ON EWMA PROCEDURE FOR DETECTION OF A CHANGE IN OBSERVATIONS VIA MARTINGALE APPROACH

Saowanit Sukparungsee*, Alexander Novikov

Department of Mathematical Sciences,
University of Technology, Sydney,
P.O. Box 123, Broadway, Sydney, NSW 2007, Australia

ABSTRACT

Using martingale technique, we present analytic approximations and exact lower bounds for the expectation of the first passage times of an Exponentially Weighted Moving Average (EWMA) procedure used for monitoring changes in distributions. Based on these results, a simple numerical procedure for finding optimal parameters of EWMA for small changes in the means of observation processes is established.

KEYWORDS: EWMA, Martingale Approach, First Passage Time, AR (1) Process, Average Run Length, Average Delay, Overshoot.

1. INTRODUCTION

Statistical process control (SPC) charts are widely used for monitoring and improving quality in manufacturing and industrial statistics and as well as in finance, medicine, epidemiology, environmental statistics, and other fields of applications. It is generally supposed that a mean of independent observations is to be sustained at its target value but in reality, this mean could change at any unanticipated time. To detect this change as soon as possible, one needs to apply statistical techniques and constrains. There could be many different settings for such constraints but the most used ones are: a mean of false alarm time (or, Average Run Length (ARL)) should be sufficiently large if the observation process is in control and a mean of average delay time (AD) should be small if the observation process is out of control. Of course, there must be a trade-off between these two conflicting requirements.

In practice the so-called Shewhart $\overline{x}$ (mean) charts are still mostly used, though it is known that they are not efficient in the monitoring of small changes in the means. The so-called Cumulative Sums (CUSUM) and Exponentially Weighted Moving Average (EWMA) charts are known to be essentially more sensitive to the detection of small changes. Unfortunately, the analysis of CUSUM and EWMA charts is much more complicated compared with Shewhart charts. CUSUM is known as an efficient tool (see e.g. Lucas and Saccucci [5]) but EWMA charts are inherently simpler than CUSUM and are believed to be more robust with respect to the assumptions compared with CUSUM.

In this paper we present some new analytic tools for finding bounds and approximations for ARL and AD for both one and two-sided EWMA procedures. A derivation of these bounds and approximations is based on a martingale approach (see some details in Appendix). In Section 2 we review the EWMA procedure and its properties. We compare our results with Monte Carlo simulation and results with Crowder [3], Lucas and Saccucci [5] in Section 3.

* Corresponding author. Tel: +61-2-9514-2256. Fax: +61-2-9514-2260
E-mail: Saowanit.Sukparungsee@student.uts.edu.au

2. THE EWMA PROCEDURE FOR MONITORING PROCESS MEAN AND ITS PROPERTIES.

Let $X_1, X_2, ..., X_t, ...$ be observed independent random variables. The martingale approach can be used for different distributions but we restrict out attention in this paper to the most important case of normally distributed random variables. The model we study is as follows:

$$\begin{cases} X_t \sim N(\mu_0, \sigma^2) & ; t = 1, 2, ..., \theta - 1 \\ X_t \sim N(\mu_1, \sigma^2) & ; t = \theta, \theta + 1, ... \;, \quad \mu_1 \neq \mu_0 \end{cases}$$

We use notation $\theta = \infty$ for the case when there is no change in the distribution of observed data. Note that if $\theta = 1$ then the change occurs at the very beginning. Results for the continuous time case $(t \in [0, \infty))$ and other distributions (including the exponential one) are planned for a future paper.

2.1 The EWMA procedure.

The Exponentially Weighted Moving Average (EWMA) statistics for the discrete time case has the following form

$$Z_t \;=\; \lambda(X_t - \mu_0) + (1-\lambda)Z_{t-1}, \qquad 0 < \lambda < 1, \qquad (1)$$

where λ is a weighting factor for previous observations. When $\lambda = 1$, it is actually the Shewhart $\bar{x}$ chart. Without loss of generality, we can assume that the initial value mean $\mu_0 = 0$ and variance of X is one. If the anticipated shift in the mean value is positive (that is $\mu_1 > 0$), then we take the decision that the process is out of control when for the first time $Z_t > H$. In another words, we use the stopping time for one-sided case:

$$\tau_1 = \inf\{t \in N : Z_t > H\}. \qquad (2)$$

If the sign of change in the means is defined in advance, we use the stopping time for two-sided case:

$$\tau_2 = \inf\{t \in N : |Z_t| > H\}. \qquad (3)$$

Note that, traditionally, $H = L\sqrt{\frac{\lambda}{2-\lambda}}$, L is a parameter to be chosen.

2.2 The properties of EWMA procedure

As mentioned earlier, we have to find a balance between the quick detection of process changes and at the same time to keep enough large false alarm time (ARL). To alleviate these conflicts, a possible trade-off is to choose the level L such that ARL is not less then a prescribed number T that is

$$ARL \equiv E(\tau) \geq T, \; |(\theta = \infty) \qquad (4)$$

and minimize AD as a function of parameters λ and L:

$$AD \equiv E(\tau - \theta | \tau \geq \theta) \;\rightarrow \min, \quad |1 \leq \theta < \infty. \qquad (5)$$

where τ is the stopping time and θ is a moment of change-point.

Our goal is to find the optimal combination of parameter (λ^*, L^*) which satisfies to these criterions.

Though the constraint (5) depends on the parameter θ, some empirical studies (see e.g. (5)) demonstrated that this dependence is not very essential. By this reason and in order to be able to compare our results with previous ones we assumed that in (5) $\theta = 1$ (that indicates the change point occurs from the beginning).

Crowder [3] used a system of Fredholm integral equation for numerical calculations of ARL and AD. Next, Brook and Evans [2] used the Markov Chain approach to approximate these characteristics. Lucas and Saccucci [5] intensively studied the different pairs (λ, L) to find AD for different magnitudes of change in mean process; we shall use their results for comparison.

2.3 The closed-form formula for the ARL and AD.

Proposition 1.

1). For one-sided EWMA:

$$ARL_1 = E(\tau_1) = \frac{1}{|\ln(1-\lambda)|}\int_0^\infty u^{-1}\left(Ee^{uZ_{\tau_1}} - 1\right)e^{\frac{-\lambda u^2}{4-2\lambda}}du. \quad (6)$$

$$AD_1 = E(\tau_1) = \frac{1}{|\ln(1-\lambda)|}\int_0^\infty u^{-1}\left(Ee^{uZ_{\tau_1}} - 1\right)e^{-\mu u-\frac{\lambda u^2}{4-2\lambda}}du., \quad \mu \geq 0 \quad (7)$$

where $Z_{\tau_1} = H + \chi^{(1)}$, $\chi^{(1)}$ is *an overshoot of one-sided EWMA statistics.*

2). For two-sided EWMA:

$$ARL_2 = E(\tau_2) = \frac{1}{|\ln(1-\lambda)|}\int_0^\infty u^{-1}\left(ECosh(uZ_{\tau_2}) - 1\right)e^{\frac{-\lambda u^2}{4-2\lambda}}du. \quad (8)$$

$$AD_2 = E(\tau_2) = \frac{1}{|\ln(1-\lambda)|}\int_0^\infty u^{-1}\left(Ee^{uZ_{\tau_2}} - 1\right)e^{-\mu u-\frac{\lambda u^2}{4-2\lambda}}du., \quad \mu \geq 0 \quad (9)$$

where $|Z_{\tau_2}| = H + \chi^{(2)}$, $\chi^{(2)}$ is *an overshoot of two-sided EWMA statistics,*

Formulas (6) and (8) contain the overshoot $\chi^{(i)}$ whose distribution is unknown. However, taking into account the fact that the overshoot is nonnegative we promptly get the following explicit lower-bound for ARL and AD of the one-sided procedure:

$$E[\tau_1] \geq g(H,\mu) = \frac{1}{|\ln(1-\lambda)|}\int_0^\infty u^{-1}\left(e^{uH} - 1\right)e^{-\mu u-\frac{\lambda u^2}{4-2\lambda}}du., \quad \mu \geq 0 \quad (10)$$

For the two-sided procedure we have by the same reasoning the following bound:

$$ARL_2 \geq f(H),\ f(H) = \frac{1}{|\ln(1-\lambda)|}\int_0^\infty u^{-1}\left[Cosh(uH) - 1\right]e^{\frac{-\lambda u^2}{4-2\lambda}}du. \quad (11)$$

To study the performance of the two-sided EWMA when the process is out control we may assume that $\mu > 0$. Note that the probability of the crossing lower bound, that is $P(Z_{\tau_2} < -H)$, should be close to zero for this case (at least, for not very small μ). This suggests to approximate that

$$AD_2 \cong g(H + \lambda C, \mu) \quad (12)$$

where we use notation from (10) and C is a constant as overshoot.

The lower bound (10) and (11) can be expressed in terms of standard special functions or easily combined numerically using the package "Mathematica". To make this lower bound more accurate, we propose to use the following approximation

$$Ee^{u\chi^{(i)}} \approx e^{u\lambda C} \quad (13)$$

where the constant C can be obtained this approximation by using Monte Carlo simulation results. In fact, we suggest finding a *corrected diffusion approximation* for the level $H = L\sqrt{\frac{\lambda}{2-\lambda}}$. Note that the very similar approximation is often used in many other problems of sequential analysis, see e.g. [12]. The theoretical justification for this approximation is based on the fact that for small values of λ, EWMA statistics are closed to a random walk and the distribution of the overshoot of random walks does not essentially dependent on level H (at least, for large H). Furthermore, it well known that the constant C for random walks (that is for $\lambda = 0$) is approximately equal to $-\frac{\varsigma(1/2)}{\sqrt{2\pi}} = 0.583$, where $\varsigma(x)$ is the Riemann zeta function (see details in [4] and [12]). Of course, this value of C should be used only as a first approximation. A more accurate approximation could be found with a Monte Carlo simulation and non-linear least-square fitting.

We suggest the following procedure for obtaining numerical results for two-sided EWMA:

1. Find the constant C for ARL using a Monte Carlo simulation for some given levels L using approximation (13) and non-linear fitting. As a result obtains an approximation in the form $ARL_2 \approx f(H+\lambda C)$

2. Find the constant C for AD with Monte Carlo simulation for given change μ and level L. As a result gets an approximation in the form $AD_2 \approx g(H+\lambda C, \mu)$

3. Find a pair of optimal parameter (λ^*, L^*) for given change μ.

3. NUMERICAL RESULTS.

We compare lower bound (11) and corrected lower bound by (13) with approximations of Crowder [3] and Lucas and Saccucci [5] and Monte Carlo simulation method (with 10^6 trials) by showing a percentage difference in Table 1 that our approximations ARL are fairly well. To find the constant C we always use the results of Monte Carlo simulation at the level $L = 2.0$ that often been suggested in the literature, see e.g. [5].

Table1. A comparison of ARL.

λ	L	C	Lower -Bound (martingale) using (11) (1)	Corrected Lower-Bound using (8) (2)	ARL (Lucas & Saccucci 1990) (3)	ARL (Crowder 1987) (4)	ARL (Monte Carlo) (5)	Percentage Difference (2) - (5)
0.01	1.00	0.583	59.28	**71.67**	-	60.11	71.90 (0.06)	**-0.32**
	2.00		447.91	**526.98**	-	453.13	527.02 (0.49)	**-0.01**
	3.00		4236.14	**5282.0**	-	-	5288.46 (5.14)	**-0.12**
0.03	1.00	0.589	19.56	**27.07**	-	26.43	27.37 (0.02)	**-1.11**
	2.00		147.79	**196.46**	-	192.16	196.46 (0.18)	**0.00**
	2.437		363	**499.21**	500	-	499.33 (0.48)	**-0.02**
	2.989		1357.79	**1999.31**	2000	-	2000.00 (1.98)	**-0.03**
	3.00		1397.76	**2061.38**	-	-	2062.34 (2.04)	**-0.05**
0.05	1.00	0.597	11.61	**17.65**	-	17.12	17.89 (0.02)	**-1.36**
	2.00		87.76	**127.33**	-	127.53	127.36 (0.12)	**-0.02**
	2.615		321.05	**500.29**	500	-	499.45 (0.49)	**0.17**
	3.00		830.02	**1381.99**	-	1379.35	1379.39 (1.36)	**0.19**
0.07	1.00	0.604	8.21	**13.45**	-	13.31	13.70 (0.01)	**-1.86**
	2.00		62.03	**96.78**		-	96.78 (0.09)	**0.00**
	2.015		63.89	**99.81**	100	-	99.83 (0.09)	**-0.02**
	3.00		586.66	**1080.97**	-	1065.81	1075.61 (1.06)	**0.50**
0.10	1.00	0.613	3.67	**10.18**	-	-	10.43 (0.01)	**-2.46**
	2.00		27.70	**73.18**	-	73.28	73.20 (0.07)	**-0.03**
	3.00		404.09	**848.40**	-	842.15	841.95 (0.83)	**0.76**
	3.059		471.80	**1009.31**	1000	-	998.01 (0.99)	**1.12**
	3.283		888.47	**2018.41**	2000	-	1994.56 (1.98)	**1.18**

Note: the standard error is showed in parentheses.

Table 2. AD for two-sided EWMA procedure with fixed ARL = 5000.

λ	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.1
L	2.99	3.21	3.32	3.40	3.46	3.49	3.52	3.55	3.58	3.59
C	0.50	0.51	0.50	0.48	0.49	0.51	0.51	0.51	0.51	0.53
Shift (μ)										
0.10	**695.3**	963.7	1202.5	1418.6	1628.5	1828.3	1972.8	2134.0	2314.3	2432.2
0.20	**204.8**	242.8	294.1	351.4	414.2	480.2	538.4	602.7	673.9	732.1
0.30	**108.9**	111.1	122.3	137.8	156.8	178.3	199.2	222.9	249.9	274.0
0.40	73.0	**68.3**	69.8	74.0	80.2	87.8	95.6	104.9	115.7	125.9
0.50	54.7	48.7	**47.5**	48.2	50.1	52.9	56.0	59.8	64.5	68.9
0.60	43.7	37.7	35.7	**35.1**	35.5	36.5	37.6	39.3	41.4	43.4
0.70	36.4	30.7	28.5	27.5	**27.2**	27.4	27.7	28.4	29.4	30.3
0.80	31.1	25.9	23.7	22.5	22.0	**21.8**	**21.8**	22.0	22.4	22.8
0.90	27.2	22.4	20.2	19.1	18.4	18.1	17.9	**17.8**	18.0	18.1
1.00	24.2	19.7	17.7	16.5	15.8	15.4	15.1	**15.0**	**15.0**	**15.0**

The last table allows us to find optimal combinations of EWMA parameters for each magnitude of change μ at given level ARL. For example, when ARL = 5000 and $\mu = 0.5$, our approximation found pairs of optimal parameter $(\lambda = 0.03, \mathrm{L} = 3.32)$ with $AD = 47.5$ compared to Lucas and Saccucci [5] $(\lambda = 0.03, \mathrm{L} = 3.299)$ with $AD = 47.7$. For $\mu = 1$ our results is $(\lambda = [0.08, 0.1], \mathrm{L} = [3.55, 3.59])$ with $AD = 15.0$ compared to [5] $(\lambda = 0.09, \mathrm{L} = 3.538)$ with $AD = 15.2$. See also Figure 1. as below. Furthermore, we found that optimal EWMA parameters with fixed values for ARL = 100, 300, 500, 1000 and 2000 are also very close in Lucas and

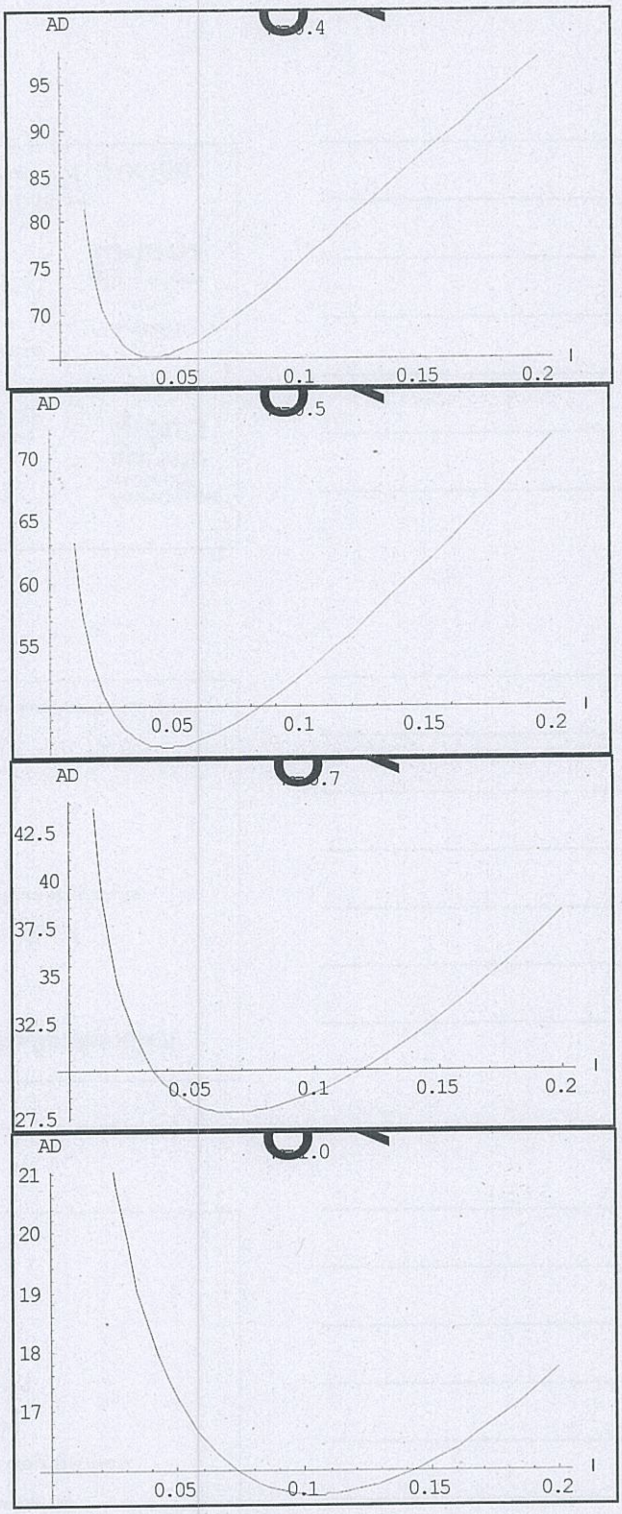

Figure 1. Curves of AD for different magnitudes of change.

4. CONCLUSIONS

We have suggested a new semi-analytical martingale approximation for studying properties of the EWMA procedure. This approach leads to closed-form formulas for ARL and AD that are easily calculated and can be computed by using the simple code of standard packages such a "Mathematica". The results from these closed-form formulas were compared favourably with those obtained by other methods.

5. ACKNOWLEDGEMENTS

We thank Dr. B. Ergashev for presenting the paper [1]. We also express our appreciations to Prof. T. Moon, Dr. N. Kordzakhia and Mr. B. Stephenson for helpful comments and providing assistance in the preparation of this paper. This paper was partially supported by Thai Government scholarship.

REFERENCES

[1] Ergashev, B.A. **2003** On a CAPM Monitoring Based on the Ewma Procedure. *Working Paper. Presented at 9-th International Conference of the Society for Computational Economics and Finance*, Seattle, July 11-13.

[2] Brook, D. and Evans, D.A. **1972** An Approach to the Probability Distribution of Cusum Run Length, *Biometrika 59*, 539-548.

[3] Crowder, S. V. **1987** A Simple Method for Studying Run-Length Distributions of Exponentially Weighted Moving Average Charts, *Technometrics, 29*, 401-407.

[4] J. Chang and Y. Peres. **1997** Ladder Heights, Gaussian Random Walks, and the Riemann Zeta Function, *Annals of Probability, 25*, 787-802.

[5] Lucas, J. M. and Saccucci, M. S. **1990** Exponentially Weighted Moving Average Control Schemes: Properties and Enhancements, With Discussion. *Technometrics, 32*, 1-29.

[6] Novikov, A. A. **1990** On the First Passage Time of an Autoregressive Process Over a Level and Application to a "Disorder Problem". *Theory of Probability and its applications, 35*, 269-279.

[7] Roberts, S. W. **1959** Control Chart Tests Based on Geometric Moving Average. *Technometrics, 1*, 239-250.

[8] Srivastava, M. S. and Wu, Y. **1993** Comparison of EWMA, CUSUM and Shiryayev-Roberts Procedures for Detecting a Shift in The Mean. *Annals of Statistics, 21*, 645-670.

[9] A.N. Shiryayev, **1996** *Probability, 2nd Edition, Springer.*, New York: Springer.

[10] Montgomery, D.C., **2005** *Introduction to Statistical Quality Control.*: New York; Chichester: Wiley.

[11] Basseville, M. and Nikiforov, I. **1993** *Detection of Abrupt Changes: Theory and Applications.*, Englewood Cliffs: Prentice-Hall.

[12] Siegmund, D. **1985** *Sequential Analysis: Tests and Confidence Intervals.*. New York: Springer - Verlag.

Appendix: The Closed-Form Formulas for ARLs and ADs

AR (1) process is defined by the recursive equation

$$Y_t = \beta Y_{t-1} + \xi_t, \quad t \in 0,1,2,...., \qquad Y_0 = y \tag{14}$$

where β is a constant, $0 < \beta < 1$. Note that EWMA statistic Z_t coincides with Y_t for a particular case of $\beta = 1 - \lambda$ and independent random variables $\xi_t = \lambda X_t$.

Expectation of First Passage Times for one-sided EWMA.

It was shown in [6] that if X_t are independent and normally distributed $N(\mu, \sigma^2)$ random variables then the process

$$H(Z_t) = \int_0^{\infty} u^{-1}[\exp(uZ_t) - \exp(uz)]\exp\left\{-\mu u - \frac{\lambda^2 u^2}{4-2\lambda}\right\}du - t\log\left(\frac{1}{1-\lambda}\right) \quad (15)$$

is a martingale (with respect to the natural filtration). By the optional stopping theorem [9], we get $E[H(Z_\tau)] = 0$ for any bounded stopping τ. This implies that for the case of EWMA

$$E(\tau) = \frac{1}{|\ln(1-\lambda)|}\int_0^{\infty} u^{-1} E\left[e^{uZ_\tau} - e^{uZ_0}\right]e^{-\mu u - \frac{\lambda u^2}{4-2\lambda}}du. \quad (16)$$

Using arguments similar to the ones used in [6] it can be shown that (16) still holds for unbounded stopping times τ_1 and τ_2 defined above. Given the initial value $Z_0 = 0$ we have

$$E(\tau_1) = \frac{1}{|\ln(1-\lambda)|}\int_0^{\infty} u^{-1}\left(Ee^{uZ_{\tau_1}} - 1\right)e^{-\mu u - \frac{\lambda u^2}{4-2\lambda}}du \quad (17)$$

Expectation of First Passage Time for two-sided case EWMA.

First, consider the symmetric case when $\mu = 0$ and $Z = 0$. Then due to the symmetry of the normal distribution both processes $H(Z_t)$ and $H(-Z_t)$ are martingales and so the process

$$\frac{H(Z_t) + H(-Z_t)}{2} = \int_0^{\infty} u^{-1}[Cosh(uZ_t) - 1]\exp\left\{-\frac{\lambda^2 u^2}{4-2\lambda}\right\}du - t\log\left(\frac{1}{1-\lambda}\right)$$

is a martingale as well. As above for the one-sided EWMA, we get that for any bounded stopping τ

$$E[H(Z_\tau) + H(-Z_\tau)] = 0$$

Again, using an analytical technique similar that in [6] it can be shown that the last identity does hold for the unbounded stopping τ_2. This implies that for the case of EWMA

$$ARL_2 = E(\tau_2) = \frac{1}{|\ln(1-\lambda)|}\int_0^{\infty} u^{-1}\left(ECosh(Z_{\tau_2}) - 1\right)e^{-\frac{\lambda u^2}{4-2\lambda}}du, \ (\mu = 0) \quad (18)$$

where $|Z_{\tau_2}| = H + \chi^{(2)} > H$. As $Cosh(Z_{\tau_2}) > Cosh(H)$ we get the lower bound (11).

We can obtain the approximation (13) by using the identity (16) with $\tau = \tau_2$. Then we have (see also formula (9) above)

$$E(\tau_2) = \frac{1}{|\ln(1-\lambda)|}\left[\begin{array}{l}\int_0^{\infty} E[I(Z_{\tau_2} > H)u^{-1}(e^{uZ_{\tau_2}} - 1)]e^{-\mu u - \frac{\lambda u^2}{4-2\lambda}}du \\ + \int_0^{\infty} E[I(Z_{\tau_2} < -H)u^{-1}(e^{-uZ_{\tau_2}} - 1)]e^{-\mu u - \frac{\lambda u^2}{4-2\lambda}}du.\end{array}\right] \quad (19)$$

where $I(A)$ is an indicator function .

As $P(Z_{\tau_2} < -H)$ is small (see discussion above) we may neglect the last term in (19) and which then leads to (9).