SUPPLEMENT TO "ASSESSMENT OF MOVEMENT DIFFICULTIES IN CHILEAN PRESCHOOLERS"

LISSETH BARRA AND FELIPE OSORIO

Supplementary Material

Appendix A introduces some notation about the multinomial distribution. Next, we present the Wald statistic to test linear hypothesis considering G independent groups. Confidence intervals for the proportions of the multinomial distribution are described in Appendix B. A computational implementation of such techniques in R language (R Core Team, 2018) is presented and is made publicly available on github.

Appendix A. Wald statistic to test linear hypothesis for the multinomial distribution

We say that a k-dimensional random vector $\boldsymbol{X}$ has a multinomial distribution if its joint density function is given by

$$f(\boldsymbol{x}; n, \boldsymbol{p}) = \frac{n!}{x_1! \cdots x_k!} \prod_{i=1}^{k} p_i^{x_i}, \tag{A.1}$$

where $\boldsymbol{x} \in \{0, 1, \dots, n\}^k$ and $\sum_{i=1}^{k} x_i = n$. In addition, it is assumed that $p_i > 0$ and $\sum_{i=1}^{k} p_i = 1$. If a random vector has a density function (A.1), we shall denote $\boldsymbol{X} \sim \mathsf{Multinomial}_k(n, \boldsymbol{p})$, with $\boldsymbol{p} = (p_1, \dots, p_k)^\top$. It is straightforward to note that

$$\mathrm{E}(\boldsymbol{X}) = n\boldsymbol{p}, \qquad \mathrm{Cov}(\boldsymbol{X}) = n(\mathrm{diag}(\boldsymbol{p}) - \boldsymbol{p}\boldsymbol{p}^\top),$$

where $\mathrm{diag}(\boldsymbol{p}) = \mathrm{diag}(p_1, \dots, p_k)$ denotes a diagonal matrix.

Consider $\boldsymbol{X}_1, \dots, \boldsymbol{X}_G$, G independent random vectors such that,

$$\boldsymbol{X}_r \sim \mathsf{Multinomial}(n_r, \boldsymbol{p}_r), \qquad r = 1, \dots, G. \tag{A.2}$$

Thus, the log-likelihood function associated with the parameter of interest $\boldsymbol{\theta} = (\boldsymbol{\theta}_1^\top, \dots, \boldsymbol{\theta}_G^\top)^\top$, with $\boldsymbol{\theta}_r = (p_{r1}, \dots, p_{r,k-1})^\top$, asumes the form:

$$\ell(\boldsymbol{\theta}) = \sum_{r=1}^{G} \log n_r! - \sum_{r=1}^{G} \sum_{i=1}^{k-1} \log x_{ri}! + \sum_{r=1}^{G} \sum_{i=1}^{k-1} x_{ri} \log p_{ri} + \sum_{r=1}^{G} x_{rk} \log\Big(1 - \sum_{i=1}^{k-1} p_{ri}\Big).$$

Following Davis and Jones (1992), we have that the maximum likelihood (ML) estimator for $\boldsymbol{\theta}$ is given by

$$\widehat{p}_{ri} = \frac{x_{ri}}{n_r}, \qquad r = 1, \dots, G; i = 1, \dots, k.$$

The Fisher information matrix for $\boldsymbol{\theta}$ based on the model defined in Equation (A.2) assumes a block diagonal form

$$\boldsymbol{\mathcal{F}}(\boldsymbol{\theta}) = \begin{pmatrix} n_1\boldsymbol{\mathcal{F}}_1(\boldsymbol{\theta}_1) & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & n_2\boldsymbol{\mathcal{F}}_2(\boldsymbol{\theta}_2) & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & n_G\boldsymbol{\mathcal{F}}_G(\boldsymbol{\theta}_G) \end{pmatrix},$$

where the Fisher information matrix for the rth group is given by

$$\boldsymbol{\mathcal{F}}_r(\boldsymbol{\theta}_r) = \operatorname{diag}^{-1}(\boldsymbol{\theta}_r) + \frac{1}{p_{rk}}\mathbf{1}\mathbf{1}^\top,$$

with $\mathbf{1} = (1,\dots,1)^\top$ being a $(k-1)\times 1$ vector of ones. Thus, $\widehat{\boldsymbol{\theta}}_1,\dots,\widehat{\boldsymbol{\theta}}_G$ are asymptotical independent with covariance matrix

$$\operatorname{Cov}(\widehat{\boldsymbol{\theta}}_r) = \frac{1}{n_r}\boldsymbol{\mathcal{F}}_r^{-1}(\boldsymbol{\theta}) = \big(\operatorname{diag}(\boldsymbol{\theta}_r) - \boldsymbol{\theta}_r\boldsymbol{\theta}_r^\top\big), \tag{A.3}$$

for $r = 1,\dots,G$. Moreover, the standard error of the ML estimator of p_{ri} is given by

$$\operatorname{SE}(\widehat{p}_{ri}) = \sqrt{\widehat{\operatorname{var}}(\widehat{p}_{ri})} = \sqrt{\widehat{p}_{ri}(1-\widehat{p}_{ri})/n_r}, \qquad r = 1,\dots,G; i = 1,\dots,k-1,$$

which allows the construction of asymptotic confidence intervals (see Appendix B).

A.1. **Wald statistic to test linear hypothesis.** Relying on results for large samples, we can test hypotheses regarding to the vector of proportions $\boldsymbol{\theta}$. Specifically, we will be interested in linear hypotheses such as

$$H_0 : \boldsymbol{H}\boldsymbol{\theta} = \boldsymbol{h} \qquad \text{against} \qquad H_1 : \boldsymbol{H}\boldsymbol{\theta} \neq \boldsymbol{h}, \tag{A.4}$$

where $\boldsymbol{H} \in \mathbb{R}^{q\times p}$ with $q = \operatorname{rank}(\boldsymbol{H})$ and $\boldsymbol{h} \in \mathbb{R}^q$. The Wald statistic to test the hypothesis stated in Equation (A.4) takes the form

$$W = (\boldsymbol{H}\widehat{\boldsymbol{\theta}} - \boldsymbol{h})^\top\{\boldsymbol{H}\boldsymbol{\mathcal{F}}^{-1}(\widehat{\boldsymbol{\theta}})\boldsymbol{H}^\top\}^{-1}(\boldsymbol{H}\widehat{\boldsymbol{\theta}} - \boldsymbol{h}),$$

which has an asymptotic chi-square distribution with q degrees of freedom.

Remark 1 (**Equality of proportions**). For the application addressed in the manuscript we have $p = G(k-1)$ with $G = 2$ groups and $k = 3$. Thus, we are interested in testing the hypothesis $H_0 : \boldsymbol{\theta}_1 = \boldsymbol{\theta}_2$ against $H_1 : \boldsymbol{\theta}_1 \neq \boldsymbol{\theta}_2$, which can be written in the form of (A.4) with $\boldsymbol{H} = (\boldsymbol{I}, -\boldsymbol{I})$. Thus, it is straightforward to note that

$$W = (\widehat{\boldsymbol{\theta}}_1 - \widehat{\boldsymbol{\theta}}_2)^\top\Big\{\frac{1}{n_1}\boldsymbol{\mathcal{F}}_1^{-1}(\widehat{\boldsymbol{\theta}}_1) + \frac{1}{n_2}\boldsymbol{\mathcal{F}}_2^{-1}(\widehat{\boldsymbol{\theta}}_2)\Big\}^{-1}(\widehat{\boldsymbol{\theta}}_1 - \widehat{\boldsymbol{\theta}}_2),$$

or equivalently as

$$W = (\widehat{\boldsymbol{\theta}}_1 - \widehat{\boldsymbol{\theta}}_2)^\top\big\{\operatorname{Cov}(\widehat{\boldsymbol{\theta}}_1) + \operatorname{Cov}(\widehat{\boldsymbol{\theta}}_2)\big\}^{-1}(\widehat{\boldsymbol{\theta}}_1 - \widehat{\boldsymbol{\theta}}_2), \tag{A.5}$$

where $\operatorname{Cov}(\widehat{\boldsymbol{\theta}}_r)$, $r = 1,2$, was presented in Equation (A.3).

Remark 2 (**Linear combinations**). To test hypotheses such as $H_0 : \boldsymbol{a}^\top\boldsymbol{\theta} = a_0$, where $\boldsymbol{a} \in \mathbb{R}^p$ and a_0 is a predefined scalar, consider the following statistic

$$Z = \frac{\boldsymbol{a}^\top\widehat{\boldsymbol{\theta}} - a_0}{\sqrt{\boldsymbol{a}^\top\boldsymbol{\mathcal{F}}^{-1}(\widehat{\boldsymbol{\theta}})\boldsymbol{a}}},$$

which is asymptotically distributed as a standard normal random variable. Thus, we reject $H_0 : \boldsymbol{a}^\top \boldsymbol{\theta} = a_0$ in favour of $H_1 : \boldsymbol{a}^\top \boldsymbol{\theta} \neq a_0$, if

$$|Z| \geq z_{1-\alpha/2}.$$

Moreover, we reject the hypothesis $H_0 : \boldsymbol{a}^\top \boldsymbol{\theta} \geq a_0$ for $H_1 : \boldsymbol{a}^\top \boldsymbol{\theta} < a_0$, if

$$Z \leq z_\alpha,$$

where $z_\alpha = \Phi^{-1}(\alpha)$ denotes a quantile of the standard normal distribution.

The following R code can be used to test the hypothesis $H_0 : \boldsymbol{\theta}_1 = \boldsymbol{\theta}_2$ against $H_1 : \boldsymbol{\theta}_1 \neq \boldsymbol{\theta}_2$ using the Wald statistic given in (A.5):[1]

```
equal.prop <- function(theta1, n1, theta2, n2) {
  # Wald statistic for equality of proportions
  k <- length(theta1)
  if (length(theta2) != k)
    stop("theta1 and theta2 must have the same length")
  #
  diff  <- theta1 - theta2
  aCov1 <- diag(theta1) - outer(theta1, theta1)
  aCov2 <- diag(theta2) - outer(theta2, theta2)
  aCov  <- aCov1 / n1 + aCov2 / n2
  s <- solve(aCov, diff)
  Wald <- sum(s * diff)
  #
  pval <- 1 - pchisq(Wald, df = k)

  ## output object
  z <- list(Wald = Wald, df = k, p.value = pval)
  z
}
```

Appendix B. Confidence intervals for multinomial proportions

Consider that the group is fixed. Thus, next we shall to remove the dependency of the index r. An asymptotic confidence interval for multinomial proportions assumes the form:

$$C_n^1(p_i) = \Big(\widehat{p}_i \mp z_{1-\alpha/2}\sqrt{\frac{\widehat{p}_i(1-\widehat{p}_i)}{n}}\Big), \qquad i = 1, \dots, k,$$

Alternatively, Quesenberry and Hurst (1964) and Bailey (1980), respectively, proposed the following confidence intervals based on large samples

$$C_n^2(p_i) = \Big(\frac{q_1 + 2x_i \mp \sqrt{q_1(q_1 + 4n_i(n - x_i)/n)}}{2(n + q_1)}\Big), \qquad i = 1, \dots, k,$$

and

$$C_n^3(p_i) = \Big(\frac{Y \mp \sqrt{C(C + 1 - Y^2)}}{C + 1}\Big)^2, \qquad i = 1, \dots, k,$$

[1] R codes are available at: https://github.com/faosorios/Multinomial-proportions

where $q_1 = \chi^2_{1-\alpha}(k-1)$ denotes a quantile of the chi-square distrbution with $k-1$ degrees of freedom, whereas

$$Y = \Big(\frac{x_i + 3/8}{n + 1/8}\Big)^{1/2}, \qquad C = \frac{\chi^2_{1-\alpha/k}(1)}{4n},$$

with $\chi^2_{1-\alpha/k}(1)$ denoting a quantile of the chi-square distribution with 1 degree of freedom.

Next, is presented the R code to obtain confidence intervals reported in the manuscript[1]

```
CI.prop <- function(p, n, level = 0.95, method = "asymp") {
  # confidence intervals for multinomial proportions
  k <- length(p)
  a <- (1 - level) / 2
  a <- c(a, 1 - a)
  qz <- qnorm(a)
  qc <- qchisq(level, df = k - 1)
  a <- 1 - level
  qb <- qchisq(1 - a / k, df = 1) # Bonferroni correction
  switch(method,
         "asymp" = { # asymptotic interval (default)
           SE <- sqrt(p * (1 - p) / n)
           ci <- p + SE %o% qz
         },
         "QH" = { # Quesenberg and Hurst (1964)
           m <- n * p
           lh <- qc + 2 * m
           SE <- sqrt(qc + 4 * m * (n - m) / n)
           ci <- lh + SE %o% qz
           ci <- ci / (2 * n + qc)
         },
         "sqrt" = { # square root transformation (Bailey, 1980)
           m <- n * p
           y <- sqrt((m + 0.375) / (n + 0.125))
           b <- 0.25 * qb / n
           SE <- 0.5 * sqrt((b + 1 - y^2) / n)
           qb <- c(-sqrt(qb), sqrt(qb))
           ci <- y + SE %o% qb
           ci <- (ci / (b + 1))^2
         })
  ci <- cbind(p, SE, ci)
  colnames(ci) <- c("prop", "SE", "lower", "upper")
  ci
}
```

References

Bailey, B.J.R. (1980). Large sample simultaneous confidence intervals for the multinomial probabilities based on trasnformations of the cell frequencies. *Technometrics* **22**, 583-589.

Davis, C.S., and Jones, M.P. (1992). Maximum likelihood estimation for the multinomial distribution. *Teaching Statistics* **14**, 9-11.

Quesenberry, C.P., and Hurst, D.C. (1964). Large sample simultaneous confidence intervals for multinomial proportions. *Technometrics* **22**, 583-589.

R Core Team (2018). *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing. Vienna, Austria. URL: www.R-project.org.

Departamento de Kinesiología, Universidad de Chile

Current address: Departamento de Kinesiología, Universidad de Chile, Av. Independencia 1027, Santiago, Chile

E-mail address: lisseth@uchile.cl

Departamento de Matemática, Universidad Técnica Federico Santa María, Chile

Current address: Departamento de Matemática, Universidad Técnica Federico Santa María, Av. España 1680, Casilla 110-V, Valparaíso, Chile

E-mail address: felipe.osorios@usm.cl