Listy Biometryczne
Vol. XXIII (1986) No. 2
Biometrical Letters

THE „REGRESSION TO THE MEAN" EFFECT AS OBSERVED FOR SOME REPEATED BLOOD PRESSURE MEASUREMENTS

ANNA BARTKOWIAK, RYSZARD RUTA

Institute of Computer Science, Wrocław University
Lower Silesia Center for Medical Diagnosis, DOLMED Wrocław

Praca wpłynęła 22 kwietnia 1986; w wersji ostatecznej 6 maja 1987

Bartkowiak A., Ruta R., 1987. The „regression to the mean" effect as observed for some repeated blood pressure measurements (Efekt „regresji do wartości średniej" na podstawie powtarzanych pomiarów ciśnienia krwi). Listy Biometryczne XXIII, z. 2. Wydawnictwo Naukowe Uniwersytetu im. Adama Mickiewicza w Poznaniu (Adam Mickiewicz University Press), pp. 75-84, 3 figs., 1 table, ISBN 83-232-0091-2, ISSN 0458-0036.

Porównujemy dwie metody zapisywania ciśnienia krwi: X_1 - automatyczną (urządzenie Avionics 1900) i X_2 - tradycyjną (za pomocą sfigmanometru). Próbujemy ustalić, czy $Y=X_2-X_1$, różnica między tymi dwoma metodami, zależy od linii odniesienia pomiaru X_1. Nachylenie linii zstępnej $Y=a+bX_1$ jest negatywne. Po dostosowaniu różnicy Y do wyniku" regresji do wartości średniej" nachylenie równa się 0.

We compare two methods of blood pressure recording: X_1-automatic (Avionics 1900 device) and X_2-traditional (by sfigmanometer). We investigate whether $Y = X_2-X_1$, the difference between these two methods, depends on the baseline measurement X_1. The slope of the regression line $Y = a+bX_1$ is negative. After adjusting the difference Y for the „regression to the mean" effect the slope becomes 0.

1. THE BIOMEDICAL PROBLEM

Traditionally blood pressure (BP) is recorded by a mercurial sfigmanometer. In DOLMED blood pressure is recorded automatically by the Avionics 1900 device connected directly to the computers core. The problem arose: are these two recordings comparable? Formally the two methods measure the same biological phenomenon, none the less practically they might give values shifted by some constant value. There are also some technical reasons by which the difference between these two methods might depend on the magnitude of the true value of the blood pressure. To clarify these doubts and to make the measurements recorded in DOLMED comparable to others recorded traditionally, an experiment was conducted. For $n_1 = 119$ men and $n_2 = 117$ women the blood pressure (systolic and diastolic) was measured in a randomized sequence 6 times: by the two methods mentionned above

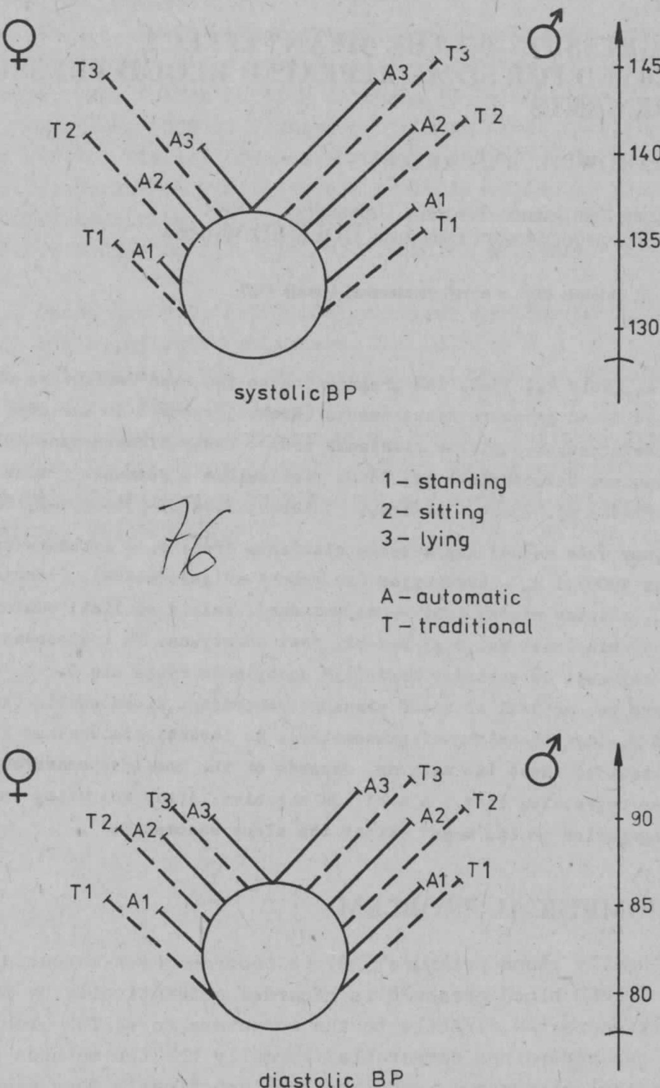

Fig. 1. Graphical presentation of mean values by glyphs

(automatic and traditional) and in 3 positions (standing, sitting, lying). The data collected in this way were subjected to statistical analysis with the aim to establish statistical dependencies between the measurements recorded automatically and traditionally. There was also an attempt to establish the main effects of methods and positions. This was done by analysis of variance (two-factor design with repeated measurements) and analysis of regression. The results are reported by Bartkowiak, Ruta, Włodarczyk (1985). The means of the 6 values of blood pressure (systolic and diastolic) recorded for men and women are shown in the form of glyphs in Figure 1.

On the right side of the glyph we have men and on the left side women. We can see a difference between men and women: the values of BP for women are generally lower than those for men. The recordings by the automatic method are marked with a solid line, and those obtained by the traditional method with a dashed line. Generally, the traditional method gives higher values than the automatic one (preference by man for rounding up to higher values?).

From an analysis of variance, as reported by Bartkowiak, Ruta, Włodarczyk (1985), it follows that the effect of the method of recording BP is statistically significant with $P < 0.01$* (except for b.p. systolic in men, where $P = 0.08$). Moreover the effect of the position is generally also statistically significant ($P < 0.001$ for BP systolic for men and women; $P = 0.04$ for BP in men and $P = 0.18$ for BP systolic in women).

Principal component analysis (performed using the program DUALNA from the SABA package, part II, developed by Bartkowiak, Krusińska (1985)) shows that the method of the recording has a stronger effect on the difference between the measurements - in comparison with the positions in which the measurements are recorded.

2. ASSESSING THE DIFFERENTIAL EFFECT OF TWO METHODS OF RECORDING ARTERIAL BLOOD PRESSURE

2.1. TESTING EQUALITY OF VARIANCES FOR PAIRWISE DEPENDENT OBSERVATIONS

Suppose X_1, X_2 are random variables denoting BP measurements recorded automatically and traditionally. Suppose further, that these variables are distributed bivariate normally:

$$\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \sim N\left\{\begin{pmatrix} \mu \\ \mu+\delta \end{pmatrix}, \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}\right\}. \tag{1}$$

---

* By P we mean the probability of rejecting under H_0 the hypothesis $H_0^{(A)}: \alpha_1 = \alpha_2$ when testing the statistical difference between methods, and hypothesis $H_0^{(B)}: \beta_1 = \beta_2 = \beta_3$ when testing the statistical difference between positions.

To test the hypothesis H_0: $\delta = 0$ we may use the classical (paired) t-test.

To test the hypothesis $H_0^{(a)}$: $\sigma_1^2 = \sigma_2^2$ one is warned against using the classical F-test (the observations of X_1 and X_2 are dependent!). We can use here the Pitman-Morgan test (Pitman (1939)), Morgan (1939)). Its idea is as follows: Let us consider the random variables U_1 and U_2, where

$$U_1 = X_1 + X_2, \quad U_2 = X_2 - X_1. \tag{2}$$

The joint distribution of U_1, U_2 is bivariate-normal:

$$\begin{pmatrix} U_1 \\ U_2 \end{pmatrix} \sim N\left\{\begin{pmatrix} 2\mu+\delta \\ \delta \end{pmatrix}, \begin{pmatrix} \sigma_1^2+\sigma_2^2+2\rho\sigma_1\sigma_2 & \sigma_2^2-\sigma_1^2 \\ \sigma_2^2-\sigma_1^2 & \sigma_1^2+\sigma_2^2-2\rho\sigma_1\sigma_2 \end{pmatrix}\right\}. \tag{3}$$

Testing $\sigma_1 = \sigma_2$ is the same as testing $\rho_{U_1,U_2} = 0$. Since U_1 and U_2 are jointly normal, the likelihood test of $\rho_{U_1,U_2} = 0$ is the usual test based on

$$T = (n-2)^{1/2} r_{U_1,U_2} / \left(1 - r_{U_1,U_2}^2\right)^{1/2}, \tag{4}$$

which, under $H_0^{(a)}$, is distributed as t_{n-2}.

Berry et al. (1984) notice, that the test (4) is algebraically equivalent to T defined as follows:

$$T = \frac{1}{2}(n-2)^{1/2}\left(\frac{S_2}{S_1} - \frac{S_1}{S_2}\right) / (1-r^2)^{1/2}, \tag{5}$$

where r, S_1, S_2 are the usual maximum likelihood estimators of ρ, σ_1, σ_2 from (1).

For our data we do not reject the hypothesis $H_0^{(a)}$. It follows, that there is no reason for rejecting the assumption that the BP measurements read automatically and in the traditional way are samples from distributions with the same variances $\sigma_1^2 = \sigma_2^2 = \sigma^2$.

A more detailed analysis of the variances in the design with pairwise dependent observations recorded in several groups of data using simultaneous test procedure will be presented in a forthcoming paper by Bartkowiak (1987).

2.2. TESTING HOMOGENEITY OF THE DIFFERENCES BETWEEN OBSERVATIONS RECORDED AUTOMATICALLY AND IN THE TRADITIONAL WAY

As was stated above in § 1, the differences between observations recorded automatically and in the traditional way are statistically significant.

We now want to investigate the difference between these measurements. It could happen that this difference depends on the magnitude of the recorded blood pressure. For example, for people with large values of arterial BP the expected difference could be larger than for people having low values of BP. Another possibility is that the differences are homogenous and do not depend from the magnitude of the recorded BP.

Suppose we observe the variables X_1 and X_2, X_1 being BP recorded automatically, and X_2 being BP recorded in the traditional way. We want to find whether the difference $Y = X_2-X_1$ depends on X_1, the BP recorded automatically. At this instance there is no preference for taking X_1 or X_2 as the dependent variable (in § 2.1 it was stated that both variables have variances not differing statistically). We plot Y against X_1 and try to estimate the functional relationship between these two variables. In the following we shall call the variable X_1 the baseline variable. Suppose this relationship is linear.

We first estimate this relationship by a regression line $Y = a+bX_1$ with the parameters a, b estimated by the least squares method. Next we look at the estimates of the parameters a, b. The parameter a (also called intercept of the regression line) determines the constant part of the difference X_2-X_1. This is the part of the difference Y that does not depend on the baseline measurement.

The parameter b (also called slope of the regression line) expresses the dependence of Y on X_1. Should the value of b be 0 then there would be no dependence on the baseline X_1. A negative value of b indicates that the (expected) difference Y decreases with increasing X_1. One would be tempted to say that for people with increased BP the difference between the two methods of BP recording is expected to be smaller than for people with low BP. But, before coming to this conclusion, we should introduce into the (estimated) slope b a correction allowing for the removal of the "regression to the mean" effect.

The name of this effect comes from Galton. He observed in 1886 the differences in hight of fathers and sons. He stated that sons of high fathers are generally smaller than would follow from the corralation between the hights of fathers and sons. He called this phenomenon „regression to the mean".

In the next section of our paper we consider this effect in more detail.

2.3. THE MATHEMATICAL BACKGROUND OF THE PHENOMENON „REGRESSION TO THE MEAN"

Let X_1, X_2 be random variables distributed bivariate normal as given by formula (1). Let $Y = X_2-X_1$ be the difference of these variables. It can be shown that the conditional expectation of the difference Y is given by the formula

$$E\{(X_2-X_1)/X_1\} = \delta-(1-\rho\theta)(X_1-\mu), \quad (6)$$

where $\theta = \sigma_2/\sigma_1$. This formula may be found in Berry et al. (1984).

It means, e.g., that for a given $x_1 > \mu$ the difference Y is expected to be lower than δ by $(1-\rho\Theta)(x_1-\mu)$, and given $x_1 < \mu$, the difference Y is expected to be higher than δ by $-(1-\rho\Theta)(x_1-\mu)$. Hence the measurement $X_1 = x$ is said to contain a regression effect $(1-\rho\Theta)(x-\mu)$. Since the regression effect is zero for $X_1 = \mu$, this effect is known as regression to the mean. It follows that, except for the case when $\rho\Theta \neq 1$, it is natural to expect that the difference Y depends on the baseline measurement X_1. For our data, with $\Theta \approx 1.0$ and ρ positive, the slope $(1-\rho\Theta)$ of the regression line $Y = a+bx$ is expected to be negative. Therefore, for people with increased BP we should generally observe smaller differences between the two methods of BP recording than for people with low BP. This follows from the assumed model with bivariate normality of the underlying distribution.

We now wish to establish the specific influence of the two methods on the observed difference Y. It follows that we should remove the „regression to the mean" effect expressed by formula (6) and observe a transformed variable $\tilde{Y}$ free from this effect. The dependence of this new variable $\tilde{Y}$ on the baseline measurement X_1 will indicate a genuine effect of the methods of recording BP on the observed difference between values of BP recorded by these methods.

To remove the regression to the mean effect we define a new random variable $\tilde{Y}$ as follows:

$$\tilde{Y} = X_2-X_1+(1-\rho\Theta)(X_1-\mu). \tag{7}$$

The conditional expectation of $\tilde{Y}$ given X_1 equals δ, i.e.

$$E(\tilde{Y}/X_1) = \delta.$$

If $H_0^{(a)}: \sigma_1^2 = \sigma_2^2$ is true, then the maximum likelihood estimators of $\sigma_1^2 = \sigma_2^2 = \sigma^2$ are

$$\hat{\sigma}^2 = \frac{1}{2}\left(S_1^2+S_2^2\right),$$

and

$$\hat{\rho} = rS_1S_2/\hat{\sigma}^2,$$

where S_1^2, S_2^2 and r are the usual unrestricted maximum likelihood estimators of σ_1^2, σ_2^2 and ρ respectively.

3. RESULTS OF CALCULATIONS FOR THE DATA

Our goal was to investigate the differential effects of two methods of BP recording. We had observations for systolic and diastolic blood pressure recorded in men and women. For each person the measurements were taken in 3 positions. In the following we denote by X_1, X_2, X_3 the BP measurements recorded automatically in the positions: standing (X_1), sitting (X_2), and lying (X_3). By X_4, X_5, X_6 we denote the measurements recorded in the traditional way. From these variables we form the differences:

$$Y_1 = X_4 - X_1, \quad Y_2 = X_5 - X_2, \quad Y_3 = X_6 - X_3.$$

We take as the baseline the variables X_1, X_2, X_3 (recordings done automatically). For each of the four groups of data (systolic blood pressure, diastolic blood pressure for women, men) we plot the regressions

Y_1 versus X_1,
Y_2 versus X_2,
Y_3 versus X_3.

An example of the plot (diastolic BP for men) is given in Figure 2.

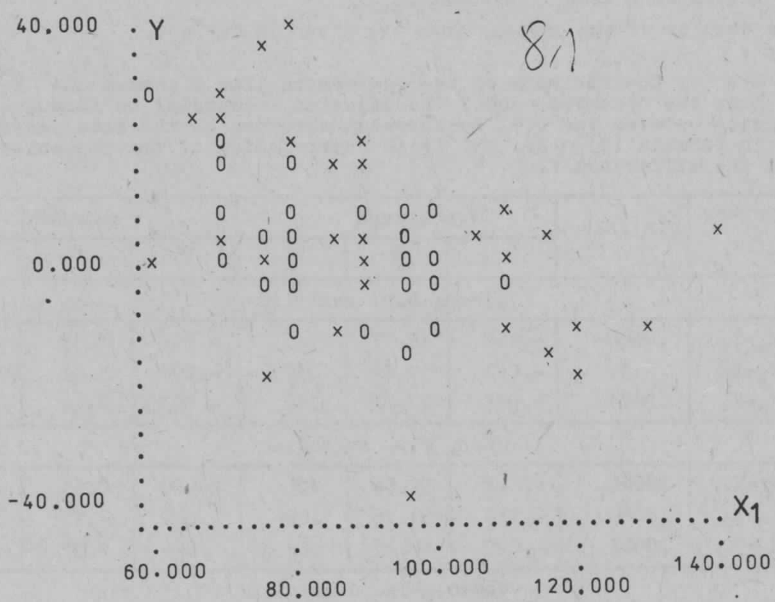

Fig. 2. Unadjusted difference $Y = X_2 - X_1$ against X_1. Multiple points signed by „0"

We should note that the observed BP values were rounded up to integers with the last digit 5. Therefore on Fig. 2 (and later also on Fig. 3) the plotted points appear on a grid. Single points are marked by „X", and multiple points by „0". The plot was made by the lineprinter of the ODRA 1305 computer using the program SCAT from the package SABA, part II, described by Bartkowiak and Krusińska (1985).

Next we calculated the estimates of the parameters a, b of the regression line Y = a+bx. Here we used the least squares method.

For each regression we tested the hypothesis $H_0 : b = 0$ using the classical F-test (see, e.g. Bartkowiak, Krusińska (1985)). First we calculted

$$F_c = SSR/(SSE/(n-2)), \quad (8)$$

where $SSR = \hat{b} \sum_{i=1}^{n} (y_i-\bar{y})(x_i-\bar{x})$, $SSE = \sum_{i=1}^{n} (y_i-\hat{a}-\hat{b}x_i)^2$, n is the sample size, $y_1, y_2, \ldots, y_n$ and $x_1, x_2, \ldots, x_n$ are the observed values of the variables Y and X. Under H_0 the statistic F given by (8) has a F distribution with $v_1 = 1$ and $v_2 = n-2$ degrees of freedom. Next we calculated

$$P = P(F > F_c/H_0). \quad (9)$$

Small values of P testify against H_0.

The results of the calculations are given in Table 1.

Table 1. Coefficients of the regression line Y = bx+a and $\tilde{Y}$ = bx+a, where Y is the observed, and $\tilde{Y}$ the adjusted (according to formula (7)) difference between two b.p. measurement recorded on the same person. P is given in formula (9). $\rho(x_i,x_j)$ is the correlation of the variables entering in the difference Y.

Difference Y	$\rho(x_i,x_j)$	Y observed			$\tilde{Y}$ adjusted		
		b	a	P	b	a	P
Women, b.p. systolic							
$Y = X_4-X_1$	.8968	-.102	16.77	.01	-.000	3.55	1.00
$Y = X_5-X_2$	.9153	-.041	10.81	.30	-.001	5.50	0.92
$Y = X_6-X_3$	.9021	-.045	11.99	.30	-.003	6.34	0.90
Men, b.p. systolic							
$Y = X_4-X_1$	.8640	-.222	30.66	.00	-.000	0.63	1.00
$Y = X_5-X_2$	.8880	-.135	21.16	.00	.002	2.11	0.92
$Y = X_6-X_3$	.8942	-.099	16.20	.02	.004	1.57	0.89
Women, b.p. diastolic							
$Y = X_4-X_1$	.7688	-.177	18.28	.01	.000	3.80	1.00
$Y = X_5-X_2$	.8147	-.025	4.99	.70	-.003	3.11	0.92
$Y = X_6-X_3$	.7866	-.205	18.28	.00	-.044	4.55	0.47
Men, b.p. diastolic							
$Y = X_4-X_1$	.6634	-.344	32.11	.00	.000	3.07	1.00
$Y = X_5-X_2$	.6739	-.329	30.99	.00	-.016	3.71	0.80
$Y = X_6-X_3$	.7738	-.247	22.80	.00	-.024	3.38	0.70

We see that all estimates of slopes of the regression lines $Y = a+bx$ are negative. For 9 out of 12 regressions the P values are very small, whence it follows that the hypothesis H_0: $b = 0$ should be rejected. As stated in § 2.3, this could be expected from the correlations between the measurements. Therefore the negative slopes of the regression lines $Y = a+bx$ are natural as due to the "regression to the mean" effect.

To clarify whether there is any genuine differential effect due to the methods of recording the measurements, we should remove the „regression to the mean" effect. We do it by calculations of the adjusted difference $\tilde{Y}$ defined by formula (7). Next we again plot $\tilde{Y}$ against the appropriate baseline measurement and calculate the coefficients a, b of the regression $\tilde{Y} = a+bX$. The estimates of the coefficients a, b are given in Table 1. An example of the plot of the adjusted difference $\tilde{Y}$ against the baseline X_1 (diastolic BP for men) is given in Figure 3.

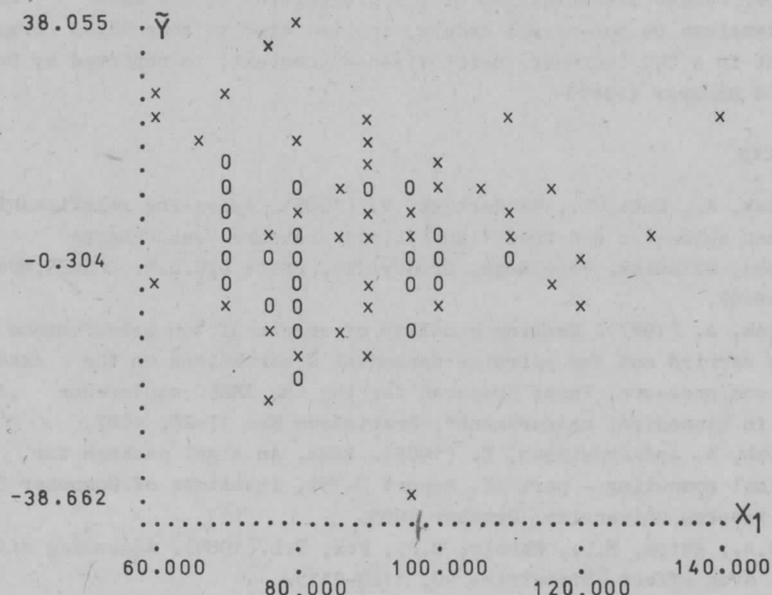

Fig. 3. Adjusted difference $\tilde{Y}$ against $\tilde{X}_1$. Multiple points signed by „0"

We see that the values of b are almost ideally equal zero. It follows, that now the adjusted differences $\tilde{Y}$ do not depend on the baseline measurements, hence the differences of the recordings obtained by both method are homogeneous with respect to the magnitude of the baseline measurements.

4. FINAL REMARKS

Our conclusions in § 3 were based on the assumption that the distributions of the variables observed in our experiment are normal. Generally this is not true for arterial blood pressure. Nonetheless we should state that our distributions do not correspond exactly to the distributions of arterial BP in the population of adult men and women. Our samples were specially enlarged in the intervals corresponding to higher values of BP The closeness of our obserwations of the BP (systolic and diastolic) to the normal distribution was investigated in the paper by Bartkowiak, Ruta, Włodarczyk (1985). They stated that the assumption of normality is not grossly violated for the underlying data. Therefore we can conclude that our data fit into the model presented in § 2 and we are justified to draw the conclusions stated in § 3.

Some considerations in the case of nonnormal underlying distributions can be found in a paper by Das and Mulder (1983).

Another recent investigation of the „regression to the mean" effect with extensions to non-normal models, applied also to some blood pressure data, but in a CHD (coronary heart disease) context, is reported by Dobson, Beath and Shaerer (1986).

REFERENCES

Bartkowiak, A., Ruta, R., Włodarczyk, W. (1985). Assessing relationships between automatic and traditional blood pressure measurements (in Polish), Człowiek, Populacja, Środowisko, Prace D.C.D.M. DOLMED, Wrocław 12, 49-92.

Bartkowiak, A. (1987). Testing equality of errors of two measurements devices carried out for pairwise dependent observations on the example of blood pressure. Paper prepared for the 4th IMEKO conference „Advances in biomedical measurement", Bratislava May 17-20, 1987.

Bartkowiak, A. and Krusińska, E. (1985). SABA, An Algol package for statistical computing - part II. Report N-151, Institute of Computer Science, Wrocław University, October 1985.

Berry, D.A., Eaton, M.L., Ekholm, B.P., Fox, T.L.(1984). Assessing differential drug effect, Biometrics 40, 1109-1115.

Das, P. and Mulder, P.G.H. (1983). Regression to the mode. Statistica Neerlandica 37, 15-20.

Dobson, A.J., Beath, K.J., Shearer, W. (1986). Regression to the mean.In: Francis, Manly, Lam (eds.), Pacific Statistical Congress, Elsevier Science Publishers, North Holland, 1986, 62-64.

Morgan, W.A. (1939). A test for the significance of the difference between the two variances in a sample from a normal bivariate population, Biometrika 31, 13-39. (reference after Berry et al.)

Pitman E.J.G. (1939). A note on normal correlation, Biometrika 31, 9-12. (reference after Berry et al.)