Journal of the American Statistical Association (1995) 90, 1373-1379

CORRELATION ANALYSIS OF EXTREME OBSERVATIONS FROM A MULTIVARIATE NORMAL DISTRIBUTION

INGRAM OLKIN AND MARLOS VIANA

ABSTRACT. In measuring visual acuity, the extremes of a set of normally distributed measures are of concern, together with one or more covariates. This leads to a model in which $(\mathrm{X}, \mathrm{Y}_1, \mathrm{Y}_2)$ are jointly normally distributed with $\mathrm{Y}_1, \mathrm{Y}_2$ exchangeable and $(\mathrm{X}, \mathrm{Y}_\mathrm{i})$ having a common correlation. Inferential procedures are developed for correlations and linear regression parameters of X and the ordered Y-values. This requires the determination of the covariance matrix of X, $\mathrm{Y}_{(1)} = \min\{\mathrm{Y}_1, \mathrm{Y}_2\}$ and $\mathrm{Y}_{(2)} = \max\{\mathrm{Y}_1, \mathrm{Y}_2\}$. The inadequacy of certain estimates that ignore the non- normality of $\{\mathrm{X}, \mathrm{Y}_{(1)}, \mathrm{Y}_{(2)}\}$ is also discussed.

AMS 1991 Mathematics Subject Classifications: 62G30, 62J05
Key words: Order statistics, regression with order statistics.

1991 *Mathematics Subject Classification.* 62G30, 62J05.
Key words and phrases. Order statistics, regression with order statistics.

1. Introduction

Visual acuity is measured by a (Snellen) eye chart and expressed in log units of the minimum angle of resolution, or Log MAR. Smaller values of Log MAR correspond to better visual acuity. Normally, a single measure of visual acuity is made in each eye, say $\mathrm{Y}_1, \mathrm{Y}_2$, together with one or more covariates X, such as the subject's age, physical condition, etc. However, because visual acuities generally are unequal in the presence of certain macular lesions, of interest are not the measures $\mathrm{Y}_1, \mathrm{Y}_2$ but rather the extreme visual acuities, the *best* acuity $\mathrm{Y}_{(1)} = \min\{\mathrm{Y}_1, \mathrm{Y}_2\}$ and the *worst* acuity $\mathrm{Y}_{(2)} = \max\{\mathrm{Y}_1, \mathrm{Y}_2\}$. Another common example is the adopted criterion for an unrestricted driver's license, which in the majority of states is based on the visual acuity of the best eye [see Fishman et al. (1993) , Szlyk et al. (1993)]. Other applications include the assessment of defective hearing in adult mentally retarded persons based on the ear with best hearing (Parving and Christensen 1990); the predictive value of the worst vision following filtration operation in the aphakic eyes of the elderly patients with glaucoma (Frenkel and Shin 1986); the comparison between best vision achieved following cataract surgery and two methods of secondary posterior capsulotomy described by Konelle (1985); sports injury data on reduction of best vision on damaged eyes (Aburn 1990), or the analysis of worst vision among patients treated for macular edema (Rehak and Vymazal 1989).

Although we can obtain maximum likelihood or Bayes estimates of the parameters of the underlying distribution of $(\mathrm{X}, \mathrm{Y}_1, \mathrm{Y}_2)$, our current interest is with the parameters of the distribution of $(\mathrm{X}, \mathrm{Y}_{(1)}, \mathrm{Y}_{(2)})$. More specifically, we are interested in making inferences on the covariance structure defined by

$$\Psi = \mathrm{Cov}(\mathrm{X},\ \mathrm{Y}_{(1)},\ \mathrm{Y}_{(2)}),$$

with particular interest in assessing the correlations η_{i} between X and $\mathrm{Y}_{(\mathrm{i})}$, the correlation θ between $\mathrm{Y}_{(1)}$ and $\mathrm{Y}_{(2)}$ and the parameters defining the best linear predictors between X and the ordered **Y**-values. Because the assumption of a multivariate normal distribution for $\{\mathrm{X}, \mathrm{Y}_1, \mathrm{Y}_2\}$ does not extend to that for $(\mathrm{X}, \mathrm{Y}_{(1)}, \mathrm{Y}_{(2)})$, we need to determine how well the correlation parameters η_{i} and θ reflect the actual dependence between the corresponding variables.
Although the model of interest relates to two Y-values, some of the results are generalizable to p variables $\mathrm{Y}_1, \mathrm{Y}_2, \ldots, \mathrm{Y}_{\mathrm{p}}$ and to the order statistics $\mathrm{Y}_{(1)} \leq \mathrm{Y}_{(2)} \leq \ldots \leq \mathrm{Y}(\mathrm{p})$. Also, the normality assumption can be weakened to that of an elliptically contoured distribution. In some instances we state the results for the more general case.

2. Models, assumptions and main results

Because of the natural symmetry between responses from fellow eyes as well as between the additional measurement X and the response of either eye, we assume that the distribution of $(\mathrm{X}, \mathrm{Y}_1, \mathrm{Y}_2)$ is multivariate normal jointly exchangeable in $(\mathrm{Y}_1, \mathrm{Y}_2)$, that is, $(\mathrm{X}, \mathrm{Y}_1, \mathrm{Y}_2)$ and $(\mathrm{X}, \mathrm{Y}_2, \mathrm{Y}_1)$ are equally distributed. This assumption translates the vector of means μ and covariance matrix Σ to be

$$\mu^{'} = (\mu_0, \mu_1, \mu_1), \quad \Sigma = \begin{bmatrix} \sigma^2 & \gamma\sigma\tau & \gamma\sigma\tau \\ \gamma\sigma\tau & \tau^2 & \rho\tau^2 \\ \gamma\sigma\tau & \rho\tau^2 & \tau^2 \end{bmatrix}, \qquad \gamma^2 < \frac{1+\rho}{2}, \quad \rho^2 < 1, \tag{2.1}$$

where the range of the parameters is such as to guarantee that the covariance matrix Σ is positive definite. When there are p Y-values, the restriction is $\gamma^2 \leq [1 + (\mathrm{p}-1)\rho]/\mathrm{p}$. Because of the context of Y_1 and Y_2 in this model the correlation ρ is restricted to the range $[0, 1)$ rather then the full range. We now provide a summary of the results concerning the ordered $\mathrm{Y}_{(\mathrm{i})}$, for which proofs are given in the Appendix.

The covariance matrix Ψ of $\mathrm{X}, \mathrm{Y}_{(1)}, \mathrm{Y}_{(2)}$ is given by

$$\Psi = \begin{bmatrix} \sigma^2 & \gamma\sigma\tau & \gamma\sigma\tau \\ \gamma\sigma\tau & \tau^2[\rho + (1-\rho)\mathrm{c}_{11}] & \tau^2[\rho + (1-\rho)\mathrm{c}_{12}] \\ \gamma\sigma\tau & \tau^2[\rho + (1-\rho)\mathrm{c}_{21}] & \tau^2[\rho + (1-\rho)\mathrm{c}_{22}] \end{bmatrix}, \tag{2.2}$$

where

$$\mathcal{C} = \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix} = \begin{bmatrix} 0.6817 & 0.3183 \\ 0.3183 & 0.6817 \end{bmatrix} \tag{2.3}$$

is the covariance matrix between the largest and the smallest of two independent standard normal variables (Beyer 1990, Table VII.2 pg. 243). The matrix $\mathcal{C}$ is tabulated for values of $p \geq 2$.

A key result is that the covariance between X and $Y_{(i)}$ is equal to the covariance between X and Y_i, which is surprising. This is a consequence of the exchangeability properties of the normality model described by (2.1). We comment further about this fact in the Appendix, where a proof of a more general result is provided.

2.1 Correlations and linear regressions. From the symmetries in (2.2), the correlations between X and $Y_{(i)}$ are common,

$$\eta = \mathrm{corr}(X, Y_{(i)}) = \frac{\gamma}{\sqrt{\rho + (1-\rho)c_{ii}}}, \qquad i = 1, 2, \tag{2.4}$$

where $-1 \leq \eta \leq 1$. Note that $\eta = 0$ if and only if $\gamma = 0$, in which case X and (Y_1, Y_2) are independent. However, it is also true that $\eta = 0$ implies that X and $(Y_{(1)}, Y_{(2)})$ are independent; this fact is not obvious since $(X, Y_{(1)}, Y_{(2)})$ are no longer jointly normally distributed. A proof is provided in the Appendix. The regressions of $Y_{(i)}$ on X are linear and parallel, and determined by

$$y_{(i)} = (\mu_1 - \frac{\gamma\tau}{\sigma}\mu_0 + c_i\tau\sqrt{1-\rho}) + \frac{\gamma\tau}{\sigma}x, \quad i = 1, 2, \tag{2.5}$$

where

$$\mathbf{c}' = (c_1, c_2) = (-0.56419, \;\; 0.56419) \tag{2.6}$$

is the expected value of the smallest and largest of two independent standard normal variables. The vertical distance between the two lines defined by (2.5) is equal to $2c_2\tau\sqrt{1-\rho} = 1.12838\,\tau\sqrt{1-\rho}$. The coefficients of the best linear predictors $x = \beta_{0i} + \beta_{1i}y_{(i)}$ of X on $Y_{(i)}$ are given by

$$\beta_{0i} = \mu_0 - \beta_1[\mu_1 + c_i\tau\sqrt{1-\rho}\;], \quad \beta_1 = \beta_{1i} = \frac{\sigma}{\tau}\,\frac{\gamma}{\rho + (1-\rho)c_{ii}} \quad i = 1, 2, \tag{2.7}$$

where the constants are given by (2.3) and (2.6). Similarly, these two lines are also parallel, with a vertical distance of $2c_2\tau\sqrt{1-\rho}\;|\;\beta_1\;|$ one from the other. The correlation θ between the extreme values $Y_{(1)}$ and $Y_{(2)}$ is

$$\theta = \mathrm{corr}(Y_{(1)}, Y_{(2)}) = \frac{\rho + (1-\rho)c_{12}}{\rho + (1-\rho)c_{22}}. \tag{2.8}$$

In the range $\rho \in (0, 1)$, $0.4669 = \frac{c_{12}}{c_{22}} < \theta < 1$, whereas the partial correlation of $Y_{(1)}$ and $Y_{(2)}$ given X is

$$\theta_{12|0} = \frac{\rho + (1-\rho)c_{12} - \gamma^2}{\rho + (1-\rho)c_{22} - \gamma^2} \;\leq\; \theta. \tag{2.9}$$

Thus, the partial correlation is always a contraction of the product moment correlation, regardless of the composition of the covariate.

3. Maximum likelihood and large-sample estimates

Given a sample $(x_\alpha, y_{1\alpha}, y_{2\alpha})$, $\alpha = 1, \ldots, N$ of size N with means $\bar{x}$, $\bar{y}_i$, cross-product matrix $A = (a_{ij})$,

$$a_{00} = \sum_{\alpha=1}^{N}(x_\alpha - \bar{x})^2, \quad a_{0j} = \sum_{\alpha=1}^{N}(x_\alpha - \bar{x})(y_{j\alpha} - \bar{y}_j), \quad a_{ij} = \sum_{\alpha=1}^{N}(y_{i\alpha} - \bar{y}_i)(y_{j\alpha} - \bar{y}_j),$$

$i, j = 1, 2$, and sample covariance matrix $S = A/n$, where $n = N - 1$, the maximum likelihood estimates of σ^2, τ^2, ρ and γ are given by

$$\hat{\sigma}^2 = \frac{a_{00}}{N}, \quad \hat{\tau}^2 = \frac{1}{2N}(a_{11} + a_{22}), \quad \hat{\rho} = \frac{a_{12}}{\frac{1}{2}(a_{11} + a_{22})}, \quad \hat{\gamma} = \frac{\frac{1}{2}(a_{01} + a_{02})}{\sqrt{\frac{1}{2}(a_{11} + a_{22})}\sqrt{a_{00}}}.$$

The standard delta method derivation [e.g. (Anderson 1985, pg. 120)] shows that the large-sample joint distribution of $\sqrt{\mathrm{N}}\,(\hat\rho-\rho,\ \hat\gamma-\gamma)$ is bivariate normal with mean vector $(0,0)$ and covariance matrix given by

$$\left[\begin{array}{cc} (1-\rho^2)^2 & \gamma(1-\rho)[-\gamma^2+\frac{1}{2}(2-\rho)(1+\rho)] \\ \\ \gamma(1-\rho)[-\gamma^2+\frac{1}{2}(2-\rho)(1+\rho)] & \frac{1}{4}[4\gamma^4-\gamma^2(5-\rho)(1+\rho)+2(1+\rho)] \end{array}\right]. \tag{3.1}$$

The asymptotic variance of $\hat\rho$ depends only on ρ, whereas the asymptotic variance of $\hat\gamma$ depends on both ρ and γ. In particular, note that when $\gamma=0$, from (3.1),

$$\mathrm{ACov}(\hat\rho,\hat\gamma\mid\gamma=0)=\left[\begin{array}{cc}(1-\rho^2)^2 & 0\\ 0 & \frac{1}{2}(1+\rho)\end{array}\right],$$

so that $\hat\rho$ and $\hat\gamma$ are asymptotically independent when X is uncorrelated with $\mathrm{Y}_1,\ \mathrm{Y}_2$. Also,

$$\mathrm{ACov}(\hat\rho,\hat\gamma\mid\rho=0)=\left[\begin{array}{cc}1 & \gamma(1-\gamma^2)\\ \\ \gamma(1-\gamma^2) & \frac{1}{4}[1+(1-\gamma^2)(1-4\gamma^2)]\end{array}\right].$$

As indicated earlier, our concern is with the parameters of the distribution of $(\mathrm{X},\mathrm{Y}_{(1)},\mathrm{Y}_{(2)})$, given by (2.4) and (2.8), from which maximum likelihood estimates of η and θ are given by

$$\hat\eta=\frac{\hat\gamma}{\sqrt{\hat\rho+(1-\hat\rho)\mathrm{c}_{11}}},\quad \hat\theta=\frac{\hat\rho+(1-\hat\rho)\mathrm{c}_{12}}{\hat\rho+(1-\hat\rho)\mathrm{c}_{11}}. \tag{3.2}$$

The asymptotic joint distribution of $\sqrt{\mathrm{N}}\,(\hat\eta-\eta,\ \hat\theta-\theta)$ is normal with means zero, variances

$$\begin{aligned}\mathrm{AVar}(\hat\eta) &= \frac{1}{4(\rho+\bar\rho\mathrm{c}_{11})}\{4\gamma^4-\gamma^2(\rho+1)(5-\rho)+2(\rho+1)+\frac{\bar{\mathrm{c}}_{11}^2(1-\rho^2)^2}{(\rho+\bar\rho\mathrm{c}_{11})^2}\\ &\qquad -\frac{4\bar{\mathrm{c}}_{11}\gamma\bar\rho}{\rho+\bar\rho\mathrm{c}_{11}}[-\rho^2+\frac{1}{2}(2-\rho)(1+\rho)]\}, \qquad (3.3)\\ \mathrm{Avar}(\hat\theta) &= \frac{(\mathrm{c}_{11}-\mathrm{c}_{12})^2(1-\rho^2)^2}{(\rho+\bar\rho\mathrm{c}_{11})^4}, \qquad (3.4)\end{aligned}$$

where $\bar\rho=1-\rho$, $\bar{\mathrm{c}}_{11}=1-\mathrm{c}_{11}$, and covariance

$$\mathrm{ACov}(\hat\eta,\ \hat\theta)=\frac{(\mathrm{c}_{11}-\mathrm{c}_{12})\gamma}{(\rho+\bar\rho\mathrm{c}_{11})^{5/2}}\{\frac{-\bar{\mathrm{c}}_{11}(1-\rho^2)^2}{2(\rho+\bar\rho\mathrm{c}_{11})}+(1-\rho)[-\rho^2+\frac{1}{2}(2-\rho)(1+\rho)]\}. \tag{3.5}$$

In particular, note that

$$\mathrm{ACov}(\hat\eta,\ \hat\theta\mid\gamma=0)=\left[\begin{array}{cc}1/(2\mathrm{c}_{11}) & 0\\ 0 & (\mathrm{c}_{11}-\mathrm{c}_{12})^2/\mathrm{c}_{11}^4\end{array}\right],$$

so that $\hat\eta$ and $\hat\theta$ are asymptotically independent when X and Y_i are independent.

Tests for correlations: an exact test for $\gamma=0$. As pointed out earlier, the hypothesis $\gamma=0$ of independence between X and $\mathbf{Y}=(\mathrm{Y}_1,\ldots,\mathrm{Y}_\mathrm{p})$ is equivalent to the hypothesis of independence between X and the corresponding vector $\mathcal{Y}$ of order statistics and to the hypothesis of null slopes in the regression lines between $\mathcal{Y}$ and X. The hypothesis $\gamma=0$ can be assessed from the large sample distribution of $\hat\gamma$ indicated by (3.1), or alternatively as follows: Let r_{11} indicate the sample equicorrelation coefficient

$$\mathrm{r}_{11}=\frac{2\sum_{\mathrm{i}<\mathrm{j}}(\mathrm{A}_{11})_{\mathrm{ij}}}{(\mathrm{p}-1)\mathrm{tr}\ \mathrm{A}_{11}}$$

associated with the sample p×p matrix of cross- products A_{11} which has a Wishart distribution $\mathrm{W}_\mathrm{p}(\tau^2[\rho_{11}\mathbf{e}\mathbf{e}'+(1-\rho_{11})\mathbf{I}],\mathrm{n})$. Similarly, let $\mathrm{r}_{11|0}$ indicated the sample equicorrelation coefficient based on

$$\mathrm{A}_{11|0}=\mathrm{A}_{11}-\mathrm{A}_{10}\mathrm{A}_{00}^{-1}\mathrm{A}_{01}\ \sim\ \mathrm{W}_\mathrm{p}(\tau^2(1-\gamma^2)[\rho_{11|0}\mathbf{e}\mathbf{e}'+(1-\rho_{11|0})\mathbf{I}],\mathrm{n}),$$

where $\rho_{11|0} = (\rho - \gamma^2)/(1-\gamma^2)$. The canonical representations of $A_{11|0}$ and $A_{11} \sim W_p(\Sigma_{11}, n)$ are determined by

$$
\begin{aligned}
U_1 &= \frac{\mathrm{tr}\ nS_{11|0}}{p}[1+(p-1)r_{11|0}] \sim \tau^2[1+(p-1)\rho - p\gamma^2]\chi^2_{n-p},\\
U_2 &= (p-1)\frac{\mathrm{tr}\ nS_{11|0}}{p}[1-r_{11|0}] \sim \tau^2(1-\rho)\chi^2_{(p-1)(n-p)},\\
V_1 &= \frac{\mathrm{tr}\ nS_{11}}{p}[1+(p-1)r_{11}] \sim \tau^2[1+(p-1)\rho]\chi^2_{n},\\
V_2 &= (p-1)\frac{\mathrm{tr}\ nS_{11}}{p}[1-r_{11}] \sim \tau^2(1-\rho)\chi^2_{(p-1)n}.
\end{aligned}
$$

Given ρ, τ, γ, U_1 is independent of U_2, and V_1 is independent of V_2. In addition, when $\gamma = 0$, from (Anderson 1985, corollary 4.3.2.)] it follows that

$$V_1 - U_1 \ \sim\ \tau^2[1+(p-1)\rho]\chi^2_p,$$

independently of V_1. Consequently, when $\gamma = 0$,

$$\frac{n}{p}\,\frac{V_1 - U_1}{V_1} = \frac{n}{p}[1 - \frac{\mathrm{tr}\ S_{11|0}[1+(p-1)r_{11|0}]}{\mathrm{tr}\ S_{11}[1+(p-1)r_{11}]} \ \sim\ F_{p,n}. \tag{3.6}$$

Similarly, when $\rho = 0$, directly from the canonical representation of A_{11},

$$(p-1)\frac{V_1}{V_2} = \frac{1+(p-1)r_{11}}{1-r_{11}} \ \sim\ F_{n,(p-1)n}, \tag{3.7}$$

so that (3.6) and (3.7) can be used to assess the corresponding hypotheses. Note that when γ is actually different from zero we should expect larger values of (3.6); when ρ is actually positive we should expect larger values of (3.7), and expect smaller values when ρ is actually negative.

4. The visual acuity data

The following results are based on $N = 43$ subjects participating in a larger experiment reported by (Fishman, Baca, Alexander, Derlacki, Glenn and Viana 1993), in which patients with Best's vitelliform macular dystrophy were evaluated for bilateral visual acuity loss (Y_1, Y_2) and age X. This is an autosomal dominant disorder that begins in childhood and is characterized by an egg-yolk-like lesion in the macula, resulting in extensive pigmentary macular degeneration. Starting with the sample means $(\bar{x}, \bar{y}_1, \bar{y}_2) = (28.833, 0.412, 0.437)$, covariance matrix S and corresponding correlation matrix R,

$$S = \begin{bmatrix} 367.996 & 4.419 & 4.200 \\ 4.419 & 0.135 & 0.074 \\ 4.200 & 0.074 & 0.163 \end{bmatrix}, \quad R = \begin{bmatrix} 1.000 & 0.627 & 0.542 \\ 0.627 & 1.000 & 0.499 \\ 0.542 & 0.499 & 1.000 \end{bmatrix},$$

one of the questions of interest is the correlation between age, X, and the extreme visual acuities $Y_{(1)}, Y_{(2)}$. Because the visual acuity Y_1, Y_2 on (respectively right and left) fellow eyes are expected to be about the same, to have about the same variability and to be equally correlated with age, we use the model defined in Section 1 to describe the data on (X, Y_1, Y_2).

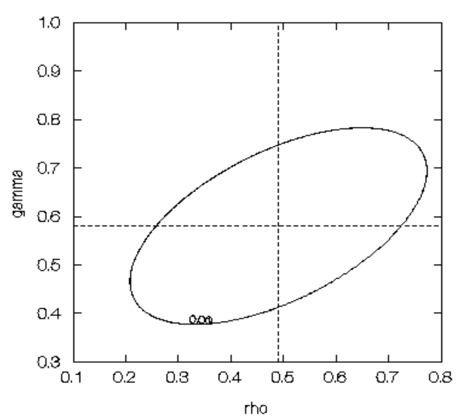
From Section 3, the maximum likelihood estimate of the correlation ρ between vision on fellow eyes is $\hat{\rho} = 0.496$, and the correlation between age and vision in either eye is $\hat{\gamma} = 0.581$. In addition, the estimated standard deviation of vision in either eye is $\hat{\tau} = 0.386$, the age standard deviation is $\hat{\sigma} = 19.182$, the mean vision and age are $\hat{\mu}_1 = 0.424$ and $\hat{\mu}_0 = 28.83$, respectively. Maximum likelihood estimates (3.2) of the correlation η between age and extreme vision and the correlation θ between extreme vision are $\hat{\eta} = 0.634$ and $\hat{\theta} = 0.782$.

The large-sample joint distributions of $\sqrt{\mathrm{N}}\,(\hat\rho-\rho,\ \hat\gamma-\gamma)$ and of $\sqrt{\mathrm{N}}\,(\hat\eta-\eta,\ \hat\theta-\theta)$ are bivariate normal (see (3.1)) with mean vector $(0,0)$ and respective covariance matrices

$$\mathrm{ACov}(\hat\rho,\hat\gamma)\approx\begin{bmatrix}0.567 & 0.230\\ 0.230 & 0.293\end{bmatrix},\quad \mathrm{ACov}(\hat\eta,\hat\theta)\approx\begin{bmatrix}0.108 & 0.109\\ 0.109 & 0.150\end{bmatrix},\tag{4.1}$$

as determined from (3.3), (3.4), and (3.5). Approximate 95% confidence regions for (ρ,γ) and for (η,θ) shown in Figures 1 and 2 indicate a significant correlation between fellow eyes, between age and vision, between age and extreme vision, and between best and worst vision. As pointed out earlier, under the present model (2.1), evidence in favor of non-null correlation γ between age and vision coincides with evidence in favor of dependence between age and extreme vision. The exact test statistic for testing $\gamma=0$ is $\mathrm{F}_{\mathrm{p,n}}=9.48$, which supports the large-sample conclusion of a non-null correlation γ between age and vision. The test statistic for testing $\rho=0$ is $\mathrm{F}_{\mathrm{n,(p-1)n}}=2.97$ which supports the earlier claim of a positive correlation ρ between vision of fellow eyes [see (3.6) and (3.7) for details of the tests]. The best linear predictor of the best vision,

FIGURE 1. Approximate 95% confidence regions for ρ and γ.



$\mathrm{Y}_{(1)}$, and the worst vision, $\mathrm{Y}_{(2)}$, given age, obtained from (2.5) are

$$\mathrm{y}_{(2)}=0.241+0.0117\mathrm{x},\quad \mathrm{y}_{(1)}=-0.067+0.0117\mathrm{x}.$$

Figure 3 shows the estimated lines. Because the regression coefficient of age is zero only when $\gamma=0$, we are confident that age is relevant to linearly predicting extreme vision. Note that the vertical distance between the two parallel lines can be estimated directly from $2\mathrm{c}_2\hat\tau\sqrt{1-\hat\rho}$, which in the present case is 0.308 log MAR units. The paralellism of the two lines indicates that there may perhaps be, on average, a constant deficit (of 0.308 log MAR units) between fellow eyes that remains constant with age. That fact was not expected. Moreover, because the macular dystrophy considered in the study is an inherited disorder, the fact that the predicted best vision $\mathrm{y}_{(1)}$ at birth is normal (log MAR=0) suggests a number of interesting questions to be investigated with a follow-up study of age and vision for patients with Best dystrophy. However, note that the paralellism is a necessary property of the model and would hold regardless what $\mathrm{x},\mathrm{y}_{(1)}$ and $\mathrm{y}_{(2)}$ are. The estimated mean-squared error matrix is

$$\hat\Psi_{11|0}=\hat\tau^2(1-\hat\gamma^2)[\hat\rho_{11|0}\begin{bmatrix}1 & 1\\ 1 & 1\end{bmatrix}+(1-\hat\rho_{11|0})\begin{bmatrix}0.6817 & 0.3183\\ 0.3183 & 0.6817\end{bmatrix}]=\begin{bmatrix}0.0747 & 0.0475\\ 0.0475 & 0.0747\end{bmatrix},$$

Figure 2. Approximate 95% confidence regions for η and θ.

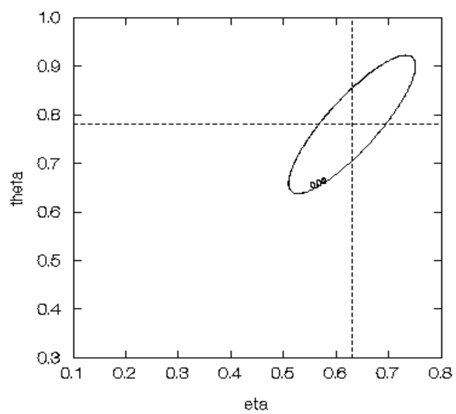
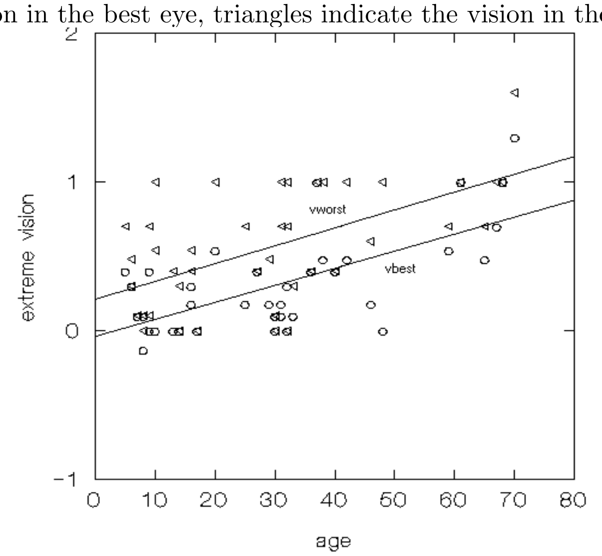


Figure 3. Estimated linear regression of extreme visual acuity on age. Dots indicate the vision in the best eye, triangles indicate the vision in the best eye.



whereas the estimated multiple correlation coefficient η^2 in each equation is 0.402.

From (2.7), the estimated best linear predictors of age given extreme vision acuities are determined by

$$\mathrm{x} = 8.93 + 34.389\mathrm{y}_{(1)}, \quad \mathrm{x} = 19.563 + 34.389\mathrm{y}_{(2)},$$

the estimated mean-squared error is 220.06, whereas the estimated multiple correlation coefficient is 0.27. The significance of vision in the regression model above is equivalent to the significance of the correlation γ between age and vision, which was found to be present.

The multiple linear regressions predicting one extreme vision from age and the other extreme, (determined from equations (A.6) and (A.7)), are

$$\mathrm{y}_{(1)} = -0.21 + 0.6351\,\mathrm{y}_{(2)} + 0.0042\,\mathrm{x}, \qquad \mathrm{y}_{(2)} = 0.28 + 0.6351\,\mathrm{y}_{(1)} + 0.0042\,\mathrm{x}. \tag{4.2}$$

These planes are parallel. Because testing the relevance of the age component coincides with testing the hypothesis $\gamma = 0$, age was found to be relevant. The relevance of the extreme vision, however, is equivalent to testing the hypothesis $\theta_{12|0} = 0$ of null partial correlation. From (2.9) it follows that the equivalent hypothesis is $\gamma^2 = \mathrm{c}_{12} + \mathrm{c}_{22}\rho$. Since the 95% confidence region for ρ, γ shown in Figure 1 does not intersect the curve determined by the hypothesis $\theta_{12|0} = 0$, we are inclined to accept the significance of the extreme vision to predict, along with age, the other extreme vision.

This result is surprising. A casual analysis (ignoring that age, best vision, and worst vision are not jointly normally distributed) of the same regression model just described leads to the conclusion that age in the presence of best vision is not relevant to predict the worst vision. However, in the presence of worst vision, age becomes relevant to predict the best vision. The ratio of the respective t statistics associated with the significance of age, obtained by casually entering age, best and worst vision into a multiple regression program and executing the two regressions, turned out to be larger than 2.5.

5. Conclusions

We have described the covariance structure between the vector $\mathcal{Y}$ of ordered components of $\mathbf{Y}$ and another covariate X and the covariance structure of $\mathcal{Y}$, assuming that the joint distribution of $\mathrm{X}, \mathbf{Y}$ is multivariate normal exchangeable in $\mathbf{Y}$. This assumption implies the parametric structure defined by (2.1). The correlations and linear regressions based on the covariance structureof $(\mathrm{X}, \mathcal{Y})$ show that X is equicorrelated (η) with the components of $\mathcal{Y}$, the regression lines of each component of $\mathcal{Y}$ on X are linear and parallel, there is a positive dependence (θ) between the components of $\mathcal{Y}$ as well as between the components of $\mathcal{Y}$ conditional on X. Moreover, the maximum likelihood estimates of η and θ are asymptotically independent when X and $\mathbf{Y}$ are independent.

Appendix. Theoretical Results and Proofs

Let $\mathcal{Y}$ denote the vector of order statistics associated with the random vector $\mathbf{Y}$. The result of importance to this study is the fact that, under the model (2.1), $\mathrm{Cov}(\mathrm{X}, \mathbf{Y}) = \mathrm{Cov}(\mathrm{X}, \mathcal{Y})$. This fact is surprising, and is intimately tied to a property of the normal distribution. When $\mathrm{X}, \mathbf{Y}$ are independent and identically distributed, the result is due to Govindarajulu (1966). In a recent paper, Siegel (1993) proves that if $\mathbf{Y}$ has a p-variate normal distribution, $\mathrm{Z} = \min\{\mathrm{Y}_1, \ldots, \mathrm{Y}_\mathrm{p}\}$, then

$$\mathrm{Cov}(\mathrm{Y}_1, \mathrm{Z}) = \sum_{\mathrm{i}=1}^{\mathrm{p}} \mathrm{Cov}(\mathrm{Y}_1, \mathrm{Y}_\mathrm{i})\mathrm{P}[\mathrm{Y}_\mathrm{i} = \mathrm{Z}].$$

In an unpublished paper, Rinott and Samuel-Cahn (1993) extend this result to the r-th order statistic [see also Anderson (1993)]. Our proof differs from these and provides some additional insight in why the result holds. It also provides as extension to elliptically contoured distributions.

To motivate our proof, consider the case $\mathrm{p} = 2$. Because $\mathrm{Y}_{(2)} = \frac{1}{2} \mid \mathrm{Y}_1 - \mathrm{Y}_2 \mid + \frac{1}{2}(\mathrm{Y}_1 + \mathrm{Y}_2)$, we have,

$$\begin{aligned} \mathrm{cov}(\mathrm{X}, \mathrm{Y}_{(2)}) - \mathrm{cov}(\mathrm{X}, \mathrm{Y}_2) = \frac{1}{2}\mathrm{cov}(\mathrm{X}, \mid \mathrm{Y}_1 - \mathrm{Y}_2 \mid) &= \int_{\mathrm{y}_2 \le \mathrm{y}_1} (\mathrm{x} - \mu_0)(\mathrm{y}_1 - \mathrm{y}_2)\mathrm{dP} \\ &= \int_{\mathrm{x}} (\mathrm{x} - \mu_0)\mathrm{R}(\mathrm{x})\mathrm{dP}_\mathrm{X}, \end{aligned}$$

where the next to last equality follows from the fact that the distribution $\mathrm{P}(\mathrm{x}, \mathbf{y})$ is symmetric in $\mathbf{y}$ and the last equality from defining

$$\mathrm{R}(\mathrm{x}) = \int_{\mathrm{y}_2 \le \mathrm{y}_1} (\mathrm{y}_1 - \mathrm{y}_2)\mathrm{dP}_{\mathrm{Y}|\mathrm{x}}.$$

However, under the assumption of an exchangeable bivariate normal distribution, the expected conditional range $\mathrm{R}(\mathrm{x})$ of $\mathbf{Y}$ given x is constant in x and can be factored out of the integral, showing that $\mathrm{cov}(\mathrm{X},\mathrm{Y}_{(2)}) = \mathrm{cov}(\mathrm{X},\mathrm{Y}_2)$.

Theorem 1. If the distributuion of $\mathrm{X},\mathrm{Y}_1,\ldots,\mathrm{Y}_\mathrm{p}$ is multivariate normal jointly exchangeable in $\mathbf{Y}$ then $\mathrm{Cov}(\mathrm{X},\mathcal{Y}) = \mathrm{Cov}(\mathrm{X},\mathbf{Y})$.

Proof: Fix $1 \le \mathrm{r} \le \mathrm{p}$ and let $\mathbf{Y} \in \mathrm{E_j}$ indicate the event $\mathrm{Y}_{(\mathrm{r})} = \mathrm{Y_j}$, $\mathrm{j} = 1,\ldots,\mathrm{p}$. Then

$$\mathrm{Cov}(\mathrm{X},\mathrm{Y}_{(\mathrm{r})}) = \sum_{\mathrm{j}=1}^{\mathrm{p}} \mathrm{Cov}(\mathrm{X},\mathrm{Y_j} \mid \mathrm{E_j})\mathrm{P_y}(\mathrm{E_j}), \quad \mathrm{Cov}(\mathrm{X},\mathrm{Y_r}) = \sum_{\mathrm{j}=1}^{\mathrm{p}} \mathrm{Cov}(\mathrm{X},\mathrm{Y_r} \mid \mathrm{E_j})\mathrm{P_y}(\mathrm{E_j}),$$

and to show that $\mathrm{Cov}(\mathrm{X},\mathrm{Y}_{(\mathrm{r})}) = \mathrm{Cov}(\mathrm{X},\mathrm{Y_r})$ is sufficient to show that $\mathrm{Cov}(\mathrm{X},\mathrm{Y_j}-\mathrm{Y_r} \mid \mathrm{E_j}) = 0$, $\mathrm{j} = 1,\ldots,\mathrm{p}$. The assumption of a multivariate normal distribution for $(\mathrm{x},\mathbf{y})$ jointly exchangeable in $\mathbf{y}$ is a sufficient condition to obtain the result. In fact, note that the conditional distribution of $\mathbf{Y}$ given $\mathrm{X} = \mathrm{x}$ is exchangeable p-variate normal with a common mean $\mu(\mathrm{x})$ and covariance matrix $\Sigma_{11|0}$ independent of X. In addition, the events $\mathbf{Y} \in \mathrm{E_j}$ and $\mathrm{Y}_{(\mathrm{r})} - \mu(\mathrm{x}) = \mathrm{Y_j} - \mu(\mathrm{x})$ are equivalent. Consequently,

$$\mathrm{R_j}(\mathrm{x}) = \int_{\mathrm{E_j}} (\mathrm{y_r}-\mathrm{y_j})\mathrm{dP}_{\mathrm{Y|x}} = \int_{\mathrm{E_j}} [(\mathrm{y_r}-\mu(\mathrm{x})) - (\mathrm{y_j}-\mu(\mathrm{x}))]\mathrm{dP}_{\mathrm{Y|x}} \tag{A.1}$$

exists and is constant in x. Therefore,

$$\mathrm{Cov}(\mathrm{X},\mathrm{Y_r}-\mathrm{Y_j} \mid \mathrm{E_j}) = \frac{1}{\mathrm{P}(\mathrm{E_j})}\int_{\mathrm{X}} (\mathrm{x}-\mu_0)\mathrm{R_j}(\mathrm{x})\mathrm{dP_X} = 0, \tag{A.2}$$

concluding the proof.

Note that the normality assumption is used to guarantee that the conditional covariance matrix of $\mathbf{Y} \mid \mathrm{X} = \mathrm{x}$ is constant in x. This is not necessary and in fact restricts the extension to elliptically contoured distributions (within this class only the normal distribution has a conditional covariance matrix constant in x).

Theorem 2. If the distributuion of $\mathrm{X},\mathrm{Y}_1,\ldots,\mathrm{Y}_\mathrm{p}$ is elliptically contoured jointly exchangeable in $\mathbf{Y}$ then $\mathrm{Cov}(\mathrm{X},\mathcal{Y}) = \mathrm{Cov}(\mathrm{X},\mathbf{Y})$.

Proof: If $\mathrm{g}[(\mathrm{x}-\mu_0,\mathbf{y}-\mu_1\mathbf{e})'\Lambda(\mathrm{x}-\mu_0,\mathbf{y}-\mu_1\mathbf{e})]$ represents an elliptically contoured density, the invariance of the joint distribution of $(\mathrm{x},\mathbf{y})$ under permutation of $\mathbf{y}$ and an expansion of $\Lambda = \Sigma^{-1}$ shows that the density can be expressed as

$$\mathrm{g}(\frac{(\mathrm{x}-\mu_0)^2}{\sigma^2} + \mathbf{z}'\Lambda_{11.2}^{-1}\mathbf{z}),$$

where $\mathbf{z} = \mathbf{y} - \mu_1\mathbf{e} - \mathrm{x}\Lambda_{22}^{-1}\Lambda_{21}$. If $\mathbf{q} = \Lambda_{22}^{-1}\Lambda_{21}$, the events $\mathbf{Y} \in \mathrm{E_j}$ and $\mathrm{Y}_{(\mathrm{r})} - \mu_1 - \mathrm{xq_j} = \mathrm{Y_j} - \mu_1 - \mathrm{xq_j}$ are equivalent and the integral in (A.2) can be expressed as

$$\int_{\mathrm{E_j}}\int_{\mathrm{X}} (\mathrm{x}-\mu_0)(\mathrm{z_r}-\mathrm{z_j})\mathrm{g}(\frac{(\mathrm{x}-\mu_0)^2}{\sigma^2} + \mathbf{z}'\Lambda_{11.2}^{-1}\mathbf{z})\mathrm{dxd}\mathbf{z},$$

which is equal to zero by virtue of the fact that g is an even function of $\mathrm{x}-\mu_0$, concluding the proof. For more details about elliptically contoured distributions see (Tong 1990, pg. 62).

Theorem 3. If $\mathrm{Y}_1,\ldots,\mathrm{Y}_\mathrm{p}$ has a p-variate normal distribution with common mean ν, common variance τ^2 and common correlation ρ, then the covariance matrix of the order statistics $\mathcal{Y} = (\mathrm{Y}_{(1)},\ldots,\mathrm{Y}_{(\mathrm{p})})$ is

$$\mathrm{Cov}(\mathcal{Y}) = \tau^2[\rho\mathbf{e}\mathbf{e}^{'} + (1-\rho)\mathcal{C}],$$

where $\mathbf{e}' = (1,\ldots,1)$ and $\mathcal{C}$ is the covariance matrix of the order statistics of p independent standard normal random variables.

Proof: First note (David 1981, pg. 108) that the distribution of $\mathbf{Y}$ is the distribution of $\nu\mathbf{e} + \tau\mathbf{T}\mathbf{U}$, where $\mathbf{U} = (\mathrm{U}_0, \mathrm{U}_1, \ldots, \mathrm{U}_\mathrm{p}) \sim \mathrm{N}_{\mathrm{p}+1}(\mathbf{0}, \mathbf{Q})$, $\mathbf{T}$ is the $\mathrm{p} \times (\mathrm{p}+1)$ matrix with first column $\sqrt{|\,\rho\,|}\,\mathbf{e}$ and remaining entries equal to $\sqrt{1-\rho}\ \mathbf{I}$, and

$$\mathbf{Q} = \mathbf{I}, \quad \text{for } \rho > 0, \quad \mathbf{Q} = \begin{bmatrix} 1 & \frac{-\sqrt{-\rho}}{\sqrt{1-\rho}}\mathbf{e}' \\ \frac{-\sqrt{-\rho}}{\sqrt{1-\rho}}\mathbf{e} & \mathbf{I} \end{bmatrix}, \quad \text{for } \rho < 0.$$

Since $\mathrm{E}[\mathbf{Y}] = \nu\mathbf{e}$ is symmetric and consequently preserves the rankings of $\mathbf{U}_1$, (this is essential to the argument), the distribution of $\mathcal{Y}$ is the distribution of

$$\nu\mathbf{e} + \tau\mathbf{T}[\mathrm{U}_0, \mathcal{U}_1], \tag{A.3}$$

where $\mathcal{U}_1$ is the ordered vector of p independent standard normal variables. In addition, because the joint distribution of $(\mathrm{U}_0, \mathrm{U}_1, \ldots, \mathrm{U}_\mathrm{p})$ satisfies the assumptions of Theorem 1, the covariance matrix Ψ for $(\mathrm{U}_0, \mathcal{U}_1)$ is

$$\Psi = \begin{bmatrix} 1 & \mathbf{0}' \\ \mathbf{0} & \mathcal{C} \end{bmatrix}, \quad \text{for } \rho > 0, \quad \Psi = \begin{bmatrix} 1 & \frac{-\sqrt{-\rho}}{\sqrt{1-\rho}}\mathbf{e}' \\ \frac{-\sqrt{-\rho}}{\sqrt{1-\rho}}\mathbf{e} & \mathcal{C} \end{bmatrix}, \quad \text{for } \rho < 0,$$

where $\mathcal{C} = \mathrm{Cov}(\mathcal{U}_1)$. A direct computation shows that $\mathrm{Cov}(\mathcal{Y}) = \tau^2\mathbf{T}\Psi\mathbf{T}' = \tau^2[\rho\mathbf{e}\mathbf{e}' + (1-\rho)\mathcal{C}]$, conluding the proof.

Linear regressions. The best linear predictors between $\mathcal{Y}$ and X are completely determined by Ψ, E(X) and $\mathrm{E}(\mathcal{Y})$. The linear regression of $\mathcal{Y}$ on X is obtained from (A.3), which shows that $\mathrm{E}[\mathcal{Y}] = \mu\mathbf{e} + \tau\sqrt{1-\rho}\mathbf{c}$, where $\mathbf{c}$ is the expected value of p ordered independent standard normal variables, and from the fact that $\Psi_{10}\Psi_{00}^{-1} = (\gamma\tau/\sigma)\mathbf{e}$. The corresponding mean-squared error matrix can be expressed as

$$\Psi_{11|0} = \tau^2(1-\gamma^2)[\rho_{11|0}\mathbf{e}\mathbf{e}' + (1-\rho_{11|0})\mathcal{C}], \quad \rho_{11|0} = \frac{\rho - \gamma^2}{1-\gamma^2},$$

whereas the (component) multiple correlation coefficient is equal to η^2, the squared correlation between the components of $\mathcal{Y}$ and X.

Note that the slope of $\mathcal{Y}$ on X coincides with the slope of $\mathbf{Y}$ on X, the intercept parameters differ by the vertical distance $2\mathrm{c}_2\tau\sqrt{1-\rho}$ between the two lines, whereas the overall mean-squared error ratio is

$$\frac{\mathrm{tr}\ \Psi_{11|0}}{\mathrm{tr}\ \Sigma_{11|0}} = \rho + (1-\rho)\mathrm{tr}\ \frac{\mathcal{C}}{\mathrm{p}} \ \leq\ 1.$$

The best linear predictor $\mu_0 + \Psi_{01}\Psi_{11}^{-1}[\mathcal{Y} - \mathrm{E}(\mathcal{Y})]$ of X on $\mathcal{Y}$ is determined by

$$\begin{aligned} \Psi_{01}\Psi_{11}^{-1} &= \gamma\tau\sigma\mathbf{e}'\{\tau^2[\rho\mathbf{e}\mathbf{e}' + (1-\rho)\mathcal{C}]\}^{-1} \\ &= \frac{\gamma\sigma\mathbf{e}'}{\tau(1-\rho)}[\mathcal{C}^{-1} - \frac{\rho(\mathbf{e}'\mathcal{C}^{-1}\mathbf{e})\mathcal{C}^{-1}}{1+\rho[(\mathbf{e}'\mathcal{C}^{-1}\mathbf{e})-1]}] = \frac{\gamma\sigma}{\tau[1+\rho(\mathbf{e}'\mathcal{C}^{-1}\mathbf{e}-1)]}\mathbf{e}'\mathcal{C}^{-1}. \end{aligned} \tag{A.4}$$

Because $\mathcal{C}$ is stochastic, $\mathcal{C}\mathbf{e} = \mathbf{e}$, $\mathcal{C}^{-1}\mathbf{e} = \mathbf{e}$ and $\mathbf{e}'\mathcal{C}\mathbf{e} = \mathrm{p}$, so that

$$\Psi_{01}\Psi_{11}^{-1} = \frac{\gamma\sigma}{\tau[1+(\mathrm{p}-1)\rho]}\mathbf{e}'.$$

The intercept parameter is

$$\begin{aligned} \mu_0 - \Psi_{01}\Psi_{11}^{-1}\mathrm{E}(\mathcal{Y}) &= \mu_0 - \frac{\gamma\sigma}{\tau[1+((\mathbf{e}'\mathcal{C}^{-1}\mathbf{e})-1)\rho]}\mathbf{e}'\mathcal{C}^{-1}\ [\mu_1\mathbf{e} + \tau\sqrt{1-\rho}\ \mathbf{c}] \\ &= \mu_0 - \frac{\gamma\sigma}{\tau[1+(\mathrm{p}-1)\rho]}\mathrm{p}\mu_1, \end{aligned} \tag{A.5}$$

because of the fact that $\mathbf{e}'\mathbf{c} = 0$. The corresponding mean-squared error is

$$\Psi_{00} - \Psi_{01}\Psi_{11}^{-1}\Psi_{10} = \sigma^2[1 - \frac{\mathrm{p}\gamma^2}{1 + (\mathrm{p}-1)\rho}].$$

Note that $\Psi_{01}\Psi_{11}^{-1} = \Sigma_{01}\Sigma_{11}^{-1}$, so that the slope coefficients of X on $\mathbf{Y}$ and of X on $\mathcal{Y}$ coincide. Similarly, the mean-squared error and the intercept parameter also coincide. That is, the two regression equations are precisely the same. However, the linear regression of X on a subvector $\mathcal{Y}_0$ of $\mathcal{Y}$ with $1 \leq \mathrm{q} < \mathrm{p}$ components is specified by equations (A.4) and (A.5) with $\mathcal{C}$ and $\mathbf{c}$ replaced by the corresponding submatrices $\mathcal{C}_0$ and $\mathbf{c}_0$ associated with the q selected order statistics, and with $\mathbf{e}$ adjusted to the same number q of components. The mean-squared error can be expressed as

$$\sigma^2[1 - \frac{\mathrm{p}_0\gamma^2}{1 + (\mathrm{p}_0 - 1)\rho}], \quad \mathrm{p}_0 = \mathbf{e}'\mathcal{C}_0^{-1}\mathbf{e},$$

whereas the multiple correlation coefficient is determined by $\gamma^2/[1 + (\mathrm{p}_0 - 1)\rho]$.

Finally, we consider the best linear prediction of one extreme of $\mathbf{Y}$ based on X and the other extreme of $\mathbf{Y}$. The appropriate partitioning of Ψ shows that the two linear regressions

$$\begin{aligned} \mathrm{Y}_{(1)} &= \mathrm{b}_{1|2} + \mathrm{b}_1\mathrm{Y}_{(2)} + \mathrm{b}_2\mathrm{X}, \\ \mathrm{Y}_{(2)} &= \mathrm{b}_{2|1} + \mathrm{b}_1\mathrm{Y}_{(1)} + \mathrm{b}_2\mathrm{X}, \end{aligned}$$

are defined by parallel planes in which the coefficient b_1 is the partial correlation $\theta_{12|0}$ given by (2.9), the coefficient of X is

$$\mathrm{b}_2 = \frac{\tau\gamma(1-\rho)(\mathrm{c}_{22} - \mathrm{c}_{11})}{\sigma[\rho + (1-\rho)\mathrm{c}_{22} - \gamma^2]}, \tag{A.6}$$

whereas the intercept coefficients are, respectively,

$$\begin{aligned} \mathrm{b}_{1|2} &= \mu_1(1 - \mathrm{b}_1) - \mathrm{b}_2\mu_0 - (\mathrm{c}_2 - \mathrm{b}_1\mathrm{c}_1)\tau\sqrt{1-\rho}, \\ \mathrm{b}_{2|1} &= \mu_1(1 - \mathrm{b}_1) - \mathrm{b}_{\mathrm{i}2}\mu_0 + (\mathrm{c}_2 - \mathrm{b}_1\mathrm{c}_1)\tau\sqrt{1-\rho}, \end{aligned} \tag{A.7}$$

which are at a vertical destance $2\mathrm{c}_2\tau\sqrt{1-\rho}\,\mathrm{i}(1+\mathrm{b}_1)$ one from the other. In addition, the model mean-squared error and corresponding multiple correlation coefficient R^2 can be expressed as

$$\text{m.s.e.} = \frac{\tau^2(1-\rho^2)(\mathrm{c}_{22} - \mathrm{c}_{12})}{\rho + (1-\rho)\mathrm{c}_{22} - \gamma^2}, \qquad \mathrm{R}^2 = 1 - \frac{\text{m.s.e.}}{\tau^2[\rho + (1-\rho)\mathrm{c}_{22}]}.$$

References

Aburn, N. (1990), 'Eye injuries in indoor cricket at wellington hospital: A survey january 1987 to june 1989', *New Zealand Medical Journal* **103**(898), 454–6.

Anderson, C. L. (1993), 'Extension of surprising covariances', *Journal of the American Statistical Association - Letters to the Editor* **88**(424), 1478.

Anderson, T. W. (1985), *An Introduction to Multivariate Statistical Analysis*, second edn, John Wiley, New York.

Beyer, W. (1990), *Standard Probability and Statistics - Tables and Formulae*, CRC Press, Boca Raton.

David, H. A. (1981), *Order Statistics*, Wiley, New York.

Fishman, G. A., Baca, W., Alexander, K. R., Derlacki, D. J., Glenn, A. M. and Viana, M. A. G. (1993), 'Visual acuity in patients with best vitelliform macular dystrophy', *Ophthalmology* **100**(11), 1665–70.

Frenkel, R. and Shin, D. (1986), 'Prevention and management of delayed suprachoroidal hemorrhage after filtration surgery', *Archives of Ophthalmology* **104**(10), 1459–63.

Govindarajulu, Z. (1966), 'Characterization of normal and generalized truncated normal distributions using order statistics', *The Annals of Mathematical Statistics* **37**, 1011–1015.

Jr., G. K. (1985), 'Knife versus neodymium:yag laser posterior capsulotomy: A one-year follow-up', *Journal - American Intra-Ocular Implant Society* **11**(5), 448–55.

Parving, A. and Christensen, B. (1990), 'Hearing of the mentally retarded living at home', *Ugeskrift For Laeger* **152**(43), 3161–4.

Rehak, J. and Vymazal, M. (1989), 'Treatment of branch retinal vein occlusion with argon laser photocoagulation', *Acta Universitatis Palackianae Olomucensis Facultatis Medicae* **123**, 231–6.

Rinott, Y. and Samuel-Cahn, E. (1993), 'Covariance between variables and their order statistics for multivariate normal variables', *Unpublished*.

Siegel, A. F. (1993), 'A surprising covariance involving the minimum of multivariate normal variables', *Journal of the American Statistical Association* **88**(421), 77–80.

Szlyk, J. P., Fishman, G. A., Sovering, K., Alexander, K. R. and Viana, M. A. G. (1993), 'Evaluation of driving performance in patients with juvenile macular dystrophies', *Archives of Ophthalmology* **111**, 207–212.

Tong, Y. L. (1990), *The Multivariate Normal Distribution*, Springer-Verlag, New York.