How to Explain the Definition of Stochastic Affiliation to Economics Students

Tonghui Wang\(^1\), Olga Kosheleva\(^2\), Vladik Kreinovich\(^3\),\(^*\)

\(^1\)Department of Mathematical Sciences, New Mexico State University, Las Cruces, NM 80003, USA
\(^2\)College of Education, University of Texas at El Paso, 500 W. University, El Paso, TX 79968, USA
\(^3\)Department of Computer Science, University of Texas at El Paso, 500 W. University, El Paso, TX 79968, USA

Received 15 January 2014; Revised 15 June 2014

Abstract

To formally describe the intuitive idea of “positive correlation” between two quantities, it is often helpful to use the notion of stochastic affiliation. While this notion is useful, its usual definition is not intuitively clear – which makes it difficult to explain this notion to, e.g., economics students. To help students understand this notion, in this paper, we show how the notion of stochastic affiliation can be explained in clear probabilistic terms.

\(c\)\(2015\) World Academic Press, UK. All rights reserved.

Keywords: stochastic affiliation, economics, teaching statistics

1 Formulation of the Problem: The Notion of Stochastic Affiliation is Difficult to Explain

Positive correlation as an important intuitive idea. In the statistical analysis of economic situations, it is often important to check which pairs of quantities \(x\) and \(y\) are “positively correlated” – in the sense that:

- the increase in \(x\) makes it more probable that \(y\) increases, and, vice versa,
- the increase in \(y\) makes it more probable that \(x\) increases.

The notion of stochastic affiliation. One of the most useful formalizations of the idea of positive correlation is the notion of stochastic affiliation; see, e.g., \([1,2]\). Two random variables \(X\) and \(Y\) whose joint distribution is described by a probability density function \(f(x, y)\) are called stochastically affiliated if for every \(x_1 \leq x_2\) and \(y_1 \leq y_2\), we have \(f(x_1, y_2) \cdot f(x_2, y_1) \leq f(x_1, y_1) \cdot f(x_2, y_2)\).

A pedagogical problem. The above definition leads to many interesting mathematical results (see, e.g., \([1,2]\)). However, this definition is difficult to understand: why do we compare the two products? This lack of intuitive understanding makes it especially difficult to teach this concept to non-mathematics students – e.g., to economics students.

What we do in this paper. In this paper, we explain how the above intuitive idea of positive correlation leads to the above formal definition. This explanations makes the definition of stochastic affiliation easier to understand for economics students.

\(^*\)Corresponding author.

Emails: twang@nmsu.edu (T. Wang), olgak@utep.edu (O. Kosheleva), vladik@utep.edu (V. Kreinovich).
2 Stochastic Affiliation: Our Explanation

Main idea behind the proposed explanation. Suppose that we have selected values $x_1 \leq x_2$ and we consider only situations in which the variable $X$ takes one of these two values. Suppose also that we have selected values $y_1 \leq y_2$ and we consider only situations in which the variable $Y$ takes one of these two values.

For each of the two values $y = y_1$ and $y = y_2$, we can consider the probabilities $p(X = x_1 \mid Y = y)$ and $p(X = x_2 \mid Y = y)$. Positive correlation means that when we increase the $y$-value from $y_1$ to $y_2$, then the probability of the larger $x$-value also increases, i.e., that

$$p(x_2 \mid Y = y_2) \geq p(x_2 \mid Y = y_1).$$

Let us show that this intuitively clear definition leads to the original (seemingly unclear) definition of stochastic affiliation.

Intuitive meaning of conditional probability: reminder. The above idea is based on the notion of stochastic affiliation. From the above idea to the formal definition of stochastic affiliation.

By definition of the probability density function, for each pair $(x, y)$, the probability that $X$ is between $x$ and $x + dx$ and that $Y$ is between $y$ and $y + dy$ is equal to $f(x, y) \cdot dx \cdot dy$. Thus, the conditional probability that $X$ is between $x$ and $x + dx$ under the condition that $Y$ is between $y$ and $y + dy$ is equal to the ratio

$$\frac{f(x, y) \cdot dx \cdot dy}{p(y \leq Y \leq y + dy)}.$$

By definition of the marginal probability density $f_Y(y)$, the probability in the denominator is equal to $f_Y(y) \cdot dy$. Thus, the desired conditional probability $p_y(X = x) \triangleq p(x \mid y)$ is equal to the ratio $f(x, y) \cdot dx / f_Y(y)$. In particular:
Multiplying both numerator and denominator by $f_Y(y)$ and dividing by $dx$, we get

$$p(X = x_2 | Y = y) = \frac{f(x_2, y)\cdot dx}{f_Y(y)\cdot dx} = \frac{f(x_2, y)\cdot dx}{f_Y(y)}.$$

The requirement that this probability increases with $y$ means that

$$p(X = x_2 | Y = y_2) \geq p(X = x_2 | Y = y_1),$$

i.e., that

$$\frac{f(x_2, y_2)}{f(x_1, y_2) + f(x_2, y_2)} \geq \frac{f(x_2, y_1)}{f(x_1, y_1) + f(x_2, y_1)}.$$

This inequality leads to the following inequality between the inverses:

$$\frac{f(x_1, y_2) + f(x_2, y_2)}{f(x_2, y_2)} \leq \frac{f(x_1, y_1) + f(x_2, y_1)}{f(x_2, y_1)},$$

i.e., to

$$\frac{f(x_1, y_2)}{f(x_2, y_2)} + 1 \leq \frac{f(x_1, y_1)}{f(x_2, y_1)} + 1.$$

Subtracting 1 from both sides of this inequality, we get

$$\frac{f(x_1, y_2)}{f(x_2, y_2)} \leq \frac{f(x_1, y_1)}{f(x_2, y_1)}.$$

Multiplying both sides by both (non-negative) denominators, we get the desired inequality

$$f(x_1, y_2) \cdot f(x_2, y_1) \leq f(x_1, y_1) \cdot f(x_2, y_2).$$

The formula that forms the traditional definition of stochastic affiliation has thus been derived.

**Acknowledgments**

This work was supported in part by the National Science Foundation grants HRD-0734825 and HRD-1242122 (Cyber-ShARE Center of Excellence) and DUE-0926721, and by a travel grant from Chiang Mai University, Thailand.

The authors are thankful to all the participants of the 7th International Conference of Thailand Econometric Society TES’2014 (Chiang Mai, Thailand, January 8–10, 2014) for valuable discussions.

**References**
