High-Dimensional Data Analysis and Statistical Inference **June 6, 2015**

Homework 3: large-scale hypothesis testing

Student: Jun Yu *ID : 1401110054*

1. *Multiple testing:* In a multiple testing situation with both N_0 and N_1 positive, show that

$$E(\frac{a}{N_0}) = \bar{\alpha} \qquad \text{and} \qquad E(\frac{b}{N_1}) = \bar{\beta}$$

where $\bar{\alpha}$ and $\bar{\beta}$ are the average size and power of the null and non-null cases, respectively.
Solution:
Note that

$$a = \sum_{i=1}^{N} 1_{(a_i=1|N_{0i}=1)} \qquad \text{and} \qquad b = \sum_{i=1}^{N} (b_i = 1|N_{1i} = 1)$$

where a_i, b_i, N_{0i}, N_{1i} are the same mean as a, b, N_0, N_1 in fig 4.1 for the test problem H_{0i}.

Note the facts, for the classical single-case testing situation, we have

$$\mathrm{P}(a = 1|N_0 = 1) = \alpha \qquad \text{and} \qquad P(b = 1|N_1 = 1) = \beta,$$

the result is obvious.

2. *Benjamini - Hochberg procedure:* Verify (4.19) in Efron's book:

$$E\{A(s)|A(t)\} = A(t) \text{ for } s \le t.$$

Solution:
Denote $\mathcal{F}_s = \sigma\{a(s), R(s), s \le 1\}$, we have

$$E\{A(s)|A(t)\} = E\{\frac{a(s)}{s}|\mathcal{F}_t\} = \frac{1}{s}E\{a(s)|\mathcal{F}_t\} = \frac{1}{s}a(t)\frac{s}{t} = \frac{a(t)}{t}$$

where we used that under $\mathcal{F}_s, a(s) = \#\{p_i^0 : p_i^0 \le s\}$ where p_i^0 stands for the true hypothesi's p-value and these p_i^0 are uniformly distributed on $[0, t]$ and are independent.

3. *Two-sample mean test:* Consider the test statistic T_n in Chen and Qin (2010),Show that under H_1 and the condition in (3.4),

$$\mathrm{Var}(T_n) = \{\frac{2}{n_1(n_1-1)}\mathrm{tr}(\Sigma_1^2) + \frac{2}{n_2(n_2-1)}\mathrm{tr}(\Sigma_2^2) + \frac{4}{n_1 n_2}\mathrm{tr}(\Sigma_1\Sigma_2)\}\{1 + o(1)\}$$

where the o(1) term vanishes under H_0.

Solution:
Denote $P_1 = \frac{\sum_{i\neq j}^{n_1} X_{1i}' X_{1j}}{n_1(n_1-1)}$, $P_2 = \frac{\sum_{i\neq j}^{n_2} X_{2i}' X_{2j}}{n_2(n_2-1)}$, and $P_3 = -2\frac{\sum_{i=1}^{n_1}\sum_{j=1}^{n_2} X_{1i}' X_{2j}}{n_1 n_2}$. It can be shown that

$$\begin{aligned}
\mathrm{Var}(P_1) &= \frac{2}{n_1(n_1-1)}\mathrm{tr}(\Sigma_1^2) + \frac{4\mu_1'\Sigma_1\mu_1}{n_1}, \\
\mathrm{Var}(P_2) &= \frac{2}{n_2(n_2-1)}\mathrm{tr}(\Sigma_2^2) + \frac{4\mu_2'\Sigma_2\mu_2}{n_2}, \\
\mathrm{Var}(P_3) &= \frac{4}{n_1 n_2}\mathrm{tr}(\Sigma_1\Sigma_2) + \frac{4\mu_1'\Sigma_1\mu_1}{n_1} + \frac{4\mu_2'\Sigma_2\mu_2}{n_2}.
\end{aligned}$$

Because the two samples are independent, $\mathrm{Cov}(P_1, P_2) = 0$. Also,

$$\mathrm{Cov}(P_1, P_3) = \frac{4\mu_1'\Sigma_1\mu_2}{n_1} \text{ and } \mathrm{Cov}(P_2, P_3) = \frac{4\mu_1'\Sigma_2\mu_2}{n_2}.$$

In summary,

$$\begin{aligned}
\mathrm{Var}(T_n) = &\frac{2}{n_1(n_1-1)}\mathrm{tr}(\Sigma_1^2) + \frac{2}{n_2(n_2-1)}\mathrm{tr}(\Sigma_2^2) + \frac{4}{n_1 n_2}\mathrm{tr}(\Sigma_1\Sigma_2) \\
&+ \frac{4(\mu_1-\mu_2)'\Sigma_1(\mu_1-\mu_2)}{n_1} + \frac{4(\mu_1-\mu_2)'\Sigma_2(\mu_1-\mu_2)}{n_2}.
\end{aligned}$$

Hence we get the desired results.

4. *Scaled Lasso:* Show that the loss function:

$$L_{\lambda_0}(\boldsymbol{\beta}, \sigma) = \frac{\|\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta}\|_2^2}{2n\sigma} + \frac{\sigma}{2} + \lambda_0\|\boldsymbol{\beta}\|_1$$

is jointly convex in $(\boldsymbol{\beta}, \sigma)$.

Solution:
For $0 \leq \phi \leq 1$, we have

$$\begin{aligned}
L_{\lambda_0}(\boldsymbol{\phi\beta + (1-\phi)\beta}, \phi\sigma + (1-\phi)\sigma) \leq\ & \frac{\|\boldsymbol{y} - \boldsymbol{X\phi\beta}\|_2^2}{2n\sigma} + \frac{\phi\sigma}{2} + \lambda_0\|\boldsymbol{\phi\beta}\|_1 \\
&+ \frac{\|\boldsymbol{y} - \boldsymbol{X(1-\phi)\beta}\|_2^2}{2n\sigma} + \frac{(1-\phi)\sigma}{2} + \lambda_0\|\boldsymbol{(1-\phi)\beta}\|_1 \\
\leq\ & \frac{\|\boldsymbol{y} - \boldsymbol{X\phi\beta}\|_2^2}{2n\phi\sigma} + \frac{\phi\sigma}{2} + \lambda_0\|\boldsymbol{\phi\beta}\|_1 \\
&+ \frac{\|\boldsymbol{y} - \boldsymbol{X(1-\phi)\beta}\|_2^2}{2n(1-\phi)\sigma} + \frac{(1-\phi)\sigma}{2} + \lambda_0\|\boldsymbol{(1-\phi)\beta}\|_1.
\end{aligned}$$

The first inequality due to the convexity of the function $L_{\lambda_0,\sigma}(\boldsymbol{\beta}) = \frac{\|\boldsymbol{y}-\boldsymbol{X\beta}\|_2^2}{2n\sigma} + \lambda_0\|\boldsymbol{\beta}\|_1$. The desired results follows by the definition.