# Normal's deconvolution and the independence of sample mean and variance

Problem 03.4.1

Karim M. Abadir and Jan R. Magnus

*Econometric Theory*, 19, 691 (2003)

*Econometric Theory*, **19**, 2003, 691–705. Printed in the United States of America.
DOI: 10.1017/S0266466603004122

# PROBLEMS AND SOLUTIONS

## PROBLEMS

### 03.4.1. Normal's Deconvolution and the Independence of Sample Mean and Variance

Karim M. Abadir
*University of York, UK*

Jan R. Magnus
*Tilburg University, The Netherlands*

(a) Let $x_1$ and $x_2$ be independent variates having m.g.f.s $m_1(t_1)$ and $m_2(t_2)$, respectively, and define $y := x_1 + x_2$. Prove that $y$ is normal if and only if $x_1$ and $x_2$ are both normal. Is the existence of m.g.f.s necessary for this result?

(b) Let $\boldsymbol{x} := (x_1, \dots, x_n)'$ be a vector of independent (but not necessarily identically distributed) components, where $2 \le n < \infty$. Define $\bar{x} := (1/n)\sum_{i=1}^{n} x_i$ and $z := \sum_{i=1}^{n}(x_i - \bar{x})^2$. It is well known that if $\boldsymbol{x} \sim \mathrm{N}(\mu\boldsymbol{\imath}, \sigma^2 \boldsymbol{I}_n)$, then $\bar{x} \sim \mathrm{N}(\mu, \sigma^2/n)$ independently from $z/\sigma^2 \sim \chi^2(n-1)$. For $n \ge 3$, prove that if $\bar{x} \sim \mathrm{N}(\mu, \sigma^2/n)$ and $z/\sigma^2 \sim \chi^2(n-1)$, then $\boldsymbol{x} \sim \mathrm{N}(\mu\boldsymbol{\imath}, \sigma^2 \boldsymbol{I}_n)$.

(c) Why is the last statement in (b) not necessarily true for $n = 2$? What additional conditions are needed to make it hold for $n = 2$?

*ACKNOWLEDGMENT*

We thank Stéphane Gregoir, Paolo Paruolo, and two referees for their help and acknowledge support from the British ESRC grant R000239538.

### 03.4.2. The Asymptotic Distribution of the Dickey–Fuller Statistic under Nonnegativity Constraint

Giuseppe Cavaliere
*University of Bologna, Italy*

Consider the Dickey–Fuller statistic $Z_T := T\hat{\phi}$, where $\hat{\phi} := (\sum_{t=1}^{T} X_{t-1}^2)^{-1} \times \sum_{t=1}^{T} X_{t-1}\Delta X_t$. It is well known (see, e.g., Phillips, 1987) that if the data-generating process (d.g.p.) is the random walk $X_t = X_{t-1} + \varepsilon_t$, with initial condition $X_0 = 0$ and with $\{\varepsilon_t\}$ i.i.d. $N(0, \sigma^2)$, then

$$Z_T \xrightarrow{w} \left(2\int_0^1 B(s)^2\, ds\right)^{-1} (B(1)^2 - 1), \tag{1}$$

where $B$ is a standard Brownian motion and $\xrightarrow{w}$ denotes weak convergence with respect to the uniform metric.

Solution

*Econometric Theory*, 20, 805–807 (2004)

*Econometric Theory*, **20**, 2004, 805–810. Printed in the United States of America.
DOI: 10.1017/S026646660420408X

# PROBLEMS AND SOLUTIONS

## SOLUTIONS

### 03.3.1. Normal's Deconvolution and the Independence of Sample Mean and Variance—Solution[1]

Karim M. Abadir
*University of York, UK*

Jan R. Magnus
*Tilburg University, The Netherlands*

(the posers of the problem)

(a) The "if" part is easy to prove. If $x_i \sim \mathrm{N}(\mu_i, \sigma_i^2)$ for $i = 1,2$, then their independence simplifies the moment generating function (m.g.f.) of $y$ to

$$m_y(t) \equiv \mathrm{E}(\mathrm{e}^{t(x_1+x_2)}) = \mathrm{E}(\mathrm{e}^{tx_1})\mathrm{E}(\mathrm{e}^{tx_2}) = \mathrm{e}^{(\mu_1+\mu_2)t+(\sigma_1^2+\sigma_2^2)t^2/2},$$

so that $y \sim \mathrm{N}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$.

The "only if" part is less obvious. We will assume that $y \sim \mathrm{N}(0,1)$, without loss of generality (the usual extension to $\mu + \sigma y$ applies). Then, the characteristic function (c.f.) of $y$ is

$$\mathrm{e}^{-t^2/2} = \mathrm{E}(\mathrm{e}^{\mathrm{i}ty}) = \mathrm{E}(\mathrm{e}^{\mathrm{i}t(x_1+x_2)}) = \mathrm{E}(\mathrm{e}^{\mathrm{i}tx_1})\mathrm{E}(\mathrm{e}^{\mathrm{i}tx_2}) \equiv m_1(\mathrm{i}t)m_2(\mathrm{i}t),$$

by independence of $x_1$ from $x_2$. Note that, for $t$ real-valued,

$$|\mathrm{E}(\mathrm{e}^{\mathrm{i}tx_2})| \leq \mathrm{E}(|\mathrm{e}^{\mathrm{i}tx_2}|) = \mathrm{E}(1) = 1,$$

so we have $\mathrm{e}^{-t^2/2} \leq |m_1(\mathrm{i}t)|$ or equivalently $-2\log|m_1(\mathrm{i}t)| \leq t^2$. Because the m.g.f. of $x_1$ exists, all the derivatives of $m_1(\mathrm{i}t)$ are finite at $t = 0$ and $\log|m_1(\mathrm{i}t)|$ has a Taylor-series representation as a polynomial in $t$. From the previous inequality, the maximal power of this polynomial is 2. As a result, $m_1(t) = \exp(\alpha_1 t + \alpha_2 t^2)$ for suitably chosen constants $\alpha_1$ and $\alpha_2$ (recall that $m_1(0)$ is set to 1, by definition). This establishes normality for $x_1$ and, by symmetry of the argument, for $x_2$ too.

Cramér's (1936) deconvolution theorem is actually more general than is stated in part (a), because it does not presume the existence of m.g.f.s for $x_1$ and $x_2$, at the cost of a further complication of the proof. In our proof, we have used (without needing to resort to the language of complex analysis) the fact that the existence of the m.g.f. implies that it is analytic (satisfies the Cauchy–Riemann equations) and is thus differentiable infinitely many times in an open neighborhood of $t = 0$ in the complex plane. On the other hand, if one did not assume the existence of m.g.f.s, then one would require some theorem from complex function theory. One such requisite would be the "principle of isolated zeros" or "uniqueness theorem for analytic functions." Another alternative requisite would be "Hadamard's factorization theorem," used in Loève (1977, p. 284).

(b) For $n < \infty$, Cramér's deconvolution theorem (see part (a)) can be used $n-1$ times to tell us that $\bar{x} \sim \mathrm{N}(\mu, \sigma^2/n)$ decomposes into the sum of $n$ independent normals, so that $\operatorname{var}(\boldsymbol{x}) = \boldsymbol{\Sigma}$ is a diagonal matrix satisfying $\operatorname{tr}(\boldsymbol{\Sigma}) = n\sigma^2$. However, the theorem does not imply that the components of the decomposition have identical variances and means, and we need to derive these two results, respectively.

Define the idempotent matrix $\boldsymbol{A} = (a_{ij}) := \boldsymbol{I}_n - \boldsymbol{\imath\imath}'/n$. Then, because $\boldsymbol{x}'(\sigma^{-2}\boldsymbol{A})\boldsymbol{x} \sim \chi^2(n-1)$, we have $\boldsymbol{A} = \sigma^{-2}\boldsymbol{A\Sigma A}$. The fact that $\boldsymbol{A}$ is idempotent implies that $\boldsymbol{ADA} = \mathbf{O}$, where

$$\boldsymbol{D} = \operatorname{diag}(d_1, \dots, d_n) := \boldsymbol{I}_n - \sigma^{-2}\boldsymbol{\Sigma}$$

with

$$\operatorname{tr}(\boldsymbol{D}) = n - \sigma^{-2}\operatorname{tr}(\boldsymbol{\Sigma}) = n - \sigma^{-2}n\sigma^2 = 0. \qquad \textbf{(1)}$$

The diagonal elements of $\boldsymbol{ADA}$ are given by

$$(\boldsymbol{ADA})_{jj} = d_j - \frac{2}{n}d_j + \frac{1}{n^2}\operatorname{tr}(\boldsymbol{D}) = \left(1 - \frac{2}{n}\right)d_j. \qquad \textbf{(2)}$$

For $n \geq 3$, the equation $\boldsymbol{ADA} = \mathbf{O}$ thus gives $d_j = 0$ for $j = 1, \dots, n$, and hence $\boldsymbol{\Sigma} = \sigma^2\boldsymbol{I}_n$.

To obtain the mean, we note that the noncentrality parameter of $\boldsymbol{x}'(\sigma^{-2}\boldsymbol{A})\boldsymbol{x}$ is given by $\boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1/2}(\sigma^{-2}\boldsymbol{A})\boldsymbol{\Sigma}^{-1/2}\boldsymbol{\mu}$. Because our quadratic form has a central $\chi^2$-distribution and $\boldsymbol{\Sigma} = \sigma^2\boldsymbol{I}_n$, we obtain $\boldsymbol{A\mu} = \mathbf{0}$ and hence

$$\boldsymbol{\mu} = \boldsymbol{\imath}\,\frac{\boldsymbol{\imath}'\boldsymbol{\mu}}{n}.$$

Then, $\mathrm{E}(\boldsymbol{x}) = \mu\boldsymbol{\imath}$ follows by

$$\mu = \mathrm{E}(\bar{x}) = \mathrm{E}\left(\frac{\boldsymbol{\imath}'\boldsymbol{x}}{n}\right) = \frac{\boldsymbol{\imath}'\boldsymbol{\mu}}{n}.$$

(c) When $n = 2$,

$$\boldsymbol{A} = \frac{1}{2}\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$$

and

$$\boldsymbol{ADA} = \frac{1}{4}\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} d_1 & 0 \\ 0 & d_2 \end{pmatrix}\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} = \frac{d_1 + d_2}{4}\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}.$$

Equating the latter to zero, as in (2), provides no further information on the variance of the two normal components of $\boldsymbol{x}$, beyond what was already known from (1). In this case, result (b) does not hold.

As a counterexample, let

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \sim \mathrm{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{3}{2} \end{pmatrix}\right).$$

Then, it is still the case that $\bar{x} \sim \mathrm{N}(0, \frac{1}{2})$ and

$$z = \frac{1}{2}(x_1, \quad x_2)\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \frac{1}{2}(x_1 - x_2)^2 \sim \chi^2(1).$$

Notice, however, that $\mathrm{cov}(x_1 + x_2, x_1 - x_2) = \mathrm{var}(x_1) - \mathrm{var}(x_2) \neq 0$, so that $\bar{x}$ is not independent of $z$. We will now show that assuming independence of $\bar{x}$ from $z$ makes the statement in (b) hold for $n = 2$ also.

Independence of the linear form $\boldsymbol{\imath}'\boldsymbol{x}/n$ from the quadratic form $\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x}/\sigma^2$ occurs if and only if $\boldsymbol{A\Sigma\imath} = \boldsymbol{0}$. For $n = 2$, setting

$$\frac{1}{2}\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} \sigma_1^2 - \sigma_2^2 \\ \sigma_2^2 - \sigma_1^2 \end{pmatrix}$$

equal to zero ensures that $\sigma_1^2 = \sigma_2^2$.

A variation on part (c) is proved by a different approach in Zinger (1958, Theorem 6). There, independence of $\bar{x}$ from $z$ is assumed but not the normality of $\boldsymbol{x}$. In fact, for $2 \leq n < \infty$, normality of $\boldsymbol{x}$ is obtained there as a result of one of two alternative assumptions on the components of $\boldsymbol{x}$ being pairwise identically distributed or being decomposable further as independent and identically distributed (i.i.d.) variates.

## NOTE

1. An independent solution has been proposed by Luc Lauwers, KU Leuven, Belgium.

## REFERENCES

Cramér, H. (1936) Über eine eigenschaft der normalen verteilungfunktion. *Mathematische Zeitschrift* 41, 405–414.

Loève, M. (1977) *Probability Theory I*. 4th ed. Springer-Verlag.

Zinger, A.A. (1958) Independence of quasi-polynomial statistics and analytical properties of distributions. *Theory of Probability and Its Applications* 3, 247–265.