Testing Paired Data

Kevin Hayes, Anthony Kinsella,
and Kevin O'Brien

Department of Mathematics and Statistics,
University *of* Limerick.

The 2010 Conference of Applied Statistics in Ireland, May 16*th* to 18*th*.

Outline

Outline

1. A Look at Some Tests for Paired-Samples
2. An Alternative Paired-Sampes t-Test
3. Power Functions
4. Concluding Remarks

Outline

Outline

1. A Look at Some Tests for Paired-Samples
2. An Alternative Paired-Sampes t-Test
3. Power Functions
4. Concluding Remarks

Let Y_1 and Y_2 be distributed bivariate normal: $\mathrm{E}(Y_1) = \mu_1$, $\mathrm{E}(Y_2) = \mu_2$, $\mathrm{var}(Y_1) = \sigma_1^2$, $\mathrm{var}(Y_2) = \sigma_2^2$, and $-1 < \rho_{12} < 1$.

Of interest are the tests of the hypotheses:

1. H$'$: $\mu_1 - \mu_2 = 0$, ?;
2. H$''$: $\sigma_1^2 = \sigma_2^2$, ?, ?;
3. H$^{\mathrm{J}}$: $\mu_1 - \mu_2 = 0$ and $\sigma_1^2 = \sigma_2^2$, ?.

Define $D = Y_1 - Y_2$ and $S = Y_1 + Y_2$

$$
\begin{aligned}
\mathrm{E}(D) &= \mu_D = \mu_1 - \mu_2 \\
\mathrm{E}(S) &= \mu_S = \mu_1 + \mu_2 \\
\mathrm{var}(D) &= \sigma_D^2 = \sigma_1^2 + \sigma_2^2 - 2\rho_{12}\sigma_1\sigma_2 \\
\mathrm{var}(S) &= \sigma_S^2 = \sigma_1^2 + \sigma_2^2 + 2\rho_{12}\sigma_1\sigma_2 \\
\mathrm{cov}(D, S) &= \sigma_1^2 - \sigma_2^2
\end{aligned}
$$

Let Y_1 and Y_2 be distributed bivariate normal: $\mathrm{E}(Y_1) = \mu_1$, $\mathrm{E}(Y_2) = \mu_2$, $\mathrm{var}(Y_1) = \sigma_1^2$, $\mathrm{var}(Y_2) = \sigma_2^2$, and $-1 < \rho_{12} < 1$.

Of interest are the tests of the hypotheses:

1. H': $\mu_1 - \mu_2 = 0$, ?;
2. H'': $\sigma_1^2 = \sigma_2^2$, ?, ?;
3. H^{J}: $\mu_1 - \mu_2 = 0$ **and** $\sigma_1^2 = \sigma_2^2$, ?.

Define $D = Y_1 - Y_2$ and $S = Y_1 + Y_2$

$$
\begin{aligned}
\mathrm{E}(D) &= \mu_D = \mu_1 - \mu_2 \\
\mathrm{E}(S) &= \mu_S = \mu_1 + \mu_2,
\end{aligned}
$$

Let Y_1 and Y_2 be distributed bivariate normal: $\mathrm{E}(Y_1) = \mu_1$, $\mathrm{E}(Y_2) = \mu_2$, $\mathrm{var}(Y_1) = \sigma_1^2$, $\mathrm{var}(Y_2) = \sigma_2^2$, and $-1 < \rho_{12} < 1$.

Of interest are the tests of the hypotheses:

1. H': $\mu_1 - \mu_2 = 0$, ?;
2. H'': $\sigma_1^2 = \sigma_2^2$, ?, ?;
3. H^{J}: $\mu_1 - \mu_2 = 0$ **and** $\sigma_1^2 = \sigma_2^2$, ?.

Define $D = Y_1 - Y_2$ and $S = Y_1 + Y_2$

$$
\begin{aligned}
\mathrm{E}(D) &= \mu_D = \mu_1 - \mu_2 \\
\mathrm{E}(S) &= \mu_S = \mu_1 + \mu_2, \\
\mathrm{var}(D) &= \sigma_D^2 = \sigma_1^2 + \sigma_2^2 - 2\rho_{12}\sigma_1\sigma_2, \\
\mathrm{var}(S) &= \sigma_S^2 = \sigma_1^2 + \sigma_2^2 + 2\rho_{12}\sigma_1\sigma_2, \\
\mathrm{cov}(D, S) &= \sigma_1^2 - \sigma_2^2.
\end{aligned}
$$

Let Y_1 and Y_2 be distributed bivariate normal: $\mathrm{E}(Y_1) = \mu_1$, $\mathrm{E}(Y_2) = \mu_2$, $\mathrm{var}(Y_1) = \sigma_1^2$, $\mathrm{var}(Y_2) = \sigma_2^2$, and $-1 < \rho_{12} < 1$.

Of interest are the tests of the hypotheses:

1. H': $\mu_1 - \mu_2 = 0$, ?;
2. H'': $\sigma_1^2 = \sigma_2^2$, ?, ?;
3. H^{J}: $\mu_1 - \mu_2 = 0$ **and** $\sigma_1^2 = \sigma_2^2$, ?.

Define $D = Y_1 - Y_2$ and $S = Y_1 + Y_2$

$$\begin{aligned}
\mathrm{E}(D) &= \mu_D = \mu_1 - \mu_2 \\
\mathrm{E}(S) &= \mu_S = \mu_1 + \mu_2, \\
\mathrm{var}(D) &= \sigma_D^2 = \sigma_1^2 + \sigma_2^2 - 2\rho_{12}\sigma_1\sigma_2, \\
\mathrm{var}(S) &= \sigma_S^2 = \sigma_1^2 + \sigma_2^2 + 2\rho_{12}\sigma_1\sigma_2, \\
\mathrm{cov}(D, S) &= \sigma_1^2 - \sigma_2^2.
\end{aligned}$$

Let Y_1 and Y_2 be distributed bivariate normal: $\mathrm{E}(Y_1) = \mu_1$, $\mathrm{E}(Y_2) = \mu_2$, $\mathrm{var}(Y_1) = \sigma_1^2$, $\mathrm{var}(Y_2) = \sigma_2^2$, and $-1 < \rho_{12} < 1$.

Of interest are the tests of the hypotheses:

1. H': $\mu_1 - \mu_2 = 0$, ?;
2. H'': $\sigma_1^2 = \sigma_2^2$, ?, ?;
3. H^{J}: $\mu_1 - \mu_2 = 0$ **and** $\sigma_1^2 = \sigma_2^2$, ?.

Define $D = Y_1 - Y_2$ and $S = Y_1 + Y_2$

$$
\begin{aligned}
\mathrm{E}(D) &= \mu_D = \mu_1 - \mu_2 \\
\mathrm{E}(S) &= \mu_S = \mu_1 + \mu_2, \\
\mathrm{var}(D) &= \sigma_D^2 = \sigma_1^2 + \sigma_2^2 - 2\rho_{12}\sigma_1\sigma_2, \\
\mathrm{var}(S) &= \sigma_S^2 = \sigma_1^2 + \sigma_2^2 + 2\rho_{12}\sigma_1\sigma_2, \\
\mathrm{cov}(D, S) &= \sigma_1^2 - \sigma_2^2.
\end{aligned}
$$

Let Y_1 and Y_2 be distributed bivariate normal: $\mathrm{E}(Y_1) = \mu_1$, $\mathrm{E}(Y_2) = \mu_2$, $\mathrm{var}(Y_1) = \sigma_1^2$, $\mathrm{var}(Y_2) = \sigma_2^2$, and $-1 < \rho_{12} < 1$.

Of interest are the tests of the hypotheses:

1. H': $\mu_1 - \mu_2 = 0$, ?;
2. H'': $\sigma_1^2 = \sigma_2^2$, ?, ?;
3. H^{J}: $\mu_1 - \mu_2 = 0$ and $\sigma_1^2 = \sigma_2^2$, ?.

Define $D = Y_1 - Y_2$ and $S = Y_1 + Y_2$

$$
\begin{aligned}
\mathrm{E}(D) &= \mu_D = \mu_1 - \mu_2 \\
\mathrm{E}(S) &= \mu_S = \mu_1 + \mu_2, \\
\mathrm{var}(D) &= \sigma_D^2 = \sigma_1^2 + \sigma_2^2 - 2\rho_{12}\sigma_1\sigma_2, \\
\mathrm{var}(S) &= \sigma_S^2 = \sigma_1^2 + \sigma_2^2 + 2\rho_{12}\sigma_1\sigma_2, \\
\mathrm{cov}(D, S) &= \sigma_1^2 - \sigma_2^2.
\end{aligned}
$$

Let Y_1 and Y_2 be distributed bivariate normal: $\mathrm{E}(Y_1) = \mu_1$, $\mathrm{E}(Y_2) = \mu_2$, $\mathrm{var}(Y_1) = \sigma_1^2$, $\mathrm{var}(Y_2) = \sigma_2^2$, and $-1 < \rho_{12} < 1$.

Of interest are the tests of the hypotheses:

1. H′: $\mu_1 - \mu_2 = 0$, ?;
2. H″: $\sigma_1^2 = \sigma_2^2$, ?, ?;
3. H^{J}: $\mu_1 - \mu_2 = 0$ **and** $\sigma_1^2 = \sigma_2^2$, ?.

Define $D = Y_1 - Y_2$ and $S = Y_1 + Y_2$

$$
\begin{aligned}
\mathrm{E}(D) &= \mu_D = \mu_1 - \mu_2 \\
\mathrm{E}(S) &= \mu_S = \mu_1 + \mu_2, \\
\mathrm{var}(D) &= \sigma_D^2 = \sigma_1^2 + \sigma_2^2 - 2\rho_{12}\sigma_1\sigma_2, \\
\mathrm{var}(S) &= \sigma_S^2 = \sigma_1^2 + \sigma_2^2 + 2\rho_{12}\sigma_1\sigma_2, \\
\mathrm{cov}(D, S) &= \sigma_1^2 - \sigma_2^2.
\end{aligned}
$$

Let Y_1 and Y_2 be distributed bivariate normal: $\mathrm{E}(Y_1) = \mu_1$, $\mathrm{E}(Y_2) = \mu_2$, $\mathrm{var}(Y_1) = \sigma_1^2$, $\mathrm{var}(Y_2) = \sigma_2^2$, and $-1 < \rho_{12} < 1$.

Of interest are the tests of the hypotheses:

1. H': $\mu_1 - \mu_2 = 0$, ?;
2. H'': $\sigma_1^2 = \sigma_2^2$, ?, ?;
3. H^{J}: $\mu_1 - \mu_2 = 0$ **and** $\sigma_1^2 = \sigma_2^2$, ?.

Define $D = Y_1 - Y_2$ and $S = Y_1 + Y_2$

$$
\begin{aligned}
\mathrm{E}(D) &= \mu_D = \mu_1 - \mu_2 \\
\mathrm{E}(S) &= \mu_S = \mu_1 + \mu_2, \\
\mathrm{var}(D) &= \sigma_D^2 = \sigma_1^2 + \sigma_2^2 - 2\rho_{12}\sigma_1\sigma_2, \\
\mathrm{var}(S) &= \sigma_S^2 = \sigma_1^2 + \sigma_2^2 + 2\rho_{12}\sigma_1\sigma_2, \\
\mathrm{cov}(D, S) &= \sigma_1^2 - \sigma_2^2.
\end{aligned}
$$

Let Y_1 and Y_2 be distributed bivariate normal: $\mathrm{E}(Y_1) = \mu_1$, $\mathrm{E}(Y_2) = \mu_2$, $\mathrm{var}(Y_1) = \sigma_1^2$, $\mathrm{var}(Y_2) = \sigma_2^2$, and $-1 < \rho_{12} < 1$.

Of interest are the tests of the hypotheses:

1. H': $\mu_1 - \mu_2 = 0$, **?**;
2. H'': $\sigma_1^2 = \sigma_2^2$, **?**, **?**;
3. H^{J}: $\mu_1 - \mu_2 = 0$ **and** $\sigma_1^2 = \sigma_2^2$, **?**.

Define $D = Y_1 - Y_2$ and $S = Y_1 + Y_2$

$$
\begin{aligned}
\mathrm{E}(D) &= \mu_D = \mu_1 - \mu_2 \\
\mathrm{E}(S) &= \mu_S = \mu_1 + \mu_2, \\
\mathrm{var}(D) &= \sigma_D^2 = \sigma_1^2 + \sigma_2^2 - 2\rho_{12}\sigma_1\sigma_2, \\
\mathrm{var}(S) &= \sigma_S^2 = \sigma_1^2 + \sigma_2^2 + 2\rho_{12}\sigma_1\sigma_2, \\
\mathrm{cov}(D, S) &= \sigma_1^2 - \sigma_2^2.
\end{aligned}
$$

H′: $\mu_1 - \mu_2 = 0$, ?

The test of the hypothesis H′: $\mu_D = 0$ due to ? is based on the t-ratio

$$T^* = \frac{\bar{D}}{\hat{\sigma}_D/\sqrt{n}} \sim t_{(n-1)\mathrm{df}}, \tag{1}$$

where $\bar{D}$ is the mean of the n case-wise differences $D_i = Y_{i1} - Y_{i2}$ of the pairs (Y_{i1}, Y_{i2}), $i = 1, \ldots, n$ and

$$\hat{\sigma}_D^2 = \frac{1}{n-1} \sum_{i=1}^{n} (D_i - \bar{D})^2$$

is the usual estimate of the variance of the differences σ_D^2.

This classical test makes no assumptions about the equality (or otherwise) of the variance parameters σ_1^2 and σ_2^2.

H′: $\mu_1 - \mu_2 = 0$, ?

The test of the hypothesis H′: $\mu_D = 0$ due to ? is based on the t-ratio

$$T^* = \frac{\bar{D}}{\hat{\sigma}_D/\sqrt{n}} \sim t_{(n-1)\mathrm{df}}, \tag{1}$$

where $\bar{D}$ is the mean of the n case-wise differences $D_i = Y_{i1} - Y_{i2}$ of the pairs (Y_{i1}, Y_{i2}), $i = 1, \ldots, n$ and

$$\hat{\sigma}_D^2 = \frac{1}{n-1} \sum_{i=1}^{n} (D_i - \bar{D})^2$$

is the usual estimate of the variance of the differences σ_D^2.

This classical test makes no assumptions about the equality (or otherwise) of the variance parameters σ_1^2 and σ_2^2.

H′′: $\sigma_1^2 = \sigma_2^2$, ?, ?

The test of the hypothesis that the variances σ_1^2 and σ_2^2 are equal, which was devised concurrently by ? and ?, is based on the correlation of D with S, the coefficient being $\rho_{DS} = (\sigma_1^2 - \sigma_2^2)/(\sigma_D \sigma_S)$, which is zero if, and only if, $\sigma_1^2 = \sigma_2^2$. Consequently a test of H'': $\sigma_1^2 = \sigma_2^2$ is equivalent to a test of H'': $\rho_{DS} = 0$ and the test statistic is the familiar t-test for a correlation coefficient with $(n-2)$ degrees of freedom:

$$T^*_{\mathrm{PM}} = R\sqrt{\frac{n-2}{1-R^2}}, \tag{2}$$

where

$$R = \frac{\sum(D_i - \bar{D})(S_i - \bar{S})}{[\sum(D_i - \bar{D})^2 \sum(S_i - \bar{S})^2]^{\frac{1}{2}}} \tag{3}$$

is the sample correlation coefficient of the n case-wise differences D_i and sums $S_i = Y_{i1} + Y_{i2}$.

H^{J}: $\mu_1 - \mu_2 = 0$ and $\sigma_1^2 = \sigma_2^2$, ?

They write the conditional expectation of D given S as

$$\begin{aligned} \mathrm{E}(D|S) &= \mu_D + \mathrm{cov}(D,S)\mathrm{var}(S)^{-1}(S - \mu_S), \\ &= \mu_D + [(\sigma_1^2 - \sigma_2^2)/\sigma_S^2](S - \mu_S), \\ &= \beta_0 + \beta_1 S, \end{aligned}$$

where $\beta_0 = \mu_D - [(\sigma_1^2 - \sigma_2^2)/\sigma_S^2]\mu_S$ and $\beta_1 = (\sigma_1^2 - \sigma_2^2)/\sigma_S^2$.

The hypothesis H^{J}: $\mu_D = 0$ and $\sigma_1^2 = \sigma_2^2$, is true *iff* $\beta_0 = \beta_1 = 0$.

The test statistic is

$$F^* = \left(\frac{n-2}{2}\right)\left(\frac{\sum D_i^2 - \mathrm{SSE}}{\mathrm{SSE}}\right) \sim F_{(2,n-2)\mathrm{df}}, \qquad (4)$$

where SSE is the residual error sum-of-squares from $\hat{D}_i = \hat{\beta}_0 + \hat{\beta}_1 s_i$.

$\mathrm{H^J}$: $\mu_1 - \mu_2 = 0$ **and** $\sigma_1^2 = \sigma_2^2$, ?

They write the conditional expectation of D given S as

$$\begin{aligned}
\mathrm{E}(D|S) &= \mu_D + \mathrm{cov}(D,S)\mathrm{var}(S)^{-1}(S - \mu_S), \\
&= \mu_D + [(\sigma_1^2 - \sigma_2^2)/\sigma_S^2](S - \mu_S), \\
&= \beta_0 + \beta_1 S,
\end{aligned}$$

where $\beta_0 = \mu_D - [(\sigma_1^2 - \sigma_2^2)/\sigma_S^2]\mu_S$ and $\beta_1 = (\sigma_1^2 - \sigma_2^2)/\sigma_S^2$.

The hypothesis $\mathrm{H^J}$: $\mu_D = 0$ and $\sigma_1^2 = \sigma_2^2$, is true *iff* $\beta_0 = \beta_1 = 0$.

The test statistic is

$$F^* = \left(\frac{n-2}{2}\right)\left(\frac{\sum D_i^2 - \mathrm{SSE}}{\mathrm{SSE}}\right) \sim F_{(2,n-2)\mathrm{df}}, \tag{4}$$

where SSE is the residual error sum-of-squares from $\hat{D}_i = \hat{\beta}_0 + \hat{\beta}_1 s_i$.

H^{J}: $\mu_1 - \mu_2 = 0$ **and** $\sigma_1^2 = \sigma_2^2$, **?**

They write the conditional expectation of D given S as

$$\begin{aligned} \mathrm{E}(D|S) &= \mu_D + \mathrm{cov}(D,S)\mathrm{var}(S)^{-1}(S - \mu_S), \\ &= \mu_D + [(\sigma_1^2 - \sigma_2^2)/\sigma_S^2](S - \mu_S), \\ &= \beta_0 + \beta_1 S, \end{aligned}$$

where $\beta_0 = \mu_D - [(\sigma_1^2 - \sigma_2^2)/\sigma_S^2]\mu_S$ and $\beta_1 = (\sigma_1^2 - \sigma_2^2)/\sigma_S^2$.

The hypothesis H^{J}: $\mu_D = 0$ and $\sigma_1^2 = \sigma_2^2$, is true *iff* $\beta_0 = \beta_1 = 0$.

The test statistic is

$$F^* = \left(\frac{n-2}{2}\right)\left(\frac{\sum D_i^2 - \mathrm{SSE}}{\mathrm{SSE}}\right) \sim F_{(2,n-2)\mathrm{df}}, \tag{4}$$

where SSE is the residual error sum-of-squares from $\hat{D}_i = \hat{\beta}_0 + \hat{\beta}_1 s_i$.

H^{J}: $\mu_1 - \mu_2 = 0$ and $\sigma_1^2 = \sigma_2^2$, ?

They write the conditional expectation of D given S as

$$\begin{aligned} \mathrm{E}(D|S) &= \mu_D + \mathrm{cov}(D,S)\mathrm{var}(S)^{-1}(S-\mu_S), \\ &= \mu_D + [(\sigma_1^2 - \sigma_2^2)/\sigma_S^2](S-\mu_S), \\ &= \beta_0 + \beta_1 S, \end{aligned}$$

where $\beta_0 = \mu_D - [(\sigma_1^2 - \sigma_2^2)/\sigma_S^2]\mu_S$ and $\beta_1 = (\sigma_1^2 - \sigma_2^2)/\sigma_S^2$.

The hypothesis H^{J}: $\mu_D = 0$ and $\sigma_1^2 = \sigma_2^2$, is true *iff* $\beta_0 = \beta_1 = 0$.

The test statistic is

$$F^* = (\frac{n-2}{2})(\frac{\sum D_i^2 - \mathrm{SSE}}{\mathrm{SSE}}) \sim F_{(2,n-2)\mathrm{df}}, \tag{4}$$

where SSE is the residual error sum-of-squares from $\hat{D}_i = \hat{\beta}_0 + \hat{\beta}_1 s_i$.

Outline

Write

$$\begin{aligned}\mu_{D|S=s} &= \mu_D + [(\sigma_1^2 - \sigma_2^2)/\sigma_S^2](s - \mu_S) \\ &= \beta_0 + \beta_1 s\end{aligned}$$

where $\beta_1 = (\sigma_1^2 - \sigma_2^2)/\sigma_S^2$ and $\beta_0 = \mu_D - \beta_1\mu_S$ with $\sigma_{D|S}^2 = \sigma_D^2 - (\sigma_1^2 - \sigma_2^2)^2/\sigma_S^2$.

Can also be written $D_i = \beta_0 + \beta_1 s_i + \varepsilon_i$ $(i = 1, \ldots, n)$ where ε_i are IID $\mathrm{N}(0, \sigma_\varepsilon^2)$ distribution having $\sigma_\varepsilon^2 = \sigma_{D|S}^2$, and the s_i terms are the values in the conditioning set $S_1 = s_1, S_2 = s_2, \ldots, S_n = s_n$ and are restricted only by the exclusion of the trivial case $s_1 = s_2 = \cdots = s_n$.

Can be reformulated as $D_i = \alpha + \beta_1(s_i - \bar{s}) + \varepsilon_i$ where the intercept parameters are related by $\alpha = \mu_D + \beta_1(\bar{s} - \mu_S)$, and where the slope parameters and the error structures of the two regression models are identical.

Write

$$\begin{aligned} \mu_{D|S=s} &= \mu_D + [(\sigma_1^2 - \sigma_2^2)/\sigma_S^2](s - \mu_S) \\ &= \beta_0 + \beta_1 s \end{aligned}$$

where $\beta_1 = (\sigma_1^2 - \sigma_2^2)/\sigma_S^2$ and $\beta_0 = \mu_D - \beta_1 \mu_S$ with $\sigma_{D|S}^2 = \sigma_D^2 - (\sigma_1^2 - \sigma_2^2)^2/\sigma_S^2$.

Can also be written $D_i = \beta_0 + \beta_1 s_i + \varepsilon_i$ $(i = 1, \ldots, n)$ where ε_i are IID $\mathrm{N}(0, \sigma_\varepsilon^2)$ distribution having $\sigma_\varepsilon^2 = \sigma_{D|S}^2$, and the s_i terms are the values in the conditioning set $S_1 = s_1, S_2 = s_2, \ldots, S_n = s_n$ and are restricted only by the exclusion of the trivial case $s_1 = s_2 = \cdots = s_n$.

Can be reformulated as $D_i = \alpha + \beta_1(s_i - \bar{s}) + \varepsilon_i$ where the intercept parameters are related by $\alpha = \mu_D + \beta_1(\bar{s} - \mu_S)$, and where the slope parameters and the error structures of the two regression models are identical.

Write

$$\begin{aligned}\mu_{D|S=s} &= \mu_D + [(\sigma_1^2 - \sigma_2^2)/\sigma_S^2](s - \mu_S) \\ &= \beta_0 + \beta_1 s\end{aligned}$$

where $\beta_1 = (\sigma_1^2 - \sigma_2^2)/\sigma_S^2$ and $\beta_0 = \mu_D - \beta_1\mu_S$ with $\sigma_{D|S}^2 = \sigma_D^2 - (\sigma_1^2 - \sigma_2^2)^2/\sigma_S^2$.

Can also be written $D_i = \beta_0 + \beta_1 s_i + \varepsilon_i$ $(i = 1, \ldots, n)$ where ε_i are IID $\mathrm{N}(0, \sigma_\varepsilon^2)$ distribution having $\sigma_\varepsilon^2 = \sigma_{D|S}^2$, and the s_i terms are the values in the conditioning set $S_1 = s_1, S_2 = s_2, \ldots, S_n = s_n$ and are restricted only by the exclusion of the trivial case $s_1 = s_2 = \cdots = s_n$.

Can be reformulated as $D_i = \alpha + \beta_1(s_i - \bar{s}) + \varepsilon_i$ where the intercept parameters are related by $\alpha = \mu_D + \beta_1(\bar{s} - \mu_S)$, and where the slope parameters and the error structures of the two regression models are identical.

Treating the s_i terms as fixed known constants the MLEs are:

$$\begin{aligned}\hat{\alpha} &= \bar{D},\\ \hat{\beta}_1 &= \frac{\sum(D_i - \bar{D})(s_i - \bar{s})}{\sum(s_i - \bar{s})^2}\\ &= \frac{\sum D_i(s_i - \bar{s})}{\sum(s_i - \bar{s})^2},\\ \hat{\sigma}_\varepsilon^2 &= \frac{1}{n}\Big(\sum(D_i - \bar{D})^2 - \frac{[\sum(s_i - \bar{s})D_i]^2}{\sum(s_i - \bar{s})^2} \Big)\end{aligned}$$

The estimates $\hat{\alpha}$ and $\hat{\beta}_1$ are independent and have distributions $\mathrm{N}(\alpha, n^{-1}\sigma_\varepsilon^2)$ and $\mathrm{N}(\beta_1, [\sum(s_i - \bar{s})^2]^{-1}\sigma_\varepsilon^2)$.

Treating the s_i terms as fixed known constants the MLEs are:

$$
\begin{aligned}
\hat{\alpha} &= \bar{D}, \\
\hat{\beta}_1 &= \frac{\sum(D_i - \bar{D})(s_i - \bar{s})}{\sum(s_i - \bar{s})^2} \\
&= \frac{\sum D_i(s_i - \bar{s})}{\sum(s_i - \bar{s})^2}, \\
\hat{\sigma}_\varepsilon^2 &= \frac{1}{n}\Big(\sum(D_i - \bar{D})^2 - \frac{[\sum(s_i - \bar{s})D_i]^2}{\sum(s_i - \bar{s})^2} \Big)
\end{aligned}
$$

The estimates $\hat{\alpha}$ and $\hat{\beta}_1$ are independent and have distributions $\mathrm{N}(\alpha, n^{-1}\sigma_\varepsilon^2)$ and $\mathrm{N}(\beta_1, [\sum(s_i - \bar{s})^2]^{-1}\sigma_\varepsilon^2)$.

Treating the s_i terms as fixed known constants the MLEs are:

$$\hat{\alpha} = \bar{D},$$

$$\begin{aligned}\hat{\beta}_1 &= \frac{\sum(D_i - \bar{D})(s_i - \bar{s})}{\sum(s_i - \bar{s})^2} \\ &= \frac{\sum D_i(s_i - \bar{s})}{\sum(s_i - \bar{s})^2},\end{aligned}$$

$$\hat{\sigma}_\varepsilon^2 = \frac{1}{n}\left(\sum(D_i - \bar{D})^2 - \frac{[\sum(s_i - \bar{s})D_i]^2}{\sum(s_i - \bar{s})^2} \right)$$

The estimates $\hat{\alpha}$ and $\hat{\beta}_1$ are independent and have distributions $\mathrm{N}(\alpha, n^{-1}\sigma_\varepsilon^2)$ and $\mathrm{N}(\beta_1, [\sum(s_i - \bar{s})^2]^{-1}\sigma_\varepsilon^2)$.

The random variables

$$\begin{aligned} Q_0 &= n(\bar{D}-\alpha)^2/\sigma_\varepsilon^2 \ \sim\ \chi^2_{1\mathrm{df}}, \\ Q_1 &= (\hat{\beta}_1-\beta_1)^2\sum(s_i-\bar{s})^2/\sigma_\varepsilon^2 \ \sim\ \chi^2_{1\mathrm{df}} \\ Q_2 &= n\hat{\sigma}_\varepsilon^2/\sigma_\varepsilon^2 \sim \chi^2_{(n-2)\mathrm{df}}, \end{aligned}$$

are independent and

$$Q = \sum[D_i-\alpha-\beta(s_i-\bar{s})]^2 \sim \chi^2_{n\mathrm{df}}.$$

Writing $\tilde{\sigma}_{\varepsilon}^2 = n\hat{\sigma}_{\varepsilon}^2/(n-2)$, there are, consequently, two independent F ratio tests from the full model:

$$F_0^* = \frac{n\bar{D}^2}{\tilde{\sigma}_{\varepsilon}^2} \sim F_{(1,n-2)\mathrm{df}},$$

testing $\mathrm{H}\colon\ \alpha = 0$ in favour of a reduced model with no intercept parameter; and

$$F_1^* = \frac{\hat{\beta}_1^2 \sum_{i=1}^{n} (s_i - \bar{s})^2}{\tilde{\sigma}_{\varepsilon}^2} \sim F_{(1,n-2)\mathrm{df}},$$

testing $\mathrm{H}\colon\ \beta_1 = 0$ in favour of a model with no slope parameter.

Since $(T_{k\mathrm{df}})^2 \equiv F_{(1,k)\mathrm{df}}$ these F ratio tests can be reexpressed as two Student t-tests, each with $(n-2)$ degrees of freedom:

$$T_0^* = \frac{\bar{D}}{\tilde{\sigma}_\varepsilon/\sqrt{n}} \sim t_{(n-2)\mathrm{df}} \tag{5}$$

and

$$T_1^* = \frac{\hat{\beta}_1}{\tilde{\sigma}_\varepsilon/\sqrt{\sum(s_i-\bar{s})^2}} \sim t_{(n-2)\mathrm{df}}. \tag{6}$$

As Q_0 and Q_1 are independent it also follows that

$$\frac{1}{2}(Q_0 + Q_1) \sim \chi^2_{2\mathrm{df}},$$

and a third F-ratio testing $\mathrm{H}\colon \alpha = \beta_1 = 0$ in favour of a model with no intercept parameter and no slope parameter can be formed as

$$F^* = \frac{(T_0^*)^2 + (T_1^*)^2}{2} \sim F_{(2,n-2)\mathrm{df}}, \tag{7}$$

and is an alternative representation of the Bradley-Blackwood test statistic given in (4) above.

Outline

- The power function for the test of the marginal hypothesis H: $\mu_D = 0$ using Student's paired t-test is characterized by a non-central F-distribution with degrees of freedom parameters 1 and $(n-1)$ in the numerator and denominator, and non-centrality parameter $\delta = n(\mu_D)^2/\sigma_D^2$.
- The power function for the test of the marginal hypothesis H: $\mu_D = 0$ using the t-ratio T_0^* is a non-central F-distribution with degrees of freedom parameters 1 and $(n-2)$ in the numerator and denominator, and non-centrality parameter $\delta_0 = n(\mu_D)^2/\sigma_{D|S}^2$, where $\sigma_{D|S}^2 = \sigma_D^2 - (\sigma_1^2 - \sigma_2^2)^2/\sigma_S^2 = \sigma_D^2(1-\rho_{DS}^2)$.
- These results follow as a special cases of power functions for tests of the general linear hypothesis in univariate linear models (?).

- The power function for the test of the marginal hypothesis H: $\mu_D = 0$ using Student's paired t-test is characterized by a non-central F-distribution with degrees of freedom parameters 1 and $(n-1)$ in the numerator and denominator, and non-centrality parameter $\delta = n(\mu_D)^2/\sigma_D^2$.
- The power function for the test of the marginal hypothesis H: $\mu_D = 0$ using the t-ratio T_0^* is a non-central F-distribution with degrees of freedom parameters 1 and $(n-2)$ in the numerator and denominator, and non-centrality parameter $\delta_0 = n(\mu_D)^2/\sigma_{D|S}^2$, where $\sigma_{D|S}^2 = \sigma_D^2 - (\sigma_1^2 - \sigma_2^2)^2/\sigma_S^2 = \sigma_D^2(1-\rho_{DS}^2)$.
- These results follow as a special cases of power functions for tests of the general linear hypothesis in univariate linear models (?).

- The power function for the test of the marginal hypothesis H: $\mu_D = 0$ using Student's paired t-test is characterized by a non-central F-distribution with degrees of freedom parameters 1 and $(n-1)$ in the numerator and denominator, and non-centrality parameter $\delta = n(\mu_D)^2/\sigma_D^2$.
- The power function for the test of the marginal hypothesis H: $\mu_D = 0$ using the t-ratio T_0^* is a non-central F-distribution with degrees of freedom parameters 1 and $(n-2)$ in the numerator and denominator, and non-centrality parameter $\delta_0 = n(\mu_D)^2/\sigma_{D|S}^2$, where $\sigma_{D|S}^2 = \sigma_D^2 - (\sigma_1^2 - \sigma_2^2)^2/\sigma_S^2 = \sigma_D^2(1-\rho_{DS}^2)$.
- These results follow as a special cases of power functions for tests of the general linear hypothesis in univariate linear models (?).

Outline

- A closed test procedure based on the maximum and minimum t-ratios.
- Hsu's paper "On samples from a normal bivariate population" discusses 7 tests of hypotheses relevant to this discussion.
- The Student t-test is UMP.
- Question the extent to which the results depend on Normality?
- Model selection and Bayes factors.
- kevin.hayes@ul.ie