View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Lancaster E-Prints

Partial stochastic dominance for the multivariate Gaussian distribution

Amanda Turner[a,*], John Whitehead[a]

[a] *Department of Mathematics and Statistics, Lancaster University, Lancaster, LA1 4YF, United Kingdom*

Abstract

We establish a partial stochastic dominance result for the maximum of a multivariate Gaussian random vector with positive intraclass correlation coefficient and negative expectation. Specifically, we show that the distribution function intersects that of a standard Gaussian exactly once.

Keywords: Gaussian comparison inequalities, stochastic dominance, multivariate normal distribution, positive intraclass correlation coefficient

2010 MSC: Primary: 60E15, secondary: 62E17, 62G30

1. Introduction

Gaussian comparison inequalities provide a useful tool in probability and statistics, with applications in areas including Gaussian processes and extreme value theory. A survey of results and applications can be found in the books by Ledoux and Talagrand [1] and Lifshits [2]. Suppose that $\boldsymbol{X} = (X_1, \dots, X_k)$ and $\boldsymbol{Y} = (Y_1, \dots, Y_k)$ are two multivariate Gaussian vectors. Comparison inequalities typically involve finding conditions on the correlation structures of $\boldsymbol{X}$ and $\boldsymbol{Y}$ from which it can be deduced that $\mathbb{P}(\boldsymbol{X} \in C) \leq \mathbb{P}(\boldsymbol{Y} \in C)$ for some suitable class of sets $C \in \mathbb{R}^k$, usually of the form $\prod_{i=1}^{k}(-\infty, x_i]$. An important example is Slepian's inequality [3] which states that if $\mathbb{E}(\boldsymbol{X}) = \mathbb{E}(\boldsymbol{Y})$, $\mathbb{E}(X_i^2) = \mathbb{E}(Y_i^2)$ for all i and $\mathbb{E}(X_i X_j) \leq \mathbb{E}(Y_i Y_j)$ for all $i \neq j$, then $\mathbb{P}(X_1 \leq x_1, \dots, X_k \leq x_k) \leq$

*Corresponding author

Email address: `a.g.turner@lancaster.ac.uk` (Amanda Turner)

$\mathbb{P}(Y_1 \leq x_1, \ldots, Y_k \leq x_k)$ for all $(x_1, \ldots, x_k) \in \mathbb{R}^k$.

A direct consequence of Slepian's inequality is that $F_{X^*}(x) \leq F_{Y^*}(x)$ for all $x \in \mathbb{R}$, where $X^* = \max\{X_1, \ldots, X_k\}$ and $Y^* = \max\{Y_1, \ldots, Y_k\}$, so X^* stochastically dominates Y^* and the distribution functions of X^* and Y^* never cross each other. In this paper, by contrast, we obtain a *partial* stochastic dominance result by showing that, under certain assumptions on $\boldsymbol{X}$, $F_{X^*}(x)$ intersects the standard Gaussian distribution function $\Phi(x)$ *exactly once*. Suppose that $\boldsymbol{X}$ is a multivariate normal random vector with expectation $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_k)$ and all variances equal to 1. It is easy to see that if $\mu_i \geq 0$ for some i, then $F_{X^*}(x) < \Phi(x - \mu_i) \leq \Phi(x)$. Therefore, X^* dominates the standard Gaussian. Our result shows that when $\mu_i < 0$ for all i, if the covariances of $\boldsymbol{X}$ are equal and positive (that is, $\boldsymbol{X}$ has the intraclass correlation structure with positive correlation coefficient), then $F_{X^*}(x)$ intersects $\Phi(x)$ exactly once and this is from below. Therefore there exists some value $x_0 \in \mathbb{R}$ such that X^* dominates the standard Gaussian on the interval $(-\infty, x_0)$ but the standard Gaussian dominates X^* on (x_0, ∞).

Multivariate normal random vectors with the intraclass correlation structure occur in random effects models in which the error in a measurement arises as a combination of a class-specific error and an individual-specific error. More precisely, $X_i = \mu_i + \sqrt{\rho}Y_0 + \sqrt{1-\rho}Y_i$ for $i = 1, \ldots, k$, where $\rho \in (0, 1)$ and the $Y_0, \ldots, Y_k$ are independent standard normal random variables. Our motivation for this work was an application to the Bayesian design of exploratory clinical trials in which k experimental treatments are compared to a single control [4]. In that paper, one or more of the treatments is suitable to be developed further in a phase III trial if there is a sufficiently high probability that at least one treatment out-performs the control by a given threshold. Corollary 1 enables us to quantify the effect of increasing the threshold on that probability. This is then used to recommend an appropriate sample size for the trial.

The main results are stated in Section 2 and proved in Section 3. The proof is surprisingly long and technical, as well as being very sensitive to the assumptions. We are not aware of any simplifications to the argument, however, nor of other results in the literature that enable a comparison of this form.

2. Statement of results

In this section, we state our main theorem, which is then proved in Section 3. We also state and prove the corollary of this result that is used in [4].

We begin with some notation. For $\rho \in (0,1)$ and $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_k) \in \mathbb{R}^k$, let $(X_1, \ldots, X_k) \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ be a multivariate Gaussian random vector with $\Sigma_{ij} = \rho + (1-\rho)\delta_{ij}$, where δ_{ij} is the Kronecker delta. Let $X^* = \max\{X_1, \ldots, X_k\}$. For any random variable Y, we denote the distribution function of Y by F_Y and the density function of Y by f_Y. In the special case when $Y \sim N(0,1)$, we set $\Phi = F_Y$ and $\phi = f_Y$.

Theorem 1. *For any $\rho \in (0,1)$ and $\boldsymbol{\mu} \in (-\infty, 0)^k$, the distribution functions $F_{X^*}(x)$ and $\Phi(x)$ intersect exactly once. Furthermore, if $x_0 \in \mathbb{R}$ is the intersection point, then $f_{X^*}(x_0) > \phi(x_0)$.*

A direct consequence of this result is that

$$\begin{aligned} F_{X^*}(x) &> \Phi(x) \quad \text{for all } x > x_0; \\ F_{X^*}(x) &< \Phi(x) \quad \text{for all } x < x_0. \end{aligned}$$

Equivalently, if $Z \sim N(0,1)$, then the conditional distribution of $[X^* | X^* > x_0]$ is stochastically dominated by the conditional distribution of $[Z | Z > x_0]$ and the conditional distribution of $[X^* | X^* < x_0]$ stochastically dominates the conditional distribution of $[Z | Z < x_0]$.

Corollary 1. *For any $\rho \in (0,1)$ and $\boldsymbol{\mu} \in \mathbb{R}^k$, if $\mathbb{P}(X_i < 0$ for all $i = 1, \ldots, k) \geq \kappa$ for some $\kappa \in (0,1)$, then $\mathbb{P}(X_i < \Phi^{-1}(\zeta) - \Phi^{-1}(\kappa)$ for all $i = 1, \ldots, k) > \zeta$ for all $\zeta \in (\kappa, 1)$.*

Proof of Corollary 1. We have

$$\begin{aligned}\Phi(\Phi^{-1}(\kappa)) &\leq \mathbb{P}(X_i < 0 \text{ for all } i = 1, \ldots, k) \\ &= \mathbb{P}(X_i + \Phi^{-1}(\kappa) < \Phi^{-1}(\kappa) \text{ for all } i = 1, \ldots, k) \\ &= F_{Y^*}(\Phi^{-1}(\kappa)),\end{aligned}$$

where $Y^* = \max\{X_1 + \Phi^{-1}(\kappa), \ldots, X_k + \Phi^{-1}(\kappa)\}$. If $\zeta > \kappa$, then $\Phi^{-1}(\zeta) > \Phi^{-1}(\kappa)$ and so, applying Theorem 1 to $(X_1 + \Phi^{-1}(\kappa), \ldots, X_k + \Phi^{-1}(\kappa))$, gives $F_{Y^*}(\Phi^{-1}(\zeta)) > \Phi(\Phi^{-1}(\zeta)) = \zeta$, as required. (Since $F_{X^*}(x) < \Phi(x)$ for all $x \in \mathbb{R}$ when $\mu_i \geq 0$ for any i, the existence of κ guarantees that $\mathbb{E}(X_i + \Phi^{-1}(\kappa)) < 0$ for all i.)

3. Proof of main result

In this section we provide a proof of Theorem 1. We begin by expressing $F_{X^*}(x)$ in terms of independent identically distributed (i.i.d.) Gaussian random variables. We then show that the distribution functions intersect at most once provided that a quantity expressed in terms of these i.i.d. variables can be shown to be strictly positive. This quantity is obtained as the solution to a first order linear differential equation with variable linear coefficient. Positivity follows by showing that the linear coefficient is negative. The coefficient is expressed in terms of standard univariate Gaussian density functions which then enables us to deduce positivity as a consequence of properties of the inverse Mill's ratio. Finally, we show that the distribution functions must intersect at least once, by considering their relative values in the limit as $x \to \pm\infty$.

For $\rho \in (0, 1)$, $\nu_0 \in \mathbb{R}$ and $\boldsymbol{\nu} = (\nu_1, \ldots, \nu_k) \in \mathbb{R}^k$, let $Y_0, \ldots, Y_k$ be independent Gaussian random variables with $Y_0 \sim N(\nu_0, \rho)$ and $Y_i = N(\nu_i, 1 - \rho)$ for $i = 1, \ldots, k$. Let

$$G(\nu_0, \boldsymbol{\nu}) = \mathbb{P}(\max\{Y_1 - Y_0, \ldots, Y_k - Y_0\} < 0).$$

Observe that $(Y_1 - Y_0, \ldots, Y_k - Y_0) \sim N(\boldsymbol{\nu} - \nu_0, \boldsymbol{\Sigma})$ where $\boldsymbol{\Sigma}$ is as defined in

Section 2 and hence $F_{X^*}(x) = G(x, \boldsymbol{\mu}) = G(0, \boldsymbol{\mu} - x)$.

As $G(\nu_0, \boldsymbol{\nu})$ is strictly increasing in $\nu_0 \in (-\infty, \infty)$ from 0 to 1, there exists a unique function $g : (0,1) \times \mathbb{R}^k \to \mathbb{R}$ such that $G(g(\zeta, \boldsymbol{\nu}), \boldsymbol{\nu}) = \zeta$. So $F_{X^*}(x) = \Phi(x)$ for some $x \in \mathbb{R}$ if and only if $g(\zeta, \boldsymbol{\mu}) = \Phi^{-1}(\zeta)$ for some $\zeta \in (0,1)$. In order to show that there is at most one value of x for which $F_{X^*}(x) = \Phi(x)$, it is enough to show that $h(\zeta, \boldsymbol{\nu}) = \Phi^{-1}(\zeta) - g(\zeta, \boldsymbol{\nu})$ is strictly increasing in ζ or equivalently that

$$z(\zeta, \boldsymbol{\nu}) = \frac{\partial h}{\partial \zeta}(\zeta, \boldsymbol{\nu}) = \left(\phi(\Phi^{-1}(\zeta))\right)^{-1} - \frac{\partial g}{\partial \zeta}(\zeta, \boldsymbol{\nu}) > 0.$$

The second statement of the theorem follows directly from this result by the argument below. Since $G(g(\zeta, \boldsymbol{\nu}), \boldsymbol{\nu}) = \zeta$, differentiating with respect to ζ gives

$$\frac{\partial G}{\partial \nu_0}(g(\zeta, \boldsymbol{\nu}), \boldsymbol{\nu}) \frac{\partial g}{\partial \zeta}(\zeta, \boldsymbol{\nu}) = 1.$$

Suppose there exists some $x_0 \in \mathbb{R}$ such that $F_{X^*}(x_0) = \Phi(x_0)$. Let $\zeta_0 = \Phi(x_0)$ so $g(\zeta_0, \boldsymbol{\mu}) = x_0 = \Phi^{-1}(\zeta_0)$. Then

$$f_{X^*}(x_0) = \frac{\partial G}{\partial \nu_0}(x_0, \boldsymbol{\mu}) = \left(\frac{\partial g}{\partial \zeta}(\zeta_0, \boldsymbol{\mu})\right)^{-1} > \phi(\Phi^{-1}(\zeta_0)) = \phi(x_0)$$

as required.

By symmetry, it is sufficient to prove that $z(\zeta, \boldsymbol{\nu}) > 0$ in the case $\nu_1 \geq \cdots \geq \nu_k$. Now $G(\nu_0, \boldsymbol{\nu}) \to \mathbb{P}(X_1 \leq \nu_0) = \Phi(\nu_0 - \nu_1)$ as $\nu_k \leq \cdots \leq \nu_2 \to -\infty$. Since G and its derivatives are equicontinuous in all variables, it follows that $g(\zeta, \boldsymbol{\nu}) \to \nu_1 + \Phi^{-1}(\zeta)$ and $z(\zeta, \boldsymbol{\nu}) \to \frac{\partial \nu_1}{\partial \zeta} = 0$ as $\nu_k \leq \cdots \leq \nu_2 \to -\infty$. We abusively use the notation $f(\boldsymbol{\nu}^i)$ to denote $\lim_{\nu_k \leq \cdots \leq \nu_{i+1} \to -\infty} f(\boldsymbol{\nu})$, so $z(\zeta, \boldsymbol{\nu}^1) = 0$ for all values of $\zeta \in (0,1)$ and $\nu_1 \in \mathbb{R}$.

Recall that $G(g(\zeta, \boldsymbol{\nu}), \boldsymbol{\nu}) = \zeta$. Differentiating both sides with respect to ν_i, $i = 1, \ldots, k$, gives

$$\frac{\partial G}{\partial \nu_i}(g(\zeta, \boldsymbol{\nu}), \boldsymbol{\nu}) + \frac{\partial G}{\partial \nu_0}(g(\zeta, \boldsymbol{\nu}), \boldsymbol{\nu}) \frac{\partial g}{\partial \nu_i}(\zeta, \boldsymbol{\nu}) = 0$$

and hence

$$\frac{\partial g}{\partial \nu_i}(\zeta, \boldsymbol{\nu}) = -Q_i(g(\zeta, \boldsymbol{\nu}), \boldsymbol{\nu})$$

where

$$Q_i(\nu_0, \boldsymbol{\nu}) = \frac{\frac{\partial G}{\partial \nu_i}}{\frac{\partial G}{\partial \nu_0}}(\nu_0, \boldsymbol{\nu}).$$

It follows that

$$\begin{aligned}
\frac{\partial}{\partial \nu_i}\left(z(\zeta, \boldsymbol{\nu}) - \left(\phi(\Phi^{-1}(\zeta))\right)^{-1}\right) &= \frac{\partial}{\partial \zeta}\frac{\partial h}{\partial \nu_i}(\zeta, \boldsymbol{\nu}) \\
&= \frac{\partial}{\partial \zeta}\left(Q_i(g(\zeta, \boldsymbol{\nu}), \boldsymbol{\nu})\right) \\
&= \frac{\partial Q_i}{\partial \nu_0}(g(\zeta, \boldsymbol{\nu}), \boldsymbol{\nu})\frac{\partial g}{\partial \zeta}(\zeta, \boldsymbol{\nu}) \\
&= -\frac{\partial Q_i}{\partial \nu_0}(g(\zeta, \boldsymbol{\nu}), \boldsymbol{\nu})\left(z(\zeta, \boldsymbol{\nu}) - \left(\phi(\Phi^{-1}(\zeta))\right)^{-1}\right).
\end{aligned}$$

Therefore, for each $i = 1, \ldots, k$, if ν_j is fixed for all $j \neq i$, $z(\zeta, \boldsymbol{\nu}) - \left(\phi(\Phi^{-1}(\zeta))\right)^{-1}$ is the solution to a first order linear differential equation in ν_i and hence we can evaluate $z(\zeta, \boldsymbol{\nu}^i)$ inductively for $i = 2, \ldots, k$ by

$$z(\zeta, \boldsymbol{\nu}^i) - \left(\phi(\Phi^{-1}(\zeta))\right)^{-1} = \left(z(\zeta, \boldsymbol{\nu}^{i-1}) - \left(\phi(\Phi^{-1}(\zeta))\right)^{-1}\right)\exp\left(-\int_{-\infty}^{\nu_i} A_i(t)dt\right),$$

where $A_i(t)$ is the limit of

$$\frac{\partial Q_i}{\partial \nu_0}(g(\zeta, \nu_1, \ldots, \nu_{i-1}, t, \nu_{i+1}, \ldots, \nu_k), \nu_1, \ldots, \nu_{i-1}, t, \nu_{i+1}, \ldots, \nu_k)$$

as $\nu_k \leq \cdots \leq \nu_{i+1} \to -\infty$. Hence

$$\begin{aligned}
z(\zeta, \boldsymbol{\nu}^i) = z(\zeta, \boldsymbol{\nu}^{i-1}) &\exp\left(-\int_{-\infty}^{\nu_i} A_i(t)dt\right) \\
&+ \left(\phi(\Phi^{-1}(\zeta))\right)^{-1}\left(1 - \exp\left(-\int_{-\infty}^{\nu_i} A_i(t)dt\right)\right).
\end{aligned}$$

Suppose that it is true that $z(\zeta, \boldsymbol{\nu}^{i-1}) \geq 0$, for some $i \geq 2$. Then, if

$$\frac{\partial Q_i}{\partial \nu_0}(\nu_0, \boldsymbol{\nu}^i) \geq 0 \text{ for all } \nu_0 \in \mathbb{R} \text{ and } \nu_1 \geq \cdots \geq \nu_i, \tag{1}$$

with strict inequality unless $\nu_1 = \nu_2 = \cdots = \nu_i$, then it follows that $z(\zeta, \boldsymbol{\nu}^i) > 0$, and by induction the theorem will be proven.

We now show that (1) holds for $i = 1, \ldots, k$. Since

$$\frac{\partial Q_i}{\partial \nu_0} = \frac{\partial}{\partial \nu_0}\left(\frac{\frac{\partial G}{\partial \nu_i}}{\frac{\partial G}{\partial \nu_0}}\right) = \frac{\frac{\partial^2 G}{\partial \nu_0 \partial \nu_i}\frac{\partial G}{\partial \nu_0} - \frac{\partial^2 G}{\partial \nu_0^2}\frac{\partial G}{\partial \nu_i}}{\left(\frac{\partial G}{\partial \nu_0}\right)^2},$$

it is sufficient to show that

$$\Delta_i = \left(\frac{\partial^2 G}{\partial \nu_0 \partial \nu_i}\frac{\partial G}{\partial \nu_0} - \frac{\partial^2 G}{\partial \nu_0^2}\frac{\partial G}{\partial \nu_i}\right)(\nu_0, \boldsymbol{\nu}^i) \geq 0$$

for all $i = 1, \ldots, k$, $\nu_0 \in \mathbb{R}$ and $\nu_1 \geq \cdots \geq \nu_i$ with strict inequality unless $\nu_1 = \nu_2 = \cdots = \nu_i$.

Now

$$\begin{aligned} G(\nu_0, \boldsymbol{\nu}) &= \mathbb{P}(\max\{Y_1, \ldots, Y_k\} < Y_0) \\ &= \int_{-\infty}^{\infty} f_{Y_0}(t)\mathbb{P}(\max\{Y_1, \ldots, Y_k\} < t)dt \\ &= \frac{1}{\sqrt{\rho}}\int_{-\infty}^{\infty} \phi\left(\frac{t-\nu_0}{\sqrt{\rho}}\right)\prod_{j=1}^{k}\Phi\left(\frac{t-\nu_j}{\sqrt{1-\rho}}\right)dt. \end{aligned}$$

Hence

$$\begin{aligned} \frac{\partial G}{\partial \nu_0}(\nu_0, \boldsymbol{\nu}) =& -\frac{1}{\rho}\int_{-\infty}^{\infty} \phi'\left(\frac{t-\nu_0}{\sqrt{\rho}}\right)\prod_{j=1}^{k}\Phi\left(\frac{t-\nu_j}{\sqrt{1-\rho}}\right)dt \\ =& -\frac{1}{\sqrt{\rho}}\left[\phi\left(\frac{t-\nu_0}{\sqrt{\rho}}\right)\prod_{j=1}^{k}\Phi\left(\frac{t-\nu_j}{\sqrt{1-\rho}}\right)\right]_{-\infty}^{\infty} \\ &+ \frac{1}{\sqrt{\rho}}\int_{-\infty}^{\infty}\phi\left(\frac{t-\nu_0}{\sqrt{\rho}}\right)\frac{\partial}{\partial t}\left(\prod_{j=1}^{k}\Phi\left(\frac{t-\nu_j}{\sqrt{1-\rho}}\right)\right)dt \\ =& \frac{1}{\sqrt{\rho}}\int_{-\infty}^{\infty}\phi\left(\frac{t-\nu_0}{\sqrt{\rho}}\right)\frac{\partial}{\partial t}\left(\prod_{j=1}^{k}\Phi\left(\frac{t-\nu_j}{\sqrt{1-\rho}}\right)\right)dt, \end{aligned}$$

$$\frac{\partial G}{\partial \nu_i}(\nu_0, \boldsymbol{\nu}) = -\frac{1}{\sqrt{\rho(1-\rho)}} \int_{-\infty}^{\infty} \phi\left(\frac{t-\nu_0}{\sqrt{\rho}}\right) \phi\left(\frac{t-\nu_i}{\sqrt{1-\rho}}\right) \prod_{j \neq i} \Phi\left(\frac{t-\nu_j}{\sqrt{1-\rho}}\right) dt,$$

$$\frac{\partial^2 G}{\partial \nu_0 \partial \nu_i}(\nu_0, \boldsymbol{\nu}) = \frac{1}{\sqrt{\rho(1-\rho)}} \int_{-\infty}^{\infty} \phi'\left(\frac{t-\nu_0}{\sqrt{\rho}}\right) \phi\left(\frac{t-\nu_i}{\sqrt{1-\rho}}\right) \prod_{j \neq i} \Phi\left(\frac{t-\nu_j}{\sqrt{1-\rho}}\right) dt,$$

and

$$\frac{\partial^2 G}{\partial \nu_0^2}(\nu_0, \boldsymbol{\nu}) = \frac{1}{\rho} \int_{-\infty}^{\infty} \phi'\left(\frac{t-\nu_0}{\sqrt{\rho}}\right) \frac{\partial}{\partial t}\left(\prod_{j=1}^{k} \Phi\left(\frac{t-\nu_j}{\sqrt{1-\rho}}\right)\right) dt.$$

It follows that

$$\begin{aligned}
\Delta_i &= \frac{1}{\sqrt{\rho(1-\rho)}} \int_{-\infty}^{\infty} \phi'\left(\frac{s-\nu_0}{\sqrt{\rho}}\right) \phi\left(\frac{s-\nu_i}{\sqrt{1-\rho}}\right) \prod_{j \neq i} \Phi\left(\frac{s-\nu_j}{\sqrt{1-\rho}}\right) ds \\
&\quad \times \frac{1}{\sqrt{\rho}} \int_{-\infty}^{\infty} \phi\left(\frac{t-\nu_0}{\sqrt{\rho}}\right) \frac{\partial}{\partial t}\left(\prod_{j=1}^{k} \Phi\left(\frac{t-\nu_j}{\sqrt{1-\rho}}\right)\right) dt \\
&\quad - \frac{1}{\sqrt{\rho(1-\rho)}} \int_{-\infty}^{\infty} \phi\left(\frac{s-\nu_0}{\sqrt{\rho}}\right) \phi\left(\frac{s-\nu_i}{\sqrt{1-\rho}}\right) \prod_{j \neq i} \Phi\left(\frac{s-\nu_j}{\sqrt{1-\rho}}\right) ds \\
&\quad \times \frac{1}{\rho} \int_{-\infty}^{\infty} \phi'\left(\frac{t-\nu_0}{\sqrt{\rho}}\right) \frac{\partial}{\partial t}\left(\prod_{j=1}^{k} \Phi\left(\frac{t-\nu_j}{\sqrt{1-\rho}}\right)\right) dt \\
&= \frac{1}{\rho^2\sqrt{1-\rho}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left\{ (s-t)\phi\left(\frac{s-\nu_0}{\sqrt{\rho}}\right) \phi\left(\frac{t-\nu_0}{\sqrt{\rho}}\right) \phi\left(\frac{s-\nu_i}{\sqrt{1-\rho}}\right) \right. \\
&\quad \left. \times \prod_{j \neq i} \Phi\left(\frac{s-\nu_j}{\sqrt{1-\rho}}\right) \frac{\partial}{\partial t}\left(\prod_{j=1}^{k} \Phi\left(\frac{t-\nu_j}{\sqrt{1-\rho}}\right)\right) \right\} ds dt.
\end{aligned}$$

Interchanging s and t, we also have

$$\begin{aligned}
\Delta_i &= \frac{1}{\rho^2\sqrt{1-\rho}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left\{ (t-s)\phi\left(\frac{s-\nu_0}{\sqrt{\rho}}\right) \phi\left(\frac{t-\nu_0}{\sqrt{\rho}}\right) \phi\left(\frac{t-\nu_i}{\sqrt{1-\rho}}\right) \right. \\
&\quad \left. \times \prod_{j \neq i} \Phi\left(\frac{t-\nu_j}{\sqrt{1-\rho}}\right) \frac{\partial}{\partial s}\left(\prod_{j=1}^{k} \Phi\left(\frac{s-\nu_j}{\sqrt{1-\rho}}\right)\right) \right\} ds dt.
\end{aligned}$$

Hence, Δ_i can be expressed as the average of these two forms, giving

$$\Delta_i = \frac{1}{2\rho^2\sqrt{1-\rho}} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \left\{ \phi\left(\frac{s-\nu_0}{\sqrt{\rho}}\right) \phi\left(\frac{t-\nu_0}{\sqrt{\rho}}\right) \phi\left(\frac{s-\nu_i}{\sqrt{1-\rho}}\right) \phi\left(\frac{t-\nu_i}{\sqrt{1-\rho}}\right) \right.$$
$$\left. \times \prod_{j\neq i} \Phi\left(\frac{s-\nu_j}{\sqrt{1-\rho}}\right) \prod_{j\neq i} \Phi\left(\frac{t-\nu_j}{\sqrt{1-\rho}}\right) (s-t)\left(H_i(s,\boldsymbol{\nu}) - H_i(t,\boldsymbol{\nu})\right) \right\} dsdt,$$

where

$$H_i(t,\boldsymbol{\nu}) = \frac{\Phi\left(\frac{t-\nu_i}{\sqrt{1-\rho}}\right)}{\phi\left(\frac{t-\nu_i}{\sqrt{1-\rho}}\right)} \frac{\frac{\partial}{\partial t}\left(\prod_{j=1}^k \Phi\left(\frac{t-\nu_j}{\sqrt{1-\rho}}\right)\right)}{\prod_{j=1}^k \Phi\left(\frac{t-\nu_j}{\sqrt{1-\rho}}\right)}.$$

Now

$$(s-t)\left(H_i(s,\boldsymbol{\nu}) - H_i(t,\boldsymbol{\nu})\right) = (s-t)\int_t^s \frac{\partial H_i}{\partial u}(u,\boldsymbol{\nu})du$$

which is positive for all s and t if the integrand is positive for all u. In order to show that $\Delta_i \geq 0$, it is therefore sufficient to show that $\frac{\partial H_i}{\partial t}(t,\boldsymbol{\nu}^i) \geq 0$ for all $i = 1,\ldots,k$, and $\nu_1 \geq \cdots \geq \nu_i$ with strict inequality unless $\nu_1 = \nu_2 = \cdots = \nu_i$.

Let $m(x) = \phi(x)/\Phi(x)$ be the inverse Mill's ratio. Then

$$H_i(t,\boldsymbol{\nu}) = \frac{1}{\sqrt{1-\rho}} \sum_{j=1}^k \frac{m\left(\frac{t-\nu_j}{\sqrt{1-\rho}}\right)}{m\left(\frac{t-\nu_i}{\sqrt{1-\rho}}\right)}.$$

Using the fact that $m'(x) = -m(x)(x+m(x))$, we obtain

$$\begin{aligned}
\frac{\partial H_i}{\partial t}(t,\boldsymbol{\nu}^i) &= \frac{1}{1-\rho}\sum_{j=1}^k \left(\frac{m'\left(\frac{t-\nu_j}{\sqrt{1-\rho}}\right)}{m\left(\frac{t-\nu_i}{\sqrt{1-\rho}}\right)} - \frac{m\left(\frac{t-\nu_j}{\sqrt{1-\rho}}\right) m'\left(\frac{t-\nu_i}{\sqrt{1-\rho}}\right)}{m\left(\frac{t-\nu_j}{\sqrt{1-\rho}}\right)^2} \right) \\
&= \frac{1}{1-\rho}\sum_{j=1}^k \frac{m\left(\frac{t-\nu_j}{\sqrt{1-\rho}}\right)}{m\left(\frac{t-\nu_i}{\sqrt{1-\rho}}\right)} \left(\frac{\nu_j-\nu_i}{\sqrt{1-\rho}} + m\left(\frac{t-\nu_i}{\sqrt{1-\rho}}\right) - m\left(\frac{t-\nu_j}{\sqrt{1-\rho}}\right) \right) \\
&= \frac{1}{(1-\rho)^{3/2}}\sum_{j=1}^k \frac{m\left(\frac{t-\nu_j}{\sqrt{1-\rho}}\right)}{m\left(\frac{t-\nu_i}{\sqrt{1-\rho}}\right)} \int_{\nu_i}^{\nu_j} \left(1 + m'\left(\frac{t-u}{\sqrt{1-\rho}}\right)\right) du.
\end{aligned}$$

But $m'(x) > -1$ for all $x \in \mathbb{R}$ (see Sampford [5]) and hence, letting $\nu_k \leq \cdots \leq$

$\nu_{i+1} \to -\infty$,

$$\frac{\partial H_i}{\partial t}(t, \boldsymbol{\nu}^i) \geq 0$$

for all $i = 1, \ldots, k$, and $\nu_1 \geq \cdots \geq \nu_i$ with strict inequality unless $\nu_1 = \nu_2 = \cdots = \nu_i$, as required.

Finally, we show that the distribution functions must intersect at least once. Since $F_{X^*}(x)$ is strictly decreasing in each μ_i, it is sufficient to do this in the case when $\boldsymbol{\mu} = \mu\mathbf{1}$ for some $\mu < 0$ where $\mathbf{1} = (1, 1, \ldots, 1)$. To see this, set $\mu^+ = \max_i \mu_i$ and $\mu^- = \min_i \mu_i$ and let x^+ (respectively x^-) be the intersection point of $\Phi(x)$ and the distribution function corresponding to $\mu^+\mathbf{1}$ (respectively $\mu^-\mathbf{1}$). Then $F_{X^*}(x^-) \leq \Phi(x^-)$ and $F_{X^*}(x^+) \geq \Phi(x^+)$, and hence there must exist some $x_0 \in [x^-, x^+]$ for which $F_{X^*}(x_0) = \Phi(x_0)$. (A similar argument, where Slepian's inequality provides the required monotonicity, can be used to show that $F_{X^*}(x)$ intersects $\Phi(x)$ at least once when Σ is *any* correlation matrix with $\Sigma_{ii} = 1$ for all i and $\Sigma_{ij} \in [0, 1)$ for all $i \neq j$.)

Suppose from now on that $\boldsymbol{\mu} = \mu\mathbf{1}$ for some $\mu < 0$. By using the asymptotic expansions as $x \to -\infty$ of $\Phi(x)$ and $F_{X^*}(x)$ (see, for example, [6]), we have that $\Phi(x) \asymp x^{-1}\phi(x)$ and

$$F_{X^*}(x) \asymp (x - \mu)^{-k} \phi\left(\sqrt{\frac{k}{1 + (k-1)\rho}}(x - \mu)\right).$$

Since $k/(1 + (k-1)\rho) > 1$, $\lim_{x \to -\infty}(F_{X^*}(x)/\Phi(x)) = 0$ and hence $F_{X^*}(x) < \Phi(x)$ for x sufficiently large and negative.

By Slepian's inequality, $F_{X^*}(x) \geq \Phi(x - \mu)^k$. Therefore,

$$\begin{aligned}
\lim_{x \to \infty} \frac{1 - F_{X^*}(x)}{1 - \Phi(x)} &\leq \lim_{x \to \infty} \frac{1 - \Phi(x - \mu)^k}{1 - \Phi(x)} \\
&= \lim_{x \to \infty} \frac{k\phi(x - \mu)\Phi(x - \mu)^{k-1}}{\phi(x)} \\
&= k \exp(-\mu^2/2) \lim_{x \to \infty} \exp(\mu x) \\
&= 0.
\end{aligned}$$

The final equality is the only place in the proof where we have required negative expectations. It follows that $F_{X^*}(x) > \Phi(x)$ for x sufficiently large and positive, and so the distribution functions must intersect at least once.

Remark 1. In the case when $\boldsymbol{\mu} = \mu\mathbf{1}$ for some $\mu < 0$, $x_0 = \Phi^{-1}(\zeta_0)$ where $\zeta_0 \in (0,1)$ is the solution to the equation $\Phi^{-1}(\zeta_0) - g(\zeta_0, \mathbf{0}) = \mu$. Here $g(\zeta, \mathbf{0})$ is the equicoordinate quantile function of $N(\mathbf{0}, \Sigma)$, the standard k-dimensional multivariate Gaussian with all correlations equal to ρ. It follows from the proof above that $\Phi^{-1}(\zeta) - g(\zeta, \mathbf{0})$ is strictly increasing from $-\infty$ to 0. Numerical estimates for ζ_0, and hence for x_0, can be found relatively easily as a consequence of this monotonicity. For general $\boldsymbol{\mu}$, this approach can be used to find estimates for x^- and x^+ (defined above), which give the end-points of a finite interval in which to search for x_0.

References

[1] M. Ledoux, M. Talagrand, Probability in Banach spaces: Isoperimetry and processes, Vol. 23 of Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)], Springer-Verlag, Berlin, 1991. doi:10.1007/978-3-642-20212-4.
URL `http://dx.doi.org/10.1007/978-3-642-20212-4`

[2] M. A. Lifshits, Gaussian random functions, Vol. 322 of Mathematics and its Applications, Kluwer Academic Publishers, Dordrecht, 1995. doi:10.1007/978-94-015-8474-6.
URL `http://dx.doi.org/10.1007/978-94-015-8474-6`

[3] D. Slepian, The one-sided barrier problem for Gaussian noise, Bell System Tech. J. 41 (1962) 463–501.

[4] J. Whitehead, F. Cleary, A. Turner, Bayesian sample sizes for exploratory clinical trials comparing multiple experimental treatments with a control, Statistics in Medicine (to appear).
URL `http://arxiv.org/abs/1408.6211`

[5] M. R. Sampford, Some inequalities on Mill's ratio and related functions, Ann. Math. Statistics 24 (1953) 130–132.

[6] H. Ruben, An asymptotic expansion for a class of multivariate normal integrals, J. Austral. Math. Soc. 2 (1961/1962) 253–264.