ATINER CONFERENCE PRESENTATION SERIES No: EMS2017-0050

ATINER's Conference Paper Proceedings Series
EMS2017-0050
Athens, 15 March 2018

Teaching of Expected Value of a Random Variable Calculation: The Darth Vader Rule

Krzysztof Ostaszewski

Athens Institute for Education and Research
8 Valaoritou Street, Kolonaki, 10683 Athens, Greece

ATINER's Conference Paper Proceedings Series
EMS2017-0050
Athens, 15 March 2018
ISSN: 2529-167X

Krzysztof Ostaszewski, Professor of Mathematics and Actuarial Program Director, Illinois State University, Department of Mathematics, U.S.A

Teaching of Expected Value of a Random Variable Calculation: The Darth Vader Rule

ABSTRACT

The concept of an expected value is presented in teaching of probability in a way that is dramatically different than most practical calculations in insurance are done. In this work, we show that standard instruction in probability would be greatly enriched by adding the approach of calculating expected value as the integral of the survival function (assuming the random variable cosnidered is non-negative almost surely). This simple rule, which we call *The Darth Vader Rule,* empowers practical calculations in insurance applications, inckuding insurance contract modifications such as deductible, or policy limit, and also including reinsurance contracts.

Keywords: expected value, survival function, random variable

Acknowledgment:

Introduction

A standard method of calculation of the expected value of a future lifetime of a person in actuarial mathematics is:

$$E(T) = \overset{\circ}{e}_x = \int_0^{\infty} {}_t p_x \, dt,$$

where T is the random future lifespan of a person aged x, and ${}_t p_x = s_T(t) = \Pr(T > t)$ is the *survival function* of the random variable T. Similarly, when K is random future number of future complete years lived by a person aged x, then

$$E(K) = e_x = \sum_{k=1}^{+\infty} {}_k p_x = {}_1 p_x + {}_2 p_x + {}_3 p_x + \dots .$$

But this does not look much like the standard definition of an expected value of a random variable. The standard definition is, for a continuous random variable T with probability density function $f_T(t)$,

$$E(T) = \int_{-\infty}^{\infty} t \cdot f_T(t)\, dt,$$

while for a discrete random variable T with probability function $f_T(t)$,

$$E(T) = \sum_{t: f_T(t) > 0} t \cdot f_T(t).$$

Why do Actuaries do this Strange thing with the Survival Function?

When an event has probability one, we say that it happens *almost surely*. Consider a random variable X that is non-negative almost surely, whose expected value exists. If X is continuous, then by using integration by parts, we obtain

$$E(X) = \int_{-\infty}^{+\infty} x \cdot f_X(x)\, dx = \int_0^{+\infty} x \cdot f_X(x)\, dx = \underbrace{\left| \begin{array}{ll} u = x & v = -s_X(x) \\ du = dx & dv = f_X(x)\, dx \end{array} \right|}_{\text{INTEGRATION BY PARTS}} =$$

$$= x \cdot \left(-s_X(x)\right)\Big|_{x=0}^{x \to +\infty} + \int_0^{+\infty} s_X(x)\, dx = -\lim_{x \to \infty} x \cdot s_X(x) + \int_0^{+\infty} s_X(x)\, dx = \int_0^{+\infty} s_X(x)\, dx.$$

Note that

$$0 \le \lim_{x \to \infty} x \cdot \underbrace{s_X(x)}_{= \int_x^{+\infty} f_X(t)\, dt} = \lim_{x \to \infty} x \cdot \underbrace{\int_x^{+\infty} f_X(t)\, dt}_{\text{Note that } x < t < +\infty} \le \underbrace{\lim_{x \to \infty} \int_x^{+\infty} t \cdot f_X(t)\, dt = 0}_{\text{because } E(X) = \int_0^{+\infty} t \cdot f_X(t)\, dt \text{ exists}},$$

so that we can conclude in the above reasoning that

$$\lim_{x\to\infty} x \cdot s_X(x) = 0.$$

This implies that $E(X) = \int_0^{+\infty} s_X(x)dx$ as long as X is continuous and non-negative almost surely. What if X is discrete and non-negative almost surely? Then, by definition,

$$E(X) = \sum_{x \in R} x \cdot f_X(x).$$

Assume first that we can form

$$x_0 = 0 < x_1 < x_2 < \cdots$$

the sequence of values where the probability function f_X is positive, but put $x_0 = 0$ at the beginning of this sequence regardless of whether X attains that value with probability zero or positive probability. Then, because of the step-function structure of the survival function:

$$\begin{aligned}
E(X) &= x_0 \cdot f_X(x_0) + x_1 \cdot f_X(x_1) + x_2 \cdot f_X(x_2) + x_3 \cdot f_X(x_3) + \ldots = \\
&= \underbrace{x_0}_{=0} \cdot \underbrace{\left(f_X(x_0) + f_X(x_1) + f_X(x_2) + f_X(x_3) + \ldots\right)}_{=1} + \\
&\quad + (x_1 - x_0) \cdot \left(f_X(x_1) + f_X(x_2) + f_X(x_3) + \ldots\right) + \\
&\quad + (x_2 - x_1) \cdot \left(f_X(x_2) + f_X(x_3) + f_X(x_4) + \ldots\right) + \\
&\quad + (x_3 - x_2) \cdot \left(f_X(x_3) + f_X(x_4) + f_X(x_5) + \ldots\right) + \ldots \\
&= 0 \cdot 1 + (x_1 - x_0) \cdot s_X(x_0) + (x_2 - x_1) \cdot s_X(x_1) + (x_3 - x_2) \cdot s_X(x_2) + \ldots = \\
&= \int_0^{+\infty} s_X(x)dx.
\end{aligned}$$

The above proof assumes that the point masses can be put in an increasing sequence, and there are discrete distributions that violate that assumption. So the proof is not complete for discrete random variables under this approach. However, Edwin Hewitt (1960) proved integration by parts formula for the Lebesgue-Stieltjes integral. The formula is given in the following form: Let μ and ν be measures defined for Borel subsets of $[a,b]$ and let

$$M(t) = \frac{\mu([a,t]) + \mu([a,t))}{2},$$

$$N(t) = \frac{\nu([a,t]) + \nu([a,t))}{2}.$$

Then

$$\int_a^b M(t)d\nu(t) + \int_a^b N(t)d\mu(t) = \mu([a,b]) \cdot \nu([a,b]).$$

Consider a random variable X defined on the interval $\left[0,+\infty\right)$. Assume that $E\left(X\right)$ exists and f_X is well-defined. Defined the measures μ and ν by the following conditions:

$$\mu\left(\left[0,x\right]\right)=1-s_X\left(x\right),$$

$$\nu\left(\left[0,x\right]\right)=x.$$

Then

$$M\left(x\right)=\frac{1}{2}\left(\mu\left(\left[0,x\right]\right)+\mu\left(\left[0,x\right)\right)\right)=1-s_X\left(x\right),$$

$$N\left(x\right)=\frac{1}{2}\left(\nu\left(\left[0,x\right]\right)+\nu\left(\left[0,x\right)\right)\right)=x.$$

Then Hewitt's formula implies, on an interval of the form [0, b]

$$\int_0^b\left(1-s_X\left(x\right)\right)dx+\int_0^b xf_X\left(x\right)dx=\left(\int_0^b f_X\left(x\right)dx\right)\cdot\left(b-0\right),$$

or

$$\int_0^b xf_X\left(x\right)dx=\int_0^b s_X\left(x\right)dx+b\int_0^b f_X\left(x\right)dx-b=\int_0^b s_X\left(x\right)dx-b\int_b^{+\infty} f_X\left(x\right)dx.$$

Note that

$$0\le b\int_b^{+\infty} f_X\left(x\right)dx=\int_b^{+\infty} bf_X\left(x\right)dx\le\int_b^{+\infty} xf_X\left(x\right)dx.$$

Because $E\left(X\right)$ exists

$$\lim_{b\to\infty}\int_b^{+\infty} xf_X\left(x\right)dx=0.$$

We conclude that

$$\int_0^b xf_X\left(x\right)dx=\int_0^b s_X\left(x\right)dx.$$

Muldowney, Ostaszewski and Wojdowski (2012) also provide a proof based on a generalized integration by parts formula for the Henstock-Stieltjes integral.

Mixed Random Variables

One more group of random variables, which are a source of confusion for students, but are of great importance in insurance applications, are *mixed random variables.* The simplest case of this phenomenon is when we have a random variable T that is equal to a variable T_1 with probability α and a variable T_2 with probability $1-\alpha$. For such a random variable

$$s_T(t) = \Pr(T > t) = \alpha \Pr(T_1 > t) + (1-\alpha)\Pr(T_2 > t) = \alpha s_{T_1}(t) + (1-\alpha) s_{T_2}(t),$$

so that for a random variable that is non-negative almost surely

$$\alpha E(T_1) + (1-\alpha) E(T_2) =$$

$$= \alpha \int_0^{\infty} s_{T_1}(t)\,dt + (1-\alpha)\int_0^{\infty} s_{T_2}(t)\,dt = \int_0^{\infty} s_T(t)\,dt = E(T).$$

The more general structure of a mixed random variable is when a continuous random variable T is a continuous random variable with probability density function (PDF) $f_T(t,\Lambda)$, where Λ is a random variable with PDF $h_\Lambda(\lambda)$ then

$$E(T) = \int_{-\infty}^{+\infty}\left(\int_{-\infty}^{+\infty} t \cdot f_T(t,\lambda)\,dt\right)d\lambda = \int_{-\infty}^{+\infty}\underbrace{\left(\int_{-\infty}^{+\infty} t \cdot f_T(t|\Lambda=\lambda)\,dt\right)}_{=E(T|\Lambda=\lambda)} \cdot h_\Lambda(\lambda)\,d\lambda.$$

In the case of a random variable that is non-negative almost surely, assuming that for the conditional distribution the rule that the expected value is the integral of the survival function results in the same rule for the mixed distribution:

$$E(T) = \int_{-\infty}^{+\infty}\underbrace{\left(\int_{0}^{+\infty} t \cdot f_T(t|\Lambda=\lambda)\,dt\right)}_{=E(T|\Lambda=\lambda)} \cdot h_\Lambda(\lambda)\,d\lambda =$$

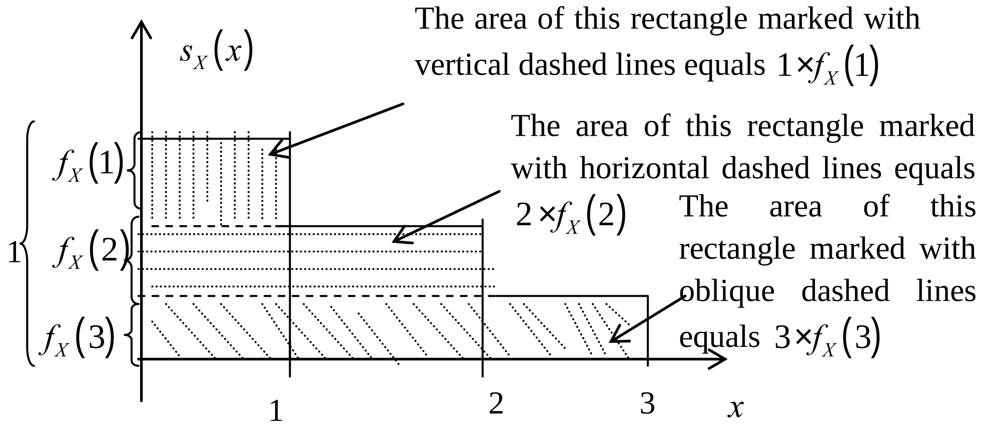
$$= \int_{-\infty}^{+\infty}\int_{0}^{+\infty} s_T(T|\Lambda=\lambda)\,dt \cdot h_\Lambda(\lambda)\,d\lambda =$$

$$= \int_{0}^{+\infty}\underbrace{\int_{-\infty}^{+\infty} \Pr(T > t|\Lambda=\lambda)\, h_\Lambda(\lambda)\,d\lambda}_{=\Pr(T>t)} \cdot dt = \int_{0}^{+\infty} s_T(t)\,dt.$$

This means that for any random variable X, which is nonnegative almost surely, and whose expected value exists, the expected value equals the integral of the survival function. We will call this important rule: *The Darth Vader Rule* (term not commonly used in other books). In the case when X is discrete and assumes only positive integer values, we have the following special rule:

$$E(X) = 1 \cdot s_X(0) + 1 \cdot s_X(1) + 1 \cdot s_X(2) + \cdots = \sum_{n=0}^{+\infty} \Pr(X > n) = \sum_{n=1}^{+\infty} \Pr(X \geq n).$$

In this special case, we can also see how Darth Vader Rule makes an intuitive sense from the figure below, showing a calculation of the expected value of a random variable that assumes only three values: 1, 2, and 3. In the figure how the area under the graph of the survival function, which equals the integral of that survival function, can be decomposed into pieces that add up to those used in the standard calculation of the expected value of a discrete random variable.



Let us not also that Feller (1968) also notes that for a random variable whose n-th moment exists

$$E\left(X^n\right) = n\int_0^{+\infty} x^{n-1} s_X(x)\,dx.$$

Using the Fubini Theorem

We should note that the presentation of the Darth Vader Rule for the mixed distribution involved the use of the Fubini Theorem. In fact, the Darth Vader Rule is a direct consequence of the Fubini Theorem. For a continuous random variable T, non-negative almost surely, with PDF $f_T(T)$, we have

$$E(X) = \int_0^{+\infty} x f_X(x)\,dx = \int_0^{+\infty}\left(\int_0^x dt\, f_X(x)\right)dx = \int_0^{+\infty}\left(\int_t^{+\infty} f_X(x)\,dx\right)dt = \int_0^{+\infty} s_X(t)\,dt.$$

For a discrete random variable with probability function $f_T(T)$, we have

$$E(X) = \sum_{i=1}^{\infty} x_i \cdot f_X(x_i) = \sum_{i=1}^{\infty}\left(\int_0^{x_i} dt\right)\cdot f_X(x_i) = \int_t^{\infty}\left(\sum_{x_i>t} f_X(x_i)\right)dt = \int_0^{+\infty} s_X(t)\,dt.$$

For the mixed distributions, the reasoning we presented before applies, but we can also use the generalized Fubini Theorem for the Henstock Integral, proved by Ostaszewski (1986), also see Kurzweil (1957) and Henstock (1988).

Example

A dart is thrown at a dartboard with radius of 7 centimeters. The point that the dart hits is uniformly distributed on the circular dartboard. Find the expected distance, in centimeters, of that point from the center of the dartboard.

Solution.
T is nonnegative with probability one and for $t < 7$

$$s_T(t) = \Pr(T > t) = \Pr\left(X^2 + Y^2 > t^2\right) =$$

$$= \frac{\text{Area of circle with radius 7} - \text{Area of circle with radius } t}{\text{Area of circle with radius 7}} =$$

$$= \frac{49 - t^2}{49} = 1 - \frac{t^2}{49}.$$

while

$$s_T(t) = \Pr(T > t) = 0$$

for $t \geq 7$. Therefore, using the Darth Vader Rule, we obtain

$$E(T) = \int_0^7 s_T(t)\,dt = \int_0^7 \left(1 - \frac{t^2}{49}\right) dt = \left(t - \frac{t^3}{3 \cdot 49}\right)\Bigg|_{t=0}^{t=7} = 7 - \frac{7^3}{3 \cdot 49} = \frac{14}{3}.$$

Of course, this problem can also be solved the traditional way, with the use of polar coordinates transformation, but that is a far more laborious approach.

Example

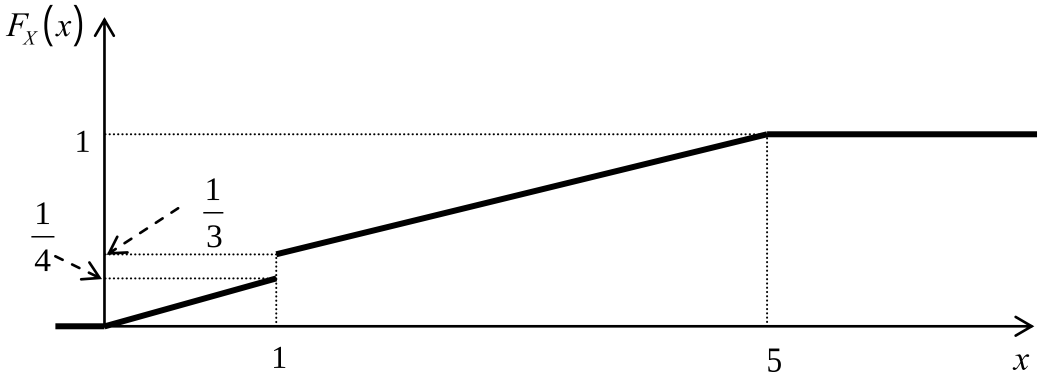
You are given the cumulative distribution function of a random variable X is 0 for $x \leq 0$, and for positive values of x:

$$F_X(x) = \begin{cases} \frac{1}{4}x & \text{for } 0 \leq x < 1, \\ \frac{1}{3} & \text{for } x = 1, \\ \frac{x+1}{6} & \text{for } 1 < x < 5, \\ 1 & \text{for } x \geq 5. \end{cases}$$

Find the expected value of X.

Solution.

Note that X is nonnegative almost surely and the graph of CDF is



The expected value is calculated by noting that $s_X(x)=1-F_X(x)$, and then evaluating areas between the CDF graph and the horizontal line at the level of 1

$$E(X)=1\times\frac{1+\frac{3}{4}}{2}+4\times\frac{\frac{2}{3}}{2}=\frac{7}{8}+\frac{4}{3}=\frac{21}{24}+\frac{32}{24}=\frac{53}{24}.$$

Conclusions

The approach in calculating expected value of a random variable presented here, and termed the *Darth Vader Rule,* is a relative simple rearrangements implied by the generalized integration by parts, or the generalized Fubini Theorem, but it is a nice didactic tool in teaching about the expected value, as it provides a relative simple and quick way of performing many calculations. It is also widely used in insurance applications.

References

Bowers, N.L, Gerber, H.U., Hickman, J.C., Jones, D.A., and Nesbitt, C.J. 1997. *Actuarial Mathematics*, 2nd Edition. Society of Actuaries; Schaumburg, IL.

Feller, W. 1966. *An Introduction to Probability Theory and its Applications.* Vol. 2, John Wiley & Sons, Inc. New York, 1966.

Henstock, R. 1988. *Lectures in the Theory of Integration.* World Scientific, Singapore.

Hewitt, E. 1960. Integration by parts for Stieltjes integrals. *Amer.Math.Monthly* 67, 419–423. http://www.jstor.org/stable/2309287.

Kurzweil, J. 1957. Generalized ordinary differential equations and continuous dependence on a parameter. *Czechoslovak Math. J.* 82, 418–446.

Muldowney, P., Ostaszewski, K., and Wojdowski, W. 2012. The Darth Vader Rule, *Tatra Mountains Mathematical Publications,* Volume 52, Issue 1, November 2012, pp. 53-63.

Ostaszewski, K. 1986. *Henstock Integration in the Plane.* Memoirs of the American Mathematical Society No. 353, Providence, RI.