cemmap

centre for microdata methods and practice

Convolution without independence

Susanne Schennach

The Institute for Fiscal Studies
Department of Economics, UCL

cemmap working paper CWP46/13

E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL

An ESRC Research Centre

Convolution without independence

Susanne M. Schennach*

August 25, 2013

Abstract

Widely used convolutions and deconvolutions techniques traditionally rely on the assumption of independence, an assumption often criticized as being very strong. We observe that independence is, in fact, not necessary for the convolution theorem to hold. Instead, a much weaker notion, known as subindependence, is the appropriate necessary and sufficient condition. We motivate the usefulness of the subindependence concept by showing that is arguably as weak as a conditional mean assumption. We also provide an equivalent definition of subindependence that does not involve Fourier transforms and devise a constructive method to generate pairs of subindependent random variables.

1 Introduction

Convolutions and deconvolutions play a central role in the identification and the estimation of measurement error models (Fan (1991), Fan and Truong (1993), Li (2002), Li and Vuong (1998), Wang and Hsiao (2011), Taupin (2001), Hu and Ridder (2012), Hu and Ridder (2010), Bonhomme and Robin (2010), Carrasco and Florens (2011), Wilhelm (2010), Schennach (2004), Schennach (2007), Schennach (2008), Schennach (2013b)) and, more generally, in any problem involving sums of independent random variables. The use of convolution techniques in this context yields very computationally and conceptually convenient methods. However, the requirement that the variables (e.g. the true quantity of interest and its measurement error) be independent is often criticized as being too strong (Bound, Brown, and Mathiowetz (2001), Hu and Schennach (2008)). In this note, we observe that independence is, in fact, not necessary for the convolution theorem to hold. Instead, a much weaker notion, known as subindependence, is the appropriate necessary and sufficient condition.

Although the concept of subindependence and its relation to convolutions is known (Hamedani and Volkmer (2009), Ebrahimi, Hamedani, Soofi, and Volkmer (2010)), it has

*The author acknowledges support from NSF grants SES-0752699 and SES-1061263.

received surprisingly little attention. This paper contributes to this literature (i) by motivating the usefulness of this concept by showing that it is as weak as a conditional mean assumption in a well-defined sense, (ii) by providing an equivalent and general definition of subindependence that does not involve Fourier transforms and (iii) by devising a simple and general method to generate pairs of subindependent random variables.

2 Main Results

Let X and Y denote two scalar real-valued random variables and let $Z = X + Y$. Let the characteristic functions (c.f.) of some random variable X be denoted by $\phi_X(\xi) \equiv E\left[e^{i\xi X}\right]$ and let the joint c.f. of two variables X and Y be denoted by $\phi_{XY}(\xi, \gamma) = E\left[e^{i\xi X} e^{i\gamma Y}\right]$. We denote the density of a random variable X (with respect the Lebesgue measure) by f_X while its cdf is denoted by F_X, and similarly for joint densities and cdf.

The convolution theorem (Loève (1977), Lukacs (1970)) states that, under independence of X and Y, we have the convenient factorization $\phi_Z(\xi) = \phi_{X+Y}(\xi) = E\left[e^{i\chi X+\gamma Y}\right] = E\left[e^{i\chi X}\right] E\left[e^{i\gamma Y}\right] = \phi_X(\xi)\,\phi_Y(\xi)$. Such a result does not actually require full independence, because the latter is equivalent to the following assumption (by Theorem 16-B in Loève (1977)):

Assumption 1 *(Independence)* $E\left[e^{i(\chi X+\gamma Y)}\right] = E\left[e^{i\chi X}\right] E\left[e^{i\gamma Y}\right]$ *for any* $\chi \in \mathbb{R}$ *and* $\gamma \in \mathbb{R}$.

Note that independence requires the factorization to hold for any χ and any γ when the convolution theorem only needs the factorization to hold for $\chi = \gamma$. This observation leads to the following weaker assumption:

Assumption 2 *(Subindependence)* $E\left[e^{i\xi(X+Y)}\right] = E\left[e^{i\xi X}\right] E\left[e^{i\xi Y}\right]$ *for all* $\xi \in \mathbb{R}$.

Note that the number of restrictions imposed by subindependence is considerably less than for independence: Only a one-dimensional subset of the domain of the joint c.f. of X and Y is constrained instead of its whole two-dimensional domain. For comparison, this is as little constraint on the joint c.f. as a conditional mean assumption $E[Y|X = x] = 0$,[1] which can also be expressed as a constraint on the c.f. on a one-dimensional subset (see Proposition 2 in Schennach (2013a)):

Assumption 3 *(Conditional mean)* $[\partial\phi_{XY}(\chi, \gamma)/\partial\gamma]_{\gamma=0} = E\left[Y e^{i\chi X}\right] = 0$ *for all* $\chi \in \mathbb{R}$.

[1] For almost every $x \in \mathbb{R}$, and assuming that $E[|Y|\,|X = x] < \infty$ and $E[|Y|] < \infty$.

Informally, one could interpret these observations as follows. If one were to select a generating process for X, Y at "random", the chances that it satisfies subindependence are of the same order as the chances that it satisfies a conditional mean assumption, while the chances of satisfying independence are considerably smaller. Another interpretation is that, in a case where the convolution theorem does not hold, the error made in using it anyway is related to the "distance" to the nearest generating process satisfying subindependence, which is much "closer" than the nearest model satisfying independence.

One could argue that, regardless of favorable dimensionality arguments, the notion of subindependence is less intuitive than the one of conditional mean, because the former is apparently tied to a Fourier representation. To address this concern, we now provide simple ways to characterize and generate subindependent pairs of random variables that does not involve Fourier transforms. First, here is an equivalent characterization of subindependence:

Lemma 1 *Two scalar real-valued random variables X and Y are subindependent iff*

$$F_{X+Y}(z) = \int\int 1(x + y \leq z)\, dF_X(x)\, dF_Y(y). \tag{1}$$

Note that the left-hand side of (1) is the distribution of the sum $X + Y$ (accounting for possible dependence between X and Y), while the right-hand side is the convolution of the marginal distributions of X and Y, expressed in a form that allows for general probability measures. The two expressions are obviously equal under independence, but this lemma shows that it holds under the weaker conditions of subindependence. For densities (instead of general probability measures), Lemma 2 in Ebrahimi, Hamedani, Soofi, and Volkmer (2010) gives a lengthier but more transparent characterization of subindependence:

Lemma 2 *Two scalar real-valued random variables X and Y with continuous joint density $f_{XY}(x, y)$ (and marginals $f_X(x)$ and $f_Y(y)$, respectively) are subindependent iff*

$$\Delta f_{XY}(x, y) \equiv f_{XY}(x, y) - f_X(x) f_Y(y)$$

satisfies

$$\int_{-\infty}^{\infty} \Delta f_{XY}(x, y)\, dx = 0 \text{ for any } y \in \mathbb{R} \tag{2}$$

$$\int_{-\infty}^{\infty} \Delta f_{XY}(x, y)\, dy = 0 \text{ for any } x \in \mathbb{R} \tag{3}$$

$$\int_{-\infty}^{\infty} \Delta f_{XY}(x, z - x)\, dx = 0 \text{ for any } z \in \mathbb{R}. \tag{4}$$

Although it is straightforward to find functions $\Delta f_{XY}(x, y)$ satisfying (2) and (3), it is more difficult to do so while at the same time satisfying (4). For this reason, we provide a simple construction to generate pairs of subindependent random variables.

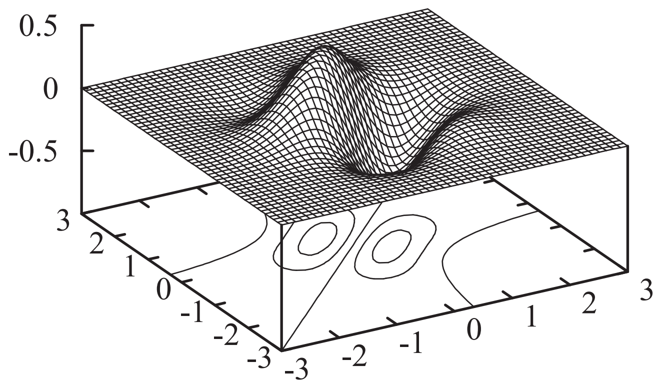

Figure 1: Function $\Delta f_{XY}(x,y)$ from Example 1.

Theorem 1 *Let X and Y be scalar real-valued random variables with marginal density $f_X(x)$ and $f_Y(y)$, respectively, and satisfying $E[|X|] < \infty$ and $E[|Y|] < \infty$. Any joint density $f_{XY}(x,y)$ such that X and Y are subindependent can be written in the form*

$$f_{XY}(x,y) = f_X(x) f_Y(y) + \left(\frac{\partial}{\partial x} - \frac{\partial}{\partial y}\right) p(x,y) \tag{5}$$

for some function $p : \mathbb{R}^2 \mapsto \mathbb{R}$ such that

$$\int_{-\infty}^{\infty} p(x,y)\, dx = 0 \text{ and } \int_{-\infty}^{\infty} p(x,y)\, dy = 0 \tag{6}$$

and

$$\lim_{|x|\to\infty} p(x, z-x) = 0 \tag{7}$$

for any $z \in \mathbb{R}$.

Remark 1 *This theorem does not guarantee that, for any choice of $p(x,y)$, the resulting $f_{XY}(x,y)$ is a valid probability density. However, it does guarantee that if one considered every possible $p(x,y)$ satisfying the restrictions and such that (5) is a well-defined density, one would have covered all possible joint densities that satisfy subindependence. There are essentially two ways in which $f_{XY}(x,y)$ could fail to be a valid density: (i) if $p(x,y)$ is not sufficiently differentiable, which is easy to avoid and (ii) if the resulting function $f_{XY}(x,y)$ reaches negative values, in which case one can merely rescale $p(x,y)$ so that $f_{XY}(x,y) \geq 0$ everywhere.*

We can use Theorem 1 to construct simple examples that provide graphical intuition into the concept of subindependence.

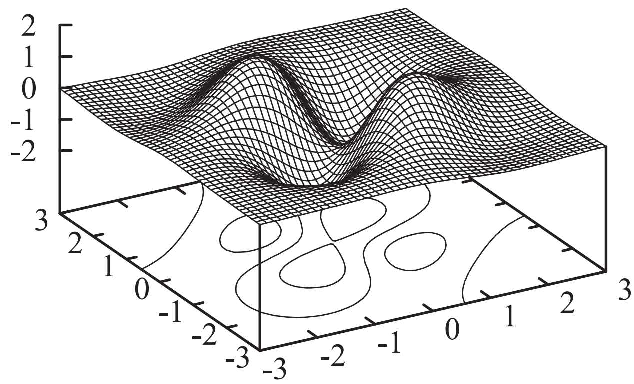

Figure 2: Function $\Delta f_{XY}(x,y)$ from Example 2.

Example 1 *Taking* $p(x,y) = xye^{-\left(x^2+y^2\right)/2}$ *yield* $f_{XY}(x,y) = f_X(x)\, f_Y(y) + \Delta f_{XY}(x,y)$ *with* $\Delta f_{XY}(x,y) = \left(y - x + xy^2 - yx^2\right) e^{-\frac{1}{2}x^2 - \frac{1}{2}y^2}$.

The deviation $\Delta f_{XY}(x,y)$ is shown in Figure 1 and illustrates perhaps the simplest general shape of a deviation from independence that will preserve subindependence. One can also easily construct an example (illustrated in Figure 2) where independence is violated but subindependence and conditional mean $E[Y|X] = 0$ hold. This is useful to see that subindependence is not incompatible with the natural conditional mean assumption.

Example 2 *Taking* $p(x,y) = \left(y^2 - 1\right) xe^{-\left(x^2+y^2\right)/2}$ *yield* $f_{XY}(x,y) = f_X(x)\, f_Y(y) + \Delta f_{XY}(x,y)$ *with* $\Delta f_{XY}(x,y) = \left(-1 + x^2 + y^2 - 3xy + xy^3 - x^2y^2\right) e^{-\frac{1}{2}x^2 - \frac{1}{2}y^2}$. *Note that* $\int y f_{XY}(x,y)\, dy = 0$ *in this case.*

3 Conclusion

This paper's aim is not to try to argue that economic models should be stated in terms of subindependence, which would admittedly be an unnatural assumption. Rather, we are arguing that inferences made under the assumption of independence are robust to large deviations from independence that maintain subindependence. This considerably expands the scope of validity of the wide range of methods developed under independence, because subindependence is arguably just as weak an assumption as conditional mean. Indeed, both conditions, when phrased in terms of c.f., impose constraints on a subset of its domain that is of the same dimension. We provide explicit examples that illustrate that deviations from independence that maintain subindependence have quite simple and plausible shapes.

A Proofs

Proof of Lemma 1. To show the equivalence of (1) and assumption 2, we observe that the equality of the cdf $F_{X+Y}(z)$ and $\tilde{F}_Z(z) \equiv \int \int 1(x+y \leq z)\, dF_X(x)\, dF_Y(y)$ is equivalent to the equality of the corresponding probability measures $dF_{X+Y}(z)$ and $d\tilde{F}_Z(z)$. By the well-known uniqueness of c.f. (Loève (1977), Lukacs (1970)), this is equivalent to the equality between the corresponding Fourier transforms:

$$\int e^{i\xi z} dF_{X+Y}(z) = \int e^{i\xi z} d\tilde{F}_Z(z). \tag{8}$$

The left-hand side of (8) is obviously

$$\int e^{i\xi z} dF_{X+Y}(z) = E\left[e^{i\xi(X+Y)}\right] \tag{9}$$

by construction. Evaluating the right-hand-side yields:

$$\begin{aligned} \int e^{i\xi z} d\tilde{F}_Z(z) &= \int e^{i\xi z} d\left(\int \int \mu_{x+x}(z)\, dF_X(x)\, dF_Y(y)\right) \\ &= \int \int \int e^{i\xi z} d\mu_{x+y}(z)\, dF_X(x)\, dF_Y(y) \end{aligned}$$

where $\mu_{z_0}(z) \equiv 1(z_0 \leq z)$ and where the second equality follows from Fubini's theorem for finite measures (see Chapter 5 in Bhattacharya and Rao (2010)). Since $d\mu_{x+y}(z)$ represents a unit point mass at $z = x + y$, we have

$$\begin{aligned} \int e^{i\xi z} d\tilde{F}_Z(z) &= \int \int e^{i\xi(x+y)} dF_X(x)\, dF_Y(y) \\ &= \int \int e^{i\xi x} e^{i\xi y} dF_X(x)\, dF_Y(y) \\ &= \int e^{i\xi x} dF_X(x) \int e^{i\xi y} dF_Y(y) \\ &= E\left[e^{i\xi x}\right] E\left[e^{i\xi z}\right] \end{aligned} \tag{10}$$

where we have again used Fubini's theorem for finite (complex) measures. Equating (9) and (10) for any $\xi \in \mathbb{R}$ yields Assumption 2. ■

Proof of Theorem 1. Subindependence of X and Y requires that $\phi_{XY}(\chi,\chi) = \phi_X(\chi)\,\phi_Y(\chi) = \phi_{XY}(\chi,0)\,\phi_{XY}(0,\chi)$. Therefore, $\phi_{XY}(\chi,\gamma)$ can be written as:

$$\phi_{XY}(\chi,\gamma) = \phi_X(\chi)\,\phi_Y(\gamma) + \Delta\phi_{XY}(\chi,\gamma) \tag{11}$$

where $\Delta\phi_{XY}(\chi,\gamma) = 0$ if either $\chi = 0$, $\gamma = 0$ or $\chi = \gamma$. Since $\Delta\phi_{XY}(\chi,\gamma)$ is a difference of c.f., which are always bounded, it is also bounded. Since $E[|X|]$ and $E[|Y|]$ are finite,

$\phi_X(\chi)$, $\phi_Y(\gamma)$ and $\phi_{XY}(\chi,\gamma)$ are everywhere continuously differentiable and, therefore, so is $\Delta\phi_{XY}(\chi,\gamma)$. In particular, this implies that near the line $\chi=\gamma$ (where $\Delta\phi_{XY}(\chi,\gamma)$ vanishes), $\Delta\phi_{XY}(\chi,\gamma)$ behaves linearly and the ratio $\frac{\Delta\phi_{XY}\ (\chi,\gamma)}{i(\chi-\gamma)}$ does not diverge ($i=\sqrt{-1}$ is a constant introduced for convenience). Away from this line, $(\chi-\gamma)$ is nonzero, so the ratio does not diverge either and $\Delta\phi_{XY}(\chi,\gamma)$ can be written in the form:

$$\Delta\phi_{XY}(\chi,\gamma)=i(\chi-\gamma)\,\psi(\chi,\gamma) \tag{12}$$

where $\psi(\chi,\gamma)$ is finite at each $(\chi,\gamma)\in\mathbb{R}^2$ and such that $\psi(\chi,0)=0$ and $\psi(0,\gamma)=0$. Equation (12). Since $(\chi-\gamma)$ is nonzero along the lines $\chi=0$ and $\gamma=0$ (except at $\chi=\gamma=0$) the constraints $\Delta\phi_{XY}(\chi,0)=0$ and $\Delta\phi_{XY}(0,\gamma)=0$ translate directly into the constraint that $\psi(\chi,0)=0$ and $\psi(0,\gamma)=0$. The inverse Fourier transform of these constraints yield (6). (The value at $\psi(0,0)$ is irrelevant since it is finite and multiplied by $(\chi-\gamma)=0$.) Since $\Delta\phi_{XY}(\chi,\gamma)$ and $\psi(\xi,\gamma)$ are bounded, they are a special case of tempered distributions and admit an inverse Fourier transform, given by (Lighthill (1962)):

$$\Delta f_{XY}(x,y)=\left(\frac{\partial}{\partial x}-\frac{\partial}{\partial y}\right)p(x,y) \tag{13}$$

where $\Delta f_{XY}(x,y)$ and $p(x,y)$ denote the inverse Fourier transforms of $\Delta\phi_{XY}(\chi,\gamma)$ and $\psi(x,\zeta)$, respectively. Combining (13) with the inverse Fourier transform of (11) yields (5). We can further restrict the behavior of $p(x,y)$ by invoking Equation (4) from Lemma 2 with $\Delta f_{XY}(x,y)$ given by (13):

$$\int_{-\infty}^{\infty}\left[\left(\frac{\partial}{\partial x}-\frac{\partial}{\partial y}\right)p(x,y)\right]_{y=z-x}dx=0$$

for any $z\in\mathbb{R}$. Letting superscripts denote orders of derivatives with respect to each arguments, we can write:

$$\begin{aligned}0 &= \int_{-\infty}^{\infty}\left(p^{(10)}(x,z-x)-p^{(01)}(x,z-x)\right)dx=\int_{-\infty}^{\infty}\frac{d}{dx}p(x,z-x)\,dx\\ &= \lim_{x\to\infty}p(x,z-x)-\lim_{x\to-\infty}p(x,z-x)\end{aligned}$$

Hence $\lim_{x\to\infty}p(x,z-x)=\lim_{x\to-\infty}p(x,z-x)$. These limits must also be zero because otherwise the constraints (6) would diverge, thus implying (7). ∎

References

Bhattacharya, R. N., and R. R. Rao (2010): *Normal Approximation and Asymptotic Expansions*. SIAM, Philadelphia.

Bonhomme, S., and J.-M. Robin (2010): "Generalized Non-Parametric Deconvolution with an Application to Earnings Dynamics," *Review of Economic Studies*, 77, 491–533.

Bound, J., C. Brown, and N. Mathiowetz (2001): "Measurement Error in Survey Data," in *Handbook of Econometrics*, ed. by J. J. Heckman, and E. Leamer, vol. V. Elsevier Science.

Carrasco, M., and J.-P. Florens (2011): "Spectral method for deconvolving a density," *Econometric Theory*, 27, 546–581.

Ebrahimi, N., G. Hamedani, E. S. Soofi, and H. Volkmer (2010): "A class of models for uncorrelated random variables," *Journal of Multivariate Analysis*, 101, 1859–1871.

Fan, J. (1991): "On the Optimal Rates of Convergence for Nonparametric Deconvolution Problems," *Annals of Statistics*, 19(3), 1257–1272.

Fan, J., and Y. K. Truong (1993): "Nonparametric Regression with Errors in Variables," *Annals of Statistics*, 21(4), 1900–1925.

Hamedani, G. G., and H. W. Volkmer (2009): "Comment on 'Conditional Moments and Independence' by A. De Paula.," *The American Statistician*, 63, 295–295.

Hu, Y., and G. Ridder (2010): "On Deconvolution as a First Stage Nonparametric Estimator," *Econometric Reviews*, 29, 1–32.

Hu, Y., and G. Ridder (2012): "Estimation of Nonlinear Models with Measurement Error Using Marginal Information," *Journal of Applied Econometrics*, 27, 347–85.

Hu, Y., and S. M. Schennach (2008): "Instrumental Variable Treatment of Nonclassical Measurement Error Models," *Econometrica*, 76, 195–216.

Li, T. (2002): "Robust and consistent estimation of nonlinear errors-in-variables models," *Journal of Econometrics*, 110, 1–26.

Li, T., and Q. Vuong (1998): "Nonparametric Estimation of the Measurement Error Model Using Multiple Indicators," *Journal of Multivariate Analysis*, 65, 139–165.

Lighthill, M. J. (1962): *Introduction to Fourier Analysis and Generalized Function*. London: Cambridge University Press.

LOÈVE, M. (1977): *Probability Theory I.* New York: Springer.

LUKACS, E. (1970): *Characteristic Functions.* Griffin, London, second edn.

SCHENNACH, S. M. (2004): "Estimation of Nonlinear Models with Measurement Error," *Econometrica*, 72, 33–75.

SCHENNACH, S. M. (2007): "Instrumental Variable Estimation of Nonlinear Errors-in-Variables Models," *Econometrica*, 75, 201–239.

SCHENNACH, S. M. (2008): "Quantile regression with mismeasured covariates," *Econometric Theory*, 24, 1010–1043.

SCHENNACH, S. M. (2013a): "Entropic Latent Variable Integration via Simulation," *Econometrica*, forthcoming, http://www.cemmap.ac.uk/cemmap/publication/id/6814.

SCHENNACH, S. M. (2013b): "Measurement error in nonlinear models — A review," in *Advances in Economics and Econometrics*, ed. by D. Acemoglu, M. Arellano, and E. Dekel, vol. 3. Cambridge University Press.

TAUPIN, M.-L. (2001): "Semi-Parametric Estimation in the Nonlinear Structural Errors-in-Variables Model," *Annals of Statistics*, 29, 66–93.

WANG, L., AND C. HSIAO (2011): "Method of moments estimation and identifiability of nonlinear semiparametric errors-in-variables models," *Journal of Econometrics*, 165, 30–44.

WILHELM, D. (2010): "Identification and Estimation of Nonparametric Panel Data Regressions with Measurement Error," Working Paper, University of Chicago.