ANOVA, still

\{Y_{ti}\} independent with \(Y_{ti} \sim \text{normal}(\mu_t, \sigma)\) for \(t = 1 \ldots k\).

Test \(H_0: \mu_1 = \mu_2 = \cdots = \mu_k\)

The usual statistic:

\[
F = \frac{M_B}{M_W} = \frac{\frac{1}{k} \sum_t n_t (\bar{Y}_t - \bar{Y}_\cdot)^2}{\frac{1}{\sum_t \sum_i (Y_{ti} - \bar{Y}_t)^2} / (\sum_t n_t - k)}
\]

P-values:

(a) Use the F(k, \(\sum n_t - k\)) distribution.
(b) Use a permutation test.

Assumptions:

(a) Underlying dist'ns are normal with common SD.
(b) Underlying dist'ns are the same.

Non-parametric ANOVA

An alternative approach: the Kruskal-Wallis test.

Rank all of the observations from 1, 2, \ldots, N.

Let \(R_{ti}\) = the rank for observation \(Y_{ti}\).

Let \(R_t = \sum_i R_{ti} / n_t\) = the average rank for group \(t\).

Null hypothesis, \(H_0\): the underlying distributions are all the same.

\[
E(R_t \mid H_0) = \frac{N+1}{2}
\]

\[
SD(R_t \mid H_0) = \sqrt{\frac{(N+1)(N-n_t)}{12 n_t}}
\]
Kruskal-Wallis test statistic

\[
H = \sum_t \left( \frac{N - n_t}{N} \right) \times \left[ \frac{\bar{R}_t - E(\bar{R}_t \mid H_0)}{SD(\bar{R}_t \mid H_0)} \right]^2
\]

\[
= \cdots = \frac{12}{N(N+1)} \sum_t n_t \left[ \bar{R}_t - \left( \frac{N+1}{2} \right) \right]^2
\]

Under \( H_0 \), and if the sample sizes are large, \( H \sim \chi^2(\text{df} = k - 1) \).

Alternatively, we could use a permutation test to estimate a P-value.

The function \texttt{kruskal.test()} in R will calculate the statistic.

Note

In the case of two groups, the Kruskal-Wallis test reduces exactly to the Wilcoxon rank-sum test.

This is just like how ANOVA is equivalent to the two-sample t test.
ANOVA Tables

Original scale / 1000:

<table>
<thead>
<tr>
<th>source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>between strains</td>
<td>33</td>
<td>20</td>
<td>1.69</td>
<td>1.70</td>
<td>0.042</td>
</tr>
<tr>
<td>within strains</td>
<td>124</td>
<td>125</td>
<td>0.99</td>
<td></td>
<td></td>
</tr>
<tr>
<td>total</td>
<td>157</td>
<td>145</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

permutation P-value = 0.043

log₁₀ scale:

<table>
<thead>
<tr>
<th>source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>between strains</td>
<td>3.35</td>
<td>20</td>
<td>0.167</td>
<td>2.25</td>
<td>0.0036</td>
</tr>
<tr>
<td>within strains</td>
<td>9.29</td>
<td>125</td>
<td>0.074</td>
<td></td>
<td></td>
</tr>
<tr>
<td>total</td>
<td>12.63</td>
<td>145</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

permutation P-value = 0.003
K-W results

The observed Kruskal-Wallis statistic for these data was \(41.32\). (Note that it doesn’t matter whether you take logs.)

Since there were 21 strains, we can compare this to a \(\chi^2\) distribution with 20 degrees of freedom. Thus we obtain the \(P\)-value = 0.003.

With a permutation test, I got \(\hat{P} = 0.0015\) (on the basis of 10,000 simulations).

In the case of ties...

In the case of ties, we assign the average rank to each.

Example:

\[
\begin{array}{cccccc}
A: & 3.5 & 3.7 & 4.0 & 4.2 & 4.3 \\
B: & & 3.9 & 4.3 & 4.5 \\
C: & 3.1 & 3.6 & 4.0 & 4.3 \\
\end{array}
\]

(1) (2) (3) (4) (5) (6/7) (8) (9/10/11) (12)

\[
\begin{array}{c}
6.5 \\
10 \\
\end{array}
\]

Then we apply a correction factor.

Let \(N = \sum_i n_i\) and \(T_i\) = no. observations in the ith set of ties (can be 1).

Let \(D = 1 - \sum_i (T_i^3 - T_i) / (N^3 - N)\)

Use the statistic \(H' = H/D\).

Note that \(D \leq 1\) and so \(H' \geq H\). For the example, \(D = 1 - \frac{(2^3 - 2) + (3^3 - 3)}{12^3 - 12} \approx 0.983\).
Blood coagulation time

Example (continued)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th>rank</th>
<th>avg rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>65</td>
<td>15</td>
<td>15</td>
<td>15</td>
<td>15.5</td>
<td></td>
</tr>
<tr>
<td>66</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>16.5</td>
<td></td>
</tr>
<tr>
<td>66</td>
<td>17</td>
<td>17</td>
<td>17</td>
<td>17.5</td>
<td></td>
</tr>
<tr>
<td>67</td>
<td>18</td>
<td>18</td>
<td>18</td>
<td>18.5</td>
<td></td>
</tr>
<tr>
<td>67</td>
<td>19</td>
<td>19</td>
<td>19</td>
<td>19.5</td>
<td></td>
</tr>
<tr>
<td>68</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>20.5</td>
<td></td>
</tr>
<tr>
<td>68</td>
<td>21</td>
<td>21</td>
<td>21</td>
<td>21.5</td>
<td></td>
</tr>
<tr>
<td>68</td>
<td>22</td>
<td>22</td>
<td>22</td>
<td>22.5</td>
<td></td>
</tr>
<tr>
<td>71</td>
<td>23</td>
<td>23</td>
<td>23</td>
<td>23.5</td>
<td></td>
</tr>
<tr>
<td>71</td>
<td>24</td>
<td>24</td>
<td>24</td>
<td>24.5</td>
<td></td>
</tr>
</tbody>
</table>
### Example (continued)

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>4</td>
<td>6</td>
<td>6</td>
<td>8</td>
<td>24</td>
</tr>
<tr>
<td>R</td>
<td>6.25</td>
<td>16.25</td>
<td>20.25</td>
<td>7.00</td>
<td></td>
</tr>
</tbody>
</table>

\[
H = \frac{12}{N(N+1)} \sum_{t} n_t \left[ \bar{R}_t - \left( \frac{N+1}{2} \right) \right]^2
\]

\[
= \frac{12}{24 \times 25} \left\{ 4 \times (6.25 - 12.5)^2 + \cdots + 8 \times (7.00 - 12.5)^2 \right\}
\]

\[
= 16.86
\]

The ties: \( T_i = (1 2 2 1 2 4 2 1 2 2 3 2) \)

\[
D = 1 - \sum_i (T_i^3 - T_i)/(N^3 - N) = \ldots = 0.991
\]

\[
H' = \frac{H}{D} = 16.86 / 0.991 = 17.02 \quad \text{[df = 3]} \quad \text{P-value} \approx 0.0007
\]
A few points

- **Calculation of P-values:** (avoiding type I errors)
  - F statistic: F distribution (requires normality)
  - K-W statistic: $\chi^2$ distribution (requires large samples)
  - Either statistic: Permutation tests

- **Power:** (avoiding type II errors)
  - K-W statistic more resistant to outliers
  - F statistic more powerful in the case of normality

- K-W statistic: don’t need to worry about transformations.

A fake example
## Results

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Method</th>
<th>Statistic</th>
<th>nom. P-value</th>
<th>Perm.’n P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>ANOVA</td>
<td>5.48</td>
<td>0.020</td>
<td>0.017</td>
</tr>
<tr>
<td></td>
<td>K-W</td>
<td>7.64</td>
<td>0.022</td>
<td>0.012</td>
</tr>
<tr>
<td>B</td>
<td>ANOVA</td>
<td>2.64</td>
<td>0.112</td>
<td>0.023</td>
</tr>
<tr>
<td></td>
<td>K-W</td>
<td>7.64</td>
<td>0.022</td>
<td>0.012</td>
</tr>
</tbody>
</table>

### Distributions

**ANOVA**
- **Dataset A**
  - F statistic
- **Dataset B**
  - F statistic

**K–W test**
- K–W statistic