Order Statistics - Handbook of Statistics, Vol. 16, Chapter 19, 1998.

LINEAR COMBINATIONS OF ORDERED SYMMETRIC OBSERVATIONS WITH APPLICATIONS TO VISUAL ACUITY

MARLOS VIANA

1. Introduction

In vision research, the Snellen chart is commonly used to assess visual acuity and is made up of letters of graduated sizes. By combining letter size and chart distance it is possible to determine the minimum visual angle of retinal resolution. A visual acuity of 20/30 means that at 20 feet the minimum angle of resolution is $\frac{30}{20}$-times the resolution standard (about 5 minutes of arc). The vision of a normal eye is recorded as 20/20 and corresponds to zero in the scale determined by the logarithm of the minimum angle of resolution (Log MAR). Normally, a single measure of visual acuity is made in each eye, say Y_1, Y_2, together with one or more covariates X, such as the subject's age, reading performance, physical condition, etc. Because smaller values of Log MAR correspond to better visual acuity, the extremes of visual acuity are defined in terms of the "best" acuity $Y_{(1)} = \min\{Y_1, Y_2\}$ and the "worst" acuity $Y_{(2)} = \max\{Y_1, Y_2\}$.

Ordered acuity measurements are also required to determine the person's total vision impairment, defined as

$$\text{Total Impairment} = \frac{3Y_{(1)} + Y_{(2)}}{4},$$

[e.g., Rubin, Munoz, Fried and West (1984)]. Consequently, there is interest in making inferences on the covariance structure

$$\Delta = \text{Cov}(X,\ w_1 Y_{(1)} + w_2 Y_{(2)}),$$

which includes the assessment of the correlation and linear predictors between X and linear combinations $w_1 Y_{(1)} + w_2 Y_{(2)}$ of the extreme acuity measurements. In particular, the correlations among average vision ($w_1 = w_2 = 0.5$), best vision ($w_1 = 1, w_2 = 0$), worst vision ($w_1 = 0, w_2 = 1$), acuity range ($w_1 = -1, w_2 = 1$), vision impairment ($w_1 = 0.75, w_2 = 0.25$) and one or more of the patient's conditions can be assessed.

Other applications of extreme bivariate measurements include the current 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)]; the assessment of defective hearing in mentally retarded adults based on the ear with best hearing (Parving and Christensen 1990); the predictive value of the worst vision following surgery in the eyes of glaucoma patients (Frenkel and Shin 1986); sports injury data on the reduction of best vision in damaged eyes (Aburn 1990);

Date: October 17, 2003.
1991 *Mathematics Subject Classification.* 62G30, 62J05.
Key words and phrases. Order Statistics, Correlation.

and the analysis of worst vision among patients treated for macular edema (Rehak and Vymazal 1989).

2. Models and Basic 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, it is assumed that the vector of means associated with (X, Y_1, Y_2) is given by $\mu^{'} = (\mu_0, \mu_1, \mu_1)$ and the corresponding covariance matrix Σ by

$$\Sigma = \begin{bmatrix} \sigma^2 & \gamma\sigma\tau & \gamma\sigma\tau \\ \gamma\sigma\tau & \tau^2 & \rho\tau^2 \\ \gamma\sigma\tau & \rho\tau^2 & \tau^2 \end{bmatrix}, \qquad \gamma^2 \le \frac{1+\rho}{2}, \quad \rho^2 \le 1, \tag{2.1}$$

where the range of the parameters is necessary and sufficient for Σ to be positive semi-definite. When there are p Y-values, the restriction is $\gamma^2 \le [1 + (p-1)\rho]/p$.

In general, the correlation ρ between Y_1 and Y_2 is in the interval $[-1, 1]$. However, in the present context in which Y_1 and Y_2 represent measurements on each eye, the correlation may be assumed to be non-negative.

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. In the bivariate case, because

$$Y_{(2)} = \frac{1}{2} \mid Y_1 - Y_2 \mid + \frac{1}{2}(Y_1 + Y_2),$$

it obtains that,

$$\begin{aligned} &\operatorname{cov}(X, Y_{(2)}) - \operatorname{cov}(X, Y_2) = \frac{1}{2}\operatorname{cov}(X, \mid Y_1 - Y_2 \mid) \\ &= \int_x \int_{y_2 \le y_1} (x - \mu_0)(y_1 - y_2) dP = \int_x (x - \mu_0) R(x) dP_X, \end{aligned}$$

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

$$R(x) = \int_{y_2 \le y_1} (y_1 - y_2) dP_{Y|x}.$$

However, under the assumption of an exchangeable bivariate normal distribution, the expected conditional range $R(x)$ of $\mathbf{Y}$ given x is constant in x and can be factored out of the integral, showing that $\operatorname{cov}(X, Y_{(2)}) = \operatorname{cov}(X, Y_2)$. This consequence of the exchangeability properties of the normality model described by (2.1) is discussed further in Olkin and Viana (1995), where the following propositions are proved (see also David (1996) and Viana and Olkin (1997)).

Proposition 2.1. If $Y_1, \ldots, Y_p$ are normally distributed with common mean ν, common variance τ^2 and common correlation ρ, then the covariance matrix of the order statistics $\mathcal{Y}' = (Y_{(1)}, \ldots, Y_{(p)})$ is

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

where $\mathcal{C}$ is the covariance matrix of the order statistics of p independent standard normal random variables.

Proposition 2.2. If the distribution of $X, Y_1, \ldots, Y_p$ is multivariate normal with $Y_1, \ldots, Y_p$ exchangeable and with (X, Y_i) and (X, Y_j) equally distributed, then $\mathrm{Cov}(X, \mathcal{Y}) = \mathrm{Cov}(X, \mathbf{Y})$.

As a consequence, the covariance matrix Δ of $(X, \mathbf{w}^{'}\mathcal{Y})$, where $\mathbf{w}$ indicates the column vector of real coefficients (w_1, w_2), is given by

$$\Delta = \mathrm{Cov}(X, \mathbf{w}^{'}\mathcal{Y}) = \begin{bmatrix} \sigma^2 & \gamma\sigma\tau\mathbf{w}\mathbf{e}^{'} \\ \gamma\sigma\tau\mathbf{w}^{'}\mathbf{e} & \tau^2[\rho(\mathbf{w}^{'}\mathbf{e})^2 + (1-\rho)\mathbf{w}^{'}\mathcal{C}\mathbf{w}] \end{bmatrix}. \tag{2.2}$$

In the bivariate case, the covariance matrix Ψ of $(X, Y_{(1)}, Y_{(2)})$ is

$$\Psi = \begin{bmatrix} \sigma^2 & \gamma\sigma\tau & \gamma\sigma\tau \\ \gamma\sigma\tau & \tau^2[\rho + (1-\rho)c_{11}] & \tau^2[\rho + (1-\rho)c_{12}] \\ \gamma\sigma\tau & \tau^2[\rho + (1-\rho)c_{21}] & \tau^2[\rho + (1-\rho)c_{22}] \end{bmatrix}, \tag{2.3}$$

where

$$\mathcal{C} = \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix} = \begin{bmatrix} 0.6817 & 0.3183 \\ 0.3183 & 0.6817 \end{bmatrix} \tag{2.4}$$

is the covariance matrix between the largest and the smallest of two independent standard normal variables (Beyer 1990, Table VII.2 p. 243). Also note that, when $p = 2$,

$$\mathbf{w}^{'}\mathcal{C}\mathbf{w} = (w_1^2 + w_2^2)c_{11} + 2w_1w_2c_{12}, \quad \mathbf{w}^{'}\mathbf{e} = w_1 + w_2.$$

3. Correlations and linear regressions

From (2.2), the correlation δ between X and a non-null linear combination $\mathbf{w}^{'}\mathcal{Y}$ of $\mathcal{Y}$ is,

$$\delta = \mathrm{Corr}(X, \mathbf{w}^{'}\mathcal{Y}) = \frac{\gamma\mathbf{w}^{'}\mathbf{e}}{\sqrt{\rho(\mathbf{w}^{'}\mathbf{e})^2 + (1-\rho)\mathbf{w}^{'}\mathcal{C}\mathbf{w}}}, \tag{3.1}$$

In addition,

$$\delta^2 \leq \frac{(1+\rho)/2}{\rho(\mathbf{w}^{'}\mathbf{e})^2 + (1-\rho)\mathbf{w}^{'}\mathcal{C}\mathbf{w}} \in [\frac{1}{2\mathbf{w}^{'}\mathcal{C}\mathbf{w}}, \frac{1}{(\mathbf{w}^{'}\mathbf{e})^2}],$$

holds for $\rho \in [0,1]$ and all $\mathbf{w} \in \mathbb{R}^2$ such that $\mathbf{w}^{'}\mathbf{e} \neq 0$.

Note that the correlation γ between X and Y_i is zero if [this is a direct consequence of Proposition 2.2] and only if the correlation between X and $Y_{(i)}$ is also zero. Therefore, $\delta = 0$ implies $\gamma = 0$ which implies (because of normality) that X and Y_i are independent. It then follows that X and $\mathbf{w}^{'}\mathcal{Y}$ are also independent. The fact that $\delta = 0$ implies the independence of X and $\mathbf{w}^{'}\mathcal{Y}$ is not obvious, because the joint distribution of $(X, \mathcal{Y})$ is no longer multivariate normal. When $\mathbf{w}^{'}\mathbf{e} = 0$ then $\delta = 0$. This implies that the correlation between X and the range of $\mathbf{Y}$ is necessarily zero under the equicorrelated-exchangeable model (2.1). Conversely, because of the normality assumption, $\delta = 0$ implies the independence of X and the range of $\mathbf{Y}$.

Proposition 3.1.

$$\underset{\mathbf{w}^{'}\mathbf{e}=1,\gamma>0,\rho>0}{\mathrm{Sup}} \mathrm{Corr}(X, \mathbf{w}^{'}\mathcal{Y}) = \frac{\gamma}{\sqrt{(1+\rho)/2}}.$$

The maximum value is obtained when $\mathbf{w}^{'} = (1/2, 1/2)$, in which case $\mathbf{w}^{'}\mathcal{Y}$ is the average of the components of $\mathbf{Y}$.

Proof. Write equation (3.1) as

$$\mathrm{Corr}(X, \mathbf{w}^{'}\mathcal{Y}) = \frac{\gamma}{\sqrt{\rho + (1-\rho)f}}, \quad f = \frac{\mathbf{w}^{'}\mathcal{C}\mathbf{w}}{(\mathbf{w}^{'}\mathbf{e})^2}.$$

Because $\mathcal{C}$ is positive definite, a solution $\mathbf{w}$ to the constrained minimization problem for $\mathbf{w}^{'}\mathcal{C}\mathbf{w}$, and equivalently for f, needs to satisfy $\mathcal{C}\mathbf{w} = \lambda\mathbf{e}$, where λ is a Lagrangian multiplier. The fact that $\mathcal{C}$ is stochastic shows that the unique constrained solution is $\mathbf{w}^{'} = (1/2, 1/2)$. □

The correlation θ between the extreme values $Y_{(1)}$ and $Y_{(2)}$ is

$$\theta = \mathrm{Corr}(Y_{(1)}, Y_{(2)}) = \frac{\rho + (1-\rho)c_{12}}{\rho + (1-\rho)c_{22}}. \tag{3.2}$$

For non-negative values of ρ, it holds that

$$0.4669 = \frac{c_{12}}{c_{22}} \le \theta \le 1,$$

whereas the partial correlation of $Y_{(1)}$ and $Y_{(2)}$ given X is

$$\theta_{12|0} = \mathrm{Corr}(Y_{(1)}, Y_{(2)} \mid X) = \frac{\rho + (1-\rho)c_{12} - \gamma^2}{\rho + (1-\rho)c_{22} - \gamma^2} \le \theta. \tag{3.3}$$

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

The minimum mean-squared error (m.s.e.) linear predictor of $\mathbf{w}^{'}\mathcal{Y}$ from X follows from the fact that

$$E[\mathbf{w}^{'}\mathcal{Y}] = \mu_1\mathbf{w}^{'}\mathbf{e} + \tau\sqrt{1-\rho}\ \mathbf{w}^{'}\mathbf{c},$$

where

$$\mathbf{c}^{'} = (c_1, c_2) = (-0.56419,\ \ 0.56419) \tag{3.4}$$

is the expected value of the smallest and largest of two independent standard normal variables [e.g., Beyer (1990)], and from the fact that $\Delta_{10}\Delta_{00}^{-1} = \gamma\tau\mathbf{w}^{'}\mathbf{e}/\sigma$. The resulting equation is

$$\mathbf{w}^{'}\mathcal{Y} = \mu_1\mathbf{w}^{'}\mathbf{e} + \tau\sqrt{1-\rho}\ \mathbf{w}^{'}\mathbf{c} + \frac{\tau}{\sigma}\gamma(x-\mu_0)\mathbf{w}^{'}\mathbf{e}. \tag{3.5}$$

The corresponding mean-squared error can be expressed as

$$\Delta_{11|0} = \tau^2(1-\gamma^2)[\rho_{\cdot|0}(\mathbf{w}^{'}\mathbf{e})^2 + (1-\rho_{\cdot|0})\mathbf{w}^{'}\mathcal{C}\mathbf{w}], \quad \rho_{\cdot|0} = \frac{\rho-\gamma^2}{1-\gamma^2}, \tag{3.6}$$

whereas the multiple correlation coefficient is equal to δ^2, the squared correlation between $\mathbf{w}^{'}\mathcal{Y}$ and X.

Similarly, the best (minimum m.s.e.) linear regression of X on $\mathbf{w}^{'}\mathcal{Y}$ is described by

$$x = \mu_0 + \frac{\sigma\mathbf{w}^{'}\mathbf{e}}{\tau[\rho(\mathbf{w}^{'}\mathbf{e})^2 + (1-\rho)\mathbf{w}^{'}\mathcal{C}\mathbf{w}]}\gamma\mathbf{w}^{'}(\mathcal{Y} - \mu_1\mathbf{e} - \tau\sqrt{1-\rho}\ \mathbf{c}), \tag{3.7}$$

with corresponding mean-squared error

$$\Delta_{00|1} = \sigma^2[1 - \frac{\gamma^2(\mathbf{w}^{'}\mathbf{e})^2}{\rho(\mathbf{w}^{'}\mathbf{e})^2 + (1-\rho)\mathbf{w}^{'}\mathcal{C}\mathbf{w}}]. \tag{3.8}$$

Also of interest are best linear predictors 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 regression equations

$$\begin{aligned} Y_{(1)} &= b_{1|2} + b_1 Y_{(2)} + b_2 X, \\ Y_{(2)} &= b_{2|1} + b_1 Y_{(1)} + b_2 X, \end{aligned}$$

are defined by parallel planes in which the coefficient b_1 is the partial correlation $\theta_{12|0}$ given by (3.3), the coefficient of X is

$$b_2 = \frac{\tau\gamma(1-\rho)(c_{22}-c_{12})}{\sigma[\rho + (1-\rho)c_{22} - \gamma^2]}, \tag{3.9}$$

whereas the intercept coefficients are, respectively,

$$\begin{aligned} b_{1|2} &= \mu_1(1-b_1) - b_2\mu_0 - (c_2 - b_1 c_1)\tau\sqrt{1-\rho}, \\ b_{2|1} &= \mu_1(1-b_1) - b_2\mu_0 + (c_2 - b_1 c_1)\tau\sqrt{1-\rho}. \end{aligned} \tag{3.10}$$

Because $c_1 = -c_2$, the vertical distance between the planes is

$$2c_2\tau\sqrt{1-\rho}\,(1+b_1).$$

In addition, the model mean-squared error and corresponding multiple correlation coefficient R^2 can be expressed as

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

4. Maximum likelihood and large-sample estimates

Given a sample $(x_\alpha, y_{1\alpha}, y_{2\alpha})$, $\alpha = 1, \dots, N$ of size N with means $\bar{x}$, $\bar{y}_i$, cross-product matrix $A = (a_{ij})$, $i, j = 0, 1, 2$, the maximum likelihood estimates of δ and θ are given by

$$\hat{\delta} = \frac{\hat{\gamma}\mathbf{w}'\mathbf{e}}{\sqrt{\hat{\rho}(\mathbf{w}'\mathbf{e})^2 + (1-\hat{\rho})\mathbf{w}'\mathcal{C}\mathbf{w}}}, \qquad \hat{\theta} = \frac{\hat{\rho} + (1-\hat{\rho})c_{12}}{\hat{\rho} + (1-\hat{\rho})c_{22}}, \tag{4.1}$$

where

$$\hat{\sigma}^2 = \frac{a_{00}}{N}, \quad \hat{\tau}^2 = \frac{1}{2N}(a_{11} + a_{22}),$$

and

$$\hat{\rho} = \frac{a_{12}}{\frac{1}{2}(a_{11} + a_{22})}, \qquad \hat{\gamma} = \frac{\frac{1}{2}(a_{01} + a_{02})}{\sqrt{\frac{1}{2}(a_{11} + a_{22})}\sqrt{a_{00}}}, \tag{4.2}$$

are the maximum likelihood estimates of σ^2, τ^2, ρ and γ based on

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

and

$$a_{ij} = \sum_{\alpha=1}^{N}(y_{i\alpha} - \bar{y}_i)(y_{j\alpha} - \bar{y}_j).$$

The delta method [e.g., Anderson (1985, p. 120)] shows that the asymptotic joint distribution of $\sqrt{N}\ (\hat{\delta}-\delta,\ \hat{\theta}-\theta)$ is normal with means zero, variances

$$\begin{aligned} \mathrm{AVar}(\hat{\delta}) &= [2\,\rho^2+6\,\rho^2\gamma^2 f+\gamma^2-5\,\rho^2\gamma^2-4\,\gamma^2\rho+4\,\gamma^4\rho+2\,\rho^3 \\ &+ +4f\rho-4\,\rho^3 f+2\,\rho^3 f^2-2\,\rho^2 f^2+2\,f^2-2\,\rho\, f^2 \\ &+ 4\,\gamma^4 f-6\,\gamma^2 f-4\,\gamma^4\rho\, f]/[-4\ (f+(1-f)\rho)^3], \end{aligned} \tag{4.3}$$

where $f=[\mathbf{w}'\mathcal{C}\mathbf{w}]/(\mathbf{w}'\mathbf{e})^2$, $\mathbf{w}'\mathbf{e}\neq 0$;

$$\mathrm{AVar}(\hat{\theta})=\frac{(c_{12}-c_{22})^2\,(1-\rho)^2\,(1+\rho)^2}{(c_{22}+(1-c_{22})\rho)^4}, \tag{4.4}$$

and covariance

$$\begin{aligned} \mathrm{ACov}(\hat{\delta},\hat{\theta}) =&[\big(-2\,\rho^2+3\,\rho^2 f+2\,\gamma^2\rho-2\,\gamma^2\rho\, f-\rho+2\,\gamma^2 f-3\,f+1\big) \\ &(c_{12}-c_{22})\,\gamma\ (1-\rho)] \\ &/[2\ (f+(1-f)\rho)^{3/2}\,(c_{22}+(1-c_{22})\rho)^2]. \end{aligned} \tag{4.5}$$

In particular, note that

$$\mathrm{ACov}(\hat{\delta},\ \hat{\theta}\mid \rho=0,\ \gamma=0)=\begin{bmatrix} \frac{1}{2\,f} & 0 \\ 0 & \frac{(c_{22}-c_{12})^2}{c_{22}^4} \end{bmatrix},$$

so that $\hat{\delta}$ and $\hat{\theta}$ are asymptotically independent when X, Y_1, Y_2 are jointly independent.

5. An exact test for $\gamma = 0$

As indicated earlier in Section 3, Proposition 2.2 implies that the following conditions are equivalent under the exchangeable multivariate normal model with covariance structure indicated by (2.1):

(1) $\gamma=\mathrm{Corr}(X,Y_i)=0$
(2) $\delta=\mathrm{Corr}(X,\mathbf{w}'\mathcal{Y})=0$
(3) X and $\mathbf{Y}$ are independent
(4) X and $\mathcal{Y}$ are independent
(5) X and $\mathbf{w}'\mathcal{Y}$ are independent

The hypothesis $\gamma=0$ can be assessed as follows. Let $A_{00}=a_{00}$, $A_{01}=(a_{0j}), j=1,\ldots,p$, $A_{11}=(a_{ij}), i,j=1,\ldots,p$, and $A_{10}=A_{01}'$. Further, let r denote the sample intraclass correlation coefficient

$$r=\frac{\sum_{i<j;\ i=1}^{p} a_{ij}/[p(p-1)/2]}{\sum_{i=1}^{p} a_{ii}/p}$$

associated with the sample $p\times p$ matrix of cross-products A_{11}. The distribution of A_{11} is Wishart $W_p(\tau^2[\rho\mathbf{e}\mathbf{e}'+(1-\rho)\mathbf{I}],n)$, $n=N-1$. Further, let $r_{\cdot|0}$ denote the sample intraclass correlation coefficient based on the conditional cross-product matrix

$$A_{11|0}=A_{11}-A_{10}A_{00}^{-1}A_{01}\ \sim\ W_p(\tau^2(1-\gamma^2)[\rho_{\cdot|0}\mathbf{e}\mathbf{e}'+(1-\rho_{\cdot|0})\mathbf{I}],n),$$

where $\rho_{\cdot|0} = (\rho - \gamma^2)/(1-\gamma^2)$. It follows [e.g., **?**)] that

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

Furthermore, 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

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

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

$$\frac{n}{p}\,\frac{V_1 - U_1}{V_1} = \frac{n}{p}[1 - \frac{(1+(p-1)r_{\cdot|0})\text{tr } S_{11|0}}{(1+(p-1)r)\text{tr } S_{11}}] \sim \text{F}_{p,n}. \tag{5.1}$$

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

$$(p-1)\frac{V_1}{V_2} = \frac{1+(p-1)r}{1-r} \sim \text{F}_{n,(p-1)n}, \tag{5.2}$$

so that (5.1) and (5.2) can be used to assess the corresponding hypotheses. Note that when γ is different from zero, larger values of (5.1) are expected; when ρ is positive larger values of (5.2) are expected. In the unrestricted case, smaller values are expected when ρ is negative.

6. Numerical Examples

The following statistics are based on $N = 42$ subjects participating in a larger experiment reported by Fishman et al. (1993), in which the evaluation of patients with Best's vitelliform macular dystrophy included the measurement of their bilateral visual acuity loss, denoted by (Y_1, Y_2), and age, denoted by X.

Because the visual acuity measurements Y_1, Y_2 in (respectively left and right) fellow eyes are expected to be about the same, to have about the same variability and to be equally correlated with age, the model defined in Section 1 to describe the data on (X, Y_1, Y_2) is used. The correlation structure between age and linear combinations of extreme visual acuity outcomes will be considered next.

The starting point is the sample means

$$(\bar{x}, \bar{y}_1, \bar{y}_2) = (28.833, 0.412, 0.437),$$

covariance matrix

$$S = \frac{A}{N-1} = \begin{bmatrix} 367.996 & 4.419 & 4.200 \\ 4.419 & 0.135 & 0.074 \\ 4.200 & 0.074 & 0.163 \end{bmatrix}$$

based on (X, Y_1, Y_2) data, and corresponding correlation matrix

$$R = \begin{bmatrix} 1.000 & 0.627 & 0.542 \\ 0.627 & 1.000 & 0.499 \\ 0.542 & 0.499 & 1.000 \end{bmatrix}.$$

Age is expressed in years and visual acuity measurements are expressed Log MAR units. The maximum likelihood estimate (4.2) of the correlation ρ between vision on fellow eyes is 0.496, whereas the estimated correlation $\hat{\gamma}$ between the patients's age and vision in either eye is 0.581. In addition, the estimated standard deviation $\hat{\tau}$ of vision in either eye is 0.386, the standard deviation for age is 19.182, the estimated mean vision and age are $\hat{\mu}_1 = 0.424$ and $\hat{\mu}_0 = 28.83$, respectively. The maximum likelihood estimate (4.1) of the correlation θ between extreme acuity outcomes is 0.782.

Table 1 summarizes the coefficients needed to estimate the correlation and linear regression parameters between X and a linear combination $\mathbf{w}'\mathcal{Y}$ of extreme acuities. From (4.1), (4.3) and (4.5), the corresponding estimates of $\hat{\delta}$, Avar($\hat{\delta}$), Acov($\hat{\delta}, \hat{\theta}$) and Avar($\hat{\delta} \mid \gamma = 0$) are shown in Table 2. The estimated large sample variance of $\hat{\theta}$, given by (4.4), is 0.6115.

The value of the test statistic (5.1) for $\gamma = 0$ is $F_{p,n} = 9.48$, which supports the conclusion of a non-null correlation γ between age and vision. Consequently, there is evidence to support the hypothesis of association between the patient's age and non-null linear combinations of extreme vision measures, such as those indicated in Table 1. Note that the range of vision acuity is necessarily independent of the patient's age under the equicorrelated-exchangeable model described by (2.1). The test statistic (5.2) for $\rho = 0$ is $F_{n,(p-1)n} = 2.97$ which also supports the claim of a positive correlation ρ between vision of fellow eyes.

The estimates of the regression lines (3.5) predicting the linear combination of extreme visual acuity from the patient's age and corresponding standard errors s.e. derived from (3.6) are shown in Table 3. Similarly, the estimates of the regression lines (3.7) predicting the patient's age from the linear combination of extreme visual acuity and corresponding standard errors s.e. obtained from (3.8) are shown in Table 4. A more realistic application, in this case, is the prediction of the subject's reading performance from a linear combination of extreme acuities, such as the subject's total visual impairment $\frac{3}{4}Y_{(1)} + \frac{1}{4}Y_{(2)}$, defined earlier in Section 1 [see also Rubin et al. (1984)].

Tables 5 and 6 show the corresponding minimum m.s.e. estimates for these models, obtained from sample means and cross-products of $(X, Y_{(1)}, Y_{(2)}$. These estimates will be contrasted with those obtained from data on (X, Y_1, Y_2). The *usual* estimates obtained from $(X, Y_{(1)}, Y_{(2)}$, although optimum in the m.s.e. sense, fail to carry over the multivariate normal assumption and properties.

Under these data, the covariance matrix (2.3) would be estimated by

$$\Psi_0 = \begin{bmatrix} 367.996 & 4.207 & 4.411 \\ 4.207 & 0.105 & 0.092 \\ 4.411 & 0.092 & 0.156 \end{bmatrix},$$

with resulting correlation matrix

$$\mathrm{Corr}_0(X, Y_{(1)}, Y_{(2)}) = \begin{bmatrix} 1.000 & 0.647 & 0.581 \\ 0.674 & 1.000 & 0.716 \\ 0.581 & 0.716 & 1.000 \end{bmatrix}.$$

In contrast, the corresponding maximum likelihood estimate obtained from Section 4 under the equicorrelated-exchangeable model is

$$\widehat{\mathrm{Corr}}(X, Y_{(1)}, Y_{(2)}) = \begin{bmatrix} 1.000 & 0.634 & 0.634 \\ 0.634 & 1.000 & 0.782 \\ 0.634 & 0.782 & 1.000 \end{bmatrix}.$$

The differences can be remarkable: for example, from Table 3, the estimated range of vision is 0.3092, whereas the unrestricted estimated value from Table 5 is 0.2494. The difference is numerically nearly equivalent to the difference between normal vision (Log MAR = 0) and a reduced vision of 20/40 (Log MAR = 0.3). The unrestricted model for best vision overestimates δ^2 by about 12% and underestimates it by about 21% for the worst vision. Tables 4 and 6 show that the expected ages corresponding to a normal best vision (the model's intercept) differ by about 8 years.

Proposition 3.1 is particularly important to justify the choice of the average vision against other convex linear combinations when the purpose is to obtain the best m.s.e. linear model relating X and the convex combination $\mathbf{w}'\mathcal{Y}$, under the equicorrelated-exchangeable model. Table 3 shows that the correlation between the subject's age and the average vision dominates the correlation with best vision, worst vision or visual impairment. This is a mathematical fact and not sampling variation.

Table 1. Linear combinations of extreme vision acuity.

$\mathbf{w}'$	$\mathbf{w}'\mathcal{Y}$	$\mathbf{w}'\mathbf{e}$	$\mathbf{w}'\mathcal{C}\mathbf{w}$	$\mathbf{w}'\mathbf{c}$
(0.5,0.5)	average vision	1	0.5	0
(1,0)	best vision	1	$c_{11} = 0.6817$	$c_1 = -0.5642$
(0,1)	worst vision	1	$c_{22} = 0.6817$	$c_2 = 0.5642$
(-1,1)	range	0	$2(c_{11} - c_{22}) = 0.7268$	$2c_2 = 1.1284$
(.75,.25)	visual impairment	1	0.5454	-0.2821

Table 2. Linear combinations of extreme vision acuity and corresponding estimates.

$\mathbf{w}'\mathcal{Y}$	$\hat{\delta}$	Avar($\hat{\delta}$)	Acov($\hat{\delta}, \hat{\theta}$)	Avar($\hat{\delta} \mid \gamma = 0$)
average vision	0.671	0.105	0.104	1.000
best vision	0.634	0.155	0.236	0.733
worst vision	0.634	0.155	0.236	0.733
visual impairment	0.661	0.117	0.151	0.916

Table 3. MLE Linear regression estimates of $\mathbf{w}'\mathcal{Y}$ on age.

$\mathbf{w}'\mathcal{Y}$	constant	coefficient	X	s.e.	r^2
average vision	0.0866	0.0117	age	0.247	0.450
best vision	-0.068	0.0117	age	0.273	0.401
worst vision	0.2412	0.0117	age	0.273	0.401
range of vision	0.3092	0	age	0.233	0
visual impairment	0.0093	0.0117	age	0.254	0.436

Table 4. MLE Linear regression estimates of age on $\mathbf{w}'\mathcal{Y}$.

X	constant	coefficient	$\mathbf{w}'\mathcal{Y}$	s.e.	r^2
age	12.464	38.599	average vision	14.222	0.450
age	8.932	34.389	best vision	14.834	0.401
age	19.565	34.389	worst vision	14.834	0.401
age	28.83	0	range of vision	19.181	0
age	15.845	37.4453	vision impairment	14.393	0.436

TABLE 5. $(x, y_{(1)}, y_{(2)})$-based linear regression estimates of $\mathbf{w}'\mathcal{Y}$ on age.

$\mathbf{w}'\mathcal{Y}$	constant	coefficient	X	s.e.	r^2
average vision	0.0876	0.0117	age	0.2506	0.4516
best vision	-0.037	0.0114	age	0.2431	0.4548
worst vision	0.2124	0.0119	age	0.3260	0.3376
range of vision	0.2494	0.0005	age	0.2823	0.0381
visual impairment	0.0253	0.0115	age	0.2365	0.4743

TABLE 6. $(x, y_{(1)}, y_{(2)})$-based linear regression estimates of age on $\mathbf{w}'\mathcal{Y}$.

X	constant	coefficient	$\mathbf{w}'\mathcal{Y}$	s.e.	r^2
age	12.427	38.568	average vision	14.381	0.4519
age	17.193	39.778	best vision	14.340	0.4548
age	13.114	28.165	worst vision	15.805	0.3376
age	28.134	2.621	range of vision	19.407	0.0381
age	14.118	40.990	vision impairment	14.080	0.4743

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, 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. (1996), 'A general representation of equally correlated variates', *Journal of the American Statistical Association* **91**(436), 1756.

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.

Olkin, I. and Viana, M. A. G. (1995), 'Correlation analysis of extreme observations from a multivariate normal distribution', *Journal of the American Statistical Association* **90**, 1373–1379.

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.

Rubin, G. S., Munoz, B., Fried, L. P. and West, S. (1984), 'Monocular vs binocular visual acuity as measures of vision impairment', *Vision Science and Its Applications - OSA Technical Digest Series* **1**, 328–331.

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.

Viana, M. A. G. and Olkin, I. (1997), Correlation analysis of ordered observations from a block-equicorrelated multivariate normal distribution, *in* S. Panchapakesan and N. Balakrishnan, eds, 'Advances in Statistical Decision Theory and Applications', Birkhauser, Boston, chapter 21, pp. 305–322.

The University of Illinois at Chicago