

Stability of post mortem regression models of maximum stature over combined samples of bulgarians and hungarians

D. Radoinova¹, N. Nikolova, Y. Yordanov², K. Tenekedjiev

IT Department Naval Academy Varna 9026, Bulgaria, e-mail: nalianik@gmail.com

¹Department of Forensic Medicine Medical University – Varna Varna 9002, Bulgaria,
e-mail: dradoinova@mail.bg

²IEMA Bulgarian Academy of Science, Sofia, Bulgaria, e-mail: iemabas@bas.bg

A typical forensic problem is the *post mortem* prediction of maximum stature of individuals using the length of limb long bone remains. Here we use linear regression models of stature age-related corrected samples of Bulgarians and Hungarians on four regressors (length of *humerus*, *fibula*, *tibia*, and of *humerus* and *tibia* combined). The precision of predictions strongly depends on the amount of available measurements, which provides incentives to combine samples. This study identifies the stable sample combination which can generate regression equations with higher quality of prediction. Combinations on sex, nationality and on both are tested for stability by three methods with seven modifications each. *K*-means clustering analysis partitioned the cases into three groups of absolute stability, stability and instability. This proved it is only reasonable to unite samples on sex. *K*-means and hierarchical clustering analysis confirmed the initial partitioning, and helped identify the characteristics of the typical stability cases.

Keywords: *post-mortem* stature, age correction, limb long bones, linear regression, combining samples, clustering

Introduction

The construction of statistically plausible regression models requires large number of measurements. That guarantees the adequacy and correctness of the predictions made on the basis of that model. For that reason it is a common practice to combine samples in order to provide larger amount of data. A question arises of whether two samples can be combined in the first place, because if not then the constructed prediction model will be misleading. That is why the stability of a constructed model over the combined sample is of importance. A problem in forensic medicine is the prediction of living stature (as one of the main and most stable features of physical development) by the *post mortem* length of limb long bones. The work [15] presents age corrected regressions of maximum stature for Bulgarians with regressors – the length of *humerus*, *tibia* and *fibula*. Similar procedures are performed in [16] for Hungarians.

Modern research focuses on creation of models that apply to different populations: nations [17] and races [7]. This study utilizes the samples of lengths of *humerus* (H), *tibia* (T) and *fibula* (F) of Bulgarians and Hungarians (males and females) from [15, 16] with the intention to identify the sample combinations, where a stable regression model can be constructed and used for prediction of maximum stature. On top, sometimes bone remains are of unknown *post mortem* origin (in terms of sex and nationality). The study is conducted in accordance with the procedure for two-sample regression analysis [18].

Materials and Methods

The total number of measured grown individuals is 684, with age varying from 19 to 66 years. The sample includes 415 Bulgarians (285 males and 130 females) and 269 Hungarians (186 males and 83 females). After the age of 45-50 years the decline of maximal human stature begins, so age correction is applied, adding the decline to the measured stature according to one of the seven known methods as explained in [15, 16]. In this study, predictions are made using four regressors: *humerus* (H), *fibula* (F), *tibia* (T), and *humerus* and *tibia* ($H+T$) combined.

Regression procedures for two samples, presented in [18] are employed to test the possibility to unite the samples and build new equations to predict maximal stature of individuals. This problem is known as testing the coefficients' stability of the three regression equations – two for the separate and one for the combined sample. The separate samples are first subjected to detection and rejection of outliers in two loops by a series of t -tests for predicted residuals. Then the combined sample is checked for heteroskedasticity using four F-tests (Ramsey [19], White [20], Glejser [5], and Goldfeld-Quadt [4]), and with one χ^2 -test (Breusch and Pagan [3]) in eleven modifications (linear, square, root and reciprocal). If data is proven to be heteroskedastic then four models of the residual modules are built and tested for adequacy using an ANOVA test [1]. The chosen model is the one with the maximally adjusted coefficient of multiple determination (\bar{R}^2) that exceeds a preliminarily defined critical level (\bar{R}_{\min}^2). If no such model exists, heteroskedasticity is negligible even if statistically significant. Then three statistical tests for stability are performed: ANOVA test for coefficient equality of the two models [2], predicted Chow test based on the larger sample (Chow-1) and predicted Chow test, based on the smaller sample (Chow-2) [9]. Result interpretation of these F-tests is not simple because all three tests may be viewed as a problem for selection of regressors using dummy variables. In addition to the classical comparison with critical value F_{crit} corresponding to preliminary chosen significance level, six other types of answers of every test are given according to different critical values (from F_{crit}^2 to F_{crit}^7), corresponding to the following criteria: maximum of \bar{R}^2 , three criteria minimizing the mean squared error of prediction (C_p of Mallows [6], S_p of Hocking [10] and PC of Amemiya [8]), information criteria of Akaike [11] and Bayes posterior relation of Leamer [13].

Six pairs of samples for each of the four regressors are tested for stability: 1) Bulgarian males and Hungarian males; 2) Bulgarian females and Hungarian females; 3) Bulgarian males and Bulgarian females; 4) Hungarian males and Hungarian females; 5) Bulgarians and Hungarians; 6) males and females. In case of proven stability, a regression model is constructed, and the confidence intervals of the coefficients, the covariance matrix, the characteristics of the standard error, and the coefficients of multiple determination (R^2 and \bar{R}^2) are calculated. The confidence intervals of the predicted maximal stature are calculated with respect to the standard error in the point of prediction and the Mahalanobis distance between the last and the middle point of the sample in the space of regressors [12].

Results

I. Stability profiles

The following settings are implied in the analysis: 1) age correction according to (Borcan et al. 1983, Giles 1991); 2) significance level for rejecting outliers beyond 0.5% with a maximum of 2 loops for detecting one level of measurement errors and one level of non-typicality; 3) significance level for checking heteroskedasticity of 5%; 4) significance level for the ANOVA test with all models of 5%; 5) minimal adjusted coefficient of multiple determination $\bar{R}_{\min}^2 = 15\%$ for accepting a model eliminating the heteroskedasticity; a model explaining less than 15% of the observed variance of the residuals' module is practically insignificant which means the heteroskedasticity is negligible; 6) significance level of 5% for the t-test for the regression coefficients and of the ANOVA test checking the model adequacy; 7) confidence level of the standard error range of 95%; 8) confidence levels of the stature range and of the regression coefficients regions respectively of 95%, 99%, and 99.9%; 9) Significance level of 5% for the classical stability F-tests.

There are 24 paired samples to be analyzed, indexed as follows:

- 1 to 4 – regressions of “Bulgarian males – Hungarian males” on $H, F, T,$ and $H+T$;
- 5 to 8 – regressions of “Bulgarian females – Hungarian females” on $H, F, T,$ and $H+T$;
- 9 to 12 – regressions of “Bulgarian males – Bulgarian females” on $H, F, T,$ and $H+T$;
- 13 to 16 – regressions of “Hungarian males – Hungarian females” on $H, F, T,$ and $H+T$;
- 17 to 20 – regressions of “males – females” on $H, F, T,$ and $H+T$;
- 21 to 24 – regressions of “Bulgarians – Hungarians” on $H, F, T,$ and $H+T$.

The stability is tested by 21 F-tests (ANOVA, Chow-1 and Chow-2 tests in seven modifications). Then 24 number of 21-dimensional binary vectors, indicating the results from the stability tests in each case (with 0 for stability, and 1 for instability) are formed. Cluster analysis is applied over the binary vectors with the intention to allocate the cases into groups with similar stability profile. The K -means clustering method ^[20] with city block distance measure is used to form compact and well-separated exclusive clusters of vectors in the 21-dimensional space. The method identifies three groups, the first containing 4 cases (6, 7, 8, and 22), the second containing 8 cases (from 1 to 5, 21, 23 and 24), and the third containing 12 cases form a third group (from 9 to 20). According to the observed characteristics, these clusters can be respectively referred to as absolute stability group, stability group, and instability group. The results prove that the combinations “Bulgarian males – Bulgarian females” and “Hungarian males – Hungarian females” are instable for all the regressors. Because of that the regressions 21 to 24 for Bulgarians and Hungarians are practically useless even if proven stable, as long as their parts are unstable themselves. That is why clustering analysis is repeated over the remaining 20 cases. The K -means method again identified three groups, keeping the same distribution of the remaining cases in the three groups as before. Hierarchical clustering analysis with city block distance as dissimilarity measure and with unweighted average distance as a linkage factor between the clusters is also performed [14]. The method identifies the same three groups as the K -means method (see the dendrogram on Fig. 1), with the only exception that the sixth paired sample has much better stability than paired samples 7 and 8. The latter two are combined with the stability group (from 1 to 5) earlier than with the former. However, the stability profile of cases 7 and 8 is good enough to be considered absolutely stable. Having in mind the stability profile of the 20 cases and the confirmed results from the clustering analysis, it is now possible to identify the profile of the typical cases of stability that can be observed:

A. Absolute stability – the following is present: 1) the ANOVA test shows stability at least in its classical version and on the Leamer criterion; 2) the Chow test over the larger sample shows stability on all criteria; 3) the Chow test over the smaller sample shows stability at least on the Amemiya and Leamer criteria. Here, the data in the samples may be combined and the predictions will be much more precise than those made on the basis of the separate samples.

B. Stability – the following is present: 1) the ANOVA test shows instability on all criteria; 2) the Chow test over the larger sample shows stability on all criteria; 3) the Chow test over the smaller sample shows stability only on the Leamer criterion, the Amemiya criterion can give either result, and all other criteria show instability. The data in the samples may be combined and the predictions will be much more precise than those made on the separate samples, yet less precise than on the combined samples with absolute stability.

C. Instability – the following is present: 1) the ANOVA test shows instability on all criteria; 2) the Chow test over the larger sample shows stability only on the Leamer criterion; 3) the Chow test over the smaller sample shows stability only on the Amemiya and the Leamer criteria. The two samples should not be combined and no regressions should be constructed since the prediction error will be much higher than that on the separate samples.

II. Regression results for samples which can be unified

The absolute stable and the stable regressions (whose samples can be unified) are from 1 to 8. The results from the ANOVA, Chow-1 and Chow-2 for them are shown on Table 1. For those cases the regression equations of maximum stature, the lower (σ_{ϵ}^l) and upper (σ_{ϵ}^u) bounds of the estimated standard error, the adjusted coefficients of multiple determination ($\bar{R}^2=0$), the number of outliers from the Bulgarian (O_B) and

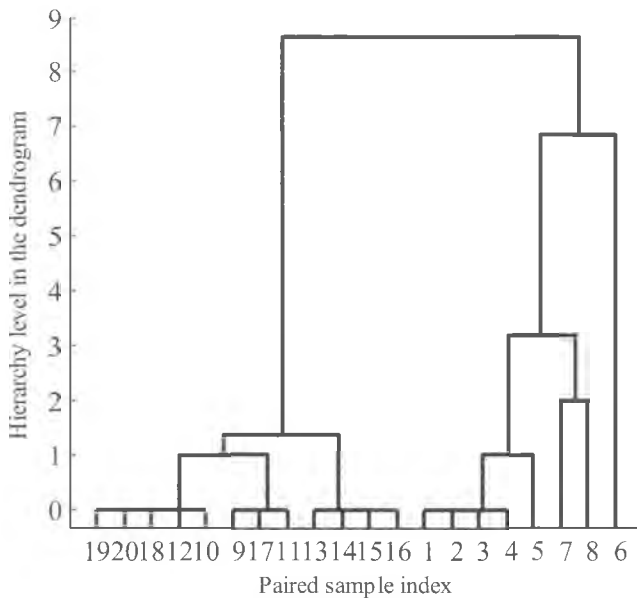


Fig. 1. Dendrogram from the hierarchical analysis of 20 stability profiles represented as 21-dimensional vectors of stability results

from the Hungarian (O_H) samples, and the stability profile (SP: A- absolute stable, B- stable) are given in Table 2. All regression coefficients are significant (all t-tests with $p_{value} < 0.0005$) and the models are adequate (ANOVA with $p_{value} < 0.0005$).

Table 1. F_{value} of the combined samples of Bulgarian and Hungarian males or females on the length of humerus (H), fibula (F), tibia (T) and on humerus and tibia ($H+T$) along with the critical values of the ANOVA, Chow-1 and Chow-2 stability tests on seven criteria. The critical values that are exceeded by F_{value} (i.e. the test that show instability) are bolded.

Bone(s) Sex	Test	F_{value}	F^1_{crit}	F^2_{crit}	F^3_{crit}	F^4_{crit}	F^5_{crit}	F^6_{crit}	F^7_{crit}	Stability Profile
H male	ANOVA	8.8	5.4	1	2	2.0	2.0	0.99	6.2	Stability
	Chow-1	0.52	1.4	1	2	2.0	2.7	0.74	15	
	Chow-2	2.3	1.4	1	2	2.0	3.6	0.54	26	
F male	ANOVA	8.0	5.4	1	2	2.0	2.0	0.99	6.2	Stability
	Chow-1	0.40	1.4	1	2	2.0	2.7	0.74	15	
	Chow-2	3.1	1.4	1	2	2.0	3.6	0.53	26	
T male	ANOVA	10	5.4	1	2	2.0	2.0	0.99	6.2	Stability
	Chow-1	0.40	1.4	1	2	2.0	2.7	0.75	16	
	Chow-2	3.3	1.4	1	2	2.0	3.6	0.543	26	
$H+T$ male	ANOVA	13	4.3	1	2	2.0	2.0	0.99	6.2	Stability
	Chow-1	0.61	1.4	1	2	2.0	2.7	0.73	16	
	Chow-2	2.4	1.4	1	2	2.0	3.6	0.53	26	
H female	ANOVA	18	5.4	1	2	2.0	2.0	0.99	5.4	Stability
	Chow-1	0.63	1.7	1	2	2.0	2.7	0.74	11	
	Chow-2	3.8	1.7	1	2	2.0	3.7	0.52	16	
F female	ANOVA	0.52	5.4	1	2	2.0	2.0	0.99	5.4	Absolute stability
	Chow-1	0.47	1.7	1	2	2.0	2.6	0.74	11	
	Chow-2	2.1	1.7	1	2	2.0	3.7	0.51	16	
T female	ANOVA	2.1	5.4	1	2	2.0	2.0	0.99	5.4	Absolute stability
	Chow-1	0.53	1.7	1	2	2.0	2.6	0.74	11	
	Chow-2	2.0	1.7	1	2	2.0	3.7	0.52	16	
$H+T$ female	ANOVA	3.1	4.4	1	2	2.0	2.0	0.98	5.4	Absolute stability
	Chow-1	0.68	1.7	1	2	2.0	2.7	0.73	11	
	Chow-2	1.7	1.7	1	2	2.0	3.7	0.51	16	

Discussion

This study showed that the combination of samples is only reasonable if done according to sex. Combining measurements of individuals of different sex shows no stability. This fact proves the sexual dimorphism and follows from the different body proportions of males and females. Since the combination of males and females shows instability, then the combination of samples of Bulgarians and Hungarians is of no practical use even if proven stable.

The results from the stability tests allow outlining the following cases in analyzing bone remains and the procedures in each.

Table 2. Regression equations for maximum stature (S) in (cm) on the combined samples of Bulgarian and Hungarian males (m) or females (f) on the length of *humerus* (H), *fibula* (F), *tibia* (T) and on *humerus* and *tibia* ($H+T$) in (cm). The standard errors of the regression coefficients are written in parentheses.

Regressors	Sex	Regression Equation	σ_c^d	σ_c^u	O_B/O_H	S P	R ²
H	m	$S=86.77(1.12)+2.530(0.033) \times H+\varepsilon$	1.51	1.72	6/4	B	0.926
F	m	$S=83.47(1.045)+2.412(0.029) \times F+\varepsilon$	1.39	1.59	7/7	B	0.940
T	m	$S=89.88(1.075)+2.240(0.029) \times T+\varepsilon$	1.53	1.75	8/5	B	0.927
$H+T$	m	$S=86.12(0.98)+1.502(0.091) \times H++0.963(0.080) \times T+\varepsilon$	1.32	1.51	7/4	B	0.944
H	f	$S=92.15(1.85)+2.183(0.057) \times H+\varepsilon$	1.63	1.98	1/3	B	0.875
F	f	$S=86.39(1.25)+2.191(0.036) \times F+\varepsilon$	1.03	1.25	2/4	A	0.947
T	f	$S=92.49(1.52)+2.013(0.044) \times T+\varepsilon$	1.33	1.62	2/3	A	0.911
$H+T$	f	$S=88.93(1.39)+0.873(0.10) \times H++1.307(0.094) \times T+\varepsilon$	1.21	1.47	1/2	A	0.931

A. Analysis of bones of males for all the regressors (H , F , T or $H+T$): 1) if it is not known whether the individual is a Bulgarian or Hungarian, then use the regressions over the combined sample of males; 2) if the individual is known to be a Bulgarian, then use the regressions over the sample of Bulgarian males; 3) if the individual is known to be a Hungarian, then use the regressions over the sample of Hungarian males.

B. Analysis of bones of females: 1) if the regressors are F , T or $H+T$ then use the regressions over the combined sample of females; 2) if the regressor is H then: a) if it is not known whether the individual is a Bulgarian or Hungarian, use the regression over the combined sample of females; b) if the individual is known to be a Bulgarian, use the regression over the sample of Bulgarian females for H ; c) if the individual is known to be a Hungarian, use the regression over the sample of Hungarian females for H .

C. Analysis of bones of Bulgarians or Hungarians for all the regressors (H , F , T or $H+T$): 1) if it is not known whether the individual is a male or a female make two conditional predictions for males and for females using I and II; 2) if the individual is known to be a male, then follow A; 3) if the individual is known to be a female, then follow B.

References

1. Applied Linear Statistical Models. NY, Irwin/McGraw-Hill, 5th ed., 2004, 119-127.
2. Applied Regression Analysis and Multivariable Methods. 4th ed., NY, Duxbury Press, 2008, 558-568.
3. Breusch, T. S., A. R. Pagan. *Econometrica*. – 47, 1979, 1287-1294.
4. *Econometric Analysis*, 6th ed., NJ, Pearson Education Inc., 2008, 1178.
5. Furno, M. – *The Indian J. Stats.* 67 /2, 2005, 335-358.
6. Gilmour, S. G. – *Statistician*. 45, 1996, 49-56.
7. Holland, T. D. – *Am. J. Phys. Anthropol.* 96, 1995, 315-320.
8. *Introduction to Statistics and Econometrics*. Boston, Harvard University Press, 1994, 299-309.
9. *Introduction to Econometrics*. Oxford, Oxford University Press, 2007, 347-348.
10. Lavergne, P. – *Economet Rev.* 17, 1998, 227-273.
11. *Model Selection and Inference: A Practical Information-Theoretic Approach.*, NY. Springer, 2002, 60-64.
12. *Numerical Recipes – The Art of Scientific Computing*, Cambridge, Cambridge University Press, 1992, 671-680.
13. Pagan, A. In: *Modelling Economic Series: Advanced Texts in Econometrics*, (ed. C. W. J. Granger), Oxford, Oxford University Press, 1991, 97-120.
14. *Pattern Classification*, NY, Wiley Interscience Publication, 2001, 9-14; 526-528; 550-556.
15. Radoinova, D., K. Tenekedjiev, Y. Yordanov. – *HOMO*, **52**, 2002, 221-232.
16. Radoinova, D., K. Tenekedjiev, Y. Yordanov. – *Annales Historico-Naturales Musei Nationalis Hungarici*, **94**, 2002, 237-251.
17. Sjoqvold, T. – *Hum. Evol.*, 5, 1990, 431-447.
18. Tenekedjiev, K., D. Radoinova. – *Acta Morphologica et Anthropologica*, **6**, 2001, 90-97.
19. *Using Econometrics – A Practical Guide*. Addison-Wesley-Longman, 2001, 193-195.
20. White, H. A. *Econometrica*, 48, 1980, 817-838.