper, *Journal of Cerebral Blood Flow and Metabolism* (In press)

Fox PT, Perlmutter JS, Raichle ME (1985) "A Stereotactic Method of Anatomical Localization for Positron Emission Tomography" *Journal of Computer Assisted Tomography* 9(1)141–153

Fox PT, Mintun MA, Reiman EM, Raichle ME (1988) "Enhanced Detection of Focal Brain Responses Using Intersubject Averaging and Change-Distribution Analysis of Subtracted PET Images" *Journal of Cerebral Blood Flow and Metabolism* **8**:642–653

Fox PT, Mintun MA (1989) "Non-Invasive Functional Brain Mapping by Change Distribution Analysis of Averaged PET Images of $H_2{}^{15}O$ Tissue Activity" *Journal of Nuclear Medicine* **30**:141–149

Friston KJ, Passingham RE, Nutt JG, Heather JD, Sawle GV, Frackowiak RSJ (1989) "Localisation in PET Images: Direct Fitting of the Intercommissural (AC-PC) Line" *Journal of Cerebral Blood Flow and Metabolism* **9**:690–695

Friston KJ, Frith CD, Liddle PF, Dolan RJ, Lammertsma AA, Frackowiak RSJ (1990) "The Relationship Between Global and Local Changes in PET Scans" *Journal of Cerebral Blood Flow and Metabolism* **10**:458–466

Friston KJ, Dolan RJ, Frackowiak RSJ (1991a) "Statistical Parametric Mapping" Review of theory for SPM Manual. Private communication.

Friston KJ, Frackowiak RSJ (1991b) "Imaging Functional Anatomy" In *Brain Work and Mental Activity. Alfred Benzon Symposium 31* NA Lassen, DH Ingvar, ME Raichle, L Friberg (eds) pp267–277. Munksgaard, Copenhagen. (Gives an overview of the analysis implemented in the SPM software.)

Friston KJ, Frith CD, Liddle PF, Frackowiak RSJ (1991c) "Plastic Transformation of PET Images" *Journal of Computer Assisted Tomography* **15**:634–639

Friston KJ, Frith CD, Liddle PF, Frackowiak RSJ (1991d) "Comparing Functional (PET) Images: The Assessment of Significant Change" *Journal of Cerebral Blood Flow and Metabolism* **11**:690–699

Friston KJ, Frith CD, Liddle PF, Frackowiak RSJF (1993a) "Functional Connectivity: The Principal Component Analysis of Large (PET) Data Sets" *Journal of Cerebral Blood Flow and Metabolism* **13**:5–14

Friston KJ, Frith CD, Frackowiak RSJF (1993b) "Principal Component Learning Algorithms: A Neurobiological Analysis" *Proceedings of the Royal Society of London* **B254**:47–54
Functional Connectivity: The Principal Component Analysis of Large (PET) Da

Friston KJ (1994a) “Statistical Parametric Mapping: Ontology and Current Issues” Submitted to *Journal of Cerebral Blood Flow and Metabolism*

Friston KJ, Holmes AP, Worsley KJ, Poline JP, Frith CD, Frackowiak RSJ (1994b) “Statistical Parametric Maps in Functional Imaging: A General Approach” Submitted to *Human Brain Mapping*

Friston KJ, Ashburner J, Poline JB, Frith CD, Heather JD, Frackowiak RSJF (1994c) “The spatial registration and normalisation of images” Submitted to *Human Brain Mapping*

Friston KJ, Worsley KJ, Frackowiak RSJ, Mazziotta JC, Evans AC (1994d) “Assessing the Significance of Focal Activations Using their Spatial Extent” *Human Brain Mapping* **1**:214–220

Geman S, Geman D (1984) “Stochastic Relaxation, Gibbs Distributions and the Bayesian Restoration of Images” *IEEE Transactions on Pattern Analysis and Machine Intelligence* **6**:721-741

Goldstein H (1986) “Multilevel Mixed Model Analysis Using Iterative Generalised Least Squares” *Biometrika* **73**:43–56

Green PJ (1985) Discussion of “A Smoothed EM Approach to Indirect Estimation Problems, with Particular Reference to Stereology and Emission Tomography” by BW Silverman *et al.* *Journal of the Royal Statistical Society* **B52**:303–304

Green PJ (1990) “Bayesian Reconstructions from Emission Tomography Data Using a Modified EM algorithm” *IEEE Transactions on Medical Imaging* **9**:84-93

Hasofer, AM (1976) “The mean number of maxima above high levels in Gaussian random fields” *Journal of Applied Probability*. **13**:377–379

Healy, MJR (1986) *Matrices for Statisticians*. Oxford.

Herman GT (1980) *Image Reconstruction from Projections* New York: Academic

Herschovitch P, Raichle ME, Kilbourn MR, Welch MJ (1987) “Positron Emission Tomographic Measurement of Cerebral Blood Flow and Permeability — Surface Area Product of Water Using [^{15}O]Water and [^{11}C]Butanol” *Journal of Cerebral Blood Flow and Metabolism* **7**:527–542

Herscovitch P, Markham J, Raichle ME (1983) “Brain Blood Flow Measured with Intravenous $H_2{}^{15}O$. I:Theory and Error Analysis” *Journal of Nuclear Medicine* **14**:782–789

Hochberg Y, Tamhane AC (1987) *Multiple Comparisons Procedures* John Wiley & Sons.

Hoeffding W (1952) “The Large Sample Power of Tests Based on Permutations of Observations” *Annals of Mathematical Statistics* **23**:196–192

Hoffman EJ, Phelps ME (1986) “Positron Emission Tomography: Principles and Quantification” In *Positron Emission Tomography and Autoradiography: Principles and Applications for the Brain and Heart* Phelps *et al.* (eds). Raven Press, New York.

Hoffman EJ, Cutler PD, Digby WM, Mazziotta JC (1990) “3D Phantom to SImulate Cerebral Blood Flow and Metabolic Images for PET” *IEEE Transactions on Nuclear Science* **37**:616–620

Holm S (1979) “A Simple Sequentially Rejective Multiple Test Procedure” *Scandanavian Journal of Statistics* **6**:54–70

Holmes AP, Ford I (1993a) “A Bayesian Approach to Significance Testing for Statistic Images from PET” *Annals of Nuclear Medicine* **7**:S106–S107

Holmes AP, Ford I (1993b) “A Bayesian Approach to Significance Testing for Statistic Images from PET” In *Quantification of Brain Function: Tracer Kinetics and Image Analysis in PET* (1993), K Uemura *et al*. (eds). Excerpta Medica International Congress Series (1030), Elsevier Science Publishers. p521–531.

Holmes AP, Blair RC, Watson JDG, Ford I (1995) “Non-Parametric Analysis of Statistic Images from Functional Mapping Experiments” *Journal of Cerebral Blood Flow and Metabolism*. In Press.

Huang S-C, Phelps ME (1986) “Principles of Tracer Kinetic Modelling in Positron Emission Tomography and Autoradiology” In *Positron Emission Tomography and Autoradiography: Principles and Applications for the Brain and Heart* Phelps *et al.* (eds). Raven Press, New York.

Ising E (1957) “Information Theory and Statistical Mechanics” *Phys. Rev.* **106**:620–630

Kaufman L (1987) “Implementing and Accelerating the EM Algorithm for Positron Emission Tomography” *IEEE Transactions on Medical Imaging* **6**:37–51

Kearfott KJ (1985) Comment on “A Statistical Model for Positron Emission Tomography” (by Y Vardi, LA Shepp, L Kaufman) *Journal of the American Statistical Association* **80**:26–28

Kety SS (1951) “The Theory and Applications of the Exchange of Inert Gas at the Lungs and Tissues” *Pharmmacol. Review* **3**:1–41

Kety SS (1991) “The Early History of the Coupling Between Cerebral Blood Flow, Metabolism and Function” In *Brain Work and Mental Activity: Alfred Benzon Symposium 31.* NA Lassen *et al.* (eds) Munksgaard, Copenhagen. pp19–28

Knuth DE (1992) “Two Notes on Notation” *The American Mathematical Monthly* **99**:403–422

Krzanowski WJ, (1988) *Principles of Multivariate Analysis: A User's Perspective*, Oxford.

Lange K, Carson R (1984) "EM Reconstruction Algorithms for Emission and Transmission Tomography" *Journal of Computer Assisted Tomography* **8**:306–316

Mazoyer BM, Tzourino N, Poline J-B, Petit L, Levrier O, Joliot M (1993) "Anatomical Regions of Interest versus Stereotactic Space: A Comparison of Two Approaches for Brain Activation Maps Analysis" In *Quantification of Brain Function: Tracer Kinetics and Image Analysis in PET* K Uemura *et al.* (eds). Excerpta Medica International Congress Series (1030), Elsevier Science Publishers.

Mazziotta JC, Huang SC Phelps ME, Carson RE, MacDonald NS, Mahoney K (1985) "A Non-Invasive Positron Computed Tomography Technique Using ^{15}O Labelled Water for the Evaluation of Neurobiological test batteries" *Journal of Cerebral Blood Flow and Metabolism* **5**:70–78

McColl JH, Holmes AP, Ford I (1994) "Statistical Methods in Neuroimaging with Particular Application to Emission Tomography" *Statistical Methods in Medical Research* **3**:63–86

McCrory SJ, Ford I (1991) "Multivariate Analysis of SPECT Images with Illustrations in Alzheimers Disease" *Statistics in Medicine* **10**:1711–1718

Miller RJ (1986) *Beyond ANOVA, Basics of Applied Statistics* Wiley, New York.

Minoshima S, Berger KL, Lee KS, Mintun MA (1992) "An Automated Method for Rotational Correction and Centering of Three-Dimensional Functional Brain Images" *Journal of Nuclear Medicine* **33**:1579–1585

Mintun MA, Gorman J, Swift AG, Snyder DL (1985) "Evaluation of the Maximum Likelihood Method for Reconstruction Images in Positron Emission Tomography" (Abstract) *Journal of Nuclear Medicine* **26**:720

Mintun MA, Lee KS (1990) "Mathematical Realignment of Paired PET Images to Enable Pixel-by-Pixel Subtraction" (Abstract) *Journal of Nuclear Medicine* **31**:816

Moeller JR, Strother SC, Sidtis JJ, Rottenberg DA (1987) "Scaled Subprofile Model: A Statistical Approach to the Analysis of Functional Patters in Positron Emission Tomographic Data" *Journal of Cerebral Blood Flow and Metabolism* **7**:649–658

Moeller JR, Strother SC (1991) "A Regional Covariance Approach to the Analysis of Functional Patterns in Positron Emission Tomographic Data" *Journal of Cerebral Blood Flow and Metabolism* **11**:A121–A135

Phelps M, Mazziotta JC, Schelbert HR (eds) (1986) *Positron Emission Tomography and Autoradiography: Principles and Applications for the Brain and Heart* Raven Press, New York.

Pitman EJG (1937a) "Significance Tests Which May Be Applied To Samples From Any Population" *Journal of the Royal Statistical Society* Supplement **4**:119–130

Pitman EJG (1937b) "Significance Tests Which May Be Applied To Samples From Any Population. II The Correlation Coefficient Test" *Journal of the Royal Statistical Society* Supplement **4**::225–232

Poline J-B, Mazoyer BM (1993) "Analysis of Individual Positron Emission Tomography Activation Maps by Detection of High Signal-to-Noise Ratio Pixel Clusters" *Journal of Cerebral Blood Flow and Metabolism* **13**:425–437

Poline J-B, Mazoyer BM (1994a) "Analysis of Individual Brain Activation Maps Using a Multi-Scale Detection Algorithm" Submitted to *IEEE Transactions in Medical Imaging*.

Poline J-B, Mazoyer BM (1994b) "Cluster Analysis in Individual Functional Brain Images: Some New Techniques to enhance the Sensitivity of Activation Detection Methods" Submitted to *Human Brain Mapping*

Radon J (1917) "Über die Bestimmung von Funktionen durch ihre Integralwerte längs gewisser Mannigfaltigkeiten" *Berichte Saechsische Akademie der Wissenschaften* **69**:262–277

Raichle ME, Martin WRW, Herscovitch P, Mintun MA, Markham J (1983a) "Brain Blood Flow Measured with Intravenous $H_2{}^{15}O$. II: Implementation and Validation" *Journal of Nuclear Medicine* **14**:790–798

Raichle ME (1983b) "Positron Emission Tomography" *Annual Review of Neuroscience* **6**:249–267

Roland PE, Levin B, Kawashima R, Åkerman S (1993) "Three-Dimensional Analysis of Clustered Voxels in ^{15}O-Butanol Brain Activation Images" *Human Brain Mapping* **1**:3–19

Scheffé H (1959) *The Analysis of Variance*. Wiley, New York.

Scientific American (1993) *Mind and Brain: Readings from Scientific American Magazine* Scientific American, Inc.

Shepp LA, Kruskal JB (1978) "Computerised Tomography: The New Medical X-Ray Technology" *American Mathematical Monthly* **85**:420–439

Shepp LA, Vardi Y (1982) "Maximum Likelihood Reconstruction for Emission Tomography" *IEEE Transactions on Medical Imaging* **1**:113–122

Shepp LA, Vardi Y, Ra JB, Hilal SK, Cho ZH (1984) "Maximum Likelihood PET with Real Data" *IEEE Transactions on Nuclear Science* **31**:910–913

Silbersweig DA, Stern E, Frith CD, Cahill C, Schnorr L, Grootoonk S, Spinks T, Clark J, Frackowiak RSJ, Jones T (1993) "Detection of Thirty-Second Cognitive Activations in Single Subjects with Positron Emission Tomography: A New Low-Dose $H_2{}^{15}O$ Regional Cerebral Blood Flow Three-Dimensional Imaging Technique *Journal of Cerebral Blood Flow and Metabolism* **13**:617–629

Silbersweig DA, Stern E, Schnorr L, Frith CD, Ashburner J, Cahill C, Frackowiak RSJ, Jones T (1994) "Imaging Transient, Randomly Occuring Neuropsychological Events in Single Subjects with Positron Emission Tomography: An Event-Related Countrate Correlational Analysis" In preperation.

Silverman BW, Jones MC, Wilson JD, Nychka (1990) “A Smoothed EM Approach to Indirect Estimation Problems, with Particular Reference to Stereology and Emission Tomography” *Journal of the Royal Statistical Society* **B52**:271–324

Smith AFM, Roberts GO (1993) “Bayesian Computation Via the Gibbs Sampler and Related Markov Chain Monte Carlo Methods” *Journal fo the Royal Statistical Society* **B55**:3–23

Snedecor GW, Cochran WG (1967) *Statistical Methods* (6th edition) Iowa State University Press

Sokolov L (1979) “Relation between Physiological Function and Energy Metabolism in the Central Nervous Systemm” *Journal of Neurochemistry* **29**:13–26

Spinks TJ, Jones T, Bailey DL, Townsend DW, Gillardi MC, Casey ME, Sipe B (1991) “Physical characteristics of a new positron tomograph with retractable septa” *IEEE Medical Imaging Conf. Rec.* 1991:1313–1317

Strother SC, Casey ME, Hoffman EJ (1990) “Measuring PET scanner sensitivity: Relating count rates to signal-to-noise ratios using noise equivalent counts” *IEEE Transactions in Nuclear Science* **37**:783–788

Strother SC, Kanno I, Rottenberg DA (1994) “Principal Component Analysis, Variance Partitioning and Functional Connectivity” Submitted to *Journal of Cerebral Blood Flow and Metabolism*

Talairach J, Tournoux P (1988) *Co-Planar Stereotactic Atlas of the Human Brain* Stuttgart:Thieme

Tennent RM (1971) *Science Data Book* Oliver and Boyd

Ter-Pogossian MM, Raichle ME, Sobel BE (1980) “Positron Emission Tomography” *Scientific American* **243**:140–155

Townsend DW, Geissbuhler A, Defrise M, Hoffman EJ, Spinks TJ, Bailey DL, Gilardi M-C, Jones T (1991) “Fully Three-Dimensional Reconstruction for a PET Camera with Retractable Septa” *IEEE Transactions on Medical Imaging* 10(4)505–512

Uemura K, Lassen NA, Jones T, Kanno I (eds) (1993) *Quantification of Brain Function: Tracer Kinetics and Image Analysis in PET*. Excerpta Medica International Congress Series (1030), Elsevier Science Publishers.

“Health” *Vogue* January 1992: 130–131

Vardi Y, Shepp LA, Kaufman L (1985) “A statistical model for positron emission tomography” *Journal of the American Statistical Association* **80**:8–37

Watson JDG, Myers R, Frackowiak RSJ, Hajnal JV, Woods RP, Mazziotta JC, Shipp S, Zeki S (1993) “Area V5 of the Human Brain: Evidence from a Combined Study Using Positron Emission Tomography and Magnetic Resonance Imaging” *Cerebral Cortex* **3**:79–94

Westfall PH, Young SS (1993) *Resampling-Based Multiple Testing: Examples and Methods for p-Value Adjustment.* Wiley, New York.

Wildt AR, Ahtola O (1978) *Analysis of Covariance.* Sage Publications 07-012, Beverley Hills / London.

Woods RP, Cherry SR, Mazziotta JC (1992) "Rapid Automated Algorithm for Aligning and Reslicing PET Images" *Journal of Computer Assisted Tomography* **16**:620–633

Woods RP, Mazziotta JC, Cherry SR (1993a) "Automated Image Registration" In *Quantification of Brain Function: Tracer Kinetics and Image Analysis in PET* K Uemura *et al.* (eds). Excerpta Medica International Congress Series (1030), Elsevier Science Publishers.

Woods RP, Mazziotta JC, Cherry SR (1993b) "MRI–PET Registration with Automated Algorithm" *Journal of Computer Assisted Tomography* 17(4)536–546

Worsley KJ, Evans AC, Marrett S, Neelin P (1992) "A Three-Dimensional Statistical Analysis for CBF Activation Studies in human Brain" *Journal of Cerebral Blood Flow and Metabolism* **12**:900–918

Worsley KJ, Evans AC, Marrett S, Neelin P (1993b) "Letters to the Editor" *Journal of Cerebral Blood Flow and Metabolism* **13**:1040–1042

Worsley KJ, Evans AC, Marrett S, Neelin P (1993) Detecting and Estimating the Regions of Activation in CBF activation studies in Human Brain. In *Quantification of Brain Function. Tracer Kinetics and Image Analysis in Brain PET.* Eds. K Uemura *et al*. Elsevier Science Publishers BV. 535–544

Worsley KJ (1994a) "Local Maxima and the Expected Euler Characteristic of Excursion Sets of χ^2, F and t Fields" *Advances in Applied Probability* **26**:13–42

Worsley KJ (1994b) Private communication from a paper in preparation.

Worsley KJ (1994c) "Quadratic tests for local changes in random fields with applications to medical images" Submitted to *Journal of the American Mathematical Association*

Yarowsky PJ, Ingevar DH (1981) "Neuronal Activity and Energy Metabolism" *Federation Proceedings* **40**:2353–2363

Zar JH (1984) *Biostatistical Analysis*. Prentice-Hall International

Zeki S, Watson JDG, Lueck CJ, Friston KJ, Kennard C, Frackowiak RSJ (1991) "A Direct Demonstration of Functional Specialisation in Human Visual Cortex" *The Journal of Neuroscience* **11**:641–649