Multiple testing and false discovery

Statistical modelling: theory and practice

Gilles Guillot

gigu@dtu.dk

October 2, 2013

Introductory example: testing $\mu = 0$ for an i.i.d normal sample

False discovery for a two-hypothesis problem

- For a test at level α with test statistic T, the threshold is chosen so that $P(|T| > t) = \alpha$. This is the type I error rate.
- If two hypotheses $H_o^{(1)}$ and $H_o^{(2)}$ are tested with test statistics $T^{(1)}$ and $T^{(2)}$, a wrong conclusion is drawn if $|T^{(1)}| > t$ or $|T^{(2)}| > t$. This occurs with probability

$$
\begin{aligned}
P(|T^{(1)}| > t \cup |T^{(2)}| > t) &= P(|T^{(1)}| > t) + P(|T^{(2)}| > t) \\
&\quad - P(|T^{(1)}| > t \cap |T^{(2)}| > t)
\end{aligned}
$$

If $T^{(1)}$ and $T^{(2)}$ are independent,
$P(|T^{(1)}| > t \cup |T^{(2)}| > t) = \alpha + \alpha - \alpha^2 \approx 2\alpha$.

Multiple testing in large dimension

Two-hypothesis testing problem:

Carrying out two tests at level α, the "global" type I error rate can be twice larger than expected.

A p-hypothesis testing problem:

Carrying out p tests at level α, we almost surely end up with at least one wrong conclusion (if p is large enough).

A placebo cures at least one disease if one make enough clinical trials!

The Bonferonni correction

For p hypotheses

$$
\begin{aligned}
P(\bigcup_k |T^{(k)}| > t) &\leq \sum_k P(|T^{(k)}| > t)) \\
&= pP(|T^{(1)}| > t)
\end{aligned}
$$

If t is chosen so that $P(|T^{(1)}| > t) = \alpha/p$ then $P(\bigcup_k |T^{(k)}| > t) \leq \alpha$.

- This is the Bonferroni correction. It is quick and easy rule to control false discovery.
- It is used e.g. when one has to make a large number of simple linear regressions or one-way ANOVA
- It is very "conservative" i.e. it tends to conserve H_o too often.
- There are many other methods (e.g. Holm, Benjamini and Hochberg)