УДК 519.814:519.226:006.86

И.Р. Шайняк

НИЦ контроля и диагностики технических систем, Нижний Новгород, Россия

ОБЪЕДИНЕНИЕ ИНФОРМАЦИИ ПРИ ФОРМИРОВАНИИ ОЦЕНКИ НЕОПРЕДЕЛЕННОСТИ ИЗМЕРЕНИЯ

Рассматривается общий подход к возможности корректировки расчета неопределенности измерения при наличии информации о точности примененного метода измерений, полученной в соответствии с ISO 5725-2 (результатами межлабораторного эксперимента), и условия, необходимые для выполнения такой корректировки. Показано, что возможность объединения информации существует, если известна структура вывода неопределенности, в частности, если неопределенность измерения получена с применением формулы Уэлча-Саттертуэйта в соответствии с Руководством по выражению неопределенности измерения (GUM).

***Ключевые слова**: неопределенность измерения, GUM, байесовский подход, объединение информации.*

Введение

В литературе известно не так много работ, посвященных уточнению результата оценки неопределенности измерения при наличии априорной (или апостериорной) информации об измеряемой величине (см., например, [1], [2]). В дополнении [3] к Руководству по выражению неопределенности измерений (GUM) [4] при наличии такой информации рекомендуется обращаться к работе [2]. Типичным примером появления дополнительной информации об измеряемой величине является проведение измерений той же величины в заданных условиях измерений, но другой лабораторией с применением других средств измерений и, возможно, другого метода измерений.

Существенно иной, однако, будет ситуация, при наличии дополнительной информации, полученной при измерениях не той же самой, но аналогичной измеряемой величины (т.е. той же физической величины, имеющей другое значение в условиях проведения измерений). Типичным примером получения такой информации является проведение межлабораторного эксперимента по [5] в целях оценки точности применяемого метода измерений. Результат измерений по [5] обычно приводят в виде оценки стандартного отклонения воспроизводимости s_R. В таком виде, как правило, сведения о точности метода приводится и в аттестованных методиках измерений. Важно отметить, что при этом обычно опускается такая важная информация, как число степеней свободы m, соответствующее полученной оценке s_R, и в дальнейших измерениях по методике s_R рассматривается как точно известная величина.

Международный стандарт [6] рассматривает вопрос выражения точности измерения в терминах неопределенности, когда имеется априорная информация о результатах наблюдений, интерпретируемая в классическом частотном смысле. Согласно [6] структура оценки стандартной неопределенности u(y) для измеряемой величины y включает в себя в общем случае оценку стандартного отклонения воспроизводимости s_R, а также еще две составляющие, связанные с применяемым методом измерений и с неполнотой варьирования возможных влияющих величин в ходе межлабораторного эксперимента (влияние двух последних источников неопределенности в настоящей статье далее учитывать не будем). При таком подходе результаты измерений, выполненные данным методом разными лабораториями в разных условиях измерений, будут включать в себя одну и ту же оценку неопределенности измерения, что в принципе нельзя признать нормальным. Международный стандарт [7], устанавливающий требования к компетентности испытательных лабораторий, предписывает им проводить собственную оценку неопределенности выполненного измерения, и нести ответственность за достоверность данной оценки.

Конечно, встречаются измерительные задачи, для которых строгое следование схеме оценивания неопределенности по GUM представляет собой трудновыполнимую задачу, и в обоснованных случаях приписывание u(y) значения s_R можно признать допустимым. Но в общем случае постановка задачи должна быть иной: лаборатория получает собственную оценку неопределенности измерения, используя при этом s_R в качестве дополнительной априорной информации. При этом, важно отметить, что для реализации такого подхода помимо значения s_R необходимо также знать и соответствующее число степеней свободы m. В противном случае, если считать характеристику, описывающую возможный разброс значений, которые могут быть обоснованно приписаны измеряемой величине, известной точно, то никакое объединения информации невозможно в принципе, и остается только следовать подходу, изложенному в [6].

Пример 1 из GUM (приложение H)

Дальнейшие рассуждения построим на основе примера 1 (калибровка концевой меры длины), приведенного в приложении H GUM.

Одной из входных величин в уравнении измерения для длины концевой меры является разность длин калибруемой концевой меры и эталона.

Выборочное стандартное отклонение, характеризующее результат сравнения двух (произвольных) концевых мер, было получено на основе 25 повторных наблюдений (возможно, в ходе межлабораторного эксперимента – это в GUM не уточняется) и было принято в качестве оценки стандартного отклонения для данной входной величины. Сама процедура калибровки включала в себя пять повторных наблюдений, по которым было рассчитано среднее, принятое в качестве наилучшей оценки входной величины. Таким образом, мы вновь сталкиваемся с рассмотренной выше ситуацией, когда стандартная неопределенность оценивается только на основе предшествующих измерений, а разброс результатов наблюдений, полученных в ходе самой калибровки, во внимание не принят. При этом никаких объяснений, почему информация от двух разных наблюдений не была объединена для получения стандартной неопределенности входной величины, в GUM не приводится. Данное оценивание не может рассматриваться как оценивание стандартной неопределенности входной величины по типу A [4], поскольку оценка стандартной неопределенности получена на основе априорных знаний. Тем не менее, сами условия проведения наблюдений допускают объединение информации как в частотном смысле, так и в смысле более адекватного для интерпретации GUM байесовского вывода [8]. Рассмотрим оба этих подхода для данной задачи, сформулировав ее предварительно в более общем виде.

Пусть по двум выборкам наблюдений, полученной в ходе измерения и априорной, объемом n и m для случайных величин X и Y, распределенных по нормальному закону $N(\mu_X;\sigma^2)$ и $N(\mu_Y;\sigma^2)$ соответственно, получены выборочные значения оценок параметров распределений: $\bar{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$; $S_X^2=\frac{1}{n}\sum_{i=1}^{n}\left(X_i-\bar{X}\right)^2$; $S_Y^2=\frac{1}{m}\sum_{i=1}^{m}\left(Y_i-\bar{Y}\right)^2$. Оценка $\bar{Y}$ неизвестна. Требуется построить на основе информации, содержащейся в двух выборках, доверительный интервал (для частотной интерпретации вероятностей) и апостериорную вероятность (на основе байесовского подхода) для μ_X.

1. Построение доверительного интервала

Найдем оценку максимального правдоподобия для параметра σ^2. Функция правдоподобия для двух выборок имеет вид

$$L = \left(2\pi\sigma^2\right)^{-(m+n)/2} \exp\left[\sum_{i=1}^{n} -\frac{\left(X_i - \mu_X\right)^2}{2\sigma^2} - \sum_{j=1}^{m} -\frac{\left(Y_j - \mu_Y\right)^2}{2\sigma^2}\right].$$

Тогда

$$\ln L = -\frac{m+n}{2}\ln 2\pi - (m+n)\ln\sigma - \\ -\sum_{i=1}^{n}\frac{\left(X_i - \mu_X\right)^2}{2\sigma^2} - \sum_{j=1}^{m} -\frac{\left(Y_j - \mu_Y\right)^2}{2\sigma^2}.$$

Оценки $\hat{\mu}_X$, $\hat{\mu}_Y$ и $\hat{\sigma}^2$ доставляют максимум функции $\ln L$. Из этого условия получаем

$$\hat{\mu}_X = \bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i\,;\; \hat{\mu}_Y = \bar{Y} = \frac{1}{m}\sum_{i=1}^{m} Y_i\,.$$

Дифференцируя $\ln L$ по σ, подставляя $\hat{\sigma}^2$ вместо σ^2 и приравнивая полученное выражение к нулю, получаем:

$$-\frac{n+m}{2\hat{\sigma}^2} + \sum_{i=1}^{n}\frac{\left(X_i - \hat{\mu}_X\right)^2}{2\left(\hat{\sigma}^2\right)^2} + \sum_{j=1}^{m}\frac{\left(Y_j - \hat{\mu}_Y\right)^2}{2\left(\hat{\sigma}^2\right)^2} = 0,$$

откуда

$$\hat{\sigma}^2 = S^2 = \frac{1}{n+m}\left[\sum_{i=1}^{n}\left(X_i - \hat{\mu}_X\right)^2 + \sum_{j=1}^{m}\left(Y_j - \hat{\mu}_Y\right)^2\right] = \\ = \frac{n}{n+m}S_X^2 + \frac{m}{n+m}S_Y^2.$$

Получена, таким образом, объединенная оценка выборочной дисперсии, не зависящая от неизвестного значения $\hat{\mu}_Y$.

Про масштабированные выборочные дисперсии nS_X^2/σ^2 и mS_Y^2/σ^2 известно, что они имеют хи-квадрат распределение с $(n-1)$ и $(m-1)$ степенями свободы соответственно: $\chi^2(n-1)$ и $\chi^2(m-1)$.

Можно видеть, что $\xi = (n+m)S^2/\sigma^2 = = nS_X^2/\sigma^2 + mS_Y^2/\sigma^2$ является суммой двух случайных величин, распределенных по законам $\chi^2(n-1)$ и $\chi^2(m-1)$, т.е. она сама распределена по закону $\chi^2(n+m-2)$.

Масштабированная и центрированная случайная величина $\eta = \dfrac{\bar{X} - \mu_X}{\sigma/\sqrt{n}}$ имеет стандартизованное нормальное распределение $N(0;1)$. Это позволяет сформировать случайную величину $\dfrac{\eta}{\sqrt{\xi/(n+m-2)}} = \sqrt{\dfrac{n(n+m-2)}{n+m}}\dfrac{\bar{X} - \mu_x}{S}$, которая будет иметь t-распределение Стьюдента с $(n+m-2)$ степенями свободы, $t(n+m-2)$.

Если $t_{\alpha,n+m-2}$ – $(1-\alpha)$ – квантиль распределения Стьюдента с числом степеней свободы $(n+m-2)$, то $\bar{X} \pm t_{\alpha,n+m-2} S\Big/\sqrt{\dfrac{n(n+m-2)}{n+m}}$ – $(1-2\alpha)$-доверительный интервал для μ_X.

2. Получение апостериорной вероятности

В байесовском подходе функцию правдоподобия формируют на основе n наблюдений, а данные по выборке объемом *m* используют для получения априорной информации о параметрах распределения, которые рассматривают как случайные величины. Априорная выборка получена для генеральной совокупности с математическим ожиданием μ_Y и не дает никакой информации о возможном значении μ_X. Поэтому в качестве априорного распределения $p(\mu_X)$ для μ_X разумно выбрать неинформативное равномерное распределение на бесконечном интервале, т.е. $\mu_X \propto 1$.

Известно, что сформированная на основе выборочной дисперсии S_Y^2 нормального распределения $N(\mu_Y;\sigma^2)$ случайная величина $m\dfrac{S_Y^2}{\sigma^2}$ распределена по закону $\chi^2(m-1)$, а обратная к ней величина $\dfrac{1}{m}\dfrac{\sigma^2}{S_Y^2}$ имеет обратное хи-квадрат распределение с $(m-1)$ степенью свободы $\mathrm{Inv}-\chi^2(m-1)$, т.е.

$$p\left(\sigma^2\right) = \frac{2^{-(m-1)/2}}{\Gamma((m-1)/2)}\left(\frac{mS_Y^2}{\sigma^2}\right)^{-(m-1)/2}\sigma^2\exp\left(-\frac{mS_Y^2}{2\sigma^2}\right).$$

Функция максимального правдоподобия в формуле Байеса будет иметь вид

$$L = \prod_{i=1}^{n}\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left[-\frac{\left(X_i - \mu_X\right)^2}{2\sigma^2}\right].$$

Тогда апостериорную совместную вероятность $p\left(\mu_X,\sigma^2\middle|X_i\right)$ для μ_X и σ^2 можно привести к виду

$$p\left(\mu_{X,}\,\sigma^2\middle|X_i\right) = k\sigma^{3-(m+n)}\exp\left(-\frac{mS_Y^2 + nS_X^2}{2\sigma^2}\right)\times \\ \times\exp\left[-\frac{n\left(\mu_X - \hat{\mu}_X\right)^2}{2\sigma^2}\right],$$

где k – масштабирующий множитель.

Апостериорную вероятность $p\left(\mu_X\middle|X_i\right)$ для μ_X находят интегрированием $p\left(\mu_X,\sigma^2\middle|X_i\right)$ по σ^2, т.е.

$$p\left(\mu_X\middle|X_i\right) = \int_0^\infty p\left(\mu_X,\sigma^2\middle|X_i\right)d\sigma^2,$$

а масштабирующий множитель k – из условия

$$\int_{-\infty}^{\infty} p\left(\mu_X\middle|X_i\right)d\mu_X = 1.$$

В результате мы снова приходим к тому же t-распределению с числом степеней свободы $n+m-2$, что было получено при анализе в рамках частотного подхода, но теперь это будет апостериорное распределение центрированной и масштабированной случайной величины, ассоциированной с измеряемой величиной.

Распространение на случай оценки суммарной стандартной неопределенности

В полученном результате для $p(\mu_X X_i)$ важно, что окончательный результат не зависит от выборочных значений X_i, а зависит только от известных значений достаточных статистик $\hat{\mu}_X$, S_X^2 и S_Y^2, а также от чисел степеней свободы n и m. Этого достаточно, чтобы распространить результат, полученный в вышерассмотренном примере для одной входной величины, на способ объединения информации при оценивании суммарной стандартной неопределенности для измеряемой величины. Для этого необходимо знать те же пять параметров: $\hat{\mu}_X$, S_X^2, S_Y^2, n и m, – но уже применительно к распределению выходной величины.

Параметры S_Y^2 и m – это априорная информация, полученная, например, в результате межлабораторного эксперимента. S_X^2 и $\hat{\mu}_X$ – оценка стандартной неопределенности и наилучшая оценка измеряемой величины, полученные в соответствии с обычной процедурой GUM. Сложнее обстоит дело с получением оценки общего числа степеней свободы n в проведенном измерении, поскольку в общем случае ни процедура GUM, ни тем более способ получения апостериорного распределения, ассоциированного с измеряемой величиной, методом Монте-Карло [3] не только не включают в себя расчет n, но даже и не подразумевают наличия этой величины. Единственным исключением, позволяющим в данный момент использовать описанную процедуру объединения информации, является описанный в разделе G.4 приложения G в GUM способ расчета эффективного числа степеней свободы n_{eff} с использованием формулы Уэлча-Саттертуэйта. В этом случае в выражении для $p(\mu_X X_i)$ значение n следует заменить на n_{eff}.

По счастливой случайности сама структура вывода с использованием формулы Уэлча-Саттертуэйта предполагает оценку степени доверия к получаемому результату, мерилом которого при частотной интерпретации вероятности служит число степеней свободы. Именно это позволяет сопоставлять результаты прошлых и настоящих измерений в смысле их «надежности»: более «надежные» оценки входят в окончательный результат с большим весом. Отсюда понятно, почему в примере из GUM пренебрегли вкладом информации о разбросе наблюдаемой величины, полученной только на основе пяти наблюдений в ходе калибровки (предварительная оценка на основе 25 наблюдений будет, конечно, гораздо более достоверной).

Однако, вообще говоря, понятие достоверности или «надежности» оценки не является неотъемлемой частью байесовского вывода, а при отсутствии характеристик достоверности информации объединять данные из разных наблюдений не представляется возможным.

Список литературы

1. Lira I., Wöger W. Bayesian evaluation of the standard uncertainty and coverage probability in a simple measurement model // Meas. Sci. Technol. – 2001. – Vol. 12. – P. 1172–1179.

2. Elster C. Calculation of uncertainty in the presence of prior knowledge // Metrologia. – 2007. – V. 44. – P. 111-116.

3. ISO/IEC Guide 98-3:2008/Supplement 1:2008 Propagation of distributions using a Monte Carlo method.

4. ISO/IEC Guide 98-3:2008 Uncertainty of measurement – Part 3: Guide to the expression of uncertainty in measurement (GUM:1995).

5. ISO 5725-2:1994 Accuracy (trueness and precision) of measurement methods and results -- Part 2: Basic method for the determination of repeatability and reproducibility of a standard measurement method.

6. ISO 21748:2010 Guidance for the use of repeatability, reproducibility and trueness estimates in measurement uncertainty estimation.

7. ISO/IEC 17025:2005 General requirements for the competence of testing and calibration laboratories.

8. Lira I., Grientschnig, D. Bayesian assessment of uncertainty in metrology: a tutorial // Metrologia. – 2010. – V. 44. – P. R1-R14/.

Поступила в редколлегию 3.04.2014

Рецензент: д-р техн. наук А.Г. Чуновкина, Всероссийский НИИ метрологии им. Д.И. Менделеева, Санкт-Петербург.

ОБ'ЄДНАННЯ ІНФОРМАЦІЇ ПРИ ФОРМУВАННІ ОЦІНКИ НЕВИЗНАЧЕНОСТІ ВИМІРЮВАННЯ

І.Р. Шайняк

Розглядається загальний підхід до можливості коригування розрахунку невизначеності вимірювання за наявності інформації про точність застосованого методу вимірювань, отриманої відповідно до ISO 5725-2 (результатами міжлабораторного експерименту), і умови, необхідні для виконання такого коригування. Показано, що можливість об'єднання інформації існує, якщо відома структура виведення невизначеності, зокрема, якщо невизначеність вимірювання отримана із застосуванням формули Уелча - Саттертуейт відповідно до Керівництва з висловом невизначеності вимірювання (GUM).

Ключові слова: *невизначеність вимірювання, GUM, баєсівський підхід, об'єднання інформації.*

COMBINING HETEROGENEOUS DATA IN MEASUREMENT UNCERTAINTY ESTIMATION

I.R. Szajniak

An approach how the measurement result could be improved in the presence of data on measurement method accuracy obtained as a consequence of an interlaboratory experiment in accordance with ISO 575-2 as well as conditions when such an approach could be applied are considered. It is shown that data could be combined if only information on the structure of inference in a suitable form is kept, particularly if the inference is based on the Welch-Satterthwaite formula.

Keywords: *measurement uncertainty, GUM, Bayesian approach, heterogeneous data combination.*