Nonparametric estimation of blood alcohol concentration from transdermal alcohol measurements using alcohol biosensor devices

Bryan Vader

In collaboration with:

Dr. Alona Kryshchenko, CSU Channel Islands
&
Dr. Melike Sirlanci, CALTECH

Channel Islands

CALIFORNIA STATE UNIVERSITY

Symposium on Data Science & Statistics

June 4, 2020

Table of Contents

Table of Contents

- Used primarily for monitoring sobriety.

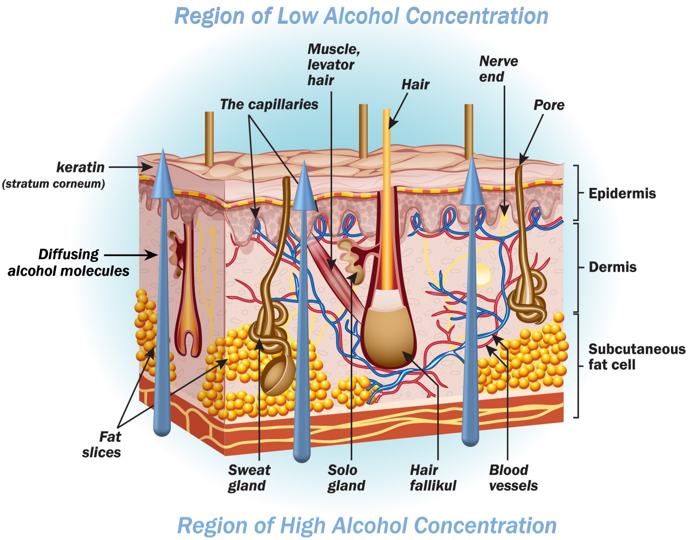
SCRAM CAM

- Used primarily for monitoring sobriety.
- Recently NIAAA called for research to construct alcohol biosensor devices which are more accurate, discreet, precise, and tamper proof.

SCRAM CAM

BACtrack Skyn

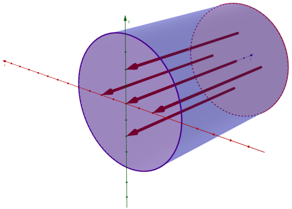
Table of Contents


Region of Low Alcohol Concentration
Muscle, levator hair
Nerve end
Hair
The capillaries
Pore
keratin
(stratum corneum)
Epidermis
Diffusing alcohol molecules
Dermis
Subcutaneous fat cell
Fat slices
Sweat gland
Solo gland
Hair fallikul
Blood vessels
Region of High Alcohol Concentration

The Heat Equation

Definition of Variables

$u(\eta, t)$	- concentration of alcohol in moles/cm^2
$y(\eta, t)$	- concentration of alcohol in moles/cm^2 including biological variation $\varepsilon \sim N(0, \sigma)$
η	- depth into dermal layer (normalized as percentage of depth)
t	- time in hours
q	- parameters of diffusivity vector in cm^2/hour
$\alpha(t)$	- input of alcohol into the system at the bottom of the dermal layer ($\eta = 1$) in BAC/BrAC units



The Heat Equation: Two Parameters

Heat Equation in Strong Formulation

$$\frac{\partial u}{\partial t}(\eta, t) = q_1 \frac{\partial^2 u}{\partial \eta^2}(\eta, t), \qquad 0 < \eta < 1, t > 0 \qquad (1)$$

$$u(\eta, 0) = 0, \qquad 0 < \eta < 1 \qquad (2)$$

$$u(0, t) = q_1 \frac{\partial u}{\partial \eta}(0, t), \qquad t > 0 \qquad (3)$$

$$\frac{q_1}{q_2} \frac{\partial u}{\partial \eta}(1, t) = \alpha(t), \qquad t > 0 \qquad (4)$$

$$y(\eta, t) = c \cdot u(\eta, t) + \varepsilon, \qquad \varepsilon \sim N(0, \sigma) \qquad (5)$$

$$q = \langle q_1, q_2 \rangle \sim F \qquad (6)$$

Problem

- Need $u(\eta, t)$ for various depths and points in time, but solving this analytically is likely very difficult or impossible.

Problem

- Need $u(\eta, t)$ for various depths and points in time, but solving this analytically is likely very difficult or impossible.
- Note: we sample from from two distributions to provide values for q at this stage, however much of the remainder of the project is concerned with estimating an appropriate value.

Discretizing the Heat Equation

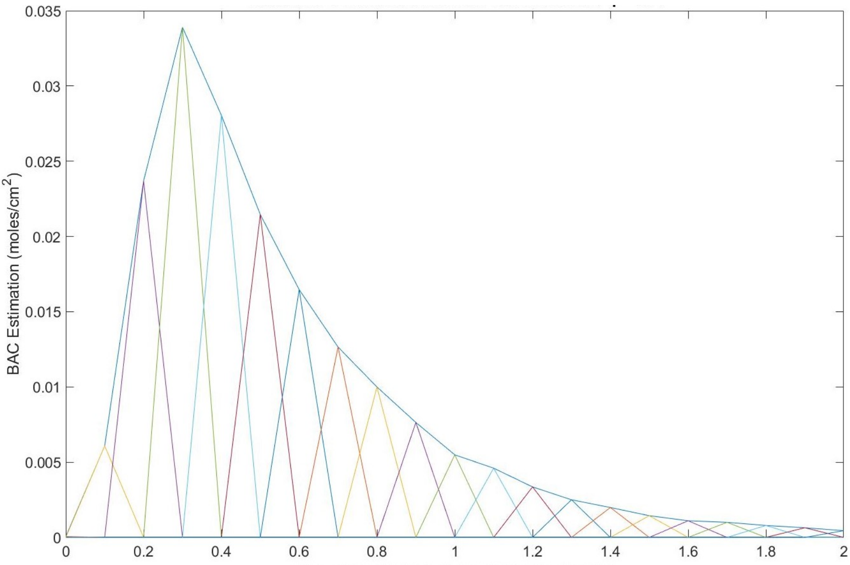
Problem

- Need $u(\eta, t)$ for various depths and points in time, but solving this analytically is likely very difficult or impossible.
- Note: we sample from from two distributions to provide values for q at this stage, however much of the remainder of the project is concerned with estimating an appropriate value.

Plan to Solve this Problem

- Use the **Finite Element Method** with **B-Splines** to **discretize** η and t and view $u(\eta, t)$ as a matrix of concentration values.

B-Splines



BAC Estimation (moles/cm^2)
0.035
0.03
0.025
0.02
0.015
0.01
0.005
0
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
2

The Final Discretized Model

Recursion With Respect to Time

Note: this vector is now only dependent on the parameter q.

$$\hat{u}^N(t_j; q) = \sum_{k=1}^{j} \left(\alpha_{k-1} \left(\hat{A}(q) \right)^{j-k} \right) \hat{B}(q)$$

Recursion With Respect to Time

Note: this vector is now only dependent on the parameter q.

$$\hat{u}^N(t_j; q) = \sum_{k=1}^{j} \left(\alpha_{k-1} \left(\hat{A}(q) \right)^{j-k} \right) \hat{B}(q)$$

The Discretized Matrix Form

Note: this matrix is now only dependent on the parameter q.

$$U^{N,T}(q) = \left[\vec{0}, \left\{ \sum_{k=1}^{j} \left(\alpha_{k-1} \left(\hat{A}(q) \right)^{j-k} \right) \hat{B}(q) \right\}_{j=1}^{T} \right]$$

The Discretized Form at $\eta = 0$ (TAC) Depth

$$y(q) = \hat{c}(q) \cdot U^{N,T}(q)$$

TAC With Measurement Error

The Discretized Form at $\eta = 0$ (TAC) Depth

$$y(q) = \hat{c}(q) \cdot U^{N,T}(q)$$

TAC with Measurement Error

$$y_\epsilon(q) = y(q) + \epsilon$$
$$\epsilon \sim N(0, \sigma)$$

Pairs of Probability Density Functions

Pair 1: $q_1 \sim \mathcal{N}(1, \sigma^2)$
$q_2 \sim \mathcal{N}(1, \sigma^2)$

Simulated TAC using Matlab

Pairs of Probability Density Functions

Pair 1:

$$q_1 \sim \mathcal{N}(1, \sigma^2)$$
$$q_2 \sim \mathcal{N}(1, \sigma^2)$$

Pair 2:

$$q_1 \sim 0.7\mathcal{N}(0.7, \sigma^2) + 0.3\mathcal{N}(1, \sigma^2)$$
$$q_2 \sim \mathcal{N}(1, \sigma^2)$$

Pairs of Probability Density Functions

Pair 1:

$$q_1 \sim \mathcal{N}(1, \sigma^2)$$
$$q_2 \sim \mathcal{N}(1, \sigma^2)$$

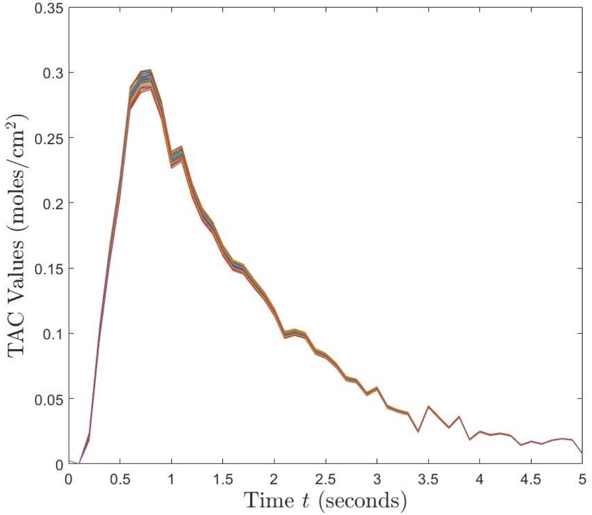
Pair 2:

$$q_1 \sim 0.7\mathcal{N}(0.7, \sigma^2) + 0.3\mathcal{N}(1, \sigma^2)$$
$$q_2 \sim \mathcal{N}(1, \sigma^2)$$

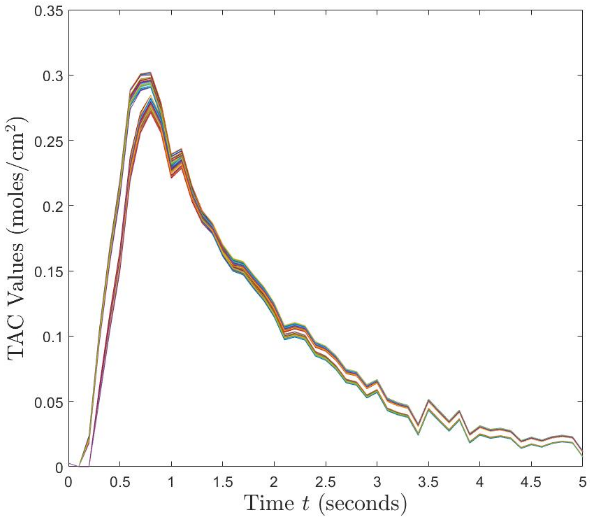
Pair 3:

$$q_1 \sim 0.7\mathcal{N}(0.7, \sigma^2) + 0.3\mathcal{N}(1, \sigma^2)$$
$$q_2 \sim 0.3\mathcal{N}(0.7, \sigma^2) + 0.7\mathcal{N}(1, \sigma^2)$$

Pair 1: Simulated TAC using Matlab



Pair 2: Simulated TAC using Matlab



Pair 3: Simulated TAC using Matlab

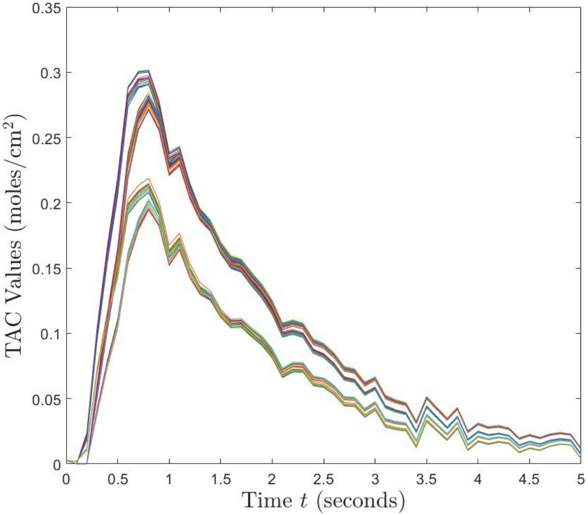


Table of Contents

Likelihood Function

Let $Y_1, Y_2, \dots, Y_P$ be independent measurements that have a joint distribution $p(Y_1, Y_2, \dots, Y_P | q)$. The parameters are unknown random variables distributed with distribution F which is also unknown. For a given F parameter the **log likelihood function** is

$$\begin{aligned}\ell(F) = \log(p(Y_1, Y_2, \dots Y_P | F)) &= \log\left(\prod_{i=1}^{P} p_i(Y_i | F)\right) \\ &= \sum_{i=1}^{P} \log(p_i(Y_i | F)) \\ &= \sum_{i=1}^{P} \log\left(\int_{\Omega} p_i(Y_i | q)\, \mathrm{d}F(q)\right).\end{aligned}$$

Maximum Likelihood Estimation

Observe that when

$$L(F^*) \geq L(F)$$

for all $F \in \mathscr{F}$, then the **maximum likelihood estimator** F^{ML} is more likely to be F^* than any other F.

Maximum Likelihood Estimation

Observe that when

$$L(F^*) \geq L(F)$$

for all $F \in \mathscr{F}$, then the **maximum likelihood estimator** F^{ML} is more likely to be F^* than any other F.

Maximum Likelihood Estimator

The **maximum likelihood estimator** of F is

$$F^{ML} = \text{argmax}\{\ell(F) : F \in \mathscr{F}\}$$

where $\mathscr{F}$ is the family of all probability distributions on Ω.

Theorem: (Lindsay, 85)

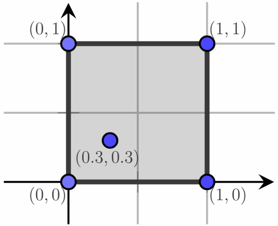
If Ω is compact, F^{ML} can be found in the class of discrete distributions on Ω with at most P support points.

Theorem: (Lindsay, 85)

If Ω is compact, F^{ML} can be found in the class of discrete distributions on Ω with at most P support points.

Carathéodory's Theorem: (Roberts, 73)

If $U \subseteq \mathbb{R}^n$ and its convex hull $H(U)$ has dimension m, then for each $z \in H(U)$, there exists $m+1$ points $x_0, x_1, \ldots, x_m$ of U such that z is a convex combination of these points.



Expectation-Maximization Optimization Problem

Maximum Likelihood Estimator

$$F^{ML} = \operatorname{argmax}\left\{\sum_{i=1}^{P}\log\left(\sum_{k=1}^{K} w_k p_i(Y_i \mid \xi_k)\right), K \leq P\right\}$$

$$w_k = \mathbb{P}(q = \xi_k) \geq 0, \text{ for all } 1 \leq k \leq K,$$

$$\sum_{k=1}^{K} w_k = 1$$

What is Needed For the Optimization Problem?

Let

$$\lambda = (\xi^K, w^K) = (\xi_1, \xi_2, \ldots, \xi_K, w_1, w_2, \ldots, w_K)$$

be the vector of weights and supports. It is now appropriate to state the question as

$$\ell(F) \approx \ell(\lambda) = \ell(\xi_1, \xi_2, \ldots, \xi_K, w_1, w_2, \ldots, w_K).$$

Table of Contents

NPAG Algorithm

Starting conditions: $\mathbf{Y}_\epsilon$, ξ^0, $\mathbf{a}$,$\mathbf{b}$,$\mathbf{\Delta_D}$

Output: $\xi, \mathbf{w}, \ell(\mathbf{w}, \xi)$

NPAG Algorithm

Starting conditions: $\mathbf{Y}_\epsilon$, ξ^0, $\mathbf{a}$,$\mathbf{b}$,$\mathbf{\Delta_D}$

Output: $\xi, \mathbf{w}, \ell(\mathbf{w}, \xi)$

Step 1: Initialize the grid of support points ξ^0,

NPAG Algorithm

Starting conditions: $\mathbf{Y}_{\epsilon}$, ξ^0, $\mathbf{a}$,$\mathbf{b}$,$\boldsymbol{\Delta}_{\mathbf{D}}$

Output: $\xi, \mathbf{w}, \ell(\mathbf{w}, \xi)$

Step 1: Initialize the grid of support points ξ^0,

Step 2: Calculate the Ψ matrix of probabilities and use the PDIP method to calculate each of the support point weights,

NPAG Algorithm

Starting conditions: $\mathbf{Y}_{\epsilon}$, ξ^0, $\mathbf{a}$,$\mathbf{b}$,$\mathbf{\Delta_D}$
Output: $\xi, \mathbf{w}, \ell(\mathbf{w}, \xi)$

Step 1: Initialize the grid of support points ξ^0,

Step 2: Calculate the Ψ matrix of probabilities and use the PDIP method to calculate each of the support point weights,

Step 3: Condense the grid by removing support points with sufficiently low probability,

NPAG Algorithm

Starting conditions: $\mathbf{Y}_\epsilon$, ξ^0, $\mathbf{a}$,$\mathbf{b}$,$\mathbf{\Delta_D}$

Output: $\xi, \mathbf{w}, \ell(\mathbf{w}, \xi)$

Step 1: Initialize the grid of support points ξ^0,

Step 2: Calculate the Ψ matrix of probabilities and use the PDIP method to calculate each of the support point weights,

Step 3: Condense the grid by removing support points with sufficiently low probability,

Step 4: Check the exit condition for the log-likelihood difference, i.e. $\left|\ell(\lambda^{(n+1)}) - \ell(\lambda^{(n)})\right| \leq 10^{-4}$, and exit if true or else proceed to Step 5,

NPAG Algorithm

Starting conditions: $\mathbf{Y}_\epsilon$, ξ^0, $\mathbf{a}$,$\mathbf{b}$,$\mathbf{\Delta_D}$

Output: $\xi, \mathbf{w}, \ell(\mathbf{w}, \xi)$

Step 1: Initialize the grid of support points ξ^0,

Step 2: Calculate the Ψ matrix of probabilities and use the PDIP method to calculate each of the support point weights,

Step 3: Condense the grid by removing support points with sufficiently low probability,

Step 4: Check the exit condition for the log-likelihood difference, i.e. $\left|\ell(\lambda^{(n+1)}) - \ell(\lambda^{(n)})\right| \leq 10^{-4}$, and exit if true or else proceed to Step 5,

Step 5: Expand the grid by adding new support points or reset the grid when support points are too close in proximity, go to Step 2.

D-Function

Define the D-Function, the directional derivative of $\ell(F)$ in the direction of the atomic density function δ_ξ, as

$$D(\xi, F) = \left(\sum_{i=1}^{P} \frac{p(Y_i \mid \xi)}{p(Y_i \mid F)} \right) - P. \tag{7}$$

Convergence of NPAG to F^{ML}

D-Function Properties

- $F^* = F^{ML}$ if and only if $\max_{\xi \in \Omega} D(\xi, F^*) = 0$.

Convergence of NPAG to F^{ML}

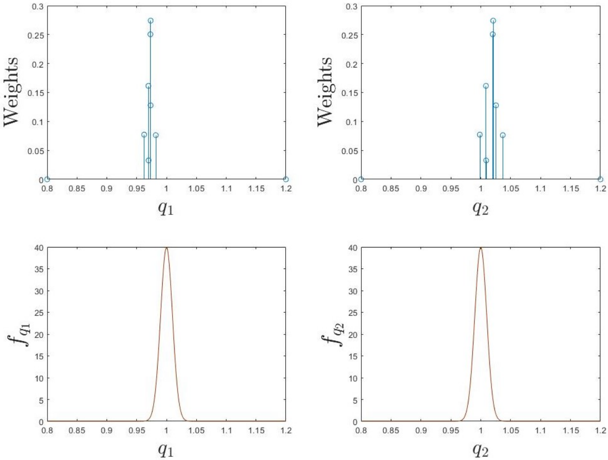
Channel Islands
CALIFORNIA STATE UNIVERSITY

D-Function Properties

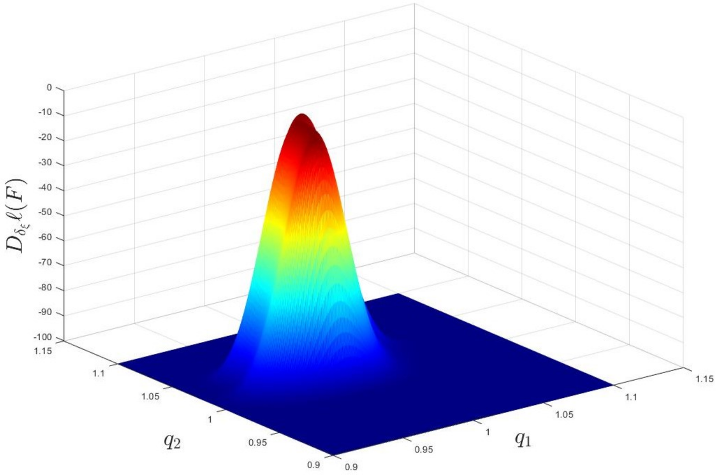
- $F^* = F^{ML}$ if and only if $\max_{\xi \in \Omega} D(\xi, F^*) = 0$.
- When $\max_{\xi \in \Omega} D(\xi, F^*) \neq 0$, it is true that

$$L(F^{ML}) - L(F^*) \leq \max_{\xi \in \Omega} D(\xi, F^*)$$

for $F^*, F^{ML} \in \mathscr{F}_K$.

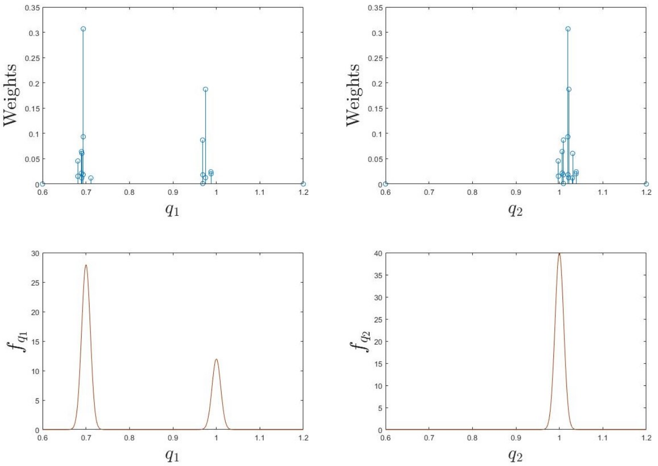


Weighted Mean: $q_1 = 0.9724$, $q_2 = 1.0183$



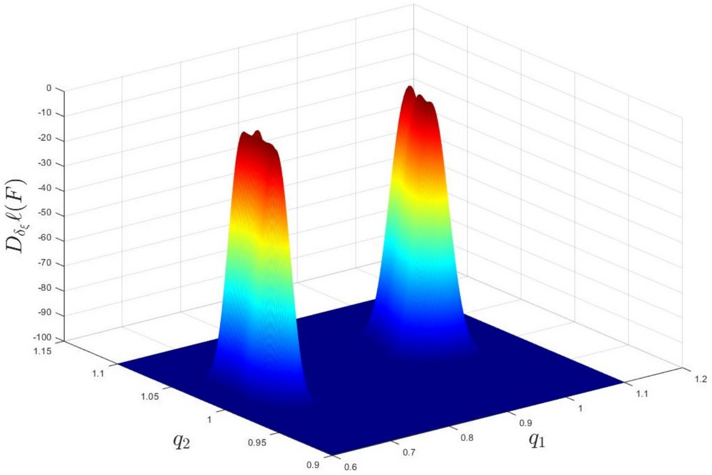
D-Function max $= -0.0019$

Pair 2: Maximum Likelihood using Matlab

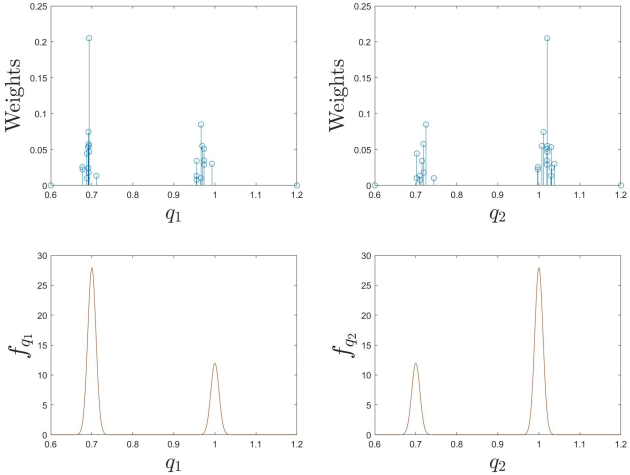


Weighted Mean: $q_1 = 0.7909$, $q_2 = 1.0183$

Pair 2: D-Function using Matlab

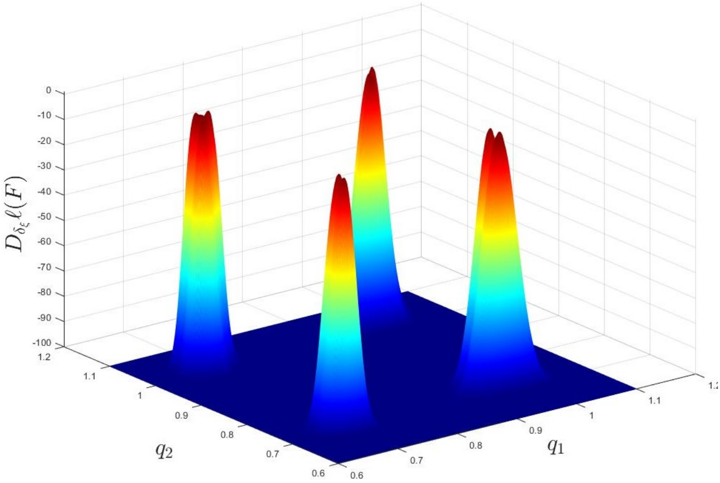


D-Function max $= -0.0026$



Weighted Mean: $q_1 = 0.7888$, $q_2 = 0.9343$

Pair 3: D-Function using Matlab



D-Function max = 0.0029

Table of Contents

Conclusions

- A relevant summary of the history of alcohol biosensors and TAC was provided.

Conclusions

- A relevant summary of the history of alcohol biosensors and TAC was provided.
- A new diffusion model was developed and tested.

Conclusions

- A relevant summary of the history of alcohol biosensors and TAC was provided.
- A new diffusion model was developed and tested.
- Theory was established to reduce the maximum likelihood estimator to finite support.

Conclusions

- A relevant summary of the history of alcohol biosensors and TAC was provided.
- A new diffusion model was developed and tested.
- Theory was established to reduce the maximum likelihood estimator to finite support.
- Simulated data tested and verified the model.

Conclusions

- A relevant summary of the history of alcohol biosensors and TAC was provided.
- A new diffusion model was developed and tested.
- Theory was established to reduce the maximum likelihood estimator to finite support.
- Simulated data tested and verified the model.
- The algorithm results for NPAG Algorithm are very encouraging and constitute further study.

Future Work

- Compare finite element method with B-splines against other discretization methods.

Future Work and Conclusions

Future Work

- Compare finite element method with B-splines against other discretization methods.
- Utilize a stochastic process for error. Probably Brownian motion.

Future Work and Conclusions

Future Work

- Compare finite element method with B-splines against other discretization methods.
- Utilize a stochastic process for error. Probably Brownian motion.
- Gather new BAC, BrAC, and TAC training data from humans.

Future Work

- Compare finite element method with B-splines against other discretization methods.
- Utilize a stochastic process for error. Probably Brownian motion.
- Gather new BAC, BrAC, and TAC training data from humans.
- Use a Bayesian approach to estimate the parameter for an individual.

Table of Contents

Acknowledgements

Special thanks to

my advisor,	**Dr. Alona Kryshchenko**
my thesis committee,	**Dr. Jorge Garcia** and
	Dr. Brian Sittinger
and our CalTech colleage	**Dr. Melike Sirlanci**

The generosity of CSU Channel Islands helped to fund this thesis research with a Mini-Grant AY 2018-2019, and for that I am extremely grateful.


Any Questions?

Selected Bibliography

1 Chubatiuk, Alona (2013), "Nonparametric estimation of an unknown probability distribution using maximum likelihood and Bayesian approaches." Dissertation, USC.

2 Lindsay, Bruce G. (1983), "The Geometry of Mixture Likelihoods: A General Theory." Annals of Statistics, 11, 1, pp. 86-94.

3 W. M. Yamada and M. Neely and J. Bartroff and D. Bayard and J. Burke and M. van Guilder and R. W. Jelliffe and A. Kryshchenko and R. Leary and T. Tatarinova and A. Schumitzky (2019), "An algorithm for nonparametric estimation of a multivariate mixing distribution with applications to population pharmacokinetics," Preprints, doi: 10.20944/preprints201909.0231.v1.

4 Multiphysics Cyclopedia (2017), "The Finite Element Method (FEM)," Comsol [online]. Available at https://www.comsol.com/multiphysics/finite-element-method.

5 Roberts, A. Wayne, and Varberg, Dale E. (1973), *Convex functions*, New York, New York: Academic Press.

6 Schultz, Martin H. (1972), *Spline Analysis*, Englewood Cliffs, New Jersey: Prentice-Hall.

7 Schumitzky, Alan (1991), "Nonparametric EM Algorithms For Estimating Prior Distributions." Applied Mathematics and Computation, 45, pp. 143-157.

8 Tuysuzoglu, Melike Sirlanci (2018), "Finite Dimensional Approximation and Convergence in the Estimation of the Distribution of, and Input to, Random Abstract Parabolic Systems with Application to the Deconvolution of Blood/Breath Alcohol Concentration from Biosensor Measured Transdermal Alcohol Level," Dissertation, USC.