

Wednesday, November 04, 2009

Derivation of isosceles trapezoidal distribution

Raghu N Kacker¹ and James F Lawrence^{2,1}

¹ National Institute of Standards and Technology, Gaithersburg, MD 20899, USA

² George Mason University, Fairfax, VA 22030, USA

Emails:

raghu.kacker@nist.gov

james.lawrence@nist.gov

Abstract

It is known that, if the mid-point of a rectangular distribution is specified, the half-width is inexactly known, and the state of knowledge about the half-width may be represented by a narrower rectangular distribution then the resulting distribution looks like an isosceles trapezoid whose sloping sides are curved (an isocurvilinear trapezoid). What sort of probability distribution must the half-width have for the resulting distribution to be an isosceles trapezoid? We show that if the distribution of the half-width is a variation of the rectangular distribution whose top is sloping (a sloping top distribution) then the resulting distribution is isosceles trapezoid.

1. Introduction

The Guide to the Expression of Uncertainty in Measurement (GUM) [1, section 4.3.9] suggests that the state of knowledge about a quantity whose range of values is inexactly known is better represented by an isosceles trapezoid than a rectangular distribution. If the mid-point of a rectangular distribution is specified, the half-width is inexactly known, and the state of knowledge about the half-width may be represented by a rectangular distribution then the resulting distribution is isocurvilinear trapezoid [2]. When the uncertainty in the specification of the half-width is not excessive, the isocurvilinear trapezoid is well approximated by an isosceles trapezoid. Still a question remains: what sort of probability distribution must the half-width have for the resulting distribution to be an isosceles trapezoid suggested in the GUM? We have determined that if the distribution of the half-width is a variation of the rectangular distribution whose top is sloping (a sloping top distribution) then the resulting distribution is isosceles trapezoid.

2. Derivation of isosceles trapezoid from a sloping top distribution

Suppose the probability distribution of a variable X is rectangular on the interval $(\mu - Z, \mu + Z)$, whose mid-point μ is specified but the half-width Z is inexactly known. Then the conditional probability density function (pdf) of X given $Z = z$ is

$$f_{X|Z}(x|z) = \frac{1}{2z} \quad \text{for } \mu - z \leq x \leq \mu + z. \quad (1)$$

Suppose the half-width Z has the pdf

$$g(z) = \frac{2z}{\beta^2 - \alpha^2} \quad \text{for } \alpha \leq z \leq \beta. \quad (2)$$

Then as discussed in the appendix, the unconditional pdf $f(x)$ of X is

$$f(x) = \begin{cases} \frac{x - (\mu - \beta)}{\beta - \alpha} \frac{1}{\beta + \alpha} & \text{for } \mu - \beta \leq x \leq \mu - \alpha \\ \frac{1}{\beta + \alpha} & \text{for } \mu - \alpha \leq x \leq \mu + \alpha \\ \frac{(\mu + \beta) - x}{\beta - \alpha} \frac{1}{\beta + \alpha} & \text{for } \mu + \alpha \leq x \leq \mu + \beta \end{cases} \quad (3)$$

The expression (3) is the pdf $f(x)$ of a trapezoidal distribution with parameters $\mu - \beta$, $\mu - \alpha$, $\mu + \alpha$, and $\mu + \beta$ [3, section 2].

In the appendix, we discuss how the pdf $g(z)$ for the half-width Z which yields the trapezoidal pdf (3) was determined. The pdf $g(z)$ is a special case of a generalized trapezoidal distribution described in [4, substitute $\alpha = b/c$ in eqn 6 on p 88]. A plot of the pdf $g(z)$ for $\alpha = 0.5$ and $\beta = 1.5$ is shown in figure 1. Included in figure 1 is a plot of the pdf $h(z)$ of the corresponding rectangular distribution on the interval $(0.5, 1.5)$. Since the top of the pdf $g(z)$ is sloping we can refer to it as a sloping top distribution.

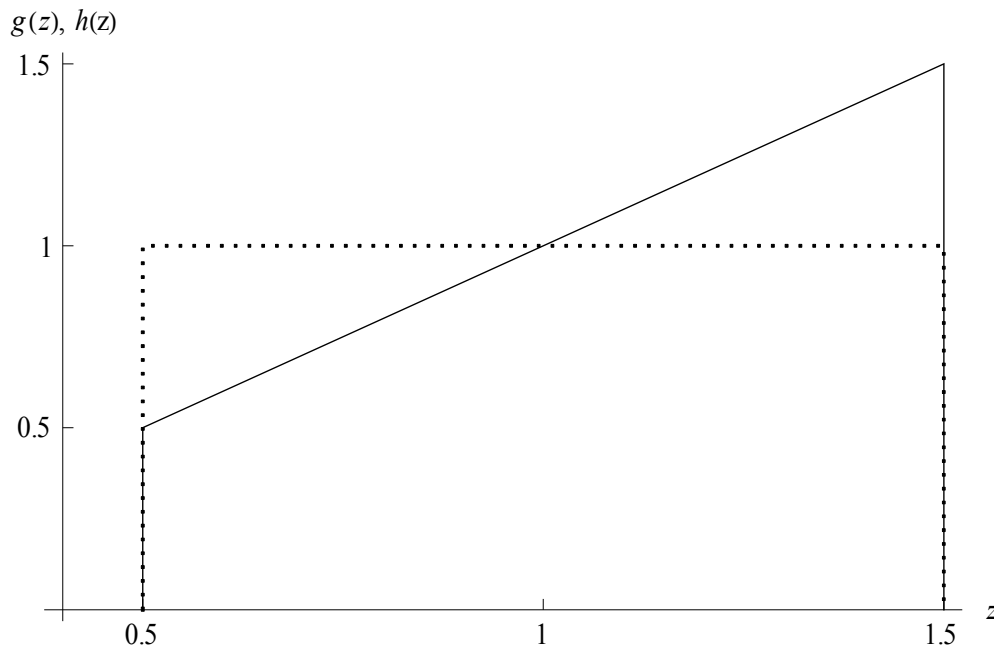


Figure 1: The pdf of a sloping top distribution with the pdf $g(z)$ eqn (2) for $\alpha = 0.5$ and $\beta = 1.5$ plotted in solid line and the pdf $h(z)$ of the corresponding rectangular distribution plotted in dotted line.

The moment generating function (mgf) of the pdf $g(z)$ is

$$M_Z(t) = \int_{\alpha}^{\beta} e^{tz} g(z) dz = 1 + \sum_{k=1}^{\infty} \left(\frac{2}{k+2} \frac{\beta^{k+2} - \alpha^{k+2}}{\beta^2 - \alpha^2} \right) \frac{t^k}{k!}. \quad (4)$$

The coefficient of $t^k/k!$ in $M_Z(t)$ is the k -th moment; thus,

$$E(Z^k) = \frac{2(\beta^{k+2} - \alpha^{k+2})}{(k+2)(\beta^2 - \alpha^2)}. \quad (5)$$

In particular, the expected value $E(Z)$ is

$$E(Z) = \frac{2(\beta^2 + \alpha^2 + \alpha\beta)}{3(\beta + \alpha)} = \frac{\beta + \alpha}{2} + \frac{(\beta - \alpha)^2}{6(\beta + \alpha)}. \quad (6)$$

and the second moment $E(Z^2)$ is

$$E(Z^2) = \frac{\beta^2 + \alpha^2}{2}. \quad (7)$$

The variance $V(Z) = E(Z^2) - (E(Z))^2$ simplifies to

$$V(Z) = \frac{(\beta - \alpha)^2(\beta^2 + 4\alpha\beta - \alpha^2)}{18(\beta + \alpha)^2} = \frac{(\beta - \alpha)^2}{12} - \frac{(\beta - \alpha)^4}{36(\beta + \alpha)^2}. \quad (8)$$

The expected value and the variance of a rectangular distribution on the interval (α, β) are $(\beta + \alpha)/2$ and $(\beta - \alpha)^2/12$, respectively. Therefore, from equations (6) and (8) we note that the mean squared deviation, $E(Z - (\beta + \alpha)/2)^2$, of the pdf $g(z)$ about its mid-point $(\beta + \alpha)/2$ is identical to the variance of the rectangular distribution on (α, β) .

Appendix

If the pdf of Z is $g(z)$ for $\alpha \leq z \leq \beta$, then from (1), the joint pdf of X and Z is

$$f_{X,Z}(x, z) = \frac{g(z)}{2z} \quad \text{for} \quad \mu - z \leq x \leq \mu + z \quad \text{and} \quad \alpha \leq z \leq \beta \quad (9)$$

The unconditional pdf $f(x)$ of X for a particular value x is the integral of the joint pdf (9) with respect to the possible values of z corresponding to that x . The joint pdf has positive value in the area bounded by the lines $z = \alpha$, $z = \beta$, $z = x - \mu$, and $z = \mu - x$ indicated in figure 2. The range of the possible values of z for a given x depends on which of the three horizontal line segments in figure 2 contains that value x . If $\mu - \beta \leq x \leq \mu - \alpha$, then $\mu - x \leq z \leq \beta$; if $\mu - \alpha \leq x \leq \mu + \alpha$, then $\alpha \leq z \leq \beta$, and if $\mu + \alpha \leq x \leq \mu + \beta$, then $x - \mu \leq z \leq \beta$. We want X to have a trapezoidal distribution with parameters $\mu - \beta$, $\mu - \alpha$, $\mu + \alpha$, and $\mu + \beta$. Therefore, when $\mu - \beta \leq x \leq \mu - \alpha$, we want the slope of

$$f(x) = \int_{\mu-x}^{\beta} \frac{g(z)}{2z} dz \quad (10)$$

to be a positive constant, when $\mu - \alpha \leq x \leq \mu + \alpha$, we want the slope of

$$f(x) = \int_{\alpha}^{\beta} \frac{g(z)}{2z} dz \quad (11)$$

to be zero, and when $\mu + \alpha \leq x \leq \mu + \beta$, we want the slope of

$$f(x) = \int_{x-\mu}^{\beta} \frac{g(z)}{2z} dz \quad (12)$$

to be a negative constant. In particular, when $\mu + \alpha \leq x \leq \mu + \beta$, we want the derivative (slope) of (12), which is

$$\frac{df(x)}{dx} = -\frac{g(x-\mu)}{2(x-\mu)} \quad (13)$$

to equal some negative constant $-C_0$. Therefore

$$g(x-\mu) = 2C_0(x-\mu) \quad \text{for } \mu + \alpha \leq x \leq \mu + \beta \quad (14)$$

Putting $x - \mu = z$, we have

$$g(z) = 2C_0(z) \quad \text{for } \alpha \leq z \leq \beta \quad (15)$$

Since $g(z)$ is a pdf, we must have

$$1 = \int_{\alpha}^{\beta} g(z) dz = 2C_0 \int_{\alpha}^{\beta} z dz = C_0(\beta^2 - \alpha^2). \quad (16)$$

Thus $C_0 = 1/(\beta^2 - \alpha^2)$ and the pdf $g(z)$ of Z is

$$g(z) = \frac{2z}{\beta^2 - \alpha^2} \quad \text{for } \alpha \leq z \leq \beta. \quad (17)$$

If we substitute (17) in (10), (11), and (12), we get the pdf $f(x)$ of X given in (3).

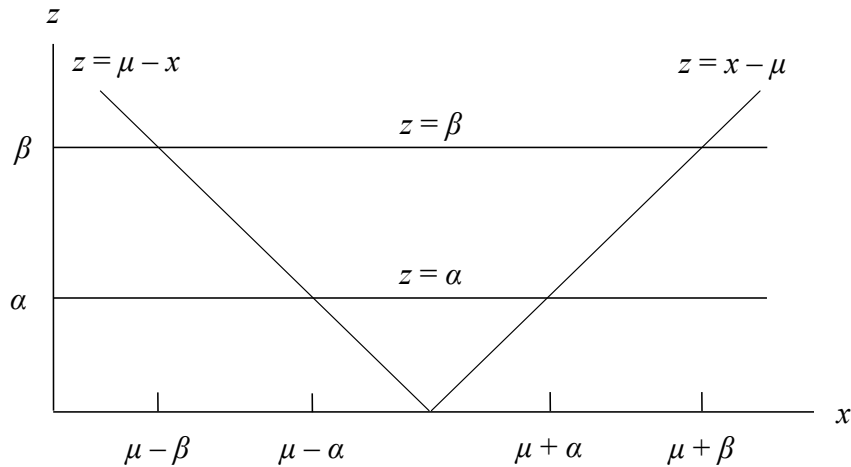


Figure 2: The joint pdf of X and Z has positive value in the area bounded by the lines $z = \alpha$, $z = \beta$, $z = x - \mu$, and $z = \mu - x$

Summary

The GUM [1, section 4.3.9, note 2] states that if the width of a rectangular distribution is inexactly known and the state of knowledge about the width can be represented by another (narrower) rectangular distribution then the resulting distribution is an isosceles trapezoid. This statement is not exactly correct but it is a reasonable approximation for many applications in metrology [2]. In this paper we have shown that if the mid-point of rectangular distribution is specified, the half-width is inexactly known, and the probability distribution about the half-width is represented by a sloping top distribution on an interval (α, β) with slope $2/(\beta^2 - \alpha^2)$ then the resulting distribution is indeed an isosceles trapezoid. The mean square deviation of the sloping top distribution about the mid-point is equal to the variance of the corresponding rectangular distribution.

References

- [1] *Guide to the Expression of Uncertainty in Measurement (GUM)* 2nd ed 1995 (Geneva: International Organization for Standardization, ISO) ISBN 92-67-10188-9 (2008 electronic version of GUM from the BIPM Joint Committee on Guides in Metrology JCGM is available at http://www.bipm.org/utis/common/documents/jcgm/JCGM_100_2008_E.pdf)
- [2] Kacker R N and Lawrence J F 2009 Rectangular distribution whose width is not exactly known: isocurvilinear trapezoidal distribution *Metrologia* **46** 254-260

- [3] Kacker R N and Lawrence J F 2007 Trapezoidal and triangular distributions for Type B evaluation of standard uncertainty *Metrologia* **44** 117-127
- [4] René van Dorp J and Kotz S 2003 Generalized trapezoidal distributions *Metrika* **58** 85-97