two-sample t-tests

Prof. Patrick Bennett

Concepts from previous lectures

- t distribution
- standard error of the mean
- degrees-of-freedom
- Null and alternative/research hypotheses (H0 vs H1)

Hand et al (1994)

- Experimental Question:
  - Is family therapy an effective treatment for anorexia?
- 17 girls participated in study
  - weighed before & after treatment
  - weights (in pounds given in Table 13.1)
- Statistical Question:
  - Does before/after weight differ?
Hand et al. (1994)

- N=17
- subjects weighed before & after family therapy
- diff = after - before

before & after measures are not independent because they come from the same subject

$t$ tests depend, in part, on $N$ (e.g., df)

should $N_{\text{total}} = N_{\text{before}} + N_{\text{after}}$?

no, because the 2 sets of measures are not independent

our analysis must take dependence into account

simple solution: analyze difference scores

### Table

<table>
<thead>
<tr>
<th>subject</th>
<th>before</th>
<th>after</th>
<th>diff</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>83.8</td>
<td>95.2</td>
<td>11.4</td>
</tr>
<tr>
<td>2</td>
<td>83.3</td>
<td>94.3</td>
<td>11</td>
</tr>
<tr>
<td>3</td>
<td>86</td>
<td>91.5</td>
<td>5.5</td>
</tr>
<tr>
<td>4</td>
<td>82.5</td>
<td>91.9</td>
<td>9.4</td>
</tr>
<tr>
<td>5</td>
<td>86.7</td>
<td>100.3</td>
<td>13.6</td>
</tr>
<tr>
<td>6</td>
<td>79.6</td>
<td>76.7</td>
<td>-2.9</td>
</tr>
<tr>
<td>7</td>
<td>76.9</td>
<td>76.8</td>
<td>-0.1</td>
</tr>
<tr>
<td>8</td>
<td>94.2</td>
<td>101.6</td>
<td>7.4</td>
</tr>
<tr>
<td>9</td>
<td>73.4</td>
<td>94.9</td>
<td>21.5</td>
</tr>
<tr>
<td>10</td>
<td>80.5</td>
<td>75.2</td>
<td>-5.3</td>
</tr>
<tr>
<td>11</td>
<td>81.6</td>
<td>77.8</td>
<td>-3.8</td>
</tr>
<tr>
<td>12</td>
<td>82.1</td>
<td>95.5</td>
<td>13.4</td>
</tr>
<tr>
<td>13</td>
<td>77.6</td>
<td>90.7</td>
<td>13.1</td>
</tr>
<tr>
<td>14</td>
<td>83.5</td>
<td>92.5</td>
<td>9</td>
</tr>
<tr>
<td>15</td>
<td>89.9</td>
<td>93.8</td>
<td>3.9</td>
</tr>
<tr>
<td>16</td>
<td>86</td>
<td>91.7</td>
<td>5.7</td>
</tr>
<tr>
<td>17</td>
<td>87.3</td>
<td>98</td>
<td>10.7</td>
</tr>
</tbody>
</table>

- $r = 0.54$
- $r_s = 0.61$

### Graph

- Difference (after-before) lbs.
- Weight After Therapy (lbs.)
- Weight Before Therapy (lbs.)

- mean: $\bar{D} = 7.26$
- standard deviation: $s_D = 7.16$
Hand et al (1994)

- Hypotheses for scores:
  - H0: $μ_{\text{Before}} = μ_{\text{After}}$
  - H1: $μ_{\text{Before}} \neq μ_{\text{After}}$
- for difference scores:
  - H0: $(μ_{\text{After}} - μ_{\text{Before}}) = μ_D = 0$
  - H1: $(μ_{\text{After}} - μ_{\text{Before}}) = μ_D \neq 0$

- what is our observed value of $t$?
- what are our critical values of $t$?

$t = \frac{\bar{D} - μ_D}{s_D} = \frac{\bar{D} - 0}{\frac{s_D}{\sqrt{N}}}$

mean: $\bar{D} = 7.26$

standard deviation: $s_D = 7.16$

$t = \frac{7.26 - 0}{7.16} = \frac{7.26}{1.74} = 4.18$

$df = N - 1 = 17 - 1 = 16$

$t = 7.26 - 0 \quad \bar{D} = 7.26 \quad \frac{7.16}{17} \quad 1.74 \quad 4.18$

$α = .05$ (2-tailed), $df=16$, $t_{\text{critical}} = \pm 2.12$

• is family therapy an effective treatment for anorexia?
• 17 participants weighed before & after therapy
• $t$ test used to evaluate H0 of no change in weight
  - rejected H0 in favour of H1 (i.e., weight change ≠ zero)
  - direction of effect (after > before) means weight gain not loss
• conclude that family therapy is/was an effective treatment?
  - can you think of an alternative explanation of result?
• alternative explanation: weight gain was due to normal growth of time
Hand et al (1994)

• is family therapy an effective treatment for anorexia?
  - rejected H0 in favour of H1 (i.e., weight change ≠ zero)
  - but this is weak evidence for an effect of therapy
  - because simple alternative explanation exists:
    ‣ weight gain was due to normal growth over time
• experiment needs a control group
  - experimental studies: a group that does not receive the treatment/procedure of interest
  - correlational studies: a group that is not exposed to or does not experience the variable of interest
• control group in Hand et al study:
  - set of individuals who do not receive therapy
  - ideally, individuals would be assigned randomly to therapy and no-therapy groups. Why?

Hand et al (1994)

• is family therapy an effective treatment for anorexia?
• original experiment did include a control group
  - experimental group: N=17, received family therapy
  - control group: N=26, did not receive family therapy
• new experimental hypothesis:
  - was weight gain different in the two groups?
  - let μ FT & μ C represent mean weight gain in 2 groups
  ‣ H0: μ FT = μ C
  ‣ H1: μ FT ≠ μ C

\[
\begin{align*}
\text{Hand et al. (1994)} \\
\end{align*}
\]

\[
\begin{align*}
\text{Hand et al. (1994)} \\
\end{align*}
\]

\[
\begin{align*}
\text{Sampling distribution of difference between means} \\
\end{align*}
\]

• therapy (N=17):
  - mean = 7.26, sd = 7.16
• control (N=26):
  - mean = -0.45, sd = 7.99
• is difference between group means unusually large if H0 (μ FT = μ C) is true?

\[
\begin{align*}
\text{Sampling distribution of difference between means} \\
\end{align*}
\]

• Each mean has a sampling distribution
• Difference between means also has a sampling distribution
  - μ D = μ 1 - μ 2, VAR D = VAR 1 + VAR 2
• Variance of difference between 2 independent variables equals the sum of their variances
  - if means are distributed normally, then difference (or sum) is distributed normally
  ‣ Central Limit Theorem
• From our samples, we can estimate distribution of difference between group means
Is our observed difference between group means unusually large given the null hypothesis that the two groups do not differ?

We will try to answer this question using a t test.

t test for 2 independent samples

calculate difference between means to a t statistic

\[
t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{s_{\bar{X}_1 - \bar{X}_2}} = \frac{\bar{X}_1 - \bar{X}_2 - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}
\]

\[
t = \frac{\hat{\mu}_D - \mu_D}{\hat{\sigma}_D}
\]

when H0 is \( \mu_D = (\mu_1 - \mu_2) = 0 \)

We use a slightly different formula when \( n_1 \neq n_2 \)

"pooled" variance estimate

\[
s_{\bar{X}_1 - \bar{X}_2} = \sqrt{\frac{s_p^2}{n_1} + \frac{s_p^2}{n_2}}
\]

\[
s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}
\]

\[
s_{\bar{X}_1 - \bar{X}_2}^2 = \sqrt{\frac{s_p^2}{n_1} + \frac{s_p^2}{n_2}}
\]

t test for independent means

\[
t = \frac{\bar{X}_1 - \bar{X}_2}{s_p} \quad df = (n_1 - 1) + (n_2 - 1) = n_1 + n_2 - 2
\]

when H0 is true (\( \mu_1 - \mu_2 = 0 \)), t statistic follows t distribution with \( n_1 + n_2 - 2 \) degrees of freedom
Hand et al. (1994)

H0: $\mu_{FT} = \mu_{C}$
H1: $\mu_{FT} \neq \mu_{C}$

$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{s^2_{1/n_1} + s^2_{2/n_2}}}$

$s^2 = \frac{25(7.99^2) + 16(7.16^2)}{26 + 17 - 2} = 58.9$

$t = \frac{7.26 - (-0.45)}{\sqrt{58.9/26 + 58.9/17}} = 7.71/2.39 = 3.22$

$df = 26 + 17 - 2 = 41$

Is the observed value of $t = 3.22$ unusual given that the Null hypothesis is true?

Control Group
mean = -0.45
sd = 7.99
n = 26

Therapy Group
mean = 7.26
sd = 7.16
n = 17

General Strategy
reject H0 if $t$ exceeds critical values of $t$

- significance level = .05
- 2-tailed test
- df = 41
- critical $t = \pm ??$

General Strategy
reject H0 if $t$ exceeds critical values of $t$

- is family therapy an effective treatment for anorexia?
- measured weight gain in control and therapy groups
- difference between groups was significant
  - $t(41) = 3.22$, p<.05, 2-tailed
  - reject null hypothesis that weight gain was the same in 2 groups
- result is consistent with hypothesis that family therapy is an effective treatment for anorexia
- next steps in research program?
  - replicate findings!
  - "Without replication, all results should be taken as preliminary." -- Gary Marcus
  * https://www.newyorker.com/news/news-desk/cleaning-up-science