

South Dakota State University
**Open PRAIRIE: Open Public Research Access Institutional
Repository and Information Exchange**

Department of Economics Staff Paper Series

Economics

9-1-1992

Teaching Introductory Statistics: A Graphical Relationship Between The Cumulative Distribution Function and the Probability Distribution Function

Dwight Adamnson
South Dakota State University

Scott Fausti
South Dakota State University

Follow this and additional works at: http://openprairie.sdstate.edu/econ_staffpaper

 Part of the [Statistics and Probability Commons](#)

Recommended Citation

Adamnson, Dwight and Fausti, Scott, "Teaching Introductory Statistics: A Graphical Relationship Between The Cumulative Distribution Function and the Probability Distribution Function" (1992). *Department of Economics Staff Paper Series*. Paper 88.
http://openprairie.sdstate.edu/econ_staffpaper/88

This Article is brought to you for free and open access by the Economics at Open PRAIRIE: Open Public Research Access Institutional Repository and Information Exchange. It has been accepted for inclusion in Department of Economics Staff Paper Series by an authorized administrator of Open PRAIRIE: Open Public Research Access Institutional Repository and Information Exchange. For more information, please contact michael.biondo@sdstate.edu.

TEACHING INTRODUCTORY STATISTICS: A GRAPHICAL
RELATIONSHIP BETWEEN THE CUMULATIVE
DISTRIBUTION FUNCTION AND THE
PROBABILITY DISTRIBUTION FUNCTION

by

Dwight W. Adamson and Scott W. Fausti¹

Economic Staff Paper No. 92-5

September 1992

¹ Dwight W. Adamson and Scott W. Fausti are assistant professors of economics at South Dakota State University. The authors wish to thank Chuck Lamberton and John Sondey for their comments and any remaining errors are the responsibility of the authors.

Sixty-one copies of this document were printed by the Economics Department at a cost of \$.84 per document.

ABSTRACT

Introductory statistics textbooks typically develop the concept of continuous random variables with a discussion of only the variables' probability distribution function and omit any discussion of the cumulative distribution function. The cumulative distribution function, however, is useful in developing the concepts of a normal distribution, the standard normal distribution and the probability of a continuous random variable falls within a specific range of values since the standard normal statistical table is derived from the cumulative distribution function. This paper develops a simple graphical relationship between a continuous random variables' cumulative distribution function and its probability distribution function that can be used as a pedagogical tool to more efficiently develop the undergraduate students' conceptual understanding of the topic.

**Teaching Introductory Statistics: A Graphical
Relationship Between the Cumulative
Distribution Function and the
Probability Distribution Function**

Introduction

In many undergraduate business and economic programs, economists often have the responsibility of teaching introductory statistics. As economists teaching statistics, we have found that many students have difficulty grasping the concept of probability distributions for continuous random variables (CRV). We find the concept is best explained by first developing the graphical relationship between CRV's cumulative distribution function and its probability density function. This relationship plays an important role in developing the concepts of: 1) a normal distribution; 2) the standard normal distribution; and 3) how to determine the probability that a normally distributed CRV lies within a specific range of values.

Textbooks developed for the Introduction to Business and Economic Statistics textbook market, such as Keller and Warrack (1991), McClave and Dietrich (1991), and Weiss and Hasset (1991) avoid discussing the cumulative distribution function of a normally distributed CRV. In their textbooks, however, these authors use a standard normal statistical table that is derived from the cumulative distribution function. This approach is not uncommon. In our experience, this "sleight of hand" confuses all but the most sophisticated student. If a student becomes confused about the role of the density function in deriving probability values, then

the concepts used in developing sampling theory and the theory of statistical inference will leave the student bewildered.

Not all textbooks developed for this market ignore the cumulative distribution function. Newbold (1991) develops the mathematical relationship between the density function and the cumulative distribution function and states that the standard normal statistical table is derived from the cumulative distribution function, however, he does not formally develop the relationship graphically. Daniel and Terrell (1992) is another text that at least mentions that the statistical table for the standard normal distribution is derived from the cumulative distribution function. This relationship is graphically developed by Thomas (1986) to complement his mathematical presentation. His textbook, however, is calculus based and targeted for the physical sciences.

For this article we develop a complete, yet simple, graphical presentation of the relationship between the cumulative distribution function and the density function for a CRV that follows a normal distribution. We then introduce the standard normal distribution.

Developing the Graphical Relationship Between the Probability Density Function and Cumulative Distribution Function

Assume X is a continuous random variable which follows a normal distribution:

$$1) X \sim N(\mu, \sigma^2).$$

The probability density function of X is defined mathematically as:

$$2) f_x = \frac{1}{\sqrt{2\pi\sigma}} \exp^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}.$$

The relationship between the variable's density function and its cumulative distribution function, denoted $F_x(x)$, is mathematically defined as:

$$3) f_x(x) = \frac{d}{dx} [F_x(x)], \quad \forall F_x(x) = \int_{-\infty}^{\infty} f_x(x) dx.$$

It is clear that the probability density function is a derived function of the cumulative distribution function. A graphical presentation of this relationship is given in Figure 1.

The mathematical relationship between a normally distributed CRV's cumulative distribution function and its probability density function allows us to find the probability that the CRV lies within a specific range of values. To continue our discussion, the following mathematical properties are now given:

$$4) \int_{-\infty}^{\infty} f_X(x) dx = 1,$$

and therefore,

$$5) F_X(\infty) = 1,$$

it follows that,

$$6) F_X(x_0) = P(X \leq x_0) = \int_{-\infty}^{x_0} f_X(x) dx.$$

In words, the probability that X takes on a value equal to or less than x_0 is given by the cumulative distribution function.

Assume we are interested in the probability of X taking on a value lying in a range between A and B ; $P(A \leq X \leq B)$. A graphical solution to this problem is given in Figure 2. The standard textbook explanation is that the probability, $P(A \leq X \leq B)$, is given by the area under the density function $f_X(x)$ from A to B . As demonstrated in Figure 2, $P(A \leq X \leq B)$ can also be shown graphically with the cumulative distribution function: $P(A \leq X \leq B) = F_X(B) - F_X(A)$.

Deriving the numerical solution, however, requires the use of integral calculus. For the introductory course, this procedure is avoided by invoking the property that any normal distribution can be transformed into a standard normal distribution, denoted $Z \sim N(0,1)$. The distribution of X can be transformed into a standard normal distribution by employing the formula: $Z = (X - \mu) / \sigma$. Transforming the CRV X into CRV Z also transforms the variable's

density function and its cumulative distribution function as graphically shown in Figure 3. The student then derives the solution by using probability values generated by the cumulative distribution function, not the probability density function.

A Graphical Linkage Between the Probability Density Function and the Cumulative Distribution Function

Students in introductory courses often fail to realize that the probability values found in the standard normal statistical table are generated by the cumulative distribution function. It is the cumulative distribution function of the CRV Z and not the probability density function which generates the probabilities used to find the solution to $P(A \leq X \leq B)$.

Assume X is a CRV following a normal distribution and the parameters of the distribution are defined as $X \sim N(16, 16)$. We are interested in the probability that X lies in a range from 12 to 20. The solution is found by transforming the distribution of X into a standard normal distribution:

$$7) P(12 \leq X \leq 20) = P\left[\frac{(12-16)}{4} \leq Z \leq \frac{(20-16)}{4}\right] = P(-1 \leq Z \leq 1) = F_z(1) - F_z(-$$

The solution to the problem is stated in terms of the cumulative distributive function, not the probability density function. This is the most widely used approach in introductory statistics textbooks found in the business and economics curriculum. The formula linkage, however, is never made in most textbooks. A simple graphical example, like the one given in Figure 4, clearly demonstrates the relationship and therefore provides the student

with an intuitive understanding of the relationship between the two functions and their respective roles in generating probability values for a CRV lying within a specific range of values.

REFERENCES

- Daniel, Wayne W., and James C. Terrell. 1992. *Business Statistics for Management and Economics*. 6th ed. Boston, MA: Houghton Mifflin Company.
- Keller, Gerald, and Brain Warrack. 1991. *Essentials of Business Statistics: A Systematic Approach*. Belmont, CA: Wadsworth Publishing Company.
- McClave, James T. and Frank H. Dietrich II. 1991. *Statistics*. 5th ed. San Francisco, CA: Dellen Publishing Company.
- Newbold, Paul. 1991. *Statistics for Business and Economics*. 3rd ed. Englewood Cliffs, NJ: Prentice-Hall, Inc.
- Thomas, John B. 1986. *Introduction to Probability*. New York: Springer-Verlag
- Weiss, Neil A., and Matthew J. Hassett. 1991. *Introductory Statistics*. 3rd ed. Reading, MA: Addison-Wesley Publishing Company, Inc.

FIGURE I

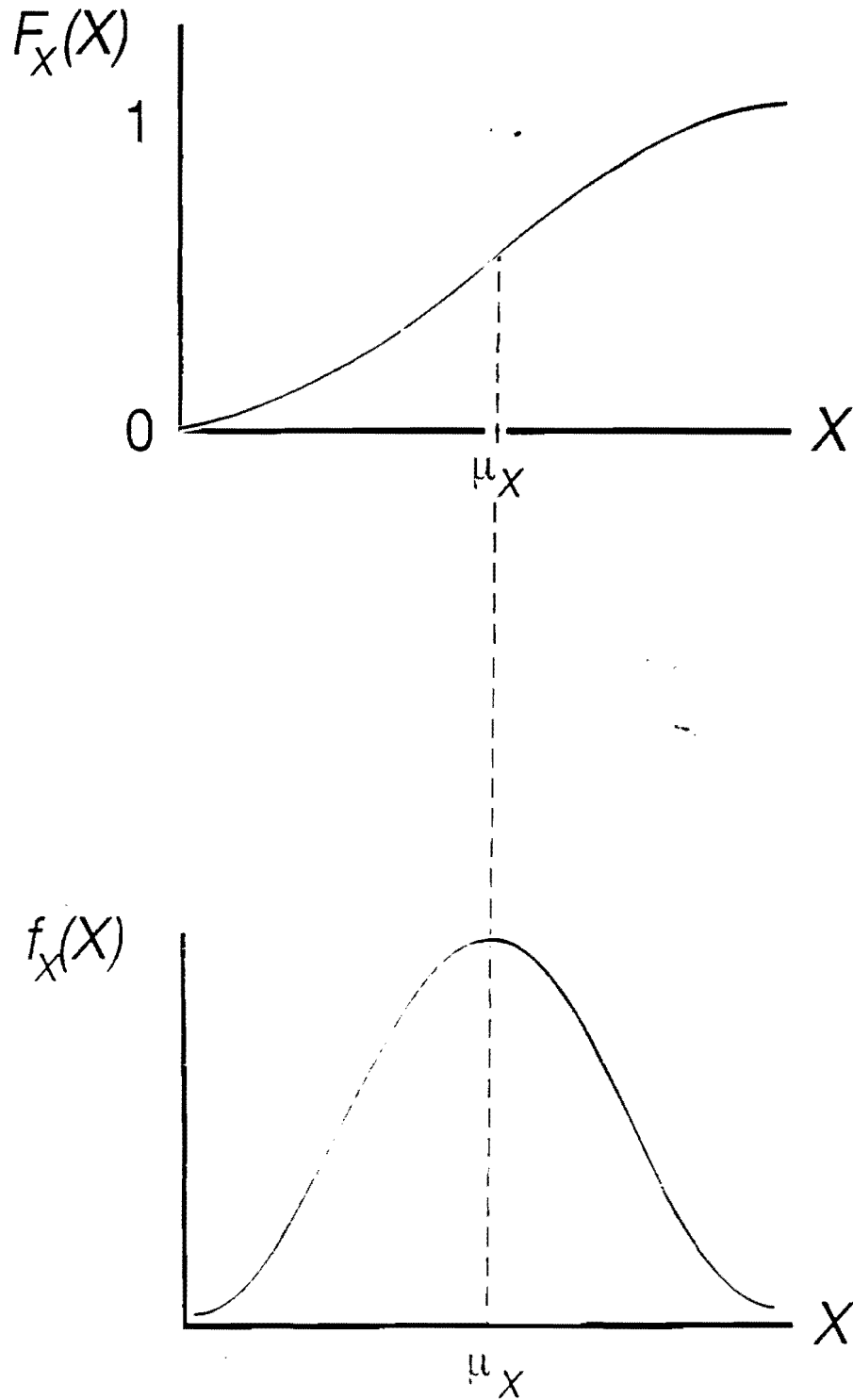


FIGURE II

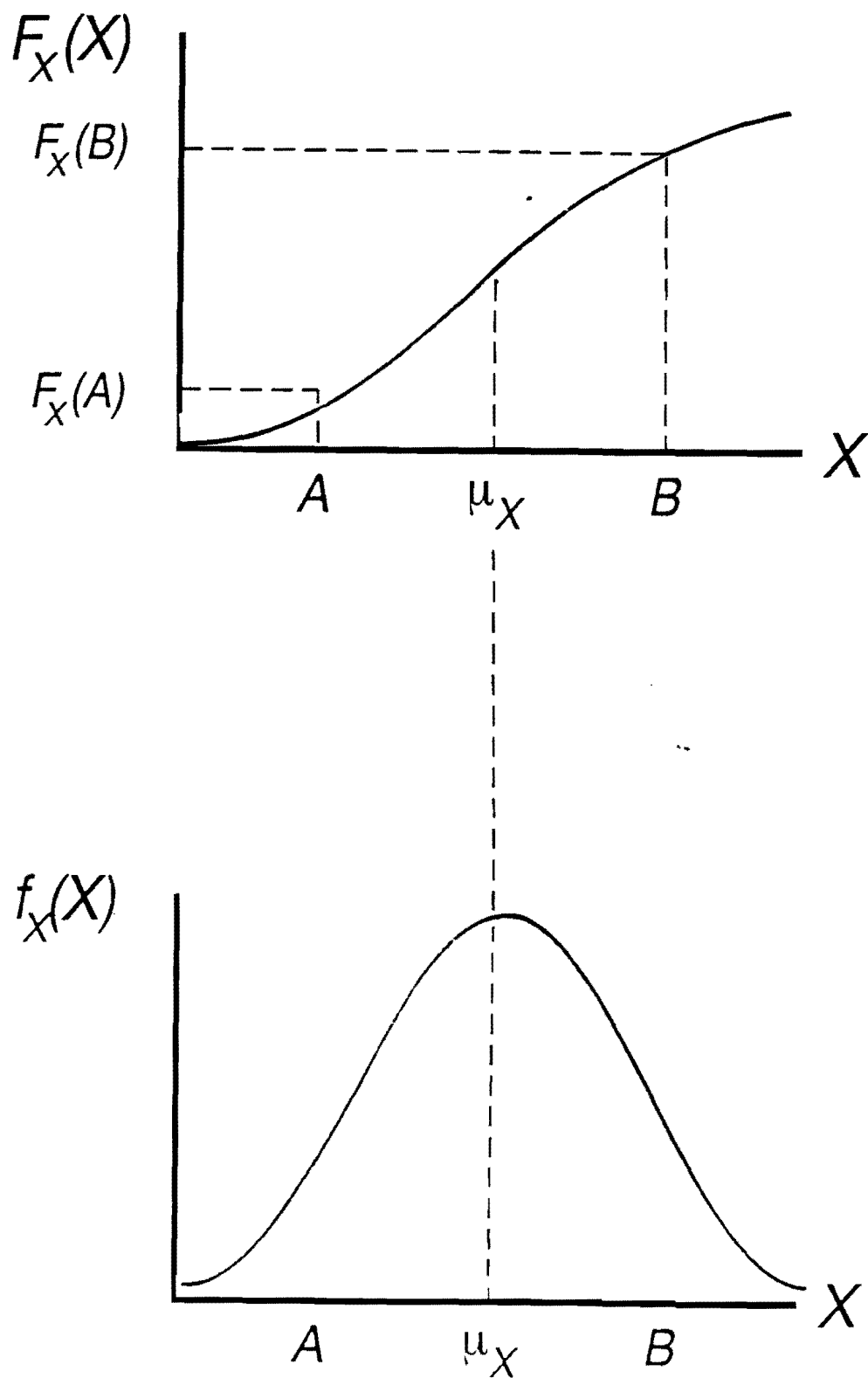


FIGURE III

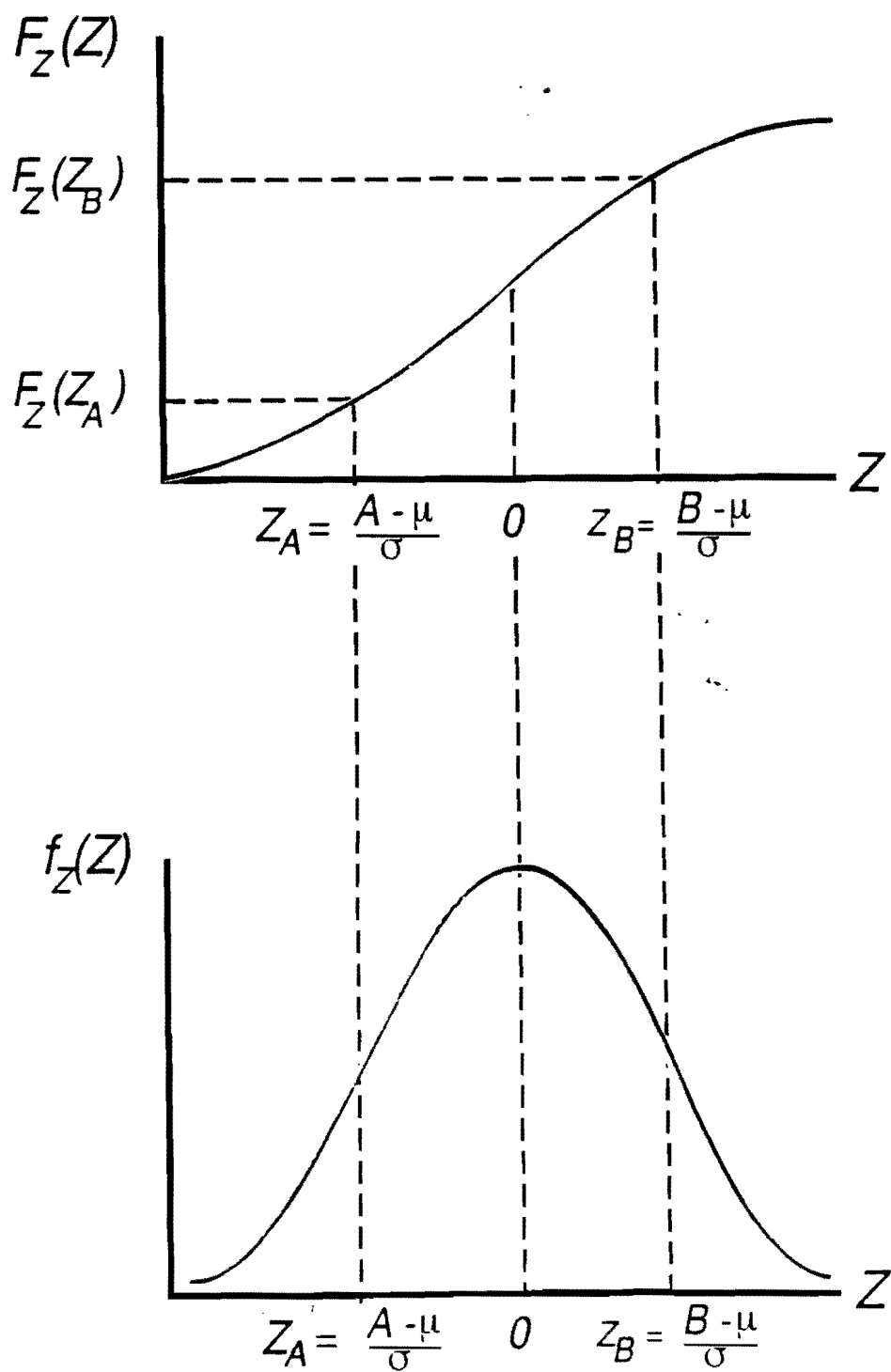


FIGURE IV

