How to perform statistical analysis?

Hypothesis testing

Marie Tremblay-Franco et Gildas Le Corguillé

31/05/2017 v 1.0.0
The univariate modules

- The "Univariate" and « Anova (N-way)" modules allow you to perform:
  - Student / Wilcoxon test
  - N-ways ANOVA / Kruskal-Wallis test
  - Pearson / Spearman correlation test

- It is available in the "Statistical Analysis" sections of LC-MS, GC-MS, and NMR
THEORY
**INTRODUCTION**

- **Hypothesis testing**: statistical method used to take a decision about the value of a parameter of interest (mean, proportion, correlation, probability distribution, ...), based on a sample.

- **Null hypothesis $H_0$**: prior value of the parameter. Usually no effect, no difference, no link.

- **Test statistic ($Z$)**: variable used to take a decision = reject $H_0$ or not
  - Is the observed difference due to hazard (sampling fluctuations) or not (*significant*)?

- **Inferential method**: extrapolation from sample to population.
## ERROR TYPES

<table>
<thead>
<tr>
<th>DECISION</th>
<th>REALITY (unknown)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$H_0$ TRUE</td>
</tr>
<tr>
<td>NO REJECT</td>
<td>No error</td>
</tr>
<tr>
<td>REJECT</td>
<td>No error</td>
</tr>
</tbody>
</table>
### ERROR TYPES

<table>
<thead>
<tr>
<th>DECISION</th>
<th>REALITY (unknown)</th>
</tr>
</thead>
<tbody>
<tr>
<td>H₀ TRUE</td>
<td>H₀ FALSE</td>
</tr>
<tr>
<td>NO REJECT</td>
<td>No error</td>
</tr>
<tr>
<td>REJECT</td>
<td>α</td>
</tr>
<tr>
<td></td>
<td>No error</td>
</tr>
</tbody>
</table>

- **Two error types:**
  - **Type I error**: asserting a difference which doesn’t exist! (false positive)
  - \( \alpha = \Pr[]{\text{reject } H₀ \mid H₀ \text{ true}} \)
ERROR TYPES

<table>
<thead>
<tr>
<th>DECISION</th>
<th>REALITY (unknown)</th>
</tr>
</thead>
<tbody>
<tr>
<td>H₀ TRUE</td>
<td>H₀ FALSE</td>
</tr>
<tr>
<td>NO REJECT</td>
<td>No error</td>
</tr>
<tr>
<td>REJECT</td>
<td>α</td>
</tr>
</tbody>
</table>

- α = P[Reject H₀ | H₀ true]
- α = significance threshold, generally 5%

Two error types:

- **Type I error**: asserting a difference which doesn’t exist! (false positive)

  - α = P[Reject H₀ | H₀ true]
ERROR TYPES

<table>
<thead>
<tr>
<th>DECISION</th>
<th>REALITY (unknown)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( H_0 ) TRUE</td>
</tr>
<tr>
<td>NO REJECT</td>
<td>No error</td>
</tr>
<tr>
<td>REJECT</td>
<td>( \alpha )</td>
</tr>
</tbody>
</table>

- Two error types:
  - **Type I error**: asserting a difference which doesn’t exist! (false positive)
    - \( \alpha = P[\text{Reject } H_0 \mid H_0 \text{ true}] \)
    - \( \alpha \): significance threshold, generally 5%
  - **Type II error**: failing to assert a difference which exists! (false negative)
    - \( \beta = P[ \mid H_0 \text{ false}] \)
**ERROR TYPES**

<table>
<thead>
<tr>
<th>DECISION</th>
<th>REALITY (unknown)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>H₀ TRUE</td>
<td>H₀ FALSE</td>
</tr>
<tr>
<td>NO REJECT</td>
<td>No error</td>
</tr>
<tr>
<td>REJECT</td>
<td>α</td>
</tr>
</tbody>
</table>

- **Two error types:**
  - **Type I error:** asserting a difference which doesn’t exist! (false positive)
    - \( \alpha = P[\text{Reject } H₀ \mid H₀ \text{ true}] \)
    - \( \alpha \) = significance threshold, generally 5%
  - **Type II error:** failing to assert a difference which exists! (false negative)
    - \( \beta = P[\text{No reject } H₀ \mid H₀ \text{ false}] \)
    - **Power lack**
• Probability when $H_0$ is false
• Probability when $H_0$ is false to reject $H_0$ (true positive)

$$\pi = P[\text{Reject } H_0 \mid H_0 \text{ false}] = 1 - \beta$$

• Depend on the sample size: $n \uparrow \Rightarrow \pi \uparrow \Rightarrow \beta \downarrow$
P-VALUE (1)

- **p-value** = statistical significance of evidence (quantifies the confidence you can have in your decision)
  - Probability of observing under $H_0$ a result equal to or more extreme than what was actually observed on the collected sample
  - The smaller the $p$-value, the larger the significance (error is less likely)
P-VALUE (2)

Observed value

\[ q_{\alpha/2} \quad q_{1-\alpha/2} \]

Observed value

\[ p \]
P-VALUE (3)

- **p-value** = statistical significance of evidence (quantifies the confidence you can have in your decision)
  - Probability of observing under $H_0$ a result equal to or more extreme than what was actually observed on the collected sample
  - The smaller the $p$-value, the larger the significance (error is less likely)
  - Reject $H_0$ when $p$-value $< \alpha$
  - $\triangledown p \neq P[H_0$ vraie$]$
STUDENT TEST (1)

- Used to compare mean value of the variable of interest for 2 populations

- Null Hypothesis $H_0 : \mu_1 = \mu_2$
  - Ex: Is brain weight identical for animals treated with 0.025 μg and 0.25 μg BPA?

- Test statistic
  \[
  z = \frac{m_1 - m_2}{\sqrt{s^2 \left( \frac{1}{n_1} + \frac{1}{n_2} \right)}}
  \]
  \[
  s^2 = \frac{(n_1 - 1)s_1^2 - (n_2 - 1)s_2^2}{n_1 + n_2 - 2}
  \]
• \( H_0 \) is rejected if \(|z_{\text{obs}}| > t_{(n_1+n_2-2);\alpha/2}\)

• Assumptions
  • Variable of interest is normally distributed
  • Equality of sample variances
ONE WAY ANALYSIS OF VARIANCE (1)

- Comparison of means of 3 or more populations: link between a qualitative factor (treatment dose, diet, ...) and a quantitative variable (metabolite intensity, ...)
  - Is brain weight identical for animals treated with DMSO (vehicle), 0.025 µg, 0.25 µg or 25µg BPA?

- Student test generalization

- Hypothesis
  - $H_0 : \mu_1 = \mu_2 = \ldots = \mu_G$
  - $H_1 : \text{at least 2 group means are different: } \mu_d \neq \mu_k$

- Assumptions
  - Variable of interest is normally distributed
  - Sample variances are equal
ONE-WAY ANALYSIS OF VARIANCE (2)

- Decomposition of total variability (SST)

\[ SST = \sum_{g=1}^{G} \sum_{i=1}^{n_g} (y_{gi} - \mu)^2 \]

\(\text{SST} = \text{distance between global mean and all observations (regardless of the level of qualitative variable)}\)
ONE-WAY ANALYSIS OF VARIANCE (3)

- Decomposition of total variability (SST)
  - SST = *Within group variability* (SSW)

\[
\sum_{g=1}^{G} \sum_{i=1}^{n_g} (y_{gi} - \mu_g)^2
\]

Biological variable

Group 1, Group 2, Group 3

Qualitative factor levels
ONE-WAY ANALYSIS OF VARIANCE (4)

- Decomposition of total variability (SST)
  - SST = Within group variability (SSW) + Between group variability (SSB)

**Biological variable**

\[ \sum_{g=1}^{G} n_g (\mu_g - \mu)^2 \]

SSB = Distance between global mean and group means
ONE-WAY ANALYSIS OF VARIANCE (6)

- ANOVA Table

<table>
<thead>
<tr>
<th>Variability</th>
<th>Sum of Squares</th>
<th>Degrees of freedom</th>
<th>Mean Squares</th>
<th>Fisher statistic test</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between</td>
<td>SSB</td>
<td>G-1</td>
<td>SSB / (G-1)</td>
<td>$F = \frac{SSB}{SSW} \cdot \frac{n-G}{G-1}$</td>
<td></td>
</tr>
<tr>
<td>Within</td>
<td>SSW</td>
<td>n-G</td>
<td>SSW / (N-G)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>SST</td>
<td>n-1</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- $H_0$ is rejected if $F_{obs} > F_{n-G;G-1}(\alpha)$: $\alpha$-order quantile of Fisher distribution with n-G et G-1 degrees of freedom

$\Rightarrow H_0$ rejected if total variability $>>$ within variability

- $H_0$ no rejected if total variability $\approx$ within variability
N-WAY ANOVA (1)

<table>
<thead>
<tr>
<th>SampleName</th>
<th>Hour</th>
<th>Treatment</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>T0_10_M</td>
<td>T0</td>
<td>10</td>
<td>85.10</td>
</tr>
<tr>
<td>T5_10_M</td>
<td>T5</td>
<td>10</td>
<td>87.60</td>
</tr>
<tr>
<td>T0_15_M</td>
<td>T0</td>
<td>15</td>
<td>84.20</td>
</tr>
<tr>
<td>T6_15_M</td>
<td>T5</td>
<td>15</td>
<td>75.90</td>
</tr>
</tbody>
</table>

- Analysis of variance:
  \[ \text{SST} = \text{SSB} + (\text{SSW}_1 + \text{SSW}_2 + \text{SSW}_{ht}) \]
N-WAY ANOVA (2)

• Several null hypothesis are tested in parallel
  – Relative to Factor 1 (here Hour):
    • $H_0$: "$\mu_{1i} = \mu_{1j} = \ldots$"
    • $H_1$: "there is at least 1 average of the F1 different from other"
  – Relative to Factor 2 (here Treatment):
    • $H_0$: "$\mu_{2i} = \mu_{2j} = \ldots$"
    • $H_1$: "there is at least 1 average of the F2 different from other"
  – Relative to the interaction between the 2 Factors:
    • $H_0$: "there is not any interaction between factors 1 and 2"
    • $H_1$: "there is interactions between factors 1 and 2"
N-WAY ANOVA (3)

- Interaction

Lines representing the two stimulations are parallel

Lines representing the two stimulations are not parallel

Lines representing the two stimulations are not parallel + crossing lines
N-WAY ANOVA (4)

<table>
<thead>
<tr>
<th>SampleName</th>
<th>Hour</th>
<th>Treatment</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>T0_10_M</td>
<td>T0</td>
<td>10</td>
<td>85.10</td>
</tr>
<tr>
<td>T5_10_M</td>
<td>T5</td>
<td>10</td>
<td>87.60</td>
</tr>
<tr>
<td>T0_15_M</td>
<td>T0</td>
<td>15</td>
<td>84.20</td>
</tr>
<tr>
<td>T6_15_M</td>
<td>T5</td>
<td>15</td>
<td>75.90</td>
</tr>
</tbody>
</table>

\[
u \frac{1}{\sum (x_i - \bar{x})^2} = \frac{1}{\sum (y_i - \bar{y})^2} = \sum \mu_{ij} = \sum \epsilon_{ijk}
\]

\[
Y_{ijk} = \mu_{ij} + \epsilon_{ijk} \quad \text{with} \quad \epsilon_{ijk} \sim \mathcal{N}(0, \sigma^2)
\]

\[
P(X \leq x, Y \leq y | Z = z) = P(X \leq x | Z = z) \cdot P(Y \leq y | Z = z)
\]

<table>
<thead>
<tr>
<th></th>
<th>Df</th>
<th>Sum Sq</th>
<th>Mean Sq</th>
<th>F value</th>
<th>Pr(&gt;F)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hour</td>
<td>2</td>
<td>31.1</td>
<td>15.55</td>
<td>5.923</td>
<td>0.005</td>
</tr>
<tr>
<td>Treatment</td>
<td>1</td>
<td>2.017</td>
<td>2.017</td>
<td>0.3772</td>
<td>0.54</td>
</tr>
<tr>
<td>Hour:Treatment</td>
<td>1</td>
<td>63.27</td>
<td>84.20</td>
<td>11.83</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Residuals</td>
<td>10</td>
<td>5.346</td>
<td>0.5346</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
PEARSON TEST (1)

• Used to test dependence between two variables $X$ and $Y$

• Null Hypothesis $H_0$: $\rho = 0$ (independence)
  
  • Ex: are intensities of buckets 3.01ppm and 1.90ppm correlated?

• Test statistic
  
  $$ Z = \frac{r}{\sqrt{(1-r^2)/(n-2)}}, \text{ with } r = \frac{\sum_{i=1}^{n} (x_i - \mu_x)(y_i - \mu_y)}{\sigma_x \sigma_y} $$

• Under $H_0$, $Z \sim t$ with n-2 degrees of freedom
PEARSON TEST (2)

- $H_0$ is rejected if $z_{obs} > t_{n-2;1-\alpha/2}$
- ! Pearson correlation = linear relationship
  - Does not measure non linear relationship
WHEN ASSUMPTIONS ARE NOT SATISFIED…

- Previous tests = parametric tests
  - Assumption about observation probability distribution (normal distribution, variance equality, ...)
- When assumptions are not satisfied?
  - Data transformation for normality
WHEN ASSUMPTIONS ARE NOT SATISFIED...

- Previous tests = parametric tests
  - Assumption about observation probability distribution (normal distribution, variance equality, ...)

- When assumptions are not satisfied?
  - If any transformation works, parametric tests are unusable

- Non parametric tests
  - No assumption about probability distribution, variance, ...
  - Based on observation ranks
MANN WHITNEY (1)

- Used to compare **probability distribution** of a quantitative variable observed on two samples

- Null hypothesis
  - $H_0$: the 2 samples come from the same population
  - $H_1$: the 2 samples are not from the same population (one sample tends to have larger values) = at least, sample medians are not equal

- Test statistic
  \[
  U_{n_1} = n_1 n_2 + \frac{n_1(n_1 + 1)}{2} - R_{n_1}, \quad U_{n_2} = n_1 n_2 + \frac{n_2(n_2 + 1)}{2} - R_{n_2} = n_1 n_2 - U_{n_1}
  \]
  \[
  U_{n_1,n_2} = \min(U_{n_1,n_2})
  \]
MANN WHITNEY (2)

- When $n_1 < 20$ and $n_2 < 20$
  - $H_0$ is rejected if $U_{n_1,n_2} > c$: tabulated values

- If $n_1 \geq 20$ and $n_2 \geq 20$
  - Normal approximation:
    $$Z_{n_1,n_2} = \frac{2U_{n_1,n_2} + 1 - n_1n_2}{\sqrt{(n_1n_2)(n_1 + n_2 + 1)/3}}$$
  - $H_0$ is rejected if $|Z_{n_1,n_2}| \geq q_{1-\alpha/2}$, $1-\alpha/2$-order quantile of the normal $N(0,1)$ distribution
KRUSKAL-WALLIS TEST (1)

• Used to compare **probability distribution** of a quantitative variable observed on three and more samples

• Null hypothesis
  • $H_0$: the samples come from the same population
  • $H_1$: sample medians are not equal for at least 2 samples

• Test statistic

$$S = \frac{12}{N(N+1)} \sum_{g=1}^{G} \left[ \frac{R_g^2}{n_g} - 3(N+1) \right], \quad R_g = \sum_{i=1}^{n_g} R_i$$
When \( n_g < 4 \) for at least one level
  - \( H_0 \) is rejected if \( S_{\text{obs}} \geq KW_\alpha \), \( \alpha \)-order quantile of Kruskal & Wallis distribution (tabulated values)
  - \( H_0 \) is not rejected if \( S_{\text{obs}} < KW_\alpha \)

When \( n_g \geq 5 \) for all \( G \) groups: approximation with \( \chi^2 \) distribution with \( G-1 \) degrees of freedom
SPEARMAN TEST

- Used to test dependence between two variables (monotonic relationship)
- Null Hypothesis $H_0: \rho = 0$ (independence)
- Test statistic
  $$Z = 1 - \frac{6 \sum_{i=1}^{n} d_i^2}{n^3 - n}, \quad d_i = R_{x_i} - R_{y_i}$$
- When $n < 11$, $H_0$ is rejected if $Z > c_{\alpha/2}$: tabulated value
- When $n \geq 11$, $H_0$ is rejected if $Z > t_{\alpha/2}$, $\alpha/2$-order quantile of Student distribution with $n-2$ degrees of freedom
MULTIPLE TESTING (1)

• Univariate test ⇒ each feature is individually tested
• Metabolomic dataset: hundreds or thousands of features ⇒ individual tests simultaneously
  • $p$ tests of type $H_{ij}^0: \mu_1 = \mu_2$ vs $H_{ij}^1: \mu_1 \neq \mu_2$ are made ($j=1...p$)
• ANOVA / Kruskal-Wallis test: $H_0$ rejected ⇒ 2 means at least are different ⇒ pairwise comparisons
  • $G(G-1)/2$ tests of type $H_0: \mu_k = \mu_l$ vs $H_1: \mu_k \neq \mu_l$ are made ($k, l \leq G$)

⇒ MULTIPLE TESTING PROBLEM

• $\uparrow \alpha$ risk: the probability of getting a significant result simply due to chance keeps going up (false positive).
  • For $k$ comparisons: $P[\text{Reject } H_0 \mid H_0 \text{ true}] = 1-(1-\alpha)^k$
MULTIPLE TESTING (2)

• $\alpha = 0.05$

$\Rightarrow$ P-value correction depends on the number of comparisons (the probability of observing at least one significant result due to chance remains below your desired significance level)

• Several correction methods proposed
  • Bonferroni
  • False Discovery Rate
  • Family-Wise Error Rate
WHICH TEST FOR WHICH DATA?

Quant. vs Qual.
- Normality, Homoscedasticity
  - Yes
    - ≥ 3 groups
      - ANOVA
    - 2 groups
      - Student
  - No
    - ≥ 3 groups
      - Kruskal-Wallis
    - 2 groups
      - Mann-Whitney

Quant. vs Quant.
- Normality
  - Yes
    - Pearson
  - No
    - Spearman

Qual. vs Qual.
- $\chi^2$
HOW TO DO WITH GALAXY?
GALAXY FORM (1)

- Data matrix
- Sample metadata
- Variable metadata
- Biological factor
- Test choice

HELP
GALAXY FORM (2)

Data matrix
Sample metadata
Variable metadata
Biological factor
Test choice

HELP
GALAXY FORM (3)

Data matrix
Sample metadata
Multiple testing correction

HELP
Data matrix
Sample metadata
Variable metadata
Biological factor
Test choice
Significance threshold
### Galaxy Result: Variable Metadata

#### Galaxy / 4 / Metabolomics

<table>
<thead>
<tr>
<th>Variable</th>
<th>PH</th>
<th>Value</th>
<th>Variable</th>
<th>PH</th>
<th>Value</th>
<th>Variable</th>
<th>PH</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>X9.495</td>
<td>7.4</td>
<td>-1.145e-06</td>
<td>X9.496</td>
<td>7.4</td>
<td>-9.9883160631688e-07</td>
<td>X9.497</td>
<td>7.4</td>
<td>-6.035844158844e-06</td>
</tr>
<tr>
<td>X9.475</td>
<td>7.4</td>
<td>1.3777272727272e-06</td>
<td>X9.476</td>
<td>7.4</td>
<td>-1.0420703207922e-06</td>
<td>X9.477</td>
<td>7.4</td>
<td>-1.6404454545454e-06</td>
</tr>
<tr>
<td>X9.465</td>
<td>7.4</td>
<td>0.15004446740039</td>
<td>X9.466</td>
<td>7.4</td>
<td>0.0826365879567458</td>
<td>X9.467</td>
<td>7.4</td>
<td>-7.9584115384157e-07</td>
</tr>
<tr>
<td>X9.455</td>
<td>7.4</td>
<td>0.155785256535413</td>
<td>X9.445</td>
<td>7.4</td>
<td>-8.026365879567458</td>
<td>X9.435</td>
<td>7.4</td>
<td>-7.0507168321610e-07</td>
</tr>
<tr>
<td>X9.425</td>
<td>7.4</td>
<td>0.137281360782203</td>
<td>X9.415</td>
<td>7.4</td>
<td>-8.94466961500091</td>
<td>X9.405</td>
<td>7.4</td>
<td>-4.4577920277922e-07</td>
</tr>
<tr>
<td>X9.395</td>
<td>7.4</td>
<td>-6.1036363636363e-07</td>
<td>X9.385</td>
<td>7.4</td>
<td>-6.190999999999999</td>
<td>X9.375</td>
<td>7.4</td>
<td>-4.85643056435e-07</td>
</tr>
<tr>
<td>X9.365</td>
<td>7.4</td>
<td>0.62406264024007</td>
<td>X9.355</td>
<td>7.4</td>
<td>0.596923457335513</td>
<td>X9.345</td>
<td>7.4</td>
<td>-1.4421940519458e-07</td>
</tr>
<tr>
<td>X9.335</td>
<td>7.4</td>
<td>0.621970294201778</td>
<td>X9.325</td>
<td>7.4</td>
<td>-2.1879220722072e-07</td>
<td>X9.315</td>
<td>7.4</td>
<td>-4.489013069106075</td>
</tr>
<tr>
<td>X9.305</td>
<td>7.4</td>
<td>0.216367581507682</td>
<td>X9.295</td>
<td>7.4</td>
<td>-7.64850104850104e-07</td>
<td>X9.285</td>
<td>7.4</td>
<td>-3.30155844555844e-06</td>
</tr>
<tr>
<td>X9.275</td>
<td>7.4</td>
<td>0.652610228858504</td>
<td>X9.265</td>
<td>7.4</td>
<td>-1.75048051948562e-06</td>
<td>X9.255</td>
<td>7.4</td>
<td>-8.0961039610396e-07</td>
</tr>
<tr>
<td>X9.245</td>
<td>7.4</td>
<td>0.4024302051362</td>
<td>X9.235</td>
<td>7.4</td>
<td>-1.3712980712087e-07</td>
<td>X9.225</td>
<td>7.4</td>
<td>-1.0404545454545e-06</td>
</tr>
<tr>
<td>X9.215</td>
<td>7.4</td>
<td>0.16210730241644</td>
<td>X9.205</td>
<td>7.4</td>
<td>-1.1977097207922e-06</td>
<td>X9.195</td>
<td>7.4</td>
<td>-1.211165531158638e-07</td>
</tr>
<tr>
<td>X9.185</td>
<td>7.4</td>
<td>-2.294155815815815</td>
<td>X9.175</td>
<td>7.4</td>
<td>-4.012078532087072e-07</td>
<td>X9.165</td>
<td>7.4</td>
<td>-5.63915584455858e-06</td>
</tr>
<tr>
<td>X9.155</td>
<td>7.4</td>
<td>0.68801953383457</td>
<td>X9.145</td>
<td>7.4</td>
<td>-7.58532467324669</td>
<td>X9.135</td>
<td>7.4</td>
<td>-1.9738116831167e-07</td>
</tr>
</tbody>
</table>

### Univariate Analysis

- Variable metadata
- PH: Treatment
test_BPA 25ug BPA 25ng
diff
- PH: Treatment_test_BPA 25ug BPA 25ng_none
- PH: Treatment_test_BPA 25ug BPA 25ng_sig

### Multivariate Analysis

- Univariate: Univariate statistics
- Multivariate PCA, PLS and OPLS
- HCA: Heatmap
- HCA: Heatmap of the dataMatrix
GALAXY RESULT: BOXPLOT
MMUSCULUS DATA DESCRIPTION
**Experimental design**

- **2 BPA doses:**
  - 25 ng
  - 250 ng/kg body weight/day

- **Timeline:**
  - **GD 1**
  - **GD 8**
  - **PND 17**
  - **F1 Birth**
  - **PND 21**

- **Stages:**
  - **F0 Mating**
  - **Vaginal plug**
  - **Pump**
  - **Period of Exposure**
Description

- 24 brain samples from mice pups
  - Mothers exposed to BPA (25 or 250 ng / kg body weight / day)
  - Pups sacrificed at 21 days: brain collection
- Extraction and NMR sample preparation
- NMR analysis:
  - Bruker DRX-600-Avance spectrometer using an inverse detection 5mm cryoprobe attached to a cryoplatform; CPMG spin-echo pulse sequence
  - Fourier transformation was applied, then all spectra were phased and baseline corrected using TopSpin
Galaxy parameters

- Bucket width: 0.01
- Left Border: 9.50
- Right Border: 0.80
- Exclusion zone(s):
  - Zone 1: 5.10-4.50
Paper reference

• Please cite:

Exercise

• Does BPA dose influence NMR variable’s relative intensity?
  • Which test would you use (cite parametric or non test)?
  • Run the chosen test with different multiple test correction methods
  • Compare results

• To help in NMR spectrum annotation, we can identify NMR buckets corresponding to the same metabolite(s)
  • Which test do you use?
  • For example, one of Glucose chemical shift is 1.89ppm
    • Run the parametric and the non parametric versions of the chosen test, using 1.89ppm as comparison variable
    • Compare results