The $t$-test and ANOVA

David L. Streiner, Ph.D.
Director, Kunin-Lunenfeld Applied Research Unit
Assistant V.P., Research
Baycrest Centre for Geriatric Care

Professor, Department of Psychiatry
University of Toronto
Question: Does a psychosocial intervention improve quality of life (QOL) following liver transplantation?

Two groups of transplant patients
- One gets treatment as usual (TAU)
- Other gets psychosocial intervention (PSI)

Measure QOL 6 months post-op
Background

• *A priori*, we say that a 10 point difference in QOL would be clinically important.

• After the study, we find a difference between the groups.

• Is the difference *statistically* significant?

• Answer: It all depends.
1. How big is the difference?
   • The bigger the difference, the more likely the results are significant
Mean Difference

TAU

PSI
Mean Difference

Scale:

0  5  10  15  20  25  30  35

TAU

PSI
• Does that mean that a large difference will be significant?
  – Again, it all depends
  – How accurately was each mean estimated?
  – If we repeated the study 100 times, how close to one another are the estimates of each group’s mean?
Spread of Mean Estimates

Wide

Narrow
It Depends On …

- The spread of means is the *Standard Error of the Mean (SEM)*
- SEM is smaller if:
  - SD of the group is small
  - $N$ is large
- The smaller the SEM, the more likely the results are significant
Summarizing

Difference between groups more likely to be significant if there is:

1. Large difference between means
2. Small SD in each group
3. Large $N$ in each group
Writing it Out as an Equation

\[ t = \frac{M_1 - M_2}{\sqrt{s_1^2 + s_2^2}} \]
Conceptually

The $t$-test (and all statistical tests) are a ratio:

\[
\frac{\text{Signal}}{\text{Noise}}
\]

where \textit{signal} is the size of effect;
\textit{noise} is the amount of error
Adding More Groups

- Let’s add a third group: occupational therapy (OT)
- How do we check for differences?
- Can do $t$-tests comparing:
  - TAU versus PSI
  - TAU versus OT
  - PSI versus OT
Problems with Many $t$-tests

• There are two problems with multiple $t$-tests:
  – Inflates probability of Type I error
  – Tests not independent of each other
    • If PSI > OT, and
    • OT > TAU, then
    • PSI must be > TAU
One-Way ANOVA

• Solution is one-way Analysis of Variance (ANOVA)
• Extension of t-test for >2 groups
• But – how do we calculate equivalent of 
  \((M_1 - M_2)\) when there are > 2 groups?
Difference for >2 Means

TAU

PSI

OT

0  5  10  15  20  25  30  35
Difference for >2 Means

• Step 1: Find the Grand Mean ($M_G$) – the mean of the 3 means; or the mean of all people, irrespective of group

  – Assuming all groups have same sample size, it is \((12.5 + 22.5 + 25.0) / 3 = 60 / 3 = 20.0\)
Difference for >2 Means

Grand Mean

TAU

PSI

OT
Difference for >2 Means

- Step 1: Find the Grand Mean ($M_G$)
- Step 2: Determine the deviation ($d$) of each mean from $M_G$
Difference for >2 Means

Grand Mean

\[ d_1 \]

\[ d_2 \]

\[ d_3 \]
Difference for >2 Means

• Step 1: Find the Grand Mean ($M_G$)
• Step 2: Determine the deviation ($d$) of each mean from $M_G$

- $d_1 = (12.5 - 20.0) = -7.5$
- $d_2 = (22.5 - 20.0) = 2.5$
- $d_3 = (25.0 - 20.0) = 5.0$
Difference for >2 Means

- We can’t just add them up
- The answer will always be 0
  - E.g., (-7.5) + (2.5) + (5.0) = 0
- Square each value first
- Divide by number of means
Difference for >2 Means

- Step 1: Find the Grand Mean \((M_G)\)
- Step 2: Determine the deviation \((d)\) of each mean from \(M_G\)
- Step 3: Determine mean (average) squared deviation
Difference for >2 Means

- This is the numerator
- It is a variance; in this case, the variance between the groups
- The more the means are spread out, the larger the between-groups variance (i.e., the signal)
Difference for >2 Means

- What’s the denominator (the noise)?
  - As with the *t*-test, it is the variance *within* each of the groups
- So, ANOVA is again signal-to-noise ratio
- For various reasons, we call the variance the *mean square*
One-Way ANOVA

• So, the ANOVA is:

\[
F = \frac{\text{Mean Square Between Groups}}{\text{Mean Square Within Groups}}
\]
Conceptually

- There is *variance* (variability) among all of the outcomes
- ANOVA tries to find the sources of this variance:
  - Due to differences between group
  - Variability within the groups (error)
  - “Error” better termed “unexplained variance”
Beyond One-Way ANOVA

- Can look for other factors to account for (explain) error variance; e.g., sex
- Design becomes more complicated
## Factorial ANOVA

<table>
<thead>
<tr>
<th></th>
<th>TAU</th>
<th>PSI</th>
<th>OT</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Males</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Females</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Factorial ANOVA

- Can add more factors
- Limited by sample size, ability to understand results
- Adding factors that don’t explain variance make the results worse
- Best guide is judgement