A new control chart for detecting linear trends

Eckhard Liebscher
University of Applied Sciences Merseburg
Department of Engineering and Natural Sciences
Eberhard-Leibnitz-Str. 2
06217 Merseburg
Germany

October 1, 2024

Abstract

In the paper we study properties of estimators ...

AMS Classification: ..

1 Introduction

The Exponentially Weighted Moving Average (EWMA) chart is a powerful tool in statistical process control. It is designed to monitor the process by use of a continuous quality variable. Unlike traditional control charts such as the Shewhart chart, the EWMA chart incorporates the entire history of the process by weighting observations in a geometrically decreasing order. Recent observations have a higher influence on the chart, making it particularly sensitive to small changes in the process. The EWMA chart is widely used in various industries to ensure consistent quality and to detect subtle trends that might indicate potential issues. The aim of this paper is to introduce a new chart which is able to detect linear trends. Here were incorporate the idea of exponentially weighted regression. Our approach differs significantly from the Holt-Winter exponential smoothing procedure which gives also estimators for the quality level and the linear trend (see Gardner and McKenzie (1985)).

Let $X_1, X_2, \ldots$ be the time series. The EWMA control chart serves as statistical tool for monitoring the production process. It tracks the exponentially weighted moving average of the quality feature values or a sample mean of it. In this way the chart gives more weight to recent observations and less weight to observations further back in the past. The values $\tilde{X}_t$ of the chart can be evaluated as follows:

$$\tilde{X}_t = \lambda X_t + (1-\lambda)\tilde{X}_{t-1} \text{ for } t = 1, 2 \ldots$$

where $\lambda \in (0,1)$ is the smoothing (learning) factor. The concept of exponentially weighted smoothing goes back to Brown (1956). This concept was introduced in the context of regression in Christiaanse (1971).

In applications, a linear trend is often present in the case of a perturbed behaviour of the production process. We develop a control chart which is especially designed to detect a slow onset of drift. The aim is to monitor the series in order to detect a possible linear trend. The idea is to estimate the slope using exponential weights. The estimator is constructed by the weighted least squares method for linear regression models. The concept of exponentially weighted regression is studied in Luxenberg and Boyd (2024). The reader finds the theory of weighted least squares in Myers et al. (2010), for example.

2 EWMA chart for slope

2.1 Theory

We consider the time series $X_1, X_2, \ldots$ The idea of our method is to fit a linear trend function and to assign exponentially weights to the observations in the quadratic loss. The quadratic loss function is given by

$$D_n(\beta) = \sum_{i=1}^{n} (1-\lambda)^{n-i} \left(X_i - a - bi\right)^2,$$

where $\beta = (a,b)^T$ and b is the slope. Define

$$\begin{aligned}
T_n &= \sum_{i=1}^{n} (1-\lambda)^{n-i} X_i, \; U_n = \sum_{i=1}^{n} i(1-\lambda)^{n-i} X_i, \\
l_n &= \sum_{i=1}^{n} (1-\lambda)^{n-i}, \; L_n = \sum_{i=1}^{n} i(1-\lambda)^{n-i}, \; K_n = \sum_{i=1}^{n} i^2(1-\lambda)^{n-i}, \\
\bar{l}_n &= \sum_{i=1}^{n} (1-\lambda)^{2n-2i}, \; \bar{L}_n = \sum_{i=1}^{n} i(1-\lambda)^{2n-2i}, \; \bar{K}_n = \sum_{i=1}^{n} i^2(1-\lambda)^{2n-2i}.
\end{aligned}$$

The minimization of D_n leads to the linear equation system

$$\begin{aligned}
\sum_{i=1}^{n} (1-\lambda)^{n-i} \left(X_i - a - bi\right) &= 0, \\
\sum_{i=1}^{n} i(1-\lambda)^{n-i} \left(X_i - a - bi\right) &= 0,
\end{aligned}$$

or equivalently,

$$\begin{aligned} T_n &= a\ l_n + b\ L_n \\ U_n &= a\ L_n + b\ K_n \end{aligned}$$

The solution of the system results in the estimators for offset and slope:

$$\begin{aligned} \hat{a} &= \frac{K_n T_n - L_n U_n}{K_n l_n - L_n^2}, \\ \hat{b} &= \frac{l_n U_n - L_n T_n}{K_n l_n - L_n^2}. \end{aligned}$$

For l_n, L_n, K_n, T_n, U_n recursion formulas are availabe

$$\begin{aligned} l_1 &= 1,\ l_{n+1} = \sum_{i=1}^{n+1} (1-\lambda)^{n+1-i} = (1-\lambda) l_n + 1 \\ L_1 &= 1,\ L_{n+1} = \sum_{i=1}^{n+1} i(1-\lambda)^{n+1-i} = (1-\lambda) L_n + n + 1 \\ K_1 &= 1,\ K_{n+1} = \sum_{i=1}^{n+1} i^2(1-\lambda)^{n+1-i} = (1-\lambda) K_n + (n+1)^2 \end{aligned}$$

$$\begin{aligned} T_1 &= X_1,\ T_{n+1} = \sum_{i=1}^{n+1} (1-\lambda)^{n+1-i} X_i = (1-\lambda) T_n + X_{n+1} \\ U_1 &= X_1,\ U_{n+1} = \sum_{i=1}^{n+1} i(1-\lambda)^{n+1-i} X_i = (1-\lambda) U_n + (n+1) X_{n+1} \end{aligned}$$

These recursion formulas allow for the development of a fast computation algorithm.

In the following we consider two situations:

Situation 1: In-control process without perturbation: $\mathbb{E}X_i = \mu_0$ where μ_0 is the target value for the quality feature.

Situation 2: Out-of-control process with shift:

$$\mathbb{E}X_i = \mu_0 \text{ for } i \leq t_0, \quad \mathbb{E}X_i = \mu_0 + \delta \text{ for } i > t_0$$

with shift $\delta \neq 0$.

Situation 3: Out-of-control process with linear drift as perturbation:

$$\mathbb{E}X_i = \mu_0 \text{ for } i \leq t_0, \quad \mathbb{E}X_i = \mu_0 + (i - t_0)\delta \text{ for } i > t_0$$

with drift $\delta \neq 0$.

Theorem 1 *Assume that* $\mathbb{E}X_i = \mu_0, \mathrm{Var}(X_i) = \sigma^2$.
(i) Then

$$\mathrm{Var}(\hat{b}) = \frac{\sigma^2}{(L_n^2 - K_n l_n)^2} v_n, \; v_n = L_n^2 \bar{l}_n - 2 L_n l_n \bar{L}_n + l_n^2 \bar{K}_n.$$

Morover,

$$\mathrm{Var}(\hat{b}) = \frac{2\sigma^2\lambda^3}{(2-\lambda)^3} + o(1),$$

(ii) in-control case:

$$\mathbb{E}(\hat{b}) = 0$$

and

$$\hat{b} \to 0 \quad a.s.$$

(iii) out-of-control case of situation 3:

$$\hat{b} \to \delta \quad a.s. \text{ and } \mathbb{E}(\hat{b}) = \delta + o(1)$$

The values can be computed recursively as

$$\begin{aligned}
\bar{l}_1 &= 1, \;\; \bar{l}_{n+1} = \sum_{i=1}^{n+1} (1-\lambda)^{2n+2-2i} = (1-\lambda)^2 \bar{l}_n + 1 \\
\bar{L}_1 &= 1, \;\; \bar{L}_{n+1} = \sum_{i=1}^{n+1} i(1-\lambda)^{2n+2-2i} = (1-\lambda)^2 \bar{L}_n + n + 1 \\
\bar{K}_1 &= 1, \;\; \bar{K}_{n+1} = \sum_{i=1}^{n+1} i^2(1-\lambda)^{2n+2-2i} = (1-\lambda)^2 \bar{K}_n + (n+1)^2.
\end{aligned}$$

The chart is realized as follows: The out-of-control signal is triggered if

$$\left|\hat{b}\right| > c\hat{\sigma}\frac{\sqrt{v_n}}{K_n l_n - L_n^2}$$

is satisfied, where $c > 0$ is a certain constant to be chosen. At all, two quantities, λ and c are to be chosen by the practitioner. The goal is to figure out values c, λ such that for process without perturbation, the ARL is at least a fixed value (370.4 for example) and for the process with linear trend is minimal.

2.2 Simulations

Let $X_1, X_2, \ldots, X_t \in \mathbb{R}$ be the time series we intend to monitor. We simulated time series items according to one of the above described situation:

Situation 1 - in-control process without perturbation: $X_t \sim \mathcal{N}(\mu_0, \sigma^2)$ are i.i.d. random variables.

Situation 2 - shift:

$$X_t = \begin{cases} Z_t + \mu_0 \text{ for } i \leq t_0, \\ Z_t + \mu_0 + \delta \text{ for } i > t_0 \end{cases}$$

with shift $\delta \neq 0$, where $Z_t \sim \mathcal{N}(0, \sigma^2)$ are i.i.d. random variables

Situation 3 - linear drift:

$$X_t = \begin{cases} Z_t + \mu_0 \text{ for } i \leq t_0, \\ Z_t + \mu_0 + (i - t_0)\delta \text{ for } i > t_0 \end{cases}$$

with drift $\delta \neq 0$. ARL_0 is the average run length for the in-control process without perturbation.

We chose $t_0 = 20$ and $ARL_0 = 370.4$. In several cases we evaluated optimized charts with specific vectors (c, λ). The optimization of the ARL w.r.t. the vectors (c, λ) is realized by a series of simulations until the minimal ARL is reached subject to $ARL_0 \geq 370.4$. We simulated the time series 100000 times.

Table 1: Optimized values of (c, λ) and ARL for several scenarios, situation 2 "shift" and situation 3 "drift" as above (* means optimized values of c, λ)

	scenario	EWMAS			classical EWMA			combined method	
δ		ARL	c	λ	ARL	c	λ	ARL_0	ARL
0.5	shift	135.88	1.585	0.0256	18.28	1.535*	0.00010		
1	shift	12.77	1.562	0.0305	9.560	2.810*	0.130		
2	shift	3.079	1.565	0.0308	3.329	2.986*	0.382		
0.05	drift	15.31	1.585*	0.0256*					
0.1	drift	10.26	1.562*	0.0305*	12.695	2.87*	0.174*	212.98	10.80
					12.746	2.810	0.130		
0.2	drift	6.942	1.565*	0.0308*	8.310	2.94*	0.260*		
					8.492	2.986	0.382		

3 Proofs

First, we prove some auxiliary statements in this section.

Lemma 2 *For $a \in (0,1)$, we have*

$$\begin{aligned}
\sum_{i=1}^{n} i^2 a^{n-i} &= \frac{n^2}{(1-a)} - \frac{2na}{(1-a)^2} + \frac{a^2+a}{(1-a)^3} + o(1), \\
\sum_{i=1}^{n} i a^{n-i} &= \frac{n}{1-a} - \frac{a}{(1-a)^2} + o(n^{-1}), \\
\sum_{i=1}^{n} a^{n-i} &= \frac{1}{1-a} + o(n^{-2}).
\end{aligned}$$

PROOF: Obviously, one obtains

$$\sum_{i=1}^{n} i^2 a^{n-i} = \frac{-a^{n+1} - a^{n+2} + n^2 + (n+1)^2 a^2 + a\,(1 - 2n(n+1))}{(1-a)^3},$$

and

$$\sum_{i=1}^{n} i a^{n-i} = \frac{-a + n - an + a^{n+1}}{(1-a)^2}, \tag{1}$$

which implies the lemma. $\boxdot$

Lemma 3 *The following identities hold true for $\lambda \in (0,1)$:*

$$\begin{aligned}
&a) & l_n &= \frac{1-(1-\lambda)^n}{\lambda}, \\
&b) & L_n &= \frac{\lambda(n+1) - 1 + (1-\lambda)^{n+1}}{\lambda^2}, \\
&c) & \sum_{i=t_0+1}^{n} (l_n i - L_n)\,(1-\lambda)^{n-i} &= o(1), \\
&d) & \sum_{i=t_0+1}^{n} i\,(l_n i - L_n)\,(1-\lambda)^{n-i} &= \frac{1-\lambda}{\lambda^4} + o(1).
\end{aligned}$$

PROOF: Assertion a) follows directly. Assertion b) is a consequence of (1).

c)

$$
\begin{aligned}
& \sum_{i=t_0+1}^{n} (l_n i - L_n)(1-\lambda)^{n-i} \\
= \; & l_n \sum_{i=1}^{n-t_0} (i+t_0)(1-\lambda)^{n-t_0-i} - L_n \sum_{i=1}^{n-t_0} (1-\lambda)^{n-t_0-i} \\
= \; & \frac{\lambda(n-t_0+1)-1}{\lambda^3} + \frac{t_0}{\lambda^2} - \frac{\lambda(n+1)-1}{\lambda^3} + o(1) \\
= \; & \lambda^{-2} - \lambda^{-3} - \lambda^{-2} + \lambda^{-3} + o(1) = o(1).
\end{aligned}
$$

d) In view of Lemma 2, we can derive

$$
\begin{aligned}
& \sum_{i=t_0+1}^{n} i\,(l_n i - L_n)(1-\lambda)^{n-i} \\
= \; & l_n \sum_{i=1}^{n-t_0} (i+t_0)^2 (1-\lambda)^{n-t_0-i} - L_n \sum_{i=1}^{n-t_0} (i+t_0)(1-\lambda)^{n-t_0-i} \\
= \; & \left(\frac{1}{\lambda} + o(n^{-1})\right)\left(\frac{(n-t_0)^2}{\lambda} - \frac{2(n-t_0)(1-\lambda)}{\lambda^2} + \frac{\lambda^2-3\lambda+2}{\lambda^3} + 2t_0\left(\frac{n-t_0}{\lambda} - \frac{1-\lambda}{\lambda^2}\right) + \frac{t_0^2}{\lambda} + o(1)\right) \\
& - \left(\frac{\lambda(n+1)-1}{\lambda^2} + o(1)\right)\left(\frac{n-t_0}{\lambda} - \frac{1-\lambda}{\lambda^2} + t_0\frac{1}{\lambda} + o(1)\right) \\
= \; & \lambda^{-4}\left(n^2\lambda^2 + 2n\lambda^2 - 2n\lambda + \lambda^2 - 3\lambda + 2\right) - \lambda^{-4}\left(\lambda(n+1)-1\right)^2 + o(1) \\
= \; & \lambda^{-4}\left(n^2\lambda^2 + 2n\lambda^2 - 2n\lambda + \lambda^2 - 3\lambda + 2\right) + o(1) \\
= \; & \frac{1-\lambda}{\lambda^4} + o(1). \quad \boxdot
\end{aligned}
$$

The following lemma is a direct consequence of Theorem 9 of Chow and Lai (1973).

Lemma 4 *Let $W_1, \ldots, W_n$ independent random variables with $\mathbb{E}W_i = 0, \mathrm{Var}(W_i) = \sigma^2$. We consider the sum $Y_n = \sum_{i=1}^n a_{ni} W_i$, where a_{ni} are non-random constants. Assume that*

$$
\limsup_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} a_{ni}^2 < \infty, \quad \mathbb{E}W_i^2 < \infty.
$$

Then, by we obtain $Y_n \to 0$ a.s..

Proof of Theorem 1: Observe that

$$
\begin{aligned}
\lim_{n\to\infty} \left(K_n l_n - L_n^2\right) &= \frac{1-\lambda}{\lambda^4}, \\
\lim_{n\to\infty} (l_n n - L_n) &= \frac{1-\lambda}{\lambda^2}.
\end{aligned}
$$

Now let $a_{ni} = (l_n i - L_n)(1-\lambda)^{n-i}$. Further by Lemma 2, it follows that

$$
\begin{aligned}
\frac{1}{n}\sum_{i=1}^{n} a_{ni}^2 &= \frac{1}{n}\left(l_n^2\left(\frac{n^2}{\lambda(2-\lambda)} - \frac{2n(1-\lambda)^2}{\lambda^2(2-\lambda)^2} + o(n)\right) - 2l_n L_n\left(\frac{n}{\lambda(2-\lambda)} - \frac{(1-\lambda)^2}{\lambda^2(2-\lambda)^2} + o(1)\right)\right. \\
&\quad \left. + L_n^2\left(\frac{1}{\lambda(2-\lambda)} + o(n^{-1})\right)\right) \\
&= n\left(\frac{1-2+1}{\lambda^3(2-\lambda)}\right) - \frac{2(1-\lambda)^2}{\lambda^4(2-\lambda)^2} + 2\frac{(1-\lambda)^2}{\lambda^4(2-\lambda)^2} - \frac{2(\lambda-1)}{\lambda^4(2-\lambda)} + \frac{2(\lambda-1)}{\lambda^4(2-\lambda)} + o(1) \\
&= o(1).
\end{aligned}
$$

Note that $\sum_{i=1}^{n} (l_n i - L_n)(1-\lambda)^{n-i} = 0$. In the in-control case, we obtain

$$
l_n U_n - L_n T_n = \sum_{i=1}^{n} (l_n i - L_n)(1-\lambda)^{n-i}(X_i - \mu_0) = o(1) \ a.s.
$$

by applying Lemma 4. It follows that

$$
\hat{b} = \frac{l_n U_n - L_n T_n}{(1-\lambda)\lambda^{-4} + o(1)} = o(1) \quad a.s..
$$

Moreover, $\mathbb{E}\hat{b} = 0$. In the out-of control case of situation 3, we have

$$
\begin{aligned}
\hat{b} &= \frac{l_n U_n - L_n T_n}{(1-\lambda)\lambda^{-4} + o(1)} \\
&= \frac{\lambda^4 + o(1)}{(1-\lambda)}\left(\sum_{i=1}^{n} (l_n i - L_n)(1-\lambda)^{n-i}\tilde{X}_i + \sum_{i=t_0+1}^{n} (l_n i - L_n)(1-\lambda)^{n-i}(i-t_0)\delta\right) \\
&= \frac{\lambda^4 + o(1)}{(1-\lambda)}\left(\sum_{i=t_0+1}^{n} (l_n i - L_n)(1-\lambda)^{n-i}(i-t_0)\delta\right) + o(1) \ a.s.
\end{aligned}
$$

by using Lemmas 3 and 4, where $\tilde{X}_i = X_i - \mu_0$ for $i \le t_0$, and $\tilde{X}_i = X_i - \mu_0 - (i-t_0)\delta$ for $i > t_0$. Further

$$
\hat{b} = \frac{\lambda^4 + o(1)}{(1-\lambda)}\left(\frac{1-\lambda}{\lambda^4}\delta + o(1)\right) + o(1) = \delta + o(1) \quad a.s.
$$

and in a similar way,

$$
\begin{aligned}
\mathbb{E}\hat{b} &= \frac{\lambda^4(1+o(1))}{1-\lambda}\left(\sum_{i=1}^{n} (l_n i - L_n)(1-\lambda)^{n-i}\mu_0 + \sum_{i=t_0+1}^{n} (l_n i - L_n)(1-\lambda)^{n-i}(i-t_0)\delta\right) \\
&= \frac{\lambda^4 + o(1)}{1-\lambda}\left(\frac{1-\lambda}{\lambda^4}\delta + o(1)\right) = \delta + o(1)
\end{aligned}
$$

Obviously, $\mathbb{E}\hat{b} = 0$ holds in the in-control case $\delta = 0$.

In the general case, the variance of the estimator is given by

$$\begin{aligned}\mathrm{Var}(\hat{b}) &= \frac{1}{\left(L_n^2 - K_n l_n\right)^2} \sum_{i=1}^{n} \mathrm{Var}\left(\left(l_n i - L_n\right)(1-\lambda)^{n-i} X_i\right) \\ &= \frac{\sigma^2}{\left(L_n^2 - K_n l_n\right)^2} v_n,\end{aligned}$$

where

$$\begin{aligned}v_n &= \sum_{i=1}^{n} (1-\lambda)^{2n-2i} \left(L_n - l_n\, i\right)^2 \\ &= L_n^2 \bar{l}_n - 2 L_n l_n \bar{L}_n + l_n^2 \bar{K}_n.\end{aligned}$$

An application of Lemma 2 leads to

$$\begin{aligned}\bar{l}_n &= \frac{1}{\lambda(2-\lambda)} + o(n^{-2}), \quad \bar{L}_n = \frac{n}{\lambda(2-\lambda)} - \frac{(1-\lambda)^2}{\lambda^2(2-\lambda)^2} + o(1), \\ \bar{K}_n &= \frac{n^2}{\lambda(2-\lambda)} - \frac{2n(1-\lambda)^2}{\lambda^2(2-\lambda)^2} + \frac{(\lambda-1)^2\left(\lambda^2 - 2\lambda + 2\right)}{\lambda^3(2-\lambda)^3} + o(1)\end{aligned}$$

Further

$$\begin{aligned}v_n &= \left(\frac{\lambda(n+1)-1}{\lambda^2}\right)^2 \cdot \left(\frac{1}{\lambda(2-\lambda)} + o(n^{-2})\right) \\ &\quad -2\frac{1}{\lambda}\left(\frac{\lambda(n+1)-1}{\lambda^2}\right)\left(\frac{n}{\lambda(2-\lambda)} - \frac{(1-\lambda)^2}{\lambda^2(2-\lambda)^2} + o(n^{-1})\right) \\ &\quad +\frac{1}{\lambda^2}\left(\frac{n^2}{\lambda(2-\lambda)} - \frac{2n(1-\lambda)^2}{\lambda^2(2-\lambda)^2} + \frac{(\lambda-1)^2\left(\lambda^2-2\lambda+2\right)}{\lambda^3(2-\lambda)^3} + o(1)\right) + o(1) \\ &= \frac{2\left(1-\lambda\right)^2}{\lambda^5\left(2-\lambda\right)^3} + o(1).\end{aligned}$$

Hence

$$\begin{aligned}\mathrm{Var}(\hat{b}) &= \frac{\sigma^2}{\left(\frac{1-\lambda}{\lambda^4}\right)^2 + o(1)} \cdot \frac{2\left(1-\lambda\right)^2}{\lambda^5\left(2-\lambda\right)^3} + o(1) \\ &= \frac{2\sigma^2\lambda^3}{\left(2-\lambda\right)^3} + o(1),\end{aligned}$$

which implies assertion (i). $\boxdot$

4 References

Brown, R. (1956) Exponential smoothing for predicting demand. In: Philip Morris Records; Master Settlement Agreement, p. 15

Christiaanse, W. (1971). "Short-Term Load Forecasting Using General Exponential Smoothing". In: IEEE Transactions on Power Apparatus and Systems PAS-90.2, pp. 900–911.

Gardner, E. S., and McKenzie, E. (1985). Forecasting Trends in Time Series. *Management Science* 31, no. 10 : 1237–46. http://www.jstor.org/stable/2631713.

Luxenberg, E; Boyd, S (2024). Exponentially Weighted Moving Models. arXiv preprint

Myers, R.H.; Montgomery, R.H.; Vining.G.G.; Robinson, T.J. (2010) Generalized Linear Models, With Applications in Engineering and the Sciences. 2nd Edition. J. Wiley

####additional material

$$\lambda \sum_{i=1}^{n} (1-\lambda)^{n-i} \left(X_i - a - bi\right)^2 \quad \longrightarrow \quad \min$$

$$\Phi(a,b) \quad = \quad \sum_{i=1}^{n} (1-\lambda)^{n-i} \left(X_i - a - bi\right)^2 \longrightarrow \min$$

$$\begin{aligned}
\Phi_a &= -2 \sum_{i=1}^{n} (1-\lambda)^{n-i} \left(X_i - a - bi\right) = 0 \\
\Phi_a &= -2 \sum_{i=1}^{n} i (1-\lambda)^{n-i} \left(X_i - a - bi\right) = 0
\end{aligned}$$

$$\begin{aligned}
T_n &= \sum_{i=1}^{n} (1-\lambda)^{n-i} X_i, \; U_n = \sum_{i=1}^{n} i (1-\lambda)^{n-i} X_i \\
l_n &= \sum_{i=1}^{n} (1-\lambda)^{n-i}, \; L_n = \sum_{i=1}^{n} i (1-\lambda)^{n-i}, \; K_n = \sum_{i=1}^{n} i^2 (1-\lambda)^{n-i} \\
\bar{l}_n &= \sum_{i=1}^{n} (1-\lambda)^{2n-2i}, \; \bar{L}_n = \sum_{i=1}^{n} i (1-\lambda)^{2n-2i}, \; \bar{K}_n = \sum_{i=1}^{n} i^2 (1-\lambda)^{2n-2i}
\end{aligned}$$

$$\begin{aligned}
T_{n+1} &= \sum_{i=1}^{n+1} (1-\lambda)^{n+1-i} X_i = (1-\lambda) T_n + X_i \\
U_{n+1} &= \sum_{i=1}^{n+1} i (1-\lambda)^{n+1-i} X_i = (1-\lambda) U_n + (n+1) sd0 * X_i
\end{aligned}$$

$\Longrightarrow$

$$\begin{aligned}
T &= al + bL \\
U &= aL + bK
\end{aligned}$$

$$\begin{aligned}
TL &= alL + bL^2 \\
Ul &= aLl + bKl
\end{aligned}$$

$$
\begin{aligned}
\hat{a} &= \frac{1}{Kl - L^2}\left(KT - LU\right), \\
\hat{b} &= \frac{1}{L^2 - Kl}\left(LT - lU\right) = \frac{1}{L^2 - Kl}\sum_{i=1}^{n}(1-\lambda)^{n-i}\left(L - l\ i\right)X_i
\end{aligned}
$$

By Khintchine-Kolmogorov convergence theorem, the sums

$$
\sum_{i=0}^{n-1}(1-\lambda)^i W_i, \quad \sum_{i=0}^{n-1} i(1-\lambda)^i W_i
$$

converge $a.s.$, since the series $\sum_{j=1}^{\infty}\mathrm{Var}((1-\lambda)^j W_j) < +\infty, \sum_{j=1}^{\infty}\mathrm{Var}(j(1-\lambda)^j W_j) < +\infty$ are finite. Hence

$$
\begin{aligned}
\sum_{i=0}^{n-1}(1-\lambda)^i\left(X_{n-i} - \mu_0\right) &\to Z_1, \quad \sum_{i=0}^{n-1} i(1-\lambda)^i\left(X_{n-i} - \mu_0\right) \to Z_2, \\
\sum_{i=0}^{n-1}\left(\frac{1-\lambda}{\lambda^2} - i\lambda^{-1}\right)(1-\lambda)^i\left(X_{n-i} - \mu_0\right) &\to Z_3
\end{aligned}
$$

with random variables Z_1, Z_2. Furthermore, in the in-control case, we obtain

$$
\begin{aligned}
l_n U_n - L_n T_n &= \sum_{i=0}^{n-1}\left(l_n(n-i) - L_n\right)(1-\lambda)^i X_{n-i} \\
&= \sum_{i=0}^{n-1}\left(\frac{1-\lambda}{\lambda^2} - i\lambda^{-1}\right)(1-\lambda)^i \mu_0 + o(1) \\
&\quad + \sum_{i=0}^{n-1}\left(\frac{1-\lambda}{\lambda^2} - i\lambda^{-1}\right)(1-\lambda)^i\left(X_{n-i} - \mu_0\right) \\
&\quad + o(1)\sum_{i=0}^{n-1}(1-\lambda)^i\left(X_{n-i} - \mu_0\right) - o(1)\sum_{i=0}^{n-1} i(1-\lambda)^i\left(X_{n-i} - \mu_0\right) \\
&= 0 + Z_3 + o(1)\ a.s.
\end{aligned}
$$