



SANTIAGO L. IPIÑA, ANA I. DURAND

SEXUAL DIMORPHISM IN MIDDLE PLEISTOCENE HUMANS

ABSTRACT: Sexual dimorphism in a Middle Pleistocene human population (Atapuerca, Sima de los Huesos, Spain) as well as in a modern human population (Coimbra, Portugal) are analyzed and compared through a measure proposed elsewhere (see Ipiña, Durand 2000) by the authors of this work. The point estimate for the sexual dimorphism in the Atapuerca population has turned out to be twice as large as the one corresponding to the modern human population here analyzed. Although such results should be considered cautiously, given that the sample size of the Atapuerca remains is small, they show disagreement with the hypothesis according to which sexual dimorphism is comparable in the above mentioned populations (see Arsuaga et al. 1997).

KEYWORDS: Sexual Dimorphism – MI Index – Middle Pleistocene humans

Arsuaga *et al.* (1997) have suggested that sexual dimorphism in the coxal (*vertical acetabular diameter*) is comparable in Middle Pleistocene (Sima de los Huesos, Sierra de Atapuerca, Spain sample) and modern human (Museum of Anthropology, University of Coimbra, Portugal sample) populations. By using a form of the so-called bootstrap method, they compared "intrapopulation variations" in the Sima de los Huesos (SH) and modern humans in order to achieve such a conclusion. The aim of this work is to show that, in the first place, the procedure used by these authors does not seem to be the appropriate one, and secondly, when applied, a procedure that has to do with sexual dimorphism, results suggest to be cautious.

There is a great deal of interest in trying to establish the degree of dimorphism between the two sexes that compose a specific population (see, e.g. Borgognini Tarli, Repetto 1986, Lovich, Gibbons 1992, Plavcan 1994, Rehg, Leigh 1999 for reviews inside the anthropological community). Most measures, coefficients or indices which assess sexual dimorphism have been constructed on the basis that the morphological trait or random variable in which we are interested, is distributed according to a specific probability law.

When random variables and probability distributions are involved, it seems reasonably clear that a study of sexual dimorphism is equivalent to a study whose main purpose is to assess differences between the distribution functions that correspond to the two sexes involved. The best way indeed to define the distribution function of a random variable is to express its analytic function, so that if the aim is to explore the degree of confusion – or overlap – between females and males, then the analysis of sexual dimorphism will be equivalent to the assessment of the existent overlapping area between surfaces delimited by the analytic functions of both sexes.

A problem arises when one asks what are the analytic functions of sexes within the population in which they coexist. In other words, if we suppose that the *femur length of the Spaniards* (X) is normally distributed among Spanish women and that this variable is likewise normally distributed among Spanish men, then the population formed by these women and men is distributed according to a probability law, denominated a mixture model with two normal components (see, e.g. McLachlan, Basford 1988), whose probability density function is defined by:

$$f(x) = \pi_1 f_1(x) + \pi_2 f_2(x), \quad -\infty < x < \infty$$

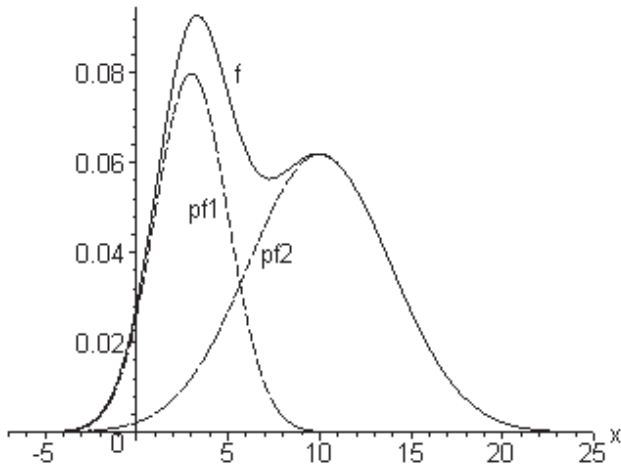


FIGURE 1. A mixture of two normal components, where $f = f_1 + f_2$, $pf1 = \pi_1 f_1$ and $pf2 = \pi_2 f_2$.

where f_i is the normal density function of the i sex and π_i is the mixing proportion in the population of the i sex, $i = 1, 2$ (female, male) (see an example in *Figure 1*). There is not clearly any normal distribution within a normal mixture since $\pi_i f_i$ is not a normal density. On the other hand, it is evident that this type of mixture depends on five parameters: two mathematical expectations (μ_i), two variances (σ_i^2) and one proportion (note that $\pi_2 = 1 - \pi_1$). Finally, when a random sample – with individuals belonging to both sexes – is extracted from a population distributed according to a mixture model with two normal components, it is clear that such a sample is not extracted from any normal population so that one can legitimately ask why then should we carry out tests or apply indices which are based upon the normal distribution.

The authors of these comments have proposed an index of sexual dimorphism that takes into account the real population composed by both sexes, an index – which we call *MI*, whose range is $(0, 0.5]$ and, as a FORTRAN program, is available on request – that is based on a mixture model with two normal components (see Ipiña, Durand 2000). In short, our index is defined by the overlap area between $\pi_1 f_1$ and $\pi_2 f_2$ because these two functions represent the contribution of each sex to the mixture; hence, our proposal is a function of the five parameters μ_i , σ_i^2 and π_i , $i = 1, 2$.

In most practical studies the five parameters that characterize a two-component normal mixture are unknown, so that an estimation procedure ought to be used. The Expectation Maximization (EM) algorithm (see McLachlan, Basford 1989; Hathaway 1985 for a constrained version of the algorithm) and the Markov Chain Monte Carlo (MCMC) Bayesian procedure (see Gilks *et al.* 1996) are the best alternatives in our opinion. The interested reader can see in Ipiña and Durand (2000) which alternative to choose as per different sampling situations.

Since the sample size of Atapuerca remains was six and they were not sexed, the MCMC alternative was used in

order to obtain the $\hat{\pi}_i, \hat{\sigma}_i^2, \hat{\mu}_i$, $i = 1, 2$ estimates (BUGS 0.5 program, MRC Biostatistics Unit, Institute of Public Health, Cambridge, UK, 1995). The size of the Atapuerca sample is so small indeed that estimates of especially female and male variances might not be reliable. As a consequence, we let these variances in the resultant MCMC Atapuerca normal mixtures – the reader is referred to Gilks *et al.* 1996, in order to understand how different normal mixtures can be obtained by simply modifying the corresponding parameters of the variance distribution – took different values and then selected from among these pair of values the one that yielded the greatest likelihood function.

In other words, suppose we have a sample with three observables (x_1, x_2, x_3) and that, e.g., $\Theta_a = (\mu_{a1}, \sigma_{a1}^2, \pi_{a1})$ and $\Theta_b = (\mu_{b1}, \sigma_{b1}^2, \pi_{b1})$, $i = 1, 2$, denote the vectors of parameters of two normal mixtures. In such a case, the likelihood function f_a of that sample for the first mixture is:

$$f_a(x_1, x_2, x_3; \Theta_a) = \prod_{j=1}^3 \left[\sum_{i=1}^2 \frac{\pi_{ai} \exp\left(-\frac{(x_j - \mu_{ai})^2}{2\sigma_{ai}^2}\right)}{\sigma_{ai} \sqrt{2\pi}} \right],$$

and similarly, f_b is:

$$f_b(x_1, x_2, x_3; \Theta_b) = \prod_{j=1}^3 \left[\sum_{i=1}^2 \frac{\pi_{bi} \exp\left(-\frac{(x_j - \mu_{bi})^2}{2\sigma_{bi}^2}\right)}{\sigma_{bi} \sqrt{2\pi}} \right].$$

In the case that $f_b > f_a$, then we select Θ_b , that is, the mixture with means μ_{bi} variances σ_{bi}^2 and proportions π_{bi} , $i = 1, 2$.

The *MI* sexual dimorphism indices obtained from the *vertical acetabular diameter* (data provided by J. L. Arsuaga) were compared in the same two human samples analyzed by Arsuaga *et al.* (1997): the first one from a human Middle Pleistocene biological population (Sima de los Huesos in Sierra de Atapuerca, Spain) and the second one corresponding to a population of modern humans (Museum of Anthropology from the University of Coimbra, Portugal).

The results obtained can be seen in *Table 1*, where $\hat{\pi}_f, \hat{\mu}_f$ and $\hat{\sigma}_f^2$ stand for the estimated values of female proportions, means and variances, respectively, the m subscript stands for male, $\hat{\sigma}^2$ is the estimate of the mixture variance and \widehat{MI} is the corresponding estimate of the *MI* index of sexual dimorphism.

In the light of these results, it is observed, in the first place, that the σ^2 normal mixture variances of the Atapuerca SH and Coimbra populations are comparable. It can be easily shown that "intrapopulation variation" is

TABLE 1. Results obtained for the estimations of the mixture parameters and MI index in the Atapuerca and Coimbra populations here analyzed.

	Coimbra population	Atapuerca SH population
$\hat{\pi}_f$	0.47	0.27
$\hat{\pi}_m$	0.53	0.73
$\hat{\mu}_f$	49.879	52.16
$\hat{\mu}_m$	55.2	58.24
$\hat{\sigma}_f^2$	7.4398	2.8833
$\hat{\sigma}_m^2$	7.6446	7.6892
$\hat{\sigma}^2$	14.5833	13.6776
\widehat{MI}	0.1659	0.0835

proportional to σ^2 so, the fact that Arsuaga *et al.* (1997), on the basis of their "intrapopulation variation" analysis solely, concluded that sexual dimorphism is comparable in the Atapuerca SH and Coimbra populations is, in this manner, easily understood.

In the second place, the hypothesis according to which sexual dimorphism in the coxal (*vertical acetabular diameter*) is comparable in Middle Pleistocene and modern humans populations seems to be in need of some further analysis. The fact is that the resultant sexual dimorphism in the Sima de los Huesos, Atapuerca, is instead twice as large as the one in the modern human population here analyzed so, in spite of the fact that some more refined inferential procedure would be necessary in order to obtain more definite conclusions, this finding throws some doubts on a sexual dimorphism that is comparable in the above mentioned populations.

We would like to point out that the estimates of the parameters that allow to compute an estimate of the MI index of sexual dimorphism, have been obtained as a consequence of having used the probability distribution of a vector of parameters conditional on the data currently encountered. This is the same as saying that our results are closely dependent upon the information supplied by such data so that we have not, on purpose, taken into account whether or not the Atapuerca sample here analyzed can be considered as a good representative of the Atapuerca SH population.

Finally, we do not think that analyzing "intrapopulation variation" alone is the appropriate procedure to assess sexual dimorphism. In effect, in a mixture scenario, the variance σ^2 of a two-component normal mixture – remember that this parameter is proportional to "intrapopulation variation" – is defined by:

$$\sigma^2 = \pi_1 \sigma_1^2 + (1 - \pi_1) \sigma_2^2 + \pi_1 (1 - \pi_1) (\mu_1 - \mu_2)^2 .$$

Let us suppose then that σ^2 takes a specific value σ_0^2 . There is clearly a variety of values for π_1 , σ_i^2 , μ_i , $i = 1, 2$, verifying that $\sigma^2 = \sigma_0^2$. If we measure sexual dimorphism through an index that depends on these five parameters,

which seems to be the more suitable alternative, it is easily seen that different values for such an index can be obtained for each σ_0^2 .

Needless to say that the same conclusion is obtained in the case that sexual dimorphism is measured through an index depending upon some, not necessarily all, of the five parameters, as for example the quotient of the two mean parameters – of course, fixing the remaining three parameters. Therefore, it does not seem appropriate to analyze solely σ^2 , or a quantity like the "intrapopulation variation", when our purpose is to analyze sexual dimorphism. Further, from the fact that different values of a sexual dimorphism measure can generate the same value of the σ^2 mixture variance, it is easily seen that "sexual dimorphism is *potentially* a major source of size variation in a population" as Arsuaga *et al.* (1997) argue.

REFERENCES

- ARSUAGA J. L., CARRETERO J. M., LORENZO C., GRACIA A., MARTINEZ L., BERMUDEZ DE CASTRO J. M., CARBONELL E., 1997: Size variation in Middle Pleistocene humans. *Science* 277: 1086–1088.
- BORGOGNINI TARLI S. M., REPETTO E., 1986: Methodological considerations on the study of sexual dimorphism in past human populations. *Hum. Evol.* 1: 51–66.
- GILKS W. R., RICHARDSON S., SPIEGELHALTER D. J., 1996: *Markov Chain Monte Carlo in Practice*. Chapman and Hall, London. Pp. 486.
- HATHAWAY R. J., 1985: A constrained EM algorithm for univariate normal mixtures. *J. Statist. Comput. Simul* 23: 211–230.
- IPÍÑA S. L., DURANDA A. I., 2000: A measure of sexual dimorphism in populations which are univariate normal mixtures. *Bull. Math. Biol.* 62: 925–941.
- LOVICH J. E., GIBBONS J. W., 1992: A review of techniques for quantifying sexual size dimorphism. *Growth Dev. Aging* 56: 269–281.
- McLACHLAN G. J., BASFORD K. E., 1988: *Mixture Models. Inference and Applications to Clustering*. Marcel Dekker Inc., New York. Pp. 253.
- PLAVCAN J. M., 1994: Comparison of four simple methods for estimating sexual dimorphism in fossils. *Amer. J. Phys. Anthropol.* 94: 465–476.
- REHG J. A., LEIGH S. R., 1999: Estimating sexual dimorphism and size differences in the fossil record: A test of methods. *Amer. J. Phys. Anthropol.* 110: 95–104.

Santiago L. Ipiña
 Matemática Aplicada
 Facultad Biología
 Universidad Complutense Madrid
 28040 Madrid
 E-mail: ipina@eucmax.sim.ucm.es
 Tel.: +34 91 394 5069
 Fax: +34 91 394 5051