Gait & Posture 39 (2014) 518–522

Contents lists available at ScienceDirect

Gait & Posture

journal homepage: www.elsevier.com/locate/gaitpost

Ellipse area calculations and their applicability in posturography

Patric Schubert [1,*], Marietta Kirchner [1]

Department of Health, Fresenius University of Applied Sciences, Limburger Straße 2, 65510 Idstein, Germany

ARTICLE INFO

Article history:
Received 18 July 2012
Received in revised form 18 June 2013
Accepted 3 September 2013

Keywords:
Posturography
Postural control
COP area calculation
Prediction ellipse
Confidence ellipse

ABSTRACT

The quantification of postural sway is considered to be an essential part of posturography and is important for research and clinical utility. A widely used method to calculate the scatter of center of pressure data is an ellipse that encloses about $100(1-\alpha)\%$ of the observations. However, underlying definitions and terminologies have been misused in many cases. Hence, outcomes of different studies are proved to be incommensurable. In order to attain inter-study comparability, standardization of calculation methods has to be advanced. This work features a comprehensive and consistent overview of the methods for elliptic area approximation contrasting general principles of confidence and prediction regions. As a result, we recommend the usage of the prediction ellipse, as far as we demonstrate that confidence ellipses emerge to be inappropriate for posturographic scatter evaluation. Furthermore, we point at problems that come along with different sample sizes.

1. Introduction

If a subject is asked to stand as still as possible on a force plate, small oscillations of the center of pressure (COP) can be observed [1]. Quantification of these fluctuations is considered to be an essential part of posturography and is important for research and clinical utility [2]. Among a great variety of parameters the computation of sway area is a traditional and widely used method. A geometrically simple figure is an ellipse that encloses about $100(1-\alpha)\%$ of the observations in the scatter plot [3]. However, as an ellipse including this characteristic cannot be univocally constructed, terminologies have to be formulated more precisely. However, with respect to literature, definitions and terminologies have been misused in many cases. For instance, Prieto et al. [4] and Prieto and Myklebust [5] use the term ‘confidence ellipse’; however, underlying calculations are conceptually referred to ‘prediction ellipses’ which was spuriously adopted by several authors [6–10]. Both terms are the bivariate analogs of the univariate views on confidence and prediction intervals with different prerequisites and underlying formulas. Moreover, some confusion prevails concerning calculation methods. As in the past computations implemented regression analysis, more recent investigations are based on the theories of principal component analysis (PCA) in order to gain uniqueness of the ellipse. The latter method is proposed by Oliveira et al. [11] which to their point of view “does not appear to have been considered for this purpose before”. First, the introduced procedure can be reduced to an Eigen value problem. Second, similar procedures have been published by others yet [5,12].

The quite simple formulas and procedures were unnecessarily complicated by many authors and it can be shown that they could be merged into each other. In addition, some formal mistakes regarding published equations [4,7,13] are revealed. Inter-study comparisons in posturography require adequate standardization of the calculation methods. This work features a comprehensive and consistent overview of the methods for elliptic area approximation contrasting the principles of confidence and prediction ellipses.

2. Basic considerations

For this purpose, we have to differentiate between the construction of confidence and prediction regions. Therefore, we present basic assumptions from a statistical point of view. Subsequently, we refer specifically to area calculation of confidence and prediction ellipses.

2.1. Confidence and prediction intervals

First, we would like to reduce the distinction to the univariate case [14]. In practice, the population mean μ and its standard deviation σ are not known. One estimates the true population mean μ from the observed data and asks how good the estimation is. When $\bar{x}$ is the sample mean, one would like to define a region around $\bar{x}$ which covers with $100(1-\alpha)\%$ of probability the true value μ. Let $\bar{X} = 1/n\sum_{i=1}^{n} X_i$ be the estimation function. According

* Corresponding author. Tel.: +49 6126 935 2293.
E-mail addresses: rough@ironsport.de, Patric.schubert@hs-fresenius.de, p.schubert@sport.uni-frankfurt.de (P. Schubert).
[1] Both authors contributed equally to this paper.

0966-6362/$ – see front matter
http://dx.doi.org/10.1016/j.gaitpost.2013.09.001

to the central limit theorem $\bar{X}$ follows a normal distribution. That is, $\bar{X} \sim N(\mu, \sigma^2/n)$ which results in $(\bar{X}-\mu) \sim N(0, \sigma^2/n)$. Denote a random variable $Z_C = \bar{X} - \mu$ with mean 0 and standard deviation $\sigma/\sqrt{n}$. Then we can define

$$P\left(-z_{(1-\alpha/2)} \le \frac{Z_C}{\sigma/\sqrt{n}} \le z_{(1-\alpha/2)}\right) = 1-\alpha, \quad (1)$$

with $z_{(1-\alpha/2)}$ being the $(1-\alpha/2)$-quantile of the standard normal distribution. This can be transformed into

$$P\left(\bar{X} - z_{(1-\alpha/2)}\frac{\sigma}{\sqrt{n}} \le \mu \le \bar{X} + z_{(1-\alpha/2)}\frac{\sigma}{\sqrt{n}}\right) = 1-\alpha. \quad (2)$$

With respect to a realization based on a sample of size n where the unknown standard deviation σ is estimated by $s = \sqrt{(1/(n-1))\sum_{i=1}^{n}(x_i - \bar{x})^2}$ the confidence interval is given by [14]

$$\left[\bar{x} - t_{(1-\alpha/2),n-1}\frac{s}{\sqrt{n}}; \bar{x} + t_{(1-\alpha/2),n-1}\frac{s}{\sqrt{n}}\right], \quad (3)$$

with $t_{(1-\alpha/2),n-1}$ being the $(1-\alpha/2)$-quantil of the t-distribution with $(n-1)$ degrees of freedom. The t-distribution converges against the normal distribution for large n. The confidence interval determines a region which covers with 100(1 $-\alpha$)% of probability the unknown population mean.

In contrast, prediction intervals estimate an unknown future observation from the statistic of the observed sample. The task is to create limits to the sample mean $\bar{x}$ so that with $100(1-\alpha)$% of probability the future observation will fall into the prediction interval. Let Z be a new single observation. We define a random variable $Z_P = Z - \bar{X}$ with mean 0 and standard deviation $\sigma \cdot \sqrt{1+(1/n)}$. As before (Eq. (1)) we have

$$P\left(-z_{(1-\alpha/2)} \le \frac{Z_P}{\sigma \cdot \sqrt{1+(1/n)}} \le z_{(1-\alpha/2)}\right) = 1-\alpha, \quad (4)$$

which gives a prediction interval [14]

$$\left[\bar{x} - t_{(1-\alpha/2),n-1}s\sqrt{1+\frac{1}{n}}; \bar{x} + t_{(1-\alpha/2),n-1}s\sqrt{1+\frac{1}{n}}\right]. \quad (5)$$

Confidence intervals will shrink as the sample size n increases because $\lim_{n\to\infty} 1/\sqrt{n} = 0$. Prediction intervals, on the other hand, will have a limit of $(z_{(1-\alpha/2)} \cdot s)$ as it is $\lim_{n\to\infty} t_{(1-\alpha/2),n-1} = z_{(1-\alpha/2)}$ and $\lim_{n\to\infty}\sqrt{1+(1/n)} = 1$. In opposition to confidence intervals, they have to cope with both the uncertainty of the true population mean and the scatter of the sample, which means that the region is on average broader.

2.2. Confidence and prediction ellipses

The theories on confidence and prediction regions can be extended to the multivariate case. In posturography a bivariate view complies with this purpose (see Appendix section c). For large n, the bivariate confidence region is a confidence ellipse, which can be seen as the contour lines of the bell-shaped mound of the bivariate normal population distribution. Two parameters are required (we denote vectors and matrices in bold letters): the population mean vector $\mu = \begin{pmatrix}\mu_x \\ \mu_y\end{pmatrix}$ for which the estimate function is $\bar{X} = \frac{1}{n}\begin{pmatrix}\sum_{i=1}^{n} X_i \\ \sum_{i=1}^{n} Y_i\end{pmatrix}$, and population covariance matrix $\Sigma = \begin{pmatrix}\sigma_x^2 & \sigma_{x,y} \\ \sigma_{x,y} & \sigma_y^2\end{pmatrix}$, which is estimated by the unbiased sample covariance matrix $S = \begin{pmatrix}s_x^2 & s_{x,y} \\ s_{x,y} & s_y^2\end{pmatrix}$. The random variable $Z_C = \bar{X} - \mu$ then has mean 0 and covariance $(1/n)\,\Sigma$ (estimated by $(1/n)\,S$). We can write (see Appendix section a):

$$P\left(n \cdot {Z_C}^T S^{-1} Z_C \le T^2_{(1-\alpha),2,n-2}\right) = 1-\alpha \quad (6)$$

where $(\cdot)^T$ is the transpose and T^2 is the Hotelling T-squared distribution with 2 and $(n-2)$ degrees of freedom which forms the multivariate generalization of the t-distribution [15]. It can be shown that $T^2_{2,n-2} = (2(n-1)/(n-2))\,F_{2,n-2}$, where $F_{2,n-2}$ is the Fisher-distribution with 2 and $(n-2)$ degrees of freedom [16]. A similar notation to Eq. (6) can be found in Sokal and Rohlf [3]. We expect the confidence ellipse covers with a $100(1-\alpha)$% of probability the true population mean μ.

With respect to posturography, we want to describe the ellipse itself. The directions x and y are now referred to the medial–lateral and the anterior–posterior component of the COP motion with sample mean $\bar{x} = \begin{pmatrix}\bar{x} \\ \bar{y}\end{pmatrix}$. After the division by n we get

$$(\bar{x}-\mu)^T S^{-1}(\bar{x}-\mu) = \frac{2(n-1)}{n(n-2)}F_{(1-\alpha),2,n-2}. \quad (7)$$

The left side can be also merged into

$$\frac{s_x^2 s_y^2}{s_x^2 s_y^2 - s_{xy}^2}\left[\frac{(\bar{x}-\mu_x)^2}{s_x^2} + \frac{(\bar{y}-\mu_y)^2}{s_y^2} - \frac{2s_{xy}(\bar{x}-\mu_x)(\bar{y}-\mu_y)}{s_x^2 s_y^2}\right] = \frac{2(n-1)}{n(n-2)}F_{(1-\alpha),2,n-2}. \quad (8)$$

This equation can now be solved to get a region around μ_x and μ_y with a given probability of $100(1-\alpha)$% as this is preset by the F-value.

An equal form can be looked-up in Rocchi et al. [17]. However, the authors mix the theories of confidence and prediction ellipses as they talk of confidence but calculate prediction ellipses. The right side of Eq. (8) demonstrates a coefficient which limits the size of the confidence ellipse. With larger n its size will shrink: $\lim_{n\to\infty} 2(n-1)/n(n-2) = 0$ (for $n > 101$ the coefficient is less than 0.01). If the experimenter wants to describe the scatter or covariation of the data, a prediction ellipse will be more appropriate. Consider **Z** as a vector of a new observation and $\boldsymbol{Z_P} = \boldsymbol{Z} - \bar{\boldsymbol{X}}$ a random variable with mean 0 and covariance $(1+1/n)\boldsymbol{\Sigma}$. Then, the probability of the prediction region can be denoted as

$$P\left(\left(1+\frac{1}{n}\right)^{-1} \cdot {Z_P}^T S^{-1} Z_P \le T^2_{(1-\alpha),2,n-2}\right) = 1-\alpha. \quad (9)$$

Multiplying the factor $(n+1)/n$ the equation for the prediction ellipse is then defined by [18]:

$$(z-\bar{x})^T S^{-1}(z-\bar{x}) = \frac{2(n+1)(n-1)}{n(n-2)}F_{(1-\alpha),2,n-2}. \quad (10)$$

For large n it is $\lim_{n\to\infty} 2(n+1)(n-1)/n(n-2)F_{(1-\alpha),2,n-2} = 2F_{(1-\alpha),2,n-2}$ (for $n > 202$ the coefficient is less than 1.01). The right side approaches the Chi square distribution with 2 degrees of freedom: $\lim_{n\to\infty} 2 \cdot F_{(1-\alpha),2,n-2} \sim \chi^2_{(1-\alpha),2}$. The bivariate prediction region is an ellipse that covers with $100(1-\alpha)$% of probability the future observation **Z**. A $100(1-\alpha)$% prediction region for the next observation is identical to a $100(1-\alpha)$% tolerance region of type 2, due to the following definition: A bivariate tolerance region is an ellipse such that the expected value of the proportion of the population contained in the ellipse is exactly $100(1-\alpha)$% [18].

The constructions of confidence and prediction regions differ in their assumptions. Both regions are located around the sample