So far we have looked at 1-sample, 2-sample, and $t$-sample problems.

In the latter we dealt with a treatment at $t$ levels.

One could call this treatment also a factor.

Now we will address experiments where several factors come into play.

First we will do this for two such factors.

How should we go about this?
We have 3 types of insecticides (I, II, and III) and 4 methods (A, B, C, D) of delivering the insecticide. (I, II, and III) and (A, B, C, D) are the levels of the respective factors.

The response $Y$ is the time to death in minutes.

We want to find the best insecticide and the best delivery method.

We have 48 experimental insects to experiment with.

Randomly divide the 48 insects into $12 = 3 \times 4$ groups of 4 insects each, assigning the respective groups to the 12 treatment combinations $(I, A), (I, B), (I, C), (I, D), \ldots, (III, C), (III, D)$.

Randomize the order of all 48 runs to eliminate order biases.

This is a factorial design, specifically a $3 \times 4$ factor design with 4 replications.
It is useful to visualize the responses in relation to the factor levels as follows:

<table>
<thead>
<tr>
<th>Insecticide Type</th>
<th>Delivery Method</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A</td>
</tr>
<tr>
<td>I</td>
<td>$Y_{I,A}$</td>
</tr>
<tr>
<td>II</td>
<td>$Y_{II,A}$</td>
</tr>
<tr>
<td>III</td>
<td>$Y_{III,A}$</td>
</tr>
</tbody>
</table>

where $Y_{I,A}$ stands short for $(Y_{I,A,1}, \ldots, Y_{I,A,4})$, and so on.

More generically we would denote the $k^{\text{th}}$ response under level $i$ from factor 1 and level $j$ from factor 2 by $Y_{i,j,k}$. This is more useful in $\Sigma$ summation notation.
First Look at Insecticide Boxplots

Applied Statistics and Experimental Design

Fritz Scholz — Fall 2006

- Boxplots showing the distribution of time to death (minutes) for different types of poison and delivery methods.
Insecticide type III and delivery method A seem to give the best combination. Is *combination* the right word here?

Are the effects of delivery consistent across types, i.e., is the delivery effectiveness order (in terms of faster response time) the same from one insecticide type to another?

It could be that delivery type A is not the fastest acting among all four when applied to insecticide type III.

Delivery method C could actually be better in combination with III.
Insecticide type III seems to have lowest response with all delivery methods.

The mean levels for each (delivery method, III) combination are $\approx$ consistent.

The scatter within each (delivery method, III) combination is quite tight.

Delivery appears to have an effect on the response under type I and II, both in absolute terms and relative to each other.

It appears that $\text{scatter} \uparrow$ as $\text{mean} \uparrow$ across all combinations.

$\Rightarrow$ variance stabilizing transformation. Deal with that first.
Linear Fit \[ \log(s_i) = a \times \log(\hat{\mu}_i) + b \]
According to our guidelines this suggests \( \alpha = 2 \) or \( \lambda = 1 - \alpha = -1 \),
i.e., \( \tilde{Y}_{ijk} = Y_{ijk}^{-1} = 1/Y_{ijk} \) a reciprocal transform for our response times.

A rationalization attempt:
Suppose the absorption rate \( R = d/t \) (of dose \( d \) over time \( t \)) under any given combination is the most variable process aspect from insect to insect.
Assume that this absorption variability (ingestion variability from insect to insect) is constant over all insecticide types and delivery methods.
Assume further, that the lethal dose \( D \) is \( \approx \) constant for all insects.
Then the time to reach lethal dose is \( T = D/R \). If we took \( 1/T = R/D \) as transformed response we would have constant variability in \( 1/T \).
Linearizing by a 1-term Taylor expansion around \( \mu_R \) and treating \( D \) as a constant

\[
T = \frac{D}{R} \approx \frac{D}{\mu_R} - (R - \mu_R) \frac{1}{\mu_R^2} \Rightarrow \mu_T \approx \frac{D}{\mu_R}, \quad \approx \text{var}(T) \approx \frac{\sigma_R^2}{\mu_R^4} \Rightarrow \sigma_T \propto \mu_T^2
\]
Reciprocal Time Boxplots

Applied Statistics and Experimental Design

Fritz Scholz — Fall 2006

reciprocal time to death (1/minute)

type of poison

A B C D

delivery method

reciprocal time to death (1/minute)
Reciprocal Time by Delivery Method

Applied Statistics and Experimental Design

Fritz Scholz — Fall 2006

Reciprocal time to death (1/minutes)

delivery method A

I
II
III

0.0 0.1 0.2 0.3 0.4 0.5

Reciprocal time to death (1/minutes)

delivery method B

I
II
III

0.0 0.1 0.2 0.3 0.4 0.5

Reciprocal time to death (1/minutes)

delivery method C

I
II
III

0.0 0.1 0.2 0.3 0.4 0.5

Reciprocal time to death (1/minutes)

delivery method D

I
II
III

0.0 0.1 0.2 0.3 0.4 0.5 0.6

12
Full Comparison of Reciprocal Times

Applied Statistics and Experimental Design
Fritz Scholz — Fall 2006

reciprocal time to death (1/minute)

0.0 0.1 0.2 0.3 0.4 0.5 ...

type.delivery combinations

I.A II.A III.A I.B II.B III.B I.C II.C III.C I.D II.D III.D
Linear Fit \[ \log(s_i) = a \times \log(\hat{\mu}_i) + b \] for Reciprocal Times
We can perform 3 different 1-way ANOVAs:

1) Focus on the 3 types alone, i.e., I, II, III
   
   We have 3 samples of 16 observations each.

2) Focus on the 4 delivery methods, i.e., A, B, C, D
   
   We have 4 samples of 12 observations each.

3) Focus on all $3 \times 4$ combinations, i.e., (I,A),..., (I,D),..., (III,A),..., (III,D)
   
   We have 12 samples of 4 observations each.
Three ANOVAs

Analysis of Variance Table

Response: recip.time

<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>type</td>
<td>2</td>
<td>0.34877</td>
<td>0.17439</td>
<td>25.621</td>
<td>3.728e-08 ***</td>
</tr>
<tr>
<td>Residuals</td>
<td>45</td>
<td>0.30628</td>
<td>0.00681</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Response: recip.time

<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>delivery</td>
<td>3</td>
<td>0.20414</td>
<td>0.06805</td>
<td>6.6401</td>
<td>0.0008496 ***</td>
</tr>
<tr>
<td>Residuals</td>
<td>44</td>
<td>0.45091</td>
<td>0.01025</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Response: recip.time

<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>type:delivery</td>
<td>11</td>
<td>0.56862</td>
<td>0.05169</td>
<td>21.531</td>
<td>1.289e-12 ***</td>
</tr>
<tr>
<td>Residuals</td>
<td>36</td>
<td>0.08643</td>
<td>0.00240</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>