Concordant Measure E and a Test Statistic

$$E = \frac{1}{2}\left\{\frac{\sum p_{1j}s_{2j}}{\sum p_{1j}} + \frac{\sum p_{2j}s_{1j}}{\sum p_{2j}}\right\}$$

Population parameters and estimates

We postulate that (p_{1j}, s_{1j}), for $j = 1, ...n$ is a random sample from a "population 1" with mean (μ_{p_1}, μ_{s_1}), variances $\sigma^2_{p_1}$, $\sigma^2_{s_1}$, and covariance γ_1. Similarly, we assume (p_{2j}, s_{2j}) are from "population 2" with mean (μ_{p_2}, μ_{s_2}), variances $\sigma^2_{p_2}$, $\sigma^2_{s_2}$, and covariance γ_2.

We can estimate the parameters of the populations using the sample data as follows. For $i = 1, 2$

$$\begin{aligned} \hat{\mu}_{p_i} = \bar{p}_i = \sum_j p_{ij}/n, \qquad & \hat{\mu}_{s_i} = \textstyle\sum_j s_{ij}/n, \\ \hat{\sigma}^2_{p_i} = \sum_j (p_{ij} - \bar{p}_i)^2/(n-1), \quad & \hat{\sigma}^2_{s_i} = \textstyle\sum_j (s_{ij} - \bar{s}_i)^2/(n-1) \qquad (1) \\ \gamma_i = \sum_j (p_{ij} - \bar{p}_i)(s_{ij} - \bar{s}_i)/(n-1) & \end{aligned}$$

Null Hypothesis of No Concordance and the Null Distribution of E

We want (estimates for) the mean and variance of E under the null hypothesis of *no concordance* between the two samples.

We assume that the null hypothesis of *no concordance* is equivalent to the condition that (p_{1i}, s_{1i})'s are independent of (p_{2i}, s_{2i})'s.

Under the null hypothesis, E is approximately normally distributed, i.e.,

$$E \sim N(\mu_E, \sigma^2_E/n) \qquad (2)$$

where μ_E, σ^2_E are given in terms of the parameters $\mu_{p_1}, \mu_{s_1}, \sigma^2_{p_1}, \sigma^2_{s_1}$, $\mu_{p_2}, \mu_{s_2}, \sigma^2_{p_2}, \sigma^2_{s_2}$, γ_1 and

γ_2, as follows:

$$\begin{aligned} \mu_E &= (\mu_{s_1} + \mu_{s_2})/2 \\ \sigma_E^2 &= \boldsymbol{\delta}' \Sigma \boldsymbol{\delta}, \end{aligned} \tag{3}$$

where

$$\boldsymbol{\delta} = 0.5 \cdot (1/\mu_{p_1}, -\mu_{s_2}/\mu_{p_1}, 1/\mu_{p_2}, -\mu_{s_1}/\mu_{p_2})', \tag{4}$$

and Σ is a symmetric matrix whose elements (on and above the diagonal) are given below

$$\Sigma = \begin{pmatrix} \sigma_{p_1}^2 \sigma_{s_2}^2 + \sigma_{p_1}^2 \mu_{s_2}^2 + \sigma_{s_2}^2 \mu_{p_1}^2 & \mu_{s_2} \sigma_{p_1}^2 & \gamma_1 \gamma_2 + \gamma_1 \mu_{p_2} \mu_{s_2} + \gamma_2 \mu_{p_1} \mu_{s_1} & \gamma_2 \mu_{p_1} \\ & \sigma_{p_1}^2 & \gamma_1 \mu_{p_2} & 0 \\ & & \sigma_{p_2}^2 \sigma_{s_1}^2 + \sigma_{p_2}^2 \mu_{s_1}^2 + \sigma_{s_1}^2 \mu_{p_2}^2 & \mu_{s_1} \sigma_{p_2}^2 \\ & & & \sigma_{p_2}^2 \end{pmatrix}$$

Test statistic

Estimates of μ_E, σ_E^2, namely, $\hat{\mu}_E$, $\hat{\sigma}_E^2$, can be obtained by substituting the estimates in (1) for the parameters in the expression for μ_E, σ_E^2, given in (3).

By (2), a test statistic for testing the null hypothesis of no concordance is a z-test statistic, given by

$$z = \sqrt{n} \left(\frac{E - \hat{\mu}_E}{\hat{\sigma}_E} . \right)$$

Details of the Derivation

$$
\begin{aligned}
E &= \frac{1}{2}\left\{\frac{\sum p_{1j}s_{2j}}{\sum p_{1j}} + \frac{\sum p_{2j}s_{1j}}{\sum p_{2j}}\right\} \\
&= \frac{1}{2}\left\{\frac{\sum p_{1j}s_{2j}/n}{\sum p_{1j}/n} + \frac{\sum p_{2j}s_{1j}/n}{\sum p_{2j}/n}\right\} \\
&= \frac{1}{2}\left\{\bar{X}_1/\bar{X}_2 + \bar{X}_3/\bar{X}_4\right\},
\end{aligned}
$$

where $X_{1j} = p_{1j}s_{2j}$, $X_{2i} = p_{1j}$, $X_{3j} = p_{2j}s_{1j}$, and $X_{4j} = p_{2j}$.

In the following we assume the null hypothesis is true.

By CLT,

$$\sqrt{n}(\bar{\boldsymbol{X}} - \boldsymbol{\mu}) \overset{approx.}{\sim} N(0, \Sigma)$$

where $\boldsymbol{X} = (X_1, X_2, X_3, X_4)'$, and $\boldsymbol{\mu}$ and $\Sigma = (\sigma_{ij})$ are the mean and variance of $\boldsymbol{X}$, given by

$$\boldsymbol{\mu} = (\mu_{p_1}\mu_{s_2}, \mu_{p_1}, \mu_{p_2}\mu_{s_1}, \mu_{p_2})$$

and Σ is a symmetric matrix given by

$$\Sigma = \begin{pmatrix} \sigma_{p_1}^2\sigma_{s_2}^2 + \sigma_{p_1}^2\mu_{s_2}^2 + \sigma_{s_2}^2\mu_{p_1}^2 & \mu_{s_2}\sigma_{p_1}^2 & \gamma_1\gamma_2 + \gamma_1\mu_{p_2}\mu_{s_2} + \gamma_2\mu_{p_1}\mu_{s_1} & \gamma_2\mu_{p_1} \\ & \sigma_{p_1}^2 & \gamma_1\mu_{p_2} & 0 \\ & & \sigma_{p_2}^2\sigma_{s_1}^2 + \sigma_{p_2}^2\mu_{s_1}^2 + \sigma_{s_1}^2\mu_{p_2}^2 & \mu_{s_1}\sigma_{p_2}^2 \\ & & & \sigma_{p_2}^2 \end{pmatrix}$$

For convenience, we will use the notation $\boldsymbol{\mu} = (\mu_1, \mu_2, \mu_3, \mu_4)'$ and $\Sigma = (\sigma_{ij})$ to refer to the components of $\boldsymbol{\mu}$ and Σ.

Now,

$$E = g(\bar{\boldsymbol{X}})$$

where $g(\boldsymbol{x}) = g(x_1, x_2, x_3, x_4) = (x_1/x_2 + x_3/x_4)/2$.

By Delta method,

$$E \overset{approx.}{\sim} N(\mu_E, \sigma_E^2/n)$$

where

$$\begin{aligned}\mu_E &= g(\boldsymbol{\mu}) = (\mu_{s_1} + \mu_{s_2})/2) \\ \sigma_E^2 &= \boldsymbol{\delta}'\Sigma\boldsymbol{\delta},\end{aligned}$$

and

$$\begin{aligned}\boldsymbol{\delta} &= \frac{\partial g(\boldsymbol{x})}{\partial \boldsymbol{x}} \; evaluated\ at\ \boldsymbol{x} = \boldsymbol{\mu} \\ &= 0.5 \cdot (1/\mu_{p_1}, -\mu_{s_2}/\mu_{p_1}, 1/\mu_{p_2}, -\mu_{s_1}/\mu_{p_2})'\end{aligned}$$

Expression $\boldsymbol{\mu}$ and Σ in terms of the population parameters

$\mu_1 = E(X_1) = E(p_1 s_2) = E(p_1)E(s_2) = \mu_{p_1}\mu_{s_2}$

$\mu_2 = \mu_{p_1}, \;\; \mu_3 = E(p_2 s_1) = \mu_{p_2}\mu_{s_1} \;\;\; \mu_4 = \mu_{p_2}$

$\sigma_{11} = var(X_1) = var(p_1 s_2)$=$E(p_1^2 s_2^2)$-$E(p_1)^2 E(s_2)^2$

$= E(p_1^2)E(s_2^2) - E(p_1^2)E(s_2)^2 + E(p_1^2)E(s_2)^2 - E(p_1)^2 E(s_2)^2$

$=E(p_1^2)V(s_2) + V(p_1)E(s_2^2)$

$=V(p_1)[V(s_2) + E(s_2)^2] + V(s_2)E(p_1)^2$

$=\sigma_{p_1}^2\sigma_{s_2}^2 + \sigma_{p_1}^2\mu_{s_2}^2 + \sigma_{s_2}^2\mu_{p_1}^2$

$\sigma_{12} = Cov(X_1, X_2) = cov(p_1 s_2, p_1) = \mu_{s_2}\sigma_{p_1}^2$

$\sigma_{13} = Cov(p_1 s_2, p_2 s_1) = \gamma_1\gamma_2 + \gamma_1\mu_{p_2}\mu_{s_2} + \gamma_2\mu_{p_1}\mu_{s_1}$

$\sigma_{14} = Cov(p_1 s_2, p_2) = \gamma_2\mu_{p_1}$

$\sigma_{22} = \sigma_{p_1}^2$

$\sigma_{23} = \gamma_1\mu_{p_2}$

$\sigma_{24} = Cov(p_1, p_2) = 0$

$\sigma_{33} = var(X_3) = var(p_2 s_1) = \sigma_{p_2}^2\sigma_{s_1}^2 + \sigma_{p_2}^2\mu_{s_1}^2 + \sigma_{s_1}^2\mu_{p_2}^2$

$\sigma_{34} = Cov(p_2 s_1, p_2) = \mu_{s_1}\sigma_{p_2}^2$

$\sigma_{44} = \sigma_{p_2}^2$