SEIO 2016

Optimal Experimental Design apply to Bioassays in Workers Exposed to Radioactive Aerosols

J. Guillermo Sánchez León
Jesús López-Fidalgo

Optimun Experimental Design Group
http://areaestadistica.uclm.es/oed/

M05 .1 Sesión Hispano - Alemana : 6 de septiembre de 2016 11:00

The problem

- **Bioassays, such as urine sampler or measurement in a lung body counter, are used to estimate the intake in accidental and routinely situations for worker exposed to radioactive aerosols.**
- **The worker is exposed each working day *i* to an unknown quantity of the isotope b_i. Periodically according to a periodic schedule (e.g., Annually) the bioassay is made producing a value m_i.**
- **An upper limit (UL) for the value of the measurement is established. When m_i>UL then the total quantity intake, $B = \sum_{i=0}^{k-1} b_i$, should be estimated.**

The procedure

For a worker exposed to a daily chronic intake during a time t, the estimated retention is :

$$y_T[t] = b_0\, r[t-0] + b_1\, r[t-1] + \ldots + b_{k-1}\, r[t-(k-1)]$$

t is the time, in days, since the first intake.

$r(t)$ is a mathematical expression that gives the retention or excretion t days after a single intake b=1 happens. It depends on its kind of isotope, its chemical and physical form. It can be calculated using a code as BIOKMOD (http://diarium.usal.es/guillermo/biokmod).

The day t_1 a bioassay is made (e.g., m Bq of U234 of lung retention or m Bq of U234 in a accumulate 24 hour of urine sample) given a value m_0, the day a new bioassay t_2 is made given a value m_1 and so on

$$m_0 \pm \in = y_T[t_1] = b_0\, r[t_1-0] + b_1\, r[t_1-1] + \ldots + b_{k-1}\, r[t_1-(k-1)]$$

$$m_1 \pm \in = y_T[t_2] = b_0\, r[t_2-0] + \ldots + b_{k-1}\, r[t_2-(k-1)] + b_k\, r[t_2-(k-1)+1] + \ldots + b_{k+j}\, r[t_2-(k-1)+j]$$

The goal is estimating $B = \sum_{i=0}^{k-1} b_i$ using: $\{m_1, \ldots, m_n\}$

Non linear regression are used to estimate B

```
Clear["Global`*"]
```

B = n $< b >$ is computed using non linear regression. Mathematica and a toolbox (Biokmod), developed by one (Sánchez) the author, is used.

```
Needs["Biokmod`Doses`"]; Needs["Biokmod`Fitmodel`"]
```

The parameters (e.g. I_i or b_i) are fitted:

By minimizing $\sum_{i=1}^{N}(F_i - Y_i)^2$ or $\sum_{i=1}^{N}(F_i - Y_i)^2/s_i^2$

Applying the following equations:

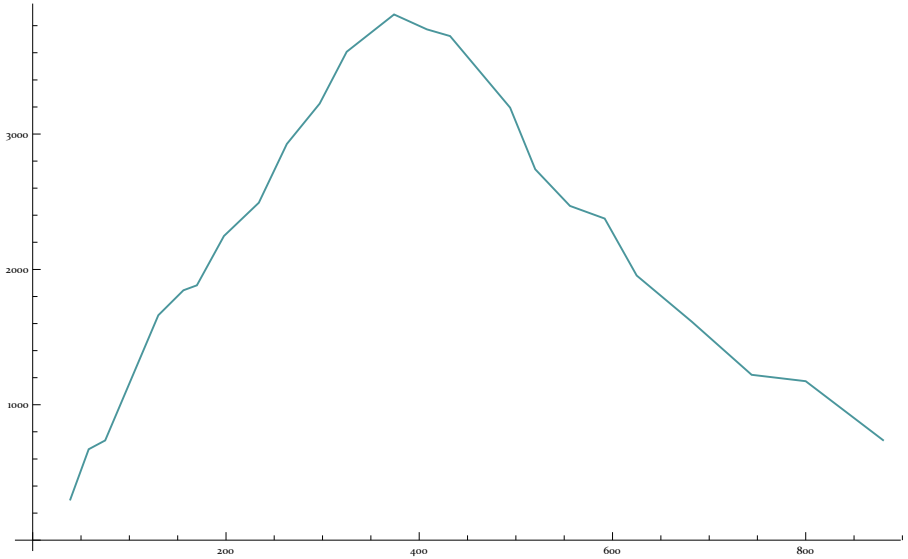
$$\hat{I}=\frac{\sum_{i=1}^{N}\frac{F_i Y_i}{s_i^2}}{\sum_{i=1}^{N}\frac{F_i^2}{s_i^2}};\ \hat{\sigma}_I = \frac{1}{\sqrt{\sum_{i=1}^{N}\frac{F_i^2}{s_i^2}}};\ \text{Log}(\hat{I}) = \frac{\sum_{i=1}^{N}\left(\text{Log}(Y_i/F_i)/\left(\text{Log}(\text{SF}_{i,j})^2\right)\right)}{\sum_{i=1}^{N}\left(1/\text{Log}(\text{SF}_{i,j})^2\right.} \Rightarrow \hat{I} = e^{\text{Log}(\hat{I})}$$

F_i denotes the values of the functions $F_i(I, t)$ associated to a measurement Y_i, s_i.

SF is the geometric standard deviation (SG) of the log-normal distribution.

Example 1: As a result of Chernobyl accident (26 April 1986) a male 39 years old and 80 kg (member of the public) was exposed to continuous and unknown ingestion of Cs-137 (This data has been supplied by Ansoborlo): {time after the accident (d), m_i(Bq Cs-137)}

```
wholeData = {{39, 300}, {58, 671}, {75, 737}, {130, 1661}, {156, 1846}, {170, 1882},
   {198, 2247}, {234, 2493}, {263, 2926}, {297, 3224}, {325, 3608}, {374, 3883}, {408, 3773}, {432, 3723}, {494, 3195},
   {520, 2740}, {556, 2469}, {592, 2375}, {625, 1954}, {682, 1614}, {744, 1221}, {800, 1174}, {880, 739}};
ListPlot[wholeData, Joined → True]
```



The retention function for the whole body of Cs-137 (that is for an acute intake "1" in t=0) is

```
qWbCs137[t1_] = qWholebody[compartMatrix[caesium], 1, 1, t1, Log[2] / (30 * 365.24)]
```

$$-0.000177381\, e^{-24.0001\, t1} + 0.00165462\, e^{-12.0001\, t1} - 0.0230333\, e^{-2.77265\, t1} + 0.0207699\, e^{-1.80006\, t1} - 0.0430482\, e^{-1.00006\, t1} + 0.139387\, e^{-0.346063\, t1} + 0.904447\, e^{-0.00636326\, t1}$$

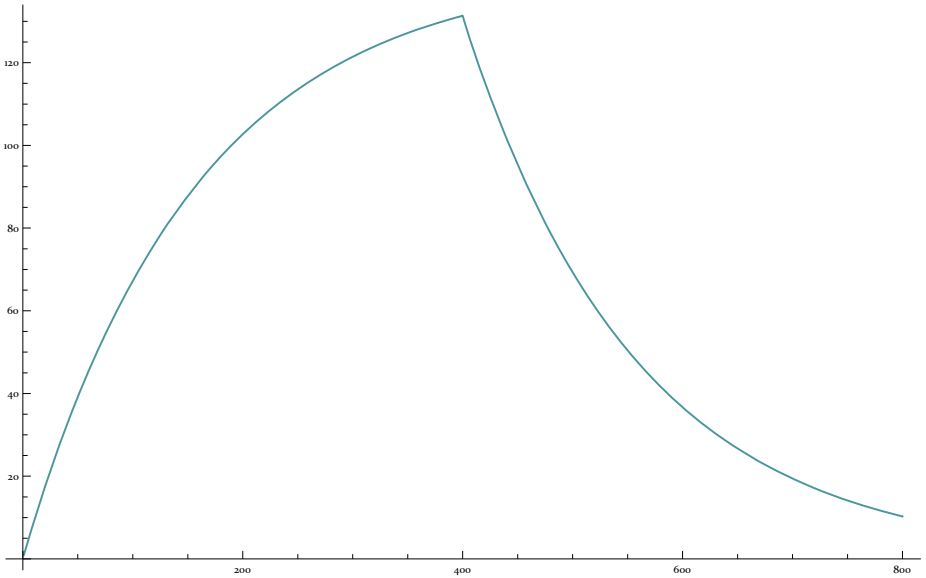
This function is applicable to a daily chronic ingestion "inp" during a time t1. The ingestion of caesium ceased in $t = T$, for $t > T$, inp = 0.

$$u_c(T) = \int_0^T u(t)\,dt$$

```
qConstant[1, {qWbCs137[t], t}, t1, 400] (*T= 400, for tutorial proposal*)
```

$$\begin{cases} \text{t1 must be non negative} & t1 < 0 \\ 142.499 + 7.39086\times10^{-6}\, e^{-24.0001\,t1} - 0.000137884\, e^{-12.0001\,t1} + 0.00830732\, e^{-2.77265\,t1} - \\ \quad 0.0115384\, e^{-1.80006\,t1} + 0.0430455\, e^{-1.00006\,t1} - 0.402779\, e^{-0.346063\,t1} - 142.136\, e^{-0.00636326\,t1} & 0 \le t1 \le 400 \\ 0. - 7.39086\times10^{-6}\, e^{-24.0001\,(t1-400)} + 0.000137884\, e^{-12.0001\,(t1-400)} - 0.00830732\, e^{-2.77265\,(t1-400)} + 0.0115384\, e^{-1.80006\,(t1-400)} - & t1 > 400 \\ \quad 0.0430455\, e^{-1.00006\,(t1-400)} + 0.402779\, e^{-0.346063\,(t1-400)} + 142.136\, e^{-0.00636326\,(t1-400)} + 7.39086\times10^{-6}\, e^{-24.0001\,t1} - 0.000137884\, e^{-12.0001\,t1} + \\ \quad 0.00830732\, e^{-2.77265\,t1} - 0.0115384\, e^{-1.80006\,t1} + 0.0430455\, e^{-1.00006\,t1} - 0.402779\, e^{-0.346063\,t1} - 142.136\, e^{-0.00636326\,t1} & \end{cases}$$

```
Plot[qConstant[1, {qWbCs137[t], t}, t1, 400], {t1, 1, 800}]
```



It can be observed that the retention was increasing until T. We can suppose that the caesium ingestion happened until T, when it ceased.

Now we can fit the experimental data bearing in mind both periods.

```
model2[t1_?NumericQ, p_?NumericQ, tt_?NumericQ] := p qConstant[1,{qWbCs137[t],t}, t1, tt]
{input, timeIntake} = {p, tt} /. FindFit[wholeData, model2[t, p, tt], {p, tt}, t] (*b,in Bq/day,and the period T,in days*)
```

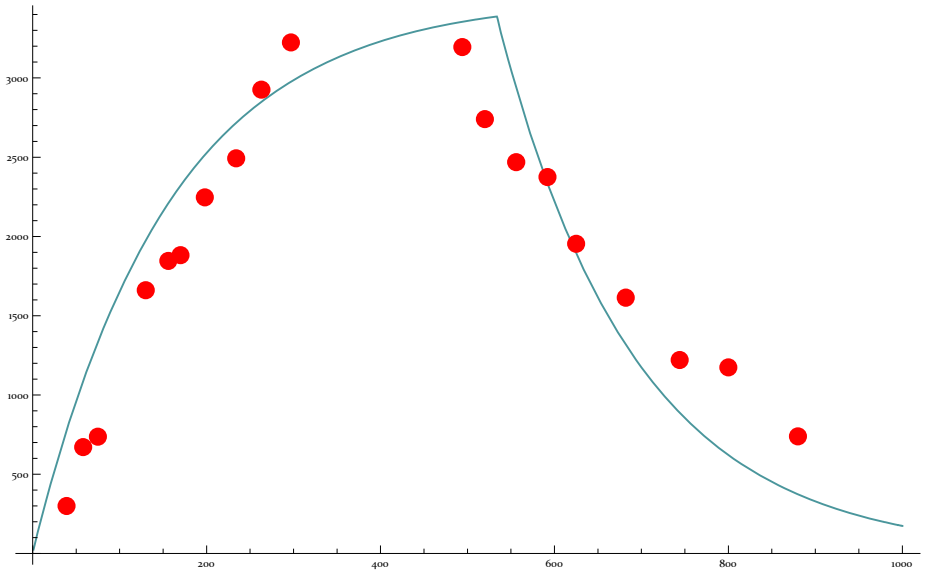
Then, the accumulated intake is

```
input timeIntake "Bq"
```

13 132.9 Bq

It can be observed the good match obtained

```
Plot[qConstant[input, {qWbCs137[t], t}, t1, timeIntake], {t1, 1, 1000}, Epilog → {Hue[0.], PointSize[0.02], Map[Point, wholeData]}]
```



Retention function mixed given by a linear combination of two retention functions.

Example: A worker was exposed from 1982-06-01 to 1990-05-14 to a chronic inhalation of enriched UO2/U3O8 radioactive aerosols of AMAD 5. Urine samples were taken before and after the intake ceased. The values of the samples are the following: {(t: day after the start intake, m: Bq/day, sd: uncertainty of the sample}:

```
sample = {{27, 0.55, 0.16}, {125, 0.125, 0.037}, {2116, 0.11, 0.033}, {2418, 0.17, 0.053},  {2575, 0.16, 0.048}, {2780, 0.11, 0.033},
   {2939, 0.057, 0.003}, {3061, 0.046, 0.003}, {3170, 0.022, 0.002}, {3249, 0.023, 0.002},  {3408, 0.026, 0.002},
   {3523, 0.026, 0.001}, {3614, 0.00893, 0.0035575}, {3680, 0.011, 0.005}, {3802, 0.01428, 0.00446}, {3886, 0.00984, 0.0053},
   {3977, 0.00856, 0.0058}, {4048, 0.01724, 0.00889}, {4156, 0.01093, 0.007064}, {4250, 0.01198, 0.0061209}, {4369, 0.0094, 0.00421}};
```

The total days of the period where the worker was exposed to the intake, in Bq, have been:

```
timecronic = QuantityMagnitude[DateDifference[{1982, 06, 01}, {1990, 05, 14}]]
```

2904

In this case the metabolism behavior of the intake radioactive aerosols can be represented by two chemical components: UO2 and U3O8 mixed: The metabolism of UO2 is described by $r_S[t]$ and the U3O8 by $r_S[t]$. Then the mixed retention function is given by

```
rSM[t_, w_] = w rS[t] + (1 - w) rS[t];
```

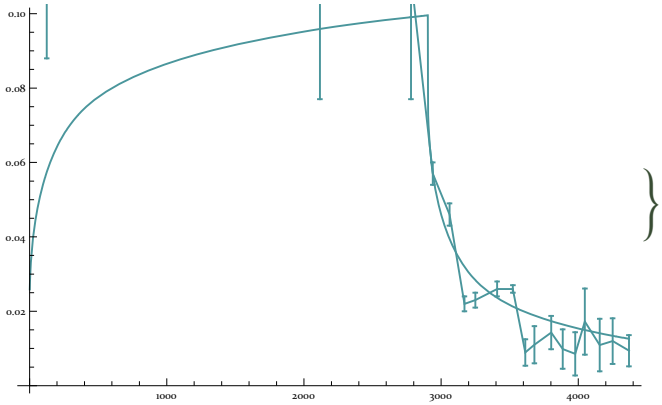
where w is a weight factor, usually unknown

The values of total intake, in Bq, and the fraction type S and M of radioactive aerosols are (standard values of f1 are used: 0.002 for S and 0.02 for M)

```
UrineBioassayEvaluation[sample, uranium, 0, timecronic, AMAD5, 0.002, 0.02]
```

{"Ji2" → 94.1652, "Total intake" → 24763., "Fraction S" → 0.896343,

"Fraction M" → 0.103657,



}

Optimal experimental design

It will be an important enhance to find the best moments where these additional bioassay measurements should be taken. With this goal it is very useful to apply optimal experimental design. The goodness of the design will depend on the number of samples and the measurement accuracy.

C - optimal design

A correlation structure usually employed when the measures are taken on the same individual is the exponential covariance (Cressie and Wikle, 2011), decreasing with the increasing distance in time between observations, $\text{Cov}(y_t, y_{t+d}) = \sigma^2 \text{Exp}(-\rho\, d)$, where σ^2 is the (assumed constant) variance of the measurements. The parameter ρ is characteristic of the individual, and for the purposes we will test with different ρ values.

Case 1.-

It is assumed each working day j ($j \in \mathbb{Z}^+$): {0, 1, ..., k} where the worker intakes an unknown single quantity b_j. In $t = k$ the exposures ceased. The first measurement is taken at $t_0 = k + d_0$ (d_0 is the time since the last intake) with a value m_1, the second one in $k + d_0 + d_1$ with results m_2, and so on,

$$B = \left(1^T \left(X^T \Sigma^{-1} X\right)^{-1} 1\right) X^T \Sigma^{-1} Y. .$$

$$\hat{\sigma}_B^2 = 1^T \left(X^T \Sigma^{-1} X\right)^{-1} 1.$$

The objective is to estimate the total intake $b = \sum_{j=1}^{k} b_j = \mathbf{1}^T B$, where $\mathbf{1}^T = (1, \ldots, 1)$ and $B^T = (b_1, \ldots, b_k)$. Optimal experimental designs have to be obtained for this estimate.

Two repeated measurements are being performed at times $t_0 = k$ and $t_1 > k$, say $Y = (y_0, y_1)^T$. Time t_1 has to be optimized for the variance of the estimator. Let X be the design matrix,

$$X = \begin{pmatrix} r(t_0) & r(t_0 - 1) & \cdots & r[t_0 - (k-1)] \\ r(t_1) & r(t_1 - 1) & \cdots & r[t_1 - (k-1)] \end{pmatrix},$$

and

$$\Sigma = \sigma^2 \begin{pmatrix} 1 & e^{-\rho(t_1 - t_0)} \\ e^{-\rho(t_1 - t_0)} & 1 \end{pmatrix}$$

be the covariance matrix of the observations. If $k = 2$ the Generalized Least Squared Estimator (GLSE) of the parameter is $\hat{B} = (X^T \Sigma^{-1} X)^{-1} X^T \Sigma^{-1} Y$ while the covariance matrix of the estimators $\hat{B}$ is $(X^T \Sigma^{-1} X)^{-1}$. The estimator of b is then $\hat{b} = \mathbf{1}^T \hat{B}$ and the variance to be minimized, $\mathbf{1}^T (X^T \Sigma^{-1} X)^{-1} 1$. If $k > 2$ then the estimator of b has to be computed using a pseudo-inverse: $\hat{b} = \mathbf{1}^T (X^T \Sigma^{-1} X)^{-} X^T \Sigma^{-1} Y$. Its variance is $\mathbf{1}^T (X^T \Sigma^{-1} X)^{-} 1$.

The Mathematica functions

The information matrix for a design ξ and normally distributed random errors is given by: $X^T \Sigma^{-1} X$. (where $\Sigma = \sigma^2 * \Gamma$, by default: $\sigma = 1$)

A two-point design

A N-point design

infMatrixG[d_0, {d_1, d_2, ... }, k, ρ]

```
infMatrixG[d0_, list_List, k_, ρ_ , σ_: 1] := Module[{X, Σ, i , j, t, tt, d, h, dd, ddd, ff, n}, n = Length[list];
   tt = Table[t_i, {i, 0, n}]; dd = Table[d_i, {i, 0, n}]; ddd = Thread[dd → Flatten[{d0, list}]];
   X = Table[r[t_j - i] /. Thread[tt -> Accumulate[dd]], {j, 0, n}, {i, 0, k - 1}] /. ddd;
ff[i_, j_] := Which[i == j, 1, i < j, e^(-ρ Σ_(h=i)^(j-1) d_h), i > j, e^(-ρ Σ_(h=j)^(i-1) d_h)];
Σ = σ^2 Array[ff, {n + 1, n + 1}] /. ddd;
   Transpose[X] . Inverse[Σ] . X];
varBG[d0_, list_List, k_, ρ_, σ_: 1] := Module[{matrizS}, matrizS = PseudoInverse[infMatrixG[d0, list, k, ρ , σ]];
  PadLeft[{1}, k, 1] . matrizS . PadLeft[{1}, k, 1]]
```

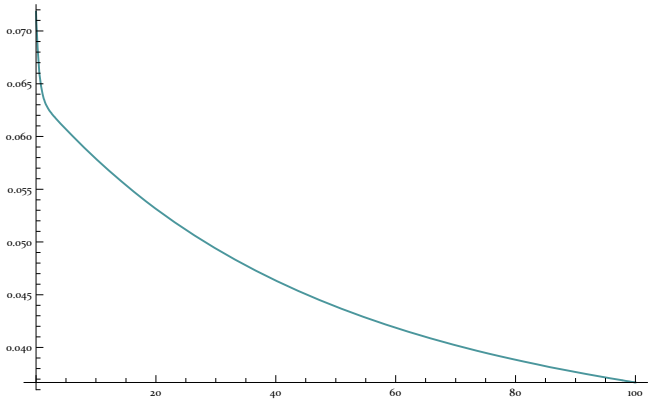
Then $\hat{\sigma}_B^2(t_i, k, \rho) > \text{Min}[\hat{\sigma}_B^2]$ is computed.

```
MinVarB[{d0_, k_, rho_, σ_: 1}, dmax_] :=
 Module[{c, d1}, c = Table[{ d1, varBG[d0, {d1}, k, rho, σ]}, {d1, 1, dmax} ]; First[MinimalBy[c, Last]]]
MinVarBG[{d0_, k_, rho_, σ_: 1}, dmax_] :=
 Module[{c, d1, d2}, c = Flatten[Table[{ d1, d2, varBG[d0, {d1, d2}, k, rho, σ]}, {d1, 5, dmax}, {d2, 5, dmax} ], 1]; First[MinimalBy[c, Last]]]
```

Case: Lung retention of UO2 by Inhalation, AMAD 5

Let's take as retention function $r_{LS}(t)$ (Uranium UO2 by inhalation AMAD 5. This function is applicable to all isotopes of enriched uranium)

```
rLS[t_] = +0.00798 e^-2.0 t + 0.0103 e^-0.03 t + 0.0161 e^-0.02 t + 0.0319 e^-0.0011 t + 0.00443 e^-0.00022 t + 0.00109 e^-0.0001 t;
r[t_] = rLS[t] (*Lung retention , metabolism S, AMAD 5*);
Plot[r[t], {t, 0, 100}, PlotRange → All]
```



A two-point design

The first measure should be taken as soon as possible (Lopez-Fidalgo J; Sánchez G; Statistical Criteria to Establish Bioassay Programs. Health Physics.. 89 (4).2005). Usually $t_0 = 1$ (in days).

inVarB[{to_, k_, rho_, σ_: 1}, tmax_]

- **We apply the functions before described to compute the C design for different k and ρ values**

```
sol1 = Table[Transpose[MinVarB2[{k, k, #}, k + 180] & /@ {0.01, 0.1, 0.5, 1, 10}][[1]] - k, {k, {3, 10, 50, 100, 500, 1000}}]
sol1 = {{1, 1, 1, 66, 66}, {1, 130, 130, 130, 130},
   {115, 145, 145, 145, 145}, {116, 149, 149, 149, 149}, {105, 161, 161, 161, 161}, {96, 159, 159, 159, 159}};
```

```
TableForm[Join[{{0.01, 0.1, 0.5, 1, 10}}, sol1], TableHeadings → { {"k↓/rho→", 3, 10, 50, 100, 500, 1000}, None}]
```

k↓/rho→	0.01	0.1	0.5	1	10
3	1	1	1	66	66
10	1	130	130	130	130
50	115	145	145	145	145
100	116	149	149	149	149
500	105	161	161	161	161
1000	96	159	159	159	159

```
SetDirectory[NotebookDirectory[]]
```

```
Export["disenos1.csv", sol1]
```

Work in progress

Case 2: The chronic intake is represented by: $u_c(t) = \int_0^t u(t)\,dt$.

Case 3: The intake follow after each bioassay.

Case 4: `u[t, w] = w u_S[t] + (1 - w) u_M[t];`

Case 5: b_i is a random variable that follow a known distribution function

Case 5

Thank you for your attention

http://diarium.usal.es/guillermo