Module 19:
FWER Correction

Family-Wise Error Rate

- The family-wise error rate (FWER) is the probability of making one or more Type I errors in a family of tests, under the null hypothesis.

- FWER controlling methods:
  - Bonferroni correction
  - Random Field Theory
  - Permutation Tests

Problem Formulation

- Let H_{0i} be the hypothesis that there is no activation in voxel i, where $i \in$ V ={1,…. m}.

- Let T_i be the value of the test statistic at voxel i.

- The family-wise null hypothesis, H_0, states that there is no activation in any of the m voxels.

$$H_0 = \bigcap_{i \in V} H_{0i}$$

- If we reject a single voxel null hypothesis, H_{0i}, we will reject the family-wise null hypothesis.

- A false positive at any voxel gives a Family-Wise Error (FWE)

- Assuming H_0 is true, we want the probability of falsely rejecting H_0 to be controlled by α, i.e.

$$P\left(\bigcup_{i \in V} \{T_i \geq u\} \mid H_0\right) \leq \alpha$$

Bonferroni Correction

- Choose the threshold so that

$$P(T_i \geq u \mid H_0) \leq \frac{\alpha}{m}$$

- Hence,

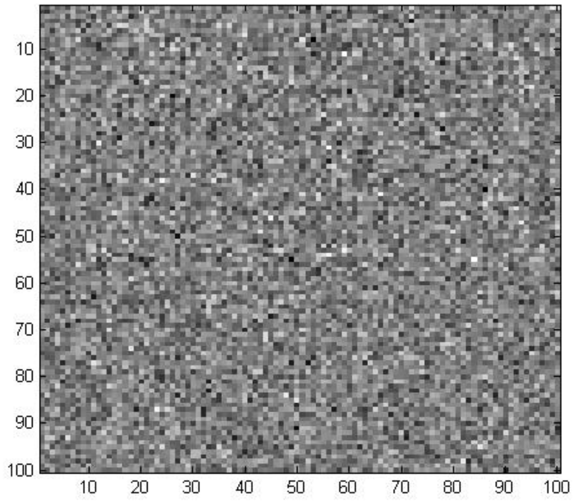
$$FWER = P\left(\bigcup_{i \in V} \{T_i \geq u\} \mid H_0\right)$$

$$\leq \sum_i P(T_i \geq u \mid H_0) \qquad \textbf{Boole's Inequality}$$

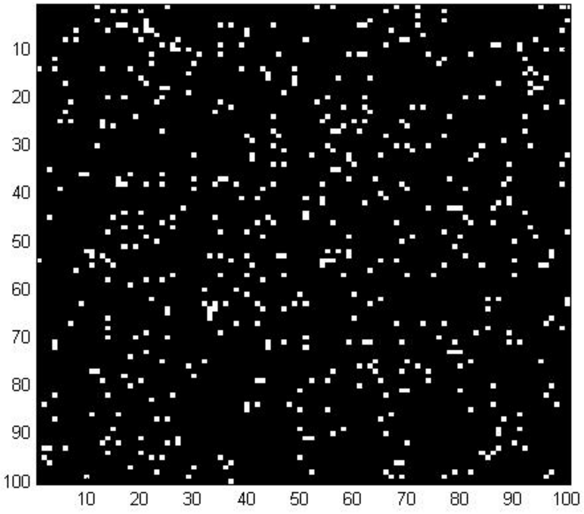
$$\leq \sum_i \frac{\alpha}{m} = \alpha$$

Example

Generate 100×100 voxels from an iid N(0,1) distribution



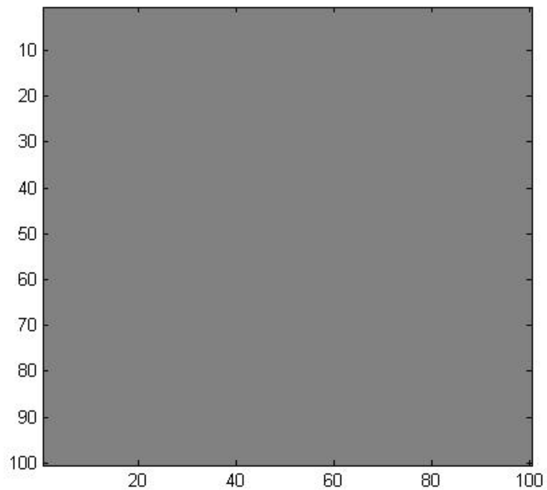
Threshold at u=1.645



Approximately 500 false positives.

To control for a FWE of 0.05, the Bonferroni correction is 0.05/10,000.

This corresponds to u=4.42.



No false positives

On average only 5 out of every 100 generated in this fashion will have one or more values above u.

Bonferroni Correction

- The Bonferroni correction is very conservative, i.e. it results in very strict significance levels.
- It decreases the power of the test (probability of correctly rejecting a false null hypothesis) and greatly increases the chance of false negatives.
- It is not optimal for correlated data, and most fMRI data has significant spatial correlation.

Spatial Correlation

- We may be able to choose a more appropriate threshold by using information about the spatial correlation in the data.
- Random field theory allows one to incorporate the correlation into the calculation of the appropriate threshold.
- It is based on approximating the distribution of the maximum statistic over the whole image.

Maximum Statistic

- Link between FWER and max statistic.

$$\text{FWER} = \text{P(FWE)}$$

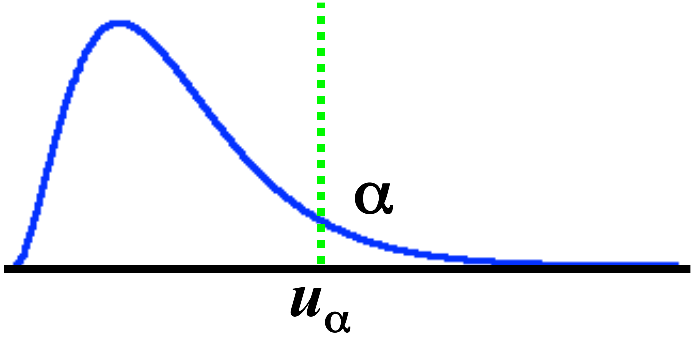
$$= \text{P}\left(\cup_i \{T_i \geq u\} \mid H_o \right)$$

P(any t-value exceeds *u* under null)

$$= \text{P}\left(\max_i T_i \geq u \mid H_o \right)$$

P(max t-value exceeds *u* under null)

Choose the threshold u such that the max only exceeds it α% of the time



Random Field Theory

- A random field is a set of random variables defined at every point in D-dimensional space.
- A Gaussian random field has a Gaussian distribution at every point and every collection of points.
- A Gaussian random field is defined by its mean function and covariance function.

Random Field Theory

- Consider a statistical image to be a lattice representation of a continuous random field.

- Random field methods are able to:

  - approximate the upper tail of the maximum distribution, which is the part needed to find the appropriate thresholds; and

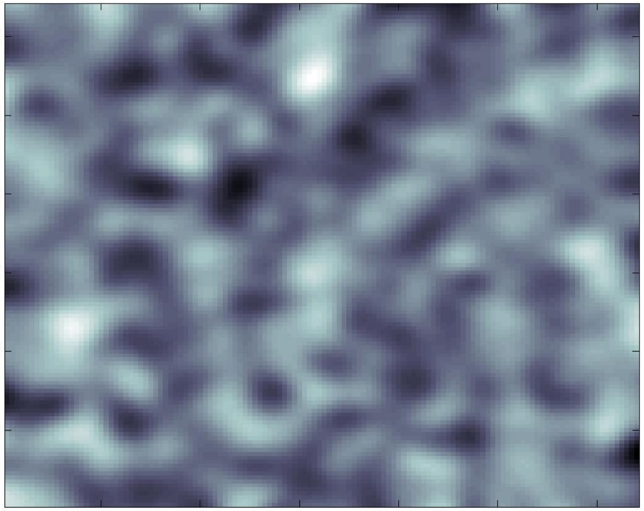
  - account for the spatial dependence in the data.

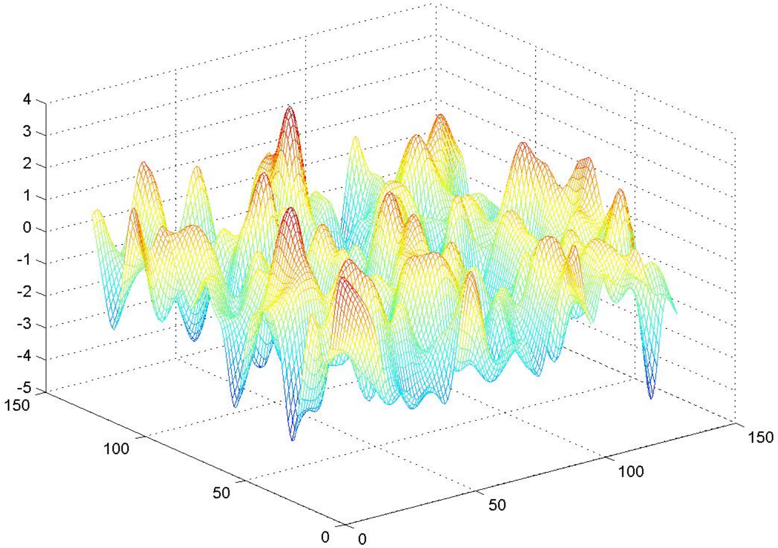
Random Field Theory

- Consider a random field $Z(s)$ defined on

$$s \in \Omega \subset R^D$$

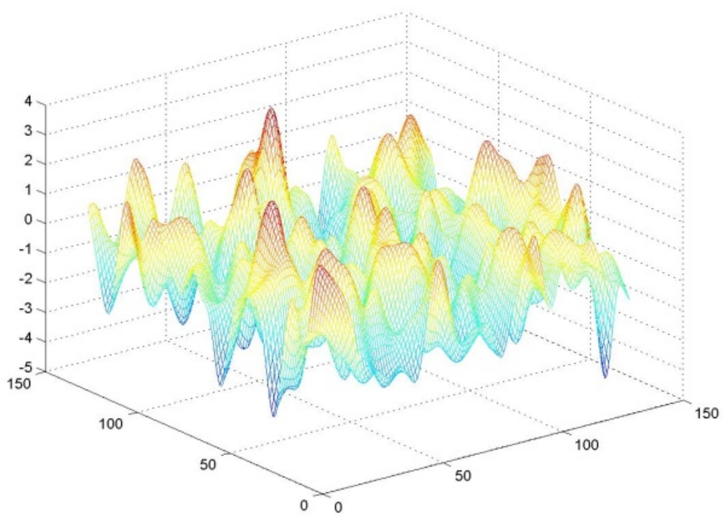
where *D* is the dimension of the process.





Euler Characteristic

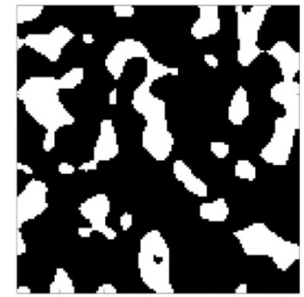
- Euler Characteristic χ_u
  - A property of an image after it has been thresholded.
  - Counts #blobs - #holes
  - At high thresholds, just counts #blobs



Random Field

$\chi_u = 28 - 1 = 27$

$\boldsymbol{u = 0.5}$

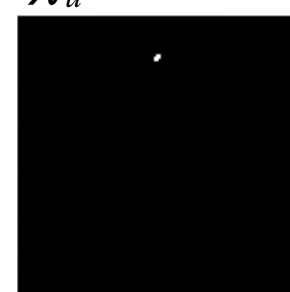


$\chi_u = 2$

$\boldsymbol{u = 2.75}$

$\chi_u = 1$

$\boldsymbol{u = 3.5}$



Threshold

Controlling the FWER

- Link between FWER and Euler Characteristic.

$$\text{FWER} = P(\max_i T_i \geq u \mid H_o)$$

$$= P(\text{One or more blobs} \mid H_o)$$

no holes exist

$$\approx P(\chi_u \geq 1 \mid H_o)$$

never more than 1 blob

$$\approx E(\chi_u \mid H_o)$$

- Closed form results exist for $E(\chi_u)$ for Z, t, F and χ^2 continuous random fields.

3D Gaussian Random Fields

For large search regions:

$$E(\chi_u) \approx R(4\log 2)^{3/2}(u^2 - 1)e^{-u^2/2}(2\pi)^{-2}$$

where

$$R = \frac{V}{FWHM_x FWHM_y FWHM_z}$$

Here V is the volume of the search region and the full width at half maximum (FWHM) represents the smoothness of the image estimated from the data.

R = Resolution Element (Resel)

Controlling the FWER

For large u:

$$FWER \approx R(4\log 2)^{3/2}(u^2 - 1)e^{-u^2/2}(2\pi)^{-2}$$

where

$$R = \frac{V}{FWHM_x FWHM_y FWHM_z}$$

Properties:

- As u increases, FWER decreases (Note u large).
- As V increases, FWER increases.
- As smoothness increases, FWER decreases.

RFT Assumptions

- The entire image is either multivariate Gaussian or derived from multivariate Gaussian images.
- The statistical image must be sufficiently smooth to approximate a continuous random field.
  - FWHM at least twice the voxel size.
  - In practice, FWHM smoothness 3-4×voxel size is preferable.
- The amount of smoothness is assumed known.
  - Estimate is biased when images not sufficiently smooth.
- Several layers of approximations.

End of Module