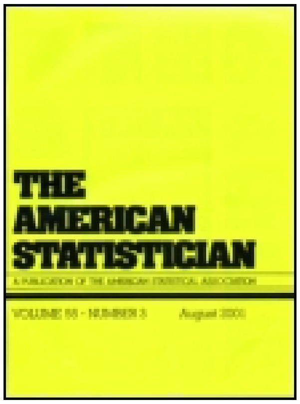

The American Statistician
Publication details, including instructions for authors and subscription information:
http://amstat.tandfonline.com/loi/utas20

A Remark on the Alternative Expectation Formula

Liang Hong [a]

[a] Department of Mathematics , Bradley University , Peoria , IL , 61625

Accepted author version posted online: 20 Sep 2012.Published online: 21 Dec 2012.

To cite this article: Liang Hong (2012) A Remark on the Alternative Expectation Formula, The American Statistician, 66:4, 232-233, DOI: 10.1080/00031305.2012.726934

To link to this article: http://dx.doi.org/10.1080/00031305.2012.726934

A Remark on the Alternative Expectation Formula

Liang Hong

Students in their first course in probability will often see the expectation formula for nonnegative continuous random variables in terms of the survival function. The traditional approach for deriving this formula (using double integrals) is well-received by students. Some students tend to approach this using integration by parts, but often get stuck. Most standard textbooks do not elaborate on this alternative approach. We present a rigorous derivation here. We hope that students and instructors of the first course in probability will find this short note helpful.

KEY WORDS: Expectation; Harmonic analysis; Integration by parts; Nonnegative continuous random variable; Survival function.

1. INTRODUCTION

For a nonnegative continuous random variable X with finite mean, its expectation is given by

$$E[X] = \int_0^{+\infty} x f(x) dx, \tag{1}$$

where $f(x)$ is the density of X. Here and throughout, we follow Ross (2010) and use continuous random variables to refer to random variables with absolutely continuous densities.

Equation (1) is usually given as a definition in a first course in probability, that is, the introductory calculus-based probability. In the same course, students will often encounter an alternative formula:

$$E[X] = \int_0^{+\infty} S(x) dx, \tag{2}$$

where $S(x) = P\{X > x\}$ is the survival function of X. We usually give Equation (2) as an in-class exercise or a homework assignment. Standard textbooks widely used in English-speaking countries such as Bean (2001), Dudewicz and Mishra (1988), Feller (1968), Hoel, Port, and Stone (1971), Hogg, McKean, and Craig (2012), Ross (2010), and Wackerly, Mendenhall, and Scheaffer (2007) often derive Equation (2) from (1) using double integrals:

$$\begin{aligned} E[X] &= \int_0^{+\infty} \left(\int_0^x dy \right) f(x) dx, \\ &= \int_0^{+\infty} \int_y^{\infty} f(x) dx dy, \\ &= \int_0^{+\infty} S(y) dy. \end{aligned}$$

We found this approach is well-received by the students. Nonetheless, each time we taught this class, some students tried to proceed as follows:

$$\begin{aligned} E[X] &= -\int_0^{+\infty} x dS(x), \\ &= -\lim_{x\to\infty} tS(t)\Big|_0^x + \int_0^{+\infty} S(x) dx, \\ &= \int_0^{+\infty} S(x) dx. \end{aligned} \tag{3}$$

This integration-by-parts approach seems to be plausible. However, a close scrutiny of Equation (3) reveals that $\lim_{x\to\infty} xS(x) = 0$ used in the third equality is unjustified. We always asked students to provide a full justification for this if they employed this approach. We found our students often had difficulty giving a tenable argument. Some students simply relied on intuition and failed to recognize the need to justify it. Some believed the survival function $S(x)$ decays faster than the linear term x as $x \to \infty$; but they couldn't give a rigorous argument. Quite a few took recourse to L'Hospital rule as follows:

$$\begin{aligned} \lim_{x\to\infty} xS(x) &= \lim_{x\to\infty} \frac{S(x)}{\frac{1}{x}} \\ &= \lim_{x\to\infty} \frac{-f(x)}{-\frac{1}{x^2}}. \end{aligned}$$

However, they couldn't continue from here. After they failed to supply a correct justification, some students started to question the validity of the approach given by Equation (3). Others tend to believe they will need some heavy machinery from advanced probability for a rigorous argument.

Indeed, Equation (3) is correct and $\lim_{x\to\infty} xS(x) = 0$ can be justified using calculus only. Note that many textbooks on advanced probability such as Billingsley (1995), Chow and Teicher

Liang Hong is Assistant Professor, Department of Mathematics, Bradley University, Peoria, IL 61625 (E-mail: *lhong@bradley.edu*). The author is grateful to Editor Ronald Christensen and the Associate Editor for their comments which led to an improvement on a previous draft.

  DOI: 10.1080/00031305.2012.726934

(1997), Chung (2001), Loève (1977), and Shiryaev (1995) do not give a rigorous argument for Equation (3) in the measure-theoretic setting. We hope students and teachers will benefit from the discussion in this note.

2. THE CORRECT ARGUMENT

We first remark that the integration-by-parts approach is not novel. This idea is widely used in harmonic analysis to study the maximal function. Indeed, the survival function $S(x)$ of a nonnegative random variable X is a special case of the distribution function in harmonic analysis. [Note that the distribution function in harmonic analysis is different from the distribution function in probability.] Interested readers can consult Stein (1970), Stein (1993), and Zhou (1999) for more details. Using a standard argument in harmonic analysis, one can argue that $\lim_{x\to\infty} xS(x) = 0$ as follows: Let P be the distribution of X. Then

$$0 \le xP\{X > x\} \le \int_{\{X>x\}} XdP. \tag{4}$$

Since X has finite mean, $\lim_{x\to\infty} xS(x) = 0$ follows from the absolute continuity of the integral. [See, for example, Royden (1988) or Hewitt and Stromberg (1965).]

However, students in the first course in probability usually don't have any background in analysis. Hence, the above argument can only serve as a reference for instructors. For students, we need to give a calculus-based argument. Such an argument can be easily "translated" from Equation (4). Specifically, we have

$$0 \le xS(x) = xP\{X > x\} = x\int_x^\infty f(t)dt \le \int_x^\infty tf(t)dt.$$

The conclusion follows by taking $x \to \infty$.

Remark. Following the same line of reasoning, one can show that $\lim_{x\to\infty} x^k S(x) = 0$ for any positive integer k, provided X has finite kth moment, that is, $E|X^k| < \infty$. This will lead to the following alternative kth moment formula:

$$E[X^k] = k\int_0^{+\infty} x^{k-1}S(x)dx. \tag{5}$$

Obviously, Equation (2) is a special case of Equation (5). Indeed, Equation (5) is just a special case of a well-known property of the distribution function in harmonic analysis. [See, for example, Stein (1970).]

[Received June 2012. Revised August 2012.]

REFERENCES

Bean, M. A. (2001), *Probability: The Science of Uncertainty With Applications to Investments, Insurance, and Engineering*, Belmont, CA: Brooks Cole. [232]

Billingsley, P. (1995), *Probability and Measure* (3rd ed.), Hoboken, NJ: Wiley. [232]

Chow, Y. S., and Teicher, H. (1997), *Probability Theory: Independence, Interchangeability, Martingales* (3rd ed.), New York: Springer. [233]

Chung, K. L. (2001), *A Course in Probability* (3rd ed.), Burlington, MA: Academic Press, Elsevier. [233]

Dudewicz, E. J., and Mishra, S. N. (1988), *Modern Mathematical Statistics*, Hoboken, NJ: Wiley. [232]

Feller, W. (1968), *An Introduction to Probability Theory and Its Applications,* (Vol. 1, 3rd ed.), Hoboken, NJ: Wiley. [232]

Hewitt, E., and Stromberg, K. (1965), *Real and Abstract Analysis*, New York: Springer. [233]

Hoel, P. G., Port, S. C., and Stone, C. J. (1971), *Introduction to Probability Theory*, Boston, MA: Houghton. [232]

Hogg, R. V., McKean, J. W., and Craig, A. T. (2012), *Introduction to Mathematical Statistics* (7th ed.), Upper Saddle River, NJ: Prentice Hall. [232]

Loève, M. (1977), *Probability Theory I, II* (4th ed.), New York: Springer. [233]

Ross, S. (2010), *A First Course in Probability* (8th ed.), Upper Saddle River, NJ: Prentice Hall. [232]

Royden, H. L. (1988), *Real Analysis* (3rd ed.), Upper Saddle River, NJ: Prentice Hall. [233]

Shiryaev, A. N. (1995), *Probability* (2nd ed.), New York: Springer. [233]

Stein, E. M. (1970), *Singular Integrals and Differentiability Properties of Functions*, Princeton, NJ: Princeton University Press. [233]

——— (1993), *Harmonic Analysis: Real-Variable Methods, Orthogonality, and Oscillatory Integrals*, Princeton, NJ: Princeton University Press. [233]

Wackerly, D., Mendenhall, W., and Scheaffer, R. L. (2007), *Mathematical Statistics With Application* (7th ed.), Belmont, CA: Brooks Cole. [232]

Zhou, M. Q. (1999), *Lecture Notes on Harmonic Analysis* [In Chinese], Beijing, China: Beijing University Press. [233]