Disney, S.M., Gaalman, G., Hedenstierna, C.P.T. and Hosoda, T., (2014), "Fill rate in a periodic review Order-Up-To policy under correlated normally distributed demand", Pre-prints of the 18th International Working Seminar of Production Economics, Innsbruck, Austria, February 24th–28th, Vol. 1, pp103–115.

Fill rate in a periodic review Order-Up-To policy under correlated normally distributed demand

Stephen M. Disney[1], Gerard Gaalman[2], Carl Philip T. Hedenstierna[1], Takamichi Hosoda[3]

1. Logistics Systems Dynamics Group, Cardiff Business School, Cardiff University, Aberconway Building, Colum Drive, Cardiff, CF10 3EU, UNITED KINGDOM.
DisneySM@cardiff.ac.uk, cpthed@gmail.com

2. Department of Operations, Faculty of Economics and Business, University of Groningen, P.O. Box 800, 9700 AV Groningen, THE NETHERLANDS.
g.j.c.gaalman@rug.nl

3. Graduate School of International Management, Aoyama Gakuin University, 4-4-25, Shibuya, Shibuya-ku, Tokyo, 150-8366, JAPAN.
hosoda@gsim.aoyama.ac.jp

Abstract

We investigate the inventory customer service metric known as the fill rate. The fill rate is defined as the proportion of demand that is immediately fulfilled from inventory. However, the task of finding analytical solutions for general cases is difficult. In the literature two approximate approaches are proposed. The first of these approximations is the traditional fill rate (or p2 service measure) that is exact in the Order-Up-To replenishment policy with Minimum Mean Squared Error forecasting, zero lead-time and independent and identically distributed (i.i.d.) demand. However, when any of these assumptions is relaxed then the traditional fill rate measure is only a lower bound. A second approximation in the literature has been proposed by Sobel (2004) that is better able to cope with non-zero lead-times as it manages the double accounting of accumulated backlogs. Sobel's approach still requires positive i.i.d. demands implying there is no correlation between demand and net stock. However the assumption of i.i.d. demand is unrealistic. Correlation may be introduced by auto-correlated demand or forecasting methods, amongst others. We propose a new fill rate measure that can handle correlated and possibly negative demand. We assume normally distributed demand, and treat negative demand as returns. The problems reduces to identifying the minimum of correlated bi-variate random variables. There is an exact solution, but it has no closed form. However, the solution is amenable to numerical techniques and we present a custom Microsoft Excel function for practical investigations.

Key words: Fill rate, Order-up-to policy, Correlated demand, Negative demand.

1. Introduction

The fill rate is a popular measure of inventory service in high volume industries as it directly measures the customer's experience of demand fulfillment. Sometimes it is called the "p2" service measure to distinguish between the "availability" or the "p1" service measure (Silver, Pyke and Peterson, 1998). The p1 service measure determines the proportion of periods that end in a backlog or stock-out. If availability is a measure of the proportion of days in which it rains, fill rate is a measure of how much rain falls on an average day. Traditionally, fill rate is defined as

$$\text{Fill rate} = \frac{\text{Average number of units of demand filled}}{\text{Average demand}} = 1 - \frac{\text{Average backlogs}}{\text{Average demand}}, \tag{1}$$

but this simple looking definition hides a lot of technical details, issues and nuances that are often overlooked in the literature. In particular there are issues with double accounting of backlogs, lead times, auto-correlation in demand, correlation between the net stock and demand, negative demand, arbitrary ordering policies, and the distribution of demand and net stock.

In this paper we will explore the fill rate in a setting with normally distributed, but correlated (ARMA(1,1)) demands. We assume the linear, discrete time, Order-Up-To (OUT) replenishment policy is used to manage the inventory and that arbitrary, but constant, lead-times exist. As we assume a linear system operates backlogs are allowed and negative demands indicate net returns from customers in a period.

2. Literature review

Specifically we are interested in the fill rate for a single item at a single location or echelon in a supply chain. Sometimes it is referred to as the: item fill rate; volume fill rate; unit fill rate, or simply just the fill rate, Guijarro, Cardós and Babiloni (2012). This is different from the so-called order fill rate, Larson and Thorstenson (2014), which applies to groups of items. Silver and Bischak (2011) provide an excellent review of the literature on various single item fill rate approaches. All papers that we have found assume i.i.d. demand and these are mostly modeled as Poisson, Erlang, Normal, Gamma, or Binomial distributions.

Johnson, Lee and Davis (1995) discuss the error in fill rate expressions in periodic inventory systems due to double counting of backorders (when lead time is longer than the review period and a replenishment order is not sufficient to cover the existing backlog demand). This double counting over-estimates the number of stock outs and leads to underestimated fill rates. They identify further issues with stochastic lead-times and order crossovers.

Sobel (2004) provides general expressions for fill rates under various demand distributions, including normally i.i.d. demand in the OUT replenishment model. This is extended by Zhang and Zhang (2007) to a case where the review period is of arbitrary length. Teunter (2009) notes that this situation can be modeled as a compound renewal process with a much more direct derivation. Zhang and Sobel (2012) assume general random demand; Larsen and Thorstenson (2008) a compound renewal process. Tyworth and O'Neill (1997) investigate the sensitivity of the shape of the lead time demand distribution in a continuous review system.

In Kwon, Kim and Baek (2006), demand is normally distributed. The authors discuss the "traditional" expression of the fill rate. Referring to Johnson, Lee and Davis (1995), they state "this expression underestimates the true fill rate… To overcome such a problem … others have developed… more accurate approximation methods. Nevertheless, all of these approximate expressions suffer from the same problems of under-estimating the true fill rate". They recommend the use of the expression derived by Sobel (2004) for normally distributed demand. Chen, Lin and Thomas (2003), show that the expected infinite horizon item fill rate is lower than the finite horizon fill rate for arbitrary non-negative stochastic i.i.d demands. Banerjee and Paul (2005) prove their conjecture that the finite horizon fill rate is a monotonic decreasing function.

Van Donselaar and Broekmeulen (2013) consider the lost sales case, providing an extensive discussion of the lost sales literature. They use the standard fill rate expression; one proposed by Tijms and Groenevelt (1984); and develop a new one. The demand is stochastic discrete random variable fitted by a negative binomial, Poisson, or geometric distribution. Samii, Pibernik and

Yadav (2011) discuss the fill rate in a case with two classes of customers with different service levels where demand is Poisson.

Cardós and Babiloni (2011) consider the fill rate in an intermittent and slow moving demand setting with an OUT replenishment system without backlog. The lead time is smaller than the review period. The exact cycle service level is calculated as the probability that the demand is lower than the initial stock given, the demand is assumed to be always positive. Guijarro, Cardós and Babiloni (2012) develop a generalized method to compute the fill rate for any discrete i.i.d. demand process in a case where unmet demand is lost.

3. Preliminaries

3.1 Mathematical foundations. As we assume that the demand is normally distributed and the inventory control system is described by linear difference equation then all system variables are also normally distributed. Thus is it useful to define certain relationships associated with the normal distribution. The probability density function (pdf) of the standard normal distribution of a variable $x, -\infty < x < \infty$ and its inverse is

$$\varphi[x] = \tfrac{1}{\sqrt{2\pi}} \exp\left[-\tfrac{x^2}{2}\right], \; \varphi^{-1}[x] = \pm\sqrt{\log[2\pi] + 2\log[x]}. \tag{2}$$

$\varphi[x] > 0 \; \forall x$. $\varphi[x]$ is concave between $-1 < x < 1$ and convex when $x < -1$ and $x > 1$ from which we obtain the definition of the standard deviation. The cumulative distribution function (cdf) and its inverse is given by

$$\Phi[z] = \int_{-\infty}^{z} \varphi[x] \mathrm{d}x = \tfrac{1}{2}\left(1 + \operatorname{erf}\left[\tfrac{z}{\sqrt{2}}\right]\right), \; \Phi^{-1}[z] = \sqrt{2}\operatorname{erf}^{-1}[2z-1]. \tag{3}$$

$\Phi[z]$ is an increasing function within the interval $0 < \Phi[z] < 1$ for $-\infty < z < \infty$. $1 - \Phi[z] = \Phi[-z]$ and $\Phi[z] - 1 = -\Phi[-z]$. $\Phi[z]$ is convex when $z < 0$ and concave when $z > 0$. Another important relation is the so-called Error function, $\operatorname{erf}[z] = 2\Phi\left[z\sqrt{2}\right] - 1$. Both $\Phi[z]$ and the $\operatorname{erf}[z]$ cannot be expressed in terms of finite additions, subtractions, multiplications and root extractions. So both must be either computed numerically or otherwise approximated, Weisstein (2012). The standard normal Loss function is given by,

$$L[z] = \int_{z}^{\infty} \varphi[x](x - z) \mathrm{d}x = \varphi[z] - z(1 - \Phi[z]). \tag{4}$$

$\forall z$, $L[z] > 0$ and is a monotonically decreasing function in *z*. We note that $L'[z] = \Phi[z] - 1$ which means that the loss function is decreasing and that $L[z]$ is convex. The inverse loss function, $L^{-1}[z]$ has no known solution. Cain (1994) and Nadarajah and Kotz (2010) give the following expression for the pdf of the minimum of two normally distributed, correlated random variables,

$$\varphi_{\min}[x] = \frac{1}{\sigma_1}\varphi\left[\frac{x-\mu_1}{\sigma_1}\right]\Phi\left[\frac{\rho(x-\mu_1)}{\sigma_1\sqrt{1-\rho^2}} - \frac{x-\mu_2}{\sigma_2\sqrt{1-\rho^2}}\right] + \frac{1}{\sigma_2}\varphi\left[\frac{x-\mu_2}{\sigma_2}\right]\Phi\left[\frac{\rho(x-\mu_2)}{\sigma_2\sqrt{1-\rho^2}} - \frac{x-\mu_1}{\sigma_1\sqrt{1-\rho^2}}\right], \tag{5}$$

where $\{\mu_1,\sigma_1\}$ are the mean and standard deviation of x_1, a normally distributed random variable and $\{\mu_2,\sigma_2\}$ are mean and standard deviation of x_2, another normally distributed random variable. ρ is the Pearson correlation co-efficient between x_1 and x_2, $-1 \le \left(\rho = \frac{\text{cov}(x_1,x_2)}{\sigma_1\sigma 2}\right) \le 1$. We also use the maximum operator, $(x)^+ = \max[x,0]$, the minimum operator $(a,b)^- = \min[a,b]$ and the expectation operator $E[x] = \bar{x}$.

3.2. The traditional fill rate measure. The most common approach the traditional fill rate or the so-called p2 service measure, Silver, Pyke and Peterson (1998, pp 299). It is defined as

$$\beta_T = 1 - \frac{\text{E}\left[(-ns_t)^+\right]}{\text{E}[d_t]} = 1 - \frac{ESPRC}{\mu_d}. \tag{6}$$

Notice that we have provided in (6) the mechanism to determine the β_T in a time based simulation of the OUT policy with ns_t being the net stock in period time period *t* (net stock is the inventory on-hand plus the amount backlogged) and d_t the demand in time period *t*. The expected demand is often taken to be the mean demand, μ_d and the expected backlogs is often defined as the Expected Shortages Per Replenishment Cycle, Silver, Pyke and Peterson (1998),

$$ESPRC = \sqrt{1+T_p}\,\sigma_d \int_{tns/\left(\sqrt{1+T_p}\sigma_d\right)}^{\infty} \varphi[x]\left(x - \frac{tns}{\sqrt{1+T_p}\sigma_d}\right)dx = \sqrt{1+T_p}\,\sigma_d L\left[\frac{tns}{\sqrt{1+T_p}\sigma_d}\right], \tag{7}$$

T_p is the replenishment lead-time in periods, *tns* is the Target Net Stock, a safety stock, and σ_d is the long run standard deviation of the demand of a period. The problem with the traditional fill rate is when the backlogs become larger than the demand in a period the missed demand can become larger than the demand in the current period. Due to this, some of the backlogs are counted more than once. It means (6) is a lower bound of the true fill rate. However, it is quite accurate when the fill rate is near 100% and when there are no negative demands. However, when the achieved service level is more modest, the errors can become quite large and in extreme cases this measure can become negative, which is obviously nonsense.

Notice the traditional fill rate measure is only valid for i.i.d. demands - it makes a prediction of the standard deviation of the net stock levels, $\sigma_{ns} = \sqrt{1+T_p}\,\sigma_d$, which is correct for i.i.d. demands. A natural extension of (7) to account for correlated demands which we will investigate later in Section 6 would be

$$\beta_T^* = 1 - \frac{\text{E}\left[(-ns_t)^+\right]}{\text{E}[d_t]} = 1 - \frac{ESPRC}{\mu_d};\; ESPRC = \sigma_{ns}\int_{tns/\sigma_{ns}}^{\infty} \varphi[x]\left(x - \frac{tns}{\sigma_{ns}}\right)dx = \sigma_{ns} L\left[\frac{tns}{\sigma_{ns}}\right]. \tag{8}$$

3.3. The Sobel fill rate. Sobel (2004) defines the fill rate as,

$$\beta_S = \lim_{T\to\infty} \text{E}\left[\frac{\sum_{t=1}^{T} \min\left\{(ns_t + d_t)^+, d_t\right\}}{\sum_{t=1}^{T} d_t}\right] = \frac{\text{E}\left[\min\left\{(ns_t + d_t)^+, d_t\right\}\right]}{\text{E}[d_t]}, \tag{9}$$

when the limit and the expectation exists. The limit always exists for the class of generalised OUT policies, regardless of whether the demand is stationary or not. Essentially the same service measure has been studied by Zhang and Zhang (2007), Teunter (2009) and Larsen and Thorstenson (2014). Notice that the Sobel fill rate assumes always that the demand is positive. ns_t is the net stock at the end of period *t*. $ns_t + d_t$ is the net stock after the orders placed $T_p + 1$ periods ago have been received but before the demand has been satisfied. d_t is the demand in period *t*. Sobel (2004) provides the following expression for fill rate for normal i.i.d. demand (translated into the notation of this paper). He also provides a rather lengthy (hence we have not repeated it here for space reasons) expression for β_S that is based solely on standard normal pdf, cdf and loss functions.

$$\beta_S = \tfrac{1}{\mu_D} \int_0^{tns+\mu_d(T_p+1)} \left(\Phi\left[\tfrac{a-T_p\mu_d}{\sqrt{T_p}\sigma_d} \right] - \Phi\left[\tfrac{a-(1+T_p)\mu_d}{\sqrt{1+T_p}\sigma_d} \right] \right) \mathrm{d}a \tag{10}$$

If demand is negative, then (10) deems this negative demand to be fulfilled and in extreme cases this results in $\beta_S > 100\%$ or $\beta_S < 0\%$. Looking at the denominators of the cdf's in the integral and the integration limit of (10), we notice that it uses $\left\{ \sqrt{T_p}\sigma_d, \sqrt{1+T_p}\sigma_d \right\}$ - the standard deviation of the net stock levels maintained by the OUT policy when demand is i.i.d and lead-times of $\{T_p - 1, T_p\}$ are present. The natural question to ask is "if demand is not i.i.d, how does the following expression perform?"

$$\beta_S^* = \tfrac{1}{\mu_D} \int_0^{tns+\sigma_{ns}|_{Tp}} \left(\Phi\left[\tfrac{a-T_p\mu_d}{\sigma_{ns}|_{Tp-1}} \right] - \Phi\left[\tfrac{a-(1+T_p)\mu_d}{\sigma_{ns}|_{Tp}} \right] \right) \mathrm{d}a \,. \tag{11}$$

In (11), $\sigma_{ns}|_{T_p-1}$ is the standard deviation of the net stock levels in an OUT policy with a lead-time of $T_p - 1$ and $\sigma_{ns}|_{T_p}$ is the standard deviation of the net stock levels in an OUT policy with a lead-time of T_p. We will investigate the performance of (10) and (11) in Section 6.

4. Fill rates with correlated normally distributed (possibly negative) demand

Guijarro, Cardós and Babiloni (2012) argue that the condition for positive demand during a cycle must be explicitly taken into account when correctly determining the fill rate. Let's relax the assumption on non-negative demand. We let negative demands denote net returns from customers in a period. We note that returns can be quite significant in some industries such as book and consumer electronic retailing. Obviously these can not count towards the fulfilled demand and we have to adapted our definition of f_t, the fulfilled demand in time period *t*, to

$$\left. \begin{aligned} f_t &= IF\left(ns_t > 0, (d_t)^+, IF\left(d_t + ns_t > 0, d_t + ns_t, 0 \right) \right) \\ &= \min\left((d_t)^+, (d_t + ns_t)^+ \right) = \left(\min\left(d_t, d_t + ns_t \right) \right)^+ \end{aligned} \right\} . \tag{12}$$

In words (12) means that if the net stock at the end of the period is positive, then demand must have been satisfied. However if demand was negative (that is there were some returns), then the fulfilled demand is zero. If the net stock was negative at the end of the period then the fulfilled demand in the period is equal to the demand in the period minus the backlog at the end of the period. If the backlog at the end of the period is larger than the demand then no part of the demand could be satisfied and the fulfilled demand is zero. As f_t is the fulfilled demand in a single period, in order to obtain the long-run average proportion of fulfilled demand, β^*, in the presence of the possibility of negative demand then we need,

$$\beta^* = \frac{E\left[\left(\min\left(d_t, d_t + ns_t\right)\right)^+\right]}{E\left[\left(d_t\right)^+\right]}. \tag{13}$$

Assuming from now on that normally distributed demand (which may be correlated) exists our task now is to find an expression for the minimum of two normally distributed random variables, $\left(d_t, d_t + ns_t\right)$. Recall from Section 3, Nadarajah and Kotz (2010) provide the pdf of the minimum of two normally distributed, correlated random variables. We simply need to insert the variables, $\mu_1 = \mu_{ns+d} = \mu_d + tns$, $\sigma_1 = \sigma_{ns+d}$, $\mu_2 = \mu_d$, $\sigma_2 = \sigma_d$ and ρ, the Pearson Correlation Coefficient between $ns_t + d_t$ and d_t into their equation. ρ can be easily obtained via the impulse response of the system using $\rho = \left(\sum_{t=0}^{\infty}\left(ns_t + d_t\right)\left(d_t\right)\right) \big/ \left(\sigma_{ns+d}\sigma_d\right)$. For the OUT policy, i.i.d. demand $ns_0 + d_0 = 0$, and $\forall t > 0,\ d_t = 0$, thus $\rho = 0$. For non-i.i.d. demand this is not the case as ρ could be anything between $-1 \le \rho \le 1$. Using (5) together with standard techniques we find the expected value of (13) is given by

$$\beta^* = \frac{\int_0^\infty \left(\frac{1}{\sigma_{ns+d}} \varphi\left[\frac{y-\mu_{ns+d}}{\sigma_{ns+d}}\right] \Phi\left[\frac{1}{\sqrt{1-\rho^2}}\left(\frac{\rho\left(y-\mu_{ns+d}\right)}{\sigma_{ns+d}} - \frac{y-\mu_d}{\sigma_d}\right)\right] + \frac{1}{\sigma_d}\varphi\left[\frac{y-\mu_d}{\sigma_d}\right]\Phi\left[\frac{1}{\sqrt{1-\rho^2}}\left(\frac{\rho\left(y-\mu_d\right)}{\sigma_d} - \frac{y-\mu_{ns+d}}{\sigma_{ns+d}}\right)\right]\right) y \mathrm{d}y}{\sigma_d \varphi\left[\frac{\mu_d}{\sigma_d}\right] + \mu_d \Phi\left[\frac{\mu_d}{\sigma_d}\right]} \tag{14}$$

There does not appear to be an easily obtainable solution to the integral in the numerator of (14). Thus we have to resort to numerical techniques to calculate β^*. This is fairly easy to do in specialist mathematical software such as Matlab or Mathematica. However we has also developed a Microsoft Excel Add-in (source code is provided in Appendix A; the Add-in is available for download from www.bullwhip.co.uk) for use without such software.

5. Setting up an investigation of fill rates in the OUT policy with ARMA(1,1) demand

In this section we will define a modelling setting and obtain the required information for evaluating the various fill rate measures. We will model the OUT policy reacting to an Auto-Regressive Moving Average demand process of the first order (ARMA(1,1)). The mean centred ARMA(1,1) demand, Box and Jenkins (1976), in time period *t*, d_t, has a difference equation of

$$d_t = \phi\left(d_{t-1} - \mu_d\right) + \varepsilon_t - \theta\varepsilon_{t-1} + \mu_d, \tag{15}$$

ε_t is an independent and identically normally distributed random variable with a mean of zero and a variance of σ_ε^2, $\varepsilon_t \in N\left(0, \sigma_\varepsilon^2\right)$. μ_d is the mean demand, $-1 < \phi < 1$ is the autoregressive parameter and $-1 < \theta < 1$ is the moving average parameter. When $\theta = \phi$ then an i.i.d. white noise demand pattern is produced. The long run variance of the demand over time is well known to be

$$\sigma_d^2 = \sigma_\varepsilon^2\left(1 + \tfrac{(\phi-\theta)^2}{1-\phi}\right). \tag{16}$$

The OUT policy generates replenishment orders at time *t*, o_t, with the following expression;

$$o_t = \hat{d}_{t,t+T_p+1} + \left(TNS + \sum_{i=1}^{T_p} \hat{d}_{t,t+i} - ns_t - \sum_{i=1}^{T_p} o_{t-i} \right). \tag{17}$$

where $T_p \in \mathbb{N}_0$ is the replenishment lead-time. The OUT replenishment policy requires two forecasts. One of these forecasts is a prediction of demand over the lead-time, $\hat{d}_{t,(t+1,t+T_p)}$, made at time *t*. $\hat{d}_{t,(t+1,t+T_p)}$ is a forecast of demand in periods $\{t+1, t+2, ..., t+T_p\}$. The other forecast is a prediction of the demand in the periods after the lead-time, $\hat{d}_{t,t+T_p+1}$, made at time *t*, Hosoda and Disney (2009),

$$\left.\begin{aligned} \hat{d}_{t,(t+1,t+T_p)} &= \sum_{i=1}^{T_p} \hat{d}_{t,t+i} = \tfrac{1-\phi^{Tp}}{1-\phi}\left(\phi\left(d_t - \mu_d\right) - \theta\varepsilon_t\right) + \mu_d T_p \\ \hat{d}_{t,t+T_p+1} &= \phi^{T_p}\left(\phi\left(d_t - \mu_d\right) - \theta\varepsilon_t\right) + \mu_d \end{aligned}\right\}. \tag{18}$$

Finally the net stock balance equation completes the OUT policy,

$$ns_t = ns_{t-1} - d_t + o_{t-T_p-1}, \tag{19}$$

where ns_t is the net stock at time t and o_{t-T_p-1} is the order place T_p+1 periods ago. The "+1" is the sequence of events delay that is always present in discrete time systems. The variance of the inventory levels is given by, Gaalman (2006)

$$\begin{aligned} \sigma_{ns}^2 &= \sigma_\varepsilon^2 \sum_{t=0}^{T_p}\left(\left(\phi-\theta\right)\left(\tfrac{\phi^t-1}{1-\phi}\right)-1\right)^2 \\ &= \sigma_\varepsilon^2\left(\frac{T_p(\theta-1)^2}{(\phi-1)^2} + \frac{2\phi\left(1+\phi^{T_p}(\theta-1)(\phi-\theta)\right)}{(\phi-1)^3} + \frac{(\theta-\phi)^2\phi^{2+2T_p} - 1 + \theta\phi\left(\theta(2+\phi)-2-4\phi\right)}{(\phi-1)^3(1+\phi)}\right) \end{aligned} \tag{20}$$

(20) is non-decreasing in T_p. As $ns_0 + d_0 = 0$, the long run variance of $ns_t + d_t$ is given by

$$\begin{aligned} \sigma_{ns+d}^2 &= \sigma_\varepsilon^2\left(\sum_{t=1}^{T_p}\left(\left(\phi-\theta\right)\left(\tfrac{\phi^t-1}{1-\phi}\right)-1+\left(\phi^{t-1}(\phi-\theta)\right)\right)^2 + \sum_{t=T_p+1}^{\infty}\left(\phi^{t-1}(\phi-\theta)\right)^2\right) \\ &= \sigma_\varepsilon^2\left(\frac{T_p(\theta-1)^2}{(\phi-1)^2} + \frac{2(\theta-1)\phi^{T_p}(\phi-\theta)}{(\phi-1)^3} - \frac{(\theta-\phi)^2\phi^{2T_p}}{\phi^2-1} + \frac{(\theta-\phi)\left(\theta-2+\phi(2\theta-1)+\phi^{2T_p}(\theta-\phi)\right)}{(\phi-1)^3(1+\phi)}\right) \end{aligned} \tag{21}$$

and the Pearson Correlation Coefficient, $-1 \le \rho \le 1$, is given by

$$\rho = \frac{\sum_{t=0}^{\infty}(ns_t+d_t)(d_t)}{(\sigma_{ns+d}\sigma_d)} = \left(\frac{\dfrac{(\theta-\phi)^2\phi^{2T_p}}{1-\phi^2} + \dfrac{(\phi-\theta)\left(\phi^{T_p}-1\right)\left(1-\phi^{T_p+1}+\theta\left(\phi^{T_p}-\phi\right)\right)}{(\phi-1)^2(1+\phi)}}{(\sigma_{ns+d}\sigma_d)}\right). \tag{22}$$

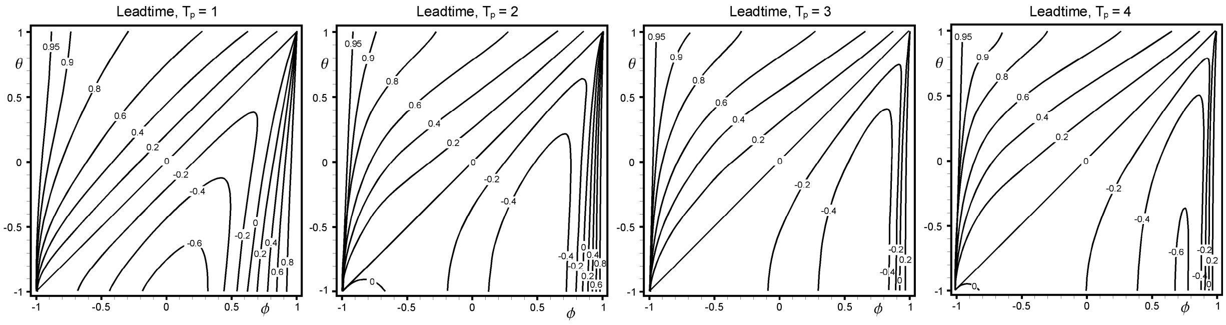

Figure 1. The Pearson Correlation Coefficient for the OUT policy under ARMA(1,1) demand

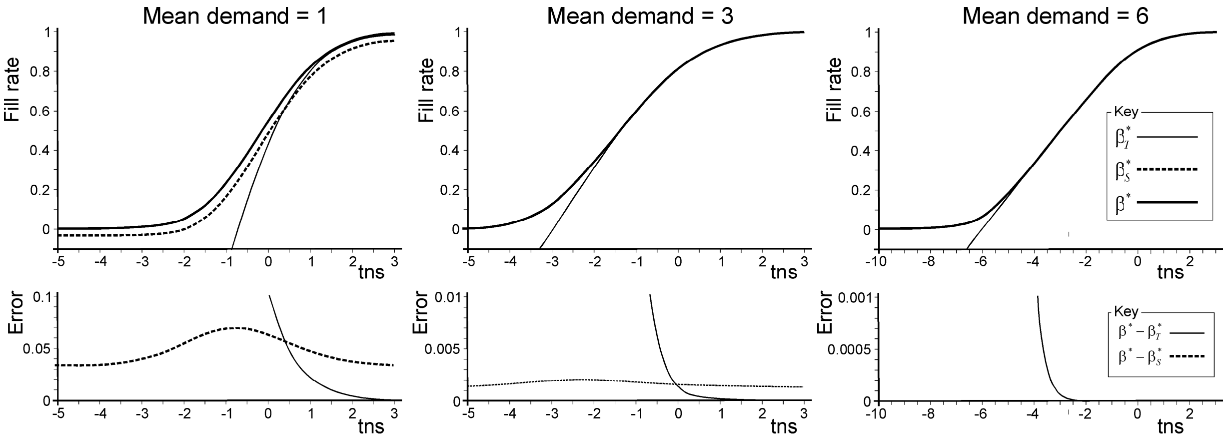

Figure 2. OUT policy fill rates with unit variance normally distributed i.i.d. demands

The Pearson Correlation Coefficient is plotted in Figure 1. Of note here is an odd-even lead time effect near $\{\phi,\theta\} \approx -1$ and the industrially prevalent ARMA(1,1) coefficients of { $\phi > 0.5$, $\theta > 0.5$ } exhibit a relatively low Pearson Correlation Coefficient.

6. Investigation of fill rate measures

Let us now investigate the performance of the various fill rate measures. Throughout this section and section 7 we assume the lead-time, $T_p = 1$, and the variance of the white noise process has been scaled to ensure the demand variance in all cases is unity. This is to allow for a fair comparison of fill rates with different correlations.

First, let us consider the case of i.i.d. demands, see Figure 2, where the upper three plots detail the fill rate with different μ_d and the lower three plots highlight the error between the exact fill rate β^* and the approximations $\{\beta_T, \beta_S\}$. In all cases the x-axis illustrates the safety stock *tns*. We can see that the traditional fill rate, β_T , is indeed a lower bound, but accuracy improves when the fill rate approaches 100%. β_S does well when the probability of negative demand is very low (when $\mu_d \geq 3$), but suffers serious errors when $\mu_d = 1$, even producing some negative fill rates.

Figure 3 illustrates the case of ARMA(1,1) demand at $\theta = \{-0.9, 0, 0.9\}$ with the mean demand $\mu_d = \{1,3\}$ and a safety stock of *tns* = 1. When $\mu_d = 3$, the three fill rate measures that change with the standard deviation of the net stock are so close together that they can not be distinguished from each other in the first row of figures. There is, however, a small slight error

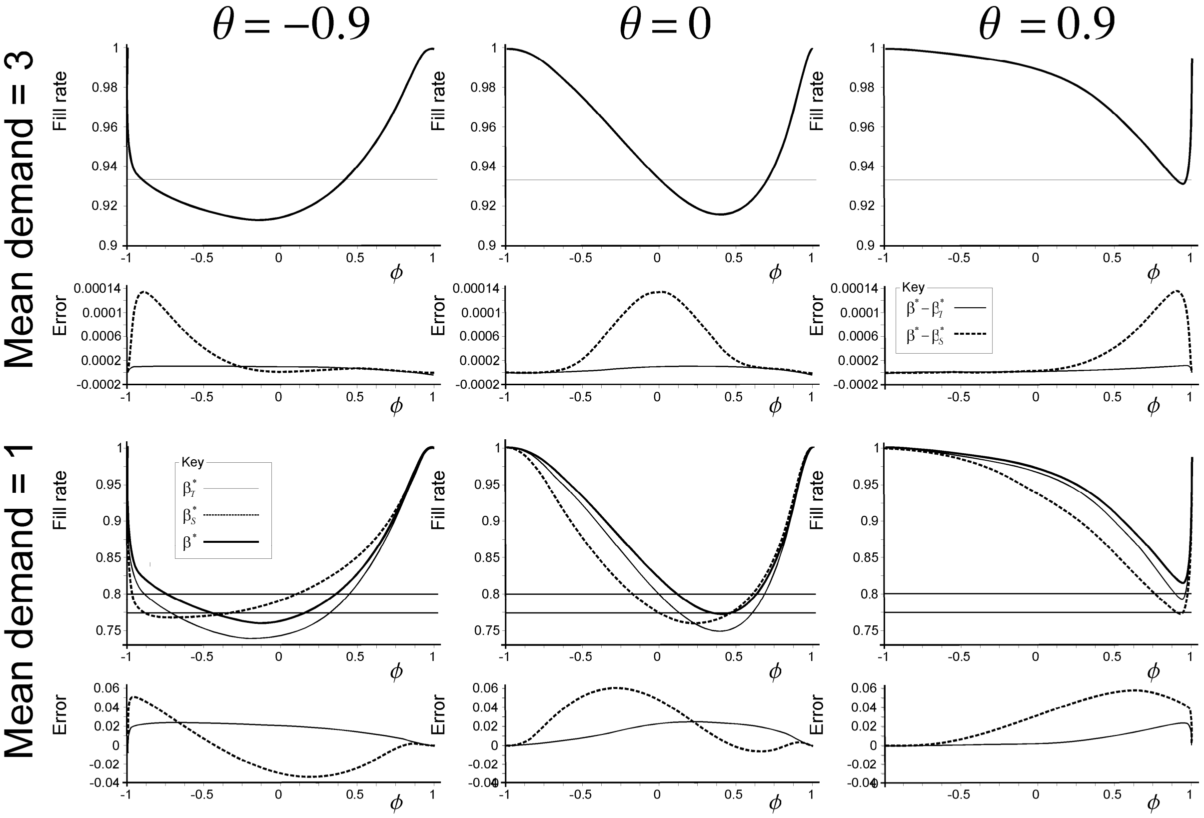

Figure 3. OUT policy fill rates with unit variance normally distributed ARMA(1,1) demands

for the two approximations $\{\beta_T^*, \beta_S^*\}$. Interestingly, the Sobel fill rate, often revered to be the "exact" solution often has the highest errors. The two fill rate measures based on incorrectly assuming i.i.d. demand, $\{\beta_T, \beta_S\}$ result in indistinguishable horizontal lines, invariant to the ARMA(1,1) parameters as they make no attempt to account for the correlation in demand. The $\{\beta_T, \beta_S\}$ errors have not been plotted in the graphs in the second (and forth) rows as they would dominate the figure. When $\mu_d = 1$, the effect of the negative demand has a much greater impact. The β_S^* measure again has often the largest errors over ϕ. The β_T^* fill rate, again dispute its simplicity and known issues for low fill rates performs rather well. The upper of the two horizontal lines in the third row of figures represent the β_T measure, the lower lines represent the β_S measure.

Figure 4 shows a contour plot of the β^* in the ARMA plane. Here *tns* = 1 and we have plotted the cases of $\mu_d = \{1,3\}$. Interestingly, there are instances of 100% fill rate when μ_d for $(\phi \approx -1,\ \theta > 0)$. Furthermore the plot shows that somewhat surprisingly the i.i.d. case has some of the lowest fill rates in the whole ARMA(1,1) parameter plane. Indeed ARMA(1,1) demand with a high Pearson Correlation Coefficient generally have high fill rates.

7. Numerical verification via simulation

We simulated the OUT policy reacting to various ARMA(1,1) demand patterns for 10,000 time periods and replicated our study 1000 times; we report the average of these 1000 replications, see Table 1. The numbers in bold are the pure $\{\beta_T, \beta_S\}$ fill rates based solely on i.i.d. demand that do not take into account the correlation in demand (see Test 4 and 8). We can see that when there is

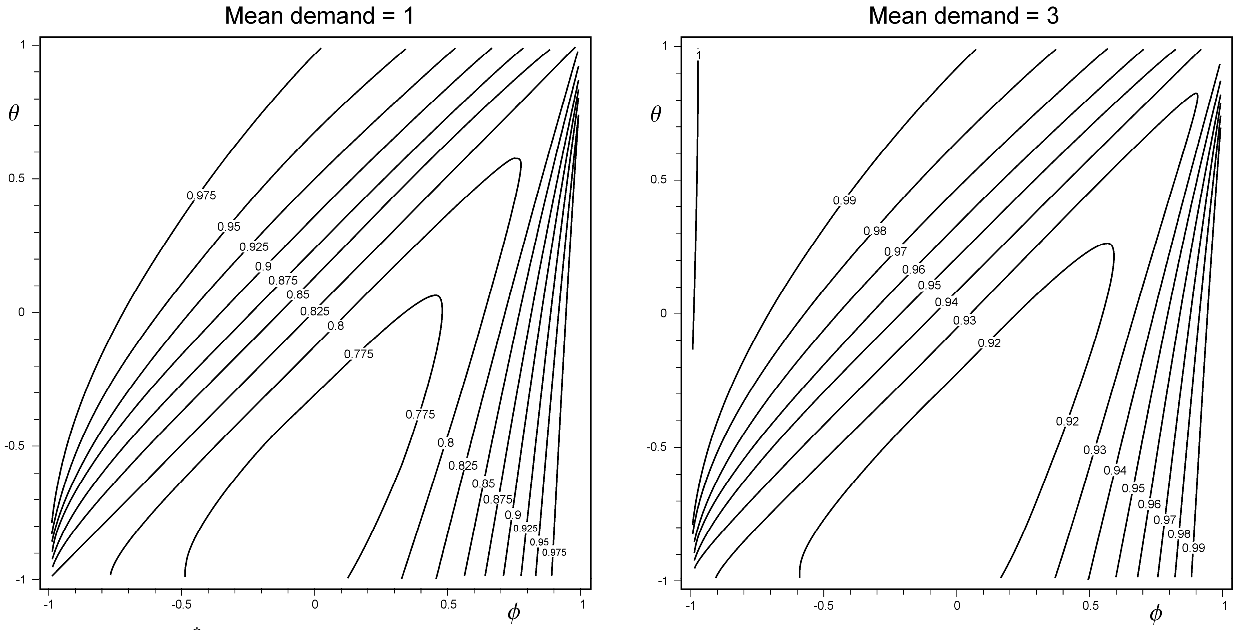

Figure 4. β^* *fill rates maintained by the OUT policy with unit variance ARMA*(1,1) *demand with a safety stock of tns* = 1

Test	μ_D	*tns*	ϕ	θ	Trad. fill rate, β_T^*		Sobel fill rate, β_S^*		Bi-variant fill rate, β^*	
					Simulation (Eq 6)	Theory (Eq 8)	Simulation (Eq 9)	Theory (Eq 11)	Simulation (Eq 13)	Theory (Eq 14)
1	-2	3	0	0	1.0042853	1.004312	1.00107	−0.033589	0.739968	0.737554
2	1	-2	0	0	-1.049922	-1.05025	-0.02522	0	0.053738	0.053713
3	1	0	0	0	0.4357787	0.43581	0.511815	0.486065	0.549456	0.54943
4	1	1	0	0	0.80062	**0.800359**	0.807976	**0.775789**	0.822962	0.82277
5	1	3	0	0	0.9913637	0.991377	0.991691	0.958323	0.992036	0.992046
6	3	-2	0	0	0.3166355	0.316582	0.344353	0.344227	0.344454	0.344423
7	3	1	-0.98	0.99	1	1	1	1	1	1
8	3	1	0	0	0.9334444	**0.933453**	0.933422	**0.933329**	0.933459	0.933464
9	3	1	0.5	-0.9	0.9383143	0.93822	0.938292	0.938221	0.938327	0.938228
10	3	1	0.7	0.5	0.9240384	0.923995	0.923994	0.923919	0.924053	0.924
11	3	1	0.9	-0.5	0.9881204	0.988115	0.987578	0.988115	0.988118	0.988117
12	3	5	0	0	0.9999749	0.999976	0.999871	0.99985	0.999975	0.999976
13	3	1	0.99	0.7	0.991278	0.991284	0.991214	0.991285	0.991275	0.991287
14	1	1	0.99	0.7	0.973914	0.973855	0.972867	0.972727	0.977039	0.977172

Table 1. Numerical verification of the three fill rate measures

a significant chance of negative demand $\{\beta_T^*, \beta_S^*\} > 1$, indicating the erroneous advice that fill rate is above 100% (see Test 1 and 11). For very low fill rates (see Test 2) , $\{\beta_T^*, \beta_S^*\} < 0$, another illogical result. Note, the order-up-to level is $S = \mu_d\left(1+T_p\right)+tns$, so for Test 2, $S = 0$, a perfectly feasible scenario. In Test 11, there is little chance of a negative demand, and the safety stock maintains high fill rates, but the Sobel fill rate is above 100%. Note {48.4%, 53.9%, 50.5%, 41.2%} of the 1000 replications of the 10,000 time period simulation resulted (in Tests 7, 11, 12 and 14 respectively) in a Sobel fill rate of above 100%. Test 13 investigates an ARMA(1,1) very close to the IMA(0,1,1) demand pattern which would be optimally forecasted by exponential smoothing. Here we can see that high fill rate achieved and the $\{\beta_T^*, \beta^*\}$ measures perform well, although β_S^* seems rather unreliable.

8. Concluding remarks

We have explored several fill rate measures from the literature. We have presented a new fill rate measure based on integrating the bi-variate normal distribution for the case of normally distributed (possibly negative) demands. When the mean demand is large (i.e. negative demand is negligible), all fill rate measures work reasonably well when operating near 100% fill rate. The impact of the demand correlation is not that significant. On the other hand, in cases where there could be negative demand, irrespective of the correlation in demand, β^* should be used. Surprising the β_T fill rate outperforms the β_S fill rate in almost all cases. For future work we could explore the use of this new fill rate measure in other replenishment policies (such as the (*R,Q*) policy or (*s,S*) policy, the Proportional OUT policy, or the Full-State policy, Gaalman (2006)). Investigations on the inverse of our new fill rate could also be undertaken, but in principle this could easily be achieved in Excel with the Solver function given that β^* is now available in Excel.

9. References

Anon, 2012. Romberg's method. www.wikipedia.org/wiki.Rombergs%27s_method. Accessed 26/8/12.

Banerjee, A. and Paul, A., 2005. Average fill rate and horizon length. *Operations Research Letters,* 33 (5), 525-530.

Box, G.E.P. and Jenkins, G.M., 1976. *Time series analysis: forecasting and control*. Holden-Day, San Francisco.

Cain, M., 1994. The moment-generating function of the minimum of bi-variate normal random variables. *The American Statistician*, 48 (2), 124-125.

Cardós, M., and Babiloni, E., 2011. Exact and approximate calculation of the cycle service level in period review inventory policies", *International Journal of Production Economics*, 131 (1), 63-68.

Chen, J., Lin, D.K.J., and Thomas, D.J. 2003. On the single item fill rate for a finite horizon, *Operations Research Letters*, 31 (2), 119-123.

Gaalman, G.J.C., 2006. Bullwhip reduction for ARMA demand: The proportional order-up-to policy versus the full-state-feedback policy. *Automatica* 42 (8), 1283-1290.

Guijarro, E., Cardós, M. and Babiloni, E., 2012. On the exact calculation of the fill rate under discrete demand patterns. *European Journal of Operational Research*, 218 (2), 442-444.

Hadley, G.F. and Whitin, T.M., 1963. *Analysis of Inventory Systems*. Prentice Hall, Englewood.

Hosoda, T. and Disney, S.M., 2009. Impact of market demand mis-specification on a two-level supply chain. *International Journal of Production Economics*, 121 (2), 739-751.

Johnson, M.E., Lee, H.L. and Davis, T., 1995. Expressions for item fill rates in periodic inventory systems. *Naval Research Logistics*, 42 (1), 57-80.

Kwon, I.-H., Kim, S.-S. and Baek, J.-G., (2006). A simulation based heuristic for serial inventory systems under fill-rate constraints. *International Journal of Advanced Manufacturing Technology.* 31 (3-4), 297-304.

Larsen, C. and Thorstenson, A., 2008. The order and volume fill rates in inventory control. *Journal of the Operational Research Society*, 59, 798-804.

Larsen, C. and Thorstenson, A., 2014. The order and volume fill rates in inventory control systems. *International Journal of Production Economics*, 147, 13-19.

Nadarajah, S. and Kotz, S., 2010. Exact Distribution of the Max/Min of Two Gaussian Random Variables. *IEEE Transactions on Very Large Scale Integration Systems*, 16 (2), 210-212.

Samii, A.-M., Pibernik, R. and Yadav, P., 2011. An inventory reservation problem with nesting and fill rate-based performance measures. *International Journal of Production Economics,* 133 (1), 393-402.

Silver, E.A, Pyke, D.F. and Peterson, R. 1998. *Inventory Management and Production Planning and Scheduling*. John Wiley and Sons, New York.

Silver, E.A. and Bischak, D.P., 2011. The exact fill rate in a periodic review base stock system under normally distributed demand. *OMEGA*, 39, 346-349.

Sobel, M.J., 2004. Fill rates of single stage and multistage supply systems. *Manufacturing and Service Operations Management*, 6 (1), 41-52.

Teunter, R., 2009. Note on the fill rate of single-stage general periodic review inventory systems. *Operations Research Letters*, 37, 67-68.

Tijms, H. C. and Groenevelt, H., 1984. Simple approximations for the reorder point in periodic and continuous review (s, S) inventory systems with service level constraints. *European Journal of Operational Research*, 17 (2), 175-190.

Tyworth, J. E. and L. O'Neill. 1997. Robustness of the normal approximation of lead time demand in a distribution setting. *Naval Research Logistics*. 44, 166–186.

Van Donselaar, K.H. and Broekmeulen, R.A.C.M. 2013. Determination of safety stocks in lost sales inventory systems with period review, positive lead time, lot sizing and target fill rate. *International Journal of Production Economics*, 143 (2), 440-448.

Weisstein, E.W., 2012. Normal Distribution. From MathWorld - A Wolfram Web Resource. http://mathworld.wolfram.com/NormalDistribution.html. Accessed 7/1/12.

Zhang, J. and Sobel, M.J., 2012. Interchangeability of fill rate constraints and backorder costs in inventory models. *International Journal of Mathematics in Operational Research*, 4 (4), 453-472.

Zhang, J. and Zhang, J., 2007. Fill rate of single stage general periodic review inventory systems. *Operations Research Letters*, 35, 503-509.

Appendix A. An Excel Add-in for determining the Fill rate

An Excel Add-in that uses the moments of the minimum of a bi-variate normal random variables (Cain, 1994), and Romberg's method (Anon, 2012), to numerically estimate the definite integral in (14) between

$$a \le f(y) \le b; \; \left\{ a = \left(\mu_{f(y)} - 6\sigma_{f(y)} \right)^{+}, \; b = \left(\mu_{f(y)} + 6\sigma_{f(y)} \right)^{+} \right\} \tag{23}$$

is given below in Table 2. Romberg's method was chosen due to its stability and accuracy. In big O notation, the error for estimate $R(n,m)$ is $O\left(\left(\frac{b-a}{2^n}\right)^{2m+2}\right)$ where $\{a,b\}$ where given in (23) and $\{n,m\}=10$ in the VB code in Table 2. When the code below is cut and pasted into an Excel function module then the expression "=Fillrate($\mu_{ns+d}, \sigma_{ns+d}, \mu_d, \sigma_d, \rho$)" is available in Excel.

VB code required to determine the fill rate

```
Option Explicit

Function fy(mu1 As Double, sigma1 As Double, mu2 As Double, sigma2 As Double, rho As Double, y As Double)
Dim f1part, f2part As Double

f1part = ((-(y - mu2) / sigma2) + rho * ((y - mu1) / sigma1)) / ((1 - rho ^ 2) ^ 0.5)
f2part = ((-(y - mu1) / sigma1) + rho * ((y - mu2) / sigma2)) / ((1 - rho ^ 2) ^ 0.5)
fy = (1 / sigma1) * WorksheetFunction.NormDist((y - mu1) / sigma1, 0, 1, False) * WorksheetFunction.NormDist(f1part, 0, 1, True) +
(1 / sigma2) * WorksheetFunction.NormDist((y - mu2) / sigma2, 0, 1, False) * WorksheetFunction.NormDist(f2part, 0, 1, True)

End Function

Function Fillrate(mu1 As Double, sigma1 As Double, mu2 As Double, sigma2 As Double, rho As Double)
Dim R(10, 10) , h, f, a, b, m1, m2, theta, y, var, s, d As Double
Dim n, m, k As Integer

theta = (sigma2 ^ 2 - 2 * rho * sigma1 * sigma2 + sigma1 ^ 2) ^ 0.5
m1 = mu1 * WorksheetFunction.NormDist((mu2 - mu1) / theta, 0, 1, True) + mu2 * WorksheetFunction.NormDist((mu1 - mu2) / theta,
0, 1, True) - theta * WorksheetFunction.NormDist((mu2 - mu1) / theta, 0, 1, False)
m2 = (sigma1 ^ 2 + mu1 ^ 2) * WorksheetFunction.NormDist((mu2 - mu1) / theta, 0, 1, True) + (sigma2 ^ 2 + mu2 ^ 2) *
WorksheetFunction.NormDist((mu1 - mu2) / theta, 0, 1, True) - (mu1 + mu2) * theta * WorksheetFunction.NormDist((mu2 - mu1) /
theta, 0, 1, False)
var = m2 - m1 ^ 2

If m1 - 6 * var ^ 0.5 < 0 Then
  a = 0
Else
  a = m1 - 6 * var ^ 0.5
End If

If m1 + 6 * var ^ 0.5 < 0 Then
  b = 0
Else
  b = m1 + 6 * var ^ 0.5
End If

For n = 0 To 10
  h = (b - a) / 2 ^ n
  If n = 0 Then
    R(0, 0) = 0.5 * (b - a) * (fy(mu1, sigma1, mu2, sigma2, rho, a) * a + fy(mu1, sigma1, mu2, sigma2, rho, b) * b)
  Else
  For m = 0 To n
    If m = 0 Then
       s = 0
       For k = 1 To 2 ^ (n - 1)
           s = s + fy(mu1, sigma1, mu2, sigma2, rho, a + (2 * k - 1) * h) * (a + (2 * k - 1) * h)
       Next k
       R(n, m) = 0.5 * R(n - 1, 0) + h * s
    Else
        R(n, m) = R(n, m - 1) + (1 / (4 ^ m - 1)) * (R(n, m - 1) - R(n - 1, m - 1))
    End If
  Next m
  End If
Next n

d = 0.5 * (mu2 + Exp(-(mu2 ^ 2 / (2 * sigma2 ^ 2))) * sigma2 * 0.797884560802865 + mu2 * (2 *
Application.WorksheetFunction.NormDist((mu2 / sigma2), 0, 1, True) - 1))

If Abs(R(9, 9) - R(10, 10)) > 0.000000001 Then
  Fillrate = "The integral has not converged to within 0.000000001"
Else
  Fillrate = R(10, 10) / d
End If

End Function
```

Table 2. VB code for the fill rate with non-negative, correlated normally distributed demands