P-values and statistical tests

4. ANOVA

Marek Gierliński
Division of Computational Biology

Hand-outs available at http://is.gd/statlec
<table>
<thead>
<tr>
<th>Section</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Introduction</td>
<td>Null hypothesis, statistical test, p-value Fisher’s test</td>
</tr>
<tr>
<td>2. Contingency tables</td>
<td>Chi-square test G-test</td>
</tr>
<tr>
<td>3. T-test</td>
<td>One- and two-sample Paired One-sample variance test</td>
</tr>
<tr>
<td>4. ANOVA</td>
<td>One-way Two-way</td>
</tr>
<tr>
<td>5. Non-parametric methods 1</td>
<td>Mann-Whitney Wilcoxon signed-rank Kruskal-Wallis</td>
</tr>
<tr>
<td>6. Non-parametric methods 2</td>
<td>Kolmogorov-Smirnov Permutation Bootstrap</td>
</tr>
<tr>
<td>7. Statistical power</td>
<td>Effect size Power in t-test Power in ANOVA</td>
</tr>
<tr>
<td>8. Multiple test corrections</td>
<td>Family-wise error rate False discovery rate Holm-Bonferroni limit Benjamini-Hochberg limit Storey method</td>
</tr>
<tr>
<td>9. What’s wrong with p-values?</td>
<td>A lot</td>
</tr>
</tbody>
</table>
One-way ANOVA
One-way ANOVA

- Extension of the t-test to more than 2 groups

- Null hypothesis: all samples came from populations with the same mean
  \[ H_0: \mu_1 = \mu_2 = \cdots = \mu_k \]

- The null hypothesis is tested by comparing variances

- ANOVA – ANalysis Of VAriance
Variance

- Sample of measurements $x_1, x_2, ..., x_n$
- Sample variance
  \[ SD_{n-1}^2 = \frac{1}{n-1} \sum (x_i - M)^2 \]

- Generalized variance: mean square
  \[ MS = \frac{SS}{\nu} \]

- where
  - $SS$ - sum of squared residuals
  - $\nu$ - number of degrees of freedom

Variance represents spread of data around their mean

Variance $= $ standard deviation$^2$
Variance between and within groups

- Variance within groups – typical variance in each group
- Variance between groups – how the sample mean varies from group to group
One-way ANOVA

- Null hypothesis: all samples came from populations with the same mean
  - $H_0: \mu_1 = \mu_2 = \cdots = \mu_k$

- Assumption: they all have common variance $\sigma^2$

- $n = 34$ data points
- $k = 4$ groups of data

- $n_g$ - number of points in group $g$
- $x_{gi}$ - body mass, group $g$, mouse $i$
- $\bar{x}_g$ - mean in group $g$
- $\bar{x}$ - grand mean, across all data points
Variance within groups

- Variance within groups is

$$MS_W = \frac{SS_W}{\nu_W}$$

- $MS_W$ estimates the common variance, $\sigma^2$, regardless of the null hypothesis

$$SS_W = \sum_{g=1}^{k} \sum_{i=1}^{n_g} (x_{gi} - \bar{x}_g)^2$$

$$\nu_W = \sum_{g=1}^{k} (n_g - 1)$$
Variance within groups

- Variance within groups is

\[ MS_W = \frac{SS_W}{\nu_W} \]

- \( MS_W \) estimates the common variance, \( \sigma^2 \), regardless of the null hypothesis

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>( SS_W )</td>
<td>524</td>
</tr>
<tr>
<td>( \nu_W )</td>
<td>30</td>
</tr>
<tr>
<td>( MS_W )</td>
<td>17.5</td>
</tr>
</tbody>
</table>

\[ SS_W = \sum_{g=1}^{k} \sum_{i=1}^{n_g} (x_{gi} - \bar{x}_g)^2 \]

mean of group \( g \)

size of group \( g \)

\[ \nu_W = \sum_{g=1}^{k} (n_g - 1) \]
Variance between groups is

\[ MS_B = \frac{SS_B}{\nu_B} \]

- sum of squared residuals
- number of degrees of freedom

\[ MS_B \] estimates the common variance, \( \sigma^2 \), **only** when the null hypothesis is true

\[
SS_B = \sum_{g=1}^{k} n_g (\bar{x}_g - \bar{x})^2
\]

\[
\nu_B = k - 1
\]
Variance between groups

- Variance between groups is

\[ MS_B = \frac{SS_B}{\nu_B} \]

**sum of squared residuals**

**number of degrees of freedom**

- \( MS_B \) estimates the common variance, \( \sigma^2 \), **only** when the null hypothesis is true.

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>( SS_W )</td>
<td>524</td>
</tr>
<tr>
<td>( \nu_W )</td>
<td>30</td>
</tr>
<tr>
<td>( MS_W )</td>
<td>17.5</td>
</tr>
<tr>
<td>( SS_B )</td>
<td>623</td>
</tr>
<tr>
<td>( \nu_B )</td>
<td>3</td>
</tr>
<tr>
<td>( MS_B )</td>
<td>208</td>
</tr>
</tbody>
</table>

\[ SS_B = \sum_{g=1}^{k} n_g (\bar{x}_g - \bar{x})^2 \]

\[ \nu_B = k - 1 \]
F-test

- $MS_W$ estimates the common variance, $\sigma^2$, regardless of the null hypothesis.

- $MS_B$ estimates the common variance, $\sigma^2$, only when the null hypothesis is true.

Test for equality of variances: F-test

$$F = \frac{MS_B}{MS_W}$$

- Degrees of freedom: $\nu_B, \nu_W$

- If $H_0$ is true, we expect $F \sim 1$

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$SS_W$</td>
<td>524</td>
</tr>
<tr>
<td>$\nu_W$</td>
<td>30</td>
</tr>
<tr>
<td>$MS_W$</td>
<td>17.5</td>
</tr>
<tr>
<td>$SS_B$</td>
<td>623</td>
</tr>
<tr>
<td>$\nu_B$</td>
<td>3</td>
</tr>
<tr>
<td>$MS_B$</td>
<td>208</td>
</tr>
<tr>
<td>$F$</td>
<td>11.9</td>
</tr>
</tbody>
</table>
Null distribution

Population of British mice
\( \mu = 20 \text{ g}, \sigma = 5 \text{ g} \)

Select four samples
size 12, 9, 8 and 5

\[ F = \frac{MS_B}{MS_W} \]

Build distribution of \( F \)

Normal population
\( \mu = 20 \text{ g}, \sigma = 5 \text{ g} \)

\[ \times 100,000 \]
Null distribution = $F$-distribution

Null distribution represents all random samples when the null hypothesis is true.

Observed $F = 11.9$

$p$-value very small
**Effect vs. no effect**

<table>
<thead>
<tr>
<th></th>
<th>$MS_W$</th>
<th>$MS_B$</th>
<th>$F$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>No effect</td>
<td>17.5 g²</td>
<td>208 g²</td>
<td>11.9</td>
<td>3×10⁻⁵</td>
</tr>
<tr>
<td>Effect</td>
<td>25.7 g²</td>
<td>24.3 g²</td>
<td>0.94</td>
<td>0.43</td>
</tr>
</tbody>
</table>

![Box plots](image)
ANOVA assumptions

- Normality – data in each group are distributed normally
  - ANOVA is quite robust against non-normality
  - if strongly not normal (e.g. log-normal) – transform to normality
  - if this fails, use non-parametric Kruskal-Wallis test

- Independence – groups are independent
  - dependence: e.g., observations of the same subjects over time
  - if groups are not independent, ANOVA is not appropriate, use other methods

- Equality of variances – groups sampled from populations with the same variance
  - sometimes called homogeneity of variances, or homoscedasticity /ˌhɒməsəˈdɛstɪsɪtɪ/
  - if variances are not equal, use Welch’s approximated test
Test to compare variances

- Null hypothesis: samples come from populations with equal variances
  - $H_0: \sigma_1^2 = \sigma_2^2 = \cdots = \sigma_k^2$

- Like ANOVA, except data $x_{gi}$ are replaced by residuals $R_{gi}$:
  - Levene’s test
  - Brown-Forsythe test

\[ R_{gi} = |x_{gi} - \bar{x}_g| \]

- Test statistic:
  \[ W = \frac{MS_B}{MS_W} \]
Test to compare variances

- Null hypothesis: samples come from populations with equal variances
  - $H_0: \sigma_1^2 = \sigma_2^2 = \cdots = \sigma_k^2$

- Test statistic:
  \[
  W = \frac{MS_B}{MS_W}
  \]

- $MS_B = 6.40 \text{ g}^2$
- $MS_W = 6.89 \text{ g}^2$
- $W = 0.930$
- $p = 0.44$
What if variances are not equal?

- B. L. Welch developed an approximated test
- Skip the details...

- Mice data

<table>
<thead>
<tr>
<th></th>
<th>$F$</th>
<th>$v_1$</th>
<th>$v_2$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANOVA</td>
<td>11.89</td>
<td>3</td>
<td>30</td>
<td>2.7×10^{-5}</td>
</tr>
<tr>
<td>Welch’s test</td>
<td>28.95</td>
<td>3</td>
<td>15.96</td>
<td>10^{-6}</td>
</tr>
</tbody>
</table>
## Post-hoc analysis: Tukey’s test

- A multiple t-test
- Finds differences and p-values for each pair of categories
- Post-hoc test, you need ANOVA first
- Skip the details...

<table>
<thead>
<tr>
<th></th>
<th>Scottish</th>
<th>Welsh</th>
<th>N.Irish</th>
</tr>
</thead>
<tbody>
<tr>
<td>Welsh</td>
<td>-1.1</td>
<td>0.95</td>
<td></td>
</tr>
<tr>
<td>N.Irish</td>
<td>-12.9</td>
<td>-11.9</td>
<td>0.00003*</td>
</tr>
<tr>
<td></td>
<td>0.0001*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>English</td>
<td>-4.9</td>
<td>-3.9</td>
<td>8.0</td>
</tr>
<tr>
<td></td>
<td>0.05</td>
<td>0.20</td>
<td>0.006*</td>
</tr>
</tbody>
</table>
How to do it in R?

# ANOVA
> mice <- read.table("http://tiny.cc/mice_1way", header=TRUE)
> mice.aov <- aov(Mass ~ Country, data=mice)
> summary(mice.aov)

             Df Sum Sq Mean Sq F value Pr(>F)
Country       3  622.7  207.56   11.89 2.67e-05 ***
Residuals    30  523.9   17.46
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

# Tukey’s Honest Significant Differences
> TukeyHSD(mice.aov)

Tukey multiple comparisons of means
95% family-wise confidence level

Fit: aov(formula = Mass ~ Country, data = mice)

$Country
            diff  lwr   upr   p adj
N.Irish-English  -8.001667 -14.04998948 -1.953344 0.0059422
Scottish-English  4.947222 -0.06331043  9.957755 0.0539580
Welsh-English  3.858333 -1.32806069  9.044727 0.2023039
Scottish-N.Irish 12.948889  6.61101070 19.286767 0.0000277
Welsh-N.Irish  11.860000  5.38219594 18.337804 0.0001394
Welsh-Scottish -1.088889 -6.61022696  4.432449 0.9494897

...
How to do it in R?

# Brown-Forsythe test for equality of variances
> library(lawstat)
> levene.test(mice$Mass, mice$Country)

modified robust Brown-Forsythe Levene-type test based on the absolute deviations from the median

data:  mice$Mass
Test Statistic = 0.92948, p-value = 0.4386

# Welch’s test for unequal variances
> oneway.test(Mass ~ Country, mice, var.equal=FALSE)

One-way analysis of means (not assuming equal variances)

data:  mass and country
F = 28.95, num df = 3.00, denom df = 15.96, p-value = 1.084e-06
Two-way ANOVA
Two variables

variable 1

variable 2
ANOVA as a linear model (one-way)

\[ x_{ir} = \mu_i + \epsilon_{ir} \]

null hypothesis

\[ H_0: \mu_1 = \mu_2 = \cdots = \mu_k \]

\[ x_{ir} = \mu + \alpha_i + \epsilon_{ir} \]

null hypothesis

\[ H_0: \alpha_1 = \alpha_2 = \cdots = \alpha_k = 0 \]

\[ \forall i: \alpha_i = 0 \]
ANOVA as a linear model (two-way)

\[
x_{ijr} = \mu_{ij} + \epsilon_{ijr}
\]

- Column means are equal:
  \[
  H_{0}^{\text{col}}: \mu_1 = \mu_2 = \cdots = \mu_{nc} \text{ or } \forall i: \alpha_i = 0
  \]

- Row means are equal:
  \[
  H_{0}^{\text{row}}: \mu_1 = \mu_2 = \cdots = \mu_{nr} \text{ or } \forall i: \beta_i = 0
  \]

- There is no interaction between rows and columns:
  \[
  H_{0}^{\text{int}}: \forall i, j: \gamma_{ij} = 0
  \]
More mice!
Two-way ANOVA – two variables
# 2-way ANOVA

```r
> mice <- read.table("http://tiny.cc/mice_2way", header=TRUE)
> mice.lm <- lm(Mass ~ Country + Colour + Country*Colour, mice)
> anova(mice.lm)
```

Analysis of Variance Table

Response: Mass

<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>Country</td>
<td>3</td>
<td>809.68</td>
<td>269.893</td>
<td>11.936</td>
<td>3.598e-06 ***</td>
</tr>
<tr>
<td>Colour</td>
<td>1</td>
<td>59.87</td>
<td>59.873</td>
<td>2.648</td>
<td>0.1092</td>
</tr>
<tr>
<td>Country:Colour</td>
<td>3</td>
<td>107.39</td>
<td>35.797</td>
<td>1.583</td>
<td>0.2034</td>
</tr>
<tr>
<td>Residuals</td>
<td>57</td>
<td>1288.80</td>
<td>22.611</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
Null hypotheses: all three true

- $x_{ijr} = \mu + \alpha_i + \beta_j + \gamma_{ij} + \epsilon_{ijr}$
- $A = \begin{pmatrix} 0 & 0 & 0 & 0 \end{pmatrix}$, $B = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, $\Gamma = \begin{pmatrix} 0 & 0 & 0 & 0 \end{pmatrix}$

<table>
<thead>
<tr>
<th></th>
<th>English</th>
<th>Scottish</th>
<th>Welsh</th>
<th>N.Irish</th>
</tr>
</thead>
<tbody>
<tr>
<td>White</td>
<td><img src="image1.png" alt="Diagram" /></td>
<td><img src="image2.png" alt="Diagram" /></td>
<td><img src="image3.png" alt="Diagram" /></td>
<td><img src="image4.png" alt="Diagram" /></td>
</tr>
<tr>
<td>Black</td>
<td><img src="image5.png" alt="Diagram" /></td>
<td><img src="image6.png" alt="Diagram" /></td>
<td><img src="image7.png" alt="Diagram" /></td>
<td><img src="image8.png" alt="Diagram" /></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>p-value</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>columns</td>
<td>0.68</td>
</tr>
<tr>
<td>rows</td>
<td>0.87</td>
</tr>
<tr>
<td>interaction</td>
<td>0.96</td>
</tr>
</tbody>
</table>
Null hypotheses: columns not equal

- \( x_{ijr} = \mu + \alpha_i + \beta_j + \gamma_{ij} + \epsilon_{ijr} \)
- \( A = \begin{pmatrix} 0 & 10 & -10 & 0 \end{pmatrix}, B = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \Gamma = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \)

<table>
<thead>
<tr>
<th></th>
<th>English</th>
<th>Scottish</th>
<th>Welsh</th>
<th>N.Irish</th>
</tr>
</thead>
<tbody>
<tr>
<td>White</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>columns</td>
<td>(2 \times 10^{-19})</td>
</tr>
<tr>
<td>rows</td>
<td>0.38</td>
</tr>
<tr>
<td>interaction</td>
<td>0.46</td>
</tr>
</tbody>
</table>
Null hypotheses: rows not equal

- $x_{ijr} = \mu + \alpha_i + \beta_j + \gamma_{ij} + \epsilon_{ijr}$
- $A = (0 \ 0 \ 0 \ 0)$, $B = \begin{pmatrix} 10 \\ -10 \end{pmatrix}$, $\Gamma = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$

<table>
<thead>
<tr>
<th></th>
<th>English</th>
<th>Scottish</th>
<th>Welsh</th>
<th>N.Irish</th>
</tr>
</thead>
<tbody>
<tr>
<td>White</td>
<td><img src="image1.png" alt="Diagram" /></td>
<td><img src="image2.png" alt="Diagram" /></td>
<td><img src="image3.png" alt="Diagram" /></td>
<td><img src="image4.png" alt="Diagram" /></td>
</tr>
<tr>
<td>Black</td>
<td><img src="image5.png" alt="Diagram" /></td>
<td><img src="image6.png" alt="Diagram" /></td>
<td><img src="image7.png" alt="Diagram" /></td>
<td><img src="image8.png" alt="Diagram" /></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>columns</td>
<td>0.75</td>
</tr>
<tr>
<td>rows</td>
<td>$10^{-27}$</td>
</tr>
<tr>
<td>interaction</td>
<td>0.56</td>
</tr>
</tbody>
</table>
Null hypotheses: interaction

- $x_{ijr} = \mu + \alpha_i + \beta_j + \gamma_{ij} + \epsilon_{ijr}$

- $A = \begin{pmatrix} 0 & 0 & 0 & 0 \end{pmatrix}, B = \begin{pmatrix} 0 \end{pmatrix}, \Gamma = \begin{pmatrix} -10 & 10 & 0 & 0 \\ 10 & -10 & 0 & 0 \end{pmatrix}$

<table>
<thead>
<tr>
<th></th>
<th>English</th>
<th>Scottish</th>
<th>Welsh</th>
<th>N.Irish</th>
</tr>
</thead>
<tbody>
<tr>
<td>White</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>White</td>
<td>Black</td>
</tr>
</tbody>
</table>

$p$

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>columns</td>
<td>0.78</td>
</tr>
<tr>
<td>rows</td>
<td>0.67</td>
</tr>
<tr>
<td>interaction</td>
<td>$10^{-20}$</td>
</tr>
</tbody>
</table>
Drug effect

<table>
<thead>
<tr>
<th></th>
<th>Drug</th>
<th>Placebo</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>18.5, 21.2, 19.9</td>
<td>3.5, 4.7, 8.3</td>
</tr>
<tr>
<td>Female</td>
<td>21.3, 20.1, 18.6</td>
<td>4.5, 5.6, 3.9</td>
</tr>
</tbody>
</table>

Measurements: a score representing response to the drug
Drug effect

Drug effect

Drug and gender effect

Interaction effect

Gender

Score

Drug

Placebo

Male

Female

Gender

Score

Drug

Placebo

Male

Female

Gender

Score

Drug

Placebo

Male

Female

Gender
Time-course experiments
Time-course experiments

- Obesity study in mice
- Two groups:
  - untreated
  - treated with a drug
- Feed them a lot
- Observe body mass over time

- Is there a difference between the two groups?
Time-course experiments

- You can do ANOVA
- \( p = 5 \times 10^{-5} \)

But

- Data are correlated
- ANOVA doesn’t recognize numerical variables (time)

```r
> dat <- read.table('http://tiny.cc/time_course', header=TRUE)
> dat.lm <- lm(Mass ~ Treatment + Time + Treatment*Time, dat)
> anova(dat.lm)

             Df  Sum Sq Mean Sq  F value    Pr(>F)
Treatment     1  85.538  85.538 20.1508 4.481e-05
Time          7 272.465  38.924  9.1694 3.825e-07
Treatment:Time 7 230.738  32.963  7.7652 2.907e-06
```
Time-course experiments

- What about t-test at each time point?
- Works well!
- Three time points are significantly different
- But: misses point-to-point correlation
Data transformation

- Original data: untreated and treated
- Normalized to first time point
- Mean of normalized data

Graphs show the change in body mass over time for untreated and treated groups, both normalized to the first time point and with mean and standard error.
Better approach: build a model

- First: understand your data
- Build a model and reduce time-course curves to just one number
- Do a t-test or similar test on these numbers

- Very simple: area under each curve
- This gives us 4 vs. 3 areas
Compare area under the curve

Time 0-10 w
FC = 1.14
P = 0.051

Time 6-10 w
FC = 1.33
P = 0.014

Time 0-5 w
FC = 0.99
P = 0.82
## Chi-square or G-test vs. ANOVA

<table>
<thead>
<tr>
<th></th>
<th>WT</th>
<th>KO1</th>
<th>KO2</th>
<th>KO3</th>
</tr>
</thead>
<tbody>
<tr>
<td>G1</td>
<td>50, 54, 48</td>
<td>61, 75, 69</td>
<td>78, 77, 80</td>
<td>43, 34, 49</td>
</tr>
<tr>
<td>S</td>
<td>172, 180, 172</td>
<td>175, 168, 166</td>
<td>162, 167, 180</td>
<td>178, 173, 168</td>
</tr>
<tr>
<td>G2</td>
<td>55, 50, 63</td>
<td>45, 41, 38</td>
<td>47, 49, 43</td>
<td>59, 50, 45</td>
</tr>
</tbody>
</table>

### Fisher’s test / Chi-square test / G-test

Experiment outcome: category
Table contains counts

<table>
<thead>
<tr>
<th></th>
<th>English</th>
<th>Scottish</th>
<th>Welsh</th>
<th>N. Irish</th>
</tr>
</thead>
<tbody>
<tr>
<td>White</td>
<td>19.1, 20, 21</td>
<td>22.3, 21.2, 25.6</td>
<td>18.1, 19.2, 22.7</td>
<td>15.6, 16.7, 15</td>
</tr>
<tr>
<td>Grey</td>
<td>20, 21, 17</td>
<td>18.6, 20.1, 19.7</td>
<td>15, 18, 22</td>
<td>12, 18.1, 20.3</td>
</tr>
</tbody>
</table>

### ANOVA

Experiment outcome: measurement (could be counts)
Table contains measurements
Bacterial antibiotic resistance

- Four strains
- Grown in normal medium and two antibiotic concentrations
- Dilution plating, count colonies

<table>
<thead>
<tr>
<th></th>
<th>WT</th>
<th>KO1</th>
<th>KO2</th>
<th>KO3</th>
</tr>
</thead>
<tbody>
<tr>
<td>No antibiotic</td>
<td>77, 51, 92</td>
<td>50, 83, 16</td>
<td>70, 111, 78</td>
<td>121, 147, 110</td>
</tr>
<tr>
<td>Conc. 1</td>
<td>83, 51, 40</td>
<td>66, 18, 49</td>
<td>95, 109, 52</td>
<td>75, 116, 109</td>
</tr>
<tr>
<td>Conc. 2</td>
<td>11, 7, 31</td>
<td>69, 41, 21</td>
<td>85, 51, 60</td>
<td>95, 128, 116</td>
</tr>
</tbody>
</table>

Outcome is measurement, not category
This is not a contingency table!

Use ANOVA
Need to check normality
Hand-outs available at http://tiny.cc/statlec