THE APPLICATION OF BAYES´ THEOREM FOR THE NEED OF FIRE TECHNICAL EXPERTISES

Otto DVOŘÁK[1]

Research article

Abstract:	The article describes literary knowledge briefly and presents an example of the application of a statistical, probability approach to the estimations of the uncertainties of the results of the qualitative GC/MS chemical analyses of the samples of the fire debris from the seat of fire for the content of accelerants.
Keywords:	Bayes´ theorem, fire technical expertise, qualitative chemical analyses, GC/MS, uncertainty estimation, fire debris, samples, accelerants.

Introduction

ČSN EN ISO/IEC 17025 (2005) demands accredited testing laboratories to have and use procedures for the assessment of the measurements of uncertainties. During their estimation, they must consider all the components of the uncertainties while using proper methods of the analysis. The introduction of a conception of the assessment of uncertainties consisting of test data according to mentioned standards is specified, among others, by the document ILAC No. G 17:2002 (Document, 2002). It is perceptible that the judgement associated to a test result and characterizing the interval of values about which it is claimed that the accurate value lies inside it (ČSN, 1994) relates to the quantitative measurement/test result according to the definition of the uncertainty. The definitions and procedures of the estimation of uncertainties are stated by e.g. the so-called GUM (1995), ČSN P ENV 13005 standard (1997), and a series of other documents, e.g. (ČSN, 2005; ČSN, 2003; Eurachem, 2000). The estimation procedure can be characterized by setting the components of uncertainties: - by the procedure A including the random statistical mistakes, - by the procedure B expressing components of uncertainties from their known sources. Consequently, the combined standard uncertainty of measurement is calculated according to the law of spreading uncertainties from its particular components and the so-called extended uncertainty, usually with the coverage factor, $k = 2$ estimating the interval around a measured result of such a size that the correct result lies in it with the 95 % certainty.

The purpose of qualitative fire test determinations, e.g. chemical analyses, is principally to determine/verify the material nature of an analysed substance/material by the identification of one or more components or to assess whether the material product inclines to spontaneous ignition, whether it has pyrophoric properties, oxidizing abilities and the like with a sufficient result e.g. yes/no, false/true. The above-mentioned document (Document, 2002) states that it is always considered how to apply the uncertainty of the measurement with qualitative tests. One of the accesses is the determination of the probability of wrongly expressed positive or negative results. An international working group that is to prepare a relevant document was established for this purpose. A brief study in this field is the article of S.L.R. Ellison et al. (Dvořák, 1997). It cites, among others: - the qualitative chem. analyses can be understood as much more important than the quantitative ones which work with the presumption of the rightness of the identification of substances/materials, which are the subject of the quantification, - the selectivity, the specificity, the detection limit, the falsely positive and the falsely negative assessment are relevant characteristics for interpreting qualitative results, - it boots the term „the identification certainty" as a parameter quantifying the degree of the confidence of the following classification, - it recaps publicized works shortly with the conclusion that the Bayes´ theorem provides a fair frame for the assessment of the uncertainty of the classification (designation), which is usually done on the basis of a qualitative test result, - it cites an example of the application of this theorem in the forensic science e.g. for the identification of the type of blood, glass, DNA (Boef and Hulanicki, 1983; Ginneken and Smeulders, 1991; Evvet and Gill, 1991).

It is evident that these procedures can also be applied in the fire science and during fire technical expertises. For example, chromatographic methods are used frequently by fire and forensic laboratories for chemical analyses of the samples of the fire debris taken for the content of the accelerants in the seat of the fire. Arsonists use petrol (gasoline),

---
[1] MV - GŘ HZS ČR, Fire Technical Institute, Prague, Czech Republic, odvorak@mvcr.cz

oil, kerosene, turpentine and oil fractions, further acetone, methanol and the like as accelerants most often. Their identification consists of recognizing pikes (their number, positions and relative size) in the chromatogram. Experience is required when interpreting the results of mixed samples and samples changed as to components (compared to the original product) in a smaller or larger degree by the effect of the fire warmth, namely according to typical components and their mutual relationship. ASTM (Evvet and Gill, 1991) offers the classification/identification of an accelerant into one of eight classes of flammable/combustible liquids and into three subclasses according to the number of carbons of n-alkanes with the exception of the petrol class according to the results of a chemical analysis. Further it mentions the fact that because the method is qualitative, its accuracy is not stated with it, and neither is the standard deviation of the results.

Materials and methods

Bayes´ theorem

The author has already described (Dvořák, 2005) the possibilities of the statistical evaluation of quantitative results of laboratorial qualitative parameters of the fire technical equipment and extinguishing agents for the need of the certification. The procedure of the decision came from the judgement of the confidence interval to the measured selective (arithmetical) average providing the standard normal distribution of errors and the selected confidence coefficient. The procedure is also applicable for the results of quantitative chemical analyses. It is possible to derive the primary relation of Bayes´ theorem according to equations (3) and (4) from known relations, see e.g. (Hebák and Kahounová, 1988), for the conditional probability of the H hypothesis considering the E experiment/evidence, ($P(H/E)$, equation (1) and reversely the conditional probability of the E experimental evidence regarding the H hypothesis ($P(E/H)$, equation (2).

$$P(H / E) = \frac{P(H \cap E)}{P(E)} \qquad (1)$$

$$P(E / H) = \frac{P(H \cap E)}{P(H)} \qquad (2)$$

where:

$P(H \cap E)$ is the probability of the penetration of the H hypothesis and the E experiment/evidence,

$P(H) \text{ and } P(E)$ are the unconditional so-called „a-priori" probability of the H

hypothesis and the E evidence and they are different from zero. $P(H/E) \text{ and } P(E/H)$ are the conditional, so-called „a-posteriori" probabilities.

$$P(H / E) = \frac{P(H) \cdot P(E / H)}{P(E)} \qquad (3)$$

$$P(E / H) = \frac{P(E) \cdot P(H / E)}{P(H)} \qquad (4)$$

When we express $P(E)$ with the help of the known relation (5), it is possible to derive another practical form of the Bayes´ theorem by its substitution into the equation (3), see equation (6).

$$P(E) = P(H \cap E) + P(\bar{H} \cap E) = P(E / H) \cdot \qquad (5)$$
$$\cdot P(H) + P(E / \bar{H}) . P(H)$$

where:

$\bar{H}$ is the supplement of the H hypothesis/event (opposite hypothesis/event) and it is true that the $P(\bar{H}) + P(H) = 1$

$$P(H / E) = \frac{P(H) \cdot P(E / H)}{P(H) . P(E / H) + P(E / \bar{H}) . P(\bar{H})} \qquad (6)$$

The odds form of Bayes´ equation is also used practically, see equation (7).

$$\frac{P(H / E)}{P(\bar{H} / E)} = \frac{P(H)}{P(\bar{H})} \cdot \frac{P(E / H)}{P(E / \bar{H})} \qquad (7)$$

Where the fraction on the left side of the equation $\dfrac{P(H / E)}{P(\bar{H} / E)}$ is the a-posteriori odds of the chances/expectations of the H hypothesis regarding the E experiment (a-posteriori odds),

the fraction $\dfrac{P(H)}{P(\bar{H})}$ is the a-priori odds of the chances/expectations of the hypothesis H regarding its negation (a-priori odds)

and the term $\dfrac{P(E / H)}{P(E / \bar{H})}$ is the so-called likelihood ratio (LR).

Results

The application of Bayes´ theorem for the need of fire technical expertises (FTE)

The example of frequently occurring chemical analyses of samples from the seat of a fire for the content of the accelerants (largely motor oil (MO) and automobile petrol (AP) for the confirmation or the negation of the arson hypothesis is chosen for

the application of the above-mentioned relations for the need of PTE from the illustrative reason. The procedure of the assessment of the uncertainty of this hypothesis can be expressed by the next steps on the basis of the result of the tests:

1   Securing the input data:

1.1 Chemical composition of accelerants:

- **to NM** (Regulation EC, 2006):

Boiling range, [°C] -	(180 - 360)
The composition of hydrocarbons, [% v/v]	- Aromates (20 - 30), from which polyaromates max. 8
	- Saturated hydrocarbons (n+i+cyclo) = the rest up to 100, from which n-alkanes (10 - 25)
	- Characteristic analytes: Pristane, Phytane, FAME (MEŘO): max. 7
	- The content of sulphur, [mg.kg⁻¹]: max. 10

Further see Fig. 1 and the Tab. 1. It can generally be stated that the characteristic sections of both fuels change by the vaporization of a sample, degradation changes and the presence of combustion products. Therefore, the interpretation of chromatograms requires experience.
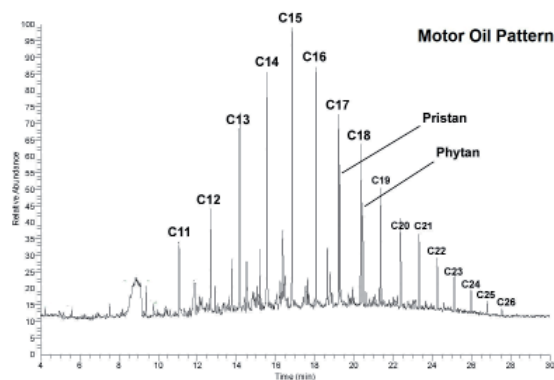

Fig. 1 An example of a motor oil pattern (TÚPO)

Tab. 1 Characteristic components found by the GC- MS method in a motor oil sample (TÚPO)

Retentive time [min]	Name of Components
11,0	C11: n-Undekane
12,6	C12: n-Dodekane
14,1	C13: n-Tridekane
15,0	C14: n-Tetradekane
16,7	C15: n-Pentadekane
18,1	C16: n-Hexadekane
19,1 - 19,2	double peak C17 + Pristane (CAS:1921-70-6)
20,3 - 20,5	double peak C18 + Phytane (CAS: 638-36-8)
21,3	C19: n-Nonadekane
22,4	C20: Eikosane
23,3	C21: Heneikosane
24,2	C22: Dokosane
etc. (MoNa oil fraction - totally approx. 1200 components)	

- **to AP** (Regulation EC, 2006):

Boiling range, [°C]:	-	(30 - 210)
The composition of hydrocarbons, [% v/v]:	-	Olefins: max. 18, typically cca 10
	-	Aromates: max. 35, typically (30 - 35), from which benzene: max. 1, typically 0,7
	-	Saturated hydrocarbons (n+i+cyclo): the rest up to 100
	-	Charakteristické analyty: Ethery (obvykle MTBE):
	-	Characteristic analytes: Ethers (usually MTBE): max. 15, (ETBE can also be) and Ethanol (bioethanol) max. 5
	-	The general content of oxygen, [% m.m⁻¹]: max. 2,7
	-	The general content of oxygen, [mg.kg⁻¹]: max. 10

Further see Fig. 2 and the Tab. 2. It can generally be stated again that the characteristic sections of both fuels change by the vaporization of a sample, by its degradation changes, by the presence of combustion products. Therefore, the interpretation of chromatograms requires experience.
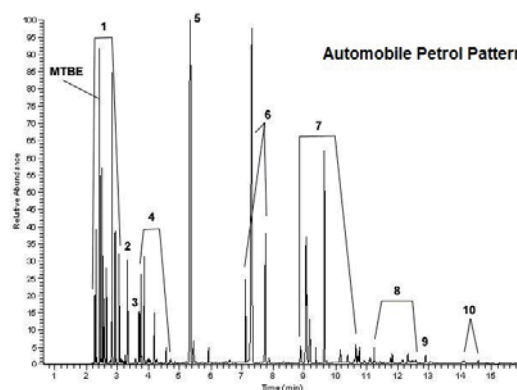

Fig. 2 An example of fresh automobile petrol pattern (GC - the straight forward spray of 0.3 µl of the sample (TÚPO)

Tab. 2 Characteristic components found in a sample of fresh automobile petrol by the GC-MS method (TÚPO)

Notation	Name of Component
1, 3, 4	Group of C5 - C8 saturated alkanes
MTBE	Tercbutyl-methylether (CAS: 1634-04-4)
2	Benzene
5	Toluene
6	Isomers of xylenes
7	Group of C3 alkyl benzenes
8	Group of C4 (C5) alkyl benzenes
9	Naphthalene
10	Isomers of 1Methyl Naphthalenes
etc. (AP- totally approx. 300 - 400 components)	

1.2 The sensitivity, detection limits for inquired substances, the linearity, the reproducibility, and the selectivity of the applied method of the chemical analysis

With the GC/MS SPME technique, see Tab. 3.

Tab. 3 The results of the validation measurement (TÚPO)

Parameter	VGR		DSQ	
analyte	o-xylene	1,2,4-TMB	m-p-xylene	1,2,4-TMB
Sensitivity	0,9922	0,9899	0,9936	0,9609
Detection limit[x]	3801087	8144842	99404691	81662592
Linearity	1,253	1,525	3,962	34,924
Repeatability r[x]	66467301	876281118	127900702	105034905

Legend:
Sensitivity = the angular coefficient of straight lines.
Detection limit = the sum of the average of a blank experiment and the triple of its standard deviation.
Linearity = the correlative coefficient of regression straight lines (TIC values vers. the substance concentration).
[x] …the numerical formulation of the surface of the pikes in a TIC figure.
Selectivity = there is a sufficient distinguishability among individual pikes in the TIC readings.

1.3 The introduction of symbols and their definitions

Let us suppose that:

H this symbol means the presence of an analyte in a sample, the positive result,

$\bar{H}$ this symbol means the absence of a searched analyte in a test sample,

$P(H)$ the H a-priori probability (before the test with the E evidence with regard to H when the relation (8) holds true was realized),

E the result of a test/chemic. analysis (evidence),

$P(E/H)$ the probability of the E experimental evidence with regard to H. If an experimental result is positive, this probability gets closer to 1. That's not the case if the measured concentration is in the proximity of the detection limit of an apparatus/an analytical method. In that case it is advantageous to calculate the probability according to the equation (8):

$$P(E/H) = 1 - P(\bar{E}/H) \qquad (8)$$

When $P(E/\bar{H})$ is the probability of a falsely negative result with regard to $H,$

$P(E/\bar{H})$ is the probability of a positive experimental result in the H absence (the falsely positive probability). The result is falsely negative in the case of the $P(E/H)$.

$P(E/\bar{H})$ the a-posteriori probability of the presence of an analyte in a test sample with regard to an experimental evidence, the E test result,

$P(E/H) = 1$ if the accelerant was found during an experiment. The exception is the case when the measured concentration of an analyte is in the proximity of a detection limit.

One can suppose for the interpretation of the equation (3):

when $P(H/E) = $ 0.5, the probability of the rightness of the hypothesis is 50 %,

$= $ 0.99, the probability of the rightness of the hypothesis is certainly 99 %.

The next scale *of values $P(H/E)$* is designed for the interpretation of the results according to the equation (6):

<0.5   the hypothesis of the arson is irrelevant,

0.5   the hypothesis of the arson isn't confirmed or disproved,

>0.5 and <0.7 the hypothesis of the arson is probable,

>0.7 and <0.8 the hypothesis of the arson is highly probable,

>0.8   the hypothesis of the arson is certainly relevant.

2 Examples of the calculation and the interpretation of their results

2.1 According to the equation (6)

Input data based on author's experience and computing results are mentioned in Tab. 4.

Tab. 4 The calculation of the P(H/E) on the basis of the result of a chemical analysis of a fire sample

P(H)	P($\bar{H}$)	P(E/H)	P(E/$\bar{H}$)	P($\bar{E}$/H)	P(H/E)
0,5	0,5	0,99	0,03	0,01	0,9706
0,95	0,05	0,99	0,03	0,01	0,9843

Results can be interpreted with the high justification certainty of the hypothesis of the arson in both cases. If the probability of the a-priori hypothesis P(H) is increased to 0.95 (e.g. because the sample was taken in the place where a trained/certified dog signed it), the certainty of the rightness of the hypothesis rises, too.

2.2 According to the equation (7)

After the substitution of data according to Tab. 4 the value:

$$\frac{P(H/E)}{P(\bar{H}/E)} = (0,5/0,5) \cdot (0,99/0,03) = 1 \cdot 33 =$$

$$=33, \text{ když } LR = 33$$

and in the second case

$= (0.95/0.5) \cdot (0.99/0.03) = 1.9 \cdot 33 = 62.7$ and LR = 33.

It is possible to interpret the results in the way that the hypothesis of the arson is relevant/has been confirmed in the first and the second case. This conclusion supposes that a combustible liquid identified as an accelerant was not used or stored in accordance with the operational regulations and/or the evidences of responsible persons/witnesses in the given space.

Conclusion

Bayes´ theorem and its possible forms of the formulation are also usable for the fire science and fire technical expertises with the help of the a-posteriori probability of a verified hypothesis/event, the a-posteriori odds, chances/expectations of the H hypothesis with regard to the E experiment and the likelihood ratio which can be estimated on the basis of the results of the used qualitative validated tests/measurements/chemical analyses. It can be supposed that they will be exploited expertly in the same way as with the evaluation of the uncertainties of the quantitative results of the tests/measurements even under the conditions of the Fire Brigades of the Czech Republic.

References

Regulation (EC) No. 1907/2006 of the European Parliament and of the Council.

BOEF, G., HULANICKI, A. (1983). *Pure Appl. Chem.*, 1983, Vol. 55, pp. 553.

ČSN EN ISO/IEC 17025:2005. *Všeobecné požadavky na způsobilost zkušebních a kalibračních laboratoř*i.

ČSN ISO 3534-1 až 3:1994. *Statistika - Slovník a značky. Část 1: Pravděpodobnost a obecné statistické termíny.*

ČSN P ISO/TS 21748:2005. *Pokyn pro použití odhadů opakovatelnosti, reprodukovatelnosti a správnosti při odhadování nejistoty měření* (Guidance for the use of repeatability, reproducibility a trueness estimates in measurement uncertainty estimation).

ČSN EN ISO 10012:2003. *Systémy managementu měřená - Požadavky na procesy měření a měřicí vybavení.*

Dokument ILAC-G17:2002. *Zavádění koncepce stanovení nejistot zkoušení v návaznosti na aplikaci normy ISO/IEC 17025*. Praha: Český institut pro akreditaci, 2004. 5 s.

DVOŘÁK, O. (1997). *Nejistota požárních testů. Dokument AZL č. 1011.2*. Praha: TÚPO, 1997.

DVOŘÁK, O. (2005). Možnosti statistického vyhodnocení výsledků laboratorních stanovení jakostních parametrů technických prostředků PO a hasiv pro potřeby certifikace. In *Požární ochrana 2005*. Ostrava: VŠB - TUO, 2005, s. 115-119.

Eurachem (2000)/CITAC Guide CG 4. *Quantifying Uncertainty in Analytical Measurement*. 2000. 120 s.

EVVET, I.W.J., GILL, P. (1991). *Electrophoresis*, 1991, Vol. 12, pp. 226.

GINNEKEN, A.M., SMEULDERS, A.W.M. (1991). *Anal. Quant. Cytol. Histol.*, 1991.

GUM (1995) - *Guide to the Expression of Uncertainty in Measurement*. Švýcarsko: BIPM, IEC, IFCC, ISO, IUPAC, IUPAP a OIML, 1995. 101 s.

HEBÁK, P., KAHOUNOVÁ, J. (1988). *Počet pravděpodobnosti v příkladech*. Praha: Polytechnická knižnice, 1988. 309 s.

ISO 5725-1:1997. *Accuracy (trueness and precision) of measurement methods and results - Part 1: General principles and definitions.*