Analysis of Microarray Data with Methods from Machine Learning and Network Theory

Summer Lecture 2015

Prof. Dr. A. B. Cremers
Dr. Jörg Zimmermann

Hypotheses Tests

Setting: You have a working hypothesis and you want to know if your data support your hypothesis or not.

Example: tea tasting Lady

A Lady claims she can distinguish between two cases

1. tea with milk, but pouring milk into the cup first
2. pouring first tea and then milk into the cup

Hypotheses Tests

Model this testing situation as a Bernoulli-Experiment:

1. What is the probability to observe *x* positive outcomes in n trials?

$$b_{n,p}(x) = \binom{n}{x} p^x (1-p)^{n-x}$$

2. What is the probability to observe *at least t* positive outcomes in n trials?

$$b_{n,p}(X \geq t) = \sum_{x=t}^{n} \binom{n}{x} p^x (1-p)^{n-x}$$

Hypotheses Tests

Null hypothesis: The lady can't distinguish the different cups of tea

This means: p = 1/2

What is the probability to observe at least *t* positive outcomes given the nullhypothesis?

$$b_{n,p}(X \geq t) = \sum_{x=t}^{n} \binom{n}{x} p^x (1-p)^{n-x} = \left(\frac{1}{2}\right)^n \sum_{x=t}^{n} \binom{n}{x}$$

Hypotheses Tests

One-sided hypothesis test:

Choose the smallest value t, so that the inequality:

$$b_{n,p}(X \geq t) \leq \alpha$$

is valid. A common value for α is 0.05.

For n = 10, the critical value is t = 9, for n = 20, we find t = 15.

Hypotheses Tests

If the observed positive outcomes are greater or equal to the critical *t*, we say:

The null hypothesis is rejected at the significance niveau α.

So in the case of our tea tasting lady, we would have the following experimental setup:

1. We choose α = 0.05 as significance niveau
2. We conduct n = 10 (n = 20) experiments
3. If the lady has classified at least 9 (15) cups of tea correct, we would conclude that she has the claimed ability at a significance niveau of 0.05.
4. Otherwise we would reject her claim.

p-values

If two experimenters use different values as significance niveaus, say 0.05 and 0.01, then the same experimental outcome could lead to rejection of the null-hypothesis in one case and acceptance (better: non-rejection) in the other case.

If a concrete result of an experiment is available, one could ask what the critical significance level is:

The critical significance level is the least value which allows a rejection of the null hypothesis.

This critical level is also called the observed significance niveau or p-value.

p-values

In the case of our tea tasting lady, the p-value for $n=20$ and

$x=14$	is	p-value = 0.0577
$x=15$	is	p-value = 0.0207
$x=16$	is	p-value = 0.0059
$x=17$	is	p-value = 0.0013
$x=18$	is	p-value = 0.0002

We see that in fact the observed significance niveau in the case $x=15$ is much better than 0.05.

p-values are widely used in reporting statistical results. Generally, a p-value less than 0.05 is regarded as a significant result.

q-values

The just considered test design is adequate in the case of testing only a few hypothesis. But if one tests a large number of hypothesis simultaneously one has to consider the possibility of false positives explicitly.

Imagine you are testing 100 null hypothesis, and all are true (but you don't know that). Then by pure chance you will get around 5 significant results, i.e. you would reject in ca. 5 cases the true null hypothesis, which is a false positive finding.

q-values

The most straightforward way to deal with this situation is the so-called Bonferroni correction:

If you are testing m independent hypotheses, reduce the local significance niveau to α/m. Then the global significance niveau, i.e. the probability of having a false positive, is (less or equal) than α.

The Bonferroni-correction assures you of an acceptable global significance level, but for the price of extremely tight local significance levels in the case of many hypotheses. This could lead to an increased number of false negatives, i.e. wrong null hypotheses would not be rejected.

q-values

A new concept to deal with this problem is the false discovery rate or q-value.

The new concept states that not the absolute number of false positives is most important, but the number of false positives relative to all positive results.

Let V be the number of false positives and S be the number of true positives. Then we want to control the following ratio:

$$V/(V+S) \leq q$$

q-values

1. Let H_1, .., H_m be the null hypotheses and p_1,..,p_m their corresponding p-values.

2. Order these values in increasing order and denote them by $p_{(1)}$,..,$p_{(m)}$.

3. For a given α, find the largest *k* such that:

$$P_{(k)} \leq \frac{k}{m}\alpha$$

4. Reject (i.e. declare positive) all $H_{(i)}$ for i=1,..,k.

q-values

Note that this will lead in general to an expected local significance niveau of:

$$\frac{\alpha(m+1)}{2m}$$

where m is the number of hypotheses and α the global significance niveau. For large m this value is significantly larger then the Bonferroni correction.

This value is also called false discovery rate adjusted α for m independent tests.