Coverage factor $k = 2$ (often) gives 95 % coverage

Maurice Cox
National Physical Laboratory, Teddington TW11 0LW, UK
Lars Nielsen
Danish Fundamental Metrology, 2970 Hørsholm, Denmark

MATHMET 2022
2 – 4 November 2022

Coverage

Coverage factor $k = 2$ often used to transform standard to expanded uncertainties

'Coverage is approximately 95 %': valid statement when probability distribution underlying the quantity is (near-)normal [GUM]

Practice criticized because of the frequent non-normality of this distribution

Example: dominant input quantity with a non-normal, say, rectangular, distribution

Demonstrate (not prove) instances in which $k = 2$ indeed provides (at least) 95 % coverage:

- Input quantities having certain symmetric unimodal (single-peaked) distributions
- Input standard uncertainties taken as the standard deviations of these distributions

Independent input quantities

Linearized measurement model

$$Y = c_1 X_1 + \cdots + c_N X_N$$

The X_i are input quantities and Y is the output quantity or measurand

$x_1, \ldots, x_N$: available estimates of the X_i

In the GUM, standard uncertainties $u(x_i)$ evaluated in either of two ways:

(A) Type A (statistical) evaluation of uncertainty used when knowledge of X_i obtained by repeated independent observations

— according probability distribution assigned to X_i

(B) Type B (non-statistical) evaluation of uncertainty applies when scientific judgement used

— X_i characterized by probability distribution based on knowledge available

Mean of distribution taken as estimate x_i of X_i and standard deviation as associated standard uncertainty $u(x_i)$: GUM-S1 (JCGM 101)

Coverage intervals for the measurand à la GUM

Coverage intervals $y \pm ku(y)$ for measurand Y

k: coverage factor corresponding to required coverage probability

Infinite DoF of $u(y)$: $k = 2$ often taken as providing 95 % coverage

Finite DoF of $u(y)$: factor from the Student's t distribution

Finite DoF of $u(y)$ consequence of finite DoF of one or more of the $u(x_i)$ resulting from the GUM's Type A uncertainty evaluations

---

Taking $u(x_i)$ as standard deviation of probability distribution for $X_i \implies$ no DoF of $u(y)$

Problem statement

Suppose each X_i characterized independently by unimodal symmetric PDF with given mean μ_i and given standard deviation σ_i having the property that the interval defined by

$$\mu_i \pm 2\sigma_i$$

covers the value of X_i with at least 95 % probability

Measurand Y has a PDF with mean μ and standard deviation σ given by

$$\mu = \mu_1 + \cdots + \mu_N$$
$$\sigma^2 = c_1^2\sigma_1^2 + \cdots + c_N^2\sigma_N^2$$

For what classes of unimodal symmetric PDFs for the input quantities can it be stated that the interval defined by

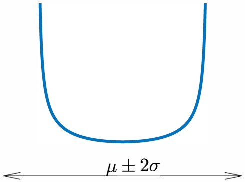
$$\mu \pm 2\sigma$$

contains Y with at least 95 % probability?

Conjecture

When the PDFs for the input quantities are symmetric and unimodal, the value of the measurand lies within 2σ of μ, where μ is the mean of Y and σ is its standard deviation, with at least 95 % probability

Unimodality is sufficient but not necessary —
probability 100 % for U-shaped (arc sin) PDF



Note: any number of symmetric unimodal probability distributions, when convolved, yield a probability distribution that is itself symmetric and unimodal and

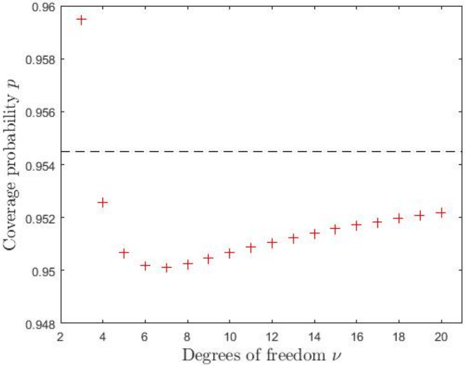
Coverage probability for input quantities provided by a coverage factor of 2

Coverage probability p provided by coverage factor $k = 2$ for commonly used distributions

PDF	Normal	Rectangular	Student's t
p	0.9545	1	> 0.9501

Coverage probability p provided by coverage factor $k = 2$ for the Student's t distribution

ν	3	4	5	6
p	0.9595	0.9526	0.9507	0.9502
ν	7	8	9	10
p	0.9501	0.9503	0.9505	0.9507



Attempts to try to disprove the conjecture

Conjecture is false in general —
instead, we have carried out many simulations to investigate when it might apply

Each simulation involves

1. Input quantities having PDFs drawn from normal, rectangular and Student's t distributions
2. Test to decide whether coverage factor $k = 2$ gives at least $95\,\%$ coverage

Simulations cover 1000s of cases, where the input PDFs are randomly drawn from above set with standard deviations from rectangular distribution $\mathrm{R}(0,\ 1]$, sensitivity coefficients from $\mathrm{R}[-1,\ 1]$, and number N of input quantities ranging from 1 to 30

Additionally, potentially 'difficult' combinations manually selected

Basis of each simulation

Sensitivity coefficients $c_1, \ldots, c_N$

PDFs with zero means (all $x_i = 0$) and given standard deviations $u(x_i)$

Standard deviation $u(y)$ of the PDF for the measurand calculated from

$$u^2(y) = c_1^2 u^2(x_1) + \ldots + c_N^2 u^2(x_n)$$

A trial consists of making a draw from each input PDF

The rth of M (large, say 10^6) trials provides values $x_1, \ldots, x_N$ of the input quantities and forms the corresponding value of the measurand:

$$y = c_1 x_1 + \cdots + c_N x_N$$

The frequentist coverage probability p is the proportion of the M y-values so obtained satisfying

$$-ku(y) < y < ku(y)$$

$\boxed{\text{If } p < 0.95, \text{ conjecture is disproved}}$

Simulation set 1: normal PDFs

Simplest, involving normal input quantities

Although Y is normal, valuable to check whether the simulation functions satisfactorily

Over all experiments, p ranged from 0.9543 to 0.9546, all exceeding 0.95, compared with the value 0.9545 from the cumulative normal distribution

Range narrower for $M = 10^7$ and more so for $M = 10^8$

Simulation set 2: as set 1 but with rectangular PDFs

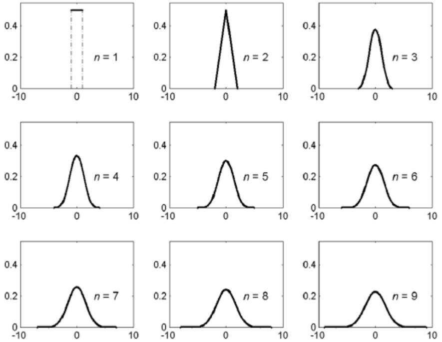
p fell from 1 to 0.9564 as N ranged from 1 to 30, the '1' corresponding to $N = 1$ as expected [GUM clause G.1.3]

(Slight) lack of monotonicity due to finite number $M = 10^6$ of Monte Carlo trials

Repetition with $M = 10^7$ came closer to exhibiting monotonicity

p would approach 0.9545: convolution of rectangular variates has this property [GUM clause G.2.2]

Identical rectangular distributions



Convolution of uniform B-splines rapidly approach normality (GUM)

Simulation set 3: Student's t PDFs

Student's t PDFs with 3 DoF

Smallest number of observations, 4, in a Type A evaluation for which the standard deviation of the Student's t distribution exists

Over many simulations of this type, proportion ranged from 0.9562 to 0.9583

Further simulations made on same basis except Student's t PDFs were assigned 7 DoF since that number came closest to the 0.9501 'limit' for a single Student's t distribution

Over many simulations of this type, proportion ranged from 0.9504 to 0.9544

Simulation set 4: mixtures of distributions plus special cases

1. Mixtures of normal, rectangular and Student's t distributions

   Many simulations made with DoFs randomly selected as integers between 3 and 10

   Over many simulations of this type, proportion ranged from 0.9500 to 1

2. Special cases, attempting to mimic practice including dominant input quantities

‘Proof by MATLAB’

Not a single instance observed that disproved conjecture for input PDFs used

‘Proof by MATLAB’ of weaker conjecture

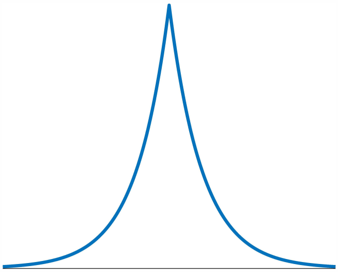
> When the PDFs for the input quantities are normal, rectangular or Student’s t, the value of the measurand lies within 2σ of μ with at least 95 % probability

Counter-example to original conjecture

Statement that the coverage factor applies to all symmetric unimodal PDFs is false

Counter-example: Laplace (double exponential) distribution: gradient discontinuity at its mean

Coverage probability is 0.9409, not quite achieving 0.95 as previously

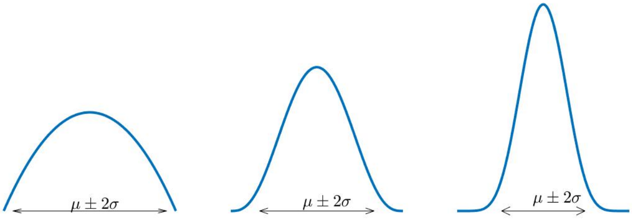


Conclusions and further remarks

Not a single instance observed that disproved conjecture for normal, rectangular or Student's t input PDFs

Even if there were exceptional cases, much evidence gathered that for the input PDFs considered coverage factor $k = 2$ delivers a coverage interval with at least 95 % coverage

Simulations being extended to include symmetric beta and other distributions, even multimodel



Symmetric beta PDFs $C[\xi(1-\xi)]^{\alpha-1}$ with $\alpha = 2,\ 4,\ 8$ from left to right

Acknowledgment

Work supported by the Data Science programme within the National Measurement System programme of the UK Government's Department for Business, Energy and Industrial Strategy

NPL
National Physical Laboratory