

Czech Technical University in Prague
Faculty of Nuclear Science and Physical Engineering
Department of Mathematics

**Advanced Bayesian Processing
of Clinical Data
in Nuclear Medicine**

Ladislav Jirsa

Ph.D. Thesis

Prague, November 1999

Acknowledgements

I would like to express my thanks to my supervisor, Doc. Ing. Jindřiška Heřmanská, CSc.. At the same time, I thank to Ing. Miroslav Kárný, DrSc. for help, advice and encouragement. Thanks also belong to Ing. Alena Halousková, CSc. and Ing. Jindřich Bůcha, CSc. for useful remarks and advice concerning this thesis. Also, I would like to thank to my colleagues in the Department of Adaptive Systems, *ÚTIA*.

This work was partially supported by these grants:

- IGA MZ ČR 4581-3,
- GA ČR 102/99/1564.

Address of the author:

Ladislav Jirsa
Institute of Information Theory and Automation
Academy of Sciences of the Czech Republic
Department of Adaptive Systems
P.O. Box 18
182 08 Prague 8
Czech Republic

e-mail: jirsa@utia.cas.cz

Contents

| | |
|--|-----------|
| Symbols and Notations | 9 |
| List of Tables | 11 |
| List of Figures | 13 |
| 1 Introduction | 15 |
| 1.1 Decision Support in Medicine | 15 |
| 1.2 Treatment of Thyroid Gland Diseases Using ^{131}I | 16 |
| 1.2.1 Clinics of Nuclear Medicine (<i>KNM</i>) | 16 |
| 1.2.2 Procedure of the Treatment | 16 |
| 1.3 Motivation — How to Process the Data | 17 |
| 1.3.1 What Is Dealt with? | 18 |
| 1.3.2 How the Problem Is Solved So Far — Examples | 18 |
| 1.4 Idea to Use Bayesian Tools for the Solutions | 22 |
| 1.4.1 Advantages of Bayesian Methodology | 22 |
| 1.4.2 Disadvantages of Bayesian Methodology | 22 |
| 1.5 Thesis Layout | 23 |
| 2 The Work — Aims and Means | 25 |
| 2.1 Aims of the Work | 25 |
| 2.2 Means and Tools Used for the Work | 25 |
| 3 Dosimetry and Biophysics | 27 |
| 3.1 Activity | 27 |
| 3.2 Effective Half-life | 28 |
| 3.3 Relative Activity | 28 |
| 3.4 Signal | 28 |
| 3.5 Background | 29 |
| 3.6 Calibration Coefficient | 29 |
| 3.7 Excretions | 29 |
| 3.8 Dose | 29 |
| 3.9 Radio-hygienic Limits | 29 |
| 3.10 Specific Irradiation of Thyroid Gland | 30 |
| 3.11 Specific Irradiation of Blood | 30 |
| 3.12 Dose Estimation — the MIRD Method | 31 |
| 4 Tools Used for the Solutions | 33 |
| 4.1 Particular Bayesian Tools | 33 |
| 4.1.1 Bayes Rule | 33 |
| 4.1.2 Chain Rule | 34 |
| 4.1.3 Likelihood Function | 34 |
| 4.1.4 Conjugated Systems of Prior <i>pdfs</i> | 34 |

| | | |
|----------|--|-----------|
| 4.1.5 | Fictitious Data | 34 |
| 4.1.6 | Sufficient Statistics and Conjugated <i>pdfs</i> | 35 |
| 4.1.7 | Symmetric Confidence Intervals | 35 |
| 4.2 | Properties of the Used Probability Density Functions | 36 |
| 4.2.1 | Logarithmic-normal <i>pdf</i> | 36 |
| 4.2.2 | Poisson Probability Function and Gamma- <i>pdf</i> | 37 |
| 4.3 | General Numerical Algorithms | 38 |
| 4.3.1 | Extreme and Root Searching | 38 |
| 4.3.2 | 1D-integration | 39 |
| 4.4 | Programming and Computing Tools | 40 |
| 5 | Solutions of the Numerical Tasks | 43 |
| 5.1 | Some Numerical Tasks | 43 |
| 5.1.1 | Confidence Interval | 43 |
| 5.1.2 | Simplex Termination Test and Convergence | 47 |
| 5.1.3 | Pseudo-normalization | 48 |
| 5.1.4 | Domain Restriction | 48 |
| 5.1.5 | Design of the Integration Table | 49 |
| 5.2 | Programming Approach | 50 |
| 5.2.1 | Numerical Objects | 50 |
| 5.2.2 | Communication with Database | 52 |
| 5.2.3 | Software Implementation | 53 |
| 6 | Solutions of the Estimation Tasks | 55 |
| 6.1 | Thyroid Gland Mass | 55 |
| 6.1.1 | Measurement | 56 |
| 6.1.2 | Model | 56 |
| 6.1.3 | Likelihood | 56 |
| 6.1.4 | Prior Information | 57 |
| 6.1.5 | Estimation | 57 |
| 6.1.6 | Implementation Remarks | 58 |
| 6.2 | Activity Kinetics | 58 |
| 6.3 | Source on the Background | 59 |
| 6.3.1 | Measurement | 59 |
| 6.3.2 | Model | 60 |
| 6.3.3 | Likelihood | 60 |
| 6.3.4 | Prior Information | 60 |
| 6.3.5 | Estimation | 61 |
| 6.3.6 | Implementation Remarks | 61 |
| 6.4 | Calibration Coefficient | 61 |
| 6.4.1 | Measurement | 62 |
| 6.4.2 | Model | 62 |
| 6.4.3 | Likelihood | 62 |
| 6.4.4 | Prior Information | 62 |
| 6.4.5 | Estimation | 62 |
| 6.4.6 | Implementation Remarks | 63 |
| 6.5 | Instantaneous Activity | 63 |
| 6.5.1 | Measurement | 63 |
| 6.5.2 | Model | 63 |
| 6.5.3 | Likelihood | 63 |
| 6.5.4 | Prior Information | 63 |
| 6.5.5 | Estimation | 63 |
| 6.5.6 | Implementation Remarks | 64 |
| 6.6 | Effective Half-life | 64 |
| 6.6.1 | Measurement | 64 |

| | | |
|----------|--|-----------|
| 6.6.2 | Model | 65 |
| 6.6.3 | Likelihood | 65 |
| 6.6.4 | Prior Information | 65 |
| 6.6.5 | Estimation | 66 |
| 6.6.6 | Implementation Remarks | 66 |
| 6.7 | Prediction of Reaching Radiohygienic Limits | 66 |
| 6.7.1 | Measurement | 66 |
| 6.7.2 | Model | 66 |
| 6.7.3 | Likelihood | 67 |
| 6.7.4 | Prior Information | 67 |
| 6.7.5 | Estimation | 67 |
| 6.7.6 | Implementation Remarks | 67 |
| 6.8 | Excretion of Activity | 67 |
| 6.8.1 | Measurement | 68 |
| 6.8.2 | Model | 68 |
| 6.8.3 | Likelihood | 68 |
| 6.8.4 | Prior Information | 68 |
| 6.8.5 | Estimation | 69 |
| 6.8.6 | Implementation Remarks | 69 |
| 6.9 | Radio-hygienic Quantities | 69 |
| 6.9.1 | Specific Irradiation of Thyroid Gland | 69 |
| 6.9.2 | Specific Irradiation of Blood | 70 |
| 6.9.3 | Medical Internal Radiation Dose — the MIRL Method | 71 |
| 7 | Experiments with Data and the Algorithms | 73 |
| 7.1 | General Testing of the Algorithms | 73 |
| 7.1.1 | Batch Processing — Test of Stability | 73 |
| 7.1.2 | Precision of QUANC8 | 73 |
| 7.1.3 | Confidence Interval | 74 |
| 7.1.4 | Simplex Method | 75 |
| 7.1.5 | Integration Table | 76 |
| 7.1.6 | Integration Table vs. Gaussian Approximation for Log-normal <i>pdf</i> | 76 |
| 7.1.7 | Run Times | 77 |
| 7.2 | Experiments with Batch Processing of Data | 78 |
| 7.2.1 | Comparison of Bayesian and Least-square T_{ef} -estimate | 78 |
| 7.2.2 | Difference Between Diagnostic and Therapeutic T_{ef} | 79 |
| 7.2.3 | Prediction of Activity in a Measurement Sequence | 82 |
| 7.2.4 | Comparison of Bayesian and Deterministic Estimate of Excretions | 83 |
| 7.2.5 | Influence of Data Quality on Thyroid Gland Mass Estimate | 83 |
| 7.3 | Experiments with Individual Estimation Tasks | 86 |
| 7.3.1 | Influence of Prior Information on the Estimate | 86 |
| 7.3.2 | Influence of Data Amount on Effective Half-life Estimate | 87 |
| 7.3.3 | Influence of t_1 on Effective Half-life Estimate | 87 |
| 8 | Conclusions | 89 |
| 8.1 | The Contributions | 89 |
| 8.2 | Open Problems | 91 |

Symbols and Notations

Mathematical symbols

| | |
|----------------------------------|--|
| α | level of confidence interval (here 0,05, see (4.7), page 35) |
| $\Gamma(x)$ | Euler gamma-function (see (4.12), page 37) |
| Δ | difference, deviation |
| Υ | see (5.19), page 50 |
| ϵ | error of a numerical method |
| ζ | prior information in some estimation tasks (see part 6.3.4, page 60 and following) |
| ϑ | parameter of a model (see part 4.1.1, page 33) |
| λ | decay constant (see (3.2), page 27) |
| ρ | fraction of diagnostic and therapeutic T_{ef} (see part 7.2.2, page 79) |
| σ | standard deviation |
| τ_s | residence time of activity in organ s (see part 3.12, page 31) |
| ψ | see (6.19), page 65 |
| Ψ | see (6.22), page 65 |
| ω_i | relative precision of thyroid gland measurement $\in (0; 1)$ (see part 6.1.1, page 56) |
| A | activity (see part 3.1, page 27) |
| A_r | relative activity in % (see part 3.3, page 28) |
| A_R | relative activity $\in (0; 1)$ (see part 3.3, page 28), $A_r = A_R \cdot 100\%$ |
| A_{apl} | applied activity (what the patient has drunk) [Bq] |
| B | mean value of background impulses (see part 6.3, page 59) |
| $\bar{D}_{s \rightarrow t}$ | dose from source organ s to target organ t (see part 3.12) |
| Ex | mean value of a random variable x |
| E_{ri} | excretions (see part 3.7, page 29) |
| $F(x)$ | distribution function of a random variable x |
| $G_\theta(x, \nu)$ | gamma- <i>pdf</i> |
| H | Bayesian estimate of thyroid gland mass [g] |
| K | normalizing constant (see $g(x)$ below) |
| \mathcal{L} | likelihood function (see part 4.1.3, page 34) |
| $L_n(x)$ | Lagrange interpolation polynome of order n (see (4.14), page 39) |
| $L_x(\mu, r)$ | logarithmic-normal <i>pdf</i> (see part 4.2.1, page 36) |
| M | mass of a patient [kg] (see part 3.11, page 30) |
| $\mathcal{N}_x(\bar{x}, \sigma)$ | normal <i>pdf</i> |
| S | activity of a standard radiation source (see part 6.4, page 61) |
| $S_{s \rightarrow t}$ | S -factor of source organ s and target organ t (see (3.12), page 31) |
| SIK | specific irradiation of blood (see part 3.11, page 30) |
| SI_{th} | specific irradiation of thyroid gland (see part 3.10, page 30) |
| T_{ef} | effective half-life (see part 3.2, page 28) |
| T_p | physical half-life (8,04 days for ^{131}I) |

| | |
|------------------------|---|
| $U(x)$ | uniform <i>pdf</i> of x |
| Z | mean value of signal without background (see part 6.3, page 59) |
| a | impulses of activity measurement |
| b | impulses of background (see part 6.3, page 59) |
| c | calibration coefficient (see part 3.6, page 29) |
| $f(x)$ | <i>pdf</i> of a random variable x |
| $f(x \vartheta)$ | <i>pdf</i> of a random variable x conditioned by a parameter ϑ |
| g_{kl} | see (5.4), page 44 |
| $g(x)$ | non-normalized <i>pdf</i> $f(x) = Kg(x)$ |
| h | integration step (see (4.15), page 40) |
| h_{palp} | palpation measurement of thyroid gland mass (see part 6.1.1, page 56) |
| h_{sono} | sonographic measurement of thyroid gland mass (see part 6.1.1, page 56) |
| $l_{\frac{1}{2}}$ | half-width (see (5.13), page 46) |
| m | No. of integration steps (see (4.16), page 40) or No. of measurements (see part 6.6.2, page 65) |
| m_{kl} | number of integration steps (see (5.2), page 44) |
| n | dimension |
| r_i | transformed relative precision of thyroid gland measurement ω_i (see (6.1), page 56) |
| s | impulses of a standard radiation source (see part 6.4, page 61) |
| t_1 | time from which $A(t)$ is considered as exponential (see (6.19), page 65) |
| $\hat{x} \pm \delta x$ | Bayesian estimate of x (see (4.8), page 36) |
| x_i | value of x in time t_i |
| $x(.n)$ | x_1, x_2, \dots, x_n |

Physical units

| | |
|----|---|
| Bq | becquerel [s^{-1}], unit of activity, number of nuclear changes per 1 s |
| Ci | curie, old unit of activity, 1 Ci = 37 GBq |
| Gy | gray [J/kg], unit of dose, energy absorbed in 1 kg of mass |

Abbreviations

| | |
|-------------|---|
| <i>KNM</i> | Clinics of Nuclear Medicine, Faculty Hospital Motol, Prague |
| <i>LS</i> | least squares |
| <i>pdf</i> | probability density function |
| RH | radio-hygienic |
| <i>ÚTIA</i> | Institute of Information Theory and Automation, Academy of Sciences of the Czech Republic, department of Adaptive Systems |

List of Tables

| | | |
|------|---|----|
| 5.1 | Coefficients g_{kl} (see (5.5)) | 44 |
| 5.2 | Values of $\mathcal{N}_x(0, 1)$ with corresponding ϵ_a | 47 |
| 7.1 | Numbers of integration steps found theoretically (m_{kl} , see (5.5), page 44) and experimentally ($m_{\epsilon_{max}}$) | 74 |
| 7.2 | Mean value and upper limit of confidence interval $\alpha = 0,04999570$ with various ϵ_{max} and ϵ_a . Exact values: $\hat{x}_{exact} = 0$, $x_{u,exact} = 1,96$ | 75 |
| 7.3 | Testing of the integration table for the maximum required error of interpolation 10^{-4} | 76 |
| 7.4 | Numbers of integration steps found theoretically (m_{kl} , see (5.5), page 44) and experimentally ($m_{\epsilon_{max}}$) for logarithmic-normal <i>pdf</i> | 76 |
| 7.5 | Confidence intervals limits for thyroid gland mass estimate. Difference of results determined using integration table and Gaussian approximation | 77 |
| 7.6 | Runtimes of numerical programs | 78 |
| 7.7 | Arithmeric mean and standard deviation of ρ — ratio of diagnostic and subsequent therapeutic T_{ef} , see (7.2) | 79 |
| 7.8 | Frequencies of different numbers of measurements m in diagnostics and therapy | 80 |
| 7.9 | Bayesian and least-squares relative differences of $t(A_l)$ prediction | 82 |
| 7.10 | Comparison of deterministic and Bayesian estimate of excretions | 83 |
| 7.11 | Relative difference of palpation and sonographic data $\Delta_r(h_{sono}, h_{palp})$ for thyroid gland mass estimation | 84 |
| 7.12 | Arithmetic means and standard deviations of $\Delta_r(H, h_{palp})$ and $\Delta_r(H, h_{sono})$ | 84 |
| 7.13 | Estimation of thyroid gland mass with sonographic priors, $\hat{h}_0=0,998$, $\nu_0 = r/3,395$ | 86 |
| 7.14 | Estimation of thyroid gland mass with palpation priors, $\hat{h}_0=7,69$, $\nu_0 = r/1,55$ | 86 |
| 7.15 | Estimation of thyroid gland mass with no prior information, $\hat{h}_0=1$, $\nu_0=0$ | 87 |
| 7.16 | Dependence of T_{ef} on number of data m | 87 |
| 7.17 | Dependence of T_{ef} on t_1 for $m = 8$ | 88 |
| 7.18 | Dependence of T_{ef} on t_1 for variable m | 88 |

List of Figures

| | | |
|------|--|----|
| 1.1 | Computation scheme for effective half-life T_{ef} and relative activities | 19 |
| 1.2 | Computation scheme for excretions E_{Ri} | 20 |
| 1.3 | Computation scheme for radio-hygienic quantities | 20 |
| 5.1 | Classes hierarchy | 51 |
| 6.1 | Bit field to control the estimation of thyroid gland mass | 58 |
| 7.1 | Comparison of Bayesian and least-squares estimate of T_{ef} , histogram of $\Delta_r(\hat{T}_{ef}, T_{ef}^{LS})$, filtered for $-1 \leq \Delta_r(\hat{T}_{ef}, T_{ef}^{LS}) \leq 1$, used 2 787 records | 79 |
| 7.2 | Histogram of T_{ef} [days], estimates in diagnostic phase, 3 412 records | 80 |
| 7.3 | Histogram of T_{ef} [days], estimates in therapeutic phase, 1 522 records | 80 |
| 7.4 | Histogram of T_{ef} [days], estimates for number of measurements $m \leq 3$, 3 348 records | 81 |
| 7.5 | Histogram of T_{ef} [days], estimates for number of measurements $m > 3$, 1 591 records | 81 |
| 7.6 | Histogram of $\Delta_r(t_l, \hat{t}_l)$ (see (7.1)), Bayesian prediction of $t(A_l)$, 989 records | 82 |
| 7.7 | Histogram of $\Delta_r(t_l, t_l^{LS})$ (see (7.1)), least-squares prediction of $t(A_l)$, 989 records | 83 |
| 7.8 | Scatterplot of h_{sono} vs. h_{palp} for thyroid gland mass measurement, 487 records | 84 |
| 7.9 | Histogram of $\Delta_r(H, h_{palp})$, 487 records | 85 |
| 7.10 | Histogram of $\Delta_r(H, h_{sono})$, 487 records | 85 |

Chapter 1

Introduction

One of the ways to treat diseases of thyroid gland is use of nuclear medicine. During the treatment, a lot of various biophysical data can be measured. Potential improvement of the treatment results can be reached, among others, by application of such a data processing methodology that corresponds to nature and properties of these data. Valid information extracted from the data can support medical decisions.

Bayesian methodology was found as a suitable tool for this purpose.

Application of the Bayesian theory to data processing includes both theoretical analysis and practical implementation. Work described in this thesis focuses mostly on practical (numerical and algorithmical) realization of Bayesian estimates derived before.

Meaningful results can be analyzed statistically. The conclusions could contribute to further improvement of the treatment.

This chapter guides through the specific problems addressed in this thesis. In section 1.1, general problem of information and decision in medicine is introduced. Section 1.2 specifies these general ideas to a frame of the Clinics of Nuclear Medicine, Motol Hospital, Prague (*KNM*), and to the treatment of thyroid gland diseases. In section 1.3, the situation of information processing at the *KNM* is outlined using some terms that are defined later in this thesis in detail. It is shown on several examples that the methods of computations on the data, that are used so far at the *KNM*, yield results that are not suitable for subsequent responsible medical decision. In section 1.4, Bayesian methodology is proposed as a tool for the solutions and both its advantages and disadvantages are mentioned. In the last section 1.5, the thesis layout is described by chapters.

1.1 Decision Support in Medicine

Every rational decision requires sufficient amount of relevant information.

In medical treatment, there are two main steps of information processing and decision making: diagnosis and therapy.

In the diagnostic step, information as complex as necessary about health state and its causes is gathered and medical conclusions formulated. In the therapeutic step, decision about some actions according to the previous conclusion is made and the appropriate actions are carried out.

Biophysical Examination as a Source of Data

To collect the facts of interest, various examination methods are used including biophysical ones. Biophysical methods rely on interaction of physical fields with biological matter. Quantitative results of direct measurements yield some data that directly or indirectly serve as information source and support for further medical decisions.

Data Processing Is a Part of Medical Decision

Mostly, the biophysical data cannot be used directly and some processing is necessary to transform the data to a desired form. This processing then becomes a part of examination and it can influence the medical decisions, consequently the therapeutic action and its result (success).

This influence, of course, depends on the importance of the given examination or measurement for the medical decision.

Reality — Data Quality Is Given

The term “data” (see part 1.3) denotes outputs of some biophysical examinations (see Chapter 3, page 27). As biophysical examination is a measurement of some physical quantity, it is loaded by uncertainty given e.g. by measurement error or other reasons.

The effort should be made to eliminate or decrease the data uncertainty as much as possible. If it is not possible, either principally or practically, such methods of data processing must be used that can treat the data uncertainty so that

- the uncertainty is not increased, preferably it is decreased,
- the uncertainty is quantified so that reliability of the result is known.

This thesis deals with processing of limited amount of noise-polluted data that serve as one of information sources for medical decisions in the treatment procedure.

1.2 Treatment of Thyroid Gland Diseases Using ^{131}I

Here, the adopted practice of the *KNM* is reviewed that should be supported by the work described in this thesis.

1.2.1 Clinics of Nuclear Medicine (*KNM*)

Clinics of Nuclear Medicine, Faculty Hospital Motol, Prague (*KNM*), is the world famous center for treatment of thyroid gland diseases using ^{131}I . For more than 30 years it treats successfully thyroid gland carcinoma, thyreo-toxicoses and other illnesses.

Few years ago, a team of engineers and researchers from the *KNM* and Institute of Information Theory and Automation, Academy of Sciences of the Czech Republic (*ÚTIA*), started to focus on data, obtained by measurement of various quantities during patients' examinations, and on the results given by traditional way of these data processing. They concentrated on quantification of uncertainty of the data and the results. They addressed relevance of using appropriate mathematical tools for data processing [13] and their influence on validity of the results [14].

Several projects dedicated to application of suitable mathematical methods in data processing were successfully defended, supported by Grant Agency of the Czech Republic (GA ČR 312/94/0679) and European Union (EU COST OL B2.20) in years 1994–96. Currently, a project for influence of biophysical factors on thyroid gland cancer treatment, supported by Internal Grant Agency of Ministry of Health, Czech Republic (IGA MZ ČR 4581-3), is running.

1.2.2 Procedure of the Treatment

Thyroid gland tumor belongs to diseases treated at the *KNM*.

One of the ways to diagnose and cure thyroid gland tumor is an internal administration of unsealed radioactive iodine [27], [36], mostly ^{131}I [11]. This element is selectively accumulated in thyroid gland where, if radioactive, it produces ionizing radiation. This radiation can be used for diagnosis of the organ state (e.g. using scintigraphy) or for therapy — destruction of the affected tissue.

If a tumor of thyroid gland is diagnosed, it is usually removed invasively by surgery. The role of radiation treatment is to destroy any tumor remnants that were not possible to be removed by surgery.

The procedure of radiation treatment consists of two steps.

Diagnosis. In this step, the patient is supplied orally ^{131}I as sodium iodide or potassium iodide of certain low activity, usually about 70–100 MBq. Then several biophysical measurements and medical examinations follow, yielding information about accumulation of ^{131}I in the area of thyroid gland, ^{131}I distribution in the organism etc.

Therapy. The diagnosis can indicate tumor tissue or its remnants that must be destroyed by radiation. This destruction is called therapy. If therapy is necessary, the therapeutic activity is individually determined for the patient. ^{131}I of this activity is supplied to the patient. In the thyroid gland, where this activity is accumulated, the tumor tissue is destroyed by the radiation. During therapy, biophysical measurements are also performed. Data collected during this step are processed for checking and hygienic purposes.

Therapeutic activity is usually approximately by two orders higher than diagnostic one, in the range 2–7 GBq.

The therapeutic activity of iodine must be high enough so that radiation is able to destroy the target tissue. Because of radiation influence to other tissues, the therapeutic activity must be, on the other hand, sufficiently low to decrease the radiation risk, i.e. damage of other tissues (e.g. bone marrow). In other words, the supplied therapeutic activity must be as low as possible with the effect of therapy guaranteed at the same time.

The determination of the therapeutic activity is given by two points of view: objective (dosimetric and other values estimation) and subjective (subsequent medical decision). In the first step, the values of some dosimetric quantities based on the measurements during the diagnostic phase are estimated. These values then significantly support the ultimate physician's decision, taking into account also the patient's state, disease stage etc.

The values of quantities influencing the physician's decision must be therefore calculated in the way so that imprecisions and errors do not increase or decrease them artificially, or, in other words, as precise as possible.

In clinics of nuclear medicine all over the world that deal with thyroid gland diseases, no standardized approach to therapeutic activity estimation is adopted. There are discussions if therapeutic activity should be individualized or not. One way is to individualize the therapeutic activity using various criteria (mass of thyroid gland, mass of the patient, kinetics of iodine in the organism etc.) [8], [34], [9], another way is to apply constant activity to each patient [41] or constant activity to mass unit of thyroid gland [4]. Generally, therapeutical activities applied in the world vary from 0,85 GBq to 11 GBq.

At the *KNM*, therapeutic activity is individualized just roughly. Practically, there are few values used as possible therapeutic activity (80 mCi, 100 mCi, 120 mCi etc.) and one of them is chosen to be supplied to a patient, according to medical decision. Finer variation of individual therapeutic activities has not been taken into account so far.

Use of relevant mathematical methods yields meaningful results in computation of estimates of important quantities on clinical and biophysical data. These estimates can be statistically analyzed because a large set of patients' data, collected for many years, is available. *Output of this analysis can indicate ways of finer individualization of therapeutic activity and further improvement of the treatment.*

1.3 Motivation — How to Process the Data

To illustrate the relationships between biophysical data and results of interest together with complexity of the data processing and its pitfalls, this section shows simplified schemes of quantities and computations defined in detail in Chapter 3, page 27.

1.3.1 What Is Dealt with?

The following figures show connections and relationships between measured data and biophysical or dosimetric results.

The graphical notation keeps these conventions:

- biophysical data (results of direct measurements of biophysical quantities) are framed by thin solid-dashed lines,
- intermediate biophysical quantities playing a role of inputs to other computations and carrying information that can be used for medical decisions are framed by half-thick solid lines,
- radio-hygienic quantities playing important role in medical decisions are framed by thick solid lines,
- other quantities are framed by thin solid lines.

The Figure 1.1 shows computational scheme for effective half-life and relative activities of the organ measured. The magnitude of relative activities can indicate necessity of further examination, the effective half-life is an important input quantity for doses estimations.

The Figure 1.2 shows computational scheme for so called excretions that are used as input quantity for blood dose estimation.

The Figure 1.3 shows computational scheme for radio-hygienic quantities that are results of interest carrying information about the patient's body reaction to application of radioactive iodine. This information serves as a direct criterion for further medical decisions.

A brief summary of meaning of the quantities mentioned above is given here:

Data are quantities directly measured, namely, impulse counts of background and signal on background, thyroid gland mass and patient's mass.

Radio-hygienic quantities are called the following quantities:

- Specific irradiation of thyroid gland that is a dose absorbed in thyroid gland tissue. The required destructive dose is 1200 Gy.
- Specific irradiation of blood that is a dose absorbed in blood. It must be less than 4 Gy and its anorganic fraction must be less than 1 Gy to avoid damage of blood and bone marrow (see part 3.11, page 30).
- Prediction of reaching requested activity by the a patient's body is important for planning the patient's treatment in the *KNM*.

Other quantities like relative activities, effective half-life, excretions etc. that are also used for medical decision.

1.3.2 How the Problem Is Solved So Far — Examples

Simple algebraic formulae to estimate biophysical, radio-hygienic and other quantities [21], [22], [36] are used so far at the *KNM*. The data measured on patients are substituted into these formulae.

There are two main specific features of these data:

Limited amount. Due to high number of patients, capacity of measurement devices, established measurement methodology and also economical limits of the *KNM*, only few measurements of each patient can be performed. These conditions lead to lack of data in majority of cases.

Uncertainty. Available data are corrupted with uncertainty due to (a) measurement techniques of low precision (thyroid gland mass), (b) variation of measurement conditions (calibration coefficient changes along sequence of excretions measurements) or — most frequently — (c) by the physical nature of the process (counting particles of ionizing radiation in some fixed time interval).

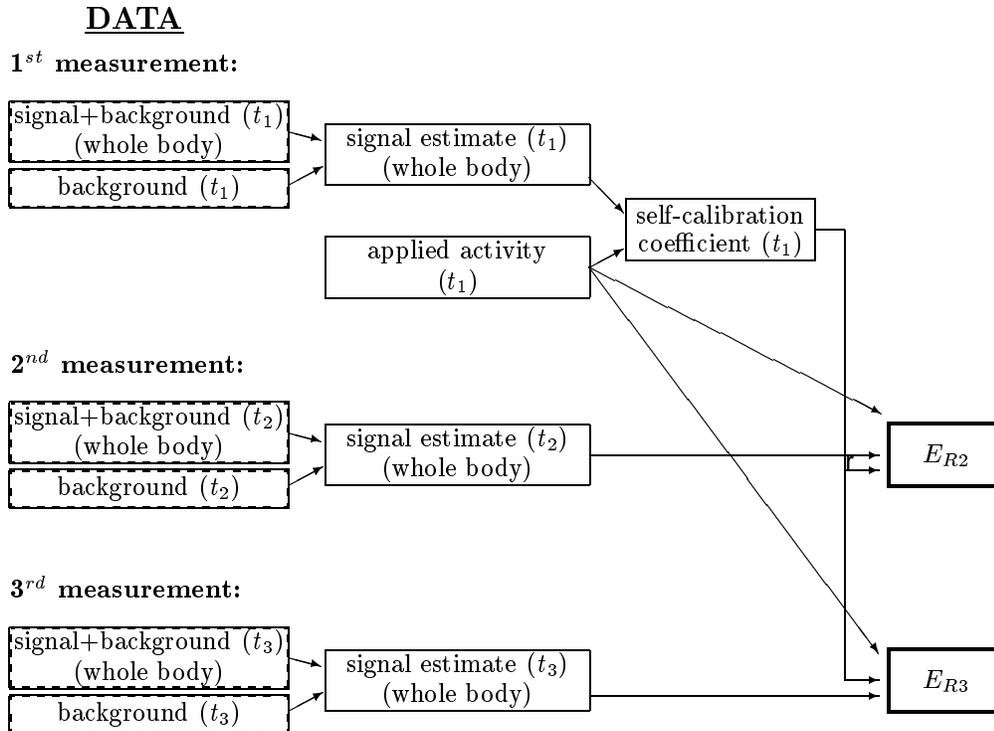
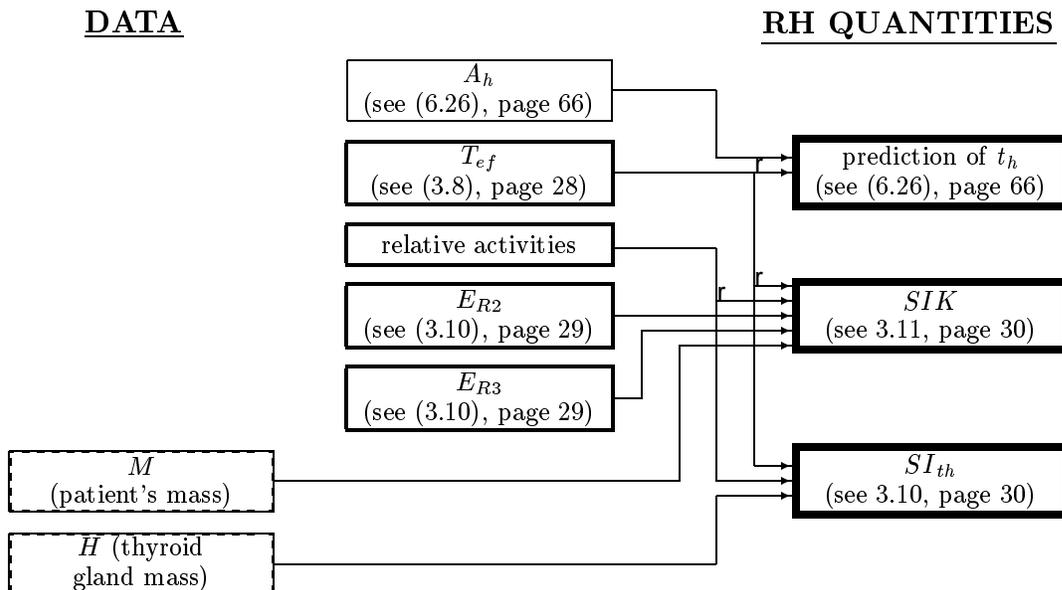
Figure 1.2: Computation scheme for excretions E_{Ri} 

Figure 1.3: Computation scheme for radio-hygienic quantities

The algebraic formulae mentioned above do not respect randomness and uncertainty of the data and take them as deterministic values. Therefore this algebraic processing is sometimes called *deterministic*.

Using the traditional simple deterministic formulae, the data uncertainties are cumulated in the uncertainty of the result by the mechanism of errors propagation. The uncertainty of the result obtained in this way is unknown and such values have sometimes even no physical meaning [11].

We will show few examples when traditional deterministic way of data processing can fail and yield meaningless values. For detailed definitions of the terms, see Chapter 3.

Example 1: Source on the Background

When signal Z of a radioactive source is measured (see parts 3.4 and 3.5, page 29), due to presence of a background B , what is obtained is the composition of the background and the signal $B + Z$. Both B and Z are random quantities. If $B + Z \gg B$ (case of higher activities), then Z can be estimated trivially in the deterministic way by subtraction $Z = (B + Z) - B$. In case of $B + Z \approx B$ (case of lower activities), the subtraction is loaded by a great error and, because of fluctuations, its result can be even negative, which is a physical nonsense [15]. This situation can be often met during staff contamination measurement [11]. Therefore some more reliable way of Z estimation is necessary.

Example 2: Excretions

Excretions are relative activities diluted by urine (see part 3.7, page 29). Activity is determined according to (3.9), page 29. For calibration coefficient c estimation (see part 3.6, page 29), two signal estimates are necessary: signal of the measured source and of the standard source of radiation. Furthermore, calibration coefficient for excretions changes in an unknown way (see part 6.8.4, page 68).

The computation scheme of excretions is shown on the Figure 1.2.

Fluctuation of signals together with the unknown change of calibration factor can cause the deterministic relative activity estimate so uncertain, that these cases can occur:

- excretions are greater than 100%,
- excretions are negative,

i.e. the patient seems to drink ^{131}I “secretly” or he seems to be ^{131}I -generator. It is obvious that these cases are physically meaningless.

Example 3: Effective Half-life

Effective half-life, according to the Figure 1.1, is obtained by fitting the model (6.19), page 65, to a sequence of activities $\{A_i\}$ in times $\{t_i\}$. Problems with activities estimation were mentioned in excretions description in the previous example. During the deterministic effective half-life estimation, a linear regression is used for points given by $(t_i, \ln A_i)$. Position of the regression line determines the effective half-life.

Due to fluctuations in the data and absence of any constraints (see (3.7), page 28), these cases can occur:

- effective half-life is greater than physical half-life,
- effective half-life is negative,

These cases have no physical meaning.

Example 4: Unknown Uncertainty

Even if the results of the estimation tasks mentioned above in the deterministic way seem to be physically meaningful ($Z > 0$, excretions and effective half-life values are within their physical bounds), it is not known how *precise* they are, i.e. what is the magnitude of their *uncertainty*.

Considering that excretions and effective half-life enter the estimations of the important radio-hygienic quantities (see Fig. 1.3), we can find out that the final deterministic estimates can be even meaningless or at least their precision is unknown. Such estimates can hardly serve as a reliable source of a responsible medical decision.

1.4 Idea to Use Bayesian Tools for the Solutions

The quantities to be estimated can be understood as results of some processes, both physical and biophysical. Majority of these processes can be successfully modelled. For some quantities, expert knowledge (e.g. physical bounds of the values) is available. This situation is suitable for using the Bayesian methodology.

As for systematic use of Bayesian methodology in this field, no such attempt was so far registered. This work is a part of effort to use Bayesian approach in nuclear medicine which is a new application area.

1.4.1 Advantages of Bayesian Methodology

Bayesian methodology has been used in *ÚTIA* for various purposes for many years. It is described thoroughly in [29]. It is a principally simple way to combine multiple information sources:

- **model** of the system we observe, i.e. how *would* the system behave if the parameters of the model *were* known,
- **data**, i.e. measured quantities, what *is* known,
- **expert knowledge** (so called *prior information*), i.e. what is known *in advance* about the model parameters, by analysis or by experience.

Formally, quantities to be estimated are unknown parameters of the model.

These sources of information are combined together in a consistent way. If a nature of the modelled system is probabilistic, it is described by a probabilistic model. Estimate of the unknown quantity (parameter) has then a form of a probability density function (*pdf*). Therefore it carries implicitly an information about the estimate uncertainty and any interval estimate can be constructed using the *pdf*. The *pdf* can be used as an input information for another estimate without information loss.

The Bayesian methodology is suitable to be used in cases of small uncertain data sets and existence of expert knowledge. Then more information can be included into the estimation process and uncertainty of the estimate can be reduced.

At the *KNM*, amounts of data related to individual patient are usually strongly limited and data themselves are uncertain, as mentioned above. At the same time a team of experienced physicians and physicists, whose knowledge can be used as a prior information, is accessible ([17], [3], [18]).

Some results (e.g. [11], [12], [16], [14]) indicate, that the Bayesian methodology is a suitable tool to solve the class of tasks formulated in this thesis.

More about the Bayesian tools used in this work can be found in Chapter 4.

1.4.2 Disadvantages of Bayesian Methodology

There are two main groups of disadvantages that occur with using the Bayesian methodology: analytical and numerical.

Among analytical disadvantages, the most annoying that one can encounter are:

- Impossibility to find a conjugated system of *pdf*s (see part 4.1.6, page 35). If such a system is not found, one cannot take advantage of small-dimensional sufficient statistics which significantly technically simplify the estimation. As an example of such a difficulty, see (6.24), page 66.
- Impossibility to integrate a posterior *pdf* analytically. This problem causes difficulties to normalize the posterior *pdf* (see (4.3), page 33) and also prevents to integrate-out redundant parameters analytically.
- “Curse of dimensionality” — problem of growing of dimension of the task with increasing amount of data [32], which fortunately does not occur in this thesis.

The most significant numerical disadvantages are:

- computational obstacles to evaluate non-normalized function (see part (5.1.3), page 48), connected with analytical integrability,
- great differences in variance with different amount of data included into estimation and subsequent evaluation problems,
- evaluation problems in some cases even if an explicit formula is given and the posterior *pdf* is normalized [6], [10], [24], [5].

Successful estimation using the Bayesian methodology represents a complex task with many kinds of technical drawbacks and problems that must be solved. The work in this thesis focuses especially on numerical part of the Bayesian estimation and production of reliable algorithms giving results with a requested numerical precision.

1.5 Thesis Layout

This thesis is divided into 8 chapters.

Chapter 1, you are just reading, gives a general introduction to problems treated in this thesis and specifies them to the particular fields — nuclear medicine and Bayesian methodology.

In Chapter 2, the specific tasks are enumerated and the tools used to reach the aim are mentioned.

The data and the results of their processing are defined and described in Chapter 3. The quantities that will be estimated are specified.

The particular tools, both theoretical and mathematical, are described in Chapter 4 with general numerical algorithms and programming ideas used for the solutions.

For the specific tasks, the general numerical methods and algorithms should have been tailored to fit the requested purpose. In Chapter 5, the procedures are specified and rules are derived with emphasis on numerical precision and keeping requested uncertainty limits. Programming approach is described.

Chapter 6 collects complete description of estimation tasks of the specific quantities, including theoretical derivation and algorithmical implementation.

In Chapter 7, the results derived in Chapter 5 are tested, together with benchmarks of some algorithms. Then experiments with selected estimation tasks derived in Chapter 6 and their results are performed both with single cases and batch data processing. For these experiments, real patients' data are used.

In Chapter 8, conclusions are made and ways of further research are outlined.

Chapter 2

The Work — Aims and Means

In this chapter, general aims and tasks are specified and tools used for the solution are outlined.

2.1 Aims of the Work

The main aim of this work is to contribute to quality improvement of the treatment of thyroid gland diseases at the *KNM*.

Estimation of radiohygienic and other quantities mentioned in Chapter 1 should be improved. This improvement should yield more reliable values of the estimated quantities, together with their precisions, as one of the sources of responsible medical decisions.

The consequences of the estimation improvement should result in a contribution to improve therapy of thyroid gland diseases.

The requested output of this work is a set of numerically stable tested real-time programs for estimation of the quantities mentioned above.

The specific tasks of the work are:

- to formulate estimation tasks from theoretical point of view,
- to elaborate the algorithms for estimation of various physical, medical and dosimetric quantities used in the field of radio-diagnosis/therapy of thyroid gland diseases,
- to test their numerical precision,
- to test their numerical stability on a large set of real clinical data,
- to implement these algorithms on the level applicable under routine conditions at the *KNM* and to substitute them for the methods of data processing used there so far,
- to explore and test the improvement of quality of the estimates,
- according to the estimation results, to propose hints where improvements of data measurement methodology is necessary.

2.2 Means and Tools Used for the Work

The main theoretical tool to be used is the Bayesian methodology.

By using the *Bayesian* methodology, estimation of radiohygienic and other quantities should be improved in the sense of physical meaning (physically meaningless estimates avoided) and precision of the results (more information involved in the estimation process), compared to the *deterministic* estimation (see part 1.3.2, page 18) used so far.

Furthermore, the “improved” values of the estimated quantities can be processed statistically. Results of this processing could indicate ways of finer individualization of therapeutic activity supplied to patients.

Specifically, a complex and interdisciplinary area is encountered where the blend of tools and subtasks is dealt with. Particularly, it is:

- modelling and identification,
- numerical algorithms,
- software and programming,
- dosimetry and radiation protection,
- databases and data management,
- data analysis.

General numerical methods and principles are adopted and tailored to the specific problems.

Chapter 3

Dosimetry and Biophysics

In this chapter, definitions of selected dosimetrical and biophysical quantities (e.g. [27]) are formulated and collected.

3.1 Activity

Activity is a physical unit describing quantitatively a physical phenomenon *radioactivity*.

Radioactive decay. If we have a sample of atoms with unstable (radioactive) nuclei, then radioactive change (decay) takes effect. Each decay results in change of the nucleus structure and consequent emission of one or more ionizing particles.

Consider a sample of atoms with $N(t)$ radioactive nuclei in time t . Due to the radioactivity, the number $N(t)$ of radioactive (undecayed) nuclei decreases in time. Therefore the number of undecayed nuclei in time $t + \Delta t$ is

$$N(t + \Delta t) = N(t) - \Delta N(t, \Delta t), \quad (3.1)$$

where $\Delta N(t, \Delta t) \geq 0$ is number of nuclei that decayed within the time interval $\langle t, t + \Delta t \rangle$. Value of Δt is chosen small.

As the radioactive decay is a random process, $\Delta N(t, \Delta t)$ is a discrete random quantity. Hence $N(t + \Delta t)$ is also a random quantity. Denoting a mean value $\mathbf{E}N(\cdot) \equiv \bar{N}(\cdot)$, the equation (3.1) can be written for these mean values as $\bar{N}(t + \Delta t) = \bar{N}(t) - \bar{\Delta N}(t, \Delta t)$. These mean values are functions continuous in time.

If we assume mutual independence of nuclear decays, we can state that the decrease of the undecayed atoms is proportional to their number, i.e. $d\bar{N}(t) = -\lambda\bar{N}(t)dt$. Hence $\bar{N}(t)$ evolves like

$$\bar{N}(t) = N_0 \cdot e^{-\lambda t}, \quad (3.2)$$

where $N_0 = \bar{N}(t_0)$ is the initial number of the nuclei in time t_0 .

The coefficient $\lambda > 0$ is called *decay constant* and it is specific for each nuclide (element with the particular nucleus structure).

Activity. The *activity* $A(t)$ of a radioactive sample is defined as

$$A(t) = -\frac{d\bar{N}(t)}{dt}, \quad (3.3)$$

which is a mean number of radioactive changes in a time unit. Then the equation (3.2) can be written as

$$A(t) = A_0 \cdot e^{-\lambda t}, \quad (3.4)$$

where A_0 is an initial activity in time t_0 .

Another possible expression for activity is $A(t) = \lambda \bar{N}(t)$.

The physical unit of the activity is Bq (*becquerel*), which is a mean number of nuclear changes per one second.

(Physical) half-life T_p of the given nuclide is a time interval in which one half of the radioactive nuclei in the sample takes a change. The half-life is related to the decay constant λ by

$$T_p = \frac{\ln 2}{\lambda}. \quad (3.5)$$

The physical half-life is always positive.

Decay law. Introducing the physical half-life (3.5) into the equation (3.4) we get

$$A(t) = A_0 \cdot \exp\left(-\frac{t}{T_p} \ln 2\right), \quad (3.6)$$

which is called *decay law*.

3.2 Effective Half-life

If some portion of a radioactive element is incorporated into a human body, the element is eliminated by natural mechanisms and its amount decreases in time. In this way the activity contained in the body decreases. Furthermore, the element decays according to the decay law (3.6). Then there are two mechanisms of the activity decrease: biological and physical.

The *effective half-life* T_{ef} of the given element in the body is a time interval, in which the activity in the body decreases to one half due to the both biological and physical mechanisms. For the effective half-life, the inequality is valid

$$0 < T_{ef} \leq T_p, \quad (3.7)$$

where T_p is a physical half-life (3.5).

Time dependence of activity in this case is modelled by various ways. The most common and simple way, used also in this thesis, is the exponential decrease

$$A(t) = A_0 \cdot \exp\left(-\frac{t}{T_{ef}} \ln 2\right). \quad (3.8)$$

The effective half-life is of the key importance for estimation of some radio-hygienic quantities (see sections 3.10, page 30, etc.).

3.3 Relative Activity

Relative activity $A_r(t)$ is the instantaneous activity $A(t)$ inside the patient's body or eliminated from his body, divided by the applied activity A_0 corrected to a physical decay and multiplied by 100%, i.e.

$$A_r(t) = A_R(t) \cdot 100\%,$$

where

$$A_R(t) = \frac{A(t)}{A_0 \exp\left(-\frac{t}{T_p} \ln 2\right)}.$$

3.4 Signal

The *signal* is a count of particles of ionizing radiation emitted by a radioactive source of interest and detected by the detector. Signal is measured within some fixed time interval. It reflects activity of the radioactive source. Signal will be usually denoted by s .

3.5 Background

During the signal measurement, there are usually more radiation sources than the one of our interest. These sources are mostly cosmic radiation, natural radioactivity and eventually some random local radioactive pollution. These “additional” sources produce disturbing signal that usually cannot be removed. Signal from such sources is called radiation *background* and it is usually denoted by b .

3.6 Calibration Coefficient

Direct measurement of activity is usually impossible, therefore two signals must be compared: one from a source of a known (standard) activity and another of the unknown (measured) activity. The coefficient relating the signal magnitude to the activity is called *calibration coefficient* and denoted by c . It is defined as

$$s = c A, \quad (3.9)$$

where s is the signal corresponding to the activity A . The calibration coefficient reflects the geometrical arrangement of the measurement, efficiency of the detection, time interval during which the measurement was performed etc.

The standard source should be of the same (or at least similar) kind as the measured one according to kind of radiation, the number of particles emitted after one radioactive change, energy of the emitted particles etc.

3.7 Excretions

“*Excretions*” are called relative activities E_{r1} and E_{r2} eliminated by urine in time intervals $\langle t_{apl}, t_{apl} + 24h \rangle$ and $\langle t_{apl} + 24h, t_{apl} + 48h \rangle$, respectively, and related to the applied activity decreased by radioactive decay in time instants $t_{apl}+24h$ or $t_{apl}+48h$, respectively. t_{apl} means time when the radioactive element was applied to the patient (application time). Mathematical definition can be written as

$$E_{ri} = \frac{A_{i-1,i}}{A_{apl} \exp\left(-\frac{t_i - t_{apl}}{T_p} \ln 2\right)} \cdot 100\%, \quad i = 1, 2, \quad (3.10)$$

where $A_{i-1,i}$ means activity outside the body that was eliminated between t_{i-1} and t_i , measured (hypothetically) in time t_i , A_{apl} is applied activity, T_p is physical half-life. If the calibration measurement is denoted by the subscript 0, then usually $t_0 \approx t_{apl}$, $t_1 \approx t_{apl} + 24$ hours and $t_2 \approx t_{apl} + 48$ hours.

These values are important for an estimation of a specific irradiation of blood (see the section 3.11, page 30).

3.8 Dose

Dose is an energy of ionizing radiation absorbed by a unit mass of a body tissue. It is a measure of radiation effects on tissue. Its unit is $\text{Jkg}^{-1} = \text{Gy}$ (gray).

3.9 Radio-hygienic Limits

There are two kinds of radio-hygienic limits that must be respected: dose limits to protect the patient and activity limits to protect the patient’s neighbourhood.

The dose limits are addressed in the parts 3.10 and 3.10 below together with the quantities definitions.

To formulate activity limits, a patient with some activity inside his body is viewed as a source of ionizing radiation. To keep radio-hygienic safety rules and protect other people from this radiation, there are several activity limits that determine the patient's daily life at the *KNM*.

If the activity is (see [11])

higher than 74 MBq, the patient's urine is stored in a special safety container, the patient has a special distinguished pyjamas and he is separated in a single-bed shielded room without permission to leave it even within the territory of the *KNM*,

between 74 MBq and 37 MBq, the patient gets ordinary pyjamas, he is permitted to leave the shielded room and to move within the territory of the *KNM*,

between 37 MBq and 7,4 MBq, the patient's urine need not be separated,

less than 7,4 MBq, the patient can be released from the *KNM*.

Due to the treatment organization, it is important to predict when the given limits will be reached by the patient.

3.10 Specific Irradiation of Thyroid Gland

Specific irradiation of thyroid gland SI_{th} is a dose absorbed in the thyroid gland. According to Marinelli's method ([21], [36]), the formula related to 100 mCi (i.e. 3,7 GBq) of applied ^{131}I is

$$SI_{th} = \frac{139 A_{rmax}}{H} T_{ef}, \quad (3.11)$$

where A_{rmax} is maximum relative activity of the thyroid gland [%], H is thyroid gland mass [g] and T_{ef} is effective half-life [days]. Unit of SI_{th} is Gy (=J/kg).

To obtain its absolute value, the formula must be multiplied by the fraction $3\,700 \text{ MBq} / A_{apl}$ [MBq]. Required dose to destroy thyroid gland tumor is 1 200 Gy [28].

3.11 Specific Irradiation of Blood

The maximum *specific irradiation of blood* SIK_{max} is an upper estimate of a dose absorbed by blood. There is distinguished *anorganic fraction* $SIK_{an,max}$ caused by iodine in anorganic chemical form and *organic fraction* $SIK_{org,max}$ caused by iodine in form of thyroxine

$$SIK_{max} = SIK_{an,max} + SIK_{org,max},$$

where

$$\begin{aligned} SIK_{an,max} &= 160 [E_{r1} + E_{r2}] / M, \\ SIK_{org,max} &= 13,9 [100 - E_{r1} - A_{r1}] T_{ef}. \end{aligned} \quad (3.12)$$

E_{r1} , E_{r2} are excretions [%] (see 3.7), M is the patient's mass [kg], A_{r1} is relative activity in thyroid gland 24 hours after application [%] and T_{ef} is effective half-life [days] (see part 3.2). The unit of SIK_{max} is mGy (=mJ/kg).

Anorganic fraction of irradiation is more dangerous for blood.

The formula (3.12) (according to [36]) is related again to 100 mCi (i.e. 3,7 GBq) of applied activity. To obtain its absolute value, the formula must be multiplied by the fraction $3\,700 \text{ MBq} / A_{apl}$ [MBq].

The anorganic fraction must not exceed 1 Gy and the total irradiation of blood must be less than 4 Gy to avoid damage of blood and bone marrow [11].

3.12 Dose Estimation — the MIRD Method

The MIRD method (Medical Internal Radiation Dose) has been developed for about 25 years [37], [38]. Aim of this method is to find out a dose to an organ if the activity of this and/or other organs is known.

If we denote target items by subscript t and source items by s , then a mean dose \bar{D}_t in a target organ from one or several source organs is (A_{apl} means applied activity)

$$\bar{D}_t = A_{apl} \sum_s \tau_s \mathcal{S}_{s \rightarrow t}.$$

Quantity τ_s is called *residence time* of activity in source organ and it is defined as

$$\tau_s = \frac{\int A_s(t) dt}{A_{apl}},$$

where Δt is a time interval within which the dose is evaluated and $A_s(t)$ is the instantaneous activity of the source organ in time t .

Quantity $\mathcal{S}_{s \rightarrow t}$ is called *S-factor* and reflects the radiation influence of the target organ by the given source organ and the radionuclide of use. Sets of *S-factors* are published for various radionuclides, combinations of source/target organs and body geometries (male/female, age etc.). The improvement of *S-factors* measurement methodologies and techniques still proceeds [35]. However, only point estimates of *S-factors* are available.

For example, value of \mathcal{S} for thyroid gland as both source and target organ for an adult male is 5,64 mGy/(MBq hour), value of \mathcal{S} for thyroid gland as source organ and bone marrow as a target for an adult male is $8,5 \cdot 10^{-3}$ mGy/(MBq hour). These *S-factors* are designed for the standard thyroid gland mass 20,7 g. To express a dose for a thyroid gland of mass H , the equation (3.12) must be multiplied by the fraction 20,7 g/ H [g].

Chapter 4

Tools Used for the Solutions

4.1 Particular Bayesian Tools

Choice of Bayesian methodics is determined especially by the following reasons:

- it combines in a consistent way all accessible information sources, specifically theoretical description of the given system, measured data and expert knowledge and experience,
- it does not rely on an asymptotic behaviour of estimates,
- it yields the information about uncertainty of the estimate.

The following part is only introductory and describes the basic terms, statements and relations used as a tool. Further information can be found e.g. in [29].

Consider a system characterized by a parameter ϑ . Let D be a symbol for measurable quantities of the system. Let f be either probability density function (*pdf* — in continuous case) or probability (in discrete case), according to the given situation. Argument of this function will indicate the random quantity in question. The variable, random quantity and its value will have the same notation, as usual.

4.1.1 Bayes Rule

Let $f(D, \vartheta)$ be a joint *pdf* of the data D and the parameter ϑ . Let $f(D|\vartheta)$ be the *pdf* of the data D *conditioned* by the parameter ϑ . Then

$$f(D, \vartheta) = f(D|\vartheta)f(\vartheta), \quad (4.1)$$

where $f(\vartheta)$ is a marginal *pdf* of the parameter ϑ . The marginal *pdf* $f(D)$ of the joint *pdf* $f(D, \vartheta)$ can be obtained by

$$f(D) = \int f(D, \vartheta) d\vartheta. \quad (4.2)$$

Combining these two relations, we obtain the *Bayes rule*

$$f(\vartheta|D) = \frac{f(D|\vartheta)f(\vartheta)}{\int f(D|\vartheta)f(\vartheta) d\vartheta} \propto f(D|\vartheta)f(\vartheta), \quad (4.3)$$

where the symbol \propto denotes a proportionality by a term independent of ϑ .

The function $f(D|\vartheta)$ is called the *parametrized model* of the system, $f(\vartheta)$ is the *prior pdf* of the parameter ϑ and the function $f(\vartheta|D)$ is the *posterior pdf* of the parameter ϑ . The posterior *pdf* is the *generalized Bayesian estimate* of the parameter ϑ conditioned on the data and using all the accessible information, including the expert knowledge — prior information.

4.1.2 Chain Rule

If we deal with a sequence of n (≥ 1) data $D \equiv D(..n) \equiv (D_1, \dots, D_n)$, the equation (4.1), page 33, can be extended as the *chain rule*

$$f(D(..n), \vartheta) = \prod_{i=1}^n f(D_i | D(..i-1), \vartheta) f(\vartheta), \quad (4.4)$$

where $f(\cdot | D(..0), \vartheta) \equiv f(\cdot | \vartheta)$ represents conditioning by the parameter and prior information only.

4.1.3 Likelihood Function

The chain rule (4.4) can be written as

$$f(D(..n), \vartheta) = \mathcal{L}(D; \vartheta) f(\vartheta),$$

where the function

$$\mathcal{L}(D; \vartheta) = \prod_{i=1}^n f(D_i | D(..i-1), \vartheta)$$

is called *likelihood function*. In contrast to the product of *pdfs*, where D means a random variable and ϑ is given, the likelihood function is considered as a function of the parameter ϑ while the data D have the given (e.g. measured) specific values and they are fixed. The Bayes rule then has a form

$$f(\vartheta | D) = \frac{\mathcal{L}(D; \vartheta) f(\vartheta)}{\int \mathcal{L}(D; \vartheta) f(\vartheta) d\vartheta} \propto \mathcal{L}(D; \vartheta) f(\vartheta). \quad (4.5)$$

The likelihood function can be multiplied by any non-zero coefficient independent of ϑ without influencing the posterior *pdf*.

4.1.4 Conjugated Systems of Prior *pdfs*

Let P be a set of prior *pdfs* $f(\vartheta)$ and let $f(D|\vartheta)$ be a parametrized model. Assume that

$$0 < \int \prod_{i=1}^n f(D_i | \vartheta) f(\vartheta) d\vartheta < +\infty.$$

If the posterior *pdf* $f(\vartheta | D(..n))$, obtained from the Bayes rule (4.3), belongs to the set P (i.e. it is of the same kind as the prior *pdf*), then the elements in P are called *conjugated pdfs* with respect to the parametrized model $f(D|\vartheta)$.

In this case, the prior *pdf* is also called *self-reproducing pdf*. For practical purposes the system P is chosen narrow to simplify the estimation.

Examples of some conjugated prior *pdfs*:

| prior | parametrized model |
|-------------------|-----------------------------------|
| <i>gamma</i> | <i>Poisson</i> |
| <i>normal</i> | <i>normal, known variance</i> |
| <i>log-normal</i> | <i>log-normal, known variance</i> |

More can be found in [2].

4.1.5 Fictitious Data

In a conjugated system of prior *pdfs*, the posterior *pdf* has the same form as the prior one. The difference is that the posterior *pdf* depends on the data which “adjust” its parameters, in contrast to the prior *pdf* that contains no data (yet). We would like to adjust the parameters of prior *pdf* of the conjugated system so that it contains the prior information on ϑ .

One of the ways is to introduce so called *fictitious data* D_f . These “data” are not really measured, but potentially observable. The prior *pdf* $f(\vartheta)$ is then conditioned by those fictitious data, i.e. $f(\vartheta) \equiv f(\vartheta|D_f)$, as the posterior *pdf* is conditioned by the real data.

The Bayes rule can be applied for estimation of the prior *pdf* using the fictitious data. Formally in the same way as (4.5) we can get

$$f(\vartheta) \equiv f(\vartheta|D_f) \propto \mathcal{L}(D_f; \vartheta) f_u(\vartheta) \propto \mathcal{L}(D_f; \vartheta),$$

where $f_u(\vartheta)$ is a flat (often uniform) *pdf*. Hence we can say that in conjugated systems, the prior *pdf* is proportional to the likelihood of the fictitious data.

As fictitious, some data of an imaginative, similar or previously performed experiment can be used.

Of course, if we are able to determine the prior *pdf* parameters more or less directly without deriving them from fictitious data, the mentioned procedure can be skipped. But the term *fictitious data* can be still used.

4.1.6 Sufficient Statistics and Conjugated *pdfs*

Statistics $S(D)$ is called a measurable vector function $S \equiv (S_1, \dots, S_m)$ of data $D \equiv (D_1, \dots, D_n)$.

Finite statistics is such a function S , dimension m of which is finite and does not increase with increasing n . For example, arithmetic mean $S = \frac{1}{n} \sum_{i=1}^n D_i$ is the finite statistics, whereas the mapping $S = D$ is not.

Sufficient statistics: Let $g(s, \vartheta)$ and $h(D)$ be non-negative functions. If a joint *pdf* $f(D, \vartheta)$ can be written as

$$f(D, \vartheta) = h(D) g(S(D), \vartheta), \quad (4.6)$$

the statistics S is called *sufficient*. Substituting (4.6) into (4.3), we get formally

$$f(\vartheta|D) \propto f(D|\vartheta) f(\vartheta) = \frac{f(D, \vartheta)}{f(\vartheta)} f(\vartheta) = h(D) g(S(D), \vartheta) \propto g(S(D), \vartheta).$$

All the information necessary for estimation of ϑ , contained in the data, is carried by the sufficient statistics.

It can be shown that if finite sufficient statistics exists, then adequately narrow system of conjugated *pdfs* exists, too [2]. The prior *pdf* then has a form $f(\vartheta) \propto g(S(D_f), \vartheta)$, where D_f are the fictitious data (see part 4.1.5). The posterior *pdf* is obtained by the change of the statistics $f(\vartheta|D(..n)) \propto g(S(D_f, D(..n)), \vartheta)$.

4.1.7 Symmetric Confidence Intervals

Confidence interval $\langle x_l, x_u \rangle$ is a form of interval estimate of an uncertain quantity. Let $f(x)$ be *pdf* of a real scalar quantity x . In case of *symmetric confidence interval*, its bounds x_l and x_u are defined as follows:

$$\int_{-\infty}^{x_l} f(x) dx = \int_{x_u}^{+\infty} f(x) dx = \frac{\alpha}{2}. \quad (4.7)$$

Hence

$$\int_{x_l}^{x_u} f(x) dx = 1 - \alpha.$$

The interval $\langle x_l, x_u \rangle$ then determines a range of the unknown value with the probability $1 - \alpha$. If we introduce

$$\hat{x} = \frac{x_u + x_l}{2}, \quad \delta x = \frac{x_u - x_l}{2},$$

the interval estimate of x can be written as

$$x = \hat{x} \pm \delta x, \quad (4.8)$$

that is appealing even for non-statisticians.

The value of α is usually chosen as 0,1 or 0,05 etc.

4.2 Properties of the Used Probability Density Functions

In several tasks, the logarithmic-normal *pdf* and Poisson probability distribution are used. This section describes some general properties of these functions from a point of view of the Bayesian estimation.

4.2.1 Logarithmic-normal *pdf*

This *pdf* represents a random quantity, logarithm of which has normal *pdf* $\mathcal{N}_z(\mu_z, r)$. If new variable x and parameter μ are introduced, where $z = \ln x$ and $\mu_z = \ln \mu$, then x has **logarithmic-normal *pdf***

$$f(x|\vartheta) = L_x(\mu, r) \equiv \frac{1}{x\sqrt{2\pi r}} \cdot e^{-\left(\ln \frac{x}{\mu}\right)^2/(2r)} \quad (4.9)$$

with **parameter** $\vartheta = (\mu, r)$.

Moments of this *pdf* are defined as follows:

$$\begin{aligned} \mathbf{E}x &\equiv \int_0^{+\infty} x L_x(\mu, r) dx = \mu e^{r/2} \\ \mathbf{var} x &\equiv \mathbf{E}[(x - \mathbf{E}x)^2] = \mu^2 e^r (e^r - 1). \end{aligned}$$

The ratio of standard deviation and mean value is independent of μ

$$\omega = \frac{\sqrt{\mathbf{var} x}}{\mathbf{E}x} = \sqrt{e^r - 1}. \quad (4.10)$$

Let us assume:

1. the data $x_i, i = 1, \dots, n$ are positive and independent,
2. the *pdf* of $x_i, i = 1, \dots, n$ is $L_x(\mu, r)$,
3. r is known,
4. μ is to be estimated.

Likelihood as a function of parameter μ with known r and data x_i is

$$\mathcal{L}(D; \mu) = \exp\left(-\frac{1}{2r} [\nu(n)(\ln \mu)^2 - 2 \ln \mu \ln \tilde{\mu}(n)]\right)$$

(term $\prod_{i=1}^n 1/x_i$ is omitted as it is independent of μ). The **sufficient statistics** is

$$\begin{aligned} \tilde{\mu}(n) &= \tilde{\mu}(n-1) x_n, & \tilde{\mu}(0) &= 1, \\ \nu(n) &= \nu(n-1) + 1, & \nu(0) &= 0. \end{aligned}$$

As the likelihood is proportional to the logarithmic-normal *pdf*, if the term $1/\mu$ is omitted, the **self-reproducing prior *pdf*** is chosen as

$$f(\mu) \propto \frac{1}{\mu} \exp\left(-\frac{1}{2r} [\nu_0(\ln \mu)^2 - 2 \ln \mu \ln \tilde{\mu}_0]\right) \propto L_\mu\left(\frac{\sqrt{\nu_0} \tilde{\mu}_0}{\nu_0}, \frac{r}{\nu_0}\right),$$

where initial values $\tilde{\mu}_0, \nu_0$ corresponding to fictitious data are positive.

Posterior pdf is $L_\mu(\hat{\mu}(n), r/\nu(n))$, where

$$\hat{\mu}(n) = \sqrt[\nu(n)]{\tilde{\mu}(n)}$$

is a geometric mean of the data.

Moments of the posterior pdf are

$$\mathbb{E}[\mu|x(..n), r] = \tilde{\mu}(n) e^{r/(2\nu(n))}, \quad \text{var} [\mu|x(..n), r] = \tilde{\mu}(n)^2 e^{r/\nu(n)} (e^{r/\nu(n)} - 1). \quad (4.11)$$

Initial conditions for the statistics $\tilde{\mu}(0) = \tilde{\mu}_0 > 0$ and $\nu(0) = \nu_0 > 0$ guarantee properness of the chosen prior *pdf*.

4.2.2 Poisson Probability Function and Gamma-*pdf*

Poisson probability function is a parametrized model with which the gamma-*pdf* is conjugated.

Poisson probability function. Poisson probability function is a distribution of a nonnegative integer random variable x with a univariate parameter $\vartheta > 0$. The form of the Poisson function is

$$f(x|\vartheta) = \mathbb{P}_\vartheta(x) \equiv \frac{\vartheta^x}{x!} e^{-\vartheta} = \frac{\vartheta^x}{\Gamma(x+1)} e^{-\vartheta}, \quad x \in \{0, 1, 2, \dots\}.$$

The used Euler gamma-function is

$$\Gamma(x) = \int_0^{+\infty} t^{x-1} e^{-t} dt, \quad \Gamma(x+1) = x \Gamma(x), \quad x > 0. \quad (4.12)$$

Moments: $\mathbb{E}x = \text{var } x = \vartheta$.

The data $\{x_i\}_{i=1}^n$ are integer, positive and independent.

Likelihood has the form

$$\mathcal{L}(D; \vartheta) = \vartheta^{\tilde{x}_n} e^{-\vartheta \nu(n)},$$

where the **sufficient statistics** are

$$\begin{aligned} \tilde{x}(n) &= \tilde{x}(n-1) + x_n, & \tilde{x}(0) &= 0, \\ \nu(n) &= \nu(n-1) + 1, & \nu(0) &= 0. \end{aligned}$$

Self-reproducing prior pdf:

$$f(\vartheta) = \frac{\nu_0^{\tilde{x}_0+1}}{\Gamma(\tilde{x}_0+1)} \vartheta^{\tilde{x}_0} e^{-\vartheta \nu_0} \equiv G_\vartheta(\tilde{x}_0, \nu_0),$$

where

$$\tilde{x}_0 \geq 0, \quad \nu_0 > 0,$$

is called *gamma-pdf* $G_\vartheta(\tilde{x}_0, \nu_0)$.

Posterior pdf is then $G_\vartheta(\tilde{x}(n), \nu(n))$ and its **moments** are

$$\mathbb{E}[\vartheta|x(..n)] = \frac{\tilde{x}(n)+1}{\nu(n)}, \quad \text{var} [\vartheta|x(..n)] = \frac{\tilde{x}(n)+1}{\nu(n)^2} = \frac{\mathbb{E}[\vartheta|x(..n)]}{\nu(n)}. \quad (4.13)$$

4.3 General Numerical Algorithms

One of the drawbacks of the Bayesian computations is numerical difficulty in practical applications. Therefore various numerical techniques were used, focused on relevance, numerical stability and reliability. Majority of the methods mentioned here concern univariate real functions.

Some general or special methods are mentioned here. These methods are usually simple. In the final applications, they are mostly combined together or modified as required.

4.3.1 Extreme and Root Searching

The procedures discussed here are used often for searching both the root and extreme, depending on accessibility of the *pdf*'s derivatives (see part 5.2.1, page 50). In case of univariate *pdf*s equipped with first derivative, the extreme searching is in some conditions equivalent to searching the root of first derivative.

Bracketting.

Before starting to search for the extreme (or root respectively) x_0 of univariate function, we should know the approximate position of this point. The reason is that the methods used here need an “initial guess” x_{init} where the searching is started. For some methods, x_{init} must be “close enough” to x_0 , for other there must be known some interval in which x_0 can be surely found, otherwise the method need not converge.

Bracketting is a procedure that yields some x_{low} and x_{high} for which is valid $x_{low} < x_0 < x_{high}$. Values of x_{low} and x_{high} can but need not fulfill some criterion. Usually $x_{init} \in \langle x_{low}, x_{high} \rangle$.

The bracketting is used in these cases:

- Root searching when no derivatives are available. Two “guess” points close to the supposed root are supplied. These points are adjusted so that the function values in these points have the opposite signs.
- Root searching when first derivative is available. Two points are found as described in the previous item. Then these points are trimmed so that the opposite signs of the function values remain and relative difference of the first derivatives in these points is below requested limit.
- Extreme searching of univariate function when no derivative is available. Two “guess” points close to the supposed minimum (resp. maximum) are supplied. These points are adjusted so that the third one between them gives the minimum (resp. maximum) value. The parabolic extrapolation and the Golden section method are used ([30]).

Simplex method.

This method [26] is a general local extreme searching method, suitable for n -variate cases. It does not require derivatives and it is mostly less sensitive to the initial guess x_{init} than other methods mentioned below. The disadvantage is that for each step it requires at least two function evaluations (in the “worst” case, typically in the end of searching, up to n evaluations) and searching is less effective with increasing n .

For the searching a *simplex* is used, that is an n -dimensional object with $n + 1$ vertices (e.g. for two-variate function it is a triangle, for three-variate a tetraeder etc.). The simplex is placed around the point x_{init} . Following some strategy, the simplex “crawls” towards the nearest local extreme, i.e. changes its geometry and position. Its vertices are projected, expanded or contracted according to their function values. The search is terminated when differences between the vertices do not exceed some chosen small value.

It was shown that this algorithm in some situation performs convergence to a non-stationary point by repeated contraction [23]. However, there exist modifications of the simplex method [33], [40].

Convergence order of iteration methods.

Let $\{x_i\}_{i=1}^{+\infty}$ be a sequence of i -th result of an iteration method where $\lim_{i \rightarrow +\infty} x_i = \alpha$. Denote $\epsilon_i = x_i - \alpha$. Then *convergence order* of the iteration method in point α is called such minimum real $p \geq 1$ for which holds

$$\lim_{i \rightarrow +\infty} \frac{|x_{i+1} - \alpha|}{|x_i - \alpha|^p} = \lim_{i \rightarrow +\infty} \frac{|\epsilon_{i+1}|}{|\epsilon_i|^p} = C \neq 0.$$

C is called *asymptotic constant of error* [31].

The higher p is, the faster the method converges.

Newton-Raphson method.

This is a method for root searching in case of univariate functions. If x_i is the i -th iteration of the search, $f(x_i)$ the function value, $f'(x_i)$ its derivative and r is multiplicity of the root, then

$$x_{i+1} = x_i - r \frac{f(x_i)}{f'(x_i)}.$$

Convergence order of this method is 2. Nevertheless, the initial guess x_{init} must be bracketed closely enough, otherwise the method can even diverge. Furthermore, the first derivative is required.

For each step, one function and one derivative evaluation is necessary.

Bisection method.

This “classical” method is also used for cases if the univariate function changes its sign in the neighbourhood of the root. The initial guess is interval with limits of opposite signs. In each iteration, this interval is halved and that half with opposite limits is chosen as a new interval.

Convergence order of this method is 1.

This method converges slower than the Newton-Raphson method, but it does not require derivatives and close bracketing. For each step, one function evaluation is necessary.

4.3.2 1D-integration

The univariate integration methods used here are based on the *closed Newton-Cotes formulae* of n -th order. The integration interval of function $f(x)$ is divided to n subintervals of the same length separated by $n + 1$ equidistant points x_i , $i = 0, \dots, n$. On these points a Lagrange interpolation polynome $L_n(x)$ of degree n

$$L_n(x) = \sum_{i=0}^n f(x_i) l_i(x),$$

$$l_i(x) = \frac{\prod_{\substack{j=0 \\ j \neq i}}^n (x - x_j)}{\prod_{\substack{j=0 \\ j \neq i}}^n (x_i - x_j)} \quad (4.14)$$

is constructed and analytically integrated. Simple formulae combining the function values $f(x_i)$ are derived, including error term. The integrated function must be continuous up to k -th derivative and the formulae are exact for polynomes up to order k , where $k = (2 \lfloor \frac{n}{2} \rfloor + 1)$ [31].

If finer division of the integration interval is necessary, the integration interval is divided to several subintervals (in the sense mentioned above) of the same length and the formula is used for each such subinterval. The overall formula is called *composed Newton-Cotes formula*.

Simpson rule.

This method is the Newton-Cotes method of second order. In this work it is used generally when the size of the integration step is known or given. The formula and its error term estimation is mentioned here.

Let f be a function continuous up to its 3rd derivative, $h > 0$ be integration step. Denote

$$\begin{aligned} f_0 &= f(x_0), \\ f_1 &= f(x_0 + h), \\ f_2 &= f(x_0 + 2h). \end{aligned}$$

Then

$$\int_{x_0}^{x_0+2h} f(x)dx = \frac{h}{3}(f_0 + 4f_1 + f_2) + \epsilon, \quad (4.15)$$

where the error term

$$\epsilon = -\frac{h^5}{90}f^{(4)}(\xi), \quad x_0 \leq \xi \leq x_0 + 2h.$$

Composed Simpson rule for integration interval $\langle a, b \rangle$ with m subintervals, m even, $f_k = f(a + kh)$ and $h = \frac{b-a}{m}$ is

$$\int_a^b f(x)dx \approx \frac{h}{3}(f_0 + 4f_1 + 2f_2 + 4f_3 + \dots + 2f_{m-2} + 4f_{m-1} + f_m) \quad (4.16)$$

and the error term

$$\epsilon = -\frac{(b-a)h^4}{180}f^{(4)}(\xi), \quad a \leq \xi \leq b. \quad (4.17)$$

QUANCS.

This algorithm, originally written in FORTRAN, uses Newton-Cotes formula of 8-th order and gives the integral value with required maximum error, supplied either in absolute or relative value, which makes it very desirable. Size of the integration step is adaptively changed according to the function curvature. Furthermore, the result reliability and eventual problem area is reported [7].

As a disadvantage, it requires a function with continuity up to 9th derivative. Eventual points of discontinuity create the mentioned “problem areas”, they decrease the result precision and dramatically (by 1–3 orders) delay the program run. But despite that, the algorithm always gives *some* result with estimation of its precision.

Except of “problem” situation mentioned above, the actual precision of the result is usually far above the required one. It takes more function evaluations than for the required precision.

4.4 Programming and Computing Tools

The algorithms and the programs elaborated as results of this work are part of software called JodNew for use at the *KNM*. The software JodNew was originally developed by MS FoxPro 2.x for MS-DOS. Later a version of JodNew for MS-Windows9x/NT was produced using MS Visual FoxPro 6.

Therefore a question of data communication between JodNew and numerical programs must have been solved since the beginning. First version of JodNew requested standalone EXE-programs for MS-DOS that directly performed read/write operations on database files. Later, because of data safety and other reasons, data were transferred between JodNew and numerical programs

through text files with one value on the line, ordered in a defined way (as FoxPro cannot create and read general binary files).

Recent version of JodNew works under 32-bit MS-Windows. The numerical programs are still written for MS-DOS. As error reporting and exception handling mechanisms are not compatible with MS-Windows programming style, it will require a portion of programming work to put it in the right way.

The desired state of the numerical programs is a DLL-library with exported estimation functions. The following software tools were used:

Programming language: C/C++

Development software: Watcom C++ 10.5

Communication with database (obsolete): Sequiter CodeBase 4.5 library translated by Watcom C++.

Some data analyses: Statistical package SPSS 8.01 for Windows9x/NT.

Chapter 5

Solutions of the Numerical Tasks

In this chapter, some specific tailoring or combinations of some numerical methods are mentioned and then some programming approaches noticed.

5.1 Some Numerical Tasks

Some specific modifications and combinations of several methods for particular cases and functions are outlined here.

5.1.1 Confidence Interval

In the applications, 95%-confidence intervals ($\alpha=0,05$, see part 4.1.7, page 35) are computed.

If the integral of the *pdf* (i.e. distribution function), on which the interval is computed, is known, the task to find the interval is not difficult. But if the integral is not implemented or known, numerical integration must be used. The general algorithm described in this part (5.1.1) deals with *pdfs* where the analytical expression of the integral is unknown on it is not useful practically.

The idea of the algorithm has several parts.

Gaussian approximation

All the univariate *pdfs* treated in this work are unimodal. To formulate some general rules and criteria for numerical precision, let us assume that their approximation by a Gaussian *pdf* is valid.

Actually, this assumption of “Gaussian approximation” does not mean that another *pdf* is replaced by a normal *pdf*. It means that some rules for numerical integration (lengths of integration steps, points where to stop the integration etc.), designed for the normal *pdf*, are applied to another *pdf* which is then integrated according to these rules.

Obviously, approximation by a normal *pdf* is invalid for logarithmic-normal *pdf* in the estimation task of thyroid gland mass (see below). This case will be treated in a different way.

Let us denote $\mu \equiv \text{E}x$ and $\sigma^2 \equiv \text{var } x$. The Gaussian *pdf* with these parameters will be denoted $\mathcal{N}_x(\mu, \sigma^2)$.

Integration algorithm

If a closed form of *pdf* integral does not exist or if its computation is practically difficult and time consuming, the integration to determine the confidence interval must be done numerically. Here, the algorithm used for this purpose is described.

The main idea of the algorithm is: divide the Gaussian *pdf*'s domain to segments of length σ and apply composed Newton-Cotes formula to each segment. The number of integration steps on each segment is chosen according to the *pdf* course and required maximum integration error.

Integration step design. For the numerical integration, the composed Simpson rule (4.16), page 40, was chosen. The error term has the form (4.17).

Let the Gaussian *pdf* domain be divided into segments $s_{kl} = \langle \mu + k\sigma; \mu + l\sigma \rangle$ where k and l are integers. Let the variable be transformed to $\mathcal{N}(0, 1)$ so that $\mu = 0$ and $\sigma = 1$, then a segment $s_{kl} = \langle k; l \rangle$.

Absolute value of the error term (4.17) for integration of $f(x)$ on a segment $s_{k,l}$ with integration step h_{kl} is then

$$\epsilon = \frac{(l-k)h_{kl}^4}{180} f^{(4)}(\xi), \quad k \leq \xi \leq l. \quad (5.1)$$

If $M_{kl} = \max_{x \in \langle k; l \rangle} |f^{(4)}(x)|$ and number of integration steps $m_{kl} = (l-k)/h_{kl}$

$$\epsilon \leq M_{kl} \frac{(l-k)h_{kl}^4}{180} = M_{kl} \frac{(l-k)^5}{180 m_{kl}^4} \leq \epsilon_{max}, \quad (5.2)$$

where ϵ_{max} is some requested maximum allowed error of integration. Hence, taking into account that m_{kl} must be an even number,

$$m_{kl} \geq 2 \left\lceil \frac{l-k}{2} \left(M_{kl} \frac{l-k}{180 \epsilon_{max}} \right)^{1/4} \right\rceil \quad (5.3)$$

is a relation for minimum number of integration steps m_{kl} of Gaussian *pdf* on the interval $\langle k; l \rangle$. Let us now calculate the coefficients M_{kl} .

For the expression of $\mathcal{N}_x(0, 1)$ as $f(x) = \sqrt{\frac{1}{2\pi}} \exp\left(-\frac{x^2}{2}\right)$, the 4th derivative is

$$f^{(4)}(x) = f(x) (x^4 - 6x^2 + 3)$$

and the 5th derivative is

$$f^{(5)}(x) = -f(x) x (x^4 - 10x^2 + 15).$$

Roots of the 5th derivative are $x_1 = 0$, $x_2 = \sqrt{5 - \sqrt{10}}$ and $x_3 = \sqrt{5 + \sqrt{10}}$ (and other two with the opposite sign). In these points the 4th derivative reaches some extreme as it is an even polynome. Then M_{kl} for any k and l can be found.

If we introduce

$$g_{kl} = (l-k) \left(M_{kl} \frac{l-k}{180} \right)^{1/4}, \quad (5.4)$$

we can write (5.3) as

$$m_{kl} \geq 2 \left\lceil \frac{g_{kl}}{2} \left(\frac{1}{\epsilon_{max}} \right)^{1/4} \right\rceil. \quad (5.5)$$

The values for several g_{kl} are mentioned in the Table 5.1.

| k | l | g_{kl} |
|-----|-----|----------|
| 0 | 1 | 0,28555 |
| 1 | 2 | 0,25321 |
| 2 | 3 | 0,16674 |
| 3 | 4 | 0,16486 |
| 4 | 5 | 0,10492 |
| 5 | 10 | 0,33331 |

Table 5.1: Coefficients g_{kl} (see (5.5))

Several comments:

1. ϵ_{max} is both absolute and relative error for normalized functions. For non-normalized functions $g(x)$, where $f(x) = K \cdot g(x)$ and $f(x)$ is normalized, it is a relative error. In this case, the relation (5.3) must be multiplied by the normalizing constant K (mostly unknown), as $f^{(i)}(x) = K \cdot g^{(i)}(x)$.
2. The integration error estimate ϵ_{max} and hence the number of steps m_{kl} is pessimistic. One reason is inequality in (5.2), another one is adjusting m_{kl} up to an even value. This ensures that the result will be of better precision than requested.
3. Usual value for ϵ_{max} here is 10^{-5} , so that the total integral over all the segments, number of which is usually about 10, was computed with error less than 10^{-4} . m_{kl} on one unit-length segment takes value from 2 to 6 (mostly often 4).
4. The value $g_{5,10}$ is rather overestimated, as on a relatively long interval $\langle 5; 10 \rangle$ the maximum value of the 4th derivative is used. This maximum is reached in point $x = 5$. For $x > 5$ its value sinks exponentially. Therefore the divisions on the interval $\langle 5; 10 \rangle$ will be rather pessimistic.
5. This procedure is cheap as for runtime requirements.

Relation between confidence interval shift and domain limitation. If *pdf* has unbounded domain (like \mathcal{N}_x has), the numerical step-by step integration cannot proceed till infinity. Hence there is a question *where* to stop the integration process to find the area below the *pdf*, i.e. *how* to limit the domain of the *pdf* to meet some requirements. Error of the area determining due to premature integration stopping will be denoted as ϵ_a .

Assume that instead of $(-\infty; +\infty)$, the interval of integration will be limited to some $\langle x_{\epsilon_{al}}; x_{\epsilon_{au}} \rangle$. Let us assume that $f(x)$ is normalized. Let $\int_{-\infty}^{x_{\epsilon_{al}}} f(x)dx = \epsilon_{al}$ and $\int_{x_{\epsilon_{au}}}^{+\infty} f(x)dx = \epsilon_{au}$. Then $\epsilon_a = \epsilon_{al} + \epsilon_{au}$ and $\int_{x_{\epsilon_{al}}}^{x_{\epsilon_{au}}} f(x)dx = 1 - \epsilon_a$. Denote $F(x) = \int_{-\infty}^x f(t)dt$. Finally, if “true” α -confidence interval, obtained by considering the whole domain, is $\langle x_l; x_u \rangle$, denote “distorted” confidence interval, obtained by limiting the domain, as $\langle \tilde{x}_l; \tilde{x}_u \rangle$. It is obvious that $x_l \leq \tilde{x}_l$ and $\tilde{x}_u \leq x_u$.

The task is (i) what ϵ_{al} and ϵ_{au} to choose so that $|x_l - \tilde{x}_l|$ and $|x_u - \tilde{x}_u|$ are kept in some requested boundaries and (ii) how to correct such shifted points.

We start with lower bound. From the definitions mentioned above results

$$\int_{-\infty}^{x_l} f(x)dx \equiv F(x_l) = \frac{\alpha}{2} \quad (5.6)$$

and

$$\frac{\int_{x_{\epsilon_{al}}}^{\tilde{x}_l} f(x)dx}{1 - \epsilon_a} = \frac{F(\tilde{x}_l) - \epsilon_{al}}{1 - \epsilon_a} = \frac{\alpha}{2}. \quad (5.7)$$

As x_l is not known, let us expand $F(x_l) \approx F(\tilde{x}_l) + f(\tilde{x}_l)(x_l - \tilde{x}_l)$ assuming that the points x_l and \tilde{x}_l are close.

Substituting the expansion into (5.6), using (5.7) and considering $\epsilon_a = \epsilon_{al} + \epsilon_{au}$, we can express x_l as

$$x_l \approx \tilde{x}_l + \frac{1}{f(\tilde{x}_l)} \left[\epsilon_{au} \frac{\alpha}{2} - \epsilon_{al} \left(1 - \frac{\alpha}{2} \right) \right]. \quad (5.8)$$

This is the expression for the lower ($\alpha/2$) bound shift.

We can repeat these steps for the upper bound. Then similar relation will be found:

$$x_u \approx \tilde{x}_u + \frac{1}{f(\tilde{x}_u)} \left[\epsilon_{au} \left(1 - \frac{\alpha}{2} \right) - \epsilon_{al} \frac{\alpha}{2} \right]. \quad (5.9)$$

These equations can be used for partial correction of the shift if ϵ_{al} and ϵ_{au} are known.

If only non-normalized $g(x)$ is available where $f(x) = K \cdot g(x)$ and K is normalizing constant, the formulae (5.8) and (5.9) must be modified. Assume again the Gaussian approximation $f(x) = \mathcal{N}_x(\mu; \sigma)$. If $\mu = x_{max}$ is mode of $g(x)$ and σ its standard deviation, then $K = g(x_{max}) \sigma \sqrt{2\pi}$. The formulae are then

$$\begin{aligned} x_l &\approx \tilde{x}_l + \frac{K}{g(\tilde{x}_l)} \left[\epsilon_{au} \frac{\alpha}{2} - \epsilon_{al} \left(1 - \frac{\alpha}{2} \right) \right], \\ x_u &\approx \tilde{x}_u + \frac{K}{g(\tilde{x}_u)} \left[\epsilon_{au} \left(1 - \frac{\alpha}{2} \right) - \epsilon_{al} \frac{\alpha}{2} \right]. \end{aligned} \quad (5.10)$$

If we want to estimate ϵ_{al} and ϵ_{au} to keep some maximum allowed shift, the procedure will be the same except the expansion that will be $F(\tilde{x}_l) \approx F(x_l) + f(x_l)(\tilde{x}_l - x_l)$. Then we will substitute for $F(\tilde{x}_l)$ to get rid of \tilde{x}_l which is now unknown.

As $\mathcal{N}_x(0, 1)$ is symmetric, we will consider only one tail. Denote $\Delta x = |x_l - \tilde{x}_l| = |\tilde{x}_u - x_u|$, assume $f(\tilde{x}_l) = f(\tilde{x}_u)$ and $\epsilon_{al} = \epsilon_{au} = \epsilon_a/2$.

The formula (5.8) modified for x_l gives the condition

$$\epsilon_a \approx 2 \Delta x \frac{f(x_l)}{1 - \alpha} \quad (5.11)$$

and the same relation is valid also for the upper bound.

If we choose $\alpha=0,05$, then $1 - \alpha/2 = F(x_u)$, where $x_u=1,9599$ and $f(x_u)=0,058439$. Relation (5.11) then has a form

$$\epsilon_a \approx 0,123 \Delta x \approx \frac{\Delta x}{10}. \quad (5.12)$$

The shift can be decreased by using the correction formulae (5.10). If the correction is used, the shift limit Δx can be chosen higher (e.g. 10^{-3}). If the correction is not used, Δx must be chosen lower (e.g. 10^{-4}). According to chosen Δx , appropriate ϵ_a is found by (5.12). The value of ϵ_a will be needed later.

Half-width searching. The integration algorithm described above requires mean value and variance of the integrated *pdf*. If these characteristics are not available, they are substituted by maximum value (mode) and “half-width”.

If $x_{max} = \arg \max f(x)$ for *pdf* $f(x)$, then the *half-width* denoted as $l_{\frac{1}{2}}$ is here defined as

$$l_{\frac{1}{2}} = \max \left\{ |x_{max} - x|; f(x) = \frac{1}{2} f(x_{max}) \right\}. \quad (5.13)$$

The specific algorithm for the half-width searching is chosen according to availability of the *pdf*'s derivatives.

The integration procedure. Following the idea of Gaussian approximation, the normal *pdf* has an unbounded domain. Symbol $F(x)$ means integral of $f(x)$ (i.e. distribution function). The integration proceeds in these steps:

1. Find $u = Ex$ and $v = \sigma$. If they are not available, find $u = x_{max}$ and $v = l_{\frac{1}{2}}$.
2. Design numbers of integration steps on the *pdf*'s segments according to the maximum requested relative error ϵ_{max} .
3. Find K_1 by integrating from u up, save the pairs $(x, F(x))$ in some array.
4. Find K_2 by integrating from u down, save the pairs $(x, F(x))$ in another array.
5. Add $K = K_1 + K_2$, link both the arrays together and find the bounds of the interval by interpolation.

The question is where to stop the integration from the point u . This question was treated above and approximate relation for appropriate ϵ_a was derived (see (5.11)).

The error of the area determining due to omitting the domain above some chosen upper bound should not be greater than a required $\epsilon_a/2$ on each side. The Table 5.2 shows some points for the Gaussian $pdf \mathcal{N}_x(0, 1)$. For $\mathcal{N}_x(0, 1)$, if the pdf is integrated within the interval $\langle 0; x_{\epsilon_a} \rangle$, the error

| ϵ_a | x_{ϵ_a} | $F(x_{\epsilon_a})$ | $F'(x_{\epsilon_a}) = f(x_{\epsilon_a})$ |
|--------------|------------------|---------------------|--|
| 10^{-1} | 1,282 | 0,900079 | $1,754 \cdot 10^{-1}$ |
| 10^{-2} | 2,327 | 0,990017 | $2,661 \cdot 10^{-2}$ |
| 10^{-3} | 3,091 | 0,999003 | $3,359 \cdot 10^{-3}$ |
| 10^{-4} | 3,719 | 0,999900 | $3,959 \cdot 10^{-4}$ |
| 10^{-5} | 4,264 | 0,999990 | $4,496 \cdot 10^{-5}$ |
| 10^{-6} | 4,744 | 0,999999 | $5,175 \cdot 10^{-6}$ |

Table 5.2: Values of $\mathcal{N}_x(0, 1)$ with corresponding ϵ_a

of determining the area is $\epsilon_a/2$. Another form of this condition is approximately (see the Table 5.2)

$$f(x) \leq \frac{4\epsilon_a}{2}$$

for $\epsilon_a \in \langle 10^{-3}; 10^{-5} \rangle$, as $f(x_{\epsilon_a}) \approx 4\epsilon_a$.

For general non-normalized $K \cdot \mathcal{N}_x(\mu, \sigma^2)$, where K is the integral over all its domain, the condition for integration from μ to some x with $\epsilon_a \in \langle 10^{-3}; 10^{-5} \rangle$ takes the form

$$\sigma \frac{2f(x)}{K} \leq \frac{4\epsilon_a}{2} \quad \text{i.e.} \quad \sigma \frac{f(x)}{K} \leq \epsilon_a.$$

During the integration, this condition is tested, where $K/2$ is the result of the numerical integration. If the condition is fulfilled, the integration can be stopped at the point x and the relative error of the area due to the bounded domain is less than $\frac{\epsilon_a}{2}$.

This formulation of the condition has an advantage that it does not explicitly contain x_{ϵ_a} but ϵ_a .

Implementation note. Value Δx is passed to the constructor of class `SymConfInterv` (see part 5.2.1, page 50) as the parameter `deltaIntervRel`. Value of ϵ_a is computed in the body of the constructor using (5.12) and it is stored in a variable `epsTolRel`. Value of ϵ_{max} (see (5.5)) is computed in the body of the constructor using $\epsilon_a/100$ (see 7.1.3, page 74) and it is stored in a variable `epsIntRel`.

5.1.2 Simplex Termination Test and Convergence

In the simplex method [26], roughly described in section 4.3.1, page 38, the extremizing simplex is shifted towards the local extreme and shrunk around it. The important moment is decision of termination of the search. The original termination test looks like this:

Let us consider n -variate function $f(x)$. The simplex has $n+1$ vertices with n coordinates. Let us denote k -th coordinate of i -th vertex as x_k^i . Let ϵ be a given small number. Then the search is terminated if

$$\epsilon^2 \geq \frac{1}{n} \sum_{\substack{i=1 \\ i \neq j}}^{n+1} (f(x^i) - f(x^j))^2.$$

To adapt the test to a scale of an optimization task, the modification was done. If $f(x^j) = 0$, keep the same test. If $f(x^j) \neq 0$, this test is used:

$$\epsilon^2 \geq \frac{1}{n(f(x^j))^2} \sum_{\substack{i=1 \\ i \neq j}}^{n+1} (f(x^i) - f(x^j))^2.$$

The number ϵ then represents some relative bound of the values difference. Value of ϵ is different according to the specific task, gradient of a function etc., here usually from 10^{-4} to 10^{-10} .

It is known that this method need not always converge to a local extreme in some family of functions [23]. It was shown that in case of bi-variate function, the subsets of domain, where the function has only up to three derivatives, can be the area of convergence even though the points are not stationary. The convergence occurs by repeated contraction of other vertices of the simplex towards the best vertex (in sense of the extreme searching) which is kept fixed in non-stationary point. The sufficient condition for elimination of this effect is existence of 4th derivative or higher that is not identically zero.

The bi-variate functions used here in estimation task of effective half-life and radio-hygienic limits prediction have sufficient number of derivatives, therefore convergence in non-stationary point should not occur. Anyway, a change in the algorithm was made to prevent repeated contraction.

In the algorithm, if count of subsequent contractions reaches some number (e.g. 10) with the best (lowest or highest, according to the searching task) vertex unchanged, a new simplex is generated on the place of the old one and the search continues.

Another hint given in [23] is to replace contraction of one vertex by shrinking all the simplex towards the best vertex after several repeated contractions. This approach was not tried here.

5.1.3 Pseudo-normalization

The majority of functions that should here play a role of *pdf* of an estimated parameter is not analytically integrable, thus they are not normalized. Only a function that is proportional to the *pdf* is available.

Furthermore, these functions usually contain exponents within orders from ± 1 to $\pm 10^6$ and their evaluation would easily exceed the range of numbers that a computer is able to operate with, even using a double precision. To evaluate these functions without numerical problems and to keep their values in a numerically reasonable range, the *pseudo-normalization* process is applied.

Consider a *pdf* $f(x) = Kg(x)$, where $K > 0$ is an unknown normalizing constant and $g(x) > 0$ is not normalized. Assume the form $g(x) = \exp(\varphi(x))$, where $\varphi(x)$ is some continuous function.

As

$$\frac{d \ln g(x)}{dx} \equiv \frac{d\varphi(x)}{dx} = \frac{1}{g(x)} \frac{dg(x)}{dx},$$

then

$$\frac{d\varphi(x)}{dx} = 0 \quad \iff \quad \frac{dg(x)}{dx} = 0.$$

Hence both $g(x)$ and $\varphi(x)$ have their local extremes in the same points, as $\ln x$ is monotonous.

Let us find mode x_{max} where $\varphi(x)$ reaches its maximum (the function $g(x)$ is unimodal). As $g(x)$ can be multiplied by any term independent of x , we will multiply it by $\exp(-\varphi(x_{max}))$. In this way we obtain the function

$$\tilde{g}(x) = e^{\varphi(x) - \varphi(x_{max})}.$$

The values of the exponent are in interval $(-\infty; 0)$. If we choose some minimum exponent q_{min} that does not cause underflow in the operation $\exp(q_{min})$, we can compare $\varphi(x) - \varphi(x_{max})$ with q_{min} . If the subtraction is less than q_{min} , the exponential is not evaluated and zero is returned. Otherwise, the exponent is in the range $\langle q_{min}; 0 \rangle$ and the operation $\exp(\cdot)$ can be easily done.

The function $\tilde{g}(x)$ is then called *pseudo-normalized*.

This procedure can be used also for functions that do not have the explicit form $g(x) = \exp(\varphi(x))$ but contain products or exponents and cause problems with their range or evaluation. The form $g(x) = \exp(\ln g(x))$ can be adopted and $\ln(\cdot)$ can simplify the function evaluation. Then the same algorithm can be used.

5.1.4 Domain Restriction

The *pdfs* of parameters estimated here have one peak. The width of this peak (variance of the *pdf*) is greater or smaller, depending on the estimation task and uncertainty in the data. In some

cases, this peak can be narrow and *pdf* values on the majority of its domain are close to zero. This situation complicates numerical integration, rough investigation of the *pdf*'s shape or searching maximum. In some situations, the execution time was increased or even numerical failure was observed.

In the cases discussed here, data of low uncertainty provide narrow peaks. These data also provide a good deterministic estimate that is very close to the *pdf*'s maximum and can be successfully used as a starting point for the maximum search to find x_{max} .

Then the domain is narrowed to contain only values greater than some number. This number can be chosen following the ideas about area error in the section about confidence interval (5.1.1). In practice, the boundary value was chosen as $f(x_{max}) \cdot 10^{-7}$.

5.1.5 Design of the Integration Table

The procedure of confidence interval determining (see part 5.1.1) cannot be used for logarithmic-normal *pdf*, i.e. Gaussian approximation cannot be applied for logarithmic-normal *pdf*. During some numerical experiments that tested validity of integration step design mentioned above, it was found out that Gauss-based procedure failed on logarithmic-normal *pdf* unexpected in the neighbourhood of the *pdf*'s maximum. The precision of the integration in this area was by two orders worse than requested. Therefore it was decided that logarithmic-normal *pdf* will be treated in a different way.

The *pdf* has two parameters μ and r (see (4.9), page 36) and transformation between random quantity x with parameters μ_x, r_x and quantity y with parameters μ_y, r_y can be easily done by the formula

$$\left(\frac{x}{\mu_x}\right)^{\sqrt{1/r_x}} = \left(\frac{y}{\mu_y}\right)^{\sqrt{1/r_y}}. \quad (5.14)$$

The table of integrals of $f(x)$ from 0 to x_i (table of distribution function $F(x)$) was designed for $\mu = r = 1$.

The idea was: construct the table with non-constant distance of adjacent points so that error of quadratic interpolation is less than some chosen value.

Interpolation of function $f(x)$ by Lagrange polynome of 2^{nd} order $L_2(x)$ (see (4.14), page 39) given by points $x_0 < x_1 < x_2$ is

$$L_2(x) = \frac{(x-x_1)(x-x_2)}{(x_0-x_1)(x_0-x_2)}f(x_0) + \frac{(x-x_0)(x-x_2)}{(x_1-x_0)(x_1-x_2)}f(x_1) + \frac{(x-x_0)(x-x_1)}{(x_2-x_0)(x_2-x_1)}f(x_2). \quad (5.15)$$

Error term for this interpolation is

$$\epsilon(x) = (x-x_0)(x-x_1)(x-x_2) \frac{f'''(\xi(x))}{6}, \quad (5.16)$$

where ξ is some number in $\langle x_0; x_2 \rangle$ dependent on x . Introducing usual "pessimistic" ϵ_{max} , we can write (5.16) as

$$\epsilon(x) \leq \max_{x \in \langle x_0; x_2 \rangle} |(x-x_0)(x-x_1)(x-x_2)| \cdot \frac{1}{6} \max_{x \in \langle x_0; x_2 \rangle} |f'''(\xi(x))| \leq \epsilon_{max}. \quad (5.17)$$

Let us denote $p(x) = |(x-x_0)(x-x_1)(x-x_2)|$ and assume equidistant points $x_0 = 0, x_1 = h$ and $x_2 = 2h$. Then points x_{max} , where $p(x)$ reaches its maximum, and $p(x_{max})$ are

$$x_{max1,2} = h \left(1 \pm \frac{\sqrt{3}}{3} \right),$$

$$p(x_{max}) = h^3 \frac{2\sqrt{3}}{9}.$$

The 3^{rd} derivative of logarithmic-normal *pdf* $f(x)$ ((4.9), page 36) is

$$f'''(x) = \frac{f(x)}{x^3} \left[\frac{3}{r} \left(2 + \frac{1}{r} \ln \frac{x}{\mu} \right) - \left(1 + \frac{1}{r} \ln \frac{x}{\mu} \right) \left(2 + \frac{1}{r} \ln \frac{x}{\mu} \right) \left(3 + \frac{1}{r} \ln \frac{x}{\mu} \right) \right]. \quad (5.18)$$

Let us denote $\Upsilon(x) = \max_{x \in \langle x_0; x_2 \rangle} |f'''(\xi(x))|$ and $\Upsilon = \max(|f'''(x_0)|, |f'''(x_1)|, |f'''(x_2)|)$. Let us assume that h is small, so $\Upsilon(x) \approx \Upsilon$. Considering this and rearranging (5.17), we will get formula for table step h

$$h \leq \left(9\sqrt{3} \frac{\epsilon_{max}}{\Upsilon}\right)^{1/3}. \quad (5.19)$$

The algorithm of the table construction is outlined below:

1. $x_0 = x_1 = x_2 = 0$
2. repeat while $0 \leq x \leq 70$
3. $x_0 = x_2$, $\Upsilon = f'''(x_0)$, compute h using (5.19)
4. $x_1 = x_0 + h$, $x_2 = x_0 + 2h$
5. $\Upsilon = \max(|f'''(x_0)|, |f'''(x_1)|, |f'''(x_2)|)$ using (5.18), compute ϵ using (5.17)
6. if $\epsilon > \epsilon_{max}$, decrease h (e.g. multiply by $\frac{2}{3}$) and go to 4
7. compute integrals from 0 to x_1 ($F(x_1)$) and from 0 to x_2 ($F(x_2)$), save the pairs $(x_1, F(x_1))$ and $(x_2, F(x_2))$ and go to 2.

If we want to find $F(x')$ with general parameters μ, r using this table, we must first transform the variable x' to the variable x with the parameters $\mu = r = 1$ according to (5.14). Then we find in the table some adjacent points $x_0 < x_1 < x_2$, where $x \in \langle x_0; x_2 \rangle$. Then we use formula (5.15) for interpolation.

The table for quadratic interpolation for x from 0 to 70 has about 80 pairs of $(x, F(x))$ for requested $\epsilon_{max} = 10^{-4}$. If the table is constructed for linear interpolation with $L_1(x)$, it has about twice as many pairs.

5.2 Programming Approach

5.2.1 Numerical Objects

Theoretical analysis of the problems and derivation of Bayesian estimates of the desired quantities provides relatively complicated formulae. Theory gives origin to various functions with different properties: some are analytically integrable or differentiable, some not (at least practically), some are easily computable, some not, some can easily yield mean value, variance, maximum value etc., some not. Some estimates are expressed by *pdfs* (the integral over their variable(s) is equal one), in some cases only functions proportional to the *pdf* are easily computable.

Development of numerical programs

Testing numerical properties of the formulae and methods to use can be done in several ways:

- Using some numerically oriented software, e.g. Matlab. Such a software provides a lot of useful tools and enables a comfortable work with them. One can focus on numerical tasks and algorithms without paying much attention to questions of programming. Unfortunately, without special accessories the results obtained by this software are not portable to other platform. In the beginning of this project, the target platform of the application JodNew was 16-bit DOS. The tools that are available now (Matlab Compiler, Matlab run-time libraries or ABET [25]) do not support this platform and were not available when this project started.
- Using some standard programming and application building tools like C or C++. Using this language, standard mathematical functions (or some basic mathematical classes) are available. Target application can be easily built practically for any platform, but the development process and numerical experiments can be very time and energy consuming. Programming also requires some “culture” to avoid mistakes that can be easily overlooked, especially in case of complex task. Both programming and numerics are tightly bound together.

- Using specific class support for object-oriented language to represent specific objects (function, algorithms...) appearing in the work. Such a class library was not available when the project started.

In the beginning, C++ development environment was available. A class library was developed that yielded some comfort with numerical work and made possible to create the application on the given platform. Object-oriented methodology enables to separate numerical tasks and programming, it solves technical questions of compatibility, specific behaviour of functions or algorithms etc.

The main idea

The main idea follows one of the object property: uniformity due to the inheritance. Each class is designed so that its object has uniform interface in communication with other objects despite its internal structure and means it uses. All the variables and/or data structures necessary for the object are encapsulated inside and their construction and destruction is done in the right way. Any name or memory conflicts between two objects are avoided. Important object's variables are protected by denied access from outside. The validity of parameters used for the object construction is tested.

Object-oriented design separates the work into two steps: (i) implement the function or algorithm (design the class), (ii) create the object and use it.

Although object-oriented program performs lower speed and higher memory requirements compared to "classical structure-oriented" program, the computational power available now makes the difference negligible for the class of tasks solved here. The advantages it brings override the loss of milliseconds and kilobytes due to using this technique.

Classes

The most important classes are shown on the scheme on Fig. 5.1. The arrow means the inherited

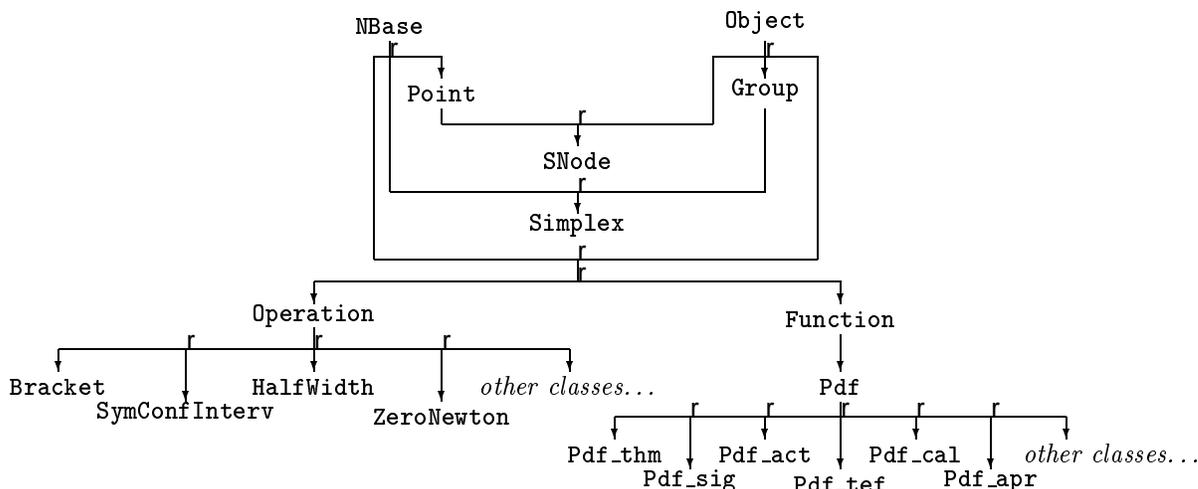


Figure 5.1: Classes hierarchy

class, the inheritance flow is oriented from upper classes to lower classes.

The class **NBase** is the base class for all classes dealing with numerics. It collects mechanism of exception handling, in practice mechanism of error handling (e.g. invalid value of parameter, unsatisfactory result of the solution etc.).

The class **Object** implements properties of ownership (object is "owned" by another object), locking/unlocking (object is "locked" by an object and cannot be modified by other objects) and event handling mechanism, that was not used here.

The class **Group** is **Object** that can contain (own) other **Objects**.

The class **Point** represents n -dimensional point with some vector operations (addition, multiplication by a scalar), equality and assignment operators etc.

The class **SNode** is a node (vertex) owned by **Simplex**. It has some extended properties compared to **Point**, e.g. it contains also function value etc.

The class **Simplex** can find local extreme of n -dimensional **Function** using the simplex method. It is a container of vertices (**SNode**).

The class **Operation** is just the successor of **NBase** and **Object**.

The classes **Bracket**, **SymConfInterv** etc. represent implementation of various numerical tasks (bracket the root/extreme, find a symmetric confidence interval for 1-dimensional **Pdf** etc). They choose the procedure according to the “equipment” (available derivatives etc.) of the **Function** they are operating on.

The class **Function** covers the basic properties of n -dimensional functions. It carries information about its domain, contains the list of singular points and supplies its value if possible (e.g. singular point of the function $(\sin x)/x$ is $x = 0$, in this case the value 1 is returned), contains additional “equipment” (derivatives or integral if available, despite they are evaluated using formula, table, communication with other **Functions** or other ways) and gives information about it. It has a uniform interface.

The class **Pdf** represents *pdf*. The features added to **Function** are normalizing constant, mean value, variance and mode, if available.

The classes **Pdf_thm**, **Pdf_sig** etc. implement the Bayesian estimates described in the Chapter 6, page 55. Their properties will be described there.

5.2.2 Communication with Database

The estimates computed by the numerical programs developed here are stored in database files for further use. The data (often up to 40–50 numbers, their number is not fixed) are also stored in database files. The communication between the database system and numerical programs is necessary.

In the previous version of JodNew designed completely for MS-DOS, the numerical programs performed a direct read/write access to database files. As a command line parameter, a database file was sent that contained information from which file and from which record the given value should be read and to which file and record the results should be written. For this access, the library Sequiter CodeBase 4.5 was used.

This way of communication had some disadvantages. Firstly, before calling the numerical programs, all the files in the database had to be closed and the FoxPro environment suspended. Then the numerical program opened the files of need, found the data, computed the result, stored it in the file and closed all the files. After that, FoxPro environment had to be resumed and all the necessary files opened. This occurred delaying in practice. Also direct access to the database brought a risk of corruption its integrity.

The next version of JodNew was written for 32-bit MS-Windows, although the numerical programs were still in MS-DOS. The new version of MS-FoxPro “upgraded” the format of database files, so they were not legible anymore by the “old” library CodeBase. The intended form of the numerical programs was a DLL-library of functions which would take parameters and return values. As some serious programming modifications of the numerical programs should have been done, the intermediate stage of data communication through files was chosen.

Unfortunately, MS-Visual FoxPro cannot create or read a general binary file, only data file or text file. Thus the input data are read from data files, converted into strings and written to the text file. This file is read by a numerical program, strings converted to binaries and passed to the algorithms. Results are converted to strings, written to the file. This file is read by FoxPro, strings converted to binaries and stored in data files.

This procedure is faster, because only one file is opened and closed and no environment is suspended and resumed. Also database cannot be affected from outside. The greatest advantage is

that data can be easily checked by simple viewing the text file. The files can be also easily created which is important in the debugging phase.

The work on DLL-library is in progress, but it is not well accepted by the rest of the development team, as any change creates only errors and problems. In this situation, it seems that the communication through text files will be kept. One reason was already mentioned. Another reason is that MS-FoxPro does not yield programming tools to control data types of memory variables and therefore binary format of a number in memory is undefined for the user. Decision of variable type is made by FoxPro during the run time and type compatibility between data files and memory variables makes communication through binary representation risky and dangerous. Furthermore, structured data types created by FoxPro appear to be incompatible with those in C++.

Strings seem to be the only really safe way of transfer, but not absolutely. MS-Windows enable to define the decimal separator. If it is not a dot ('.') but comma (','), standard conversion C-functions omit anything behind it. Therefore the decimal separator must be set in FoxPro as dot.

5.2.3 Software Implementation

The numerical programs are copied in a separate directory. They are called from FoxPro by a batch file that passes them the name of the text file containing input data. After finishing the run, the batch file checks the status of ERRORLEVEL variable. In case of error, the data are saved. The text file with the data is appended to a specific file according to the ERRORLEVEL value. In this way, data causing problems are stored for further analysis and improvements of both the numerical algorithms and the system JodNew.

The work on the DLL-library is in progress.

Chapter 6

Solutions of the Estimation Tasks

In this chapter, solutions of the estimation tasks are described from theoretical point of view. There are two kinds of tasks: those that are solved in the Bayesian way and those that are not.

In the former case, the procedure of the Bayesian estimate [11] is described in details and remarks about practical realization are attached. The sections concerning this part of tasks have the following structure:

Measurement. In this part

- in general the method of the data measurement
- the quantity to be estimated

are briefly described.

Model. Here, the model of the investigated process is formulated. The quantity to be estimated is a part of the unknown parameter of this model. The form includes

- model describing dependence of data, the quantities to be estimated and usually other parameters,
- the parameter(s) of the model.

Likelihood. If possible, likelihood function and statistics created by the measured data are described.

Prior information. In this part, everything that is known before the measurement is mentioned. The form includes

- expert knowledge, assumptions etc.,
- the corresponding prior *pdf*,
- if sufficient statistics exist, their initial conditions.

Estimation. Here, the posterior *pdf* is derived and, if possible, the mean value and the variance are evaluated.

Implementation remarks. In this part potential pitfalls of the posterior *pdf* evaluation are noted and the methods and/or tricks used in the computations are mentioned.

In the latter case when Bayesian methodology is not applied for the solution, the structure of the sections is simpler and implementation notes are omitted.

6.1 Thyroid Gland Mass

Thyroid gland mass is the key quantity for estimation of the absorbed radiation dose (see sections 3.10, 3.11 and 3.12, page 31). Unfortunately, its measurement is loaded by the largest error of all the measurements mentioned in this thesis.

6.1.1 Measurement

Thyroid gland mass measurement at the *KNM* is performed by two methods:

Palpation (estimation by a touch). This method is very subjective and the assumed relative precision of this method is about 50% ($\omega_{\text{palp}} = 0,5$).

(Ultra-)Sonography. This method is more objective and the assumed relative precision is about 20% ($\omega_{\text{sono}} = 0,2$).

Furthermore both these methods strongly depend on an experience of the physician performing the examination. Due to the stable staff we can assume personal independence of the mentioned precisions.

6.1.2 Model

As the absolute measurement error of this quantity depends on the measured magnitude, the logarithmic-Gaussian model is chosen.

Model: chosen *pdf* is

$$f(h_i|H, r_i) = \frac{1}{h_i \sqrt{2\pi r_i}} \cdot e^{-\left(\ln \frac{h_i}{H}\right)^2 / (2r_i)} \equiv \mathbf{L}_{h_i}(H, r_i),$$

where h_i is the result of the i -th measurement and r_i depends on its assumed relative precision, as mentioned in (4.10), page 36. H is the constant estimated mass. The measured values are assumed conditionally independent.

The relative precision ω_i is interpreted as a ratio of standard deviation and mean value of the logarithmic-Gaussian *pdf* $\mathbf{L}_{h_i}(H, r_i)$. It is connected with r_i by the relation (according to (4.10), page 36)

$$\omega_i = \sqrt{e^{r_i} - 1} \iff r_i = \ln(1 + \omega_i^2) \quad (6.1)$$

and hence

$$\begin{aligned} r_{\text{palp},i} &\approx 0,2 \\ r_{\text{sono},i} &\approx 0,04. \end{aligned}$$

Parameter is the estimated mass H .

6.1.3 Likelihood

in case of $n(\geq 1)$ measurements (see part 4.2.1, page 36)

$$\begin{aligned} \mathcal{L}(h(..n); H, r) &\propto \exp\left(-\frac{\ln H}{2r} \left(\nu(n) \cdot \ln H - 2 \ln \tilde{h}(n)\right)\right) \propto \\ &\propto \mathbf{L}_H\left(\hat{h}(n), \frac{r}{\nu(n)}\right), \end{aligned}$$

where $\hat{h}(n) = \tilde{h}(n)^{1/\nu(n)}$ and r is an arbitrary value chosen for a suitable scaling (it is cancelled in the *pdf*).

Sufficient statistics are expressed as follows:

$$\begin{aligned} \tilde{h}(n) &= \tilde{h}(n-1) \cdot h_n^{r/r_n}, & \tilde{h}(0) &= 1 \\ \nu(n) &= \nu(n-1) + \frac{r}{r_n}, & \nu(0) &= 0. \end{aligned} \quad (6.2)$$

6.1.4 Prior Information

Expert knowledge is given by this experience:

- the value of H can vary in a broad range, approximately from 0,1g to 500g,
- relative error for the
 - palpation method $\omega_{palp,i} = 0,5$ ($r_{palp,i} = 0,2$),
 - sonographic method $\omega_{sono,i} = 0,2$ ($r_{sono,i} = 0,04$).

Prior pdf is chosen as self-reproducing, that is in the same form as (6.1.2) with non-trivial initial conditions for the statistics \hat{h}_0 and ν_0 (6.2).

Having analyzed the available data files, we found 1 022 records with palpation measurements and 1 657 records with sonographic measurements. As there are more records with sonographic measurements which have furthermore better accuracy than palpation ones, it was decided to use sonographic data to find fictitious data for prior the *pdf*.

The results of sonographic measurements varied from 0,1 g to 470 g. The arithmetic mean of the sonographic values was $\bar{h} = 5,4$ and sample standard deviation $\bar{s} = 29,05$. Considering (4.11), page 37, we find that

$$\hat{h}_0 = \bar{h} \exp\left(-\frac{r}{2\nu_0}\right), \quad \nu_0 = \frac{r}{\ln\left[\left(\frac{\bar{s}}{\bar{h}}\right)^2 + 1\right]}.$$

The resulting fictitious data are

$$\nu_0 = \frac{r}{3,395}, \quad \hat{h}_0 = 0,998, \quad \tilde{h}_0 = 0,998^{\nu_0}. \quad (6.3)$$

where r is an arbitrary value chosen for a suitable scaling (it will be reduced in the *pdf*). The prior *pdf* of an unknown mass H is then

$$f(H) = \mathcal{L}_H\left(\hat{h}_0 = 0,998, \frac{r}{\nu_0} = 3,395\right). \quad (6.4)$$

These fictitious data correspond to a fictitious measurement where $h_i = 5,4$ and $\omega_i = 5,38$. Similarity of these two numbers is accidental. The value of r is chosen as geometric mean $r^n = \prod_{i=1}^n r_i$ of values r_i corresponding to the data h_i .

6.1.5 Estimation

Substituting the data to the statistics (6.2) with the initial conditions (6.3) and denoting $\hat{h}(n) = \sqrt[n]{\tilde{h}(n)}$, the posterior *pdf* of the mass H is

$$f(H|h(..n)) = \mathcal{L}_H\left(\hat{h}(n), \frac{r}{\nu(n)}\right). \quad (6.5)$$

with the moments

$$\begin{aligned} \mathbb{E}[H|h(..n)] &= \hat{h}(n) \cdot \exp\frac{r}{2\nu(n)} \\ \text{var}[H|h(..n)] &= (\mathbb{E}[H|h(..n)])^2 \cdot \left(\exp\frac{r}{\nu(n)} - 1\right), \end{aligned}$$

where r is the chosen “scaling” value. The expression for mode is $\hat{x} = \hat{h}(n) \exp\left(-\frac{r}{\nu(n)}\right)$.

6.1.6 Implementation Remarks

The posterior *pdf* is represented by the class `Pdf_thm`. The data pairs h_i and ω_i are passed to the constructor in an array containing information how many measurements (data pairs) are available.

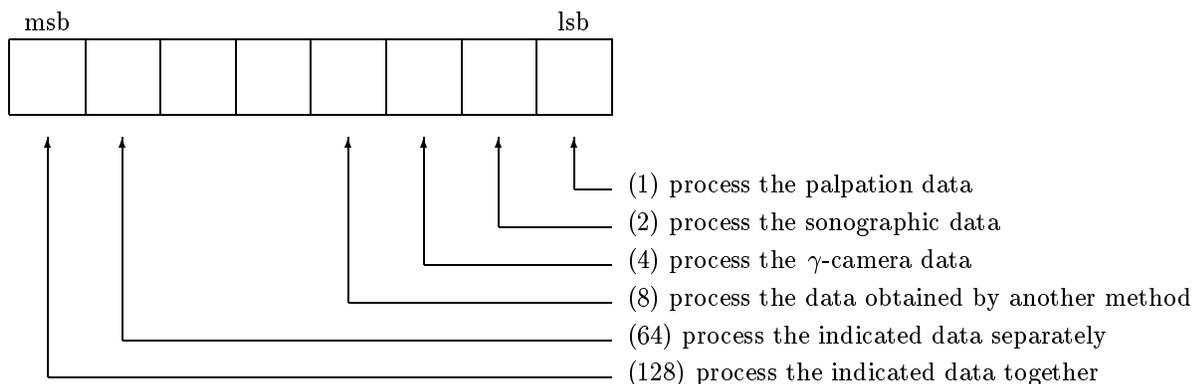
The value r is chosen as a geometric mean of all r_i (see (6.1)) corresponding to h_i .

If the exponent of (6.5) is less than logarithm of minimum value representable in double precision, zero is returned.

The posterior *pdf* formula (6.5) is numerically stable and reliable in a range far exceeding common values of thyroid gland mass. The function is normalized. The function value, mean value, dispersion, mode and integral are implemented. Because of term $1/H$, value in point 0 is defined as 0.

Because logarithmic-normal *pdf* can be hardly approximated by normal *pdf*, the confidence interval was computed in a different way than in case of other estimates. A table of integral values was generated with floating distance between the discretization points so that error of quadratic interpolation between these points does not exceed 10^{-4} . For details, see part 5.1.5, page 49.

In the *KNM* applications, estimates are requested separately using only the palpation or sonographic data and using all the data together. This request is indicated by a bit field carrying information about the data processing (see Fig. 6.1). However, this mechanism is not a part of



Example:

| <i>dec</i> | <i>bin</i> | <i>what to process</i> | <i>returned</i> |
|------------|------------|---|-----------------|
| 3 | 00000011 | nothing | nothing |
| 67 | 01000011 | palp. and sono. data separately | 2 estimates |
| 131 | 10000011 | palp. and sono. data together | 1 estimate |
| 195 | 11000011 | palp. and sono. data both separately and together | 3 estimates |

Figure 6.1: Bit field to control the estimation of thyroid gland mass

the class but of a function that manages the class construction.

6.2 Activity Kinetics

In thyroid gland diseases radiotherapy, one of the important tasks is to quantify activity kinetics in the patient's organism. This kinetics means particularly to estimate

- the *effective half-life*, in which one half of the activity is eliminated from the organ of interest (see part 3.2, page 28),
- the *prediction* of time, when some *radiohygienic limit* will be reached (see part 3.9, page 29).

The activity measurement is represented by a detection and counting the impulses (particles) caused by a radioactive decay. The amount of these impulses is the only directly measurable quantity. To determine the unknown activity, the corresponding impulses amount must be compared with the one obtained by a measurement performed on a source of known activity. Such a source is usually called *standard source*.

The following subsections (Nos. 6.3 – 6.7) are closely related together from the methodological and measurement point of view [11]. The first three tasks,

- estimation of the amount of particles coming from the source of interest (to exclude the influence of the background),
- calibration of the measurement device with a source of known activity (to convert particles amount to activity in given geometric conditions),
- estimation of the unknown activity in the given time.

can be understood as subtasks of one complex task to find out the activity kinetics. Using previous and so far usual “deterministic” methods of computation and estimation, these subtasks of activity kinetics estimation were necessary to be performed step by step. Using the Bayesian tools, the activity kinetics estimation problem is solved as whole at one time, not part by part, as outlined above. Nevertheless, the problems mentioned as “subtasks” can play an independent role, especially in these cases:

1. to examine, if simple subtraction of background impulses from impulses of source on background can yield suitable estimate of the source signal,
2. to check (test), if the geometrical and other conditions of measurement are constant in time by examining the calibration factor value,
3. to evaluate single-time activity measurement for other purposes (e.g. contamination of the staff),
4. to yield “step-by-step quantities” to the *KNM* staff which is used to check their values during the procedure of the measurement.

Therefore solutions of these steps have not only historical purpose but also direct application in auxiliary tasks.

6.3 Source on the Background

The only data directly measurable in the activity/kinetics estimation task are counts of ionizing particles detected above the places of interest in the body. These counts are random quantities “polluted” by a background radiation, also random. The measurement time must not be too long (tens of seconds), supplied activity during the diagnosis stage must be low enough to minimize the radiation risk and the measurements cannot be repeated because of the *KNM* work schedule and large amount of patients. All these reasons increase the uncertainty of the data.

Therefore the elimination of the background is a very important problem. If the source signal of interest is much higher than background, the task is very simple. If the background and the source signal are of the same level, the probabilistic approach must be used [16].

6.3.1 Measurement

Counts of particles (impulses in the measurement device) in some fixed time interval are detected. Assume that

- B is a mean value of the background particles count (without a source),
- Z is a mean value of the source particles count (without a background — so called *signal*),

- both B and Z are constant during the measurement.

Let us perform n measurements of background impulses, result of each is b_i , $i = 1, 2, \dots, n$. Denote $b(..n) = \{b_i\}_{i=1}^n$. Assume that mean value of b_i is B .

Let us perform m measurements of signal plus background impulses, result of each is z_j , $j = 1, 2, \dots, m$. Denote $z(..m) = \{z_j\}_{j=1}^m$. Assume that mean value of z_j is $B + Z$.

The mean value of source signal impulses Z is the quantity to be estimated.

6.3.2 Model

Let us assume the data are independent, Poisson-like distributed. Model of the background impulses measurement b_i with mean value B is $f(b_i|B) = P_B(b_i)$ (see part 4.2.2, page 37). Model of the source on background impulses measurement z_j with mean value $B + Z$ is $f(z_j|B, Z) = P_{B+Z}(z_j)$. Due to the independence, the joint probability function of data measurement is

$$f(b(..n), z(..m)|B, Z) = \prod_{i=1}^n P_B(b_i) \cdot \prod_{j=1}^m P_{B+Z}(z_j).$$

Z is the parameter to be estimated.

6.3.3 Likelihood

The form of likelihood is (see part 4.2.2, page 37)

$$\begin{aligned} \mathcal{L}(b(..n), z(..m); B, Z) &\propto B^{\tilde{b}(n)} e^{-[\nu(n)+\nu(m)]B} \cdot (B+Z)^{\tilde{z}(m)} e^{-\nu(m)Z}, \\ \tilde{b}(n) &= \tilde{b}(n-1) + b_n, & \tilde{b}(0) &= 0 \\ \tilde{z}(m) &= \tilde{z}(m-1) + z_m, & \tilde{z}(0) &= 0 \\ \nu(n) &= \nu(n-1) + 1, & \nu(0) &= 0. \end{aligned}$$

Terms $\tilde{b}(n)$, $\tilde{z}(m)$ and $\nu(n)$ are finite sufficient statistics.

6.3.4 Prior Information

The expert knowledge is expressed by choice of a suitable measurement range \bar{b} .

As the independence of background and source impulses is assumed, i.e.

$$f(B, Z) = f(B) \cdot f(Z),$$

prior information on background and source can be investigated independently.

As described in part 4.2.2, page 37, the prior *pdf* can be expressed by gamma-*pdf* $f(B) = G_B(\tilde{b}_0, \nu_0)$. Our requirements are (i) mean value of prior *pdf* EB is in the middle of the measurement range $\langle 0, \bar{b} \rangle$ and (ii) standard deviation $\sqrt{\text{var } B}$ of the prior *pdf* is one half of the measurement range length $\langle 0, \bar{b} \rangle$. Knowing that $EB = \frac{\tilde{b}_0 + 1}{\nu_0}$ and $\text{var } B = \frac{\tilde{b}_0 + 1}{\nu_0^2}$, we obtain

$$\tilde{b}_0 = 0, \quad \nu_0 = \frac{2}{\bar{b}} \equiv \zeta. \quad (6.6)$$

As for the source prior *pdf*, the procedure and the values for \tilde{z}_0 and ν_0 will be the same provided that the same range \bar{b} is assumed.

As the result, the prior *pdf* for the source and background is

$$\begin{aligned} f(B, Z) &\propto B^{\tilde{b}_0} (B+Z)^{\tilde{z}_0} \cdot e^{-\nu_0(2B+Z)}, \\ \tilde{b}_0 &= 0, \\ \tilde{z}_0 &= 0, \\ \nu_0 &= \frac{2}{\bar{b}} = \zeta. \end{aligned}$$

6.3.5 Estimation

According to the previous conclusions, the joint probability of both background and source impulses is

$$\begin{aligned} f(B, Z|b(..n), z(..m)) &\propto B^{\bar{b}(n)} (B + Z)^{\bar{z}(m)} \cdot e^{-\{\nu(n)+\nu(m)\}B+\nu(m)Z}, \\ \bar{b}(n) &= \bar{b}(n-1) + b_n, & \bar{b}(0) &= 0 \\ \bar{z}(m) &= \bar{z}(m-1) + z_m, & \bar{z}(0) &= 0 \\ \nu(n) &= \nu(n-1) + 1, & \nu(0) &= \zeta = \frac{2}{\bar{b}}. \end{aligned}$$

The marginal *pdf* of Z is obtained as

$$f(Z|b(..n), z(..m)) = \int_0^{+\infty} f(B, Z|b(..n), z(..m)) dB.$$

The explicit form of this formula is [6]

$$f(Z|b(..n), z(..m)) = \frac{e^{-\nu(m)Z} \sum_{j=0}^{\bar{z}(m)} \frac{\bar{z}(m)! (\bar{b}(n)+j)!}{(\bar{z}(m)-j)! j!} \frac{Z^{\bar{z}(m)-j}}{(\nu(n)+\nu(m))^{\bar{b}(n)+j+1}}}{\nu(m)^{\frac{1}{\bar{z}(m)+1}} \sum_{j=0}^{\bar{z}(m)} \frac{\bar{z}(m)! (\bar{b}(n)+j)!}{j!} \frac{\nu(m)^j}{(\nu(n)+\nu(m))^{\bar{b}(n)+j+1}}}. \quad (6.7)$$

The formula for k -th moment $\int_0^{+\infty} Z^k p(Z|b(..n), z(..m)) dZ$ has the form [6]

$$E(Z^k|b(..n), z(..m)) = \frac{\nu(m)^{\bar{b}(n)+1} \sum_{j=0}^{\bar{z}(m)} \frac{(\bar{b}(n)+j)! (\bar{z}(m)-j+k)!}{(\bar{z}(m)-j)! j!} \left(\frac{\nu(m)}{\nu(n)+\nu(m)}\right)^j}{\sum_{j=0}^{\bar{z}(m)} \frac{(\bar{b}(n)+j)!}{j!} \left(\frac{\nu(m)}{\nu(n)+\nu(m)}\right)^j}. \quad (6.8)$$

Both these formulae are numerically extremely untractable.

6.3.6 Implementation Remarks

The posterior *pdf* is represented by the class `Pdf_sig`.

Very smart and stable algorithms for the *pdf* and k -th moment evaluation were derived by [6].

Unfortunately, both the algorithms contain $\bar{z}(m)$ -times repeated cycle. This feature becomes disadvantageous for count numbers $\bar{z}(m)$ of order $10^5 - 10^6$ that are frequently used. Therefore an approximation of *pdf* (6.7) should be designed. Some attempts were done in this question [16] but not too successfully. Therefore a temporary solution was optionally accepted. The mean value EZ is taken as a center of the confidence interval. Distance of upper bound Z_u and lower bound Z_l was chosen as $Z_u - Z_l = 4 \sqrt{\text{var } Z}$. This interval in Gaussian approximation represents cca. 97% probability.

The *pdf* is normalized. Function value, mean value, variance, 1^{st} and 2^{nd} derivatives are available in the class. Computation procedure of the derivatives uses the original algorithm for *pdf* value modified in a particular way and performs complicated calls of methods inside the object, which is invisible from outside. This shows the advantage of using object-oriented tools here.

6.4 Calibration Coefficient

Calibration coefficient c , defined by (3.9), page 29, converts count of particles emitted by a source to the activity. For its estimation, a source of a known activity must be used.

The estimate of signal on background cannot be included as a part of further estimation tasks because of its computational difficulty (6.7). Therefore the impulses counts treated later in this chapter are point estimates of signal on background taken as means of confidence intervals yielded by (6.7). This simplification will cause an artificial decrease of uncertainty of the estimates mentioned below. But if $z \gg b$ (see part 6.3.1), point estimate of Z appears to be a good approximation [15].

6.4.1 Measurement

Let us have a “standard” source of the known activity S . Detecting emitted particles, let us perform n measurements with the geometrical measurement conditions unchanged. Result of j -th measurement is count s_j , where $j=1, \dots, n$. Counts s_j are independent. Let us assume that calibration coefficient c is constant.

6.4.2 Model

Data model has the form

$$f(s(..n)|c(..n), S) = \prod_{j=1}^n P_{s_j}(cS),$$

where $P_b(a)$ means Poisson distribution with mean value a and discrete variable b .

Parameter to be estimated is c .

6.4.3 Likelihood

The form of likelihood is (see part 4.2.2, page 37)

$$\begin{aligned} \mathcal{L}(s(..n); S) &\propto c^{\bar{s}(n)} e^{-c\bar{S}(n)} \\ \tilde{s}(n) &= \tilde{s}(n-1) + s_n, & \tilde{s}(0) &= 0 \\ \tilde{S}(n) &= \tilde{S}(n-1) + S_n, & \tilde{S}(0) &= 0. \end{aligned}$$

6.4.4 Prior Information

The expert knowledge is expressed by choice of a suitable value c_u that is upper limit for the geometrical conditions usually used for measurement. Lower limit is zero, as c must be non-negative. Using conjugated gamma-*pdf*

$$f(c) = G_c(\tilde{s}_0, \tilde{S}_0),$$

requesting that its mean value is $\frac{c_u}{2}$ and standard deviation is $\frac{c_u}{4}$ and using similar procedure as in (6.6), page 60, we will get

$$\tilde{s}_0 = 1, \quad \tilde{S}_0 = \frac{8}{c_u} \equiv \zeta. \quad (6.9)$$

For details see [19].

6.4.5 Estimation

According to the previous conclusions, the posterior *pdf* of the calibration coefficient is

$$\begin{aligned} f(c|s(..n), S) &= G_c(\tilde{s}(n), \tilde{S}(n)), \\ \tilde{s}(n) &= \tilde{s}(n-1) + s_n, & \tilde{s}(0) &= 1 \\ \tilde{S}(n) &= \tilde{S}(n-1) + S_n, & \tilde{S}(0) &= \zeta \equiv \frac{8}{c_u} \end{aligned} \quad (6.10)$$

and corresponding moments are

$$\mathbb{E}[c|s(..n), S] = \frac{\tilde{s}(n) + 1}{\tilde{S}(n)} \quad \text{var} [c|s(..n), S] = \frac{\mathbb{E}[c|s(..n), S]}{\tilde{S}(n)}.$$

6.4.6 Implementation Remarks

The posterior *pdf* is represented by the class `Pdf_cal`.

As gamma-*pdf* contains exponents and Γ -function arguments of high orders, enumeration problems were encountered. Therefore the posterior *pdf* is pseudo-normalized (see part 5.1.3, page 48). Because of using logarithm in pseudo-normalization, value in point 0 is defined as 0.

Function value, mean value, variance, mode, 1st and 2nd derivative are available.

6.5 Instantaneous Activity

6.5.1 Measurement

Count a of particles corresponding to the source of an unknown activity A is detected. Using the calibration information (see part 6.4, page 61), unknown activity A can be estimated.

6.5.2 Model

Model of detected particles counts a can be written as

$$f(a|c, A) \equiv f(a|c, A, s(..n), S(..n)) = P_a(cA), \quad (6.11)$$

where c is calibration factor, S is the standard source activity (known) and s is count of the corresponding standard source particles.

A is the parameter to be estimated.

6.5.3 Likelihood

Likelihood is $\mathcal{L}(a, s(..n), S(..n); c, A) \propto (cA)^a e^{-cA}$.

6.5.4 Prior Information

There is an upper bound of activity in the patient's body given by the *decay law* ((3.6), page 28), where the applied activity is substituted for A_0 . Let us denote this upper bound as A_u . Lower bound of the activity in the patient's body is 0. Estimated activity A must appear in the range $\langle 0; A_u \rangle$.

A specific form of prior *pdf* $f(A)$ will be derived later.

6.5.5 Estimation

As unknown activity A and calibration factor c are independent, the joint *pdf* can be expressed as

$$f(A, c|s(..n), S(..n), a) = f(A|a, c, s(..n), S(..n)) \cdot f(c|s(..n), S(..n)), \quad (6.12)$$

where $f(c|s(..n), S(..n))$ corresponds to (6.10) and

$$f(A|a, s(..n), S(..n)) \propto f(a|c, A, S(..n)) \cdot f(A). \quad (6.13)$$

As $f(A|s(..n), S(..n), a) = \int_0^{+\infty} f(A, c|s(..n), S(..n), a)dc$, considering (6.10) and (6.13) we get

$$f(A|s(..n), S(..n), a) = \tilde{K} f(A) \frac{A^a}{(A + \tilde{S}(n))^{a+\tilde{s}(n)+1}}, \quad (6.14)$$

where \tilde{K} is a normalizing constant and $\tilde{S}(n)$, $\tilde{s}(n)$ are described in (6.10).

The explicit form of prior *pdf* $f(A)$ is chosen as

$$f(A) \equiv f(A|s(..n), S(..n), \alpha, \beta) = K_p \frac{A^\alpha}{(A + \tilde{S}(n))^\beta},$$

where K_p is a normalizing constant of prior *pdf*, α and β are *fictitious data* (see part 4.1.5, page 34). If A_u is some upper limit that cannot be exceeded (see part 6.5.4), let us choose α and β so that $f(A)$ has mean value in $\frac{A_u}{2}$ and standard deviation is $\frac{A_u}{4}$.

Then the posterior *pdf* is

$$f(A|s(..n), S(..n), a) = K \frac{A^{\alpha+a}}{(A + \tilde{S}(n))^{\beta+a+\tilde{s}(n)+1}}, \quad (6.15)$$

where (see (6.10), page 62)

$$\begin{aligned} K &\equiv \left[\int_0^{+\infty} \frac{A^{\alpha+a}}{(A+\tilde{S}(n))^{\beta+a+\tilde{s}(n)+1}} dA \right]^{-1} = \tilde{S}(n)^{\beta-\alpha+\tilde{s}(n)} \cdot \frac{\Gamma(\beta+a+\tilde{s}(n)+1)}{\Gamma(a+\alpha+1) \Gamma(\beta-\alpha+\tilde{s}(n))} \\ \alpha &= \frac{\bar{A}^2(\bar{A}+\tilde{S}(n))^2 + D\bar{A}(2+\bar{A}-\tilde{S}(n)) - \tilde{S}(n)^2 D}{\bar{A}+\tilde{S}(n)} \\ \beta &= \frac{\bar{A}(\bar{A}+\tilde{S}(n))^2 + D(2+\bar{A}+\tilde{S}(n))}{\tilde{S}(n)D} \\ \bar{A} &= \frac{A_u}{2} \\ D &= \frac{A_u}{4}. \end{aligned} \quad (6.16)$$

For details see [19].

If we denote $[K(p, r, s)]^{-1} = \int_0^{+\infty} \frac{A^p}{(A+s)^r} dA$, then it is obvious that $E(A^j) = \frac{K(p, r, s)}{K(p+j, r, s)}$. Hence

$$E(A|\alpha, \beta, a, \tilde{S}(n), \tilde{s}(n)) = \tilde{S}(n) \frac{\alpha + a + 1}{\beta + \tilde{s}(n) - \alpha - 1}, \quad (6.17)$$

$$\text{var}(A|\alpha, \beta, a, \tilde{S}(n), \tilde{s}(n)) = [\tilde{S}(n)]^2 \frac{(\alpha + a + 1)(\beta + a + \tilde{s}(n))}{(\beta + \tilde{s}(n) - \alpha - 1)^2 (\beta + \tilde{s}(n) - \alpha - 2)}. \quad (6.18)$$

Expression for mode is $\hat{x} = \frac{\tilde{S}(n)}{\beta - a + \tilde{s}(n) + 1}$.

6.5.6 Implementation Remarks

The posterior *pdf* is represented by the class `Pdf_act`.

As the expression of the constant K contains exponents and Γ -function arguments of high orders, it is not evaluated and the posterior *pdf* is pseudo-normalized (see part 5.1.3, page 48). Because of using logarithm in pseudo-normalization, value in point 0 is defined as 0.

Function value, mean value, variance, mode, 1st and 2nd derivative are available.

6.6 Effective Half-life

Effective half-life in an important parameter characterizing activity kinetics. Time course of activity, obtained by compartment model, is a combination of polynomes, exponentials and goniometric functions with many parameters [11] estimation of which would be complicated and would suffer of lack of data. Therefore the time course of activity is modelled by 1st order exponential function that dominates over others and has only two estimated parameters.

6.6.1 Measurement

Sequence of measurements that are described in part 6.5.1, page 63, is performed. Time difference of subsequent measurements is usually about 24 hours.

6.6.2 Model

The time dependence of activity in some accumulating organ is a process of higher order. Its qualitative description is: (1) an increase from zero, (2) reaching the maximum, (3) a decrease to zero. Stages (1) and (2) take place within first few hours, so stage (3) covers majority of time and data measurements. This stage is usually modelled by mono-exponential function. Let t_1 be the time instant when activity of the organ reached its maximum (stage (2)).

Then unknown activities in t_j , $j = 1, 2, \dots, m$, follow the relation

$$\begin{aligned} A_j &\equiv A(t_j) = A_1 \cdot \exp \left[-\frac{(t_j - t_1) \cdot \ln 2}{T_{ef}} \right] \equiv A_1 \cdot \psi(T_{ef}, j) \\ \psi(T_{ef}, j) &\equiv \exp \left[-\frac{(t_j - t_1) \cdot \ln 2}{T_{ef}} \right] \end{aligned} \quad (6.19)$$

Counts $a_j = a(t_j)$, $j = 1, 2, \dots, m$ are considered Poisson-distributed and conditionally independent. Then

$$f(a(..m)|c, A_1, T_{ef}, s(..n), S(..n)) \propto \prod_{j=1}^m P_{a_j}(cA_1 \psi(T_{ef}, j)) \quad (6.20)$$

Parameters are c , A_1 and T_{ef} , where A_1 means unknown activity in time t_1 and T_{ef} is one to be estimated.

6.6.3 Likelihood

Likelihood has a form

$$\mathcal{L}(a(..m), s(..n), S(..n); c, A_1, T_{ef}) \propto c^{\bar{a}(m) + \bar{s}(n)} A_1^{\bar{a}(m)} \tilde{\Psi}(T_{ef}, m) \cdot e^{-c(A_1 \bar{\psi}(T_{ef}, m) + \bar{S}(n))}, \quad (6.21)$$

where $\bar{s}(n)$ and $\bar{S}(n)$ are described in (6.10), page 62, and

$$\begin{aligned} \bar{a}(m) &= \bar{a}(m-1) + a_m, & \bar{a}(0) &= 0, \\ \bar{\psi}(T_{ef}, m) &= \bar{\psi}(T_{ef}, m-1) + \psi(T_{ef}, m), & \bar{\psi}(T_{ef}, 0) &= 0, \\ \tilde{\Psi}(T_{ef}, m) &= \tilde{\Psi}(T_{ef}, m-1) \psi(T_{ef}, m)^{a_m}, & \tilde{\Psi}(T_{ef}, 0) &= 1. \end{aligned} \quad (6.22)$$

Note that functions $\bar{\psi}(T_{ef}, m)$ and $\tilde{\Psi}(T_{ef}, m)$ are *not* sufficient statistics as they depend on the unknown parameter T_{ef} .

6.6.4 Prior Information

There are three parameters in the model (6.20): calibration factor c , maximum activity A_1 and effective half-life T_{ef} . Let us discuss their dependence.

Calibration factor c expresses geometrical and technical conditions of the measurement, therefore it is independent of A_1 and T_{ef} . As A_1 results from complicated dynamics (as mentioned above) that is not investigated here, its dependence on T_{ef} is unknown. Hence both the parameters A_1 and T_{ef} should be considered a priori independent. It allows to discuss all three parameters separately.

Prior information on c is given by (6.10), page 62.

Lower bound of activity is 0. If we denote applied activity (that was administered into patient's body) as A_{apl} and application time as t_{apl} , it is obvious that activity in time t_1 anywhere in the body, according to the decay law (see (3.6), page 28), cannot exceed value

$$\bar{A}_1 = A_{apl} \cdot \exp \left(-\frac{t_1 - t_{apl}}{T_p} \ln 2 \right), \quad (6.23)$$

where T_p is physical half-life (see 3.5, page 28). As no other references are known, prior *pdf* of the parameter A_1 is taken as uniform density $f(A_1) = U_{A_1}(0, \bar{A}_1)$.

Considering the definition of effective half-life (3.7), page 28, prior *pdf* of T_{ef} can be expressed also uniform on $(0; T_p)$, i.e. $f(T_{ef}) = U_{T_{ef}}(0, T_p)$.

6.6.5 Estimation

Posterior *pdf* of T_{ef} can be found in two integration steps. At first, joint *pdf* of A_1 and T_{ef} is obtained by integration over c :

$$\begin{aligned} f(A_1, T_{ef} | a(..m), s(..n), S(..n)) &= \int_0^{+\infty} f(c, A_1, T_{ef} | a(..m), s(..n), S(..n)) dc \\ &\propto U_{A_1}(0, \bar{A}_1) U_{T_{ef}}(0, T_p) \frac{\tilde{\Psi}(T_{ef}, m) A_1^{\bar{a}(m)}}{\left[A_1 \tilde{\psi}(T_{ef}, m) + \tilde{S}(n) \right]^{\bar{a}(m) + \bar{s}(n) + 1}}. \end{aligned} \quad (6.24)$$

At second, the posterior *pdf* of T_{ef} is then obtained by intergration of (6.24) over A_1 in limits $(0, \bar{A}_1)$. For $\bar{A}_1 \rightarrow +\infty$, the explicit formula can be found. In this case, the calibration information is eliminated, thus unnecessary:

$$f(T_{ef} | a(..m), s(..n), S(..n)) \equiv f(T_{ef} | a(..m)) \propto U_{T_{ef}}(0, T_p) \frac{\tilde{\Psi}(T_{ef}, m)}{\tilde{\psi}(T_{ef}, m)^{\bar{a}(m) + 1}}. \quad (6.25)$$

For finite \bar{A}_1 , the integral of (6.24) must be found numerically.

6.6.6 Implementation Remarks

The posterior *pdf* is represented by the class `Pdf_tef`. The version for finite \bar{A}_1 is implemented. This class uses several other classes. Class `LPdfTef2` represents logarithm of joint *pdf* (6.24). Class `LPdfTef2` is used by class `Pdf_Tef2A1` that fixes T_{ef} and represents (6.24) as a function of A_1 . Class `Pdf_Tef` then integrates `Pdf_Tef2A1` over domain of A_1 . Class `LPdfTef2` is used also in prediction of radio-hygienic limits.

`LPdfTef2` is pseudo-normalized and offers function value and mode, classes `Pdf_Tef2A1` and `Pdf_Tef` give only function values.

6.7 Prediction of Reaching Radiohygienic Limits

After application, the patient is a source of radiation for his neighbourhood and therefore he has to be under radio-hygienic supervision. Meaning and importance of radiohygienic limits and their prediction were discussed in part 3.9, page 29.

This task is a modification of effective half-life estimation (part 6.6), therefore many steps will be the same.

6.7.1 Measurement

This part is the same like in part 6.6.1.

6.7.2 Model

Ideas of this part are the same like in 6.6.2, page 65. Furthermore, let us consider value of patient's activity A_h that determines change in his handling regulations during his hospitalization in the *KNM*. For the given A_1 and T_{ef} (see part 6.6.2, page 65), prediction of time t_h , when the patient's activity will sink to a given value A_h , can be done using the decay law

$$A_h = A_1 \exp \left[- \frac{t_h - t_1}{T_{ef}} \ln 2 \right]. \quad (6.26)$$

If we re-arrange this expression as

$$A_1 = A_h \exp \left[+ \frac{t_h - t_1}{T_{ef}} \ln 2 \right] \quad (6.27)$$

and substitute it to (6.24) for A_1 , we will get model of data of limit reaching prediction.

Parameters of this models are c , A_h and t_h , where the last one is to be estimated.

6.7.3 Likelihood

This part is the same like 6.6.3, considering the formula (6.27).

6.7.4 Prior Information

Ideas of this part are the same like in 6.6.4, page 65. As for t_h , it must be greater than t_1 and it cannot exceed value \bar{t}_h obtained by physical (the slowest possible) activity decrease

$$\bar{t}_h = t_{apl} + T_p \frac{\ln \frac{A_{apl}}{A_h}}{\ln 2}.$$

Prior *pdf* for t_h is taken as uniform $f(t_h) = U_{t_h}(t_1, \bar{t}_h)$.

6.7.5 Estimation

Substituting (6.27) into (6.24) for A_1 , we will get joint *pdf* $f(t_h, T_{ef} | A_h, a(..m), s(..n), S(..n))$. If we denote conditions of this *pdf* as \mathcal{P} , the marginal posterior *pdf* of t_h is then

$$f(t_h | \mathcal{P}) \propto U_{t_h}(t_1, \bar{t}_h) \int_0^{T_p} \frac{\tilde{\Psi}(T_{ef}, m) \exp\left[(\tilde{a}(m)+1) \frac{t_h-t_1}{T_{ef}} \ln 2\right]}{\left\{ A_h \exp\left[\frac{t_h-t_1}{T_{ef}} \ln 2\right] \tilde{\psi}(T_{ef}, m) + \tilde{S}(n) \right\}^{\tilde{a}(m)+\tilde{s}(n)+1}} \frac{dT_{ef}}{T_{ef}}. \quad (6.28)$$

This integral must be computed numerically.

6.7.6 Implementation Remarks

The posterior *pdf* is represented by the class `Pdf_apr`. This class uses several other classes. Class `LPdfTef2` represents logarithm of joint *pdf* (6.24). Class `LPdfTef2` is used by class `Pdf_AprT` that substitutes (6.27) for A_1 , fixes it and represents (6.28) as a function of T_{ef} . Class `Pdf_Apr` then integrates `Pdf_AprT` over domain of T_{ef} . Class `LPdfTef2` is used also for estimation of effective half-life T_{ef} .

`LPdfTef2` is pseudo-normalized and offers function value and mode, classes `Pdf_AprT` and `Pdf_Apr` give only function values.

6.8 Excretion of Activity

“Excretions” E_{ri} are important to determine some radio-hygienic quantities. They are defined by (3.10), page 29. Only three measurements are performed — one calibration immediately after application, when activity inside the patient is known, and two measurements with unknown patient’s activity ($i = 1, 2$).

Excretions can be defined also implicitly as

$$A_i = A_{apl} \psi(T_p, i) \left(1 - \sum_{j=1}^i E_{Rj} \right), \quad i = 1, 2, \quad (6.29)$$

where A_{apl} is applied activity, t_{apl} is application time, T_p is physical half-life, $A(i)$ means whole-body activity in time t_i and usual notation $\psi(T_p, i) = \exp\left(-\frac{t_i-t_{apl}}{T_p} \ln 2\right)$ is used. If the calibration measurement is denoted by subscript 0, then usually $t_0 \approx t_{apl}$, $t_1 \approx t_{apl} + 24$ hours and $t_2 \approx t_{apl} + 48$ hours. Note that $E_r = E_R \cdot 100\%$.

Values E_{r1} and E_{r2} are to be estimated.

Estimation will be derived for quantities E_{Ri} , change for E_{ri} can be then done by simple variable transformation (see part 3.3, page 28).

6.8.1 Measurement

To avoid direct measurements of patients' excrements activity, which is technically and methodologically complicated, unpleasant and non-hygienic, excretions are measured indirectly. Whole-body activity measurement is performed and then value of eliminated activity is estimated.

Impulses e_i in time t_i , $i = 0, 1, 2$ over the whole body are counted. The measurement in t_0 provides a calibration.

6.8.2 Model

Impulses counts are conditionally independent and Poisson-like distributed. If c is an unknown calibration factor assumed to be constant, A_i is the whole-body activity in time t_i and $m=2$, then

$$f(e_0, e(..m)|c, A(..m)) = \prod_{i=0}^m \frac{(cA_i)^{e_i}}{e_i!} \exp(-cA_i), \quad m = 2. \quad (6.30)$$

Parameters of this model are c and A_i , $i = 1, \dots, m$. Quantities to be estimated are $E_r(..m)$ that are "hidden" in $A(..m)$ through (6.29).

6.8.3 Likelihood

Substituting (6.29) into (6.30) and considering that $A_0 = A_{apl} \psi(T_p, 0)$, we get the expression for likelihood

$$\mathcal{L}(e_0, e(..m); c, E(..m)) \propto c^{\tilde{e}(m)} \exp\left[-cA_{apl} \left(\tilde{\psi}(T_p, m) - \tilde{\mathcal{E}}_R(m)\right)\right] \prod_{i=1}^m \left[1 - \tilde{E}_R(i)\right]^{e_i}, \quad (6.31)$$

where

$$\begin{aligned} \tilde{e}(m) &= \tilde{e}(m-1) + e_m, & \tilde{e}(0) &= e_0, \\ \tilde{\psi}(T_p, m) &= \tilde{\psi}(T_p, m-1) + \psi(T_p, m), & \tilde{\psi}(T_p, 0) &= \psi(T_p, 0), \\ \tilde{E}_R(m) &= \sum_{i=1}^m E_{Ri}, \\ \tilde{\mathcal{E}}_R(m) &= \sum_{i=1}^m \psi(T_p, i) \tilde{E}_R(i). \end{aligned}$$

6.8.4 Prior Information

Parameter c can change firstly during 48 hours as ^{131}I is being distributed over the organism, and secondly because of the possibility that geometrical conditions during measurements in time t_1 and t_2 can slightly differ from those in t_0 . As these changes are unknown, c must be considered constant and independent of E_{R1} and E_{R2} . Anyway, the calibration measurement is independent of the following ones.

The measurement in t_0 is simple calibration as described in section 6.4, page 61. Therefore its prior *pdf* is Gamma-*pdf* $G_c(0, \frac{1}{\bar{c}})$, where \bar{c} is some usual suitable value of c that is both mean value and standard deviation of the prior *pdf*. Because $\frac{1}{\bar{c}}$ is added to standard activity as prior information, it must be added to A_{apl} for the calibration measurement ($i = 0$). Looking at (6.31), we see that $\psi(T_p, 0)$ occurs only in sum. Then we can add $\frac{1}{\bar{c}A_{apl}}$ as initial condition for $\tilde{\psi}(T_p, 0)$:

$$\tilde{\psi}(T_p, 0) = \psi(T_p, 0) + \frac{1}{\bar{c}A_{apl}}. \quad (6.32)$$

As amount of radioactive solution diluted from the body cannot decrease, the sequence $\left\{\tilde{E}_R(i)\right\}_{i=1}^m$ is non-sinking. Furthermore $0 \leq \tilde{E}_R(i) \leq 1$ for any i , as the patient cannot dilute more than he has drunk. Putting this together, we get the condition

$$0 \leq E_{R1} \leq E_{R1} + E_{R2} \leq 1. \quad (6.33)$$

Then the prior *pdf* for E_{Ri} , $i = 1, 2$, is

$$f(E_R(..m)) = \prod_{i=1}^m U_{E_{Ri}}(0, 1 - \tilde{E}_R(i-1)) = \frac{\chi_{E_{R1}}(0, 1)}{1} \cdot \frac{\chi_{E_{R2}}(0, 1 - E_{R1})}{1 - E_{R1}} \quad (6.34)$$

for $m = 2$. Then domain of the joint posterior *pdf* is triangle, in which both E_{Ri} s are non-negative and their sum does not exceed 1.

6.8.5 Estimation

Putting together (6.31), (6.32) and (6.34), we get a joint posterior *pdf* of c , E_{R1} and E_{R2} . Integrating over c on $\langle 0, +\infty \rangle$ we get a joint *pdf* of E_{R1} and E_{R2} :

$$f(E_R(..m)|e(0), e(..m)) \propto \frac{\prod_{i=1}^m U_{E_{Ri}}(0, 1 - \tilde{E}_R(i-1)) [1 - \tilde{E}_R(i)]^{e_i}}{[\tilde{\psi}(T_p, m) - \tilde{\mathcal{E}}_R(m)]^{\tilde{e}(m)+1}}, \quad (6.35)$$

where

$$\begin{aligned} \tilde{e}(m) &= \tilde{e}(m-1) + e_m, & \tilde{e}(0) &= e_0, \\ \tilde{\psi}(T_p, m) &= \tilde{\psi}(T_p, m-1) + \psi(T_p, m), & \tilde{\psi}(T_p, 0) &= \psi(T_p, 0) + \frac{1}{\bar{c} A_{apl}}, \\ \tilde{E}_R(m) &= \sum_{i=1}^m E_{Ri}, & \tilde{E}_R(0) &= 0, \\ \tilde{\mathcal{E}}_R(m) &= \sum_{i=1}^m \psi(T_p, i) \tilde{E}_R(i) \end{aligned}$$

and $m = 2$.

Marginal *pdf*s will be obtained by numerical integration over another variable on its domain given by prior information.

6.8.6 Implementation Remarks

This estimation task is not solved in the object-oriented way. The function values of (6.35) are stored in a grid and numerically integrated. As the function values are calculated using their logarithms, the condition (6.33) must be changed to $0 \leq E_{R1} \leq E_{R1} + E_{R2} < 1$.

6.9 Radio-hygenic Quantities

Radionuclide present in the patient's body produces ionizing radiation. Necessary information to be known is which dose is absorbed in which organ, namely, if dose in the target organ (in this case thyroid gland) will be sufficiently high and if dose in other organs (e.g. blood, bone marrow, liver etc.) will not exceed some "safe" value.

In the *KNM*, two radio-hygenic quantities are being estimated: specific irradiation of thyroid gland and specific irradiation of blood.

6.9.1 Specific Irradiation of Thyroid Gland

As mentioned in part 3.10, page 30, the formula for specific irradiation of thyroid gland for 3,7 GBq of applied ^{131}I is

$$SI_{th} = \frac{139 A_{rmax}}{H} T_{ef},$$

where A_{rmax} is maximum relative activity of the thyroid gland [%], H is thyroid gland mass [g] and T_{ef} is effective half-life [days]. Unit of SI_{th} is Gy (=J/kg).

Bayesian estimate requires integration of complicated formula over two variables [11]. Furthermore, shape of the domain is irregular due to prior information, therefore some variable transformation and use of approximation integration formulae [39] is not well justified. As estimates of all input variables for the definition formula are available in their interval form, a more simple method of SI_{th} estimation was chosen. The main reasons for this decision were limited time for computation and perspective of using the MIRD method.

If x is some quantity, \hat{x} its point estimate and Δx its absolute error, the notation $x = \hat{x} \pm \Delta x$ represents some interval $(\hat{x} - \Delta x; \hat{x} + \Delta x)$, in our case the 95% symmetric confidence interval (see part 4.1.7, page 35). Let us denote lower bound of this interval $x_l = \hat{x} - \Delta x$ and upper bound $x_u = \hat{x} + \Delta x$.

Following the introduced notation, we can estimate $SI_{th,l}$ and $SI_{th,u}$ in the following “pessimistic” way:

$$\begin{aligned} SI_{th,l} &= \frac{139 A_{rmax,l}}{H_u} T_{ef,l}, \\ SI_{th,u} &= \frac{139 A_{rmax,u}}{H_l} T_{ef,u}. \end{aligned}$$

Then

$$\begin{aligned} \hat{SI}_{th} &= \frac{SI_{th,u} + SI_{th,l}}{2}, \\ \Delta SI_{th} &= \frac{SI_{th,u} - SI_{th,l}}{2} \end{aligned}$$

and $SI_{th} = \hat{SI}_{th} \pm \Delta SI_{th}$. This interval represents generally more than 95% probability [20]. The main source of uncertainty here is the estimate of thyroid gland mass.

6.9.2 Specific Irradiation of Blood

As mentioned in part 3.11, page 30, the formula for organic and anorganic fraction of maximum specific irradiation of blood for 3,7GBq of applied ^{131}I is

$$\begin{aligned} SIK_{an,max} &= \frac{160 (E_{r1} + E_{r2})}{M}, \\ SIK_{org,max} &= 13,9 (100 - E_{r1} - A_{r1}) T_{ef}. \end{aligned} \quad (6.36)$$

This formula contains four dependent variables that should be integrated off. Therefore similar approach like in the previous part was adopted for the same reason. Furthermore, data for estimation of $E_{ri}, i = 1, 2$, and A_{r1} are dependent. Integration domain is then non-trivially bounded by prior information which makes the integration task even more complicated.

Following the ideas from the previous part, we will find

$$\begin{aligned} SIK_{an,max,l} &= \frac{160 (E_{r1,l} + E_{r2,l})}{M}, \\ SIK_{an,max,u} &= \frac{160 (E_{r2,u} + E_{r2,u})}{M} \end{aligned}$$

and

$$\begin{aligned} SIK_{org,max,l} &= 13,9 (100 - E_{r1,u} - A_{r1,u}) T_{ef,l}, \\ SIK_{org,max,u} &= 13,9 (100 - E_{r1,l} - A_{r1,l}) T_{ef,u}. \end{aligned}$$

Then quantities $\hat{SIK}_{an,max}$, $\hat{SIK}_{org,max}$, $\Delta SIK_{an,max}$ and $\Delta SIK_{org,max}$ can be easily computed. The confidence intervals generally represent again the probability greater than 95%.

6.9.3 Medical Internal Radiation Dose — the MIRD Method

As mentioned in part 3.12, page 31, the formula for the mean absorbed dose in target organ is

$$\bar{D}_t = A_{apl} \sum_s \tau_s \mathcal{S}_{s \rightarrow t}. \quad (6.37)$$

If we consider thyroid gland as the only source organ of destructive β -radiation [11], we can denote τ_s explicitly as τ_{th} . To determine the residence time τ_{th} , we must compute integral below the curve $A_{th}(t)$ in the thyroid gland. In the part 6.6.2, page 65, we adopted the simplified model of single-exponential activity sinking for $t \geq t_{1,th}$ (phase (3) of the time course). To estimate the whole integral from t_{apl} to $+\infty$, we must determine the phase (1) of the time course, i.e. $A_{th}(t)$ for $t_{apl} \leq t \leq t_{1,th}$. As we know only that this phase is relatively fast, we can approximate it by line segment described by linear function $A_{th}(t) = \frac{A_{1,th}}{t_{1,th}} (t - t_{apl})$.

To summarize, we have for $A_{th}(t)$:

$$\begin{aligned} \text{phase (1):} & \quad \text{fast increase to the maximum} & \quad t_{apl} \leq t \leq t_{1,th} & \quad A_{th}(t) = \frac{A_{1,th}}{t_{1,th}} (t - t_{apl}) \\ \text{phase (3):} & \quad \text{exponential decrease} & \quad t \geq t_{1,th} & \quad A_{th}(t) = A_{1,th} \exp\left(-\frac{t-t_{1,th}}{T_{ef,th}} \ln 2\right), \end{aligned}$$

where $T_{ef,th}$ is effective half-life in thyroid gland. Then the residence time τ_{th} can be computed as

$$\tau_{th} = \frac{A_{1,th}}{A_{apl}} \left[\frac{t_{1,th}}{2} + \int_{t_{1,th}}^{+\infty} \exp\left(-\frac{t-t_{1,th}}{T_{ef,th}} \ln 2\right) dt \right].$$

The solution is

$$\tau_{th} = \frac{A_{1,th}}{A_{apl}} \left(\frac{t_{1,th}}{2} + \frac{T_{ef,th}}{\ln 2} \right).$$

Furthermore, the mean absorbed dose for thyroid gland as a target is related to some standard thyroid gland mass, which is for adult male 20,7 g [35]. If we know “true” thyroid gland mass, which is given by the Bayesian estimate H , and considering thyroid gland as the only source organ of β -radiation, we can modify the formula (6.37) for thyroid gland as

$$\bar{D}_{th \rightarrow th} = A_{apl} \frac{20,7 \text{ g}}{H} \tau_{th} \mathcal{S}_{th \rightarrow th},$$

where $\bar{D}_{th \rightarrow th}$ explicitly means the mean absorbed dose for thyroid gland.

The S-factor $\mathcal{S}_{th \rightarrow th}$ is known as number.

What we must keep in mind is that quantities $A_{1,th}$, $T_{ef,th}$ and H are Bayesian estimates represented by their confidence intervals. Following the same ideas like in previous two parts, we can write for the residence time $\tau_{th \rightarrow th}$:

$$\begin{aligned} \tau_{th,l} &= \frac{A_{1,th,l}}{A_{apl}} \left(\frac{t_{1,th}}{2} + \frac{T_{ef,th,l}}{\ln 2} \right), \\ \tau_{th,u} &= \frac{A_{1,th,u}}{A_{apl}} \left(\frac{t_{1,th}}{2} + \frac{T_{ef,th,u}}{\ln 2} \right) \end{aligned}$$

and for the mean absorbed dose for thyroid gland $\bar{D}_{th \rightarrow th}$:

$$\begin{aligned} \bar{D}_{th \rightarrow th,l} &= A_{apl} \frac{20,7 \text{ g}}{H_u} \tau_{th,l} \mathcal{S}_{th \rightarrow th}, \\ \bar{D}_{th \rightarrow th,u} &= A_{apl} \frac{20,7 \text{ g}}{H_l} \tau_{th,u} \mathcal{S}_{th \rightarrow th}. \end{aligned}$$

For adult male, $\mathcal{S}_{th \rightarrow th} = 5,64 \text{ mGy}/(\text{MBq hour})$ which corresponds to $135,36 \text{ mGy}/(\text{MBq day})$.

For application to thyroid gland as a source and bone marrow as a target, the formula (6.37) will have a form

$$\bar{D}_{th \rightarrow marrow} = A_{apl} \tau_{th} \mathcal{S}_{th \rightarrow marrow},$$

where for adult male $\mathcal{S}_{th \rightarrow marrow} = 8,5 \cdot 10^{-3}$ mGy/(MBq hour), i.e. 0,204 mGy/(MBq day).

Determining $\bar{D}_{s \rightarrow t}$ for other organs, we can use the same ideas. Note that the estimate does not include the model errors (both MIRD and approximation of time course of $A_{th}(t)$) and uncertainty of $\mathcal{S}_{th \rightarrow th}$.

Chapter 7

Experiments with Data and the Algorithms

In this chapter, some tests of the algorithms are described. Furthermore for some tasks, outputs of the Bayes-based algorithms are compared to “deterministic” results to judge the quality increase.

7.1 General Testing of the Algorithms

First, the numerical algorithms were tested for their behaviour and reliability. The focus was taken on the numerical precision.

7.1.1 Batch Processing — Test of Stability

Testing of algorithms on data is important to consider their stability and reliability. However, testing on “artificial” simulated data can hide some pitfalls, because numbers prepared artificially are “pretty” and need not reflect situations that can be met with real data, or other reasons (e.g. human mistakes etc.).

In the last five years, data of the patients treated at the *KNM* were collected. Till now, more than 5 000 data records are available, where one record means one activity application and following sequence of measurements.

The algorithms were run on all the available records of real data and no failure was detected.

7.1.2 Precision of QUANC8

Chapter 5, page 43, was dedicated to solution of numerical tasks with stress on the precision and reliability. In sections concerning integration step design, integration table design and analysis of integration interval restriction, values of distribution functions (integrals) of normal and logarithmic-normal *pdfs* were used. It is well known that these functions are not analytically integrable.

For that purpose, numerical integrals were computed using the algorithm QUANC8 (see page 40). The main advantage of this algorithm is that it can compute the integral with the required error, relative or absolute, provided that the function is continuous up to its 9th derivative.

The precision was tested in the following way. Normal *pdf* was integrated within limits 0 and x for x changing from 0 to 5 with step 0,5. The integrals were first computed by QUANC8 with required relative error 10^{-15} . Then these results were compared with ones in the tables [1]. In these tables, integrals of normal *pdf* for $0 < x \leq 3$ are calculated with 15 decimal places, integrals for $3 < x \leq 5$ with 10 decimal places.

The results for $x \in \langle 0,5; 3 \rangle$ were in agreement for all the 15 decimal places. Values of estimated error, which is a part of the QUANC8's output, were from $2,1 \cdot 10^{-17}$ to $5,3 \cdot 10^{-16}$.

As for $x \in \langle 3; 5 \rangle$ the tables yield only 10 decimal places, the decision about the algorithm precision could not be done.

The conclusion is that for sufficiently smooth functions, the QUANC8 can be used to compute numerical integrals with a precision as required. For the variable type in use, a specific lower limit must be respected. For the C-type `double` (8 bytes, 53 bits for mantissa, 9 bits for exponent, 2 sign bits), the lower bound is $\delta_m = 2^{-52}$, which is approximately $2 \cdot 10^{-16}$. For the C-type `float` (4 bytes, 24 bits for mantissa, 6 bits for exponent, 2 sign bits), the lower bound is $\delta_m = 2^{-23}$, which is approximately $1 \cdot 10^{-7}$. Requirements for a lower integration error has no meaning as δ_m is such a minimum δ , for which is valid $1 + \delta > 1$ in the given variable type. In other words, lower error, i.e. higher precision, is not possible to be reached as no more significant digits are available.

The integrals computed in this work by this algorithm can be considered reliable within the requested accuracy.

7.1.3 Confidence Interval

Here, some tests are made with algorithms building the task of confidence interval determining. Specifically, questions of integration step design, allowed restriction of *pdf*'s integration interval and the integration procedure itself are focused. The tests and experiments are performed with normal *pdf* for which the algorithms and criteria were designed.

Integration step design

Here, the validity of the formula (5.5), page 44, is tested.

The normal *pdf* $\mathcal{N}_x(0, 1)$ on the specific segment s_{kl} is integrated in $m = 2, 4, 6$ etc. integration steps and absolute error ϵ_{real} of each integration is computed. Then $m_{\epsilon_{max}}$ is found, which is such highest m , so that $|\epsilon_{real}| \leq \epsilon_{max}$. Then $m_{\epsilon_{max}}$ and ϵ_{real} are compared with m_{kl} and ϵ_{max} given by (5.5).

The results are shown in the Table 7.1.

| k | l | $\epsilon_{max} = 10^{-4}$ | | | $\epsilon_{max} = 10^{-5}$ | | | $\epsilon_{max} = 10^{-6}$ | | |
|-----|-----|----------------------------|----------------------|----------------------|----------------------------|---------------------|----------------------|----------------------------|----------------------|----------------------|
| | | m_{kl} | ϵ_{real} | $m_{\epsilon_{max}}$ | m_{kl} | ϵ_{real} | $m_{\epsilon_{max}}$ | m_{kl} | ϵ_{real} | $m_{\epsilon_{max}}$ |
| 0 | 1 | 4 | $1,1 \cdot 10^{-5}$ | 4 | 6 | $2,1 \cdot 10^{-6}$ | 6 | 10 | $2,7 \cdot 10^{-7}$ | 8 |
| 1 | 2 | 4 | $-1,3 \cdot 10^{-5}$ | 4 | 6 | $2,6 \cdot 10^{-6}$ | 6 | 8 | $-8,1 \cdot 10^{-7}$ | 8 |
| 2 | 3 | 2 | $-2,2 \cdot 10^{-5}$ | 2 | 4 | $7,9 \cdot 10^{-7}$ | 4 | 6 | $1,4 \cdot 10^{-7}$ | 4 |
| 3 | 4 | 2 | $2,5 \cdot 10^{-5}$ | 2 | 4 | $1,6 \cdot 10^{-6}$ | 4 | 6 | $3,1 \cdot 10^{-7}$ | 4 |
| 4 | 5 | 2 | $1,8 \cdot 10^{-6}$ | 2 | 2 | $1,8 \cdot 10^{-6}$ | 2 | 4 | $1,4 \cdot 10^{-7}$ | 4 |
| 5 | 10 | 4 | $3,4 \cdot 10^{-7}$ | 2 | 6 | $1,5 \cdot 10^{-7}$ | 2 | 12 | $1,9 \cdot 10^{-8}$ | 2 |

Table 7.1: Numbers of integration steps found theoretically (m_{kl} , see (5.5), page 44) and experimentally ($m_{\epsilon_{max}}$)

Values ϵ_{real} are less than ϵ_{max} , because those ϵ_{real} were selected that are *below* ϵ_{max} . It is visible that theoretical m_{kl} and experimental $m_{\epsilon_{max}}$ match for $l \leq 5$ and $\epsilon_{max} \in \langle 10^{-4}; 10^{-5} \rangle$. For $\epsilon_{max} = 10^{-6}$, theoretical m_{kl} are in majority of cases higher than $m_{\epsilon_{max}}$, but never more than by 2 for $l \leq 5$. This difference can be caused both by inequality (5.2), page 44, and limited set of values of m .

For $k = 5$ and $l = 10$, the higher precision is required, the higher is the difference between m_{kl} and $m_{\epsilon_{max}}$. This is caused by a relatively large interval $\langle 5; 10 \rangle$ on which the *maximum* of the 4th derivative in the design of m_{kl} is taken. The actual values throughout the interval are obviously much lower. If we decrease value of $g_{5,10}$ in the Table 5.1, page 44, from 0,33331 to 0,05 (i.e. approximately by 1/6), the values of m_{kl} correspond with $m_{\epsilon_{max}}$ for the interval $\langle 5; 10 \rangle$.

Anyway, numbers of integration steps found theoretically and experimentally are in a good agreement. The integration error is never underestimated and increase of m_{kl} due to overestimation is significantly higher only in interval $\langle 5; 10 \rangle$ if higher accuracy is required. It was observed that

because of the integration domain restriction (see page 45), this interval is rarely involved in the integration.

Confidence interval shift due to domain limitation

Here, the validity of the formulae (5.10) and (5.11), page 46, is tested.

The *pdf* $\mathcal{N}_x(0, 1)$ was used to find confidence interval with $\alpha = 0,04999570$ which represents probability 95,00042097%. As for the interval, $\hat{x} = 0$ and $-x_{l,exact} = x_{u,exact} = 1,96$ [1].

Various values of ϵ_{max} (see (5.2), page 44), and ϵ_a (see (5.11), page 46), were combined. Only upper bound of the confidence interval (due to symmetry) was computed in two ways: with the correction to domain limitation (x_u , see (5.10), page 46) and without it (\tilde{x}_u). Value \hat{x} is the mean of the confidence interval. The upper value of the domain x_{ϵ_a} for the given Δx is also mentioned. $\Delta\tilde{x}_u = \tilde{x}_u - x_{u,exact}$, $\Delta x_u = x_u - x_{u,exact}$. The significant results are shown in the Table 7.2.

| Δx | ϵ_{max} | x_{ϵ_a} | \hat{x} | no correction | | corrected | |
|------------|------------------|------------------|-----------------------|---------------|----------------------|------------|----------------------|
| | | | | \tilde{x}_u | $\Delta\tilde{x}_u$ | x_u | Δx_u |
| 10^{-2} | 10^{-5} | 3,5 | $2,2 \cdot 10^{-16}$ | 1,9535 | $-6,5 \cdot 10^{-3}$ | 1,9616 | $1,6 \cdot 10^{-3}$ |
| 10^{-3} | 10^{-6} | 4,5 | $1,1 \cdot 10^{-16}$ | 1,959105 | $-9,9 \cdot 10^{-4}$ | 1,959917 | $-8,3 \cdot 10^{-5}$ |
| 10^{-4} | 10^{-8} | 4,7 | $-1,0 \cdot 10^{-15}$ | 1,959907 | $-8,3 \cdot 10^{-5}$ | 1,959989 | $-2,1 \cdot 10^{-5}$ |
| 10^{-5} | 10^{-9} | 5,2 | $-4,4 \cdot 10^{-16}$ | 1,95999086 | $-8,1 \cdot 10^{-6}$ | 1,95999989 | $-2,1 \cdot 10^{-7}$ |

Table 7.2: Mean value and upper limit of confidence interval $\alpha = 0,04999570$ with various ϵ_{max} and ϵ_a . Exact values: $\hat{x}_{exact} = 0$, $x_{u,exact} = 1,96$.

The rows represent the best result for the pair Δx and ϵ_{max} in the sense of minimum $\Delta\tilde{x}_u$ and Δx_u . We can see that corrected values x_u have lower errors than \tilde{x}_u without the correction. Anyway, the errors are below the limit Δx in all the cases.

For the practice, $\Delta x = 10^{-3}$ was chosen. Value of ϵ_a is then approximately $\Delta x/10 = 10^{-4}$ (see (5.12), page 46). Value of ϵ_{max} is then chosen $\epsilon_a/100 = 10^{-6}$. This scheme can be used for ranges approximately $10^{-2} \leq \Delta x \leq 10^{-4}$. But according to our point of view, relative precision $\Delta x \approx 10^{-2}-10^{-3}$ is sufficient.

Similar results for different precisions were obtained also for some other levels of α , particularly approximately 0,34 and 0,02.

If we compare the Table 7.2 with the Table 5.2, page 47, we can notice that $x_{\epsilon_a} < x_{\epsilon_{au}}$ for the given $\epsilon_a = \Delta x/10$. This difference is caused by the fact that x_{ϵ_a} is the point where the integration can be stopped to reach “exactly” ϵ_a , whereas $x_{\epsilon_{au}}$ is the node of the numerical integration closest to x_{ϵ_a} , so that real error of area determination ϵ_{au} , that is reached during the step-by-step integration, is *less* than ϵ_a .

7.1.4 Simplex Method

It was already mentioned here that the simplex method in some conditions converges to a non-stationary point. Some modifications of the algorithm were done to prevent this case, as mentioned above. Anyway, the two-variate functions used here are smooth enough to avoid the case of unwanted convergence.

Several thousands of computations on various two-dimensional *pdfs* were done with different starting conditions. All the experiments gave correct results and no problems with convergence were detected.

This experience made us accept the simplex method for this work as an optimizing method that does not require derivatives and is not computationally demanding.

7.1.5 Integration Table

For integrating the logarithmic-normal *pdf*, the integration table was designed, so that error of quadratic interpolation is less than 10^{-4} (see part 5.1.5, page 49). Here, the error of interpolation is tested.

Integrals $I(x) = \int_0^x L_t(1,1) dt$ were computed. Values x were generated as pseudo-random numbers on the interval $(0; A)$, where A was some chosen fixed number. The integrals were computed in two ways: using algorithm QUANC8 with requested relative precision 10^{-10} (quantity $I_Q(x)$) and using the interpolation table (quantity $I_t(x)$). Then values $\Delta I(x) = I_t(x) - I_Q(x)$ were investigated.

For each A , 10000 values of x were generated and some descriptive statistics of $\Delta I(x)$ and $|\Delta I(x)|$ were calculated. The Table 7.3 shows results for $A = 2$ which represents the area with the greatest curvature of the function.

| quantity | mean | std. dev. | min. | max. |
|-------------------|-----------------------|----------------------|-----------------------|----------------------|
| $\Delta I(x)$ | $-9,56 \cdot 10^{-6}$ | $1,29 \cdot 10^{-5}$ | $-4,77 \cdot 10^{-5}$ | $5,04 \cdot 10^{-5}$ |
| $ \Delta I(x) $ | $1,11 \cdot 10^{-5}$ | $1,58 \cdot 10^{-5}$ | $3,21 \cdot 10^{-9}$ | $5,04 \cdot 10^{-5}$ |
| $\Delta I(x) > 0$ | | 7207 cases | | |
| $\Delta I(x) < 0$ | | 2739 cases | | |

Table 7.3: Testing of the integration table for the maximum required error of interpolation 10^{-4}

It is visible that maximum difference of $I_Q(x)$ and $I_t(x)$ is about half of the requested one. Similar results were obtained for different values of A . The actual error never exceeded the maximum allowed one.

7.1.6 Integration Table vs. Gaussian Approximation for Log-normal *pdf*

Posterior *pdf* of thyroid gland mass estimation is logarithmic-normal *pdf*.

It is obvious that using the integration table designed for the logarithmic-normal *pdf* in this estimation task is more correct than application of rules designed for normal *pdf* to logarithmic-normal *pdf*. In this section, thyroid gland mass confidence interval is computed using both the integration table and the rules for Gaussian approximation. Both the results are compared.

Actually, ‘‘Gaussian approximation’’ does not mean that logarithmic-normal *pdf* is replaced by normal *pdf*. It means (see the section 5.1.1, page 43) that some rules for lengths of integration steps, points where to stop the integration etc., designed for normal *pdf*, are applied to another *pdf* which is then integrated according to these rules.

Focusing on one task, design of the integration step, we can make similar experiment like in the part 7.1.3, page 74, with the Table 7.1 but the computations are done with the logarithmic-normal *pdf*. If we keep the same notation, the results are shown in the Table 7.4.

| | | $\epsilon_{max} = 10^{-5}$ | | |
|-----|-----|----------------------------|----------------------|----------------------|
| k | l | m_{kl} | ϵ_{real} | $m_{\epsilon_{max}}$ |
| 0 | 1 | 6 | $-7,9 \cdot 10^{-6}$ | 16 |
| 1 | 2 | 6 | $3,4 \cdot 10^{-6}$ | 6 |
| 2 | 3 | 4 | $2,9 \cdot 10^{-6}$ | 4 |
| 3 | 4 | 4 | $8,5 \cdot 10^{-6}$ | 2 |
| 4 | 5 | 2 | $2,1 \cdot 10^{-6}$ | 4 |
| 5 | 10 | 6 | $6,3 \cdot 10^{-6}$ | 6 |

Table 7.4: Numbers of integration steps found theoretically (m_{kl} , see (5.5), page 44) and experimentally ($m_{\epsilon_{max}}$) for logarithmic-normal *pdf*

We can see that between 0 and 1 the predicted number of integration steps ($m_{01}=6$) is much less than actual one ($m_{\epsilon_{max}}=16$) to get under the required integration error $\epsilon_{max}=10^{-5}$. We expect that error limits for confidence interval shift and its correction would disagree as well.

We performed another experiment concerning thyroid gland mass estimation more closely. We generated 1000 pairs of (h, ω) (see 6.1, page 55), estimated posterior *pdf* of H and constructed confidence interval with each pair twice on the same posterior *pdf* — once using the integration table, once using the Gaussian approximation. Values of h were generated from range $\langle 0; 10 \rangle$ and values of ω from $\langle 0; 0,6 \rangle$. We compared lower bound H_{lt} and upper bound H_{ut} of the confidence interval obtained using the table with lower H_{lg} and upper H_{ug} bounds obtained by the Gaussian approximation. Quantities

$$\Delta_{lr} = \frac{H_{lg} - H_{lt}}{H_{lt}}, \quad \Delta_{ur} = \frac{H_{ug} - H_{ut}}{H_{ut}}$$

were evaluated for each pair of (h, ω) and processed. Arithmetic mean $\bar{\Delta}_r$, sample standard deviation s_r , maximum and minimum deviations were found, see Table 7.5.

| | Lower bound | Upper bound |
|--------------------|---------------------------------------|--|
| arithmetic mean | $\bar{\Delta}_{lr} = -0,104$ | $\bar{\Delta}_{ur} = 2,25 \cdot 10^{-2}$ |
| standard deviation | $s_{lr} = 0,169$ | $s_{ur} = 1,03 \cdot 10^{-2}$ |
| minimum value | $\Delta_{lr,min} = -0,593$ | $\Delta_{ur,min} = -0,0017$ |
| maximum value | $\Delta_{lr,max} = 9,3 \cdot 10^{-4}$ | $\Delta_{ur,max} = 0,0320$ |

Table 7.5: Confidence intervals limits for thyroid gland mass estimate. Difference of results determined using integration table and Gaussian approximation

It is visible that because of using the Gaussian approximation instead of the integration table, the lower bound of the confidence interval of H is decreased by cca 10% in average with a relatively high dispersion and the upper bound is increased by approximately 2,3% in average.

According to this result, the integration table is used rather than the Gaussian approximation to determine confidence interval in the task of thyroid gland mass estimation.

7.1.7 Run Times

For illustration, run times of some computation programs are shown in the Table 7.6. The tasks of signal, effective half-life, prediction of RH limits, excretions and thyroid gland mass are tested. The programs are run 100-times on some usual data set, where run time does not depend much on data. If run time is strongly dependent on data, several cases are mentioned with different values.

The programs were run on Pentium 100MHz with MS-Windows 95. The run times include reading and writing data files.

Notes:

- Symbol “(1)” by the items of signal means “honest” computing of symmetric confidence interval on the *pdf*
- Symbol “(2)” by the items of signal means simplified confidence interval computing as mentioned in the part 6.3.6, page 61.
- Numbers “5”, “50” and “500” by the items of signal mean number of impulses (ionizing particles) detected in 1s. It is visible how computation time grows with increasing number of impulses.
- In case of thyroid gland mass estimation, only one computation is done (see part 6.1.6, page 58).

Even on as ancient machine as Pentium 100 MHz today (1999), the run times are short enough so that the programs can be used during daily work at the *KNM* with the equipment that *KNM* has.

| task | | run time [s] |
|--------------------|-----|--------------|
| signal (1) | 5 | 0,15 |
| signal (1) | 50 | 2,0 |
| signal (1) | 500 | 18,0 |
| signal (2) | 5 | 0,04 |
| signal (2) | 50 | 0,065 |
| signal (2) | 500 | 0,30 |
| T_{ef} | | 0,35 |
| RH prediction | | 2,2 |
| excretions | | 1,2 |
| thyroid gland mass | | 0,0045 |

Table 7.6: Runtimes of numerical programs

7.2 Experiments with Batch Processing of Data

In this section, results of some batch data processing is presented. Namely, Bayesian and deterministic estimates are compared for effective half-life, predictions or radio-hygienic limits and excretions. Further, influence of data amount and quality on effective half-life and thyroid gland mass estimate and other experiments is shown.

For the computations, real patients data collected for about five years were used. The data file contains about 5000 records, where one record means data of one application of ^{131}I and the sequence of various following measurements.

All the data presented here are anonymous and concern patients who are suspicious of suffering from thyroid gland carcinome.

Several times, quantities

$$\begin{aligned}\Delta_r(a, b) &= \frac{b - a}{a} \\ \delta_r(a, b) &= \frac{b}{a}\end{aligned}\tag{7.1}$$

will be used to consider relative difference of quantities a and b .

7.2.1 Comparison of Bayesian and Least-square T_{ef} -estimate

From the data file, 2978 records were selected for which an estimate of T_{ef} could be computed by least squares method. Denote the least-square estimate by T_{ef}^{LS} and the Bayesian estimate $\hat{T}_{ef} \pm \delta T_{ef}$.

According to (3.7), page 28, $0 < T_{ef} \leq T_p$, where T_p is physical half-life. After screening of T_{ef}^{LS} , it was found that

- for 96 records (3,2%), $T_{ef}^{LS} \leq 0$, where minimum value was $-97,4$ days,
- for 153 records (5,2%), $T_{ef}^{LS} > T_p = 8,04$ days, where maximum value exceeded range of the field (greater than 999,99 days), second maximum was 662,31 days,
- together, for 249 records (8,4%), estimate of T_{ef}^{LS} was out of the range (3.7), i.e. had no physical meaning.

Then variable $\Delta_r(\hat{T}_{ef}, T_{ef}^{LS})$ was introduced (see (7.1)). Histogram of $\Delta_r(\hat{T}_{ef}, T_{ef}^{LS})$ is shown on the Figure 7.1.

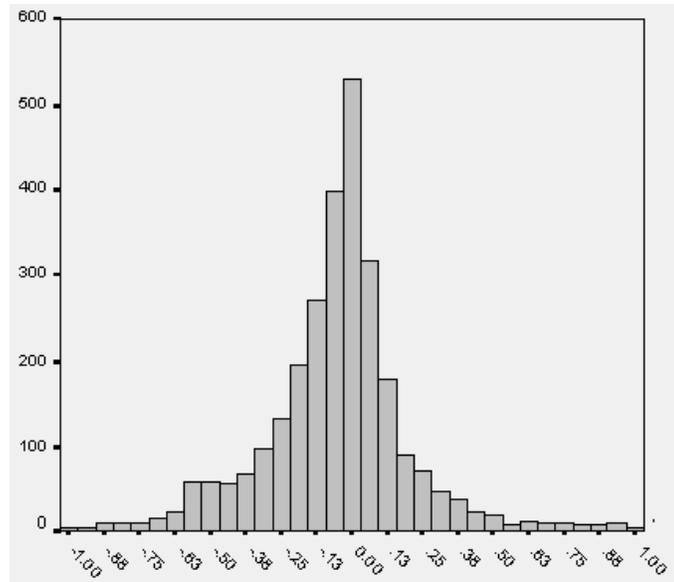


Figure 7.1: Comparison of Bayesian and least-squares estimate of T_{ef} , histogram of $\Delta_r(\hat{T}_{ef}, T_{ef}^{LS})$, filtered for $-1 \leq \Delta_r(\hat{T}_{ef}, T_{ef}^{LS}) \leq 1$, used 2 787 records

7.2.2 Difference Between Diagnostic and Therapeutic T_{ef}

It was observed that T_{ef} in diagnostic phase usually differs significantly from T_{ef} in therapeutic phase immediately following, although it is assumed that they are the same. This assumption is a key step to estimate/predict dose absorbed in an organ (see parts 3.10, page 30, 3.11 etc.).

A quantity

$$\rho = \frac{\text{diagnostic } T_{ef}}{\text{therapeutic } T_{ef}} \quad (7.2)$$

is introduced. 702 pairs of diagnostic and immediately following therapeutic applications in case of thyroid gland carcinoma were analyzed. It was found that arithmetic mean $\bar{\rho} = 2,94$ and standard deviation $\sigma_{\rho} = 1,72$. This result is repeated in the Table 7.7. It means that diagnostic T_{ef} is in

| | |
|--------------------|------------------------|
| arithmetic mean | $\bar{\rho} = 2,94$ |
| standard deviation | $\sigma_{\rho} = 1,72$ |

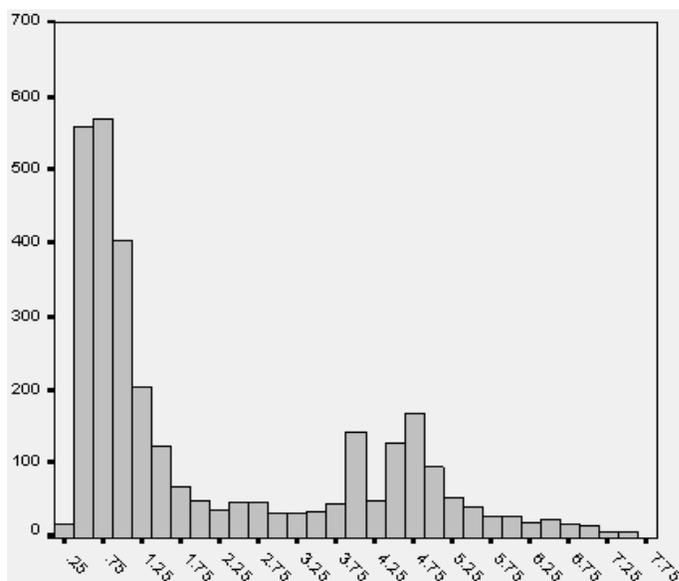
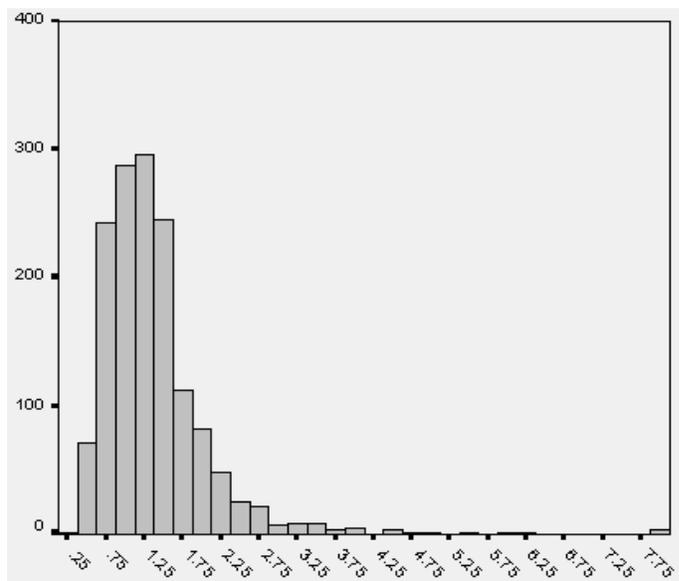
Table 7.7: Arithmetic mean and standard deviation of ρ — ratio of diagnostic and subsequent therapeutic T_{ef} , see (7.2)

average almost three-times greater than therapeutic T_{ef} .

If we compare histograms of T_{ef} s obtained in diagnostics (Figure 7.2) and therapy (Figure 7.3), we can see that therapeutic T_{ef} tends to be shorter.

Scanning the data, we can notice that number of diagnostic measurements is usually 2–3 whereas number of therapeutic measurements often exceeds 5, even reaches up to 12 or more. The question is if the difference in figures 7.2 and 7.3 is not caused by the number m of measurements and therefore by properties of model of T_{ef} (6.19). The Table 7.8 shows relations of diagnostic/therapeutic application and number of measurement m . 4939 records were used.

Let us plot histograms of T_{ef} with number of measurements $m \leq 3$ (Figure 7.4) and $m > 3$ (Figure 7.5). Although the value $m = 3$ separates diagnostic and therapeutic applications better than other values, it is not a reliable criterion. For this reason the figures 7.3 and 7.5 differ most significantly by a small peak in values of $T_{ef} > 4$ which is completely absent in therapeutic cases.

Figure 7.2: Histogram of T_{ef} [days], estimates in diagnostic phase, 3 412 recordsFigure 7.3: Histogram of T_{ef} [days], estimates in therapeutic phase, 1 522 records

| | | |
|-----------|------------|---------------|
| diagnosis | | 3 412 records |
| therapy | | 1 522 records |
| diagnosis | $m \leq 3$ | 3 057 records |
| therapy | $m > 3$ | 1 236 records |
| diagnosis | $m > 3$ | 355 records |
| therapy | $m \leq 3$ | 291 records |
| | $m \leq 3$ | 3 348 records |
| | $m > 3$ | 1 591 records |

Table 7.8: Frequencies of different numbers of measurements m in diagnostics and therapy

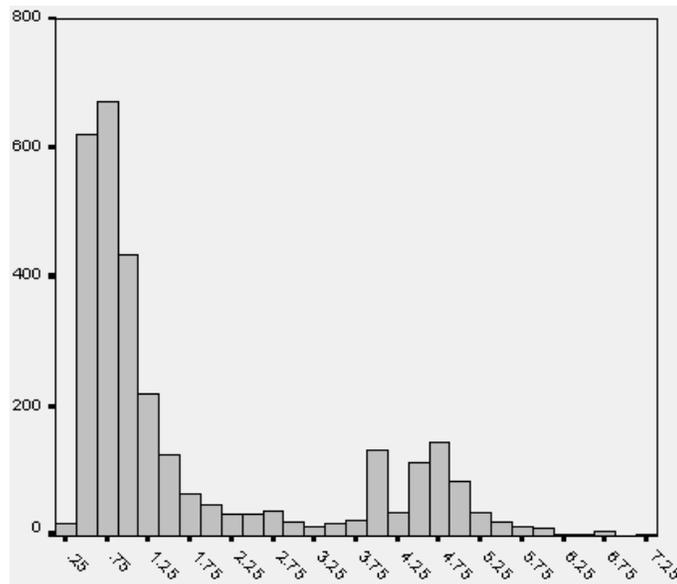


Figure 7.4: Histogram of T_{ef} [days], estimates for number of measurements $m \leq 3$, 3 348 records

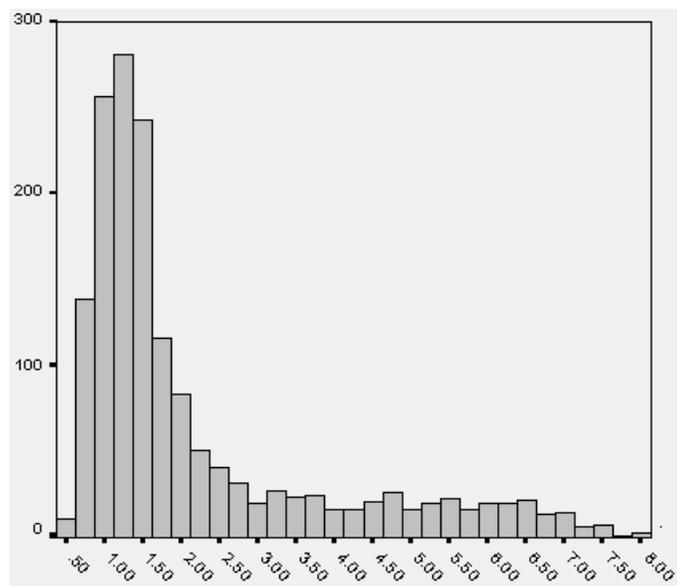


Figure 7.5: Histogram of T_{ef} [days], estimates for number of measurements $m > 3$, 1 591 records

Further, we can ask about origin of the smaller peak on the figures 7.2 and 7.4.

It can indicate:

- medical reasons (organism behaves differently with low and high activity application, grouping of diagnostic results — mixtures, etc.),
- artefacts caused by insufficient properties of the model of T_{ef} (6.19), page 65.

This problem definitely requires further analysis that falls beyond scope of this thesis.

7.2.3 Prediction of Activity in a Measurement Sequence

One of the quality tests of effective half-life estimate is this one: estimate T_{ef} with k measurements and predict time when activity of l -th measurement, where $l > k$, will be reached. In other words, we know activity A_l in time t_l and we want to “predict” t_l using data up to t_k .

For this experiment, 989 therapeutic records with number of measurements $m \geq 5$ were selected, so that data processed by least squares do not give meaningless estimate of T_{ef} . l was chosen as index of the last measurement in the sequence, k was chosen as index of at least third measurement after t_1 , at most second before k , depending on m .

Then time $t_l = t(A_l)$ was predicted both in the Bayesian way ($\hat{t}_l \pm \delta t_l$) and least-squares method (t_l^{LS}). Quantities $\Delta_r(t_l, \hat{t}_l)$ and $\Delta_r(t_l, t_l^{LS})$ were taken to compute arithmetic means and standard deviations (see Table 7.9) and plotted to the histograms (see Figure 7.6 and Figure 7.7).

| | Bayesian prediction | LS prediction |
|--------------------|--|---|
| arithmetic mean | $\overline{\Delta_r(t_l, \hat{t}_l)} = 4,81 \cdot 10^{-2}$ | $\overline{\Delta_r(t_l, t_l^{LS})} = 2,55 \cdot 10^{-3}$ |
| standard deviation | $s_B = 0,215$ | $s_{LS} = 0,132 \cdot 10^{-2}$ |

Table 7.9: Bayesian and least-squares relative differences of $t(A_l)$ prediction

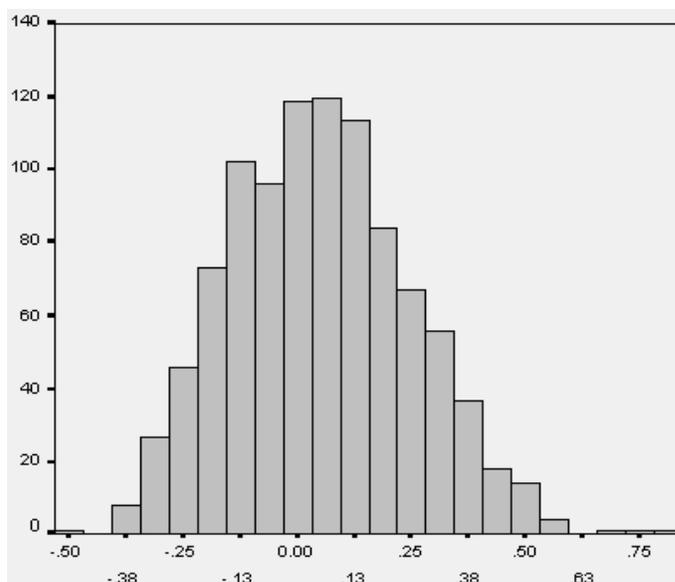


Figure 7.6: Histogram of $\Delta_r(t_l, \hat{t}_l)$ (see (7.1)), Bayesian prediction of $t(A_l)$, 989 records

It is visible that the mean of Bayesian $\Delta_r(t_l, \hat{t}_l)$ is more biased and dispersion is greater than in case of least-squares $\Delta_r(t_l, t_l^{LS})$. On the other hand, distribution of $\Delta_r(t_l, \hat{t}_l)$ seems to be closer to normal one than $\Delta_r(t_l, t_l^{LS})$. This result can again indicate insufficient model (6.19), page 65.

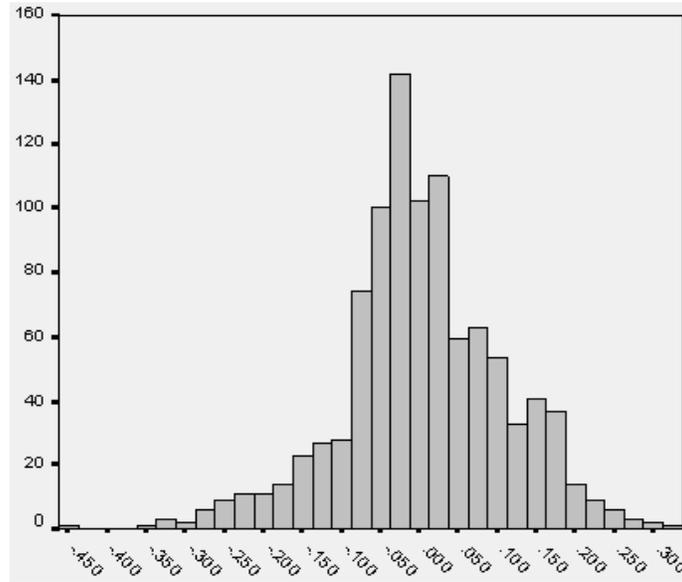


Figure 7.7: Histogram of $\Delta_r(t_l, t_l^{LS})$ (see (7.1)), least-squares prediction of $t(A_l)$, 989 records

7.2.4 Comparison of Bayesian and Deterministic Estimate of Excretions

Quantity E_{ri} called “excretion” is defined in the part 3.7, page 29.

From 3612 records, deterministic estimates of excretions E_{r2} and E_{r3} yielded 25 records (0,7% of records) where $E_{r2} + E_{r3} \leq 0$. No record where $E_{r2} + E_{r3} > 100\%$ was detected.

Using Bayesian estimates $\hat{E}_{r2} \pm \delta E_{r2}$ and $\hat{E}_{r3} \pm \delta E_{r3}$, quantities $\Delta_r(\hat{E}_{r2}, E_{r2})$, $\Delta_r(\hat{E}_{r2}, E_{r2})$, $\delta_r(\hat{E}_{r2}, \delta E_{r2})$ and $\delta_r(\hat{E}_{r3}, \delta E_{r3})$ (see (7.1)) were computed and processed in the Table 7.10.

| | arithmetic mean | standard deviation |
|---|----------------------|--------------------|
| \hat{E}_{r2} | 67,5 | 14,9 |
| \hat{E}_{r3} | 15,6 | 5,8 |
| $\delta_r(\hat{E}_{r2}, \delta E_{r2})$ | $5,6 \cdot 10^{-2}$ | 0,131 |
| $\delta_r(\hat{E}_{r3}, \delta E_{r3})$ | 0,201 | 0,210 |
| $\Delta_r(\hat{E}_{r2}, E_{r2})$ | $1,28 \cdot 10^{-2}$ | 0,285 |
| $\Delta_r(\hat{E}_{r2}, E_{r2})$ | $3,14 \cdot 10^{-2}$ | 0,405 |

Table 7.10: Comparison of deterministic and Bayesian estimate of excretions

Greater difference between Bayesian and deterministic estimate (over 3%) is for E_{r3} than for E_{r2} . At the same time, E_{r2} is estimated with lower relative uncertainty (cca. 6% in average) than E_{r3} (with almost 20%).

7.2.5 Influence of Data Quality on Thyroid Gland Mass Estimate

During data screening, it was found that thyroid gland masses measured by palpation differ in many cases significantly from those by sonography, if both the measurements are available for one thyroid gland.

Let us denote data h_{palp} as result of palpation measurement with relative precision ω_{palp} . Similarly, h_{sono} and ω_{sono} is data pair for sonographic measurement. Bayesian estimate is then $H = \hat{H} \pm \delta H$.

The differences between h_{palp} and h_{sono} are visible in the Figure 7.8. Also some characteristics

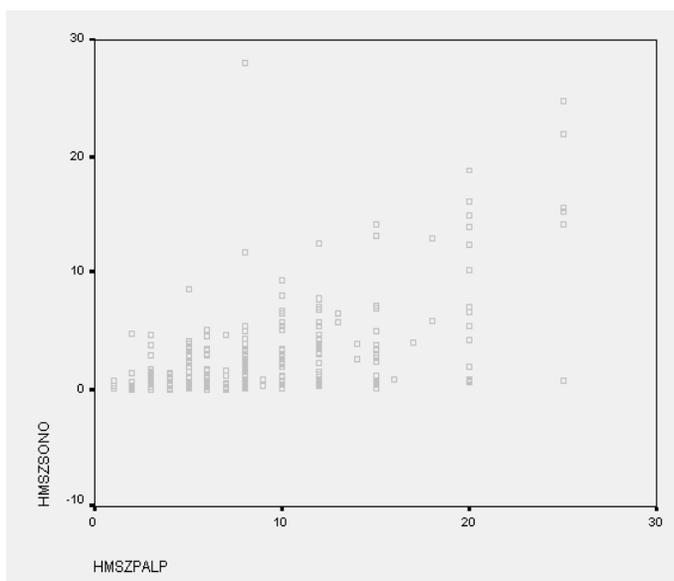


Figure 7.8: Scatterplot of h_{sono} vs. h_{palp} for thyroid gland mass measurement, 487 records

of $\Delta_r(h_{sono}, h_{palp})$ (see (7.1)) are shown in the Table 7.11.

| | | |
|--------------------|--|------|
| arithmetic mean | $\overline{\Delta_r(h_{sono}, h_{palp})} = 8,47$ | |
| standard deviation | $\sigma_{\Delta_r} = 16,31$ | |
| minimum value | $\Delta_{r,min}(h_{sono}, h_{palp}) = -0,71$ | |
| maximum value | $\Delta_{r,max}(h_{sono}, h_{palp}) = 213,29$ | (!!) |

Table 7.11: Relative difference of palpation and sonographic data $\Delta_r(h_{sono}, h_{palp})$ for thyroid gland mass estimation

We tried to consider influence of these differences on the estimate uncertainty. We chose 487 records where both palpation and sonographic measurement was finished.

We investigated dependence of $\delta_r(\hat{H}, \delta H)$ on $\Delta_r(h_{sono}, h_{palp})$ and other quantities. It was found, as property on logarithmic-normal *pdf*, that $\delta_r(\hat{H}, \delta H)$ is constant and does not depend on any data combination.

Histograms of quantities $\Delta_r(H, h_{palp})$ and $\Delta_r(H, h_{sono})$ are shown on the Figure 7.9 and Figure 7.10. The characteristics of these quantities are shown in the Table 7.12.

| | arithmetic mean | standard deviation |
|-------------------------|-----------------|--------------------|
| $\Delta_r(H, h_{palp})$ | 4,8 | 7,7 |
| $\Delta_r(H, h_{sono})$ | -0,233 | 0,14 |

Table 7.12: Arithmetic means and standard deviations of $\Delta_r(H, h_{palp})$ and $\Delta_r(H, h_{sono})$

From the figures and the tables, it is visible that relative difference between the Bayesian estimate and palpation measurement is much higher (up to 5-times in average with a very high dispersion) than in case of sonographic one (cca. 23%). It means that, according to (6.2), page 56, the lower ω is, the more it attracts a final estimate. In other words, sonographic measurement influences the final estimate much more than palpation one, if both the measurements are available.

Anyway, the methodology of the thyroid gland measurement should be revised to avoid so high average differences between palpation and sonographic data, as thyroid gland mass in a very

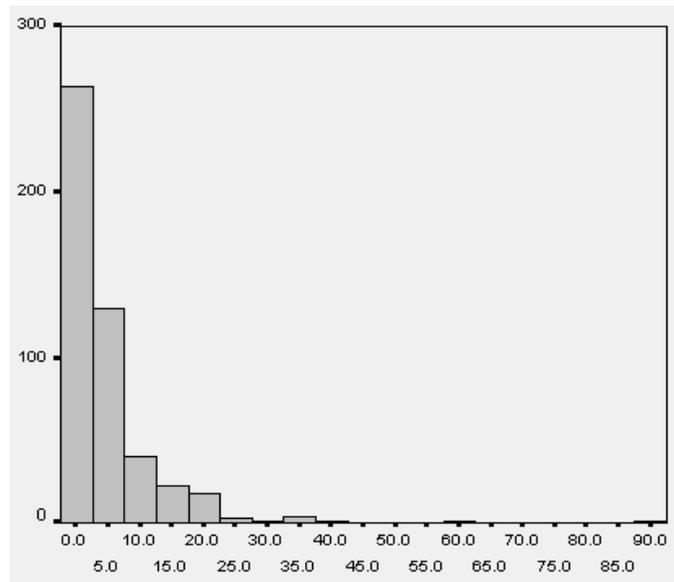


Figure 7.9: Histogram of $\Delta_r(H, h_{palp})$, 487 records

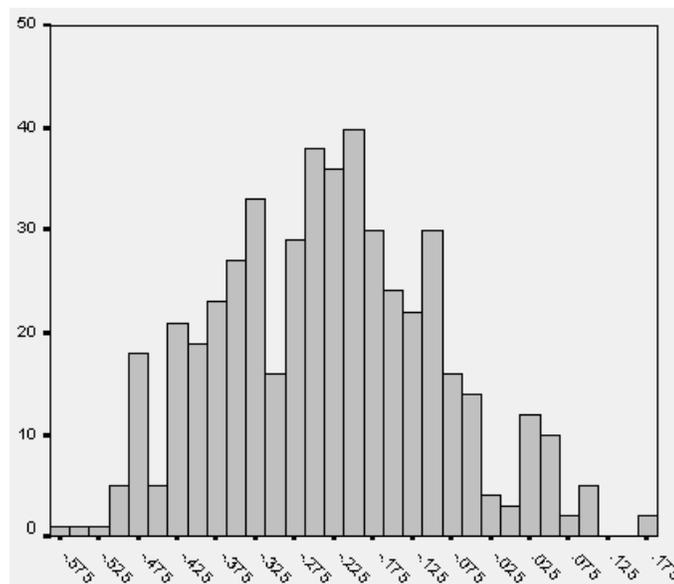


Figure 7.10: Histogram of $\Delta_r(H, h_{sono})$, 487 records

important quantity to estimate a absorbed dose (see 3.10, page 30). At least, correctness of values ω_{palp} and ω_{sono} should be tested.

7.3 Experiments with Individual Estimation Tasks

In this section, some single estimates on selected data are investigated. Influence of prior information is shown on the example of thyroid gland mass estimation. Further, quality of model for effective half-life estimation is tested by modifying a subset of data selected from two measurement sequences.

7.3.1 Influence of Prior Information on the Estimate

As derived in part (6.1.4), page 57, prior information for thyroid gland mass estimation is given by (6.3) and it was obtained from batch processing of sonographic data. The prior information represents fictitious data corresponding to $h=5,4\text{g}$ and $\omega=5,38$. In the Table 7.13, the estimates and deviations are shown for simulated data $h=2\text{ g}, 27\text{ g}, 100\text{ g}$ and 500 g . At first it is considered only palpation measurement, at second only sonographic measurement and at third combination of palpation and sonographic one with the same h . For the given *pdf*, intervals for probability 95% ($\alpha=0,05$) and 67% ($\alpha=0,23$) were found. It is visible that combination of data slightly decreases

| h | 95%-interval | | | 67%-interval | | |
|-----|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| | palpation | sonography | palp.+sono. | palpation | sonography | palp.+sono. |
| 2 | $2,74 \pm 1,96$ | $2,14 \pm 0,79$ | $2,11 \pm 0,72$ | $2,11 \pm 0,88$ | $2,02 \pm 0,38$ | $2,01 \pm 0,35$ |
| 27 | $31,5 \pm 22,5$ | $27,0 \pm 10,3$ | $27,8 \pm 9,5$ | $24,0 \pm 10,0$ | $26,5 \pm 5,0$ | $26,5 \pm 4,6$ |
| 100 | 107 ± 77 | 102 ± 37 | 102 ± 35 | 83 ± 35 | 97 ± 18 | 97 ± 17 |
| 500 | 487 ± 348 | 501 ± 184 | 501 ± 171 | 375 ± 157 | 474 ± 90 | 477 ± 83 |

Table 7.13: Estimation of thyroid gland mass with sonographic priors, $\hat{h}_0=0,998$, $\nu_0 = r/3,395$

uncertainty.

If we do batch processing of only palpation data, we find that $\bar{h}=16,7\text{ g}$ and $\bar{s}=32,18\text{ g}$. The prior is expressed by $\hat{h}_0=7,69$ and $\nu_0 = r/1,55$ which corresponds fictitious data $h=16,7\text{ g}$ and $\omega=1,93$. The results are shown in the Table 7.14. As the standard deviation is relatively less than

| h | 95%-interval | | | 67%-interval | | |
|-----|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| | palpation | sonography | palp.+sono. | palpation | sonography | palp.+sono. |
| 2 | $3,22 \pm 2,32$ | $2,22 \pm 0,81$ | $2,19 \pm 0,74$ | $2,59 \pm 1,05$ | $2,11 \pm 0,40$ | $2,09 \pm 0,36$ |
| 27 | $32,3 \pm 22,6$ | $28,1 \pm 10,3$ | $28,0 \pm 9,5$ | $25,2 \pm 10,2$ | $26,7 \pm 5,0$ | $26,7 \pm 4,6$ |
| 100 | 101 ± 71 | 101 ± 37 | 101 ± 34 | 79 ± 32 | 96 ± 18 | 96 ± 17 |
| 500 | 414 ± 290 | 485 ± 178 | 487 ± 166 | 324 ± 131 | 459 ± 86 | 465 ± 81 |

Table 7.14: Estimation of thyroid gland mass with palpation priors, $\hat{h}_0=7,69$, $\nu_0 = r/1,55$

in case of the sonographic priors, the prior information has stronger influence on the result. We can see that estimates of higher values are shifted down and estimates of lower values are shifted up. Again, with combination of data, the means of the intervals get closer to h and width of the interval decreases.

If we introduce no prior information ($\hat{h}_0=1$, $\nu_0=0$), which corresponds to $\omega \rightarrow +\infty$ and $h=\text{anything}$, we get results shown in the Table 7.15. The interval means are generally increased and procesing of more pieces of data both shifts the means to the data values and decreases the estimate uncertainty.

| h | 95%-interval | | | 67%-interval | | |
|-----|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| | palpation | sonography | palp.+sono. | palpation | sonography | palp.+sono. |
| 2 | $2,92 \pm 2,13$ | $2,15 \pm 0,80$ | $2,13 \pm 0,73$ | $2,22 \pm 0,95$ | $2,04 \pm 0,39$ | $2,03 \pm 0,36$ |
| 27 | $39,5 \pm 28,8$ | $29,0 \pm 10,8$ | $28,8 \pm 9,9$ | $29,9 \pm 12,8$ | $27,5 \pm 5,3$ | $27,4 \pm 4,8$ |
| 100 | 146 ± 107 | 108 ± 40 | 107 ± 37 | 111 ± 48 | 102 ± 19 | 101 ± 18 |
| 500 | 731 ± 533 | 538 ± 199 | 533 ± 183 | 554 ± 238 | 509 ± 97 | 507 ± 89 |

Table 7.15: Estimation of thyroid gland mass with no prior information, $\hat{h}_0=1$, $\nu_0=0$

In case of thyroid gland mass estimation, the prior information influences mostly the cases if only one measurement is available. The prior information causing the best match of data and mean of the confidence interval seems to be the one obtained by batch processing of sonographic data.

7.3.2 Influence of Data Amount on Effective Half-life Estimate

The model of effective half-life T_{ef} treats values of activities A_j measured in time instants t_j for $t_j > t_1, j = 1, \dots, m$, where m is number of measurements, according to (6.19), page 65. The question is how can m (i.e. number of measurements performed) influence the estimate of T_{ef} .

In practice, m is about 2 or 3 for diagnostic measurement, mostly over four for therapeutic measurement. Time gaps in therapeutic data sequences mean weekends, holidays etc.

We chose two sequences (red number 5138 applied 7.2.1999 and 3214 applied 19.12.1992) of therapeutic data with total number of measurements $n = 12$. We selected a sub-sequence of data with m members. We fixed t_1 as a measurement time where activity reaches its maximum and changed a number of subsequent data m . The Table 7.16 shows means of intervals \hat{T}_{ef} and deviations δT_{ef} .

| m | sequence 1 | | sequence 2 | |
|-----|----------------|-----------------|----------------|-----------------|
| | \hat{T}_{ef} | δT_{ef} | \hat{T}_{ef} | δT_{ef} |
| 1 | 4,0380 | 3,8063 | 4,0376 | 3,8030 |
| 2 | 1,5069 | 0,0047 | 4,9750 | 0,0153 |
| 3 | 6,5893 | 0,0394 | 3,4135 | 0,0055 |
| 4 | 3,9261 | 0,0069 | 2,7098 | 0,0028 |
| 5 | 2,9426 | 0,0031 | 2,3553 | 0,0018 |
| 6 | 2,4882 | 0,0019 | 2,2432 | 0,0015 |
| 7 | 2,2607 | 0,0016 | 2,1859 | 0,0013 |
| 8 | 2,1328 | 0,0015 | 2,1509 | 0,0012 |
| 9 | 2,0802 | 0,0014 | 2,1231 | 0,0011 |
| 10 | 2,0425 | 0,0013 | 2,1047 | 0,0011 |
| 11 | 2,0216 | 0,0012 | 2,0970 | 0,0010 |
| 12 | 2,0079 | 0,0012 | 2,0929 | 0,0010 |

Table 7.16: Dependence of T_{ef} on number of data m

It is visible that estimate of \hat{T}_{ef} in both cases approaches some value with increasing m . What is surprising is a very low value of δT_{ef} in all the cases except of $m = 1$.

7.3.3 Influence of t_1 on Effective Half-life Estimate

The model of effective half-life T_{ef} treats values of activities A_j measured in time instants t_j for $t_j > t_1$ according to (6.19), page 65. The time t_1 represents a time from which the activity course

$A(t)$ can be treated as exponential. The question is, how can a choice of t_1 influence the estimate of T_{ef} .

In practice, t_1 is chosen as a measurement time when the activity reaches its highest value. We check the validity of this decision and influence of this choice on the T_{ef} estimate.

The same data sequences as in the part 7.3.2 are chosen.

At first, m was set to 8 and t_1 was changed from the first to the fifth measurement. In this way, five sub-sequences were obtained and used to compute estimate of T_{ef} . The means of intervals \hat{T}_{ef} and deviations δT_{ef} are shown in the Table 7.17. Symbol n_1 denotes (absolute) index of the measurement taken (relative) as the first one (i.e. $n_1 = 3$ means 3rd measurement in the sequence is taken as the first one in this computation, 4th one as the second one etc.) Symbol $t_1(n_1)$ denotes time t_1 corresponding to the n_1 -th measurement taken as the first one.

| n_1 | sequence 1 | | | sequence 2 | | |
|-------|------------|----------------|-----------------|------------|----------------|-----------------|
| | $t_1(n_1)$ | \hat{T}_{ef} | δT_{ef} | $t_1(n_1)$ | \hat{T}_{ef} | δT_{ef} |
| 1 | 1,954 | 2,1328 | 0,0015 | 0,815 | 2,1509 | 0,0012 |
| 2 | 2,902 | 1,9899 | 0,0014 | 3,800 | 1,4721 | 0,0013 |
| 3 | 3,920 | 1,5476 | 0,0011 | 4,807 | 1,5689 | 0,0016 |
| 4 | 6,913 | 0,8643 | 0,0011 | 5,837 | 1,7001 | 0,0021 |
| 5 | 7,889 | 0,9264 | 0,0021 | 10,811 | 2,2102 | 0,0084 |

Table 7.17: Dependence of T_{ef} on t_1 for $m = 8$

Another experiment was done with both m and t_1 variable. The measurements were cut off from the beginning and the computations were done on the data that “remained”, i.e. $n_1 + m = n - 1$. The results for both the data sequences are shown in the Table 7.18.

| n_1 | m | sequence 1 | | | sequence 2 | | |
|-------|-----|------------|----------------|-----------------|------------|----------------|-----------------|
| | | $t_1(n_1)$ | \hat{T}_{ef} | δT_{ef} | $t_1(n_1)$ | \hat{T}_{ef} | δT_{ef} |
| 1 | 12 | 1,954 | 2,0079 | 0,0012 | 0,815 | 2,0929 | 0,0010 |
| 2 | 11 | 2,902 | 1,8869 | 0,0010 | 3,800 | 1,5113 | 0,0012 |
| 3 | 10 | 3,920 | 1,5181 | 0,0009 | 4,807 | 1,6070 | 0,0016 |
| 4 | 9 | 6,913 | 0,9974 | 0,0011 | 5,835 | 1,7221 | 0,0021 |
| 5 | 8 | 7,889 | 1,1921 | 0,0019 | 10,811 | 2,2099 | 0,0084 |
| 6 | 7 | 8,911 | 1,4893 | 0,0034 | 11,864 | 1,8713 | 0,0079 |
| 7 | 6 | 9,880 | 1,6970 | 0,0054 | 12,810 | 1,8196 | 0,0095 |
| 8 | 5 | 10,916 | 1,8613 | 0,0080 | 13,809 | 1,9277 | 0,0137 |
| 9 | 4 | 13,887 | 2,4612 | 0,0366 | 15,860 | 7,9947 | 0,0447 |
| 10 | 3 | 14,901 | 7,8592 | 0,1777 | 16,855 | 7,9902 | 0,0491 |
| 11 | 2 | 15,947 | 2,8026 | 0,1912 | 17,894 | 4,1365 | 0,4680 |
| 12 | 1 | 16,891 | 4,0200 | 3,8190 | 18,849 | 4,0200 | 3,8190 |

Table 7.18: Dependence of T_{ef} on t_1 for variable m

It is obvious that estimates of T_{ef} strongly depend on the subset of chosen data and deviations are too small to cover this dispersion. This observation can indicate insufficiency of the model (6.19), page 65.

Chapter 8

Conclusions

Application of the Bayesian methodology in nuclear medicine for solution of the estimation problems mentioned above and other ones seems to be promising both theoretically and practically.

At the practical level, the specific problems at the specific Clinic of nuclear medicine (connected to optimization of radiation load on patients, fulfilling more strict radiohygienic regulation measures, more economical exploitation of radiopharmaceuticals etc.) are solved.

At the theoretical and algorithmic level, problems related to applicability of Bayesian methodology, especially numerical ones, are addressed.

Within this effort, several results have been achieved. Among them, the most important are:

- The Bayesian estimates of specific biophysical quantities were derived, as summarized in Chapter 6.
- The estimates were generally and safely implemented and computer programs for Bayesian estimations were created.
- The software system JodNew that ensures complete data management of patients and contains implemented Bayesian estimates runs at the *KNM*.

The work described in this thesis is a part of effort to contribute to improvement of treatment by better data processing using sophisticated mathematical methods.

The estimates and algorithms are designed for the *KNM* but they can be used in other clinics or related fields, e.g. dosimetry, radiation protection etc.

8.1 The Contributions

The main fields to which this work contributed are as follows:

Application of the Bayesian theory in a new field

Successful application of formulae derived using some theory can be a task for itself. In this work, focus was mainly put on practical implementation of theoretical results with stress on numerical precision and stability.

1. Solutions of important numerical and programming tasks are derived in Chapter 5.
 - Construction of symmetric confidence interval had the important subtasks to derive criteria for variable integration step, confidence interval shift and its correction and termination of the step-by-step numerical integration. The task was to find the confidence interval limits with the given numerical precision.

- Design of integration table for uni-variate function with floating distance of its points is focused on determining such a distance between adjacent points, according to the function curvature, to keep the given precision of quadratic interpolation between the table points.
 - Other numerical tricks and methods are mentioned there, e.g. pseudo-normalization of non-normalized function, half-width searching, domain restriction, testing of simplex method, modification of its procedure and termination test.
 - Programming approach is described together with hierarchy of numerical classes and with other methods of solution.
2. The estimation formulae, derived in [11], were modified and extended in Chapter 6. Other contribution is in implementation remarks to each estimation task where specific features of practical realization are mentioned.
 3. The extensive tests of the algorithms are performed in Chapter 7 with the following results:
 - The algorithms are numerically stable, run on cca. 5 000 patient data records passed successfully.
 - The algorithms are numerically precise within requested accuracy.
 - Run times of the estimation programs are within fractions of seconds except of signal estimation. Therefore some simplification based on Gaussian approximation was adopted for this task that increased the computation speed 50-times.
 4. Experiments with batch processing of data and individual estimation tasks are also performed in Chapter 7. The main conclusions are:
 - Bayesian estimates are always physically meaningful compared to the deterministic ones. For example, in the investigated data sets, almost 10% of deterministic T_{ef} estimates and up to 1% of deterministic excretions estimates were physically meaningless.
 - Bayesian predictions of radio-hygienic limits yield worse results (but always meaningful) than deterministic ones. The reason can be in insufficient model of activity course in time.
 - Uncertainty of estimates depends on data uncertainty and other information used for the estimation. Results of thyroid gland mass estimation show that in case of small amount of uncertain data, prior information becomes more significant for values of the estimates.
 - Estimation of T_{ef} depends strongly on starting time t_1 and amount of data and uncertainty of the estimates are too low which is not realistic. Again, this observation can indicate insufficiency of the activity time course model.

Use of results of the Bayesian theory in practice — JodNew

The software system JodNew is used for everyday work at the *KNM*. Except of computing the Bayesian estimates for medical purposes, it completely manages patients' data including administrative ones, includes routines for preparation of radioactive solutions, supports technical dosimetric operations (calibrations of devices, measurement of standard sources...), prints reports for the physicians, updates data archives and saves data obtained during various biophysical measurements for "scientific purposes" — further processing and analyses, hopefully performed in future.

The Bayesian estimates as inputs for eventual retrospective studies for finer individualization of therapeutic activity

Estimated parameters have physical meaning and estimate methodology should respect it. Bayesian estimates on biophysical data are designed so that they are always meaningful *physically*. There

was detected an insufficiency in modelling of $A(t)$ that may cause *medical* meaninglessness of some estimates. The model improvement is already finished [13] and its algorithmical implementation should be done soon.

Results of the retrospective studies can indicate ways to refine individual therapeutic activity. However, this topic is beyond scope of this thesis.

Observations of data quality

In this context, it was observed that palpation and sonographic data in the task of thyroid gland mass differ in many cases significantly, even by 2–3 orders. This observation has brought a message to the *KNM* to check the measurement methodology and precision.

This question is very important as thyroid gland mass H is a quantity that enters formulae for estimate of dose absorbed by thyroid gland, which is a key quantity to determine a therapeutic activity. H appears both in specific irradiation estimation (see part 6.9.1, page 69) and MIRD method (see part 6.9.3, page 71). As H usually reaches low values and it is in denominator, the dose estimates are very sensitive to thyroid gland mass uncertainty.

8.2 Open Problems

Results of this thesis indicate that application of Bayesian theory in solution of the estimation tasks in nuclear medicine described here seems to be successful both in theoretical and practical level. In order to take more advantage of this approach, the following open problems should be addressed:

- algorithmical implementing of the alternative model for estimation of T_{ef} and predictions of radio-hygienic limits [13] and testing existence of differences in ^{131}I kinetics during diagnostics and therapy,
- approximating of the posterior *pdf* in the task of signal on background estimation for better further theoretical applicability and practical performance,
- looking for optimization methods with proved reliability and convergence properties for the given class of tasks and substituting them for the simplex method,
- testing of properness of Gaussian approximation for deriving of integration criteria for other *pdfs* treated here,
- deriving Bayesian estimates for the MIRD method as it is the best tool for doses estimation and contains independent input quantities (A_{apl} , τ and \mathcal{S}) in contrast to SI_{th} and SIK formulae,
- designing a retrospective study trying to discover influence of biophysical data and estimates on success of therapy.

Bibliography

- [1] Milton Abramowitz and Irene A. Stegun. *Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables*. Dover Publications, Inc., New York, 1970.
- [2] Jiří Anděl. *Mathematical Statistics*. SNTL/ALFA, 1st edition, 1978. In Czech.
- [3] J. Böhm and M. Kárný. Transformation of user's knowledge into initial values for identification. In M. Součková and J. Böhm, editors, *Preprints of the DYCOMANS workshop Industrial control and management methods: theory and practice*, pages 17–24. ÚTIA AV ČR, Prague, 1995.
- [4] H. J. Dworkin, D. A. Meier, and M. Kaplan. Advances in the management of patients with thyroid diseases. *Seminars in Nuclear Medicine*, 25(3):205–220, 1995.
- [5] Michael Evans and Tim Swartz. Methods for approximating integrals in statistics with special emphasis on Bayesian integration problems. *Statistical Science*, 10(3):254–272, 1995.
- [6] C. M. Fonseca. Bayesian estimation of the intensity of low-level radiation sources. *Jaderná energie*, 37:83–97, 1991.
- [7] G. E. Forsythe, M. A. Malcolm, and C. B. Moler. *Computer Methods for Mathematical Computations*. Prentice Hall, 1977. Russian translation, Mir, Moscow 1980.
- [8] Eli E. Furhang, Steven M. Larson, Puangrat Buranapong, and John L. Humm. Thyroid cancer dosimetry using clearance fitting. *Journal of Nuclear Medicine*, 40(1), 1999.
- [9] P. Georgi, D. Emrich, P. Heidenreich, E. Moser, C. Reiners, and H. Schicha. Radiojodtherapie des differenzierten Schilddrüsenkarzinoms. *Journal of Nuclear Medicine*, 31:151–153, 1992. In German.
- [10] T.V. Guy. Multivariate integration problem in the nuclear medicine. Technical Report 1838, ÚTIA AV ČR, P.O. Box 18, 182 08 Prague 8, Czech Republic, 1995.
- [11] J. Heřmanská. *Bayesian Approach to Dosimetric Data Evaluation for Medical Use of ¹³¹I*. Clinic of Nuclear Medicine, 2nd Medical Faculty, Charles University, Prague, 1993. Associated Professor Thesis, 103 pp. In Czech.
- [12] J. Heřmanská, T. Blažek, and M. Kárný. Beyond effective half-life characterization of radionuclide dynamics. In *Proceedings of the 18th Radiation Hygiene Days*, pages 78–85. Jáchymov, 1994.
- [13] J. Heřmanská and M. Kárný. Bayesian estimation of effective half-life in dosimetric applications. *Computational Statistics and Data Analysis*, 24(5):467–482, 1997.
- [14] J. Heřmanská, M. Kárný, T.V. Guy, L. Jirsa, T. Blažek, and J. Němec. Biokinetics of ¹³¹I in human organism. In *Proceedings of the 19th Radiation Hygiene Days*, pages 159–164. Jasná pod Chopkom, SR, 1995.

- [15] L. Jirsa, J. Heřmanská, and M. Kárný. Bayesian estimation of dosimetric quantities for low signal-to-noise ratio. In *Abstracts of the XVIIth Biophysical Days*. Hradec Králové, 1994. In Czech, non-numbered pages.
- [16] L. Jirsa, J. Heřmanská, and M. Kárný. Towards the feasible probabilistic estimation of signal on background. In P. Kneppo and M. Tyšler, editors, *Proc. of the 7th International IMEKO TC-13 Conference on Measurement in Clinical Medicine: Model Based Biomeasurements*, pages 149–151. Stará Lesná, SR, 1995.
- [17] M. Kárný, A. Halousková, and L. Zörnigová. On pooling expert opinions. In M. Blanke and T. Söderström, editors, *Preprints of the 10th IFAC Symposium on System Identification SYSID'94*, volume 2, pages 477–478,485–488. Danish Automation Society, Copenhagen, 1994.
- [18] M. Kárný and P. Nedoma. Automatic processing of prior information with application to identification of regression model. *ACASP*, 1998. Submitted.
- [19] Natalia Khailova. Bayes paradigm in practice — examples. Research report No.1901, Institute of Information Theory and Automation, Academy of Sciences, Czech Republic, ÚTIA AV ČR, POB 18, 182 08, Prague 8, CZ, June 1997.
- [20] Vladik Kreinovich, Anatoly Lakayev, Jiří Rohn, and Patrick Kahl. *Computational Complexity and Feasibility of Data Processing and Interval Computations*. Kluwer Academic Publishers, 1st edition, 1998.
- [21] L. D. Marinelli, E. H. Quimby, and G. J. Hine. Dosage determination with radioactive isotopes; II. Practical considerations in therapy and protection. *American Journal of Roentgenology and Radiotherapy*, 59:260–280, 1948.
- [22] W. V. Mayencord. *British Journal of Radiology*, Supplement 2, 1950.
- [23] Ken I. M. McKinnon. Convergence of the Nelder-Mead simplex method to a non-stationary point. *SIAM Journal of Optimization*, 1997.
- [24] J. C. Naylor and A. F. M. Smith. Econometric illustrations of novel numerical integration strategies for Bayesian inference. *Journal of Econometrics*, 38:103–125, 1988.
- [25] Petr Nedoma, Miroslav Kárný, and Josef Böhm. ABET97: Adaptive Bayesian estimation toolbox for Matlab. Technical report 1914, Institute of Information Theory and Automation, Academy of Sciences, Czech Republic, ÚTIA AV ČR, POB 18, 182 08, Prague 8, CZ, 1997.
- [26] J. A. Nelder and R. Mead. A simplex method for function minimization. *Comp. J.*, 7:308, 1965.
- [27] Jan Němec. *Treatment of Thyroid Diseases by Radioiodine*. Avicenum, Prague, 1974. In Czech.
- [28] Jan Němec. *personal message*. Prague, 1999. In Czech.
- [29] Václav Peterka. Bayesian system identification. In P. Eykhoff, editor, *Trends and Progress in System Identification*, pages 239–304. Pergamon Press, Oxford, 1981.
- [30] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. *Numerical Recipes in C*. Cambridge University Press, 2nd edition, 1992.
- [31] Anthony Ralston. *A First Course in Numerical Analysis*. McGraw-Hill, Inc., 1965. Czech translation, Academia, Prague, 1973.
- [32] J. Rojíček, M. Valečková, M. Kárný, and K. Warwick, editors. *Computer-Intensive Methods in Control and Data Processing. Can We Beat the Curse of Dimensionality? 3rd European IEEE Workshop CMP'98, Prague, CZ, ÚTIA AV ČR, POB 18, 182 08, Prague 8, CZ, September 7–9 1998*.

- [33] Aleksandr Semyonovich Rykov. *Search Optimization. Methods of Deforming Configurations*. Nauka, Moscow, 1st edition, 1993. In Russian.
- [34] M. Samuel and B. Rajashekharro. Radioiodine therapy for well-differentiated thyroid cancer: a quantitative dosimetric evaluation for remnant thyroid ablation. *Journal of Nuclear Medicine*, 35:1944–1950, 1994.
- [35] Geoffrey A. Siegel, Stephen R. Thomas, James B. Stubbs, Michael G. Stabin, and Margerite T. Hays. MIRD pamphlet No.16, Technics for quantitative radiopharmaceutical bio-distribution data aquisition and analysis. *Journal of Nuclear Medicine*, 40, suppl(2), February 1999.
- [36] K. Šilink and S. Vohnout. Use of radio-iodine in diagnostics and therapy of thyroid diseases. In J. Petrová, editor, *Radioisotopes in biology and medicine*, pages 411–462. Státní zdravotnické nakladatelství, Prague, 1960. In Czech.
- [37] W. Snyder, M. Ford, G. Warner, and S. Watson. MIRD pamphlet No.11, S-absorbed dose per unit cumulated activity for selectd radionuclides and organs. Technical report, New York, 1975.
- [38] M. G. Stabin. MIRDose: Personal computer software for internal dose assesment in nuclear medicine. *Journal of Nuclear Medicine*, 37(3):538–546, 1996.
- [39] Luke Tierny, Robert E. Kass, and Joseph B. Kadane. Approximate marginal densities of nonlinear functions. *Biometrika*, 76(3):425–433, 1989.
- [40] P. Tseng. Fortified-descent simplicial search method: a general approach. *Draft for publication*, 1996.
- [41] M. van Wyngarden and I. R. McDougall. What is the role of 1 100 MBq (<3 mCi) radioiodine ^{131}I in the treatment of patients with differentiated thyroid cancer. *Nuclear Medicine Communication*, 17:199–207, 1996.