Paper Number(s): **E4.26**
C2.2

IMPERIAL COLLEGE OF SCIENCE, TECHNOLOGY AND MEDICINE
UNIVERSITY OF LONDON

DEPARTMENT OF ELECTRICAL AND ELECTRONIC ENGINEERING
EXAMINATIONS 2001

MSc and EEE PART IV: M.Eng. and ACGI

ESTIMATION AND FAULT DETECTION

Monday, 14 May 10:00 am

There are SIX questions on this paper.

Answer FOUR questions.

Time allowed: 3:00 hours

Corrected Copy

Examiners: Clark,J.M.C. and Allwright,J.C.

Special instruction for invigilators: **None**

Information for candidates

Some formulae relevant to the questions

The normal $N(m, \sigma^2)$ density: $p(y) = \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{(y-m)^2}{2\sigma^2}\right)$

System equations:

$$\begin{aligned} x_{k+1} &= Ax_k + Bu_k + Mv_k \\ y_k &= Cx_k + Nw_k \end{aligned}$$

Here, v_k and w_k are standard white-noise sequences.

The Kalman one-step-ahead predictor:

$$\begin{aligned} \hat{x}_{k+1|k} &= A\hat{x}_{k|k-1} + K(k)\left(y_k - C\hat{x}_{k|k-1}\right) \\ K(k) &= AP_{k|k-1}C^T\left(CP_{k|k-1}C^T + NN^T\right)^{-1} \\ P_{k+1|k} &= AP_{k|k-1}A^T + MM^T - AP_{k|k-1}C^T\left(CP_{k|k-1}C^T + NN^T\right)^{-1}CP_{k|k-1}A^T \end{aligned}$$

The average quadratic cost identity:

$$\begin{aligned} E&\left[\sum_{k=0}^{N-1}\left(x_k^TQx_k + u_k^TRu_k\right) + x_N^TQ_Nx_N\right] \\ &= E\left[x_0^TS_0x_k + \sum_{k=0}^{N-1}(u_k + F_kx_k)^T(B^TS_{k+1}B + R)(u_k + F_kx_k)\right] \\ &\quad + \sum_{k=0}^{N-1} tr\left(S_{k+1}MM^T\right) \end{aligned}$$

where for $k = 0, \ldots, N-1$,

$$\begin{aligned} F_k &= \left(B^TS_{k+1}B + R\right)^{-1}B^TS_{k+1}A \\ S(k) &= A^TS_{k+1}A + Q - A^TS_{k+1}B\left(B^TS_{k+1}B + R\right)^{-1}B^TS_{k+1}A, \quad S_N = Q_N \end{aligned}$$

The algebraic Riccati equations:

$$S = A^TSA + Q - A^TSB\left(B^TSB + R\right)^{-1}B^TSA \qquad \text{(control)}$$
$$P = A^TPA + MM^T - APC^T\left(CPC^T + NN^T\right)^{-1}CPA^T \qquad \text{(filtering)}$$

1. Let $x(t)$ be the state process of the continuous time model

$$\dot{x}(t) = Ax(t) + Mv(t)$$

where $v(t)$ is a vector white noise process with covariance function $E\left[v(t)v(s)^T\right] = I\delta(t-s)$.

(a) Show that for $h > 0$ the sampled state $x_k = x(kh)$, $k = 0, 1, \ldots$, satisfies an equation of the form

$$x_{k+1} = \bar{A}x_k + \bar{v}_k$$

where $\bar{v}_k$ is discrete time white noise with covariance matrix $Q = E[\bar{v}_k\bar{v}_k^T]$. Determine how $\bar{A}$ and Q are expressed in terms of A, M and h and obtain a formula for $E[x_{k+1}x_k^T]$.

(b). In the case where $h = 200$,

$$A = \begin{bmatrix} -1 & 0 \\ 0 & -0.01 \end{bmatrix}, \qquad M = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

and the process $x(t)$ is assumed to be stationary, show that $x_k = (x_{1,k}, x_{2,k})^T$ is approximate white noise in the sense that the following cross-correlations are less than 0.2 :

$$\frac{E[x_{1,k}x_{2,k}]}{\left(Ex_{1,k}^2 Ex_{2,k}^2\right)^{\frac{1}{2}}}, \qquad \frac{E[x_{2,k+1}x_{2,k}]}{\left(Ex_{2,k+1}^2 Ex_{2,k}^2\right)^{\frac{1}{2}}}.$$

2. (a) Suppose X and Y are scalar random variables with a joint covariance matrix

$$\begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix} \quad (p_{22} > 0)\,.$$

Then the linear least squares estimate (LLSE) of X given Y is

$$\hat{X} = EX + p_{12}p_{22}^{-1}(Y - EY).$$

Prove this.

(b) Suppose $X, N_1 N_2, \ldots, N_n$ are independent uniformly distributed random variables. The density of X is constant at $\frac{1}{2a}$ on $[0, 2a]$ and that of each N_i is $\frac{1}{2}$ on [0,2]. Suppose

$$Y_i = X + N_i \quad i = 1, 2, \ldots, n\,.$$

It can be shown that the LLSE of X given $Y_1, \ldots, Y_n$ is the same as that of X given $\bar{Y}$, where $\bar{Y} = \frac{1}{n}\sum_1^n Y_i$. Use this fact to derive the LLSE of X given $Y_1, \ldots, Y_n$.

(c) Comment on whether you would expect the LLSE in (b) to be close to the conditional mean of X given Y_1 in the case when $n = 1$. Illustrate your answer with a sketch of the set on which the joint density of X and N_1 is positive.

3. A depth finder measures the depth of water beneath the keel of a ship. It produces a sequence of measurements of the form

$$y_k = x_k + w_k, \quad k = 0, 1, 2, \ldots$$

where x_k is the depth under the ship at time k and w_k is standard white noise ($Ew_k = 0$, $Ew_k w_l = \delta_{kl}$). The depth x_k is modelled by the equation

$$x_{k+1} = x_k + \sigma v_k$$

where v_k is a standard white noise process independent of w_k. The variance σ^2 of the depth noise is assume to lie between $\frac{1}{4}$ and 4.

Using the formulae at the front of the paper, express the steady-state error variance $p(\sigma^2)$ associated with the Kalman one-step-ahead predictor as a function of σ^2 and calculate it in the case where $\sigma^2 = 4$.

(b) An "observer" filter is to be employed with a fixed gain K, where $0 < K < 2$:

$$\hat{x}_{k+1} = \hat{x}_k + K(y_k - \hat{x}_k).$$

Show that the steady-state variance of the resulting estimation error $\tilde{x}_k = x_k - \hat{x}_k$ satisfies the formula

$$E[\tilde{x}_k^2] = \frac{\sigma^2 + K^2}{2K - K^2}.$$

(c) The gain K is to be chosen to make the filter of part (b) insensitive to the value of σ^2. Explain why, if K is chosen to minimize the "worst-case" error covariance

$$max\left\{E\tilde{x}_k^2 \; : \; \frac{1}{4} \leq \sigma^2 \leq 4\right\},$$

the resulting error covariance coincides with the error variance produced by the Kalman predictor for $\sigma^2 = 4$.

4. Consider the following target tracking problem. The target motion is represented by the linear system

$$x_{k+1} = Ax_k + Mv_k$$

At time k the tracker has available observations $y_0, y_1, \dots, y_{k-1}$ where y_i is given by

$$y_i = Cx_i + w_i .$$

In these equations, v_k and w_k are independent Gaussian white noise processes with identity matrix covariances. The initial state x_0 has distribution $N(m_0, P_0)$, independent of v_k, w_k. The dynamics of the tracker - in which z_k is of the same dimension as x_k - are given by

$$z_{k+1} = \bar{A}z_k + \bar{B}u_k$$

where u_k is a scalar control input and $\bar{A}$ and $\bar{B}$ are, respectively, a constant matrix and a constant vector. For each k, the control u_k is to be chosen as a function of $y_0, \dots, y_{k-1}$ and z_k to minimize

$$E[\sum_{k=0}^{N-1} \left(u_k^2 + ||\hat{x}_N - z_N||^2\right)] .$$

(a) Show that the optimal control law is the same as that for the problem of minimizing:

$$E[\sum_{k=0}^{N-1} \left(u_k^2 + ||\hat{x}_{N|N-1} - z_N||^2\right)] .$$

(b) Suppose the Kalman filter for the one-step predictor $\hat{x}_k := \hat{x}_{k|k-1}$ is time invariant with constant conditional covariance $P_0 = cov(x_k|y_{k-1}, \dots, y_0)$.

Using the formulae at the front of the paper, show that the control problem can be reformulated as a complete information LQG problem in which $q_k = \begin{pmatrix} \hat{x}_k \\ z_k \end{pmatrix}$ is regarded as as a "hyperstate" that satisfies a time invariant stochastic state equation.

Give the statistics of the noise term in this equation and formulate the criterion that is to be minimized.

5. A stochastically disturbed state process

$$x_{k+1} = Ax_k + Bu_k + Mv_k$$

is controlled by time-invariant state feedback $u = u(x_k)$ chosen to minimize - over the class of stabilizing controls for which Ex_k and $E[x_k x_k^T]$ converge to constants - the average cost rate $lim_{N\to\infty} J^{u,N}$, where

$$J^{u,N} = \frac{1}{N} E\left[\sum_{k=0}^{N-1} \left(x_k^T Q x_k + u_k^T R u_k\right)\right].$$

Here v_k is white noise with $Ev_k v_l^T = I\delta_{kl}$ and Q and R are positive definite matrices.

(a) State conditions for there to be a unique positive semidefinite solution to the "control" form of the associated algebraic Riccati equation (ARE) given at the front of the paper.

(b) Assume that the control ARE possesses a unique positive semidefinite solution S. Prove that the control law u_k that minimizes the finite-time cost

$$E[\sum_{k=0}^{N-1} \left(x_k^T Q x_k + u_k^T R u_k\right) + x_N S x_N]$$

takes the time-invariant form $u_k = u^0(x_k)$ and simultaneously minimizes the average cost rate $lim_{N\to\infty} J^{u,N}$.

(c) Suppose x_k is scalar, $A = a$ and $B = M = Q = R = 1$. Determine the optimal control law.

6. In a simple test between two hypotheses: the null hypothesis, "index $J = 0$" and the alternative, "$J = 1$", N independent measurements are made.

If $J = 0$, the probability density of each y_k is $p_0(y)$; if $J = 1$ the probability density of each y_k is $p_1(y)$.

If $\pi_0 = P(J = 0)$ and $\pi_1 = P(J = 1)$ are the prior probabilities of the two alternatives, then the minimum probability-of-error Bayes test is

$$\text{Choose} \quad J = 1 \quad \text{if} \quad \frac{\pi_1}{\pi_0} \prod_{k=1}^{N} \frac{p_1(y_k)}{p_0(y_k)} \geq 1\,;$$

$$\text{Choose} \quad J = 0 \quad \text{otherwise.}$$

(a) Show that this Bayes test has the interpretation of picking the "conditionally most likely" of the two alternatives, given the measurements.

(b) In a test to detect radiation, the N "inter-click" times y_k between $N + 1$ clicks produced by a Geiger counter are independent and identically distributed as

$$p_0(y) = 0.1e^{-0.1y}, \; y \geq 0 \quad \text{if there is no radiation}$$

$$p_1(y) = 0.5e^{-0.5y}, \; y \geq 0 \quad \text{if radiation is present.}$$

The prior probability of radiation being present is 0.5.

Derive a Bayes test in this case, and show that it depends only on the sufficient statistic $Y_N = \sum_1^N y_k$.

If

$$N = 4 \;\text{ and }\; y_1 = 6, \quad y_2 = 3, \quad y_3 = 5, \quad y_4 = 4\,,$$

what are the conditional "odds" of the presence of radiation against its absence?

END

1 Solution - ESTIMATION and FAULT DETECTION 2001

(a) By the variation-of-constants formula $\frac{1}{7}$

$$x(t) = e^{A(t-r)}x(r) + \int_r^t e^{A(t-s)} M\, v(s)\,ds.$$

So if $r = kh$, $t = (k+1)h$

$$x_{k+1} = e^{Ah}x_k + \int_{kh}^{(k+1)h} e^{A((k+1)h-s)} M\, v(s)\,ds.$$

$$= \bar{A}x_k + \bar{v}_k$$

As $v(s)$ is white clearly the $\bar{v}_k$ are uncorrelated for different k. Further, $E\bar{v}_k = \underline{0}$ and

$$E[\bar{v}_k\bar{v}_k^T] = \int_{kh}^{(k+1)h} e^{A((k+1)h-s)} M M^T e^{A^T((k+1)h-s)}\,ds.$$

10
$$= \int_0^h e^{As} M M^T e^{A^T s}\,ds = Q$$

Also $E[x_{k+1}x_k^T] = \bar{A}\,E[x_k x_k^T]$

(b) For stationary processes

$$x_k = \int_{-\infty}^{kh} e^{A(kh-s)} M\, v(s)\,ds$$

$$\therefore E[x_k x_k^T] = \int_0^\infty e^{At} M M^T e^{A^T t}\,dt$$

$$= \int_0^\infty \begin{bmatrix} e^{-t} & 0 \\ 0 & e^{-0.01t} \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} e^{-t} & 0 \\ 0 & e^{-0.01t} \end{bmatrix} dt$$

$$= \int_0^\infty \begin{bmatrix} e^{-2t} & e^{-1.01t} \\ e^{-1.01t} & e^{-0.02t} \end{bmatrix} dt.$$

$$= \begin{bmatrix} \frac{1}{2} & \frac{1}{1.01} \\ \frac{1}{1.01} & 50 \end{bmatrix}$$

5 So $$\frac{E(x_{1,k}x_{2,k})}{(Ex_{1,k}^2\; Ex_{2,k}^2)^{1/2}} = \frac{.99}{25^{1/2}} = .198$$

1(b) Solution continued

2/7

As $h = 200$ and $E[x_k x_k^T] = E[x_{k+1} x_{k+1}^T]$

$$E[x_{k+1} x_k^T] = e^{Ah} E[x_k x_k^T] = \begin{bmatrix} e^{-200} & 0 \\ 0 & e^{-2} \end{bmatrix} \begin{bmatrix} \frac{1}{2} & .99 \\ .99 & 50 \end{bmatrix}$$

$$\approx \begin{bmatrix} 0 & 0 \\ .99e^{-2} & 50e^{-2} \end{bmatrix}$$

5 So $$\frac{E[x_{2,k+1} x_{2,k}]}{(E x_{2,k+1}^2 \, E x_{2,k}^2)^{1/2}} = \frac{50e^{-2}}{(50^2)^{1/2}} = e^{-2} = 0.135$$

7/7

2 Solution

(a) A linear estimate is of the affine form $\hat{X} = a + bY$ which can be written

$$\hat{X} = EX + a' + b\,(Y - EY)$$

for suitable a'

Minimise over a', b

$$J = E[\,(X - (EX + a') - b(Y - EY))^2\,]$$

8

$$= E[\,(X-EX)^2 + a'^2 - 2b(X-EX)(Y-EY) + b^2(Y-EY)^2\,]$$

$$\frac{\partial J}{\partial a'} = 0 \Rightarrow 2a' = 0$$

$$\frac{\partial J}{\partial b} = 0 \Rightarrow -2\rho_{12} + b\rho_2 = 0 \sim b = \rho_{12}\rho_{22}^{-1}.$$

Hence the LLSE $\hat{X} = EX + \rho_{12}\rho_{22}^{-1}(Y - EY)$

(b) Clearly $EX = a$, $EX^2 = \frac{1}{2a}\int_0^{2a} x^2\,dx = \frac{(2a)^2}{3} = \frac{4}{3}a^2$

So $\operatorname{Var} X = (\frac{4}{3} - 1)a^2 = \frac{a^2}{3}$. Similarly $EN_i = 1$, $\operatorname{Var} N_i = \frac{1}{3}$

$\bar{Y} = X + \frac{1}{n}\sum_1^n N_i$ So $E\bar{Y} = a + 1$, $\operatorname{Var}\bar{Y} = \frac{a^2}{3} + n\frac{1/3}{n^2} = \frac{1}{3}(a^2 + \frac{1}{n})$

$$\rho_{12} = \operatorname{Cov}(X, \bar{Y}) = \operatorname{Var} X = \frac{a^2}{3}$$

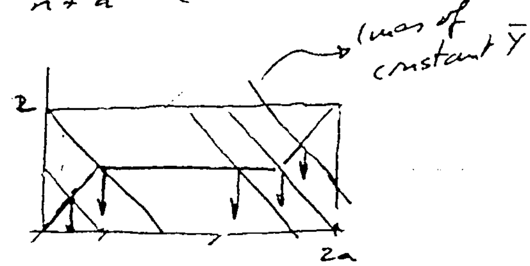
8 So the LLSE of X given $\bar{Y}$ (+ therefore $Y_1, \dots Y_n$) is

$$\hat{X} = a + \frac{a^2}{3}\frac{3n}{(na^2+1)}(\bar{Y} - a - 1)$$

$$= a + \frac{n}{n + a^{-2}}(\bar{Y} - a - 1)$$

(c) $E(X|\bar{Y})$ has discontinuous derivatives, as a fn of $\bar{Y}$, at $\bar{Y} = 2$, $\bar{Y} = 2a$.

4 The LLSE is smooth + so must differ.



3. solution

(a) The algebraic Riccati equation for this filter is

$$p = p + \sigma^2 - \frac{p^2}{p+1} \quad \text{or} \quad p^2 - \sigma^2 p - \sigma^2 = 0$$

The steady-state variance is the positive solution:

$$p_\infty(\sigma^2) = \tfrac{1}{2}\left(\sigma^2 + \sqrt{\sigma^4 + 4\sigma^2} \right)$$

4 and the Kalman gain $K(\sigma^2) = \dfrac{p_\infty}{1 + p_\infty} = \dfrac{\sigma^2 + \sqrt{\sigma^4 + 4\sigma^2}}{2 + \sigma^2 + \sqrt{\sigma^4 + 4\sigma^2}}$

So if $\sigma^2 = 4$

3 $$p_\infty(4) = \tfrac{1}{2}(4 + 4\sqrt{2}) = 2(1 + \sqrt{2}) = 4.28$$

(b) Since $x_{k+1} = x_k + \sigma v_k$

and $\hat{x}_{k+1} = \hat{x}_k + K(x_k + w_k - \hat{x}_k)$

$$\tilde{x}_{k+1} = \tilde{x}_k + \sigma v_k - K w_k - K \tilde{x}_k$$
$$= (1-K)\tilde{x}_k + \sigma v_k - K w_k$$

7 As $0 < K < 2$ the process $\tilde{x}_k$ is asymptotically stationary with steady-state 2nd moment

$$E\tilde{x}_{k+1}^2 = (1-K)^2 E\tilde{x}_k^2 + \sigma^2 + K^2$$

$$\text{or} \quad E\tilde{x}_k^2 = \frac{\sigma^2 + K^2}{2K - K^2}$$

(c) For fixed K the largest $E\tilde{x}_k^2$ is given by $\sigma^2 = 4$. (as $0 < K < 2$)

6 But the Kalman predictor minimises the variance for fixed σ^2. So the minimized worst-case variance $= p_\infty(4) = 4.28$.

4 Solution

(a) Target equation $\quad x_{k+1} = A x_k + M v_k \quad (1)$

Tracker $\quad z_{k+1} = \bar{A} z_k + \bar{B} u_k \quad (2)$

Observations $\quad y_k = C x_k + w_k \quad (3)$

From (1) + (3), which are unaffected by the tracker, we can calculate a Kalman filter for $\hat{x}_{k/k-1}$. Since z_k is a function of $y_0, \cdots y_{k-1}$,

10 $$E[\| x_N - z_N \|^2] = E[E \| x_N - \hat{x}_{N/N-1} + \hat{x}_{N/N-1} - z_N \|^2 / (y_{N-1}, \cdots)]]$$

$$= E[\| x_N - \hat{x}_{N/N-1} \|^2] + E[\| \hat{x}_{N/N-1} - z_N \|^2]$$

The first term on the right is unaffected by the control. So the optimal law also minimizes

$$E[\Sigma_0^{N-1} u_k^2 + \| \hat{x}_{N/N-1} - z_N \|^2].$$

(b) $\hat{x}_k$ satisfies $\quad \hat{x}_{k+1} = A \hat{x}_k + K(y_k - C\hat{x}_{k/k-1}) \quad K = \text{Kalman gain}$

But the term $\bar{v}_k = K(y_k - C\hat{x}_{k/k-1})$ has covariance.

$$Q_0 = APC^T(I + CPC^T)CPA^T$$

10 So we can write

$$q_{k+1} = \begin{pmatrix} \hat{x}_{k+1} \\ z_{k+1} \end{pmatrix} = \begin{bmatrix} A & 0 \\ 0 & \bar{A} \end{bmatrix} q_k + \begin{bmatrix} 0 \\ \bar{B} \end{bmatrix} u_k + \begin{bmatrix} I \\ 0 \end{bmatrix} \bar{v}_k$$

where $\bar{v}_k$ is white noise with $E \bar{v}_k \bar{v}_k^T = Q_0$

The criterion becomes $\quad E[\Sigma_0^{N-1} u_k^2 + q_N^T \begin{bmatrix} I & -I \\ -I & I \end{bmatrix} q_N]$

5 Solution

(a) The control ARE has a unique positive semidefinite solution if

4 (A, B) is a stabilizable pair
$(Q^{1/2}, A)$ is a detectable pair.

(b) The finite-time optimal control law is given by the quadratic identity (see front of paper)

$$u_k^o = -F_k x_k \quad , F_k = (B^T S_{k+1} B + R)^{-1} B^T S_{k+1} A$$

But as $Q_N = S$, the solution of the ARE $S_k = S$ for all k, and so $u_k^o = -F x_k$ where $F = (B^T S B + R)^{-1} B^T S A$.

So, for any other stabilizing law u',

$$N J^{u';N} + E[x_N'^T S x_N'] \geq N J^{u^o;N} + E[x_N^{oT} S x_N^o]$$

9

and so $J^{u';N} \geq J^{u^o;N} + \frac{1}{N}\left[E(x_N^{oT} S x_N^o) - E(x_N'^T S x_N')\right]$

Since both $E x_N^{oT} S x_N^o$, $E x_N'^T S x_N'$ converge to constants (as u^o, u' are stabilizing). So $\lim_N J^{u';N} \geq \lim_N J^{u^o;N}$

This proves that $u_k^o = -F x_k$ is optimal.

(c) S solves the ARE

$$S = a^2 S + 1 - \frac{a^2 S^2}{1+S}$$

or

$$S^2 - a^2 S - 1 = 0$$

5 Since $S \geq 0$, $S = +\frac{1}{2}a^2 + \frac{1}{2}\sqrt{(a^4+4)}$

$$u_k^o = -F x_k = \frac{-aS}{1+S} x_k = -\frac{a^3 + a\sqrt{a^4+4}}{2 + a^2 + \sqrt{a^4+4}} x_k .$$

6. Solution 7/7

(a) The density - probability function of $J=0$ and y_k, $k=1,\dots,N$

is $\pi_0 \cdot \prod_{k=1}^{N} p_0(y_k)$

+ that of $J=1$, + $y_k : k=1,\dots,N$ is $\pi_1 \cdot \prod_{k=1}^{N} p_1(y_k)$.

So the conditional density of $y_1, \dots y_N$ given $J=0$, $J=1$ are

10 $$p(y_1, \cdots y_N \mid J=0) = \frac{\pi_0 \prod_{k=1}^{N} p_0(y_k)}{K}$$

$$p(y_1 \cdots y_N \mid J=1) = \frac{\pi_1 \prod_{k=1}^{N} p_1(y_k)}{K}$$

where $K = \pi_0 \prod_{k=1}^{N} p_0(y_k) + \pi_1 \prod_{k=1}^{N} p_1(y_k)$

The ratio is $\dfrac{\pi_1}{\pi_0} \dfrac{\prod_{k=1}^{N} p_1(y_k)}{\prod_{k=1}^{N} p_0(y_k)}$, which is just the test ratio.

(b) The likelihood ratio is

$$\frac{\frac{1}{2}}{\frac{1}{2}} \cdot \frac{(0.5)^N e^{-0.5 \sum_1^N y_k}}{(0.1)^N e^{-0.1 \sum_1^N y_k}}$$

$$= e^{-0.4 Y_N + N \log_e 5}$$

7

The test reduces to $Y_N \leq \dfrac{N \log_e 5}{0.4} = 4.02\,N$
$\Rightarrow$ radiation.

In the special case, $N=4$, $Y_4 = 18$

3

The odds become $e^{-0.4 \times 18 + 4 \times 1.609}$
$= e^{-4.(1.8 - 1.609)} = .466$

So "no radiation" is more than twice as likely as "radiation".