Methodol Comput Appl Probab (2019) 21:889–906
https://doi.org/10.1007/s11009-018-9622-7

A Double Recursion for Calculating Moments of the Truncated Normal Distribution and its Connection to Change Detection

Moshe Pollak[1] · **Michal Shauly-Aharonov**[1]

Received: 7 August 2016 / Revised: 15 February 2018 /
Accepted: 18 February 2018 / Published online: 2 March 2018

Abstract The integral $\int_0^\infty x^m e^{-\frac{1}{2}(x-a)^2} dx$ appears in likelihood ratios used to detect a change in the parameters of a normal distribution. As part of the m^{th} moment of a truncated normal distribution, this integral is known to satisfy a recursion relation, which has been used to calculate the first four moments of a truncated normal. Use of higher order moments was rare. In more recent times, this integral has found important applications in methods of changepoint detection, with m going up to the thousands. The standard recursion formula entails numbers whose values grow quickly with m, rendering a low cap on computational feasibility. We present various aspects of dealing with the computational issues: asymptotics, recursion and approximation. We provide an example in a changepoint detection setting.

Keywords Changepoint · On-line · Shiryaev–Roberts · Surveillance

Mathematics Subject Classification (2010) 62L10 · 62E15 · 60E05

1 Introduction

The objective of the present paper is to propose a computationally feasible way to evaluate the integral $\int_0^\infty x^m e^{-\frac{1}{2}(x-a)^2} dx$ even when m is very large.

Obviously, the integral $\int_0^\infty x^m e^{-\frac{1}{2}(x-a)^2} dx$ is a constant multiple of the m^{th} order moment of a truncated normal distribution. In the past, interest has been mostly relegated to

✉ Moshe Pollak
moshe.pollak@mail.huji.ac.il

Michal Shauly-Aharonov
michal.shauly@mail.huji.ac.il

1 Department of Statistics, The Hebrew University of Jerusalem, 91905 Jerusalem, Israel

low values of m (see Remark 6 in Appendix D). When m is large, precise evaluation of the integral becomes difficult. The motivating case for its evaluation when m is large is a context of on-line monitoring for a change of parameter(s) of a Normal(μ, σ^2) distribution based on a sequence of (normally-distributed) independent observations $X_1, X_2, \ldots$ where both parameters are initially unknown. Classical methods for detecting a change are Shewhart, Cusum and Shiryaev-Roberts. These methods invariably assume that the baseline parameters μ and σ are known at the onset of surveillance. Often, they are not. For example, if regarding a hitherto unexplored phenomenon (that can be assumed by, say, the central limit theorem to yield normally distributed observations), neither of the parameters is initially known. In such instances, the classical fix is to use a learning sample to estimate unknown parameters and plug them into the classical procedures. The downside of this is that, even if a learning sample is available, operating characteristics of these procedures are notoriously sensitive to misspecified parameters (van Dobben de Bruyn 1968) so a huge learning sample is needed to obtain a good handle on operating characteristics (such as average run length to false alarm). A more modern approach is the use of invariant statistics as the basis for surveillance instead of the original observations (cf. Quesenberry 1991, for Shewhart; Pollak and Siegmund 1991, for Cusum and Shiryaev-Roberts; Gordon and Pollak 1995, for nonparametrics).

If one defines $Y_i = (X_i - \bar{X}_{i-1})\sqrt{\frac{i-1}{i}}$ (where $\bar{X}_{i-1} = \frac{1}{i-1}\sum_{j=1}^{i-1} X_j$) and $Z_i = \frac{Y_i}{|Y_2|}$, then the distribution of the sequence $Z_2, Z_3, \ldots$ is invariant with respect to μ and σ (hence can be calculated explicitly, since it is the same for all μ, σ). Therefore a detection scheme can be based on a respective sequence of likelihood ratios. The fact that invariant statistics are dependent is of no hindrance, as long as the joint distribution is known when the process is in control. The integral $\int_0^\infty x^m e^{-\frac{1}{2}(x-a)^2} dx$ appears in the joint density, both when the process is in control and when it is not (see Appendix C). After observing $n \geq 2$ observations, $m = n - 2$. So, as the sample size grows, m becomes large. The integral crops up in various scenarios of change detection in which the baseline parameters are unknown (cf. Pollak et al. 1993) and in detecting a change in regression parameters when the residuals are normally distributed (cf. Krieger et al. 2003).

The usual recursion formula involved in calculation of the integral has the form (cf. Horrace 2015; Dhrymes 2005)

$$I_m(h) = -h^{m-1}\frac{\phi(h)}{\Phi(h)} + (m-1)I_{m-2}(h)$$

where ϕ and Φ are the standard normal pdf and cdf, respectively, and $I_m(h) = \frac{\int_{-\infty}^h t^m \phi(t)dt}{\Phi(h)}$. $I_m(h)$ can grow quite quickly with m, putting a relatively low cap on m that allows calculation on standard computers, especially since the right hand side has both a positive and a negative term .

In the changepoint context, Pollak et al. (1993) circumvented this when the sample size m is moderate by a recursion method. In a similar context, Krieger et al. (2003) employed a Taylor series approach. The latter approach works better for larger m when $a > 0$ (since it operates on the log scale), but both suffer from the aforementioned computational difficulty when $a < 0$.

The paper is organized as follows. We first present an asymptotic formula (Theorem 1). This formula is pivotal in suggesting a double recursion (8–10) that extends computational feasibility considerably. Armed with an asymptotic approximation and a computationally feasible recursion, we examine how quickly the asymptotic formula kicks in. To complete the picture, we restate the method used by Krieger et al. (2003). In two appendices, we provide a proof of the asymptotic formula and a MATLAB program for executing the recursions 8–10. In a third appendix we give an example of an application. We conclude with a number of remarks.

2 Asymptotics and an Approximation

We first present an asymptotic formula for $\frac{\int_0^\infty x^m e^{-\frac{1}{2}(x-a)^2}dx}{\int_0^\infty x^m e^{-\frac{1}{2}x^2dx}}$.

Theorem 1 *Let* $0 \leq \delta < 1$ *and* $\max\{\frac{2}{3}\delta, \frac{1}{2}\} < \eta < 1$. *For* $|a| = O\left(\sqrt{m}^\delta\right)$, *as* $m \to \infty$

$$\frac{\int_0^\infty x^m e^{-\frac{1}{2}(x-a)^2}dx}{\int_0^\infty x^m e^{-\frac{1}{2}x^2dx}} = e^{a\sqrt{m}-\frac{1}{4}a^2\left(1+O((a\sqrt{m})^{\eta-1})\right)}.$$

The integral $\int_0^\infty x^m e^{-\frac{1}{2}x^2}dx$ will play a central role in our calculations, so we evaluate it first. Transforming $y = -\frac{1}{2}x^2$ yields

$$\begin{aligned}\int_0^\infty x^m e^{-\frac{1}{2}x^2}dx &= \int_0^\infty (\sqrt{2y})^m e^{-y}\frac{1}{\sqrt{2}}y^{-\frac{1}{2}}dy \\ &= 2^{\frac{1}{2}(m-1)}\int_0^\infty y^{\frac{1}{2}(m-1)}e^{-y}dy = 2^{\frac{1}{2}(m-1)}\Gamma\left(\frac{m+1}{2}\right) \qquad (1)\end{aligned}$$

where Γ is the gamma function. By Stirling's approximation, for large values of m this is approximately

$$2^{\frac{1}{2}(m-1)}e^{-\frac{m+1}{2}}\left(\frac{m+1}{2}\right)^{\frac{m}{2}}\sqrt{2\pi}\left(1+O\left(\frac{1}{m}\right)\right) = e^{-\frac{m+1}{2}}(m+1)^{\frac{m}{2}}\sqrt{\pi}\left(1+O\left(\frac{1}{m}\right)\right), \qquad (2)$$

an expression that tends to ∞ as $m \to \infty$. Theorem 1 leads to the approximation

$$\int_0^\infty x^m e^{-\frac{1}{2}(x-a)^2}dx \sim \sqrt{\pi}e^{-\frac{m+1}{2}}(m+1)^{\frac{m}{2}}e^{a\sqrt{m}-\frac{1}{4}a^2}. \qquad (3)$$

Note that $|\int_0^c x^m e^{-\frac{1}{2}(x-a)^2}dx| < \int_0^{|c|} x^m dx = \frac{|c|^{m+1}}{m+1}$, which is negligible with respect to $\int_0^\infty x^m e^{-\frac{1}{2}x^2}dx$, so that also

$$\int_c^\infty x^m \frac{\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}(x-a)^2}}{1-\Phi(-c)}dx \sim \frac{1}{\Phi(c)}e^{-\frac{m+1}{2}}(m+1)^{\frac{m}{2}}\sqrt{\pi}e^{a\sqrt{m}-\frac{1}{4}a^2}. \qquad (4)$$

3 A Recursion

Although for the changepoint problem the integral of interest is related only to a special case ($c = 0$ in Eq. 5) of the truncated normal distribution, we present here the general case. The chief computational issue is dealing with the integral

$$g_{c,a}(m) = \int_c^\infty x^m e^{-\frac{1}{2}(x-a)^2}dx. \qquad (5)$$

Integration by parts yields

$$\begin{aligned} g_{c,a}(m) &= \int_c^\infty x^{m-1}(x-a+a)e^{-\frac{1}{2}(x-a)^2}dx \\ &= \int_c^\infty x^{m-1}(x-a)e^{-\frac{1}{2}(x-a)^2}dx + a g_{c,a}(m-1) \\ &= -x^{m-1}e^{-\frac{1}{2}(x-a)^2}|_c^\infty + (m-1)\int_c^\infty x^{m-2}e^{-\frac{1}{2}(x-a)^2}dx + a g_{c,a}(m-1) \\ &= c^{m-1}e^{-\frac{1}{2}(c-a)^2} + (m-1)g_{c,a}(m-2) + a g_{c,a}(m-1). \end{aligned} \tag{6}$$

This recursion works well when m is small, but breaks down for large values of m, as the expressions involved become very large, passing beyond limits of precision if m grows beyond a certain limit. To overcome this, we present the following recursion. The idea is to scale the problem so that objects do not grow too fast nor decrease too rapidly. The specific scaling is motivated by Theorem 1.

Note that $\int_0^\infty x^m e^{-\frac{1}{2}x^2}dx = g_{0,0}(m)$, which obeys the recursion

$$g_{0,0}(m) = (m-1)g_{0,0}(m-2) \tag{7}$$

and can be expressed analytically as in Eq. 1. Denote

$$\psi_{c,a}(m) = \frac{g_{c,a}(m)}{g_{0,0}(m)}e^{-a\sqrt{m}} \qquad \text{and} \qquad \xi_{c,a}(m) = \frac{g_{c,a}(m)}{g_{0,0}(m-1)}e^{-a\sqrt{m}}. \tag{8}$$

It follows that

$$\begin{aligned} \psi_{c,a}(m) &= \frac{a}{m-1}\xi_{c,a}(m-1)e^{-a(\sqrt{m}-\sqrt{m-1})} \\ &\quad + \psi_{c,a}(m-2)e^{-a(\sqrt{m}-\sqrt{m-2})} \\ &\quad + \frac{c^{m-1}e^{-\frac{(c-a)^2}{2}}e^{-a\sqrt{m}}}{2^{\frac{1}{2}(m-1)}\Gamma(\frac{m+1}{2})} \end{aligned} \tag{9}$$

and

$$\begin{aligned} \xi_{c,a}(m) &= a\psi_{c,a}(m-1)e^{-a(\sqrt{m}-\sqrt{m-1})} \\ &\quad + \frac{m-1}{m-2}\xi_{c,a}(m-2)e^{-a(\sqrt{m}-\sqrt{m-2})} \\ &\quad + \frac{c^{m-1}e^{-\frac{(c-a)^2}{2}}e^{-a\sqrt{m}}}{2^{\frac{1}{2}(m-2)}\Gamma(\frac{m}{2})}. \end{aligned} \tag{10}$$

Thus the pair $(\psi_{c,a}(m), \xi_{c,a}(m))$ depends recursively on $(\psi_{c,a}(m-1), \xi_{c,a}(m-1))$ and $(\psi_{c,a}(m-2), \xi_{c,a}(m-2))$ and two terms that are easily computed at each stage, both of which tend to 0 as $m \to \infty$. The recursion can be started after calculating $g_{0,0}(1)$, $g_{0,0}(2)$ and $g_{c,a}(1)$, $g_{c,a}(2)$. Now

$$g_{0,0}(0) = \int_0^\infty e^{-\frac{1}{2}x^2}dx = \sqrt{\frac{\pi}{2}}$$

$$g_{0,0}(1) = \int_0^\infty x e^{-\frac{1}{2}x^2}dx = 1$$

$$g_{0,0}(2) = \int_0^\infty x^2 e^{-\frac{1}{2}x^2}dx = \sqrt{\frac{\pi}{2}}$$

$$g_{c,a}(0) = \int_c^\infty e^{-\frac{1}{2}(x-a)^2} dx = \sqrt{2\pi}[1 - \Phi(c-a)]$$
$$g_{c,a}(1) = \int_c^\infty x e^{-\frac{1}{2}(x-a)^2} dx = e^{-\frac{1}{2}(c-a)^2} + a\sqrt{2\pi}[1 - \Phi(c-a)]$$
$$g_{c,a}(2) = \int_c^\infty x^2 e^{-\frac{1}{2}(x-a)^2} dx = (c+a)e^{-\frac{1}{2}(c-a)^2} + (a^2+1)\sqrt{2\pi}[1 - \Phi(c-a)]$$

so the two pairs $(\psi_{c,a}(1), \xi_{c,a}(1))$, $(\psi_{c,a}(2), \xi_{c,a}(2))$ can be calculated, and the recursion can be applied from $m = 3$ onwards. See Appendix B for a MATLAB program.

Finally, the m^{th} moment of the left-curtailed normal distribution can be calculated as

$$\frac{\frac{1}{\sqrt{2\pi}} \int_c^\infty x^m e^{-\frac{(x-a)^2}{2}} dx}{1 - \Phi(c-a)} = \frac{1}{\sqrt{2\pi}} \frac{\psi_{c,a}(m) e^{a\sqrt{m}} g_{0,0}(m)}{1 - \Phi(c-a)} = \frac{1}{\sqrt{2\pi}} \frac{\psi_{c,a}(m) e^{a\sqrt{m}} 2^{\frac{1}{2}(m-1)} \Gamma(\frac{m+1}{2})}{1 - \Phi(c-a)}.$$

4 Discussion

Theorem 1 implies the approximation

$$\log\left(\frac{\int_0^\infty x^m e^{-\frac{1}{2}(x-a)^2} dx}{\int_0^\infty x^m e^{-\frac{1}{2}x^2} dx}\right) \sim a\sqrt{m} - \frac{1}{4}a^2. \tag{11}$$

Figure 1 is a depiction of this. The accuracy of this approximation depends on the value of a; the closer $|a|$ is to 0 the better the approximation. This is borne out by Table 1. The values in the body of the table are lower bounds on m for which the discrepancy between true and approximate values does not exceed a given bound.

For calculating the values in Table 1, $\frac{\int_0^\infty x^m e^{-\frac{1}{2}(x-a)^2} dx}{\int_0^\infty x^m e^{-\frac{1}{2}x^2} dx}$ was calculated via the recursions (8)–(10) by the MATLAB program of Appendix B. Higher values of m required more than

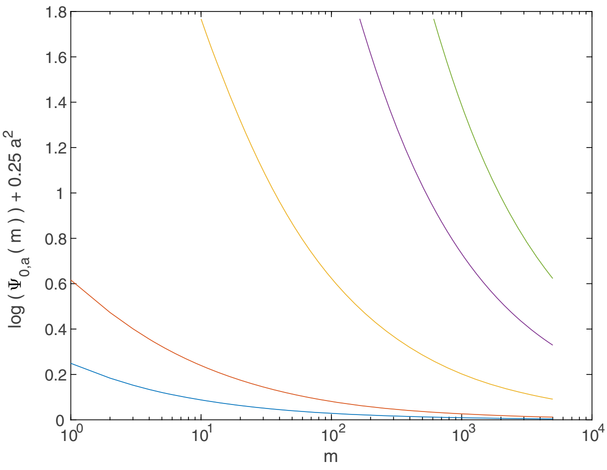

Fig. 1 $\log \psi_{0,a}(m) + \frac{1}{4}a^2$, for (clockwise) $a = 1$, $a = 2$, $a = 5$, $a = 8$, $a = 10$

Table 1 The smallest value of m with given bound on the absolute difference between $\log\left(\frac{\int_0^\infty x^m e^{-\frac{1}{2}(x-a)^2}dx}{\int_0^\infty x^m e^{-\frac{1}{2}x^2}dx}\right)$ and $\{a\sqrt{m} - \frac{1}{4}a^2\}$, for various values of a

a\bound	.05	.025	.01	.005
−2	288	1199		
−1	37	142	865	3431
−.5	8	29	178	685
.5	7	27	167	673
1	32	132	839	3378
2	268	1091	6895	27678
5	16620	66610	416787	1667776

double precision when $a < 0$ and were calculated with the aid of the Multiprecision Computing Toolbox (2016) for MATLAB. In double precision, for most practical purposes, the recursion (6) will produce inaccurate results for $m > 300$ when $a < 0$ and the double recursion (8)–(10) is not much better: the values of $\frac{\int_0^\infty x^m e^{-\frac{1}{2}(x-a)^2}dx}{\int_0^\infty x^m e^{-\frac{1}{2}x^2 dx}}$ become small very quickly as m grows; since the recursion has a negative component ($a < 0$), the cap on a computer's precision does not allow accuracy for large values of m. For $a > 0$ double precision suffices and computation time (with an Intel(R) Xeon(R) processor CPU E5-2609,@ 2.50 GHz) was mostly in fractions or single digits of a second (very high values of m took longer, but even 1667776 took only 75 seconds to obtain). When m becomes very large, the approximation suggested by Theorem 1 is a reasonable alternative to going through the recursions.

In monitoring schemes where the recursion may be applied, m is (approximately) the sample size. Since monitoring (for instance: heartbeat in an ICU) may go on for a long time, producing many observations, the sample size n may become very large. Although the recursions (8)–(10) take only a number of seconds to calculate, when monitoring a sequence such as described in Appendix C, one has to go through $O(n^3)$ calculations of $\psi_{0,\delta U}$ by the n^{th} observation; the n^{th} observation necessitates $n-1$ calculations of ψ to calculate the surveillance statistic (R_n), each of which is based on $n-2$ iterations of the recursion. This can get to be time consuming, especially if one is interested in doing simulations. A way to circumvent this is to construct a table of $\log(\psi)$ that can be used with interpolation for large sequences or simulations. When the sample size n becomes very large, the asymptotics stated in Theorem 1 kicks in, and can be used safely without a time-consuming calculation.

5 A Different Method of Calculation

Another way to evaluate

$$\frac{\int_0^\infty x^m e^{-\frac{1}{2}(x-a)^2}dx}{\int_0^\infty x^m e^{-\frac{1}{2}x^2}dx}$$

is given by Krieger et al. (2003). For the sake of completeness we provide it here. It should be noted that this method is not more successful than the (8)–(10) recursion method in calculating ψ when $a < 0$.

Recall that

$$\int_0^\infty x^m e^{-\frac{1}{2}x^2}dx = 2^{\frac{m-1}{2}}\Gamma\Big(\frac{m+1}{2}\Big)$$

and

$$\int_0^\infty x^m e^{-\frac{1}{2}(x-a)^2}dx = e^{-\frac{1}{2}a^2}\int_0^\infty e^{ax}x^m e^{-\frac{1}{2}x^2}dx = e^{-\frac{1}{2}a^2}\sum_{j=0}^{\infty}\frac{a^j}{j!}\int_0^\infty x^{m+j}e^{-\frac{1}{2}x^2}dx$$

so

$$\frac{\int_0^\infty x^m e^{-\frac{1}{2}(x-a)^2}dx}{\int_0^\infty x^m e^{-\frac{1}{2}x^2}dx} = e^{-\frac{1}{2}a^2}\sum_{j=0}^{\infty}\frac{a^j}{j!}2^{\frac{m+j-1}{2}}\frac{\Gamma(\frac{m+j+1}{2})}{2^{\frac{m-1}{2}}\Gamma(\frac{m+1}{2})} = e^{-\frac{1}{2}a^2}\sum_{j=0}^{\infty}\frac{(\sqrt{2}a)^j}{j!}\frac{\Gamma(\frac{m+j+1}{2})}{\Gamma(\frac{m+1}{2})}.$$

Fix an integer J such that $\sum_{j=J+1}^{\infty}\frac{(\sqrt{2}a)^j}{j!}\frac{\Gamma(\frac{m+j+1}{2})}{\Gamma(\frac{m+1}{2})}$ is negligible and define and calculate

$$\mu_j = j\log(\sqrt{2}a) + \log\left(\Gamma\Big(\frac{m+j+1}{2}\Big)\right) - \log\left(\Gamma\Big(\frac{m+1}{2}\Big)\right) - \log(j!)$$

$$\mu_{\max} = \max_{1\le j\le J}\mu_j$$

(J and $\mu_{\max}$ depend on a). Now calculate

$$\frac{\int_0^\infty x^m e^{-\frac{1}{2}(x-a)^2}dx}{\int_0^\infty x^m e^{-\frac{1}{2}x^2}dx} \sim e^{-\frac{1}{2}a^2}\left(\sum_{j=0}^{J}e^{\mu_j-\mu_{\max}}\right)e^{\mu_{\max}}.$$

(For negative values of a, separate the even n's from the odd ones, and handle each separately in a manner like the above.)

Note that for the general truncated normal distribution (i.e. $c \neq 0$) this approach will not work well if m is not large (since $\int_c^\infty e^{ax}x^m e^{-\frac{1}{2}x^2}dx$ is not expressable in terms of a complete Gamma function). However, as mentioned in Section 2, $\frac{\int_0^c x^m e^{-\frac{1}{2}(x-a)^2}dx}{\int_0^\infty x^m e^{-\frac{1}{2}x^2}dx}$ is negligible with respect to $\frac{\int_0^\infty x^m e^{-\frac{1}{2}(x-a)^2}dx}{\int_0^\infty x^m e^{-\frac{1}{2}x^2}dx}$ for large values of m, so also

$$\frac{\int_c^\infty x^m e^{-\frac{1}{2}(x-a)^2}dx}{\int_0^\infty x^m e^{-\frac{1}{2}x^2}dx} \sim e^{-\frac{1}{2}a^2}\left(\sum_{j=0}^{J}e^{\mu_j-\mu_{\max}}\right)e^{\mu_{\max}}.$$

Acknowledgements This work was supported by grant 1450/13 from the Israel Science Foundation and by the Marcy Bogen Chair of Statistics at the Department of Statistics, The Hebrew University of Jerusalem. The authors would like to thank the referees for their salient comments, which greatly improved the paper.

Appendix A: Proof of Theorem 1

The case $a = 0$ is trivial. In the following we assume $a \neq 0$. Recall that

$$\int_0^\infty x^m e^{-\frac{1}{2}x^2}dx = 2^{\frac{1}{2}(m-1)}\Gamma\Big(\frac{m+1}{2}\Big).$$

Hence by Stirling's approximation

$$\begin{aligned}\frac{\int_0^\infty x^{m+1}e^{-\frac{1}{2}x^2}dx}{\int_0^\infty x^m e^{-\frac{1}{2}x^2}dx} &= \frac{2^{\frac{1}{2}m}\Gamma(\frac{m+2}{2})}{2^{\frac{1}{2}(m-1)}\Gamma(\frac{m+1}{2})}\\ &= 2^{\frac{1}{2}}\frac{e^{-\frac{m+2}{2}}(\frac{m+2}{2})^{\frac{m+1}{2}}}{e^{-\frac{m+1}{2}}(\frac{m+1}{2})^{\frac{m}{2}}}\Big(1+O\Big(\frac{1}{m^2}\Big)\Big)\\ &= e^{-\frac{1}{2}}\Big(\frac{m+2}{m+1}\Big)^{(m+1)\frac{1}{2}}(m+1)^{\frac{1}{2}}\Big(1+O\Big(\frac{1}{m^2}\Big)\Big)\\ &= \sqrt{m+1}\Big(1+O\Big(\frac{1}{m}\Big)\Big)\end{aligned}$$

and so

$$\begin{aligned}e^{\frac{1}{2}a^2}\frac{\int_0^\infty x^m e^{-\frac{1}{2}(x-a)^2}dx}{\int_0^\infty x^m e^{-\frac{1}{2}x^2}dx} &= \frac{\int_0^\infty e^{ax}x^m e^{-\frac{1}{2}x^2}dx}{\int_0^\infty x^m e^{-\frac{1}{2}x^2}dx}\\ &= \sum_{j=0}^\infty \frac{\int_0^\infty \frac{(ax)^j}{j!}x^m e^{-\frac{1}{2}x^2}dx}{\int_0^\infty x^m e^{-\frac{1}{2}x^2dx}}\\ &= \sum_{j=0}^\infty \frac{a^j}{j!}\frac{\int_0^\infty x^{m+j}e^{-\frac{1}{2}x^2}dx}{\int_0^\infty x^m e^{-\frac{1}{2}x^2}dx}\\ &= \sum_{j=0}^\infty \frac{a^j}{j!}\prod_{i=1}^j \frac{\int_0^\infty x^{m+i}e^{-\frac{1}{2}x^2}dx}{\int_0^\infty x^{m+i-1}e^{-\frac{1}{2}x^2}dx}\\ &= \sum_{j=0}^\infty \frac{a^j}{j!}\prod_{i=1}^j\Big(\sqrt{m+i}\Big(1+O\Big(\frac{1}{m}\Big)\Big)\Big)\\ &= \sum_{j=0}^\infty \frac{\Big(|a|\sqrt{m}\Big(1+O\big(\frac{1}{m}\big)\Big)\Big)^j}{j!}\sqrt{\prod_{i=1}^j\Big(1+\frac{i}{m}\Big)}. \qquad (12)\end{aligned}$$

We break the sum in Eq. 12 into pieces and analyze them separately.

Let $\gamma > e$ and let $M = a\sqrt{m}\gamma$. We first show that the sum from M to ∞ in Eq. 12 is negligible. Again, by Stirling's approximation

$$\begin{aligned}\prod_{i=1}^j\Big(1+\frac{i}{m}\Big) = \frac{(m+j)!}{m!m^j} &= \frac{e^{-(m+j)}(m+j)^{m+j+\frac{1}{2}}\Big(1+O\big(\frac{1}{m+j}\big)\Big)}{m^j e^{-m}m^{m+\frac{1}{2}}\Big(1+O\big(\frac{1}{m}\big)\Big)}\\ &= e^{-j}\Big(\frac{m+j}{m}\Big)^{m+j+\frac{1}{2}}\Big(1+O\Big(\frac{1}{m}\Big)\Big)\end{aligned}$$

and so

$$\begin{aligned}\frac{\Big(|a|\sqrt{m}\Big(1+O\big(\frac{1}{m}\big)\Big)\Big)^j}{j!}\sqrt{\prod_{i=1}^j\Big(1+\frac{i}{m}\Big)} &= \frac{(|a|\sqrt{m})^j e^{-\frac{1}{2}j}\sqrt{\big(\frac{m+j}{m}\big)^{m+j+\frac{1}{2}}}\Big(1+O\big(\frac{1}{m}\big)\Big)^j}{\sqrt{2\pi}e^{-j}j^{j+\frac{1}{2}}}\\ &\quad\times\Big(1+O\Big(\frac{1}{m}\Big)\Big).\end{aligned}$$

Hence

$$\log\left\{\frac{\left(|a|\sqrt{m}\left(1+O(\frac{1}{m})\right)\right)^j}{j!}\sqrt{\prod_{i=1}^{j}\left(1+\frac{i}{m}\right)}\right\}$$

$$= j\log(|a|\sqrt{m}) + \frac{1}{2}(m+j+\frac{1}{2})\log\left(\frac{m+j}{m}\right) - \frac{1}{2}\log(2\pi) + \frac{1}{2}j - (j+\frac{1}{2})\log(j) + j\log\left(1+o\Big(\frac{1}{m}\Big)\right)$$

$$= -\Big(j+\frac{1}{2}\Big)\log(j) + \frac{1}{2}\Big(m+j+\frac{1}{2}\Big)\log(m+j) - \frac{1}{2}\Big(m+j+\frac{1}{2}\Big)\log(m) + j\log(|a|\sqrt{m}) + \frac{1}{2}j + j\log\left(1+o\Big(\frac{1}{m}\Big)\right) - \frac{1}{2}\log(2\pi).$$

For large enough m and $j = M = |a|\sqrt{m}\gamma$, this expression equals

$$-|a|\sqrt{m}\gamma\big(\log(\gamma) - 1 + o(1)\big) + \frac{1}{4}a^2\gamma^2\big(1+o(1)\big) \quad < \quad -\frac{1}{2}a\sqrt{m}\gamma\big(\log(\gamma)-1\big).$$

Let $0 < \varepsilon$. For $j \geq M$ the derivative with respect to j of

$$-\Big(j+\frac{1}{2}\Big)\log(j) + \frac{1}{2}\Big(m+j+\frac{1}{2}\Big)\log(m+j) - \frac{1}{2}\Big(m+j+\frac{1}{2}\Big)\log(m) + j\log(|a|\sqrt{m}) + \frac{1}{2}j + j\log\left(1+\frac{\varepsilon}{m}\right)$$

equals

$$-\log\Big(\frac{j}{|a|\sqrt{m}}\Big) - \frac{2m+j}{4j(m+j)} + \frac{1}{2}\log\left(1+\frac{j}{m}\right) + \log\left(1+\frac{\varepsilon}{m}\right) \quad \leq \quad -\frac{1}{2}\log(\gamma).$$

It follows that as $m \to \infty$

$$\sum_{j=M}^{\infty}\frac{\big(a\sqrt{m}\big(1+O\big(\frac{1}{m}\big)\big)\big)^j}{j!}\sqrt{\prod_{i=1}^{j}\left(1+\frac{i}{m}\right)} \leq e^{-\frac{1}{2}a\sqrt{m}\gamma\left(\log(\gamma)-1\right)}\sum_{j=0}^{\infty}e^{-\frac{1}{2}\log(\gamma)j}$$

$$= \frac{e^{-\frac{1}{2}a\sqrt{m}\gamma\left(\log(\gamma)-1\right)}}{1-e^{-\frac{1}{2}\log(\gamma)}} \to 0. \tag{13}$$

For $j \leq M$

$$\log\left(\prod_{i=1}^{j}\left(1+\frac{i}{m}\right)\right) = \sum_{i=1}^{j}\left[\frac{i}{m} - \frac{1}{2}\frac{i^2}{m^2} + \frac{1}{3}\frac{i^3}{m^3}\cdots\right]$$

$$= \frac{j(j+1)}{2m} - \sum_{i=1}^{j}\frac{i^2}{m^2}\left(\frac{1}{2} - \frac{1}{3}\frac{i}{m} + \frac{1}{4}\frac{i^2}{m^2}\cdots\right)$$

$$= j\Big(\frac{1}{2}\frac{j}{m}\Big) - O\Big(\frac{|a|^3}{\sqrt{m}}\Big).$$

It follows that

$$\sum_{j=0}^{M}\frac{\big(a\sqrt{m}\big(1+O\big(\frac{1}{m}\big)\big)\big)^j}{j!}\sqrt{\prod_{i=1}^{j}\left(1+\frac{i}{m}\right)} = \sum_{j=0}^{M}\frac{\big(a\sqrt{m}\big(1+O\big(\frac{1}{m}\big)\big)\big)^j}{j!}e^{\frac{1}{4}\frac{j^2}{m} - O\left(\frac{|a|^3}{\sqrt{m}}\right)}. \tag{14}$$

Note that if $X \sim \text{Poisson}(\lambda)$ then $P(X > y) \le \frac{Ee^{tX}}{e^{ty}} = \frac{e^{\lambda(e^t-1)}}{e^{ty}}$ for all $t > 0$, which is minimal at $t = \log\left(\frac{y}{\lambda}\right)$; hence $\log\left(P(X > y)\right) < e^{y-\lambda}/\left(\frac{y}{\lambda}\right)^y$. Recall that $\frac{1}{2} < \eta < 1$ and denote $\lambda = a\sqrt{m} + \frac{1}{4}\gamma a^2$, $y = a\sqrt{m} + (a\sqrt{m})^\eta$. Thus

$$\sum_{j=a\sqrt{m}+(a\sqrt{m})^\eta}^{M=\gamma a\sqrt{m}} \frac{\left(a\sqrt{m}\left(1+O\left(\frac{1}{m}\right)\right)\right)^j}{j!} e^{\frac{1}{4}\frac{j^2}{m}}$$

$$\begin{aligned}
&< e^{a\sqrt{m}+\frac{1}{4}a^2} e^{\frac{1}{4}a^2(\gamma-1)} \mathrm{P}(\text{Poisson}(\lambda) > \mathrm{y}) \\
&\le e^{a\sqrt{m}+\frac{1}{4}a^2} e^{\frac{1}{4}a^2(\gamma-1)} e^{\log((a\sqrt{m})^\eta - \frac{1}{4}\gamma a^2) - (a\sqrt{m}+(a\sqrt{m})^\eta)\log\left(\frac{a\sqrt{m}+(a\sqrt{m})^\eta}{a\sqrt{m}+\frac{1}{4}\gamma a^2}\right)} \\
&\le e^{a\sqrt{m}+\frac{1}{4}a^2} O\left(e^{-(a\sqrt{m})^\eta}\right).
\end{aligned} \tag{15}$$

If $X \sim \text{Poisson}(\lambda)$ then $P(X < y) \le \frac{Ee^{-tX}}{e^{-ty}} = \frac{e^{\lambda(e^{-t}-1)}}{e^{-ty}}$ for all $t > 0$, which is minimal at $t = -\log\left(\frac{y}{\lambda}\right)$; hence $\log\left(P(X < y)\right) < e^{y-\lambda}/\left(\frac{y}{\lambda}\right)^y$. Hence for $j < a\sqrt{m} - (a\sqrt{m})^\eta$

$$\sum_{j=0}^{a\sqrt{m}-(a\sqrt{m})^\eta} \frac{\left(a\sqrt{m}\left(1+O\left(\frac{1}{m}\right)\right)\right)^j}{j!} e^{\frac{1}{4}\frac{j^2}{m}}$$

$$\begin{aligned}
&< e^{a\sqrt{m}+\frac{1}{4}a^2} \times \mathrm{P}\left(\text{Poisson}\left(a\sqrt{m}\left(1+O\left(\frac{1}{m}\right)\right)\right) < a\sqrt{m} - (a\sqrt{m})^\eta\right) \\
&= e^{a\sqrt{m}+\frac{1}{4}a^2} O\left(e^{-(a\sqrt{m})^{2\eta-1}}\right).
\end{aligned} \tag{16}$$

Finally, for $a\sqrt{m} - (a\sqrt{m})^\eta \le j \le a\sqrt{m} + (a\sqrt{m})^\eta$

$$a^2[1 - 2a^{\eta-1}m^{\frac{1}{2}(\eta-1)} + a^{2(\eta-1)}m^{(\eta-1)}] \le \frac{j^2}{m} \le a^2[1 + 2a^{\eta-1}m^{\frac{1}{2}(\eta-1)} + a^{2(\eta-1)}m^{(\eta-1)}].$$

Hence $\frac{j^2}{m} = a^2\left[1 + O\left((a\sqrt{m})^{\eta-1}\right)\right]$ and

$$\sum_{j=a\sqrt{m}-(a\sqrt{m})^\eta}^{a\sqrt{m}+(a\sqrt{m})^\eta} \frac{\left(a\sqrt{m}\left(1+O\left(\frac{1}{m}\right)\right)\right)^j}{j!} e^{\frac{1}{4}\frac{j^2}{m}} = e^{a\sqrt{m}+\frac{1}{4}a^2\left[1+O\left((a\sqrt{m})^{\eta-1}\right)\right]}. \tag{17}$$

Combining Equations (12)–(17) accounts for Theorem 1.

Appendix B: Computing

This is a MATLAB (2014) program for calculating

$$\frac{\int_c^\infty x^m e^{-\frac{1}{2}(x-a)^2} dx}{\int_0^\infty x^m e^{-\frac{1}{2}x^2} dx} e^{-a\sqrt{m}}$$

Input: $m > 2, a, c$. Output: psi $=\psi_{a,c} = \frac{\int_c^\infty x^m e^{-\frac{1}{2}(x-a)^2} dx}{\int_0^\infty x^m e^{-\frac{1}{2}x^2} dx} e^{-a\sqrt{m}}$.

```
xi=0;  psi=0;
d=a-c;
g0=sqrt(2*pi)*normcdf(d);
g(1)=exp(-.5*d^2)+a*g0;
g(2)=(a+c)*exp(-.5*d^2)+(a^2+1)*g0;
psi(1)=g(1)*exp(-a);
psi(2)=g(2)* sqrt(2/pi)*exp(-a*sqrt(2));
xi(1)= sqrt(2/pi)*psi(1);
xi(2)= sqrt(pi/2)*psi(2);

for n=3:m
  psi(n)  ...
   =   a*(xi(n-1)/(n-1))*exp(-a*(sqrt(n)-sqrt(n-1)))
       + psi(n-2)*exp(-a*(sqrt(n)-sqrt(n-2)))  ...
       + exp(-a*sqrt(n))*((c/sqrt(2))^(n-1))  ...
       *exp(-.5*d^2)/gamma((n+1)/2);
  xi(n)  ...
   =   a*psi(n-1)*exp(-a*(sqrt(n)-sqrt(n-1)))  ...
       + ((n-1)/(n-2))*xi(n-2)*exp(-a*(sqrt(n)-sqrt(n-2)))  ...
       + exp(-a*sqrt(n))*c  ...
       *((c/sqrt(2))^(n-2))*exp(-.5*d^2)/gamma(n/2);
end
```

Remark The program above is intended to cover all cases of c, including $c = 0$. When n becomes large, $\log(\Gamma(n/2))$ is calculated more precisely than $\Gamma(n/2)$, so that when $c \neq 0$ the last summands in both $psi(n)$ and $xi(n)$ should be completely calculated first on the log scale and only then exponentiated.

Appendix C: A Sequential Changepoint Detection Context

The need for calculating $\psi_{0,a}(m)$ appears in a number of changepoint problems (cf. Pollak et al. 1993 and Krieger et al. 2003). For example, consider independent normally distributed random variables observed sequentially whose mean may (or may not) increase at an unknown time ν and one monitors the sequence to detect such a change. Formally,

$$X_1, \ldots, X_{\nu-1} \sim \text{Normal}(\mu, \sigma^2)$$
$$X_\nu, \ldots \quad \sim \text{Normal}(\mu + \delta\sigma, \sigma^2)$$

where $X_1, X_2, \ldots$ are independent. Consider the case that μ and σ are unknown and one considers an increase of δ standard deviations to be of import. Define:

$$Y_i = (X_i - \bar{X}_{i-1})\sqrt{\frac{i-1}{i}}$$
$$Z_i = \frac{Y_i}{|Y_2|}.$$

The sequence $\{Z_i\}$ is invariant with respect to the unknown parameters μ and σ. Therefore, the likelihood ratio

$$\Lambda_k^n = \frac{f_{\nu=k}(Z_2, \ldots, Z_n)}{f_{\nu=\infty}(Z_2, \ldots, Z_n)}$$

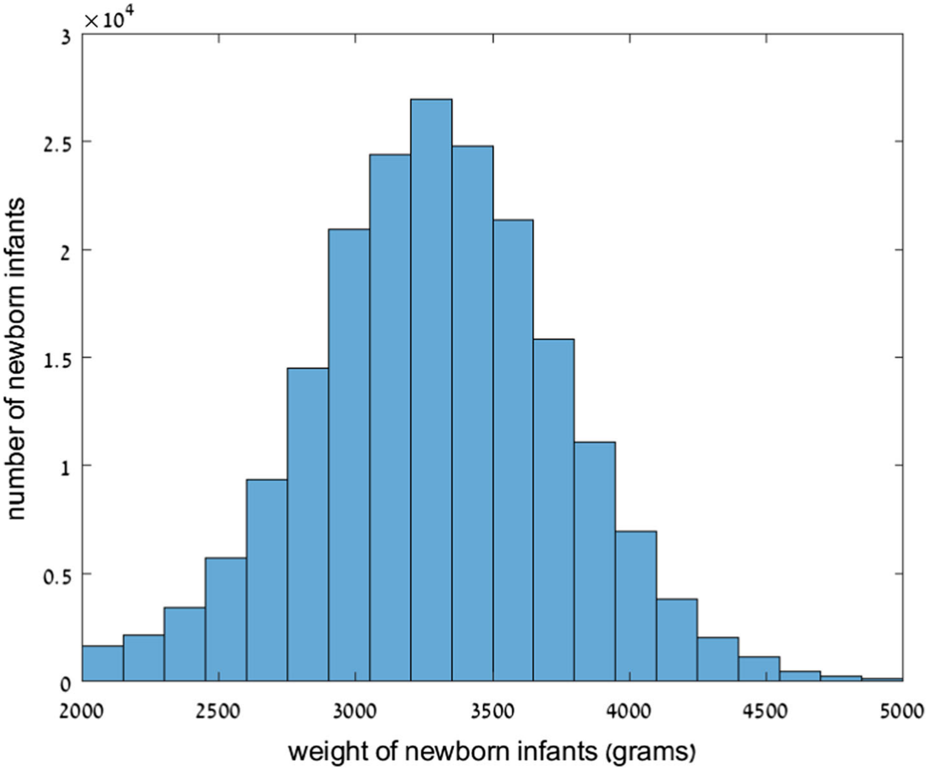

Fig. 2 Weight (in grams) at birth of 196710 infants born at a large hospital in Israel between 13/10/2004 and 19/5/2017

of the first $n-1$ invariant statistics $Z_2, \ldots, Z_n$ (for a change occurring at the $\nu = k^{\text{th}}$ observation vs. no change ever occurring) can be calculated. Once done, the Shiryaev–Roberts statistic

$$R_n = \sum_{k=1}^{n} \Lambda_k^n$$

can be used to declare that a change is in effect; an alarm would be raised if R_n crosses a pre-specified threshold A.

Since one cannot differentiate between the case that there is no change ever and the case that a change occurred at the very beginning, necessarily $\Lambda_1^n = 1$. Clearly,

$$\Lambda_2^2(\delta) = \left\{\Phi\left(\frac{\delta}{\sqrt{2}}\right) I(Z_2 = 1) + \left[1 - \Phi\left(\frac{\delta}{\sqrt{2}}\right)\right] I(Z_2 = -1)\right\} \Big/ \frac{1}{2}$$
$$= 1 + \left[2\Phi\left(\frac{\delta}{\sqrt{2}}\right) - 1\right] \operatorname{sign}(Z_2).$$

Letting

$$U_k(n) = (k-1)\frac{\sum_{i=k}^{n} \frac{Z_i}{\sqrt{i(i-1)}}}{\sqrt{\sum_{i=2}^{n} Z_i^2}} = (k-1)\frac{\sum_{i=k}^{n} \frac{Y_i}{\sqrt{i(i-1)}}}{\sqrt{\sum_{i=2}^{n} Y_i^2}}$$

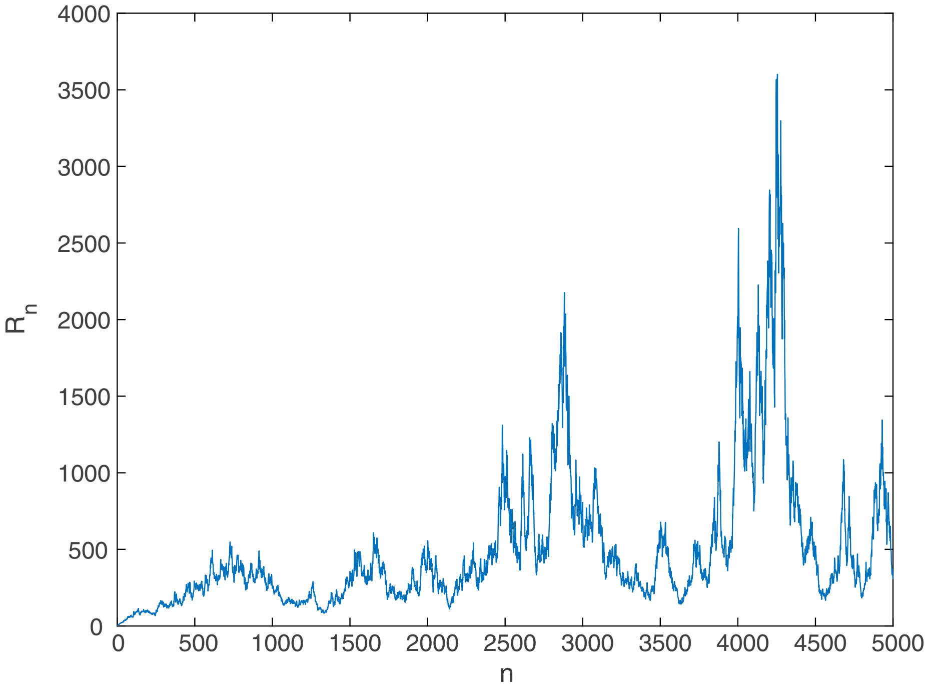

Fig. 3 A Shiryaev–Roberts control chart for the weights of the first 5000 babies described in Fig. 2 born after 13/10/2004

a lengthy calculation (similar to that in Pollak et al. 1993) yields for $2 \le k \le n$

$$\Lambda_k^n(\delta) = \{\Lambda_2^2 I\{k=2\} + I\{k \neq 2\}\} e^{-\frac{1}{2}(k-1)^2\delta^2[\frac{1}{k-1}-\frac{1}{n}]} e^{\frac{1}{2}\delta^2 U_k^2(n)} \frac{g_{0,\delta U_k(n)}(n-2)}{g_{0,0}(n-2)}.$$

(Note that $\frac{g_{0,a}(m)}{g_{0,0}(m)} = \psi_{0,a}(m)e^{a\sqrt{m}}$. For the sake of precision, it is advisable to first calculate the log of the components of $\Lambda_k^n(\delta)$ and then exponentiate their sum.)

As an example where m can be very large, we consider the list of weights at birth of 196710 babies born at a large hospital in Israel between 13/10/2004 and 18/5/2017 whose weight at birth was between 2000 and 5000 grams. Figure 2 is a histogram of the data. A normal distribution seems to fit the data well.

During the last decades, worldwide, obesity and macrosomia have been on the rise. It would have been reasonable to monitor the weight of newborn infants for an increase in mean. Since a rise could be gradual, it would be reasonable to start by trying to detect a small change — for example's sake, we choose $\delta = 0.1$, an increase of one tenth of a standard deviation (ca. 46 grams). Figure 3 presents the sequence of surveillance statistics R_n for the first 5000 observations and Fig. 4 for the first 65000 observations.

The choice of the threshold A is made in light of the risk one is willing to take regarding a false alarm. Suppose one were willing to tolerate a false alarm on the average once in 20 years. With an average of roughly 15000 babies born each year, this would mean a false alarm on the average once in 300000 observations. Using a renewal-theoretic approximation for the ARL2FA (Pollak 1987), this means that the cutoff level A should be $300000/1.06 \sim 283000$. It is clear from Figs. 3 and 4 that such a change would not been detected within the first 65000 observations.

Figure 5 presents the sequence of surveillance statistics R_n for the entire sequence. Clearly, R_n exceeds $A = 283000$ a short while after the 70000^{th} observation. In fact,the 70440^{th} is the first to carry R_n over the threshold. This means that it would have been declared after the 70440th newborn that the mean weight has increased. Figure 6 depicts the loglikelihood function $\log \Lambda_k^{70440}(\delta = 0.1)$, $k = 1, \ldots, 70440$. This function attains its maximum at $k = 65834$, so 65834 can be regarded at the time of stopping (70400) as the maximum likelihood estimate of the changepoint. This could be interpreted as the increase being detected 4606 observations (ca. 4 months) after its occurrence.

In fact, the average weight of the first 65833 newborns is 3286 grams, whereas the average weight of newborns #65834 – #67000 is 3332 grams (an increase of approximately 0.1 standard deviations).

Continuing with the ensuing newborns, R_n drops again, and until the end of the sequence does not cross the $A = 283000$ level again. In fact, the average weight of the newborns from babies #67001 – 196710 is 3291 grams. So, it seems as if the increase was either temporary, or apparent. Was the alarm false? After observing 67000 newborns, it would have seemed that the increase was real. With hindsight, one may either believe that the increase was real, but a change (a decrease) took place thereafter and the mean weight reverted to its original level, or one may interpret the episode as having been a false alarm.

The calculation of R_n was done in the following way. $R_1, \ldots, R_{5000}$ were calculated using the recursions of Section 3; ensuing $R_n{}'$s were calculated by applying Theorem 1. Rather than calculating $\frac{g_{0,\delta U_k(n)}(n-2)}{g_{0,0}(n-2)}$ by the recursions separately for each $\delta U_k(n)$, calculation time drops immensely by creating a fine enough grid of $\frac{g_{0,x}(n)}{g_{0,0}(n)}$ and interpolating. Creation of such a grid (in our example for $n = 1, \ldots, 5000$ and approximately 500 values of x, spaced so that between adjacent x's the function $\frac{g_{0,x}(n)}{g_{0,0}(n)}$ is almost perfectly linear). This grid is calculated much faster by the recursions of Section 3 than by the method of Section 5 (after all, the recursion that generated $\frac{g_{0,x}(5000)}{g_{0,0}(5000)}$ created all of $\frac{g_{0,x}(n)}{g_{0,0}(n)}$, $n = 1, \ldots, 5000$ along the way, whereas by the method of Section 5 each $\frac{g_{0,x}(n)}{g_{0,0}(n)}$ has to be calculated separately.)

Appendix D: Remarks

Remark 1 A lower bound can be obtained by Jensen's inequality:

$$\begin{aligned}
\log\left(e^{\frac{1}{2}a^2}\frac{\int_0^\infty x^m e^{-\frac{1}{2}(x-a)^2}dx}{\int_0^\infty x^m e^{-\frac{1}{2}x^2}dx}\right) &= \log\left(\int_0^\infty e^{ax}\frac{x^m e^{-\frac{1}{2}x^2}dx}{\int_0^\infty x^m e^{-\frac{1}{2}x^2}dx}\right)\\
&\geq \int_0^\infty ax\frac{x^m e^{-\frac{1}{2}x^2}dx}{\int_0^\infty x^m e^{-\frac{1}{2}x^2}dx}\\
&= a\frac{\int_0^\infty x^{m+1}e^{-\frac{1}{2}x^2}dx}{\int_0^\infty x^m e^{-\frac{1}{2}x^2}dx}\\
&= a\frac{2^{\frac{1}{2}m}\Gamma(\frac{m+2}{2})}{2^{\frac{1}{2}(m-1)}\Gamma(\frac{m+1}{2})} = a\frac{\sqrt{2}\Gamma(\frac{m+2}{2})}{\Gamma(\frac{m+1}{2})}\\
&= a\sqrt{m}\left(1+O\left(\frac{1}{m^2}\right)\right) \quad .
\end{aligned}$$

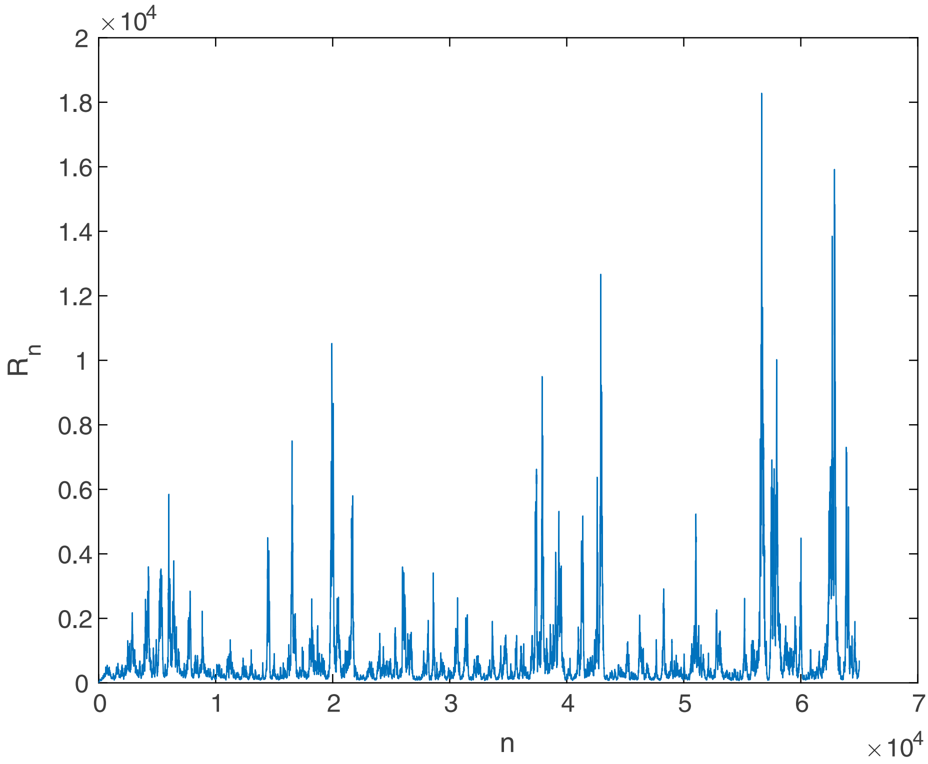

Fig. 4 A Shiryaev–Roberts control chart for the weights of the first 65000 babies described in Fig. 2 born after 13/10/2004

Remark 2 Cook (2010) presented an upper bound

$$\int_c^\infty x^m e^{-\frac{1}{2}(x-a)^2} dx \le \frac{\pi}{2\sqrt{2}} e^{-c^2-m+\frac{ma}{c}+\frac{m^2}{2c^2}} \frac{c^m}{c+\sqrt{c^2+\frac{4}{\pi}}} \tag{18}$$

for $c > 0$. On the log scale, Cook's upper bound has an asymptotic ($m \to \infty$) order of magnitude m^2, whereas Theorem 1 posits an order of magnitude $m \log m$.

Remark 3 A similar type of analysis can be done for $\int_{-\infty}^\infty x^m e^{-\frac{1}{2}(x-a)^2} dx$ (cf. Pollak et al. 1993). Note that

$$\begin{aligned}\int_{-\infty}^\infty x^m e^{-\frac{1}{2}(x-a)^2} dx &= \int_{-\infty}^0 x^m e^{-\frac{1}{2}(x-a)^2} dx + \int_0^\infty x^m e^{-\frac{1}{2}(x-a)^2} dx \\ &= (-1)^m \int_0^\infty x^m e^{-\frac{1}{2}(x+a)^2} dx + \int_0^\infty x^m e^{-\frac{1}{2}(x-a)^2} dx\end{aligned}$$

and (unless $m = 0$) that (depending on a) one of the two integrals is asymptotically negligible with respect to the other. Thus Theorem 1 can be applied to obtain an approximation to the m^{th} moment of a (non-truncated) normal distribution.

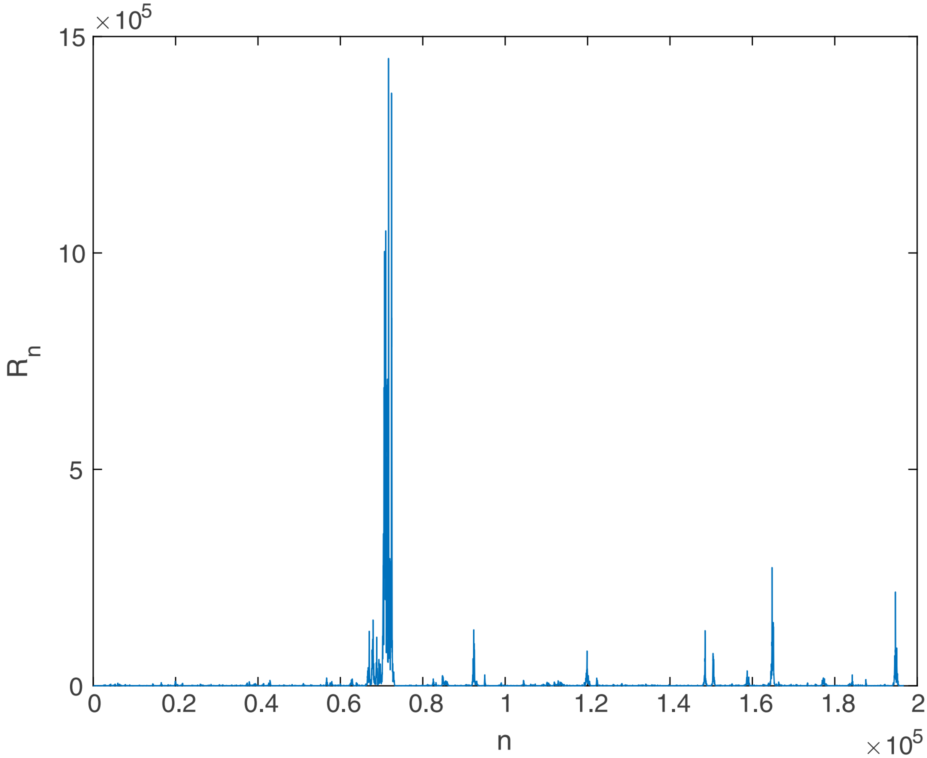

Fig. 5 A Shiryaev–Roberts control chart for the weights all of the babies born described in Fig. 2 between 13/10/2004 and 18/5/2017

Remark 4 A similar type of analysis can be done when the truncation is from both ends. To see this, it suffices to consider $\int_0^c x^m e^{-\frac{1}{2}(x-a)^2} dx$ for $c > 0$. Clearly, $\frac{\int_0^c x^m e^{-\frac{1}{2}(x-a)^2} dx}{\int_0^c x^m dx} \to e^{-\frac{1}{2}(c-a)^2}$ as $m \to \infty$. It follows that

$$\int_0^c x^m e^{-\frac{1}{2}(x-a)^2} dx = e^{-\frac{1}{2}(c-a)^2} \frac{c^{m+1}}{m+1}(1 + o(1)).$$

Without loss of generality, assume that $c > |b|$. A recursion for $\int_b^c x^m e^{-\frac{1}{2}(x-a)^2} dx$ that builds on this is

$$h_{a,b,c}(m) = \frac{m+1}{c^2}\left[h_{a,b,c}(m-2) + \frac{ac}{m} h_{a,b,c}(m-1) + \left(\frac{b}{c}\right)^{m-1} e^{-\frac{1}{2}(b-a)^2} - e^{-\frac{1}{2}(c-a)^2}\right]$$

where $h_{a,b,c}(m) = \frac{m+1}{c^{m+1}} \int_b^c x^m e^{-\frac{1}{2}(x-a)^2} dx$. (Note that for $|b| > c > b$, $\int_b^c x^m e^{-\frac{1}{2}(x-a)^2} dx = (-1)^m \int_{-c}^{-b} x^m e^{-\frac{1}{2}(x+a)^2} dx$.)

Remark 5 A similar recursion can be constructed for densities proportional to $e^{\text{const} \times x^2} I(b < x < c)$ where const > 0 and b, c are finite.

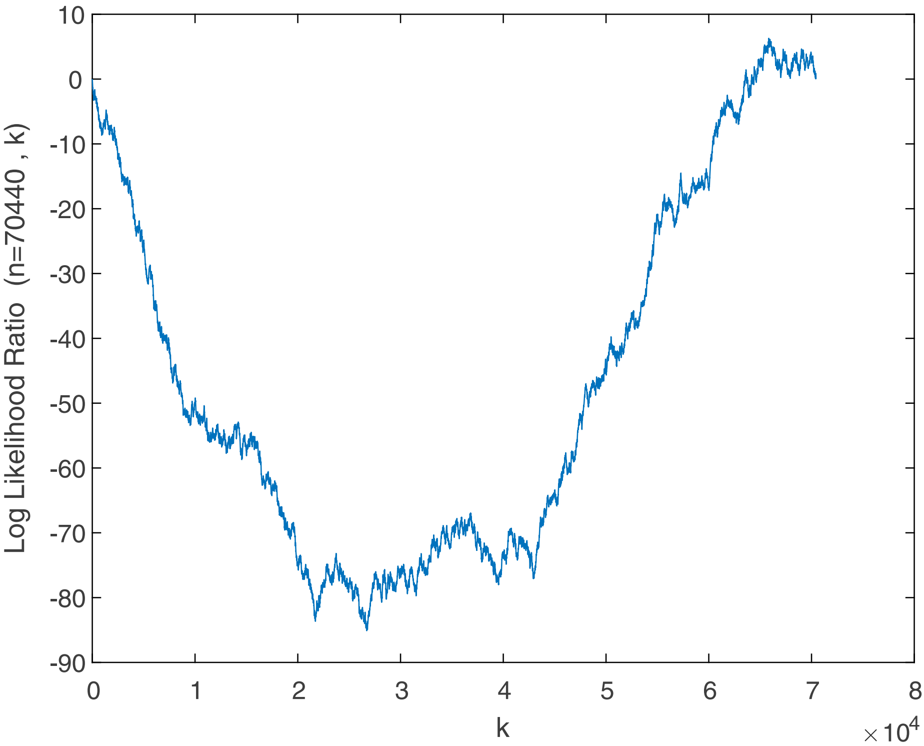

Fig. 6 The loglikelihood ratio of {increase at k} vs. {no increase} after declaring at observation #70440 that an increase is in effect

Remark 6 Tables relevant to the truncated normal distribution have been around for over a century (cf. Pearson and Lee 1908 and Lee 1914). Even today it is considered as part and parcel of applied distributions (cf. O'Connor 2011) and papers have been devoted to estimation of its parameters (cf. Barr and Sherrill 1999; Horrace 2015; Liquet and Nazarathy 2015). For most practical purposes, the first four moments of the distribution (mean, variance, skewness and kurtosis) have been of applied interest. Higher order moments may appear in quadrature methods (cf. Burkardt 2014); their order of magnitude would seem to be in the tens at most.

References

Barr DR, Sherrill ET (1999) Mean and variance of truncated normal distributions. Amer Statist 53:357–361

Burkardt J (2014) The truncated normal distribution. https://people.sc.fsu.edu/~jburkardt/presentations/truncated_normal.pdf

Cook JD (2010) Upper bounds on non-central chi-square tails and truncated normal moments. www.johndcook.com/non_central_chi_square.pdf

Dhrymes PJ (2005) Moments of truncated (normal) distributions. http://www.columbia.edu/~pjd1/papers.html

Gordon L, Pollak M (1995) A robust surveillance scheme for stochastically ordered alternatives. Ann Statist 23:1350–1375

Horrace WC (2015) Moments of the truncated normal distribution. J Prod Anal 43:133–138

Krieger AM, Pollak M, Yakir B (2003) Surveillance of a simple linear regression. J Amer Statist Assoc 98:1–15

Lee A (1914) Table of the Gaussian "tail" tunctions; when the "tail" is larger than the body. Biometrika 10:208–214

Liquet B, Nazarathy Y (2015) A dynamic view to moment matching of truncated distributions. Statist Prob Letts 104(53):87–93

MATLAB (2014) MATLAB and statistics toolbox release 2014b. The Mathworks Inc., Natick

Toolbox (2016) MATLAB multiprecision computing toolbox (ADVANPIX Release 2016). The MathWorks, Inc., Natick. http://www.advanpix

O'Connor AN (2011) Probability distributions used in reliability engineering. Reliability information analysis center (RIAC)

Pearson K, Lee A (1908) On the generalized probable error in multiple normal correlation. Biometrika 6:59–68

Pollak M (1987) Average run lengths of an optimal method of detecting a change in distribution. Ann Statist 15:749–779

Pollak M, Siegmund D (1991) Sequential detection of a change in a normal mean when the initial value is unknown. Ann Statist 19:394–416

Pollak M, Croarkin C, Hagwood C (1993) Surveillance schemes with application to mass calibration. NISTIR 5158 Technical Report, Statistical Division, The National Institute of Standards and Technology, Gaithersburg, MD 20899, USA

Quesenberry C (1991) SPC Q processes for start-up processes and short or long runs. J Qual Tech 23:213–224

van Dobben de Bruyn CS (1968) Cumulative Sum Tests: Theory and Practice. Griffin, London