YET ANOTHER WAY TO IDENTIFY THE NORMAL DISTRIBUTION

ERIC S. KEY

1. Introduction

We look at some new ways to identify normally distributed random variables. One of these is related to some problems in abnormal returns in stock prices.

2. Main results

We begin with the following lemma.

Lemma 2.1. *Suppose that* ϕ *is a characteristic function with* $\phi'(0) = 0$, $\phi''(0) = -\sigma^2 < 0$ *and in some open neighborhood of* 0,

$$a\phi'(t)\phi(at) = \phi'(at)\phi(t) \tag{2.1}$$

for some a, *where* $a \neq 0, \pm 1$. *Then* $\phi(t) = \exp(-\sigma^2 t^2/2)$ *for all real numbers* t.

Proof. Without loss of generality we may assume that (2.1) holds for some a with $0 < |a| < 1$. It follows from (2.1) that for all t in some open neighborhood of 0, and for all positive integers n

$$\frac{\phi'(t)}{\phi(t)} = \frac{\phi'(a^n t)}{a^n \phi(a^n t)} \tag{2.2}$$

For such t which are different from 0, divide both sides of (2.2) by t and let $n \to \infty$ to obtain

$$\frac{\phi'(t)}{t\phi(t)} = -\sigma^2.$$

Therefore in an open neighborhood of 0 we have

$$\phi'(t) + \sigma^2 t\phi(t) = 0.$$

Therefore in an open neighborhood of 0 we have $\phi(t) = \exp(-\sigma^2 t^2/2)$. This implies that $\phi(t) = \exp(-\sigma^2 t^2/2)$ for all real numbers t since the moments of the normal distribution uniquely characterize that distribution. □

Theorem 2.2. *Suppose that X and Y are independent random variables with standard deviations $\sigma_X > 0$ and $\sigma_Y > 0$ respectively, and $(X - \mathrm{E[X]})/\sigma_{\mathrm{X}}$ and $(Y - \mathrm{E[Y]})/\sigma_{\mathrm{Y}}$ are identically distributed. Then there exists a constant C such that*

$$\mathrm{E[(CX - (1-C)Y)I_{X+Y\leq t}] = 0} \tag{2.3}$$

for all t iff $C = \sigma_Y^2/(\sigma_X^2 + \sigma_Y^2)$, $C\mathrm{E[X]} = (1 - \mathrm{C})\mathrm{E[Y]}$, and either X and Y are identically distributed or X and Y are normally distributed.

Proof. We consider the easy direction first. Suppose that $C = \sigma_Y^2/(\sigma_X^2 + \sigma_Y^2)$ and $C\mathrm{E[X]} = (1 - \mathrm{C})\mathrm{E[Y]}$. If X and Y are identically distributed then (2.3) is trivially true. If X and Y are normally distributed, then $C(X - \mathrm{E[X]}) - (1 - \mathrm{C})(\mathrm{Y} - \mathrm{E[Y]})$ and $X + Y$ are uncorrelated, and, therefore, independent. Therefore

$$\mathrm{E[(CX - (1-C)Y)I_{X+Y\leq t}] = (CE[X] - (1-C)E[Y]) \cdot Pr(X + aY \leq t) = 0}$$

as required.

Now for the interesting direction. Suppose that (2.3) holds. Letting $t \to \infty$ we see that $C\mathrm{E[X]} - (1 - \mathrm{C})\mathrm{E[Y]} = 0$, so we have

$$\mathrm{E[(C(X - E[X]) - (1-C)(Y - E[Y])I_{X+Y\leq t}] = 0.} \tag{2.4}$$

It follows from (2.4) that for all bounded continuous functions f,

$$\mathrm{E[(C(X - E[X]) - (1-C)(Y - E[Y])f(X + Y)] = 0.} \tag{2.5}$$

If we take $f(u) = \exp(it(u - \mathrm{E[X]} - \mathrm{E[Y]}))$ for any real number t, and let ϕ_X denote the characteristic function of $X - \mathrm{E[X]}$ and let ϕ_Y denote the characteristic function of $Y - \mathrm{E[Y]}$ it follows from (2.4) that

$$C\phi_X'(t)\phi_Y(t) = (1 - C)\phi_Y'(t)\phi_X(t). \tag{2.6}$$

since X and Y have finite mean. Dividing both sides of (2.6) by t and letting $t \to 0$ gives $C\sigma_X^2 = (1 - C)\sigma_Y^2$ as required. If $\sigma_X = \sigma_Y$ then X and Y have the same mean and variance, so they are identically distributed. Suppose that $\sigma_X \neq \sigma_Y$. Let ϕ denote the common characteristic function of $(X - \mathrm{E[X]})/\sigma_{\mathrm{X}}$ and $(Y - \mathrm{E[Y]})/\sigma_{\mathrm{Y}}$. Then from (2.6) we have

$$\sigma_Y^2\sigma_X\phi'(\sigma_X t)\phi(\sigma_Y t) = \sigma_X^2\sigma_Y\phi'(\sigma_Y t)\phi(\sigma_X t). \tag{2.7}$$

If we set $u = \sigma_X t$ and $a = \sigma_Y/\sigma_X$ in (2.7) we get

$$a\phi'(u)\phi(au) = \phi'(au)\phi(u)$$

and normality follows from the Lemma. □

As an application of this theorem, we consider the following model for the return on a stock, R, given in Vachadze[2001]. A, B, and J are independent random variables. A and B are normal, and $\Pr(J = 1) = 1 - \Pr(J = 0) = p \in (0, 1)$, and it was assumed that $R = A + JB$. The idea is that under usual circumstances R is distributed like A, but from time to time there is unusual information which changes the return distribution to be that of $A + B$.

It was subsequently shown in Vachadze[2002] that if we let f_X denote the density of a random variable X then

$$\mathrm{E[J|R] = p\frac{f_{A+B}(R)}{f_R(R)}} \tag{2.8}$$

and

$$\mathrm{E[J(B - E[B])|R] = \frac{Var[B]}{Var[A + B]}(R - E[A + B])E[J|R]} \tag{2.9}$$

It is easy to check that (2.8) remains true so long as A and B have densities. The same is not true for (2.9). In fact, we have

Theorem 2.3. *Suppose that A, B and J are independent random variables, that A and B have (absolutely) continuous distributions with variances σ_A^2 and σ_B^2 respectively, that $(A - \mathrm{E[A]})/\sigma_\mathrm{A}$ and $(B - \mathrm{E[B]})/\sigma_\mathrm{B}$ are identically distributed, and that $\Pr(J = 1) = 1 - \Pr(J = 0) = p \in (0, 1)$. Put $R = A + JB$. If*

$$\mathrm{E[J(B - E[B])|R] = C(R - D)p\frac{f_{A+B}(R)}{f_R(R)}} \tag{2.10}$$

for some constants C and D then $C = \sigma_B^2/(\sigma_A^2+\sigma_B^2)$, $D = \mathrm{E[A+B]}$, and if $\sigma_A^2 \neq \sigma_B^2$ then A and B are normally distributed.

Proof. It follows from the definition of conditional expectation that if F is any Borel subset of the real line then

$$\begin{aligned} p\mathrm{E}[(\mathrm{B}-\mathrm{E}[\mathrm{B}])\mathrm{I}_{\mathrm{A+B}\in\mathrm{F}}] &= \mathrm{E}[(\mathrm{B}-\mathrm{E}[\mathrm{B}])\mathrm{I}_{\mathrm{A+B}\in\mathrm{F}}\mathrm{I}_{\mathrm{J}=1}] \\ &= \mathrm{E}[\mathrm{J}(\mathrm{B}-\mathrm{E}[\mathrm{B}])\mathrm{I}_{\mathrm{A+JB}\in\mathrm{F}}\mathrm{I}_{\mathrm{J}=1}] \\ &= \mathrm{E}[\mathrm{J}(\mathrm{B}-\mathrm{E}[\mathrm{B}])\mathrm{I}_{\mathrm{R}\in\mathrm{F}}] \\ &= pC\mathrm{E}[(\mathrm{R}-\mathrm{D})\frac{\mathrm{f}_{\mathrm{A+B}}(\mathrm{R})}{\mathrm{f}_{\mathrm{R}}(\mathrm{R})}\mathrm{I}_{\mathrm{R}\in\mathrm{F}}] \\ &= pC\int_{\mathbf{R}}(R-D)f_{A+B}(R)I_F(R)\,dR] \\ &= pC\mathrm{E}[(\mathrm{A}+\mathrm{B}-\mathrm{D})\mathrm{I}_{\mathrm{A+B}\in\mathrm{F}}] \end{aligned}$$

so

$$\mathrm{E}[(\mathrm{C}(\mathrm{A}+\mathrm{E}[\mathrm{B}]-\mathrm{D})-(1-\mathrm{C})(\mathrm{B}-\mathrm{E}[\mathrm{B}]))\mathrm{I}_{\mathrm{A+B}\in\mathrm{F}}]=0 \tag{2.11}$$

for each Borel subset of the real line. Taking $F=(-\infty,t+D]$ we see that Theorem 2.3 follows from Theorem 2.2. □

3. References

1. Vachadze, G. [2001] *Recovery of hidden information from stock price data: a semiparametric approach.* Journal of Economics and Finance 25(3) pages 243 - 258.
2. Vachadze, G. [2002] *Mean Square Optimal Estimator of Abnormal Stock Return.* Unpublished Manuscript.