Appendix B

Inverse sequential detection of parameter changes in developing time series

by Uwe Radok and Timothy J. Brown

CIRES, Campus Box 449
University of Colorado
Boulder, CO 80309

Abstract

Progressive values of two probabilities are obtained for parameter estimates derived from an existing set of values and from the same set enlarged by one or more new values, respectively. One probability is that of erroneously preferring the second of these estimates for the existing data ("type I error"), while the second probability is that of erroneously accepting their estimates for the enlarged test ("type II error"). A more stable combined "no change" probability which always falls between 0.5 and 0 is derived from the (logarithmic) width of the uncertainty region of an equivalent "inverted" sequential probability ratio test (SPRT, Wald 1945) in which the error probabilities are calculated rather than prescribed.

A parameter change is indicated when the compound probability undergoes a progressive decrease. The tests is explicitly formulated and exemplified for Gaussian samples.

Inverse sequential detection of parameter changes in developing time series

by Uwe Radok and Timothy J. Brown

1. Introduction

The series will be assumed to consist of values from a population with probability function

$$f(x) = p(x;\theta;\theta';\theta'';\cdots)\,dx \tag{1}$$

The existing, or most recent, m values of the series provide parameter estimates θ_m; θ'_m; θ''_m; $\cdots$ which change to θ_{m+j}; θ'_{m+j}; θ''_{m+j}; $\cdots$ after a further 1,2, $\cdots j$ values have been added to the series. These parameter estimates define four different likelihoods:

$$L_m(m) = \prod_1^m p_m \qquad L_{m+j}(m) = \prod_1^m p_{m+j}$$

$$L_{m+j}(m+j) = \prod_1^{m+j} p_{m+j} \qquad L_m(m+j) = \prod_1^{m+j} p_m \tag{2}$$

Throughout bracketed symbols indicate the numbers of values used to calculate the likelihoods, and subscripts the numbers of values used for the parameter estimates.

The four equations (2) have been combined for a formal test of the hypotheses $H(m)$: $\theta = \theta_m$, $\theta' = \theta'_m$, $\theta'' = \theta''_m$ $\cdots$, and $H(m+j)$: $\theta = \theta_{m+j}$, $\theta' = \theta_{m+j}$, $\theta'' = \theta''_{mj}$ $\cdots$, in terms of the Neyman-Pearson theory (see e.g., Hoel 1962, chapter 9). this showed (Radok and Brown, submitted) that the likelihood *ratios* $q(m) = L_{m+j}(m)/L_m(m)$ and $q(m+j) = L_{m+j}(m+j)/L_m(m+j)$ can be estimated by the well-known decision limits of Wald's (1945) "sequential probability ratio test (SPRT)":

$$q(m) = \frac{\beta}{1-\alpha} \qquad q(m+j) = \frac{1-\beta}{\alpha} \tag{3}$$

where α is the type I error probability of rejecting the estimates θ_m, θ'_m, θ''_m $\cdots$ for the exist-

ing m values, and β is the type II error probability of accepting the same estimates for the augmented set of $m+j$ values; β also represents the type I error of not accepting the estimates θ_{m+j}, θ'_{m+j}, θ''_{m+j}; for that augmented set of values. Our test in effect places the likelihood ratio $q(m)$ on Wald's lower deicision limit, and the ratio $q(m+j)$ on upper limits which change as further values become available. But in contrast to the SPRT the "inverse" procedure involves *known* likelihood ratios and *unknown* probabilities which can be found from (3) as

$$\alpha = (1 - q(m))/(q(m+j) - q(m))$$
$$\beta = (q(m+j)q(m) - q(m))/(q(m+j) - q(m)) \tag{4}$$

Changes in one or several of the parameters are indicated when these error probabilities show decreasing trends. New parameter values can then be estimated from the new values alone. Since the individual probabilities can be somewhat irregular due to rounding-off errors, trends become more clearly visible in their average or in a compound "no-change" probability γ defined by the logarithmic width of the SPRT's indecision region,

$$\log_e[(1-\beta)/\alpha] - \log_e[\beta/(1-\alpha)] = \log_e[q(m+j)/q(m)] = \log_e Q \quad ;$$

writing this as

$$\log_e Q = \log_e \left[\frac{1-\gamma}{\gamma}\right]^2 \tag{5}$$

leads to

$$\gamma = (1 + \sqrt{Q})^{-1} \tag{6}$$

γ falls between 0 and 0.5 as long as $q(m+j) > (q(m))$ and between the arithmetic and geometric averages of α and β (see appendix A).

2. Inverse sequential formulae for Gaussian means and variances.

The basic probability for the Gaussian distribution,

$$p = (2\pi\sigma^2)^{-1/2} \exp\left[-(x-\mu)^2/2\sigma^2\right] \tag{7}$$

involves two parameters (μ, σ^2) which can only be tested jointly since in the present context neither is known a priori. We use the sample mean as estimate for the population mean μ, and $\sigma^2 = [(n/(n-1)]\, s^2$ where s^2 is the sample variance and $n\,(= m$ or $m+j)$ is the number of values used. The products (2) are converted to sums by taking logarithms; thus

$$\log_e L_m(m) = \frac{m}{2}\log_e 2\pi - \frac{m}{2}\log_e \sigma_m^2 - \sum_1^m (x-\mu_m)^2/2\sigma_m^2 \quad ; \tag{8}$$

with, $\sum_1^m (x-\mu)^2 = (m-1)\sigma_m^2$ the last term reduces to $-(m-1)/2$.

Proceeding in the same way for $L_{m+j}(m)$ leads to

$$\log L_{m+j}(m) - \frac{m}{2}\log_e 2\pi - \frac{m}{2}\log_e \sigma_{m+j}^2 - \sum_1^m (x-\mu_{m+j})^2/2\sigma_{m+j}^2 \quad ; \tag{9}$$

with $\mu_{m+j} - \mu_m = \Delta\mu$ the numerator of the last term can be written

$$-\sum_1^m (x-(\mu_m+\Delta\mu))^2 = -\sum_1^m (x-\mu_m)^2 - \sum_1^m (-2x\Delta\mu + 2\mu_m\Delta\mu + \Delta\mu^2)$$

$$= -(m-1)\sigma_m^2 - 2m\mu_m\Delta\mu + 2m\mu_m\Delta\mu - m(\Delta\mu)^2$$

so that (9) becomes

$$\log_e L_{m+j} = -\frac{m}{2}\log 2\pi - \frac{m}{2}\log_e \sigma_{m+j}^2 - \frac{(m-1)\sigma_m^2}{2\sigma_{m+j}^2} - \frac{m(\Delta\mu)^2}{2\sigma_{m+j}^2} \tag{9'}$$

Finally subtracting (8) from (9') gives the log likelihood ratio

$$\log_e q(m) = \log_e \frac{L_{m+j}(m)}{L_m(m)} = \frac{m}{2}\log_e \frac{\sigma_m^2}{\sigma_{m+j}^2} + \frac{m-1}{2}\left[1 - \frac{\sigma_m^2}{\sigma_{m+j}^2}\right] - \frac{m(\Delta\mu)^2}{2\sigma_{m+j}^2} \tag{10}$$

Proceeding the same way for the augmented set of $m+j$ values yields first

$$\log_e m+j(m+j) = -\frac{m+j}{2}\log_e 2\pi - \frac{m+j}{2}\log_e \sigma^2_{m+j} - \sum_1^{m+j}(x-\mu_{m+j})^2/2\sigma^2_{m+j} \qquad (11)$$

where the last term, with $\sum(x-\mu_{m+j})^2 = (m+j-1)\,\sigma^2_{m+j}$ reduces to $(m+j-1)/2$. Next

$$\log_e L_m(m+j) = -\frac{m+j}{2}\log_e 2\pi - \frac{m+j}{2}\log_e \sigma^2_m - \sum_1^{m+j}(x-(\mu_{m+j}-\Delta\mu))^2/2\sigma^2_{m+j}\ . \qquad (12)$$

Expanding the last term as before yields

$$\log_e L_m(m+j) = -\frac{m+j}{2}\log_e 2\pi - \frac{m+j}{2}\log_e \sigma^2_m - \frac{(m+j-1)\sigma^2_{m+j}}{\sigma^2_m} - \frac{(m+j)\Delta\mu^2}{2\sigma^2_m} \qquad (12')$$

Subtracting (12') from (11) we obtain the second log likelihood radio

$$\log_e q(m+j) = \frac{m+j}{2}\log_e \frac{\sigma^2_m}{\sigma^2_{m+j}} + \frac{m\ j-1}{2}\left[\frac{\sigma^2_{m+j}}{\sigma^2_m} - 1\right] + \frac{(m+j)(\Delta\mu)^2}{2\sigma^2_m} \qquad (13)$$

The final formulae therefore are

$$q(m) = \exp\left[m\log_e \frac{\sigma_m}{\sigma_{m+j}} + \frac{m-1}{2}\left[1-\frac{\sigma^2_m}{\sigma^2_{m+j}}\right] - \frac{m}{2\sigma_{m+j}}(\mu_{m+j}-\mu_m)^2\right]$$

$$q(m+j) = \exp\left[(m+j)\log_e \frac{\sigma_m}{\sigma_{m+1}} + \frac{m+j-1}{2}\left[\frac{\sigma^2_{m+j}}{\sigma^2_m} - 1\right] + \frac{m+j}{2\sigma^2_m}(\mu_{m+j}-\mu_m)^2\right] \qquad (14)$$

The first two exponents in each formula reflect solely changes in variance, while the third exponents depend primarily on changes in the mean.

3. Example.

Fig. 1a shows a series of values drawn at random from Gaussian populations with different means and variances. The different probabilities are shown in fig. 1b and suggest a change starting with the 7th value and accelerating after the 9th value, the actual beginning of a new section in "statistical control" (Shewhart 1939). A spurious change is suggested briefly after the 14th

value, but this is reversed until a real change becomes visible after the 26th value, just inside the next controlled section. It is noteworthy that these control changes are revealed even though the population parameters differed considerably from their (small) sample estimates employed by the test, as the only information available in practice.

4. Conclusion.

We plan to carry out more extensive experiments to establish the full properties of the inverse sequential procedure, and to test its efficacy on different geophysical time series. Additional parameters for which the procedure has been formulated, but not yet adequately tested, include Poisson means (variances), chi-square means (representing also variances and degrees of freedom) and linear and exponential trends. These will be reported in a further paper.

Acknowledgement. The work here described has been supported by NASA Grant NAGW-2706. Partial support for the second author was provided by NOAA's Climate and Global Change Program.

References:

Hoel, P.G., 1962: *Introduction to mathematical statistics*. Wiley, 428 pp.

Radok, U. and Brown, T.J., submitted: Detecting change as it occurs. Submitted to Climatic Change.

Shewhart, W.A., 1939: Statistical method from the viewpoint of quality control. U.S. Dep. of Agriculture, Washington, D.C.

Wald, A., 1945: Sequential tests of statistical hypotheses. Annals of Math. Stats. *16*(1945), 117-186.

Captions:

Fig. 1: An inverse sequential detection of parameter changes.

a) Values drawn at random from three Gaussian populations with means and variances indicated.

b) Probabilities that no parameter changes are occurring.

For symbols see text.

Appendix A: The no-change probability y

The probability γ as defined by equation (6) is,

$$\gamma = (1 + \sqrt{Q})^{-1} , \tag{A1}$$

and lies between the arithmetic and geometric means of the probabilities α and β. This can be shown by alternatively substituting these means for α and β in the expanded equation (A1),

$$Q = \frac{[1 - (\alpha+\beta) + \alpha\beta]}{\alpha\beta} . \tag{A2}$$

When $\alpha = \beta = (\alpha + \beta)/2$, equation (A3) becomes

$$Q_1 = \frac{\left[1 - (\alpha + \beta) + \alpha\beta + \frac{\alpha^2 + \beta^2}{2}\right]}{\left[\alpha\beta + \frac{\alpha^2 + \beta^2}{2}\right]} \tag{A3}$$

This shows that the numerator N and the denominator D of Q both have been increased by $\varepsilon = (\alpha^2 + \beta^2)/2 > 0$. Since $Q > 1$ (i.e., $N > D$), then $Q = N/D > Q_1 = (N + \varepsilon)/(D + \varepsilon)$ since $ND + N\varepsilon > ND + D\varepsilon$, or $N > D$, the initial condition.

Again with $\alpha = \beta = (\alpha\beta)^{1/2}$, equation (A3) becomes

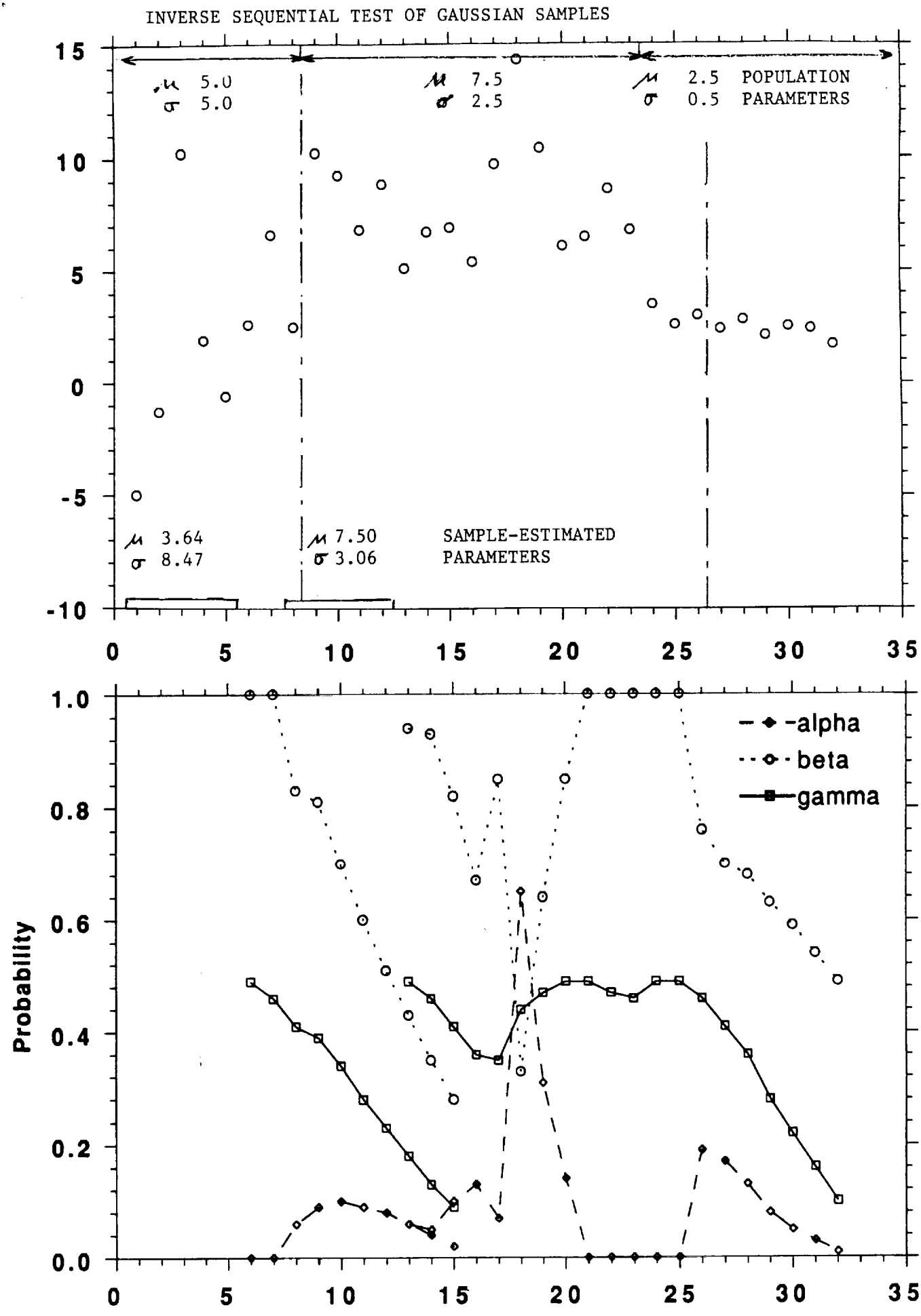
$$Q_2 = \frac{\left[1 - 2(\alpha\beta)^{1/2} + \alpha\beta\right]}{\alpha\beta} , \tag{A4}$$

so that $Q_2 - Q = -2(\alpha\beta)^{1/2} + (\alpha + \beta) > 0$ since $\alpha + \beta > 2(\alpha\beta)^{1/2}$; this can be seen by squaring both sides giving

$$(\alpha + \beta)^2 + \alpha^2 + \beta^2 + 2\alpha\beta > 4\alpha\beta , \textit{ or } (\alpha - \beta)^2 > 0 . \tag{A5}$$

Finally, with $Q_2 > Q > Q_1$,

$$\left[1 + \sqrt{Q_2}\right]^{-1} = \gamma_{geometric} < \gamma = \left[1 + \sqrt{Q}\right]^{-1} < \gamma_{arithmetic} = \left[1 + \sqrt{Q_1}\right]^{-1} . \tag{A6}$$


INVERSE SEQUENTIAL TEST OF GAUSSIAN SAMPLES
μ 5.0
σ 5.0
μ 7.5
σ 2.5
μ 2.5 POPULATION
σ 0.5 PARAMETERS
μ 3.64
σ 8.47
μ 7.50
σ 3.06
SAMPLE-ESTIMATED PARAMETERS
15
10
5
0
-5
-10
0
5
10
15
20
25
30
35
1.0
0.8
0.6
0.4
0.2
0.0
Probability
alpha
beta
gamma