Computational Modeling for Biomedical Imaging

Gary Hui Zhang

gary.zhang@ucl.ac.uk

February 2nd, 2015

What has the course covered in Part 1?

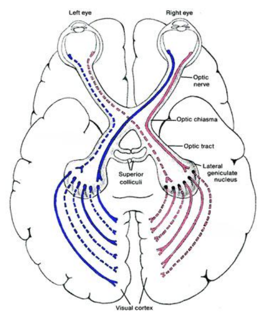
- Estimate model parameters
- Determine uncertainties in parameter estimation
- Choose the best model
- Design the most economical experiment

What will be covered in Part 2?

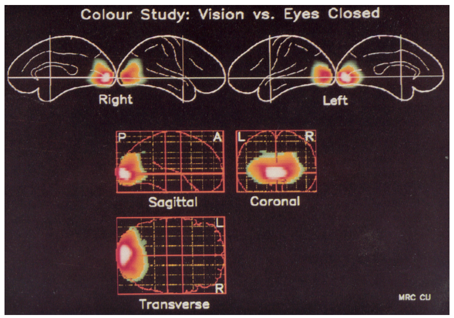
- Models for statistical inference in imaging
  - Spatially-local models -- univariate analysis
    - Statistical parametric and non-parametric mappings
- Model fitting for ill-posed problems
  - L_1 and L_2 regularization techniques

Examples of interesting questions in neuroimaging

- Which parts of our brain activate to receive and process visual stimulus?



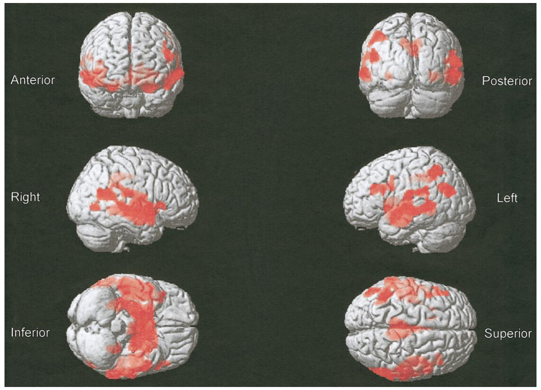
Source: senseofvision.wikispaces.com



Zeki et al, J Neuroscience, 11(3): 641-649 (1991)

Examples of interesting questions in neuroimaging

- Which area(s) of the brain are mostly affected with the onset of Alzheimer's disease?



Baron et al, NeuroImage 14, 298-309 (2001)

5

A Direct Demonstration of Functional Specialization in Human Visual Cortex

S. Zeki,[1] J. D. G. Watson,[1,2,3] C. J. Lueck,[4] K. J. Friston,[2] C. Kennard,[4] and R. S. J. Frackowiak[2,3]
[1]Department of Anatomy, University College, London WC1E 6BT, United Kingdom, [2]MRC Cyclotron Unit, Hammersmith

Context

The concept of functional specialization in the visual cortex is therefore more recent (Zeki, 1974a, 1978a). It is based on studies of the macaque monkey, which show that, of the many visual areas (Cragg, 1969; Zeki, 1969, 1971) that lie outside the striate cortex, one area (V5) is specialized for visual motion,

while another, anatomically distinct area (V4) is specialized for color (Zeki, 1973, 1974b, 1977). Following this discovery, re-

Objectives

In this study, we wanted to demonstrate functional specialization in the normal human visual cortex directly and to chart the anatomy of the areas involved. Obviously, this requires the use of at least 2 different visual stimuli, representing 2 different submodalities, in the same study. We used the same 2 submodalities, motion and color, that were used to establish functional specialization in the monkey visual cortex (Zeki, 1973, 1974b). Our approach was to use positron emission tomography (PET) to detect significant regional changes in cerebral blood flow (rCBF) in the brains of normal human subjects when they viewed stimuli chosen to emphasize color or motion. We were

6

A Direct Demonstration of Functional Specialization in Human Visual Cortex

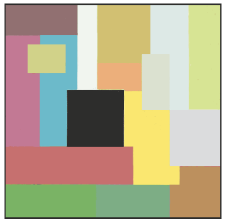
S. Zeki,[1] J. D. G. Watson,[1,2,3] C. J. Lueck,[4] K. J. Friston,[2] C. Kennard,[4] and R. S. J. Frackowiak[2,3]
[1]Department of Anatomy, University College, London WC1E 6BT, United Kingdom, [2]MRC Cyclotron Unit, Hammersmith

Experiment

Nine normal volunteers were studied. All were male, 7 were right handed, and their ages ranged from 21 to 43 yr (mean, 29 yr). Two experiments were performed. In the first, 6 subjects were stimulated with colored and isoluminant gray-shaded abstract displays occupying the central 40° of the field of view. In the second, the other 3 subjects were

Experiment 1—Stimulus A: Eyes closed. Stimulus B: Eyes open, with the central 40° occupied by an abstract display of 15 multicolored squares and rectangles (Land color Mondrian; Land, 1974). Stimulus C: Eyes open, with the central field occupied by an isoluminant gray shaded version of the colored Mondrian presented in stimulus B.

Blood flow measurement. Cerebral blood flow was measured with the dynamic/integral technique described by Lammertsma et al. (1990), of which a brief account follows. Subjects inhaled trace quantities of $C^{15}O_2$ provided at a concentration of 6 MBq/ml and a flow rate of 500 ml/min through a standard oxygen face mask for a period of 2 min. The effective dose equivalent of each administration was 1.2 mSv. Twenty-one successive dynamic PET scans were collected within a period of 3.5 min, starting 0.5 min before administration of the flow tracer ($C^{15}O_2$).



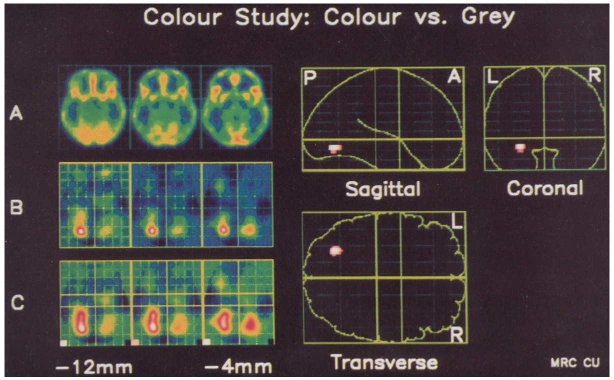
Land color Mondrian

7

A Direct Demonstration of Functional Specialization in Human Visual Cortex

S. Zeki,[1] J. D. G. Watson,[1,2,3] C. J. Lueck,[4] K. J. Friston,[2] C. Kennard,[4] and R. S. J. Frackowiak[2,3]
[1]Department of Anatomy, University College, London WC1E 6BT, United Kingdom, [2]MRC Cyclotron Unit, Hammersmith

Result



8-1

A Direct Demonstration of Functional Specialization in Human Visual Cortex

S. Zeki,[1] J. D. G. Watson,[1,2,3] C. J. Lueck,[4] K. J. Friston,[2] C. Kennard,[4] and R. S. J. Frackowiak[2,3]

[1]Department of Anatomy, University College, London WC1E 6BT, United Kingdom, [2]MRC Cyclotron Unit, Hammersmith

Result

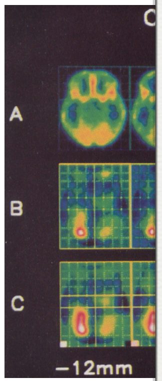


Table 1. Experiment 1: color versus gray

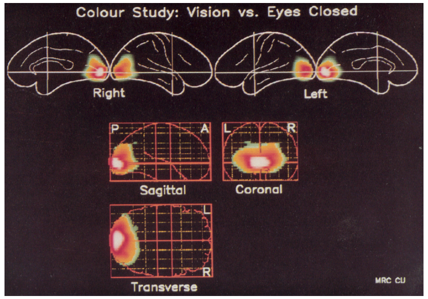
Region	CBF (adjusted group means; ml dl^{-1} min^{-1}) Rest	Gray	Color	Talairach coordinates[a]	Z score
Right V4	59	62	64	+20, −66, −4	2.90*
Left V4	65	65	68	−26, −68, −8	3.93
Mean	62	63	66		
Right V5	51	51	52		
Left V5	46	49	48		
Mean	49	50	50		
V1/V2	62	72	72	−4, −90, 0	13.0

Adjusted group mean rCBF for each condition (averaged from 2 estimations for each of the 3 physiological states) in the areas indicated in the SPMs to have shown significant rCBF change (at $p < 0.05$ level) for color versus gray. Flow values have been normalized by ANCOVA to a mean flow of 50 ml dl^{-1} min^{-1}. Loci were identified on the SPM in which V4 was activated. V1/V2 was identified

Statistical mapping

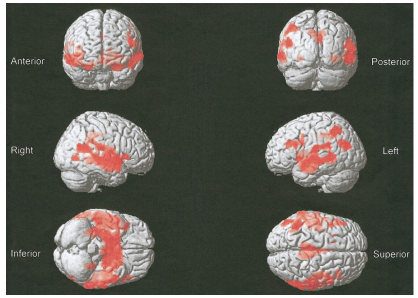
- A method for locating statistically-significant *focal or regional* effects in images.
- Application examples:

Functional Specialization of Visual Cortex



Zeki et al, J Neuroscience, 11(3): 641-649 (1991)

Brain atrophy in mild Alzheimer's disease



Baron et al, NeuroImage 14, 298-309 (2001)

Statistical mapping: a massive univariate analysis

- Step one: the formation of a statistic map , i.e., *ask the same question for each bit of the brain*
  - Compute, at **each** voxel, a **statistic** indicating evidence of the experimental **effect of interest**, at that voxel.
  - Statistic map a.k.a. *statistic image* or *statistical parametric map*
- Step two: the assessment of the statistic map
  - Assess the entire or part of the statistic image for significant experimental effects
  - Require a method that accounts for massive simultaneous multiple comparisons.

Statistical parametric mapping (SPM) vs Statistical non-parametric mapping (SnPM)

- SPM makes *strong assumptions* about data
  - Data are normally distributed, with mean parametrized by a *general linear model* (GLM).
  - Computation underpinned by classical parametric statistical framework
  - Multiple comparison correction uses the theory of Gaussian random fields
- SnPM relies on *minimal assumptions* about data
  - Permutation-based and simpler to understand
  - Handles multiple comparison correction naturally

General Linear Model (GLM)

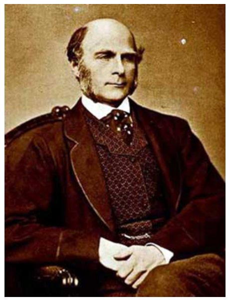
- Model definition
- Parameter estimation
- Hypothesis testing
- Applications in Neuroimaging

General Linear Model (GLM)

- A unifying framework for *regression analysis* and *analysis of variance*
  - Objectives: understand the relationship between one set of variables (*response* or *outcome* or *dependent* variables) and another (*predictor* or *independent* variables)
  - Predictor variables can be
    - Quantitative: Regression analysis or linear modeling
    - Categorical: Analysis of Variance (ANOVA)
    - Both: Analysis of Covariance (ANCOVA)

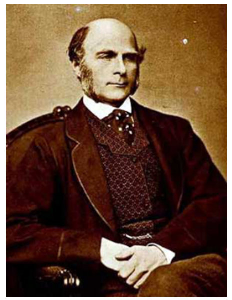
Regression Analysis

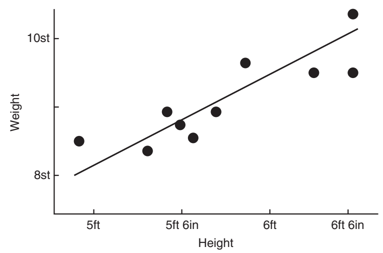
- Pioneered by Sir Francis Galton (1886; 1888)
  - Regression towards the mean: from his studies of the relationship between the characteristics of the parents and their offsprings
- Predict the average values of the response variable for each value of the predictors



Regression Analysis

- Pioneered by Sir Francis Galton (1886; 1888)
  - Regression towards the mean: from his studies of the relationship between the characteristics of the parents and their offsprings
- Predict the average values of the response variable for each value of the predictors





Weight: the response or dependent variable

Height: the predictor or independent variable

Analysis of Variance (ANOVA)

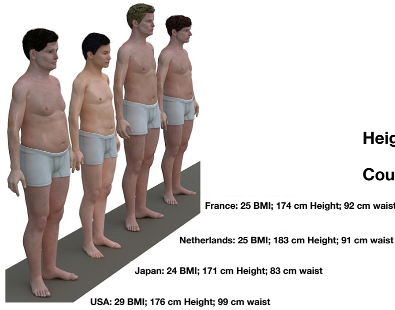
- Pioneered by Sir Ronald Fisher (1924, 1932, 1935)
  - Designed to analyse experiments where data are acquired under specific controlled conditions
- Compare the average values of the response variable between each condition (the values of the predictors)

15-1

Analysis of Variance (ANOVA)

- Pioneered by Sir Ronald Fisher (1924, 1932, 1935)
  - Designed to analyse experiments where data are acquired under specific controlled conditions
- Compare the average values of the response variable between each condition (the values of the predictors)

Source: Daily Mail, october 2nd, 2013



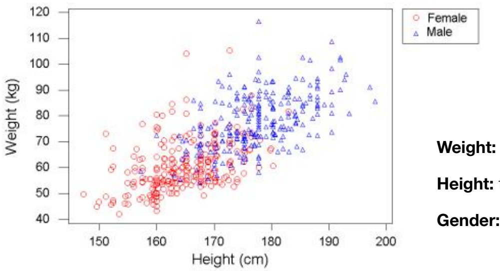
Height: the response or dependent variable

Country of Origin: the predictor or independent variable

15-2

Analysis of Covariance (ANCOVA)

- Less clear history
  - Extend ANOVA to analyse experiments where some variables can not be controlled
- Compare the average values of the response variable between conditions after removing the effect of the variables of no interest (covariates).



Heinz et al, J Statistical Education, 11(2) (2003)

Weight: the response or dependent variable

Height: the predictor or independent variable

Gender: the covariate

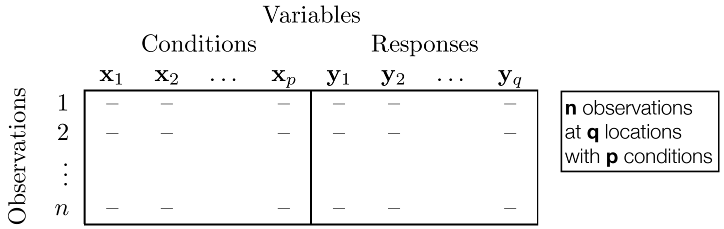
16

The conception of GLM

- In matrix algebra, the same equation can summarize all three kinds of analyses
  - All are linear models: data = model + error
- Matrix algebra provides
  - the most convenient common notation
  - a common way to implement all the analyses on computers
- Ubiquity of personal computers accelerate the adoption of the idea
  - Straightforward to implement on Matlab

17

Data for a GLM: the general structure



Wickens, Mathematics in Brain Imaging, UCLA 2004

- Observations: different time points, subjects, and/or brain locations
- Condition variables: describing groups, subjects, response conditions, or other covariates
- Response variables: observed measurements, such as brain activations

Data for a GLM: an example

t	Stim	stimulus x_0	x_1	x_2	locations y_1	y_2	y_3	y_4
1	0	1	0	0	9.45	13.25	11.23	16.48
2	0	1	0	0	9.86	10.26	11.13	13.62
3	0	1	0	0	10.17	13.90	11.74	15.13
4	1	1	1	0	12.97	11.76	10.97	16.63
5	1	1	1	0	11.31	13.83	10.65	16.42
6	1	1	1	0	12.70	10.96	10.12	17.85
7	0	1	0	0	11.38	12.95	11.15	13.65
8	0	1	0	0	10.29	12.12	11.56	15.96
9	0	1	0	0	11.82	10.29	12.73	14.27
10	2	1	0	1	10.27	12.45	14.15	19.39
11	2	1	0	1	11.54	13.25	14.33	18.49
12	2	1	0	1	8.93	8.93	14.32	16.73
13	0	1	0	0	11.01	11.69	10.40	17.31
14	0	1	0	0	8.92	11.52	10.87	14.62
15	0	1	0	0	11.04	12.85	11.09	14.00
16	2	1	0	1	9.45	11.65	15.50	17.54
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

X_0: background activity

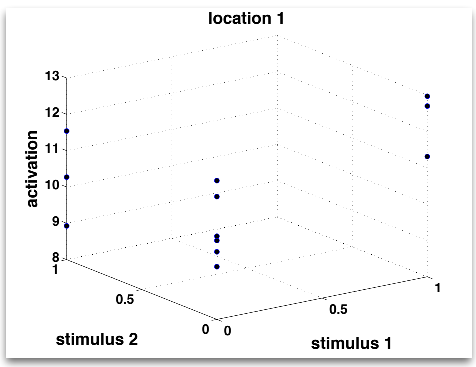
X_1: stimulus 1

X_2: stimulus 2

Wickens, Mathematics in Brain Imaging, UCLA 2004

Data for a GLM: Visualisation in variable space

Typical scatterplot



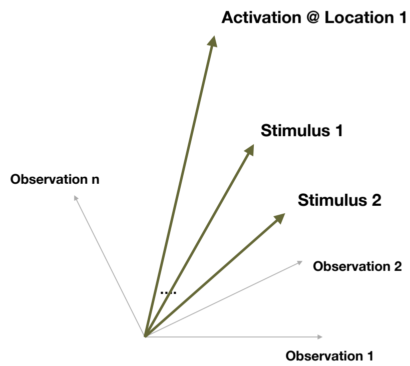
The **variable space** put the emphasis is on the **observations**

Space is defined by the **variables**:
# of axes = # of variables

Observations are **points** in the space:
# of points = # of observations

Data for a GLM: Visualisation in subject space

Subject space view



The **subject space** put the emphasis is on the **variables**

Space is defined by the **observations**:
# of axes = # of observations

Variables are **points** in the space:
# of points = # of variables

Each variable as a “spatial” vector

Length of a vector, after **centering** (removing the mean), estimates the variability of the variable

Angle between two vectors, after **centering**, equals the correlation between the variables

GLM: the definition

- A model that represents a single *response* variable by a combination of *condition* or *explanatory* variables

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_p x_{ip} + e_i$$

or

$$\mathbf{y} = \mathbf{X}\beta + \mathbf{e}$$

  - **y**: the random variable for the observations
  - **X**: the design matrix
  - **β**: the model parameters
  - **e**: the random variable for error that is independent and identically distributed, a normal distribution with zero mean and variance σ^2

GLM: the definition

- A model that represents a single *response* variable by a combination of *condition* or *explanatory* variables

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_p x_{ip} + e_i$$

or

$$\mathbf{y} = \mathbf{X}\beta + \mathbf{e}$$

- Estimates of the unobserved parameters predict measurements

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_{i1} + \hat{\beta}_2 x_{i2} + \cdots + \hat{\beta}_p x_{ip}$$

or

$$\hat{\mathbf{y}} = \mathbf{X}\hat{\beta}$$

Two Simple Examples of GLM

Bivariate Regression

	X	Y
1	1	4
2	1	7
3	3	9
4	3	12
5	4	11
6	4	12
7	5	17
8	6	13
9	6	18
10	7	17

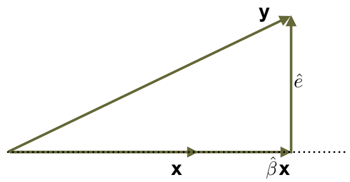
Differences between the means

	X	Y
1	-1	2
1	-1	4
2	1	5
2	1	6
2	1	6
2	1	7

Remember to centre about the mean first!

Two Simple Examples of GLM

Bivariate Regression

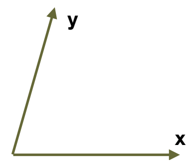


$$0 = \hat{e} \cdot x = (y - \hat{\beta}x) \cdot x$$

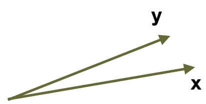
$$\hat{\beta} = \frac{y \cdot x}{x \cdot x} \qquad \hat{y} = \hat{\beta}x$$

$$\hat{e} = y - \hat{y}$$

Differences between the means



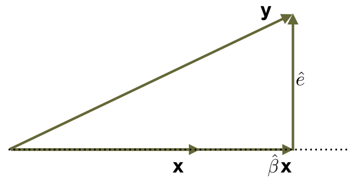
Likely with the same means



Likely with different means

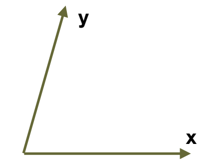
Two Simple Examples of GLM

Bivariate Regression

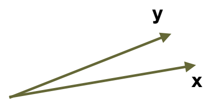


$$\hat{e}'x = 0 = (y - \hat{\beta}x)'x$$
$$\implies 0 = y'x - x'\hat{\beta}'x$$
$$\implies 0 = y'x - x'x\hat{\beta}$$
$$\hat{\beta} = \frac{x'y}{x'x} \qquad \hat{y} = \hat{\beta}x \qquad \hat{e} = y - \hat{y}$$

Differences between the means



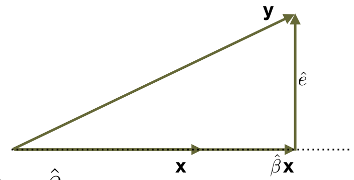
Likely with the same means



Likely with different means

Two Simple Examples of GLM

Bivariate Regression

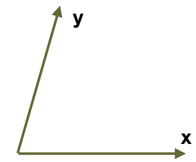


$$\hat{y} = \hat{\beta}x = \frac{x'y}{x'x}x = \frac{xx'y}{x'x} = \frac{xx'}{x'x}y$$

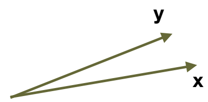
$$P_x = \frac{xx'}{x'x}$$

$$\hat{e} = y - \hat{y} = (I - P_x)y$$

Differences between the means



Likely with the same means



Likely with different means

General Linear Model (GLM)

- Model definition
- Parameter estimation
- Hypothesis testing
- Applications in Neuroimaging

Parameter estimation of GLM

- Least-square fit of the predicted measurements to the actual observations
  - Maximum-likelihood interpretation: normal distribution as the error distribution

$$P(y; \beta, \sigma^2) = \prod_{i=1}^{n} \mathcal{N}(y_i; \beta, \sigma^2)$$

$$\mathcal{L}(\boldsymbol{\beta}, \sigma^2) = \log P(\mathbf{y}; \boldsymbol{\beta}, \sigma^2).$$

  - Solved via the normal equations: compare to the bivariate regression example

$$X_i : \text{the i-th column of } X \in \mathbb{R}^{n \times p}$$

$$\sum_{i=1}^{p} X_i \alpha_i = X\alpha : \text{ any vector in the column space of X}$$

Parameter estimation of GLM

- Least-square fit of the predicted measurements to the actual observations
  - Maximum-likelihood interpretation: normal distribution as the error distribution

$$P(y;\beta,\sigma^2) = \prod_{i=1}^{n} \mathcal{N}(y_i;\beta,\sigma^2)$$

$$\mathcal{L}(\boldsymbol{\beta},\sigma^2) = \log P(\mathbf{y};\boldsymbol{\beta},\sigma^2).$$

  - Solved via the normal equations:

$$\begin{aligned} \hat{e}'(X\alpha) &= 0 = (y - X\hat{\beta})'(X\alpha) \\ \implies & 0 = (y'X - \hat{\beta}'X'X)\alpha \\ \implies & 0 = y'X - \hat{\beta}'X'X \\ \implies & 0 = X'y - X'X\hat{\beta} \end{aligned}$$

Parameter estimation of GLM

- Least-square fit of the predicted measurements to the actual observations
  - Maximum-likelihood interpretation: normal distribution as the error distribution

$$P(y;\beta,\sigma^2) = \prod_{i=1}^{n} \mathcal{N}(y_i;\beta,\sigma^2)$$

$$\mathcal{L}(\boldsymbol{\beta},\sigma^2) = \log P(\mathbf{y};\boldsymbol{\beta},\sigma^2).$$

  - Solved via the normal equations:

$$\hat{\beta} = (X'X)^{-1}X'y \qquad \hat{e} = y - X\hat{\beta} = (I - P_X)y$$

$$\hat{y} = X\hat{\beta} = X(X'X)^{-1}X'y \qquad \hat{\sigma}^2 = \frac{\|\hat{e}\|^2}{n-p}$$

$$\implies P_X = X(X'X)^{-1}X'$$

Parameter estimation of GLM

- Least-square fit of the predicted measurements to the actual observations
  - Maximum-likelihood interpretation: normal distribution as the error distribution

$$P(y;\beta,\sigma^2) = \prod_{i=1}^{n} \mathcal{N}(y_i;\beta,\sigma^2)$$

$$\mathcal{L}(\boldsymbol{\beta},\sigma^2) = \log P(\mathbf{y};\boldsymbol{\beta},\sigma^2).$$

  - Solved via the normal equations: when $r(X) < p$

$$\hat{\beta} = (X'X)^{-}X'y \qquad \hat{e} = y - X\hat{\beta} = (I - P_X)y$$

$$P_X = X(X'X)^{-}X' \qquad \hat{\sigma}^2 = \frac{\|\hat{e}\|^2}{n - r(X)}$$

An Example of Multiple Regression

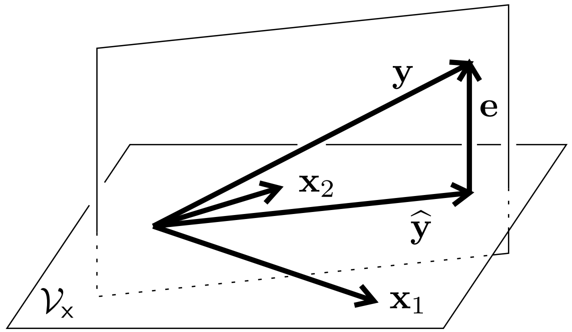
	X_1	X_2	Y
1	2	3	15
2	2	5	16
3	4	4	15
4	4	7	10
5	5	5	13
6	5	8	9
7	5	9	8
8	6	8	7
9	7	7	8
10	7	10	5
11	8	11	4

An Example of Multiple Regression

	Uncentered		
	X_1	X_2	Y
1	2	3	15
2	2	5	16
3	4	4	15
4	4	7	10
5	5	5	13
6	5	8	9
7	5	9	8
8	6	8	7
9	7	7	8
10	7	10	5
11	8	11	4

	Centered		
	X_1	X_2	Y
1	-3	-4	5
2	-3	-2	6
3	-1	-3	5
4	-1	0	0
5	0	-2	3
6	0	1	-1
7	0	2	-2
8	1	1	-3
9	2	0	-2
10	2	3	-5
11	3	4	-6

GLM: a geometric interpretation



Wickens, MBI, UCLA 2004

Parameter estimation for the toy example

	Stim	stimulus			locations			
t		x_0	x_1	x_2	y_1	y_2	y_3	y_4
1	0	1	0	0	9.45	13.25	11.23	16.48
2	0	1	0	0	9.86	10.26	11.13	13.62
3	0	1	0	0	10.17	13.90	11.74	15.13
4	1	1	1	0	12.97	11.76	10.97	16.63
5	1	1	1	0	11.31	13.83	10.65	16.42
6	1	1	1	0	12.70	10.96	10.12	17.85
7	0	1	0	0	11.38	12.95	11.15	13.65
8	0	1	0	0	10.29	12.12	11.56	15.96
9	0	1	0	0	11.82	10.29	12.73	14.27
10	2	1	0	1	10.27	12.45	14.15	19.39
11	2	1	0	1	11.54	13.25	14.33	18.49
12	2	1	0	1	8.93	8.93	14.32	16.73
13	0	1	0	0	11.01	11.69	10.40	17.31
14	0	1	0	0	8.92	11.52	10.87	14.62
15	0	1	0	0	11.04	12.85	11.09	14.00
16	2	1	0	1	9.45	11.65	15.50	17.54
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

Activation measurements at **different** locations are modelled **independently**: In other words, we fit one GLM model for each location separately.

	$\mathbf{y}_1$	$\mathbf{y}_2$	$\mathbf{y}_3$	$\mathbf{y}_4$
b_0	10.44	12.09	11.32	15.00
b_1	1.89	0.09	−0.74	1.96
b_2	−0.39	−0.52	3.25	3.04

Wickens, Mathematics in Brain Imaging, UCLA 2004

General Linear Model (GLM)

- Model definition
- Parameter estimation
- Hypothesis testing
- Applications in Neuroimaging

Hypothesis testing: some motivating examples

- Are the average height of male UCL students different from that of female students?
- Are male UCL students on average higher than their female counterparts?
- Is the variation in heights among male UCL students a lot different from that of female students?

Sampling theory: population vs sample

- We want to draw conclusions about large group of individuals
- The whole group is the *population*.
- In practice, we can not measure the entire population.
- We examine a (often small) subset of the population called a *sample.*
- Statistical inference is about inferring facts about an entire population from just a sample of the population.

Sampling theory: sampling distribution of mean

- For a normally distributed population with mean μ and variance σ^2
  - the sampling distribution of the mean is also normal
  - the distribution has the same mean μ but a different variance of σ^2/n (n is the size of the sample)
- For a general population that has well-defined mean and variance
  - the sampling distribution is asymptotically normal, a result of central limit theorem.

A primer on hypothesis testing

- The “Lady tasting tea” problem
  - The lady claims that she can tell whether the milk or tea was first added to a cup of hot water.
  - See http://en.wikipedia.org/wiki/Lady_tasting_tea

Null hypothesis and the alternative hypothesis

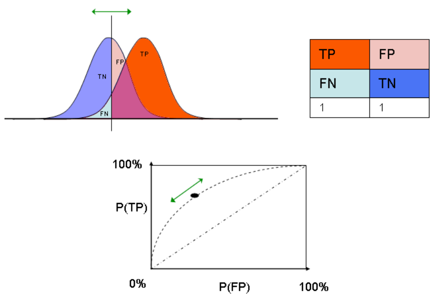
- Null hypothesis corresponds to the default situation, the converse of the situation that we look for.
  - Average heights of UCL male and female students are identical.
  - Variations in heights for UCL male and female students are identical.
- The alternative hypothesis, naturally, refers to the converse, the result we wish to confirm.

Type I and Type II errors

- Statistical terms for describing types of flaws in hypothesis testing
- Type I error: incorrectly rejecting a true null hypothesis
  - a.k.a. false-positive
- Type II error: fails to reject a false null hypothesis
  - a.k.a. false-negative

Level of significance and p-value

- Level of significance: the proportion of true null hypothesis rejected falsely
  - a.k.a. the false positive rate
  - equal to (1 - specificity)
- Related: the false negative rate
  - the proportion of false null hypothesis not rejected
  - equal to (1 - sensitivity)
- p-value: the probability that the test statistic would be equal or more extreme than the observed value under the null hypothesis.



wikipedia: ROC

Hypothesis testing: the principle

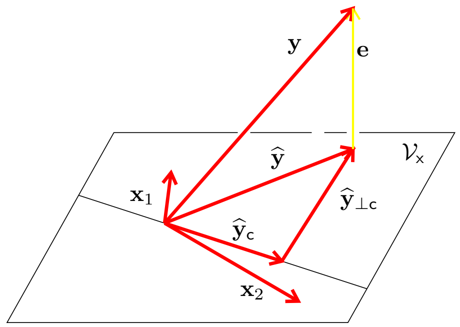
- Write down the precise hypothesis to be tested: the *null hypothesis*
- Choose the appropriate statistical model: the *test statistic*
- Choose the acceptable level of significance: Type I error
- Compute the *sampling distribution* of the relevant test statistic under the null hypothesis
- Determine the likelihood that *the current sample statistic* is drawn from the above sampling distribution
- Accept or reject the hypothesis

GLM hypothesis testing

- Requires the covariance matrix of the estimates $\mathbf{S_b} = \hat{\sigma}^2(\mathbf{X'X})^{-1}$.
- Specify *contrast of interest* by imposing linear constraints on β: $C'\beta = 0$
  - Estimable contrast must satisfy $C'\beta = P'X\beta$
- Hypothesis tested by splitting the predicted response into two orthogonal parts:
  - a part consistent with the hypothesis and a part that violates it.

$$Y \sim N(X\beta, \sigma^2 I) \qquad C'\hat{\beta} \sim N(C'\beta, \sigma^2 P'P_X P)$$
$$C'\hat{\beta} = P'X\hat{\beta} = P'P_X Y \qquad = N(C'\beta, \sigma^2 C'(X'X)^- C)$$

GLM hypothesis testing: a geometric interpretation



Wickens, MBI, UCLA 2004

Possible hypotheses using the toy example

- Null hypothesis 1: the response is not affected by the stimulus 2.

$$\mathbf{C} = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}$$

- Null hypothesis 2: the responses to both stimuli are identical.

$$\mathbf{C} = \begin{bmatrix} 0 & 1 & -1 \end{bmatrix}$$

- Null hypothesis 3: the response is not affected by both stimulus.

$$\mathbf{C} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Possible test statistics for GLM

- When the hypothesis is one-dimensional, it can be tested using a t-statistic
  - Examples: Null hypotheses 1 and 2

$$V = \mathbf{Cb} \quad \text{and} \quad s_v^2 = \mathbf{CS}_b\mathbf{C}'. \qquad t = V/s_v$$

- When the hypothesis has more than one dimension, it can be tested using a F-statistic (the number of rows in C denoted by **r**)
  - Examples: Null hypothesis 3

$$F = \frac{|\hat{\mathbf{y}}_{\perp c}|^2/r}{|\hat{\mathbf{e}}|^2/(N-p)} = \frac{MS_{\text{restriction}}}{MS_{\text{error}}}.$$

General Linear Model (GLM)

- Model definition
- Parameter estimation
- Hypothesis testing
- Applications in Neuroimaging

Issue with multiple comparisons

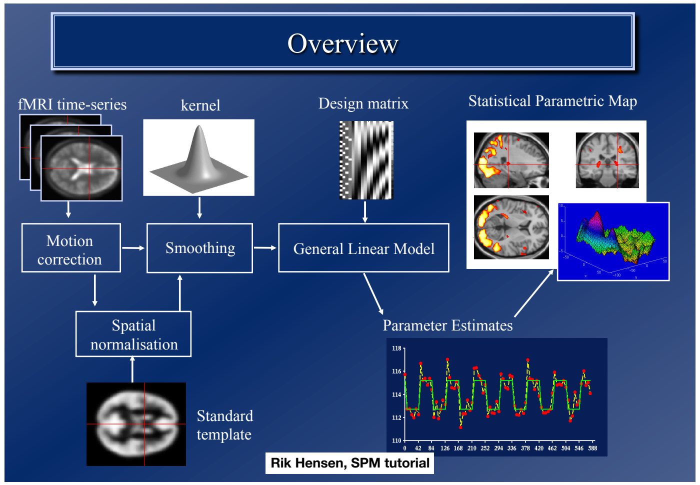
- Simultaneous testing or multiple comparisons lead to an increase in the type I error.
- If n independent comparisons are performed, the experiment-wide level of significance, a.k.a., family-wise error rate, is equal to

$$\alpha = 1 - (1 - \alpha_{\text{per comparison}})^n$$

Correction for multiple comparisons

- Bonferroni correction: α=α/n
  - Too conservative
- Gaussian random fields
  - Exploit spatial coherence in the data

Applications to NeuroImaging



References

- Thomas D. Wickens, "The General Linear Model", Mathematics in Brain Imaging, Graduate Summer School Program, UCLA, 2004

  http://www.ipam.ucla.edu/publications/mbi2004/mbi2004_5017.pdf

- Rik Hensen, "The General Linear Model and Statistical Parametric Mapping"

  http://www.mrc-cbu.cam.ac.uk/Imaging/Common/rikSPM-GLM.ppt