DRAFT

Evaluation of Mean and Variance Integrals without Integration

A. H. Joarder and M. H. Omar*

Department of Mathematical Sciences, King Fahd University of Petroleum & Minerals, Dhahran 31261, Saudi Arabia,

The mean and variance of some continuous distributions, especially exponentially decreasing probability distribution and normal distribution are considered. Since they involve integration by parts, many students do not feel comfortable. In this note we demonstrate a technique of deriving mean and variance only mostly through differential calculus. The general nature of the technique exhibits its potential for wider applications.

1. Introduction

In the service courses in engineering statistics, sometimes we need to evaluate mean and variance of some continuous distributions, especially exponentially decreasing probability density function and normal probability density function. Since these involve integration by parts and/or the use of L'Hospital's rule, many students face difficulty. In addition, instructors also face some sort of a teaching digression. In this note, we assume that the probability density function (pdf) in question has at least one continuous parameter. Since any probability density function integrates to unity, we call this integral a density identity (DI) in the parameters of the distribution. We derive the mean and variance integrals by repeatedly differentiating the DI with respect to the parameters.

The probability density function of an exponential random variable X, say the life time of a battery, is given by

$$f(x) = \lambda e^{-\lambda x}, \quad 0 < x < \infty, 0 < \lambda < \infty. \tag{1}$$

For the motivation of this distribution see Scheaffer and McClave [2]. Suppose that we want to know the expected life time of the batteries and also the variance of the life times of the batteries. The mean and variance integrals are given by

$$\begin{aligned} E(X) &= \int_0^\infty x f(x) dx \\ &= \int_0^\infty x \left(\lambda e^{-\lambda x}\right) dx \end{aligned} \tag{2}$$

and

*Corresponding author.
Emails: anwarj@kfupm.edu.sa, omarmh@kfupm.edu.sa

$$V(X) = \int_0^\infty \left(x - E(X)\right)^2 f(x)dx$$

$$= \int_0^\infty x^2 f(x)dx - \left(E(X)\right)^2 \qquad (3)$$

respectively.

Since the mean and variance integrals given by (2) and (3) involve integration by parts, students find these difficult. Since the pdf in (1) integrates to 1,

$$\int_0^\infty f(x)dx = 1$$

$$i.e. \int_0^\infty \lambda e^{-\lambda x} dx = 1. \qquad (4)$$

The above integral will be called the density identity, and will be used to evaluate the mean and variance integrals given by (2) and (3). We repeatedly differentiate the density identity with respect to λ to get new identities that are in turn exploited to evaluate the mean and variance integrals.

2. The Method

The method is described by the exponential and normal distributions. The following lemma is obvious.

Lemma 2.1 Let x and λ be nonnegative. Then we have

$$(i)\ \frac{d}{d\lambda}\left(e^{-\lambda x}\right) = -xe^{-\lambda x},$$

$$(ii)\ \frac{d}{d\lambda}\left(\lambda e^{-\lambda x}\right) = (-\lambda x + 1)e^{-\lambda x}.$$

Example 2.1 Let X have the exponential pdf given by (1). Then we have

$$(i)\ \int_0^\infty x\left(\lambda e^{-\lambda x}\right)dx = \frac{1}{\lambda},$$

$$(ii)\ \int_0^\infty x^2\left(\lambda e^{-\lambda x}\right)dx = \frac{2}{\lambda^2}.$$

Proof. (i) Differentiating the density identity (4) with respect to λ (see Lemma 2.1 (ii)), we have

$$\int_0^\infty (-\lambda x + 1)e^{-\lambda x} dx = 0$$

which simplifies to

$$\int_0^\infty \lambda x e^{-\lambda x} dx = \int_0^\infty e^{-\lambda x} dx$$

$$= \frac{1}{\lambda} \quad \text{by (4),}$$

which is part (i).

That is, $E(X) = \frac{1}{\lambda}$, which is the mean of the exponential distribution.

(ii) Differentiating the identity (i) in this example again (See Lemma 2.1 (ii)) with respect to λ, we have

$$\int_0^\infty x\left((-\lambda x + 1)e^{-\lambda x}\right)dx = \frac{-1}{\lambda^2}.$$

Then by Example 2.1 (i), we have

$$\int_0^\infty x^2 \lambda e^{-\lambda x} dx = \int_0^\infty x e^{-\lambda x} dx + \frac{1}{\lambda^2}$$

$$= \frac{1}{\lambda^2} + \frac{1}{\lambda^2}.$$

$i.e. \quad E\left(X^2\right) = \frac{2}{\lambda^2}.$

Given these results, the variance integral (3) is given by

$V(X) = E(X^2) - \left(E(X)\right)^2$

i.e. $\quad \sigma^2 = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}.$

Example 2.2 Let X have the normal distribution $N(\mu, 1)$ with pdf

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x-\mu)^2}, \ -\infty < x < \infty, \ -\infty < \mu < \infty.$$

Then $E(X) = \mu.$

Solution. The density identity is given by

$$\int_{-\infty}^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x-\mu)^2} dx = 1.$$

Differentiating both sides of the above identity with respect to μ, we have

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x-\mu)^2} \left(-(2)\frac{1}{2}(x-\mu)(-1) \right) dx = 0,$$

i.e. $\int_{-\infty}^{\infty} (x-\mu) f(x) dx = 0$,

or, $E(X-\mu) = 0.$

Example 2.3 For a general normal distribution $N(\mu, \sigma^2)$ with pdf

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/(2\sigma^2)}, \ -\infty < x < \infty, \ -\infty < \mu < \infty, \ 0 < \sigma < \infty,$$

find the mean and variance.

Solution: The density identity is given by

$$\int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/(2\sigma^2)} dx = 1. \tag{5}$$

Differentiating both sides of the above identity with respect to μ, we have

$$\int_{-\infty}^{\infty} \left(\frac{x-\mu}{\sigma^2} \right) f(x) dx = 0,$$

or, $E(X-\mu) = 0$.
That is, $E(X) = \mu$.

Again differentiating both sides of the density identity (5) with respect to σ we have

$$\int_{-\infty}^{\infty} \left(\frac{(x-\mu)^2 - \sigma^2}{\sigma^3} \right) f(x) dx = 0,$$

or, $E\left((X-\mu)^2 - \sigma^2 \right) = 0$.
That is, $V(X) = \sigma^2$.

Thus, the mean and variance of a normal distribution $N(\mu,\sigma^2)$ are given by $E(X)=\mu$ and $V(X)=\sigma^2$ respectively. In particular, when X has the normal distribution $N(0,\sigma^2)$ with pdf

$$f(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{-x^2/(2\sigma^2)},\ -\infty<x<\infty,\ 0<\sigma<\infty,$$

it follows from Example 2.3 that $E(X)=0$ and $V(X)=\sigma^2$.

3. An Application

Example 3.1 Let the continuous random variable Y denote the diameter of a hole drilled in a sheet metal component. The target diameter is 12.5 millimeters. Most random disturbances to the process result in larger diameters. Historical data show that the distribution of Y can be modeled by a probability density function

$$f(y)=20e^{-20(y-12.5)},\ \ y\geq 12.5 \qquad (6)$$

(Montgomery, Runger, and Hubele [1], p. 59). If a part with a diameter larger than 12.6 millimeters is scrapped, the probability that a part is scrapped is given by

$$\begin{aligned} P(Y>12.60) &= \int_{12.6}^{\infty} f(y)dy \\ &= e^{250}e^{-252} \\ &\approx 0.135. \end{aligned}$$

Thus for a sample of size n, approximately $0.135n$ parts should be scrapped. Note that the number of scrapped parts with larger diameters holes will have a binomial distribution $B(n,p)$ where $p\approx 0.135$ is the probability that a part with larger diameters (Y > 12.60) will be scrapped.

Suppose that we want to know the expected value and variance of the diameters of holes drilled in a sheet metal component. The mean is given by

$$\begin{aligned} E(Y) &= \int_{12.5}^{\infty} yf(y)dy \\ &= \int_{12.5}^{\infty} y\left(20e^{-20(y-12.5)}\right)dy \end{aligned} \qquad (7)$$

Letting $y = x + 12.5$, we have

$$E(Y) = \int_0^\infty (x+12.5)\left(20e^{-20x}\right)dx$$
$$= \int_0^\infty x\left(20e^{-20x}\right)dx + 12.5\int_0^\infty \left(20e^{-20x}\right)dx$$
$$= \int_0^\infty x\left(20e^{-20x}\right)dx + 12.5.$$

Then by Lemma 2.1, we have

$$E(Y) = \frac{1}{20} + 12.5 = 12.55.$$

The variance is given by

$$V(Y) = \int_0^\infty y^2 f(y)dy - 12.55^2 \tag{8}$$

But $E(Y^2) = \int_{12.5}^\infty y^2 f(y)dy$

$$= \int_{12.5}^\infty y^2\left(20e^{-20(y-12.5)}\right)dy$$
$$= \int_0^\infty (x+12.5)^2\left(20e^{-20x}\right)dx. \tag{9}$$

By Example 2.1, the integral in (9) is evaluated as

$$\int_0^\infty x^2\left(20e^{-20x}\right)dx + 2(12.5)\int_0^\infty x\left(20e^{-20x}\right)dx + (12.5)^2\int_0^\infty \left(20e^{-20x}\right)dx$$
$$= \frac{2}{20^2} + 2(12.5)\frac{1}{20} + 12.5^2(1)$$
$$= 157.505$$

Therefore the variance in (8) is given by $V(Y) = 157.505 - 12.5^2 = 1.255$ (cf. Montgomery, Runger, and Hubele [1], pp. 59-62).

We remark that the method discussed here is easily applied to most other continuous distributions. In case a pdf does not explicitly have a continuous parameter, we can formally insert it into the pdf and apply the technique discussed.

Acknowledgement

The authors gratefully acknowledge King Fahd University of Petroleum & Minerals, Saudi Arabia for providing excellent research facilities.

References

[1] Montgomery, D.C; Runger G.C. and Hubele, N.F. (1998). *Engineering Statistics*. John Wiley, New York.

[2] Scheaffer, R.L. and McClave, J.T. (1995). *Probability and Statistics for Engineering*. Duxbury Press.