A Bayesian approach to False Discovery Rate and Power in Multiple Testing

Agnese Maria Di Brisco[1,3], Manuela Berlingeri[1,2], and Piero Quatto[1,3]

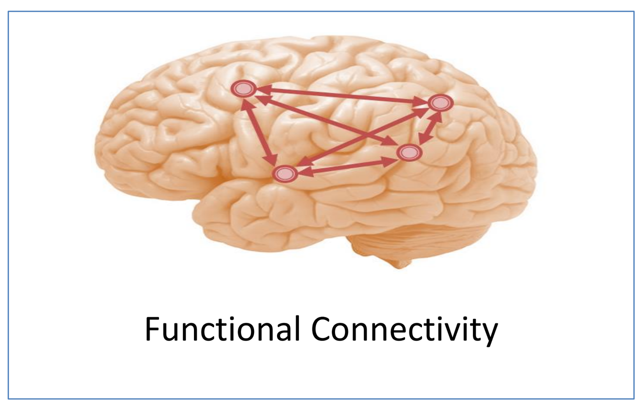
1. Milan Center for Neuroscience, Milan, Italy; 2. Psychology Department of Humanistic Studies, University of Urbino Carlo Bo, Urbino, Italy; 3. Department of Economics, Management and Statistics, University of Milano-Bicocca, Milan, Italy ;

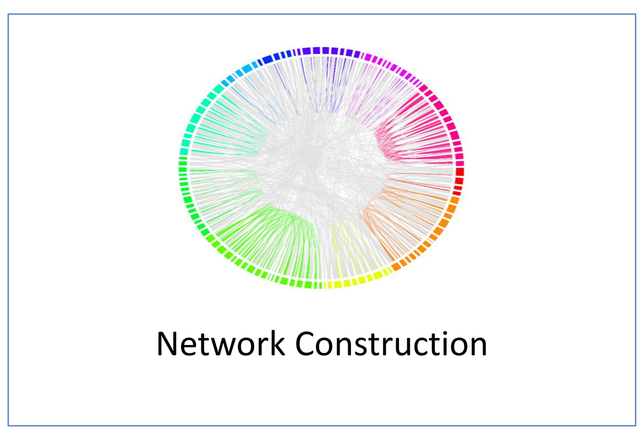
email to: agnese.dibrisco@unimib.it

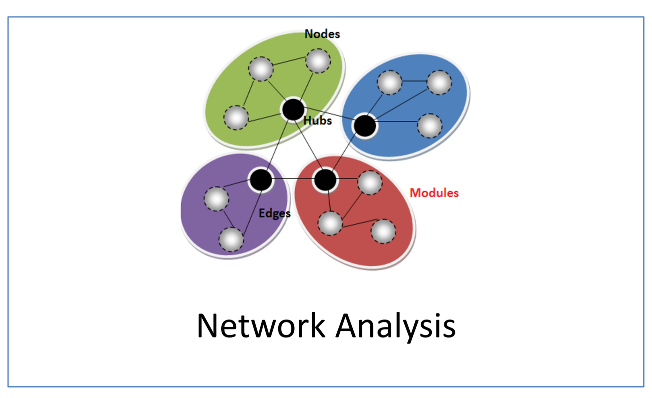
Neuromi III - Scientific Council - 13 September 2017

Functional Connectivity

The statistical analysis of fMRI is increasingly common in neuroscience. Recently, many studies on brain functional connectivity have been carried out which describe how brain regions **covariate**, both in a resting state condition or in experimental. The fMRI data analysis can be split into at lest three subsequent steps:







The pattern of functional connectivity is quantified through correlation among couples of brain regions.
Here we chose the Spearman's rank correlation coefficient since it avoids any distributional assumption on the data.
Then, the pattern of functional connectivity is summarized into a map of the brain containing only significant signals among brain regions of interest.
From a statistical point of view, the determination of significant signals results in a multiple testing problem.

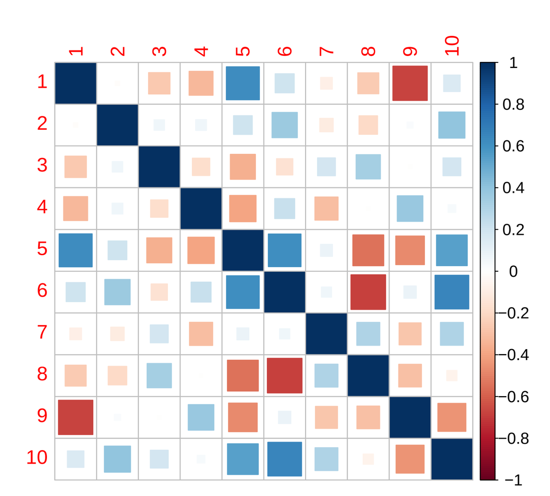


Figure: From a correlation (covariance) matrix, that captures the brain functional connectivity …

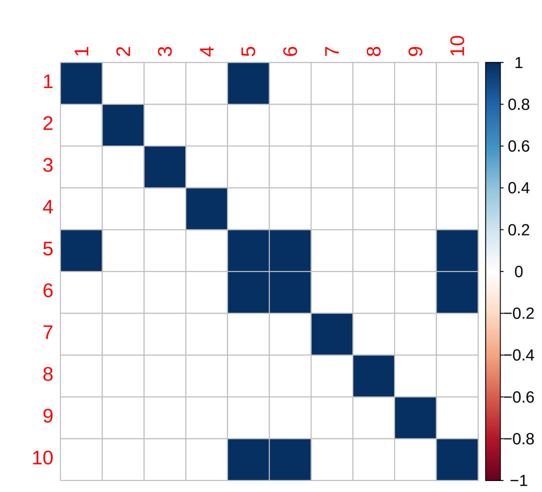


Figure: … To an adjacency matrix given by means of a multiple testing procedure

Multiple Comparisons

Let us consider m simultaneous comparisons and the corresponding Spearman's coefficients $r_1, \ldots, r_j, \ldots, r_m$; we are interested in the simultaneous testing of the following hypotheses for each correlation parameter ρ_j, for $j = 1, \ldots, m$, corresponding to r_j,

$$H_0 : \rho_j = 0 \quad vs \quad H_1 = \overline{H}_0 : \rho_j \neq 0.$$

with a priori probabilities $\pi_0 = pr(H_0)$ and $\pi_1 = pr(H_1) = 1 - \pi_0$.
The test statistic t_j is

$$t_j = \frac{r_j}{\sqrt{1 - r_j^2}} \sqrt{n-2}\,, \qquad j = 1, \ldots, m.$$

The corresponding m p-values are given by

$$p_j = 2 - 2F_{n-2}\left(|t_j|\right),$$

where F_{n-2} is the Student's cumulative distribution function (cdf) with $n-2$ degrees of freedom.
We can define the same rejection region for each test, equivalently with respect to the test statistic or its associated p-value,

$$\{|r_j| \geq \tau\} = \{p_j \leq \gamma\}$$

where $\gamma = 2 - 2F_{n-2}\left(\sqrt{\frac{(n-2)\tau^2}{1-\tau^2}}\right)$.
The τ parameter identifies the rejection threshold by which we constructed the brain network: the higher the τ, close to 1, the more the resulting brain network is conservative.
As τ increases, the number of links included in the brain network decreases.
A unified approach to identify a suitable threshold in brain networks construction with a multiple comparison perspective does not yet exist. So, we propose a method of threshold's choice that controls for the probability of false discoveries and the power of the comparisons and we provide an application of this method to real data.

Bayes False Discovery Rate (FDR) and Bayes Power (BP)

In the perspective of a brain network construction, we define the **False Discovery Rate** and the **Power** as:

$$FDR(\gamma) = pr(H_0 \,|p_j \leq \gamma) = \frac{\gamma \pi_0}{F(\gamma)},$$

$$BP(\gamma) = pr(p_j \leq \gamma \,|\overline{H}_0) = \frac{[1 - FDR(\gamma)]F(\gamma)}{1 - \pi_0},$$

where F represents the p-values cumulative distribution function.

Empirical Bayesian Estimates

The cumulative distribution function of the p-values is estimated as

$$\hat{F}(\gamma) = \frac{\#\{p_j \leq \gamma\}}{m},$$

which has expected value

$$\mathbb{E}\left[\hat{F}(\gamma)\right] = \gamma\pi_0 + BP(\gamma)\pi_1,$$

The probability π_0 is estimated conservatively through

$$\hat{\pi}_0(\lambda) = \frac{1 - \hat{F}(\lambda)}{1 - \lambda},$$

which has expected value

$$\mathbb{E}\left[\hat{\pi}_0(\lambda)\right] = \pi_0 + \frac{1 - BP(\lambda)}{1 - \lambda}\pi_1 \geq \pi_0.$$

When $\hat{\pi}_0(\lambda) = 1$, the estimate of the power is not yet well defined so we suggest $\widehat{BP}_{\lambda_2}(\gamma) = \hat{F}(\gamma)$ as a proper estimate for the power.

Balancing FDR and BP

To identify the optimal values of the tuning parameters λ_1 and λ_2 we had to:

- Bootstrap the p-values B times
- Calculate the Bootstrap versions of $\widehat{FDR}_{\lambda_1}(\gamma)$ and of $\widehat{BP}_{\lambda_2}(\gamma)$ over a range of λ values
- Choose the tuning parameters λ_1 and λ_2 which minimize the MSEs of FDR and BP, respectively:

$$\lambda_1^{opt} = \arg\min \frac{1}{B}\sum_{b=1}^{B}\left\{\widehat{FDR}^{b}_{\lambda_1}(\gamma) - \min_{\lambda_1}\widehat{FDR}_{\lambda_1}(\gamma)\right\}^2$$

$$\lambda_2^{opt} = \arg\min \frac{1}{B}\sum_{b=1}^{B}\left\{\widehat{BP}^{b}_{\lambda_2}(\gamma) - \min_{\lambda_2}\widehat{BP}_{\lambda_2}(\gamma)\right\}^2$$

Lats, we chose a suitable value of the threshold γ such that the Bayes FDR is low and the BP is reasonably high.

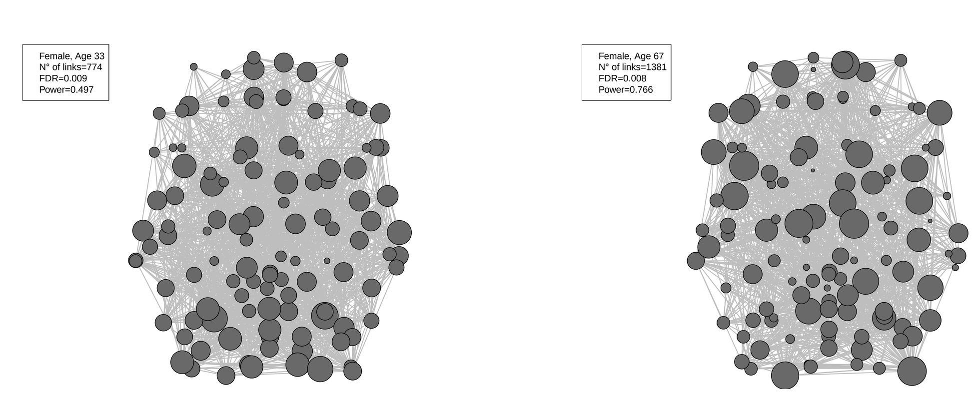
Data acquisition and preprocessing

- We acquired MRI scans of two healthy women, a 33 and a 67 years old, both with a normal cognitive profile and no evidence of medical disorders.
- The MRI acquisition was carried out by means of a 1.5 T General Electrics Scanner at the Neuroradiology Department of the Niguarda Cá Granda Hospital, Milan, Italy. The entire MRI scan lasted for 10 minutes.
- Resting-state functional images, with a voxels size of $1 \times 1 \times 1$ mm, were acquired using an Echo Planar Imaging Sequence sensitive to BOLD contrast, and afterwards were analysed using the Matlab toolbox “DPARSF-A”.
- Finally, we defined 116 non-overlapping anatomical ROIs according to the automated anatomical labelling (AAL) atlas and we extracted 116 time-series representing the mean low-frequency fluctuations.

Data analysis

We selected $\tau = 0.2$ for the elderly participant and $\tau = 0.21$ for the young one as the values which ensure the best balance between the Bayes FDR and the BP. In fact, the trade-off between the Bayes FDR and the BP implies that as the probability of false discoveries increases, the power also increases. The balanced γ is equal to the maximum value of γ value such that the Bayes FDR is smaller than the maximum error considered being acceptable. In our application, the maximum acceptable error was set equal to 0.01: if the maximum γ that guarantees $FDR \leq 0.01$ is chosen, then the BP is maximum, under such a constraint. On the contrary, if, instead of choosing the maximum γ that guarantees $FDR \leq 0.01$, we select a smaller γ, still $FDR \leq 0.01$ but the BP would be smaller.

Figure: Graphical representation in stereotactic coordinates of the brain networks of the healthy participants (woman, 33 years old (left panel) and woman, 67 years old (right panel)). Each grey circle corresponds to a ROI: the wider the diameter of the circle, the higher the degree associated to the ROI (i.e. the number of edges starting from that node); each gray line, otherwise, corresponds to a significant link between couple of ROIs.



We may compare these networks with respect to some network measures. For example, the number of links suggests a less dense network for the young participant. Since the scans have been conducted at resting state, this result is a preliminary signal of a ageing process characterized by the activation of an undifferentiated neural activity. Further studies on larger samples are necessary to confirm this hypothesis.

Simulation Study with Independent p-values

We simulated $m = 1000$ independent and normally distributed random variables $Z_i \sim N(\mu, 1)$ where $\mu = 0$ under the null hypothesis while $\mu = 2$ under the alternative hypothesis. For each i-th multiple comparison, $i = 1, \ldots, m$, we computed the associated p-value $p_i = 2 - 2\Phi(|z_i|)$, where z_i is the observed value of the i-th normal random variable. To provide Monte Carlo estimates we iterated our simulation $N = 1000$ times for $\pi_0 = 0.1, \ldots, 0.9$ and given $\gamma = 0.01$.

π_0	FDR	λ_1	$\widehat{FDR}$	MSE	BP	λ_2	$\widehat{BP}$	MSE
0.1	0.0039	0.8	0.0089	2.60e-05	0.2828	0.8	0.3271	0.0022
0.2	0.0088	0.8	0.0137	2.68e-05	0.2816	0.7	0.3311	0.0025
0.3	0.0150	0.7	0.0203	3.03e-05	0.2826	0.7	0.3307	0.0027
0.4	0.0232	0.7	0.0286	3.31e-05	0.2820	0.7	0.3310	0.0030
0.5	0.0345	0.7	0.0395	3.48e-05	0.2814	0.6	0.3382	0.0036
0.6	0.0506	0.7	0.0556	3.90e-05	0.2830	0.6	0.3370	0.0041
0.7	0.0769	0.7	0.0815	5.24e-05	0.2826	0.5	0.3474	0.0059
0.8	0.1251	0.6	0.1306	1.05e-04	0.2830	0.4	0.3599	0.0093
0.9	0.2462	0.8	0.2463	7.17e-04	0.2813	0.9	0.3275	0.0847

Simulation Studies with Dependent p-values

We assumed the set of random variables $Z_1, \ldots, Z_{1000}$ to have a multivariate normal distribution. Each marginal distribution $Z_i \sim N(\mu, 1)$ has the same variance and mean equal to 0 under the null hypothesis while equal to 2 under the alternative hypothesis.
We explored three different patterns of dependency among variables:

1. **autoregressive** pattern of dependency: each correlation among pairs of random variables equals to $Corr(Z_i, Z_j) = \rho^{|i-j|}$ and we fixed $\rho = 0.8$.
2. **unstructured** pattern of dependency: given a covariance matrix made up as block matrix with 100 sub-matrices of size 10×10 on its diagonal, the correlation's values in each sub-matrix were randomly assigned in the range $(0; 0.35)$.
3. **constant** pattern of dependency: given a covariance matrix made up as block matrix with 100 sub-matrices of size 10×10 on its diagonal, we fixed each correlation within clusters equal to the others. We simulated this covariance pattern with respect to three different constant values: $c = 0.1, 0.5, 0.9$.

Table: Simulation in case of dependent p-values given an autoregressive covariance matrix, over a range of $\pi_0 = 0.1, \ldots, 0.9$ and $\gamma = 0.1, 0.05, 0.01, 0.001$; for Bayes FDR and Power we reported the true value, the optimal values of the tuning parameters λ_1 and λ_2, the Monte Carlo Empirical Bayes estimate (indicated with hat) and the true mean square error.

π_0	FDR	λ_1	$\widehat{FDR}$	MSE	BP	λ_2	$\widehat{BP}$	MSE
0.1	0.0040	0.8	0.0092	0.000033	0.2829	0.8	0.3283	0.0024
0.2	0.0090	0.8	0.0141	0.000033	0.2809	0.7	0.3282	0.0027
0.3	0.0152	0.7	0.0204	0.000036	0.2846	0.7	0.3323	0.0030
0.4	0.0239	0.7	0.0295	0.000046	0.2801	0.7	0.3290	0.0035
0.5	0.0356	0.7	0.0410	0.000055	0.2799	0.6	0.3333	0.0044
0.6	0.0524	0.7	0.0578	0.000068	0.2819	0.6	0.3382	0.0057
0.7	0.0801	0.7	0.0858	0.000125	0.2815	0.4	0.3592	0.0105
0.8	0.1325	0.6	0.1397	0.000389	0.2806	0.3	0.3789	0.0208
0.9	0.2663	0.9	0.2737	0.003671	0.2823	0.9	0.2744	0.0686

Table: Simulation in case of dependent p-values given a hub-constant covariance matrix with the constant set equal to 0.5, over a range of $\pi_0 = 0.1, \ldots, 0.9$ and $\gamma = 0.01$; for Bayes FDR and power we reported the true value, the optimal values of the tuning parameters λ_1 and λ_2, the Monte Carlo Empirical Bayes estimate (indicated with hat) and the true mean square error.

π_0	FDR	λ_1	$\widehat{FDR}$	MSE	BP	λ_2	$\widehat{BP}$	MSE
0.1	0.0040	0.8	0.0090	0.000029	0.2827	0.8	0.3276	0.0023
0.2	0.0089	0.8	0.0140	0.000031	0.2819	0.7	0.3296	0.0026
0.3	0.0151	0.7	0.0205	0.000035	0.2835	0.7	0.3321	0.0029
0.4	0.0235	0.7	0.0289	0.000039	0.2817	0.7	0.3313	0.0033
0.5	0.0348	0.7	0.0402	0.000044	0.2824	0.6	0.3362	0.0039
0.6	0.0518	0.7	0.0571	0.000055	0.2817	0.5	0.3411	0.0049
0.7	0.0784	0.7	0.0837	0.000093	0.2832	0.4	0.3576	0.0086
0.8	0.1294	0.6	0.1357	0.000245	0.2825	0.3	0.3805	0.0161
0.9	0.2535	0.9	0.2583	0.001999	0.2866	0.9	0.3001	0.0767

Table: Simulation in case of dependent p-values given a hub-unstructured covariance matrix, over a range of $\pi_0 = 0.1, \ldots, 0.9$ and $\gamma = 0.1, 0.05, 0.01, 0.001$; for Bayes FDR and power we reported the true value, the optimal values of the tuning parameters λ_1 and λ_2, the Monte Carlo Empirical Bayes estimate (indicated with hat) and the true mean square error.

π_0	FDR	λ_1	$\widehat{FDR}$	MSE	BP	λ_2	$\widehat{BP}$	MSE
0.1	0.0039	0.8	0.0090	0.000028	0.2829	0.7	0.3310	0.0025
0.2	0.0088	0.8	0.0136	0.000027	0.2835	0.7	0.3303	0.0024
0.3	0.0151	0.7	0.0203	0.000031	0.2820	0.7	0.3296	0.0026
0.4	0.0232	0.7	0.0285	0.000034	0.2824	0.7	0.3319	0.0030
0.5	0.0344	0.7	0.0396	0.000035	0.2831	0.6	0.3368	0.0035
0.6	0.0512	0.7	0.0565	0.000044	0.2813	0.6	0.3368	0.0044
0.7	0.0773	0.6	0.0829	0.000061	0.2831	0.5	0.3469	0.0059
0.8	0.1254	0.6	0.1303	0.000118	0.2851	0.4	0.3612	0.0097
0.9	0.2505	0.8	0.2561	0.001005	0.2804	0.9	0.2910	0.0723

All simulation results provided empirical evidence supporting the robustness of the proposed estimates.

References

- Efron, B. (2010) Large-scale inference: empirical Bayes methods for estimation, testing, and prediction. Cambridge University Press.
- Kolaczyk, E.D., Csárdi, G. (2014) Statistical analysis of network data with R. Springer.
- Sala, S., Quatto, P., Valsasina, P., Agosta, F., Filippi, M. (2014) pFDR and pFNR estimation for brain networks construction. Stat Med, 33:158-69.
- Sokal, R.R., Rohlf, F.J. (1995) Biometry (3rd edn). WH Freman and company: New York.
- Storey, J.D. (2002) A direct approach to false discovery rates. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 64:479-498.