Example 3. Let X_1, X_2 be two independent and identically distributed RV's having the same distribution as X in Example 2. Let the distribution of Y be a mixture of the distributions of X_1 and $-X_2$ with equal weights. Then Y has values $\pm 2^n$ with probabilities

$$P(Y = \pm 2^n) = 1/(2en!), \qquad n = 0, 1, 2, \ldots$$

All moments of Y exist, with $E(Y^{2r+1}) = 0$ for $r = 0, 1, 2, \ldots$ and

$$E(Y^{2r}) = \sum_{n=0}^{\infty} (2^n)^{2r} \cdot [1/(2en!)] + \sum_{n=0}^{\infty} (-2^n)^{2r} \cdot [1/(2en!)] = \exp(2^{2r} - 1), \qquad r = 0, 1, 2, \ldots$$

The MGF is

$$M_Y(t) = E(e^{tY}) = \sum_{n=0}^{\infty} \exp(t \cdot 2^n) \cdot [1/(2en!)] + \sum_{n=0}^{\infty} \exp[t \cdot (-2^n)] \cdot [1/(2en!)] = e^{-1} \sum_{n=0}^{\infty} (1/n!) \cosh(2^n t)$$

By d'Alembert's criterion one can easily see that this series converges for $t = 0$ only, and therefore the MGF of Y is determined at $t = 0$ only.

The reason that the MGF is a deficient tool in comparison to the CF is that the domain of a MGF depends on the distribution, while the domain of all CF's is the same—the real line. If the domain of the MGF is an open interval containing zero, then the corresponding CF is analytic in some strip containing the real line; in such a case the MGF can be used without difficulty, with some additional restrictions in limit theorems.

Among the distributions commonly used in introductory statistics courses, the F distribution with n_1, n_2 degrees of freedom, for $n_2 = 1$ or 2 is one of the distributions described in (1). For $n_2 > 2$, the moments $E(X^r)$ exist up to $r < n_2/2$, and the MGF exists, but it does not generate the moments.

The log normal distribution is one of the distributions described in (2). If X is Poisson distributed then $Y = e^X$ will also have the properties described in (2).

The difference $Y = X_1 - X_2$ of two independent RV's X_1, X_2, distributed according to any distribution described in (2) has a distribution described in (3).

The mixture of two distributions of X_1 and $-X_2$ with nonzero weights, where X_1, X_2 have distributions described in (2) is one of the distributions described in (3).

The RV X having the t distribution with n degrees of freedom has $n - 1$ moments, and no MGF (except at one point $t = 0$). But $Y = |X|$ has $n - 1$ moments and has the MGF $M_Y(t)$ which is well determined for $t \leq 0$. The $n - 1$ moments of Y can be generated by the left-hand derivatives of $M_Y(t)$ at $t = 0$.

Acknowledgement: I am much indebted to the referees and the associate editor who helped me in improving this paper.

A Note on the Geometric Representation of the Correlation Coefficients

CHI-KUN LEUNG AND KIN LAM*

Given two random variables X_1 and X_2 with finite variances σ_1^2 and σ_2^2, we have the following formula:

$$\sigma_{1+2}^2 = \sigma_1^2 + \sigma_2^2 + 2\rho(X_1, X_2)\sigma_1\sigma_2 \qquad (1)$$

where σ_{1+2}^2 stands for the variance of $X_1 + X_2$. Formula (1) provides us the possibility of an instructive and helpful geometric representation for random variables (see for example [1]). According to many elementary texts in mathematical statistics, it is possible to represent the correlation relationship between X_1 and X_2 by two plane vectors $\vec{v}_1$ and $\vec{v}_2$ whose lengths are proportional to σ_1 and σ_2, respectively, and are separated by an angle $\theta\,(0° \leqq \theta \leqq 180°)$ whose cosine is $\rho(X_1, X_2)$ (Fig. 1).
Formula (1) is then a restatement of the cosine law. It simply says that the length of the vector $\vec{v}_1 + \vec{v}_2$ is σ_{1+2}.

Apart from the above, it will also be instructive to point out that the relative orientation of $\vec{v}_1 + \vec{v}_2$ shows the correlation relationships of $X + Y$ with respect to X and Y. More precisely,

$$\rho(X_1 + X_2, X_1) = \cos\theta_1,$$

$$\rho(X_1 + X_2, X_2) = \cos\theta_2,$$

where θ_1 and θ_2 are as indicated in Fig. 1. The proof follows quite easily:

$$\rho(X_1 + X_2, X_1) = \frac{\operatorname{cov}(X_1 + X_2, X_1)}{\sigma_{1+2}\sigma_1} = \frac{\sigma_1^2 + \operatorname{cov}(X_2, X_1)}{\sigma_{1+2}\sigma_1}$$

* Statistics Dept., Univ. of Hong Kong, Hong Kong.

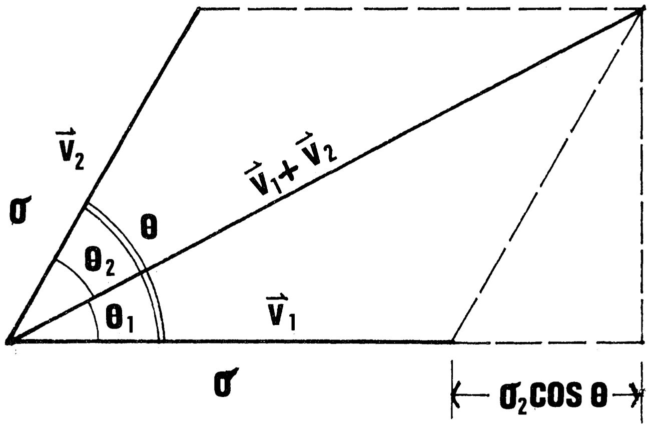

Figure 1

$$= \frac{\sigma_1 + \sigma_2 \cos\theta}{\sigma_{1+2}}.$$

A reference to Fig. 1 will show that the last expression equals $\cos\theta_1$.

Many facts concerning correlation coefficients can now be seen in terms of the geometry of the representation. For example, the statistically intuitive formula

$$\rho(X_1 + X_2, X_2) \geqq \rho(X_1, X_2) \qquad (2)$$

now corresponds to a simple fact concerning angles, i.e. a part is less than a whole. For, knowing $0 \leqq \theta_2 \leqq \theta \leqq 180°$, we can conclude $\cos\theta_2 \geqq \cos\theta$ and hence inequality (2).

In the important case when $\rho(X_1, X_2) = 0$, i.e., $\theta = \pi/2$, the vectors $\bar{v}_1$, $\bar{v}_2$ in Fig. 1 should form a rectangle. Hence we can immediately deduce that

$$\rho(X_1, X_1 + X_2) = \frac{\sigma_1}{\sigma_{1+2}}.$$

Such relationships are useful in educational statistics (see, for example, [2]).

It is equally stimulating to consider the geometric

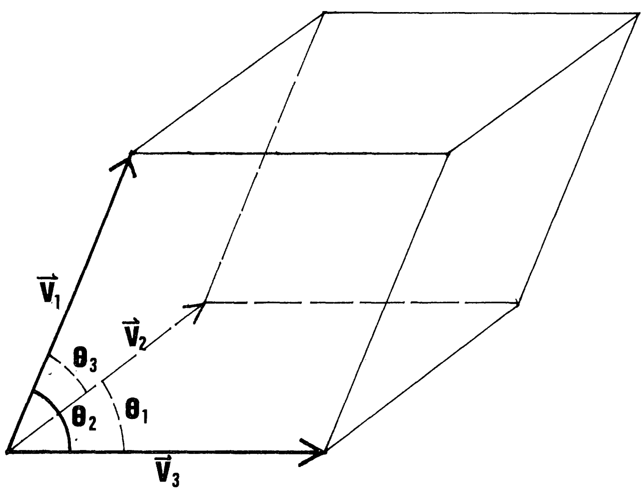

Figure 2

representation of the correlation relationship between three random variables. Given three random variables X_1, X_2 and X_3, consider three vectors $\bar{v}_1$, $\bar{v}_2$ and $\bar{v}_3$ in space, whose lengths are proportional to σ_1, σ_2 and σ_3, respectively, and are separated by angles θ_1, θ_2 and θ_3 (Fig. 2), where

$$\cos\theta_1 = \rho(X_2, X_3),$$

$$\cos\theta_2 = \rho(X_1, X_3),$$

$$\cos\theta_3 = \rho(X_1, X_2).$$

It follows from an argument similar to the above that the length of the diagonal vector of the parallelopiped formed by $\bar{v}_1$, $\bar{v}_2$ and $\bar{v}_3$ will represent the standard deviation of the random variable $X_1 + X_2 + X_3$. Furthermore, its relative orientation shows the correlation relationship between $X_1 + X_2 + X_3$ and other random variables.

We shall reap one result from this three-dimensional representation. Consider the problem of determining the possible range of values of $\rho(X_1, X_3)$ when $\rho(X_1, X_2)$ and $\rho(X_2, X_3)$ are held fixed. Those who are familiar with the partial correlation coefficient can apply the formula (see, for example, [3])

$$\rho(X_1, X_3) = \rho(X_1, X_2)\rho(X_2, X_3) + \rho_{X_1X_3.X_2}\sqrt{1 - \rho^2(X_1, X_2)}\sqrt{1 - \rho^2(X_2, X_3)}$$

to obtain

$$\rho(X_1, X_2)\rho(X_2, X_3) - \sqrt{1 - \rho^2(X_1, X_2)} \times \sqrt{1 - \rho^2(X_2, X_3)} \leqq \rho(X_1, X_3) \leqq \rho(X_1, X_2) \times \rho(X_2, X_3) + \sqrt{1 - \rho^2(X_1, X_2)}\sqrt{1 - \rho^2(X_2, X_3)} \qquad (3)$$

A resort to geometric representation is very helpful to explain the above inequality to students with no knowledge of partial correlation coefficients. Since $\rho(X_1, X_2)$ and $\rho(X_2, X_3)$ are held fixed, we can imagine a three-dimensional picture (Fig. 2), in which θ_3 and θ_1 are preassigned, while the remaining angle θ_2 at the vertex is allowed to vary. A careful study of this configuration will show that although θ_2 is free to vary, it has upper and lower bounds depending on θ_1 and θ_3. In fact,

$$\theta_1 - \theta_3 \leqq \theta_2 \leqq \theta_1 + \theta_3, \quad \text{when } \theta_1 + \theta_3 \leqq 180°$$

$$\theta_1 - \theta_3 \leqq \theta_2 \leqq 360° - (\theta_1 + \theta_3) \quad \text{when } \theta_1 + \theta_3 \geqq 180°.$$

In either case,

$$\cos(\theta_1 - \theta_3) \geqq \cos\theta_2 \geqq \cos(\theta_1 + \theta_3).$$

Equivalently,

$$\cos\theta_1 \cos\theta_3 + \sin\theta_1 \sin\theta_3 \geqq \cos\theta_2 \geqq \cos\theta_1 \cos\theta_3 - \sin\theta_1 \sin\theta_3.$$

Substituting

$$\cos\theta_2 = \rho(X_1, X_3),$$

$$\cos\theta_1 = \rho(X_2, X_3),$$

$$\cos\theta_3 = \rho(X_1, X_2),$$

$$\sin\theta_1 = \sqrt{1 - \rho^2(X_2, X_3)}, \text{ and}$$

$$\sin\theta_2 = \sqrt{1 - \rho^2(X_1, X_3)}.$$

we arrive at inequality (3).

REFERENCES

[1] B. W. Lindgren: "Statistical Theory," 2nd ed., Macmillan, London, 1968, p. 126.

[2] F. M. Lord and M. R. Novick: "Statistical Theories of Mental Test Scores," Addison-Wesley, Reading, Mass., 1968.

[3] D. F. Morrison: "Multivariate Statistical Methods," McGraw-Hill, New York, 1967, p. 89.

A Simple Proof of Local Minimum Variance Unbiasedness

S. BLUMENTHAL*

The purpose of this note is to give a compact, simple proof of the fact that a certain estimator of the common mean of two normal populations with different variances is locally minimum variance unbiased (L.M.V.U.). Specifically, if $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_n$ are mutually independent random samples from a $N(\mu, \sigma^2)$ and a $N(\mu, \rho\sigma^2)$ distribution, respectively, $(-\infty < \mu < \infty, 0 < \sigma^2 < \infty, 0 < \rho < \infty$, all unknown), then $\hat{\mu}(\rho_0) = (\rho_0\bar{X} + \bar{Y})/(1 + \rho_0)$ is L.M.V.U. at $\rho = \rho_0$. This example is given in Zacks [1] as an illustration of the theorem [1, Theorem 3.3.1] that $\phi_0(X)$ is L.M.V.U. at θ_0 if and only if ϕ_0 is uncorrelated at θ_0 with any unbiased estimator of zero (having finite variance at θ_0). We believe that our proof of this zero correlation might be desirable for classroom use because of its brevity, its use of completeness ideas developed earlier in the book, and its elimination of the need to introduce notions of translation invariance.

Let $T = (\bar{X}, \bar{Y}, Q_1, Q_2)$ be the minimal sufficient statistic comprised of the means and sums of square deviations from the two samples. Let $V = \bar{X} - \bar{Y}$. Then $U = \hat{\mu}(\rho_0)$ is $N(\mu, \sigma^2(\rho^2_0 + \rho)/n(1 + \rho_0)^2)$, V is $N(0, \sigma^2 (1 + \rho)/n)$, and $\mathrm{Cov}(U, V) = \sigma^2(\rho_0 - \rho)/n(1 + \rho_0)$. Let ω_0 be the parameter subspace with $\rho = \rho_0$. On ω_0, U, V are independent. Suppose $h(T) = h(U, V, Q_1, Q_2)$ is an unbiased estimate of zero. Write

$$G(u; \mu, \sigma^2, \rho) = \int_0^\infty \int_0^\infty \int_{-\infty}^\infty h(u, v, q_1, q_2) \times f_V(v; \sigma^2, \rho) f_{Q1}(q_1; \sigma^2) f_{Q2}(q_2; \rho\sigma^2) \times dv\,dq_1\,dq_2,$$

where f_V, f_{Q1}, f_{Q2} are the respective marginal densities. Let $G_0(u; \mu, \sigma^2) = G(u; \mu, \sigma^2, \rho_0)$. Note that on ω_0,

$$Eh(T) = \int_{-\infty}^\infty G_0(u; \mu, \sigma^2) f_U(u; \mu, \sigma^2)\,du = 0 \text{ all } (\mu, \sigma^2).$$

The completeness of the normal distributions for fixed σ^2 as μ varies, implies that

$$(*) \qquad G_0(u; \mu, \sigma^2) = 0 \text{ a.e.u. and any } \sigma^2.$$

* Dept. of Statistics, Coll. of Arts & Sciences, Univ. of Kentucky, Lexington, KY 40506.

Now note that for any parameter point θ_0 in ω_0,

$$\mathrm{Cov}_{\theta_0}(U, h(T)) = E_{\theta_0}Uh(T) = \int_{-\infty}^\infty uG_0(u; \mu, \sigma^2) f_U(u; \mu, \sigma^2)\,du$$

which is zero by (*). This completes the proof of zero correlation for any point θ_0 in ω_0.

REFERENCE

[1] Zacks, S. (1971): "The Theory of Statistical Inference," J. Wiley & Sons, Inc., New York.