Moments of a Regime-Switching Stochastic Interest Rate Model With Randomized Regimes

James G. Bridgeman, FSA
University of Connecticut 196 Auditorium Rd.U-3009 Storrs CT 06269-3009
bridgeman@math.uconn.edu

December 21, 2007

Abstract

Prior work indicates that a regime-switching stochastic model with randomized regime parameters creates a more plausible set of extreme paths than do the usual stochastic interest rate models. Generalizing the Black-Karasinski model by randomizing the mean reversion target provides an example. Better to understand such models, as well as to calibrate their parameters without trial and error runs of a stochastic model, we use an asymptotic expansion to estimate the moments of the integrated stochastic process over time. Mention is made of a family of related asymptotic expansions that could be of wider interest.

1 Background

A previous paper [1] motivated and introduced a stochastic process for interest rates that generalizes the Black-Karasinski model by making the mean-reversion target in the Black-Karasinski a random variable. That paper contained an expression for the mean (and, implicitly, higher moments) of the resulting integrated interest rate distribution at a fixed point in time but failed to simplify that expression into closed forms.

This paper develops a technique for approximating the moments in closed forms, simplifies some expressions in [1] and corrects a confusing error.

The stochastic process (introduced at (2.6.1) in [1]) is:

$$d\ln\left(\mathbf{r}_t\right) = \left[1-(1-F)^{dt}\right]\left[\sum_{j=0}^{\infty}\mathbf{1}_{[\mathbf{j},\mathbf{j+1})}(t)\ln(\mathbf{T}_j)-\ln(\mathbf{r}_{t-dt})\right] + (1-F)^{dt}D_t dt + (1-F)^{dt}\sigma\sqrt{dt}\mathbf{N}_t \tag{1.1}$$

where

$\mathbf{r}_t$ is the interest rate we want to model over time.

F is an annualized mean reversion factor between 0 and 1.

dt is a discrete time-step interval.

$\mathbf{1}_{[\mathbf{j},\mathbf{j+1})}(t)$ is the indicator for t to be in a regime, a random interval $[\mathbf{t_j}, \mathbf{t_{j+1}})$

$\{\mathbf{t_{j+1}} - \mathbf{t_j}\}_{1\leq j}$ are i.i.d. random variables with common law $gamma(\alpha, \beta)$, the inter-arrival intervals for regime-switches.

$\mathbf{t_1}$ is independent of $\{\mathbf{t_{j+1}} - \mathbf{t_j}\}_{1\leq j}$ and distributed as a randomly chosen point within a $gamma(\alpha, \beta)$ interval, and $t_0 = 0$, so the process begins at a random time within the first regime.

$\{\ln(\mathbf{T}_j)\}_{1\leq j}$ are i.i.d. normal random variables, independent of the $\{\mathbf{t_{j+1}} - \mathbf{t_j}\}_{0\leq j}$, making $\{\mathbf{T}_j\}_{1\leq j}$ a set of i.i.d. lognormal mean reversion targets for the interest rate, thus characterizing each regime by a randomly chosen mean reversion target (T_0 is a fixed initial target value during the first regime).

D_t is an annualized drift-compensation function available to be determined up front as part of the model.

σ is an annualized volatility parameter.

$\{\mathbf{N}_t\}_{0\leq t}$ are i.i.d standard normal random variables independent of all the other random variables in the process.

For a continuous model just think of $dt \to 0$ in (1.1), replace $\sqrt{dt}\mathbf{N}_t$ with a standard Wiener process $\mathbf{dW}_t$ and use the Taylor expansion of $(1-F)^{dt}$, ignoring $dt\mathbf{dW}_t$, dt^2 and higher, to get

$$d\ln(\mathbf{r}_t) = \left\{ -\ln(1-F)\left[\sum_{j=0}^{\infty} \mathbf{1}_{[j,j+1)}(t)\ln(T_j) - \ln(\mathbf{r}_t)\right] + D_t \right\} dt + \sigma \mathbf{dW}_t \tag{1.2}$$

The usual Black-Karasinski would be this continuous case, but with fixed mean-reversion target instead of this random one. See [7], for example.

[1] integrated the discrete time-step case to:

$$\begin{aligned}\ln(\mathbf{r}_t) &= \ln(r_0)\,(1-F)^t + \sigma\sqrt{dt}\sum_{s=1}^{\frac{t}{dt}}\mathbf{N}_{t-(s-1)dt}\,(1-F)^{sdt} \\ &\quad + \ln(T_0)\left[(1-F)^{(t-\mathbf{t}_1)_+} - (1-F)^t\right] \\ &\quad + \sum_{j=1}^{\infty}\ln(\mathbf{T}_j)\left[(1-F)^{(t-\mathbf{t}_{j+1})_+} - (1-F)^{(t-\mathbf{t}_j)_+}\right] \\ &\quad + dt\sum_{s=1}^{\frac{t}{dt}} D_{t-(s-1)dt}\,(1-F)^{sdt} \qquad (1.3)\end{aligned}$$

(the slightly more complicated expression (2.6.2) in [1] reduces to (1.3) upon ignoring the set of sample paths of measure zero for which some model time-step t coincides with a random regime-switch point in time $\mathbf{t}_j$.)

The problem posed at the conclusion of [1] was to find closed forms for the mean and (by implication) higher moments of the random variable $\mathbf{r}_t$. These would facilitate calibration and interpretation of the model along the lines already performed in [1] for the traditional mean-reverting lognormal model with fixed reversion target T:

$$\begin{aligned}d\ln(\mathbf{r}_t) &= \left[1-(1-F)^{dt}\right]\left[\ln(T) - \ln(\mathbf{r}_{t-dt})\right] \\ &\quad + (1-F)^{dt} D_t dt + (1-F)^{dt}\sigma\sqrt{dt}\mathbf{N}_t \qquad (1.4)\end{aligned}$$

This integrated to

$$\begin{aligned}\ln(\mathbf{r}_t) &= \ln(r_0)\,(1-F)^t + \sigma\sqrt{dt}\sum_{s=1}^{\frac{t}{dt}}\mathbf{N}_{t-(s-1)dt}\,(1-F)^{sdt} \\ &\quad + \ln(T)\left[1-(1-F)^t\right] + dt\sum_{s=1}^{\frac{t}{dt}} D_{t-(s-1)dt}\,(1-F)^{sdt} \qquad (1.5)\end{aligned}$$

(see (2.2.4) in [1]) and is analogous to (1.3.) above.

In this traditional mean-reverting lognormal model with fixed reversion target T, requiring that the model interest rate series display an observed volatility (after mean reversion) of σ^2_{obs}, based perhaps on historically observed volatility of $(\ln\mathbf{r}_t - \ln\mathbf{r}_{t-dt})$, and an observed variance V_{obs} of integrated rates $\mathbf{r}_t$ across sample paths for a fixed t in the model, based perhaps on historically observed variance of the distribution of $\mathbf{r}_t$ over a long period of time, produced a requirement that the annualized volatility parameter value σ and the annualized mean-reversion parameter value F in the model satisfy,

$$\begin{aligned}F &= 1-\left\{1-\frac{\sigma^2_{obs}dt}{\ln\left(V_{obs}+T^2\right)-\ln(T^2)}\right\}^{\frac{1}{2dt}} \text{ and} \qquad (1.6) \\ \sigma &= \frac{\sigma_{obs}}{(1-F)^{dt}} \qquad (1.7)\end{aligned}$$

For the continuous model $dt \to 0$, the identity $lim_{\epsilon \to 0} (1 + a\epsilon)^{1/\epsilon} = e^a$ applied to (1.6) produces

$$F = 1 - e^{-\frac{1}{2}\sigma_{obs}^2 \left[\ln(V_{obs}+T^2)-\ln(T^2)\right]^{-1}} \tag{1.8}$$

for the value of the annualized mean-reversion factor F required to reproduce V_{obs} and σ_{obs}^2. (1.6)-(1.8) correct errors in formulae (2.3.5) and (2.3.6) in [1], which erroneously conflated the roles of V_{obs} and σ_{obs}^2.

Another application of (1.4) developed in [1] determined the drift compensation function D_t required in the traditional mean-reverting lognormal model with reversion target T to produce the intuitively desirable relationship

$$\mathbb{E}[\mathbf{r}_t] = r_0^{(1-F)^t} T^{\left[1-(1-F)^t\right]} \tag{1.9}$$

where r_0 is the starting value of the interest rate in question. According to (2.2.8) and (2.2.10) in [1] the condition that (1.9) be true is that the drift-compensation function D_t be given the (perhaps) surprising form:

$$D_t = -\frac{1}{2}\sigma^2 \frac{(1-F)^{dt}}{1+(1-F)^{dt}} \left[1 + (1-F)^{2t-dt}\right], \text{ or} \tag{1.10}$$

$$D_t = -\frac{1}{4}\sigma^2 \left[1 + (1-F)^{2t}\right] \text{ in the continuous case.} \tag{1.11}$$

Derivation of results corresponding to (1.6)-(1.11) in the case of randomized mean-reversion targets, (1.1) or (1.2), requires knowledge of the variance and higher moments of the random variable $\mathbf{r}_t$ for that case. Such knowledge will also allow fitting to higher moment conditions analogous to those that led to (1.6)-(1.8). This will be valuable because empirical evidence to be reviewed in a later paper suggests that the historical distribution of interest rates displays lower 4th and 6th moments than the lognormal that results from integrating a mean-reverting lognormal model with fixed mean-reversion target, while the historical volatility of interest rates displays higher 4th and 6th moments than the volatility of the lognormal model with fixed mean-reversion target.

The balance of this paper will derive closed-form approximations for all of the moments of the random variable $\mathbf{r}_t$ in the case of randomized mean-reversion targets (1.1) or (1.2). A family of related approximation techniques that could be of wider interest to actuaries and financial engineers is mentioned, as well. We will concentrate on the discrete time-step model (1.1) since we can pass to (1.2) at any point by taking the limit as $dt \to 0$.

2 Integration for the Moments

Moments for the mean-reverting lognormal model (1.4) with fixed mean-reversion target T fell out easily in [1] because the expression for $\ln(\mathbf{r}_t)$ given at (1.5) is just a (complicated) sum of constants and constants times standard normal random variables, so $\ln(\mathbf{r}_t)$ is a normal random variable and $\mathbf{r}_t$ is a lognormal random variable, with μ and σ^2 parameters determined by the constants in (1.5). The resulting expressions are complicated but tractable. The standard expression

$$\mathbb{E}\left[(\mathbf{r}_t)^l\right] = e^{l\mu+\frac{1}{2}(l\sigma)^2} \tag{2.1}$$

for the moments of a lognormal thus becomes available, with complicated μ and σ parameters determined by the constants in (1.5). This was the basis for the results in [1] and in section 1 above about the mean-reverting lognormal with fixed mean-reversion target T.

This approach does not go through for the model (1.1) with random mean-reversion targets because the expression for $\ln(\mathbf{r}_t)$ given at (1.3) is not just a sum of constants and constants times normal random variables. The terms

$$\ln(\mathbf{T}_j)\left[(1-F)^{(t-\mathbf{t}_{j+1})_+} - (1-F)^{(t-\mathbf{t}_j)_+}\right]$$

have expressions involving the random variables $\mathbf{t}_{j+1}$ and $\mathbf{t}_j$, which are sums of gamma inter-arrival intervals for regime-switches, exponentiated and multiplied by the normal random variables $\ln(\mathbf{T}_j)$. The term

$$\ln(T_0)\left[(1-F)^{(t-\mathbf{t}_1)_+} - (1-F)^t\right]$$

involves a random variable $\mathbf{t}_1$ that is a random time within the first gamma regime. The cut-off at t in the $(t-\mathbf{t}_j)_+$ expressions ensures that the moments of the exponentiated expressions exist, but it adds to the challenge of actually calculating them. What to do?

First, condition on the random variables $\{\mathbf{t}_j\}_{j\geq 1}$. Then the expression for $\ln(\mathbf{r}_t)$ given at (1.3) is just a (complicated) sum of constants and constants times normal random variables, including both $\{\mathbf{N}_j\}$ and $\{\ln(\mathbf{T}_j)\}$. Conditional on $\{\mathbf{t}_j\}_{j\geq 1}$, then, $\ln(\mathbf{r}_t)$ is normal and $\mathbf{r}_t$ is lognormal, with complicated parameters involving the random variables $(1-F)^{(t-\mathbf{t}_1)_+}$ and

$$\left\{(1-F)^{(t-\mathbf{t}_{j+1})_+} - (1-F)^{(t-\mathbf{t}_j)_+}\right\}_{j\geq 1}.$$

This doesn't buy us much, because the unconditioned expectations will require a formidable analysis to unravel expectations involving those complicated parameters, derived from the gamma inter-arrival structure for regime-switches.

But next, anticipate the empirical conclusion to be presented in a later paper that $\mathbf{r}_t$ has smaller high moments (hence less tail weight) than a lognormal. In that case, an Edgeworth expansion for $\ln(\mathbf{r}_t)$ might provide a reasonable approximation for the moments of $\mathbf{r}_t$ along the lines of (2.1). In section 5 we derive that expansion and approximation. The first few terms are

$$\mathbb{E}\left[(\mathbf{r}_t)^l\right] \approx e^{l\mu+\frac{1}{2}(l\sigma)^2}\left\{1+\frac{l^4}{4!}\left[\mu_4-3\sigma^4\right]\right\} \tag{2.2}$$

$$\approx e^{l\mu+\frac{1}{2}(l\sigma)^2}\left\{1+\frac{l^4}{4!}\left[\mu_4-3\sigma^4\right]\left(1-\frac{3}{4!}(l\sigma)^2\right)+\frac{l^6}{6!}\left[\mu_6-15\sigma^6\right]\right\} \tag{2.3}$$

and the general expression is

$$\mathbb{E}\left[\left(\mathbf{r}_t\right)^l\right] =$$

$$= \quad e^{l\mu+\frac{1}{2}(l\sigma)^2}\left\{1+\lim_{N\to\infty}\sum_{j=2}^{N}\frac{l^{2j}}{(2j)!}\left[\mu_{2j}-(2j)?\sigma^{2j}\right]\sum_{n=0}^{N-j}\frac{(-1)^n\,(2n)?}{(2n)!}\,(l\sigma)^{2n}\right\} \tag{2.4}$$

where $(2n)? = (2n-1)(2n-3)\cdots(1)$ and μ, σ^2, and μ_{2j} are the mean, variance, and higher central moments of $\ln(\mathbf{r}_t)$. We hope to be able to calculate these moments by conditioning on $\{\mathbf{t}_j\}_{j\geq 1}$. Already, we have been able to exclude the odd central moments of $\ln(\mathbf{r}_t)$ that would otherwise appear in (2.2)-(2.4) because, conditional on $\{\mathbf{t}_j\}_{j\geq 1}$, $\ln(\mathbf{r}_t)$ is normal and has odd central moments equal to 0, so the unconditioned odd central moments also must be 0.

We have not yet made an estimate of the approximation error in (2.2)-(2.4) so it remains unclear how many terms are needed for good convergence. One would expect that 4th or 6th moment terms should pick up most of the departure from lognormality given the the thin tails and broad shoulders that characterize the historical distribution of $\ln(\mathbf{r}_t)$, but that needs to be verified in future work. Section 5 mentions some variants of the technique that could improve convergence if need be. They might be of interest in their own right to actuaries and financial engineers for wider application.

The parameter μ that we need in (2.2)-(2.4) is $\mathbb{E}\left[\ln(\mathbf{r}_t)\right]$. Conditioning on $\{\mathbf{t}_j\}_{j\geq 1}$ (1.3) is just constants and normals so

$$\begin{aligned}
\mathbb{E}\left[\ln(\mathbf{r}_t)\right] &= \ln(r_0)(1-F)^t + \ln(T_0)\left\{\mathbb{E}\left[(1-F)^{(t-\mathbf{t}_1)_+}\right]-(1-F)^t\right\} \\
&\quad +\mu_T\mathbb{E}\left[\sum_{j=1}^{\infty}\left[(1-F)^{(t-\mathbf{t}_{j+1})_+}-(1-F)^{(t-\mathbf{t}_j)_+}\right]\right] \\
&\quad +dt\sum_{s=1}^{\frac{t}{dt}}D_{t-(s-1)dt}\,(1-F)^{sdt} \\
&= \ln(r_0)(1-F)^t + \ln(T_0)\left\{\mathbb{E}\left[(1-F)^{(t-\mathbf{t}_1)_+}\right]-(1-F)^t\right\} \\
&\quad \mu_T\left\{1-\mathbb{E}\left[(1-F)^{(t-\mathbf{t}_1)_+}\right]\right\} \Longleftarrow \text{telescoped} \\
&\quad +dt\sum_{s=1}^{\frac{t}{dt}}D_{t-(s-1)dt}\,(1-F)^{sdt}
\end{aligned} \tag{2.5}$$

where μ_T stands for the common mean of $\{\ln(\mathbf{T}_j)\}$, the expectation of the indicated sum telescopes along with the sum by monotone convergence, and the sum telescopes (eliminating all but the first of our troublesome random variables involving $\mathbf{t}_j$) because on almost all paths $\mathbf{t}_j \geq t$ for some j and all thereafter, making the summands 0 from that point on. We'll see later that the remaining expectation $\mathbb{E}\left[(1-F)^{(t-\mathbf{t}_1)_+}\right]$ is essentially a constant times a value of a Laplace transform that we can calculate. (2.5) is the same as (2.6.5) in [1], which is as far as we got in that paper.

For the variance and higher central moments of $\ln(\mathbf{r}_t)$ that we need in (2.2)-(2.4), condition on $\{\mathbf{t}_j\}_{j\geq 1}$ so that (1.3) is just constants and independent normals and use (2.5) and the fact that central moments of a normal distribution are $(2n)?\sigma^{2n}$ where again $(2n)? = (2n-1)(2n-3)\cdots(1)$:

$$\mathbb{E}\left[\{\ln(\mathbf{r}_t) - \mathbb{E}\left[\ln(\mathbf{r}_t)\right]\}^{2n}\right] =$$

$$\begin{aligned} &= (2n)?\mathbb{E}\left[\left\{\sigma^2 dt \sum_{s=1}^{\frac{t}{dt}} (1-F)^{2sdt} + \sigma_T^2 \sum_{j=1}^{\infty} \mathbf{e}_j^2\right\}^n\right] \\ &= (2n)?\mathbb{E}\left[\left\{\sigma^2 dt\,(1-F)^{2dt}\,\frac{1-(1-F)^{2t}}{1-(1-F)^{2dt}} + \sigma_T^2 \sum_{j=1}^{\infty} \mathbf{e}_j^2\right\}^n\right] \end{aligned} \quad (2.6)$$

σ_T^2 stands for the common variance of $\{\ln(\mathbf{T}_j)\}$, for each j the random variable $\mathbf{e}_j$ is defined by ·

$$\mathbf{e}_j = \left\{(1-F)^{(t-\mathbf{t}_{j+1})_+} - (1-F)^{(t-\mathbf{t}_j)_+}\right\} \quad (2.7)$$

and we summed the geometric series that occurred. For a given application of (2.2)-(2.4) we will need only a finite number of n, maybe even just $n = 1$, 2 and 3, so it is not only correct but probably practical as well to expand the $\{\quad\}^n$ expression in (2.6) as a binomial. Then the only probabilistic calculations we will need in order to pass from conditional to unconditioned values in (2.6) are to evaluate terms of the form

$$\mathbb{E}\left[\left(\sum_{j=1}^{\infty} \mathbf{e}_j^2\right)^m\right]. \quad (2.8)$$

arising in the expectation of the binomial expansion of (2.6).

The difficulty in evaluating the expectations (2.8) is that, as mentioned earlier, $\{\mathbf{e}_j\}_{j\geq 1}$ are not normal or otherwise friendly random variables, rather they are complicated outcomes of the gamma inter-arrival structure for regime-switches. For $m > 1$ we also face a bewildering array of cross-terms (products involving different values of t_j) that would seem to defy closed form evaluation.

But we can reduce the $m > 1$ cases in (2.8) to combinations of things close to the $m = 1$ case because $\mathbf{e}_j$ and $\mathbf{e}_i$ are uniformly correlated for $i \neq j$ in the sense of the following lemma, that will be proved in section 4.5. The constants $\rho_{a_1,\ldots,a_k}$ that occur will be given closed form expression (4.5.8) that can be calculated (and, obviously, they are symmetric in the subscripts).

Lemma:

$$\mathbb{E}\left[\mathbf{e}_{j_1}^{2a_1}\cdots\mathbf{e}_{j_k}^{2a_k}\right] = \rho_{a_1,\dots,a_k}\mathbb{E}\left[\mathbf{e}_{j_1}^{2a_1}\right]\cdots\mathbb{E}\left[\mathbf{e}_{j_k}^{2a_k}\right] \text{ when no two of } \{j_1,\dots,j_k\} \text{ are equal, independent of } \{j_1,\dots,j_k\}. \quad \blacksquare \tag{2.9}$$

Consider, for example, what (2.9) does for the case $m = 2$:

$$\begin{aligned}
\mathbb{E}\left[\left(\sum_{j=1}^{\infty}\mathbf{e}_j^2\right)^2\right] &= \mathbb{E}\left[\sum_{j=1}^{\infty}\mathbf{e}_j^4 + \sum_{j=1}^{\infty}\mathbf{e}_j^2\left\{\left(\sum_{i=1}^{\infty}\mathbf{e}_i^2\right) - e_j^2\right\}\right] \\
&= \mathbb{E}\left[\sum_{j=1}^{\infty}\mathbf{e}_j^4\right] + \rho_{1,1}\sum_{j=1}^{\infty}\mathbb{E}\left[\mathbf{e}_j^2\right]\mathbb{E}\left[\left(\sum_{i=1}^{\infty}\mathbf{e}_i^2\right) - e_j^2\right], \text{ by (2.9)} \\
&= \mathbb{E}\left[\sum_{j=1}^{\infty}\mathbf{e}_j^4\right] + \rho_{1,1}\sum_{j=1}^{\infty}\mathbb{E}\left[\mathbf{e}_j^2\right]\left\{\mathbb{E}\left[\sum_{i=1}^{\infty}\mathbf{e}_i^2\right] - \mathbb{E}\left[e_j^2\right]\right\} \\
&= \mathbb{E}\left[\sum_{j=1}^{\infty}\mathbf{e}_j^4\right] + \rho_{1,1}\mathbb{E}\left[\sum_{j=1}^{\infty}\mathbf{e}_j^2\left\{\mathbb{E}\left[\sum_{i=1}^{\infty}\mathbf{e}_i^2\right] - \mathbb{E}\left[e_j^2\right]\right\}\right] \\
&= \mathbb{E}\left[\sum_{j=1}^{\infty}\mathbf{e}_j^4\right] + \rho_{1,1}\left\{\left(\mathbb{E}\left[\sum_{j=1}^{\infty}\mathbf{e}_j^2\right]\right)^2 - \mathbb{E}\left[\sum_{j=1}^{\infty}\mathbf{e}_j^2\mathbb{E}\left[e_j^2\right]\right]\right\}
\end{aligned} \tag{2.10}$$

where monotone convergence justifies moving expectations across the infinite sums.

Similarly, but already more complex, for $m = 3$:

$$\mathbb{E}\left[\left(\sum_{j=1}^{\infty}\mathbf{e}_j^2\right)^3\right] =$$

$$= \mathbb{E}\left[\begin{array}{c}\displaystyle\sum_{j=1}^{\infty}\mathbf{e}_j^6 + 3\sum_{j=1}^{\infty}\mathbf{e}_j^4\left\{\left(\sum_{i=1}^{\infty}\mathbf{e}_i^2\right) - e_j^2\right\} \\ \displaystyle + \sum_{j=1}^{\infty}\mathbf{e}_j^2\left\{\begin{array}{c}\displaystyle\left(\sum_{i=1}^{\infty}\mathbf{e}_i^2\left[\left(\sum_{k=1}^{\infty}\mathbf{e}_k^2\right) - e_i^2 - e_j^2\right]\right) \\ \displaystyle -e_j^2\left[\left(\sum_{k=1}^{\infty}\mathbf{e}_k^2\right) - e_j^2\right] + e_j^4\end{array}\right\}\end{array}\right]$$

$$
\mathbb{E}\left[\left(\sum_{j=1}^{\infty}\mathbf{e}_j^2\right)^3\right] =
$$

$$
\begin{aligned}
= \quad & \mathbb{E}\left[\sum_{j=1}^{\infty}\mathbf{e}_j^6\right] + 3\rho_{2,1}\left\{\mathbb{E}\left[\sum_{j=1}^{\infty}\mathbf{e}_j^4\right]\mathbb{E}\left[\sum_{j=1}^{\infty}\mathbf{e}_j^2\right] - \mathbb{E}\left[\sum_{j=1}^{\infty}\mathbf{e}_j^4\mathbb{E}\left[e_j^2\right]\right]\right\} \\
& +\rho_{1,1,1}\left(\mathbb{E}\left[\sum_{j=1}^{\infty}\mathbf{e}_j^2\right]\right)^3 - 3\rho_{1,1,1}\mathbb{E}\left[\sum_{j=1}^{\infty}\mathbf{e}_j^2\right]\mathbb{E}\left[\sum_{j=1}^{\infty}\mathbf{e}_j^2\mathbb{E}\left[e_j^2\right]\right] \\
& +\rho_{1,1,1}\mathbb{E}\left[\sum_{j=1}^{\infty}\mathbf{e}_j^2\left(\mathbb{E}\left[e_j^2\right]\right)^2\right] + \rho_{1,1,1}\mathbb{E}\left[\sum_{j=1}^{\infty}\mathbf{e}_j^2\mathbb{E}\left[e_j^4\right]\right] \\
= \quad & \mathbb{E}\left[\sum_{j=1}^{\infty}\mathbf{e}_j^6\right] + 3\rho_{2,1}\mathbb{E}\left[\sum_{j=1}^{\infty}\mathbf{e}_j^4\right]\mathbb{E}\left[\sum_{j=1}^{\infty}\mathbf{e}_j^2\right] - \left(3\rho_{2,1} - \rho_{1,1,1}\right)\mathbb{E}\left[\sum_{j=1}^{\infty}\mathbf{e}_j^4\mathbb{E}\left[e_j^2\right]\right] \\
& +\rho_{1,1,1}\left\{\left(\mathbb{E}\left[\sum_{j=1}^{\infty}\mathbf{e}_j^2\right]\right)^3 - 3\mathbb{E}\left[\sum_{j=1}^{\infty}\mathbf{e}_j^2\right]\mathbb{E}\left[\sum_{j=1}^{\infty}\mathbf{e}_j^2\mathbb{E}\left[e_j^2\right]\right]\right. \\
& \left. +\mathbb{E}\left[\sum_{j=1}^{\infty}\mathbf{e}_j^2\left(\mathbb{E}\left[e_j^2\right]\right)^2\right]\right\}
\end{aligned} \tag{2.11}
$$

A general formulation for any m for the corresponding configuration of

$$
\mathbb{E}\left[\left(\sum_{j=1}^{\infty}\mathbf{e}_j^2\right)^m\right]
$$

has eluded me, but it is clear that for any specific m a chain of combinations such as (2.10) or (2.11) will reduce that expectation to a combination of terms of the form

$$
\mathbb{E}\left[\sum_{j=1}^{\infty}\mathbf{e}_j^{2n}\prod_{k=1}^{n}\left(\mathbb{E}\left[\mathbf{e}_j^{2k}\right]\right)^{n_k}\right], \text{ where } \sum_{k=1}^{n} kn_k \leq m - n \tag{2.12}
$$

with constants, including the constants $\rho_{a_1,\ldots,a_k}$. Expressions (2.10) and (2.11) already give the required combinations for the cases $m = 2$ or 3 most likely to be required in practice. Note particularly that in (2.12) it is always the case that $k \leq n$.

The goal is to evaluate the expectation (2.12). Our results will allow evaluation of the slightly more general expectation (2.13) where we do not restrict ourselves to even powers. Indeed, in the proofs in section 4 it will be essential to allow odd powers (or at least the power one).

$$\mathbb{E}\left[\sum_{j=1}^{\infty}\mathbf{e}_j^n\prod_{k=1}^{n}\left(\mathbb{E}\left[\mathbf{e}_j^k\right]\right)^{n_k}\right] \tag{2.13}$$

Section 3 will set up the machinery for and then state results to be proved in section 4 that reduce the evaluation of expectations of the form (2.13) to the evaluation of Laplace transforms of the inter-arrival gamma distribution and its equilibrium distribution, together with the distribution function for the equilibrium distribution. This will allow the evaluation of expectations of the form (2.8) using (2.10), (2.11) and the corresponding expressions for higher values of m if needed for accuracy. This in turn will allow evaluation of the central moments of $\ln(\mathbf{r}_t)$ using (2.6). These central moments, together with (2.5), finally will allow a closed form approximation to the moments of $\mathbf{r}_t$ selecting the approximation from (2.2)-(2.4) depending on the accuracy desired.

3 Set Up and Statement of Results

3.1 Definitions and Notation:

Let $\mathbf{d}_1, \mathbf{d}_2, ..., \mathbf{d}_j, ...$be i.i.d inter-arrival intervals with common law $\mathbf{d}$. It is necessary to deal with the model assumption that the first model inter-arrival interval reflects a random starting point within the first i.i.d. inter-arrival interval. (At time 0 we have no way of knowing how far along we are in the current regime.) To handle this, define $\bar{\mathbf{d}}_0$ by the relationships $0 \leq \bar{\mathbf{d}}_0 \leq \mathbf{d}_1$ and $\bar{\mathbf{d}}_0 \simeq (\mathbf{d}_1 - \bar{\mathbf{d}}_0)$ using the notation $\simeq$ to mean "equal in law." Let $\bar{\mathbf{d}}$ stand for the common law of $\bar{\mathbf{d}}_0$ and $(\mathbf{d}_1 - \bar{\mathbf{d}}_0)$. Then $\bar{\mathbf{d}}$ follows the "equilibrium distribution" corresponding to the distribution of $\mathbf{d}$ (see lemma (3.7) to follow).

Now we can establish the 0 of the time parameter at the random point $\bar{\mathbf{d}}_0$ within $\mathbf{d}_1$. Define $\bar{\mathbf{d}}_1 = \mathbf{d}_1 \wedge (\bar{\mathbf{d}}_0 + t) - \bar{\mathbf{d}}_0$ and

$$\begin{aligned}
t_0 &= 0 \\
\mathbf{t}_1 &= \bar{\mathbf{d}}_1 \\
\mathbf{t}_2 &= \bar{\mathbf{d}}_1 + \mathbf{d}_2 \\
&\cdots \\
\mathbf{t}_j &= \bar{\mathbf{d}}_1 + \mathbf{d}_2 + ... + \mathbf{d}_j
\end{aligned}$$

so $\mathbf{t}_j - \mathbf{t}_{j-1} = \mathbf{d}_j$ for $j > 1$ and $\mathbf{t}_1 - t_0 = \mathbf{t}_1 = \bar{\mathbf{d}}_1$.

For a given time t define the random variable $\mathbf{J}$=min $\{j : \mathbf{t}_j \geq t\}$ (a "stopping regime") and the indicator random variables $\{\mathbf{1}_{j<\mathbf{J}}\}_{j\geq 1}$defined by $\mathbf{1}_{j<\mathbf{J}} = 0$ for

$j \geq \mathbf{J}$ and $\mathbf{1}_{j<\mathbf{J}} = 1$ for $j < \mathbf{J}$. Finally, model the fact that the time t occurs somewhere within the last inter-arrival interval by defining $\bar{\mathbf{d}}_{\mathbf{J}} = t - \mathbf{t}_{\mathbf{J}-1}$ and $\bar{\mathbf{d}}_{\mathbf{J+1}} = \mathbf{t}_{\mathbf{J}} - t$ so that

$$\mathbf{d}_1, \mathbf{d}_2, ..., \mathbf{d}_{\mathbf{J}}$$

envelops $[0, t]$ and is the minimal sequence to do so because

$$t = \bar{\mathbf{d}}_1 + \mathbf{d}_2 + ... + \mathbf{d}_{\mathbf{J}-1} + \bar{\mathbf{d}}_{\mathbf{J}}$$

where $\mathbf{d}_1$ and $\mathbf{d}_{\mathbf{J}}$ hang over the edges in the sense that $\mathbf{d}_1 = \bar{\mathbf{d}}_0 + \bar{\mathbf{d}}_1$ and $\mathbf{d}_{\mathbf{J}} = \bar{\mathbf{d}}_{\mathbf{J}} + \bar{\mathbf{d}}_{\mathbf{J+1}}$.

To save repetition whenever a specific case of (2.13) is unambiguously intended, for any expression x let $\nu(x)$ stand for $x^n \prod_{k=1}^{n} \mathbb{E}\left[x^k\right]^{n_k}$ so that (2.13) is abbreviated to $\mathbb{E}\left[\sum_{j=1}^{\infty} \nu\left(\mathbf{e}_j\right)\right]$.

Finally, define a constant G so that $\mathbb{E}\left[(1-G)^{\mathbf{d}}\right] = \mathbb{E}\left[\nu\left((1-F)^{\mathbf{d}}\right)\right]$. That is, letting $\mathcal{L}_{\mathbf{d}}(x)$ be the Laplace transform of the density for $\mathbf{d}$ evaluated at x and $\mathcal{L}_{\mathbf{d}}^{-1}(y)$ be its inverse (inverse function, not inverse transform) evaluated at y, then define

$$\begin{aligned}
(1-G) &= \exp\left\{-\mathcal{L}_{\mathbf{d}}^{-1}(\mathbb{E}\left[\nu\left((1-F)^{\mathbf{d}}\right)\right])\right\}, \text{ so} && (3.1)\\
(1-G)^{\mathbf{d}} &= \exp\left\{-\mathbf{d}\mathcal{L}_{\mathbf{d}}^{-1}(\mathbb{E}\left[\nu\left((1-F)^{\mathbf{d}}\right)\right])\right\} \\
\mathbb{E}\left[(1-G)^{\mathbf{d}}\right] &= \mathcal{L}_{\mathbf{d}}\left\{\mathcal{L}_{\mathbf{d}}^{-1}(\mathbb{E}\left[\nu\left((1-F)^{\mathbf{d}}\right)\right]\right\} \\
\mathbb{E}\left[(1-G)^{\mathbf{d}}\right] &= \mathbb{E}\left[\nu\left((1-F)^{\mathbf{d}}\right)\right] && (3.2)
\end{aligned}$$

3.2 Main Results

The goal is to evaluate the expectation (2.13).

Theorem:

$$\mathbb{E}\left[\sum_{j=1}^{\infty} \mathbf{e}_j^n \prod_{k=1}^{n} \left(\mathbb{E}\left[\mathbf{e}_j^k\right]\right)^{n_k}\right] =$$

$$\begin{aligned}
= \quad & K\left\{\mathbb{E}\left[\nu\left((1-F)^{\bar{\mathbf{d}}\wedge t}\right)\right] - \mathbb{P}\left[\bar{\mathbf{d}} \geq t\right] \nu\left((1-F)^t\right)\right\} \frac{\mathbb{E}\left[\nu\left(1-(1-F)^{\mathbf{d}}\right)\right]}{1-\mathbb{E}\left[\nu\left((1-F)^{\mathbf{d}}\right)\right]} + \\
& + \mathbb{E}\left[\nu\left(1-(1-F)^{\bar{\mathbf{d}}\wedge t}\right)\right] - \mathbb{P}\left[\bar{\mathbf{d}} \geq t\right] \nu\left(1-(1-F)^t\right) && (3.3)
\end{aligned}$$

where the constant K is

$$K = 1 - \mathbb{E}\left[\left(\mathbb{E}\left[\nu\left((1-F)^{\mathbf{d}}\right)\right]\right)^{\mathbf{J}-2} \mid \mathbf{J} > 1\right] \qquad \blacksquare$$

Lack of a closed form for the distribution of the stopping regime $\mathbf{J}$, or at least for its Laplace transform, will make it impossible to calculate with theorem (3.3). However, theorem (3.3) as a general result for all F, n, and $\{n_k\}_{k\leq n}$ will allow a proof for a more tractable corollary from a computational standpoint.

Corollary:

$$K = 1 - (1-G)^t \frac{\mathbb{E}\left[(1-G)^{-\bar{\mathbf{d}}\wedge t}\right] - \mathbb{P}\left[\bar{\mathbf{d}} \geq t\right](1-G)^{-t}}{\mathbb{E}\left[(1-G)^{\bar{\mathbf{d}}\wedge t}\right] - \mathbb{P}\left[\bar{\mathbf{d}} \geq t\right](1-G)^{t}} \blacksquare \tag{3.4}$$

$(1-G)$ is computable using (3.1) so corollary (3.4) eliminates the need to know the distribution of $\mathbf{J}$ or its Laplace transform. Putting this value for K back into theorem (3.3) we can evaluate (2.13) as

Theorem:

$$\mathbb{E}\left[\sum_{j=1}^{\infty} \mathbf{e}_j^n \prod_{k=1}^{n} \left(\mathbb{E}\left[\mathbf{e}_j^k\right]\right)^{n_k}\right] =$$

$$= \left(1 - (1-G)^t \frac{\mathbb{E}\left[(1-G)^{-\bar{\mathbf{d}}\wedge t}\right] - \mathbb{P}\left[\bar{\mathbf{d}} \geq t\right](1-G)^{-t}}{\mathbb{E}\left[(1-G)^{\bar{\mathbf{d}}\wedge t}\right] - \mathbb{P}\left[\bar{\mathbf{d}} \geq t\right](1-G)^{t}}\right) \cdot$$
$$\cdot \left\{\mathbb{E}\left[\nu\left((1-F)^{\bar{\mathbf{d}}\wedge t}\right)\right] - \mathbb{P}\left[\bar{\mathbf{d}} \geq t\right]\nu\left((1-F)^t\right)\right\} \frac{\mathbb{E}\left[\nu\left(1-(1-F)^{\mathbf{d}}\right)\right]}{1-\mathbb{E}\left[\nu\left((1-F)^{\mathbf{d}}\right)\right]} +$$
$$+ \mathbb{E}\left[\nu\left(1-(1-F)^{\bar{\mathbf{d}}\wedge t}\right)\right] - \mathbb{P}\left[\bar{\mathbf{d}} \geq t\right]\nu\left(1-(1-F)^t\right) \tag{3.5}$$

$\blacksquare$

Asymptotically this gives a more digestible result

Corollary:

$$\lim_{t\to\infty} \mathbb{E}\left[\sum_{j=1}^{\infty} \mathbf{e}_j^n \prod_{k=1}^{n} \left(\mathbb{E}\left[\mathbf{e}_j^k\right]\right)^{n_k}\right] =$$

$$= \frac{\mathbb{E}\left[\nu\left((1-F)^{\bar{\mathbf{d}}}\right)\right]\mathbb{E}\left[\nu\left(1-(1-F)^{\mathbf{d}}\right)\right]}{1-\mathbb{E}\left[\nu\left((1-F)^{\mathbf{d}}\right)\right]} + \mathbb{E}\left[\nu\left(1-(1-F)^{\bar{\mathbf{d}}}\right)\right] \tag{3.6}$$

$\blacksquare$

Corollary (3.6) has an intuitive interpretation. The numerator reflects a common factor within all of the $\mathbf{e}_j$ prior to the last regime. The first factor in the numerator reflects a part of that common factor stemming from an overhang into the last regime. The denominator comes from summing an infinite geometric series that reflects factors within the $\mathbf{e}_j$ that vary with j. The final term picks up an effect that persists into the limit from the truncation of $\mathbf{e}_j$ in the last regime when $t < \infty$.

All the additional complexity in theorem (3.5) captures the effects of stopping the process at t: (a) on almost all paths the geometric series has only a finite number of terms and (b) there is a non-zero residual likelihood that no regime change occurs.

3.3 Computation

Every element of theorem (3.5) or corollary (3.6) can be computed, as well as every element of expression (4.5.8) below for the $\rho_{a_1,\ldots,a_k}$ of lemma (2.9) needed in (2.10)-(2.11), if we know the probability distribution function for $\bar{\mathbf{d}}$ and the Laplace transforms $\mathcal{L}_{\mathbf{d}}$, $\mathcal{L}_{\bar{\mathbf{d}}}$ and $\mathcal{L}_{\bar{\mathbf{d}}\wedge t}$ for the densities of $\mathbf{d}$, $\bar{\mathbf{d}}$ and $\bar{\mathbf{d}}\wedge t$. All expressions of the form $\mathbb{E}\left[x^{\mathbf{v}}\right]$ for $\mathbf{v}$ one of those random variables can be evaluated as $\mathcal{L}_{\mathbf{v}}\left[-\ln\left(x\right)\right]$. The constant $1-G$ is defined in terms of $\mathcal{L}_{\mathbf{d}}^{-1}$ (see (3.1)).

In practical applications we are likely to know the distribution function for $\bar{\mathbf{d}}$ and the Laplace transforms $\mathcal{L}_{\mathbf{d}}$, $\mathcal{L}_{\bar{\mathbf{d}}}$ and $\mathcal{L}_{\bar{\mathbf{d}}\wedge t}$. For example, with our *gamma* (α,β) assumption for $\mathbf{d}$, the Laplace transform is well-known or can be integrated directly

$$\begin{aligned}
\mathcal{L}_{\mathbf{d}}\left(x\right) &= \frac{1}{\Gamma\left(\alpha\right)\beta^{\alpha}}\int_0^{\infty} e^{-xt}t^{\alpha-1}e^{-\frac{t}{\beta}}\,dt \\
&= \frac{1}{\Gamma\left(\alpha\right)\beta^{\alpha}}\Gamma\left(\alpha\right)\left(\frac{\beta}{\beta x+1}\right)^{\alpha} \\
&= \left(1+\beta x\right)^{-\alpha} \\
\mathcal{L}_{\mathbf{d}}^{-1}\left(y\right) &= \frac{1}{\beta}\left(y^{-\frac{1}{\alpha}}-1\right).
\end{aligned}$$

The random variable $\bar{\mathbf{d}}$ follows the equilibrium distribution corresponding to $\mathbf{d}$, i.e. $\bar{\mathbf{d}}$ has non-negative support and $\mathbf{d} \simeq \bar{\mathbf{d}} + \bar{\mathbf{d}}^{'}$, where $\bar{\mathbf{d}} \simeq \bar{\mathbf{d}}^{'}$, but obviously not independent.

Lemma: The probability density function of $\bar{\mathbf{d}}$ is given by

$$f_{\bar{\mathbf{d}}}\left(x\right) = \frac{\mathbb{P}\left[\mathbf{d} \geq x\right]}{\mathbb{E}\left[\mathbf{d}\right]} \tag{3.7}$$

Proof: This is a well known property of the equilibrium distribution proven by conditioning on the value of $\bar{\mathbf{d}}^{'} \simeq \bar{\mathbf{d}}$.

$$\begin{aligned}
f_{\bar{\mathbf{d}}}\left(x\right) &= \int_0^{\infty} f_{\bar{\mathbf{d}}'}\left(y\right) f_{\mathbf{d}}\left(x+y|\mathbf{d} \geq y\right)dy \\
&= \int_0^{\infty} f_{\bar{\mathbf{d}}}\left(y\right)\frac{f_{\mathbf{d}}\left(x+y\right)}{\mathbb{P}\left[\mathbf{d} \geq y\right]}dy, \text{ an integral equation for } f_{\bar{\mathbf{d}}}\left(x\right).
\end{aligned}$$

Since $\int_0^\infty \mathbb{P}[\mathbf{d} \geq x]\, dx = \mathbb{E}[\mathbf{d}]$ (integrate by parts) guess that the solution for the integral equation is

$$f_{\bar{\mathbf{d}}}(x) = \frac{\mathbb{P}[\mathbf{d} \geq x]}{\mathbb{E}[\mathbf{d}]}, \text{ which works.} \qquad \blacksquare$$

In our case,

$$\mathbb{P}\left[\bar{\mathbf{d}} \geq t\right] = \int_t^\infty \frac{\mathbb{P}[\mathbf{d} \geq s]}{\mathbb{E}[\mathbf{d}]} ds = 1 - \Gamma\left(\alpha + 1; \frac{t}{\beta}\right) - \frac{t}{\alpha\beta}\left[1 - \Gamma\left(\alpha; \frac{t}{\beta}\right)\right]$$

after an integration by parts, using incomplete gamma functions. $\mathcal{L}_{\bar{\mathbf{d}}\wedge t}$ is given by

$$\mathcal{L}_{\bar{\mathbf{d}}\wedge t}(x) = \frac{1}{\mathbb{E}[\mathbf{d}]}\int_0^t e^{-xs}\mathbb{P}[\mathbf{d} \geq s]\, ds + e^{-xt}\mathbb{P}\left[\bar{\mathbf{d}} \geq t\right]$$

$$\begin{aligned} = \quad & \frac{1}{\alpha\beta x}\left\{1 - e^{-xt}\left[1 - \Gamma\left(\alpha; \frac{t}{\beta}\right)\right] - (1 + \beta x)^{-\alpha}\,\Gamma\left(\alpha; \frac{(1 + \beta x)\,t}{\beta}\right)\right\} \\ & + e^{-xt}\left\{1 - \Gamma\left(\alpha + 1; \frac{t}{\beta}\right) - \frac{t}{\alpha\beta}\left[1 - \Gamma\left(\alpha; \frac{t}{\beta}\right)\right]\right\} \end{aligned}$$

after an integration by parts. Letting $t \to \infty$ gives

$$\mathcal{L}_{\bar{\mathbf{d}}}(x) = \frac{1}{\alpha\beta x}\left[1 - (1 + \beta x)^{-\alpha}\right]$$

These formulae allow us to calculate all of the values in theorem (3.5), corollary (3.6), and expression (4.5.8) so we can calculate (2.12) and all $\rho_{a_1,\ldots,a_k}$ and with them all of the values in section 2 leading to the calculation of approximate values (2.2) to (2.4) for any moments of the random variable r_t.

4 Proofs of the Main Results

Theorem (3.3) is the key. We show in section 4.2 how assuming theorem (3.3) opens the rest of the doors. Then in section 4.4 we prove theorem (3.3) itself. Sections 4.1 and 4.3 provide some some required preliminary results about the interarrival structure. Finally, section 4.5 proves lemma (2.9).

4.1 Lemmata for Corollary (3.4)

Lemma:

$$\mathbf{t}_1 \simeq \bar{\mathbf{d}} \wedge t \text{ where } " \simeq " \text{ means "equal in law"} \tag{4.1.1}$$

Proof: By definition (see section 3.1)

$$\begin{aligned} \mathbf{t}_1 &= \bar{\mathbf{d}}_1 = \mathbf{d}_1 \wedge (\bar{\mathbf{d}}_0 + t) - \bar{\mathbf{d}}_0 \\ \text{so } \mathbf{t}_1 &= (\mathbf{d}_1 - \bar{\mathbf{d}}_0) \wedge t \\ \text{but by definition } \mathbf{d}_1 - \bar{\mathbf{d}}_0 &\simeq \bar{\mathbf{d}} \qquad \blacksquare \end{aligned}$$

Lemma:

$$\mathbf{J} = 1 \Leftrightarrow \mathbf{t}_1 = t \tag{4.1.2}$$

Proof: By definition

$$\begin{aligned} \mathbf{J} &= 1 \Leftrightarrow \mathbf{t}_1 \geq t \\ \text{and (4.1.1) implies } \mathbf{t}_1 &\leq t \\ \text{so } \mathbf{J} &= 1 \Leftrightarrow \mathbf{t}_1 = t \qquad \blacksquare \end{aligned}$$

Lemma:

$$\mathbb{P}[\mathbf{J} = 1] = \mathbb{P}\left[\bar{\mathbf{d}} \geq t\right] \tag{4.1.3}$$

Proof: By (4.1.2.)

$$\begin{aligned} \mathbb{P}[\mathbf{J} = 1] &= \mathbb{P}[\mathbf{t}_1 = t] \\ &= \mathbb{P}\left[\bar{\mathbf{d}} \wedge t = t\right] \text{ by (4.1.1.)} \\ &= \mathbb{P}\left[\bar{\mathbf{d}} \geq t\right] \qquad \blacksquare \end{aligned}$$

Lemma: For any constant G:

$$\mathbb{E}\left[\sum_{j=1}^{\infty}\left\{(1-G)^{(t-\mathbf{t}_{j+1})_+} - (1-G)^{(t-\mathbf{t}_j)_+}\right\}\right] =$$

$$= 1 - \mathbb{P}\left[\bar{\mathbf{d}} \geq t\right] - (1-G)^t\left\{\mathbb{E}\left[(1-G)^{-\bar{\mathbf{d}}\wedge t}\right] - \mathbb{P}\left[\bar{\mathbf{d}} \geq t\right](1-G)^{-t}\right\} \tag{4.1.4}$$

Proof: On almost all paths, $\mathbf{t}_j \geq t$ for some $\mathbf{t}_j$ and all thereafter causing the sum to telescope. By monotone convergence the expectation also telescopes:

$$\mathbb{E}\left[\sum_{j=1}^{\infty}\left\{(1-G)^{(t-\mathbf{t}_{j+1})_+}-(1-G)^{(t-\mathbf{t}_j)_+}\right\}\right]=$$

$$\begin{aligned}
&= \mathbb{E}\left[1-(1-G)^{(t-\mathbf{t}_1)_+}\right]\\
&= 1-\mathbb{P}\left[\mathbf{J}=1\right]-\mathbb{P}\left[\mathbf{J}>1\right]\mathbb{E}\left[(1-G)^{t-\mathbf{t}_1}\,|\mathbf{J}>1\right] \text{ conditioning on } \mathbf{J}\\
&\qquad \text{and using (4.1.2)}\\
&= 1-\mathbb{P}\left[\mathbf{J}=1\right]-(1-G)^t\,\mathbb{P}\left[\mathbf{J}>1\right]\mathbb{E}\left[(1-G)^{-\mathbf{t}_1}\,|\mathbf{J}>1\right]\\
&= 1-\mathbb{P}\left[\mathbf{J}=1\right]-(1-G)^t\left\{\mathbb{E}\left[(1-G)^{-\mathbf{t}_1}\right]-\mathbb{P}\left[\mathbf{J}=1\right]\mathbb{E}\left[(1-G)^{-\mathbf{t}_1}\,|\mathbf{J}=1\right]\right\}\\
&\qquad \text{conditioning on } \mathbf{J} \text{ again}\\
&= 1-\mathbb{P}\left[\bar{\mathbf{d}}\ \geq t\right]-(1-G)^t\left\{\mathbb{E}\left[(1-G)^{-\bar{\mathbf{d}}\wedge t}\right]-\mathbb{P}\left[\bar{\mathbf{d}}\ \geq t\right](1-G)^{-t}\right\}\\
&\qquad \text{by (4.1.3), (4.1.1), and (4.1.2)} \qquad \blacksquare
\end{aligned}$$

4.2 Assuming Theorem (3.3) to be true

For this section we assume theorem (3.3) to be true for all choices of F, n, and $\{n_k\}_{k\leq n}$. Note that for all such choices the random variables $\mathbf{d}$, $\bar{\mathbf{d}}$, $\{\mathbf{t}_j\}_{j\geq 1}$, and $\mathbf{J}$ remain the same. They characterize the interarrival structure which we assume to be set once and for all throughout all of this work.

Proof of corollary (3.4). Consider the case of theorem (3.3) when $n=1$, $n_1=0$ and F takes the value G given by (3.1) for some other general choice of F, n, and $\{n_k\}_{k\leq n}$. Using definition (2.7) on the left hand side and putting the expression for K right into the statement, theorem (3.3) reads:

$$\mathbb{E}\left[\sum_{j=1}^{\infty}\left\{(1-G)^{(t-\mathbf{t}_{j+1})_+}-(1-G)^{(t-\mathbf{t}_j)_+}\right\}\right]=$$

$$\begin{aligned}
= \quad & \left(1-\mathbb{E}\left[\left(\mathbb{E}\left[(1-G)^{\mathbf{d}}\right]\right)^{\mathbf{J}-2}\mid \mathbf{J}>1\right]\right)\cdot\\
&\cdot\left\{\mathbb{E}\left[(1-G)^{\bar{\mathbf{d}}\wedge t}\right]-\mathbb{P}\left[\bar{\mathbf{d}}\geq t\right](1-G)^t\right\}\frac{\mathbb{E}\left[1-(1-G)^{\mathbf{d}}\right]}{1-\mathbb{E}\left[(1-G)^{\mathbf{d}}\right]}+\\
&+\mathbb{E}\left[1-(1-G)^{\bar{\mathbf{d}}\wedge t}\right]-\mathbb{P}\left[\bar{\mathbf{d}}\geq t\right]\left(1-(1-G)^t\right)
\end{aligned}$$

$$\mathbb{E}\left[\sum_{j=1}^{\infty}\left\{(1-G)^{(t-\mathbf{t}_{j+1})_+}-(1-G)^{(t-\mathbf{t}_j)_+}\right\}\right]=$$

$$\begin{aligned}
= \quad & \left(1-\mathbb{E}\left[\left(\mathbb{E}\left[(1-G)^{\mathbf{d}}\right]\right)^{\mathbf{J}-2}\mid \mathbf{J}>1\right]\right)\cdot \\
& \cdot\left\{\mathbb{E}\left[(1-G)^{\bar{\mathbf{d}}\wedge t}\right]-\mathbb{P}\left[\bar{\mathbf{d}}\geq t\right](1-G)^t\right\} \\
& +1-\mathbb{E}\left[(1-G)^{\bar{\mathbf{d}}\wedge t}\right]-\mathbb{P}\left[\bar{\mathbf{d}}\geq t\right]+\mathbb{P}\left[\bar{\mathbf{d}}\geq t\right](1-G)^t
\end{aligned}$$

Substitute the expression from (3.2) into the right hand side:

$$\mathbb{E}\left[\sum_{j=1}^{\infty}\left\{(1-G)^{(t-\mathbf{t}_{j+1})_+}-(1-G)^{(t-\mathbf{t}_j)_+}\right\}\right]=$$

$$\begin{aligned}
= \quad & \left(1-\mathbb{E}\left[\left(\mathbb{E}\left[\nu\left((1-F)^{\mathbf{d}}\right)\right]\right)^{\mathbf{J}-2}\mid \mathbf{J}>1\right]\right)\cdot \\
& \cdot\left\{\mathbb{E}\left[(1-G)^{\bar{\mathbf{d}}\wedge t}\right]-\mathbb{P}\left[\bar{\mathbf{d}}\geq t\right](1-G)^t\right\} \\
& +1-\mathbb{E}\left[(1-G)^{\bar{\mathbf{d}}\wedge t}\right]-\mathbb{P}\left[\bar{\mathbf{d}}\geq t\right]+\mathbb{P}\left[\bar{\mathbf{d}}> t\right](1-G)^t
\end{aligned}$$

Rearrange and simplify to isolate

$$1-\mathbb{E}\left[\left(\mathbb{E}\left[\nu\left((1-F)^{\mathbf{d}}\right)\right]\right)^{\mathbf{J}-2}\mid \mathbf{J}>1\right]=$$

$$=1+\frac{\mathbb{E}\left[\sum_{j=1}^{\infty}\left\{(1-G)^{(t-\mathbf{t}_{j+1})_+}-(1-G)^{(t-\mathbf{t}_j)_+}\right\}\right]-\left(1-\mathbb{P}\left[\bar{\mathbf{d}}\geq t\right]\right)}{\mathbb{E}\left[(1-G)^{\bar{\mathbf{d}}\wedge t}\right]-\mathbb{P}\left[\bar{\mathbf{d}}>t\right](1-G)^t}$$

and substitute the expression from (4.1.4) into the numerator

$$=1-(1-G)^t\,\frac{\mathbb{E}\left[(1-G)^{-\bar{\mathbf{d}}\wedge t}\right]-\mathbb{P}\left[\bar{\mathbf{d}}\ \geq t\right](1-G)^{-t}}{\mathbb{E}\left[(1-G)^{\bar{\mathbf{d}}\wedge t}\right]-\mathbb{P}\left[\bar{\mathbf{d}}>t\right](1-G)^t}$$

proving corollary (3.4) ■

Proof of Theorem (3.5). Substitute the expression from corollary (3.4) into the statement of theorem (3.3). ■

Proof of Corollary (3.6). Take $\lim_{t\to\infty}$ in the statement of Theorem (3.5). Note, (3.2) and $0<(1-F)<1$ imply that $0<(1-G)<1$ so $\lim_{t\longrightarrow\infty}(1-G)^t=0$. ■

4.3 Lemmata for Theorem (3.3)

Lemma:

As joint distributions $\{\mathbf{J},\, \bar{\mathbf{d}}_{\mathbf{J}}\} \;\simeq\; \{\mathbf{J},\, \bar{\mathbf{d}}_1\}$ where " $\simeq$ " means "equal in law" (4.3.1)

Proof: Consider the i.i.d. inter-arrival intervals with common law $\mathbf{d}$

$$\mathbf{d}_1,\, \mathbf{d}_2,\, \dots\, ,\, \mathbf{d}_{\mathbf{J}}$$

defined in section 3.1 to envelop the interval $[0, t]$. Model it as arising from a larger sequence of i.i.d. inter-arrival intervals with common law $\mathbf{d}$

$$\mathbf{d}'_0,\, \mathbf{d}'_1,\, \dots\, \mathbf{d}'_k,\, \dots,\, \mathbf{d}'_{\mathbf{M}'},\, \mathbf{d}'_{\mathbf{M}'+1},\, \mathbf{d}'_{\mathbf{M}'+2},\, \dots\, ,\, \mathbf{d}'_{\mathbf{M}'+\mathbf{J}-1},\, \mathbf{d}'_{\mathbf{M}'+\mathbf{J}}$$

where $\mathbf{d}'_0$ starts at some large negative value $-T$ of the time parameter and $\mathbf{M}' = \max\left\{ m : \sum_{k=0}^{m} \mathbf{d}'_k < T \right\}$. Then $\mathbf{d}'_{\mathbf{M}'+1}$ is the interval containing 0 and $\mathbf{d}'_{\mathbf{M}'+\mathbf{J}}$ is the interval containing t where $\mathbf{J}$ is as defined in section 3.1. Ignoring the set of paths of measure zero on which $\sum_{k=0}^{\mathbf{M}'+1} \mathbf{d}'_k = T$,

$$\lim_{T\to\infty} \left\{ \mathbf{d}'_{m+1}, \mathbf{d}'_{m+2},\, \dots\, ,\, \mathbf{d}'_{m+j-1},\, \mathbf{d}'_{m+j} | m = \mathbf{M}',\, j = \mathbf{J} \right\} \simeq$$

$$\simeq \{\mathbf{d}_1,\, \mathbf{d}_2,\, \dots\, ,\, \mathbf{d}_{j-1},\, \mathbf{d}_j | j = \mathbf{J}\}$$

as joint distributions enveloping the interval $[0, t]$ because in section 3.1 we established 0 at a random point within $\mathbf{d}_1$ and in the limit on the left the definition of $\mathbf{M}'$ accomplishes the same thing within $\mathbf{d}'_{m+1}$.

Alternatively, model

$$\mathbf{d}_1,\, \mathbf{d}_2,\, \dots\, ,\, \mathbf{d}_{\mathbf{J}}$$

as arising from a different sequence of i.i.d. inter-arrival intervals with common law $\mathbf{d}$

$$\mathbf{d}''_{\mathbf{M}''-1}, \mathbf{d}''_{\mathbf{M}''-2},\, \dots,\, \mathbf{d}''_{\mathbf{M}''-\mathbf{J}+1},\, \mathbf{d}''_{\mathbf{M}''-\mathbf{J}},\, \mathbf{d}''_{\mathbf{M}''-\mathbf{J}-1},\, \dots\, , \mathbf{d}''_k,\, \dots, \mathbf{d}''_1,\, \mathbf{d}''_0$$

where $\mathbf{d}''_0$ ends at some large positive value $t + T$ of the time parameter and $\mathbf{M}'' = \min\left\{ m : \sum_{k=0}^{m-1} \mathbf{d}''_k > t + T \right\}$. Then $\mathbf{d}''_{\mathbf{M}''-1}$ is the interval containing 0 and $\mathbf{d}''_{\mathbf{M}''-\mathbf{J}}$ is the interval containing t. Ignoring the set of paths of measure zero on which $\sum_{k=0}^{\mathbf{M}''-1} \mathbf{d}''_k = t + T$,

$$\lim_{T\to\infty}\left\{\mathbf{d}''_{m-1},\mathbf{d}''_{m-2},\,...\,,\,\mathbf{d}''_{m-j+1},\,\mathbf{d}''_{m-j}|m=\mathbf{M}'',\,j=\mathbf{J}\,\right\}\simeq$$

$$\simeq\{\mathbf{d}_1,\,\mathbf{d}_2,\,...\,,\,\mathbf{d}_{j-1},\,\mathbf{d}_j|j=\mathbf{J}\}$$

as joint distributions enveloping the interval $[0,t]$ because in section 3.1 we established 0 at a random point within $\mathbf{d}_1$ and in the limit on the left the definition of M'' accomplishes the same thing within $\mathbf{d}''_{m-1}$.

But conditional on $j=\mathbf{J}$ we have $\mathbf{M}''-j-1=\max\left\{m:\sum_{k=0}^{m}\mathbf{d}''_k<T\right\}$ so running backwards and interchanging the roles of 0 and t there is a symmetry with the first case that can be expressed as

$$\{\mathbf{d}_j,\,\mathbf{d}_{j-1},\,...\,,\,\mathbf{d}_2,\,\mathbf{d}_1|j=\mathbf{J}\}\simeq$$

$$\begin{aligned}
&\simeq\lim_{T\to\infty}\left\{\mathbf{d}''_{m-j},\mathbf{d}''_{m-j+1},\,...\,,\,\mathbf{d}''_{m-2},\,\mathbf{d}''_{m-1}|m=\mathbf{M}'',\,j=\mathbf{J}\,\right\}\\
&\simeq\lim_{T\to\infty}\left\{\mathbf{d}'_{m+1},\mathbf{d}'_{m+2},\,...\,,\,\mathbf{d}'_{m+j-1},\,\mathbf{d}'_{m+j}|m=\mathbf{M}',\,j=\mathbf{J}\,\right\}\\
&\simeq\{\mathbf{d}_1,\,\mathbf{d}_2,\,...\,,\,\mathbf{d}_{j-1},\,\mathbf{d}_j|j=\mathbf{J}\}
\end{aligned}$$

as joint distributions enveloping the interval $[0,t]$.

Furthermore, under this symmetry $\bar{\mathbf{d}}_{\mathbf{J}+1}$ on the left (as defined in section 3.1) corresponds on the right to $\bar{\mathbf{d}}_0$ (as defined in section 3.1). Therefore, using the definitions of $\bar{\mathbf{d}}_\mathbf{J}$ and $\bar{\mathbf{d}}_\mathbf{1}$ from section 3.1

$$\left\{\bar{\mathbf{d}}_\mathbf{J},\,\mathbf{J}\right\}=\left\{\mathbf{d}_\mathbf{J}-\bar{\mathbf{d}}_{\mathbf{J}+1},\,\mathbf{J}\right\}\simeq\left\{\mathbf{d}_\mathbf{1}-\bar{\mathbf{d}}_\mathbf{0},\,\mathbf{J}\right\}=\left\{\,\bar{\mathbf{d}}_\mathbf{1},\,\mathbf{J}\right\}\quad\blacksquare$$

Lemma: The common law of $\bar{\mathbf{d}}_\mathbf{J}$ and $\bar{\mathbf{d}}_\mathbf{1}$ is

$$\bar{\mathbf{d}}_\mathbf{J}\simeq\bar{\mathbf{d}}_\mathbf{1}\simeq\bar{\mathbf{d}}\wedge t \tag{4.3.2}$$

Proof: directly from (4.3.1) and (4.1.1) since $\mathbf{t}_1=\bar{\mathbf{d}}_\mathbf{1}$ by definition in section 3.1 ■

It is important to note that while $\bar{\mathbf{d}}_\mathbf{1}$ and $\bar{\mathbf{d}}_\mathbf{J}$ have a common law, they are not independent. The relation

$$t=\bar{\mathbf{d}}_1+\mathbf{d}_2+...+\mathbf{d}_{\mathbf{J}-1}+\bar{\mathbf{d}}_\mathbf{J},$$

not to mention the possibility that $\mathbf{J}=1$, entangles their distributions and leads to most of the complexity in theorem (3.3) and theorem (3.5). The following lemmata, however, provide enough independence to prove theorem (3.3) and resolve (12.13).

Lemma:

$$\mathbf{J}=1\Leftrightarrow\bar{\mathbf{d}}_\mathbf{J}=\bar{\mathbf{d}}_\mathbf{1}=t \tag{4.3.3}$$

Proof: By (4.1.2), (4.3.1) and the definition $\mathbf{t}_1=\bar{\mathbf{d}}_\mathbf{1}$ ■

Lemma: Conditional on $\mathbf{J} = j' > 1$ the following are each sets of independent random variables

$$\left\{\mathbf{J},\bar{\mathbf{d}}_1\right\}, \ \left\{\bar{\mathbf{d}}_1, \mathbf{d}_2, ..., \mathbf{d}_{j'-1}\right\}, \ \left\{\mathbf{d}_2, ..., \mathbf{d}_{j'-1}, \bar{\mathbf{d}}_{j'}\right\} \text{ and } \left\{\mathbf{J},\bar{\mathbf{d}}_{\mathbf{J}}\right\} \tag{4.3.4}$$

Proof: Independence of the first three sets, conditional on $\mathbf{J} = j' > 1$, follows directly from the independence of $\{\mathbf{d}_1, \mathbf{d}_2, \dots, \mathbf{d}_{j'}\}$ and the definitions of $\bar{\mathbf{d}}_1$, $\mathbf{J}$ and $\bar{\mathbf{d}}_{\mathbf{J}}$ in section 3.1. Independence of $\left\{\mathbf{J}, \bar{\mathbf{d}}_{\mathbf{J}}\right\}$, conditional on $\mathbf{J} > 1$, follows from (4.3.1) and the independence of $\left\{\mathbf{J}, \bar{\mathbf{d}}_1\right\}$ conditional on $\mathbf{J} > 1$. ■

Lemma: Conditional on $\mathbf{J} = j' > j \geq 1$ the indicator random variables $\{\mathbf{1}_{j<\mathbf{J}}\}_{j\geq 1}$ defined in section 3.1 satisfy

$$\mathbf{1}_{j<\mathbf{J}} \ \text{ is independent of } \{\mathbf{d}_{j+1}, ..., \mathbf{d}_{j'-1}\}. \tag{4.3.5}$$

Proof: Directly from (4.3.4) upon noting that by definition $\mathbf{1}_{j<\mathbf{J}}$ is determined completely by

$$\mathbf{t}_j = \bar{\mathbf{d}}_1 + \mathbf{d}_2 + ... + \mathbf{d}_j \qquad \blacksquare$$

Definition: With $\mathbf{e}_j$ as defined at (2.7) define

$$\bar{\mathbf{e}}_j = (1-F)^{-\bar{\mathbf{d}}_{\mathbf{J}}}\, \mathbf{e}_j \tag{4.3.6}$$

Lemma: $\bar{\mathbf{e}}_j$ can be expressed as follows:

$$\begin{aligned} \bar{\mathbf{e}}_j &= (1-F)^{\mathbf{d}_{j+2}+...+\mathbf{d}_{\mathbf{J}-1}} \left(1-(1-F)^{\mathbf{d}_{j+1}}\right) \text{ for } j < \mathbf{J}-1, \\ \bar{\mathbf{e}}_{\mathbf{J}-1} &= (1-F)^{-\bar{\mathbf{d}}_{\mathbf{J}}} - 1, \text{ and} \\ \bar{\mathbf{e}}_j &= 0 \text{ for } j \geq \mathbf{J} \end{aligned} \tag{4.3.7}$$

Proof: Directly from the definition of $\mathbf{J}$ in section 3.1, definitions (2.7) and (4.3.6), and the relation $t - \mathbf{t}_j = \mathbf{d}_{j+1} + ... + \mathbf{d}_{\mathbf{J}-1} + \bar{\mathbf{d}}_{\mathbf{J}}$ ■

4.4 Proof of Theorem (3.3)

The proof is inspired by the proof of Wald's equations in [2].

When $\mathbf{J} = \mathbf{1}$, the exponents $(t - \mathbf{t}_j)_+$ and $(t - \mathbf{t}_{j+1})_+$ appearing in the definition of $\mathbf{e}_j$ at (2.7) are 0 for all $j \geq 1$ making each term of $\mathbb{E}\left[\sum_{j=1}^{\infty} \nu(\mathbf{e}_j)\right]$

vanish, so

$$
\begin{aligned}
\mathbb{E}\left[\sum_{j=1}^{\infty}\nu\left(\mathbf{e}_j\right)\right] &= \mathbb{P}\left[\mathbf{J}>1\right]\mathbb{E}\left[\sum_{j=1}^{\infty}\nu\left(\mathbf{e}_j\right)|\mathbf{J}>1\right] \\
&= \mathbb{P}\left[\mathbf{J}>1\right]\mathbb{E}\left[\sum_{j=1,\neq\mathbf{J}-1}^{\infty}\nu\left((1-F)^{\bar{\mathbf{d}}_{\mathbf{J}}}\right)\nu\left(\bar{\mathbf{e}}_j\right)|\mathbf{J}>1\right]+ \\
&\quad +\mathbb{P}\left[\mathbf{J}>1\right]\mathbb{E}\left[\nu\left(\mathbf{e}_{\mathbf{J}-1}\right)|\mathbf{J}>1\right], \text{ using (4.3.6),(4.3.7)\&(2.7).} \\
&= A+B \qquad (4.4.1)
\end{aligned}
$$

Work on A first. Use (4.3.7) to justify the introduction of indicator random variables:

$$
\begin{aligned}
A &= \mathbb{P}\left[\mathbf{J}>1\right]\mathbb{E}\left[\sum_{j=1,\neq\mathbf{J}-1}^{\infty}\nu\left((1-F)^{\bar{\mathbf{d}}_{\mathbf{J}}}\right)\nu\left(\bar{\mathbf{e}}_j\right)|\mathbf{J}>1\right] \\
&= \mathbb{P}\left[\mathbf{J}>1\right]\mathbb{E}\left[\sum_{j=1,\neq\mathbf{J}-1}^{\infty}\nu\left((1-F)^{\bar{\mathbf{d}}_{\mathbf{J}}}\right)\mathbf{1}_{j<\mathbf{J}}\nu\left(\bar{\mathbf{e}}_j\right)|\mathbf{J}>1\right], \text{ by (4.3.7).}
\end{aligned}
$$

$$
\begin{aligned}
&= \mathbb{P}\left[\mathbf{J}>1\right]\sum_{j=1,\neq\mathbf{J}-1}^{\infty}\mathbb{E}\left[\nu\left((1-F)^{\bar{\mathbf{d}}_{\mathbf{J}}}\right)\mathbf{1}_{j<\mathbf{J}}|\mathbf{J}>1\right]\mathbb{E}\left[\nu\left(\bar{\mathbf{e}}_j\right)|\mathbf{J}>1\right], \text{ by} \\
&\quad \text{monotone convergence, with (4.3.7), (4.3.4) and (4.3.5) for independence.} \\
&= \mathbb{P}\left[\mathbf{J}>1\right]\mathbb{E}\left[\left\{\sum_{j=1,\neq\mathbf{J}-1}^{\infty}\nu\left((1-F)^{\bar{\mathbf{d}}_{\mathbf{J}}}\right)\mathbf{1}_{j<\mathbf{J}}\mathbb{E}\left[\nu\left(\bar{\mathbf{e}}_j\right)|\mathbf{J}>1\right]\right\}|\mathbf{J}>1\right], \text{ by} \\
&\quad \text{monotone convergence.} \\
&= \mathbb{P}\left[\mathbf{J}>1\right]\mathbb{E}\left[\left\{\nu\left((1-F)^{\bar{\mathbf{d}}_{\mathbf{J}}}\right)\sum_{j=1,\neq\mathbf{J}-1}^{\infty}\mathbf{1}_{j<\mathbf{J}}\mathbb{E}\left[\nu\left(\bar{\mathbf{e}}_j\right)|\mathbf{J}>1\right]\right\}|\mathbf{J}>1\right] \\
&= \mathbb{P}\left[\mathbf{J}>1\right]\mathbb{E}\left[\left\{\nu\left((1-F)^{\bar{\mathbf{d}}_{\mathbf{J}}}\right)\sum_{j=1}^{\mathbf{J}-2}\mathbb{E}\left[\nu\left(\bar{\mathbf{e}}_j\right)|\mathbf{J}>1\right]\right\}|\mathbf{J}>1\right] \text{ because the} \\
&\quad \text{indicators kill all terms beyond } \mathbf{J}-1 \text{ and the } \mathbf{J}-1 \text{ term was} \\
&\quad \text{removed earlier.}
\end{aligned}
$$

$$
\begin{aligned}
A &= \mathbb{P}\left[\mathbf{J}>1\right]\mathbb{E}\left[\left\{\sum_{j=1}^{\mathbf{J}-2}\mathbb{E}\left[\nu\left(\bar{\mathbf{e}}_j\right)|\mathbf{J}>\mathbf{1}\right]\right\}|\mathbf{J}>1\right]\mathbb{E}\left[\nu\left((1-F)^{\bar{\mathbf{d}}_{\mathbf{J}}}\right)|\mathbf{J}>1\right], \\
&\quad \text{by (4.3.4) for independence}
\end{aligned}
\qquad (4.4.2)
$$

Now use (4.3.7) to conclude that for $j \le \mathbf{J}-2$

$$\begin{aligned}
\mathbb{E}\left[\nu\left(\bar{\mathbf{e}}_j\right)|\mathbf{J}>1\right] &= \mathbb{E}\left[\nu\left((1-F)^{\mathbf{d}_{j+2}+\ldots+\mathbf{d}_{\mathbf{J}-1}}\left(1-(1-F)^{\mathbf{d}_{j+1}}\right)\right)|\mathbf{J}>1\right] \\
&= \mathbb{E}\left[\nu\left((1-F)^{\mathbf{d}_{j+2}+\ldots+\mathbf{d}_{\mathbf{J}-1}}\right)|\mathbf{J}>1\right]\cdot \\
&\quad \cdot\mathbb{E}\left[\nu\left(1-(1-F)^{\mathbf{d}_{j+1}}\right)|\mathbf{J}>1\right] \\
&= \left(\mathbb{E}\left[\nu\left((1-F)^{\mathbf{d}}\right)\right]\right)^{\mathbf{J}-2-j}\mathbb{E}\left[\nu\left(1-(1-F)^{\mathbf{d}}\right)\right], \text{ by}
\end{aligned}$$

the section 3.1 definition of $\{\mathbf{d}_j\}$ to be i.i.d. with law $\mathbf{d}$ (and with no dependence on $\mathbf{J}$.)

Substitute this expression into (4.4.2) and pull the $\mathbb{E}\left[\nu\left(1-(1-F)^{\mathbf{d}}\right)\right]$ factor out of the sum and the expectation giving

$$\begin{aligned}
A &= \mathbb{P}[\mathbf{J}>1]\,\mathbb{E}\left[\left\{\sum_{j=1}^{\mathbf{J}-2}\left(\mathbb{E}\left[\nu\left((1-F)^{\mathbf{d}}\right)\right]\right)^{\mathbf{J}-2-j}\right\}|\mathbf{J}>1\right]\cdot \\
&\quad \cdot\mathbb{E}\left[\nu\left((1-F)^{\bar{\mathbf{d}}_{\mathbf{J}}}\right)|\mathbf{J}>1\right]\mathbb{E}\left[\nu\left(1-(1-F)^{\mathbf{d}}\right)\right]
\end{aligned}$$

Sum the geometric series

$$\sum_{j=1}^{\mathbf{J}-2}\left(\mathbb{E}\left[\nu\left((1-F)^{\mathbf{d}}\right)\right]\right)^{\mathbf{J}-2-j} = \frac{1-\left(\mathbb{E}\left[\nu\left((1-F)^{\mathbf{d}}\right)\right]\right)^{\mathbf{J}-2}}{1-\mathbb{E}\left[\nu\left((1-F)^{\mathbf{d}}\right)\right]}$$

and pull the $1-\mathbb{E}\left[\nu\left((1-F)^{\mathbf{d}}\right)\right]$ factor out of the expectation giving

$$\begin{aligned}
A &= \mathbb{P}[\mathbf{J}>1]\,\mathbb{E}\left[\left(1-\left(\mathbb{E}\left[\nu\left((1-F)^{\mathbf{d}}\right)\right]\right)^{\mathbf{J}-2}\right)|\mathbf{J}>1\right]\cdot \\
&\quad \cdot\mathbb{E}\left[\nu\left((1-F)^{\bar{\mathbf{d}}_{\mathbf{J}}}\right)|\mathbf{J}>1\right]\frac{\mathbb{E}\left[\nu\left(1-(1-F)^{\mathbf{d}}\right)\right]}{1-\mathbb{E}\left[\nu\left((1-F)^{\mathbf{d}}\right)\right]} \\
&= K\mathbb{P}[\mathbf{J}>1]\,\mathbb{E}\left[\nu\left((1-F)^{\bar{\mathbf{d}}_{\mathbf{J}}}\right)|\mathbf{J}>1\right]\frac{\mathbb{E}\left[\nu\left(1-(1-F)^{\mathbf{d}}\right)\right]}{1-\mathbb{E}\left[\nu\left((1-F)^{\mathbf{d}}\right)\right]}
\end{aligned} \tag{4.4.3}$$

where

$$K = 1-\mathbb{E}\left[\left(\mathbb{E}\left[\nu\left((1-F)^{\mathbf{d}}\right)\right]\right)^{\mathbf{J}-2} \mid \mathbf{J}>1\right]$$

as in the statement of theorem (3.3).

But conditioning on $\mathbf{J} =$ or > 1,

$$\mathbb{P}[\mathbf{J} > 1]\,\mathbb{E}\left[\nu\left((1-F)^{\bar{\mathbf{d}}_{\mathbf{J}}}\right)|\mathbf{J} > 1\right] =$$

$$\begin{aligned}
&= \mathbb{E}\left[\nu\left((1-F)^{\bar{\mathbf{d}}_{\mathbf{J}}}\right)\right] - \mathbb{P}[\mathbf{J} = 1]\,\mathbb{E}\left[\nu\left((1-F)^{\bar{\mathbf{d}}_{\mathbf{J}}}\right)|\mathbf{J} = 1\right] \\
&= \mathbb{E}\left[\nu\left((1-F)^{\bar{\mathbf{d}}\wedge t}\right)\right] - \mathbb{P}\left[\bar{\mathbf{d}} \geq t\right]\nu\left((1-F)^{t}\right)
\end{aligned}$$

, by (4.3.2) for the first expectation, (4.1.3) for the probability, and (4.3.3) for the second expectation.

Substitute this expression into (4.4.3) to get

$$A = K\left\{\mathbb{E}\left[\nu\left((1-F)^{\bar{\mathbf{d}}\wedge t}\right)\right] - \mathbb{P}\left[\bar{\mathbf{d}} \geq t\right]\nu\left((1-F)^{\mathbf{t}}\right)\right\}\frac{\mathbb{E}\left[\nu\left(1-(1-F)^{\mathbf{d}}\right)\right]}{1-\mathbb{E}\left[\nu\left((1-F)^{\mathbf{d}}\right)\right]}$$

for the first term in (4.4.1). A similar conditioning on $\mathbf{J} =$ or > 1 eliminates $\mathbf{J}$ from the second term B making up (4.4.1):

$$\begin{aligned}
B &= \mathbb{P}[\mathbf{J} > 1]\,\mathbb{E}[\nu(\mathbf{e}_{\mathbf{J}-1})\,|\mathbf{J} > 1] \\
&= \mathbb{P}[\mathbf{J} > 1]\,\mathbb{E}\left[\nu\left(1-(1-F)^{\bar{\mathbf{d}}_{\mathbf{J}}}\right)|\mathbf{J} > 1\right],
\end{aligned}$$

by (2.7), the definition of $\mathbf{J}$ in section 3.1, and the relation $t - \mathbf{t}_{\mathbf{J}-1} = \bar{\mathbf{d}}_{\mathbf{J}}$. So

$$\begin{aligned}
B &= \mathbb{E}\left[\nu\left(1-(1-F)^{\bar{\mathbf{d}}_{\mathbf{J}}}\right)\right] - \mathbb{P}[\mathbf{J} = 1]\,\mathbb{E}\left[\nu\left(1-(1-F)^{\bar{\mathbf{d}}_{\mathbf{J}}}\right)|\mathbf{J} = 1\right] \\
&= \mathbb{E}\left[\nu\left(1-(1-F)^{\bar{\mathbf{d}}\wedge t}\right)\right] - \mathbb{P}\left[\bar{\mathbf{d}} \geq t\right]\nu\left(1-(1-F)^{t}\right)
\end{aligned}$$

, for the same reasons as in the expression for A.

Using (4.4.1) the expressions for A, B and K that we have derived establish theorem (3.3) ■

4.5 Proof of Lemma (2.9)

Lemma:

$$\mathbb{E}\left[\mathbf{e}_{j_1}^{2a_1}\cdots\mathbf{e}_{j_k}^{2a_k}\right] = \rho_{a_1,\ldots,a_k}\mathbb{E}\left[\mathbf{e}_{j_1}^{2a_1}\right]\cdots\mathbb{E}\left[\mathbf{e}_{j_k}^{2a_k}\right] \text{ when no two of } \{j_1,\ldots,j_k\} \text{ are equal, independent of } \{j_1,\ldots,j_k\}. \quad ■ \tag{2.9}$$

Proof: This proof is more easy to grasp now, with the material in sections 4.1, 4.2, 4.3 and 4.4 in hand, than it would have been when the lemma was stated. If $\mathbf{J} = \mathbf{1}$ then lemma (4.1.2) and the definition (2.7) of $\mathbf{e}_j$ make the lemma trivially true so we can assume $\mathbf{J} > \mathbf{1}$. Assume $j_1 = \max\{j_1,\ldots,j_k\}$.

Then applying lemma (4.3.7) and definition (4.3.6) to $\mathbf{e}_{j_1}^{2a_1}$, but definition (2.7) and the definition of $\mathbf{t}_k$ in section (3.1) to $\mathbf{e}_{j_i}^{2a_i}$ for all $i > 1$,

$$\mathbb{E}\left[\mathbf{e}_{j_1}^{2a_1} \cdots \mathbf{e}_{j_k}^{2a_k} | J > 1\right] =$$

$$\begin{aligned} = \quad & \mathbb{E}\left[(1-F)^{2a_1 \bar{\mathbf{d}}_{\mathbf{J}}} (1-F)^{2a_1\left(\mathbf{d}_{j_1+2}+\ldots+\mathbf{d}_{\mathbf{J}-1}\right)} \left(1-(1-F)^{\mathbf{d}_{j_1+1}}\right)^{2a_1} \cdot \right. \\ & \left. \cdot \prod_{i=2}^{k} \left\{ \left((1-F)^{\mathbf{d}_{j_i+1}} - 1\right)^{2a_i} (1-F)^{2a_i\left(\mathbf{d}_2+\ldots+\mathbf{d}_{j_i}\right)} (1-F)^{2a_i \bar{\mathbf{d}}_1} \right\} |\mathbf{J} > \mathbf{1} \right], \end{aligned}$$

which by lemma (4.3.4)

$$\begin{aligned} = \quad & \mathbb{E}\left[(1-F)^{2a_1 \bar{\mathbf{d}}_{\mathbf{J}}} \prod_{i=2}^{k} (1-F)^{2a_i \bar{\mathbf{d}}_1} |\mathbf{J} > \mathbf{1}\right] \cdot \\ & \cdot \mathbb{E}\left[(1-F)^{2a_1\left(\mathbf{d}_{j_1+2}+\ldots+\mathbf{d}_{\mathbf{J}-1}\right)} \left(1-(1-F)^{\mathbf{d}_{j_1+1}}\right)^{2a_1} |\mathbf{J} > \mathbf{1}\right] \cdot \\ & \cdot \mathbb{E}\left[\prod_{i=2}^{k} \left((1-F)^{\mathbf{d}_{j_i+1}} - 1\right)^{2a_i} (1-F)^{2a_i\left(\mathbf{d}_2+\ldots+\mathbf{d}_{j_i}\right)} |\mathbf{J} > \mathbf{1}\right]. \end{aligned} \tag{4.5.1}$$

So if we recursively define

$$\begin{aligned} \rho_{a_1,\ldots,a_k} &= \rho_{a_1,a_2+\ldots+a_k} \rho_{a_2,\ldots,a_k} \text{ where the recursion starts with} \\ \rho_{a,b} &= \frac{\mathbb{E}\left[(1-F)^{2a\bar{\mathbf{d}}_{\mathbf{J}}} (1-F)^{2b\bar{\mathbf{d}}_1} |\mathbf{J} > \mathbf{1}\right]}{\mathbb{E}\left[(1-F)^{2a\bar{\mathbf{d}}_{\mathbf{J}}} |\mathbf{J} > \mathbf{1}\right] \mathbb{E}\left[(1-F)^{2b\bar{\mathbf{d}}_1} |\mathbf{J} > \mathbf{1}\right]} \text{ for any } a, b. \end{aligned} \tag{4.5.2}$$

then successively for a_1, a_2, ... , a_k we can just reverse the process from (4.5.1) back to the statement of the lemma. ■

But we want a computable version of $\rho_{a_1,\ldots,a_k}$, so there's more work to do. Conditioning on $\mathbf{J} =$ or > 1 and using (4.3.2), (4.1.3), and (4.3.3), for any a, b

$$\mathbb{E}\left[(1-F)^{2a\bar{\mathbf{d}}_{\mathbf{J}}} |\mathbf{J} > \mathbf{1}\right] =$$

$$\begin{aligned} & \mathbb{E}\left[(1-F)^{2a\bar{\mathbf{d}}_{\mathbf{J}}}\right] - \mathbb{P}\left[\mathbf{J} = \mathbf{1}\right] \mathbb{E}\left[(1-F)^{2a\bar{\mathbf{d}}_{\mathbf{J}}} |\mathbf{J} = \mathbf{1}\right] \\ = \quad & \mathbb{E}\left[(1-F)^{2a\bar{\mathbf{d}}\wedge t}\right] - \mathbb{P}\left[\bar{\mathbf{d}} \geq t\right] (1-F)^{2at}, \end{aligned} \tag{4.5.3}$$

each term of which can be computed by the recipes in section 3.3.

In exactly the same way

$$\mathbb{E}\left[(1-F)^{2b\bar{\mathbf{d}}_{\mathbf{J}}}\,|\mathbf{J}>\mathbf{1}\right]=$$

$$\begin{aligned}
& \mathbb{E}\left[(1-F)^{2b\bar{\mathbf{d}}_{\mathbf{J}}}\right]-\mathbb{P}\left[\mathbf{J}=\mathbf{1}\right]\mathbb{E}\left[(1-F)^{2b\bar{\mathbf{d}}_{\mathbf{J}}}\,|\mathbf{J}=\mathbf{1}\right] \\
= \;& \mathbb{E}\left[(1-F)^{2b\bar{\mathbf{d}}\wedge t}\right]-\mathbb{P}\left[\bar{\mathbf{d}}\geq t\right](1-F)^{2bt}.
\end{aligned} \tag{4.5.4}$$

The numerator of (4.5.2) requires a little more. Assume for convenience that $a\geq b$ (the result has to be symmetric anyway)

$$\mathbb{E}\left[(1-F)^{2a\bar{\mathbf{d}}_{\mathbf{J}}}(1-F)^{2b\bar{\mathbf{d}}_1}\,|\mathbf{J}>\mathbf{1}\right]=$$

$$\begin{aligned}
= \;& \mathbb{E}\left[(1-F)^{2a\bar{\mathbf{d}}_{\mathbf{J}}}(1-F)^{2b\left[t-\left(\bar{\mathbf{d}}_{\mathbf{J}}+\mathbf{d}_{\mathbf{J}-1}+\ldots+\mathbf{d}_2\right)\right]}\,|\mathbf{J}>\mathbf{1}\right] \\
= \;& (1-F)^{2bt}\,\mathbb{E}\left[(1-F)^{2(a-b)\bar{\mathbf{d}}_{\mathbf{J}}}\;|\mathbf{J}>\mathbf{1}\right]\cdot \\
& \cdot\mathbb{E}\left[(1-F)^{-2b(\mathbf{d}_{\mathbf{J}-1}+\ldots+\mathbf{d}_2)}\,|\mathbf{J}>\mathbf{1}\right]\text{ by (4.3.4)} \\
= \;& (1-F)^{2bt}\left\{\mathbb{E}\left[(1-F)^{2(a-b)\bar{\mathbf{d}}\wedge t}\right]-\mathbb{P}\left[\bar{\mathbf{d}}\geq t\right](1-F)^{2(a-b)t}\right\}\cdot \\
& \cdot\mathbb{E}\left[\left(\mathbb{E}\left[(1-F)^{-2b\mathbf{d}}\right]\right)^{\mathbf{J}-2}|\mathbf{J}>\mathbf{1}\right]
\end{aligned} \tag{4.5.5}$$

by conditioning on $\mathbf{J}$ = or >1 and (4.3.2), (4.1.3), and (4.3.3) for the first expectation, and by (4.3.4) for the second expectation. We've seen something like that last factor before.

Analogous to (3.1) define H by

$$(1-H)=(1-F)^{-2b} \tag{4.5.6}$$

so as in (3.2)

$$\mathbb{E}\left[(1-H)^{\mathbf{d}}\right]=\mathbb{E}\left[(1-F)^{-2b\mathbf{d}}\right]$$

and by corollary (3.4)

$$\mathbb{E}\left[\left(\mathbb{E}\left[(1-F)^{-2b\mathbf{d}}\right]\right)^{\mathbf{J}-2}|\mathbf{J}>\mathbf{1}\right]=$$

$$= \;(1-F)^{-2bt}\,\frac{\mathbb{E}\left[(1-F)^{2b\bar{\mathbf{d}}\wedge t}\right]-\mathbb{P}\left[\bar{\mathbf{d}}\geq t\right](1-F)^{2bt}}{\mathbb{E}\left[(1-F)^{-2b\bar{\mathbf{d}}\wedge t}\right]-\mathbb{P}\left[\bar{\mathbf{d}}\geq t\right](1-F)^{-2bt}} \quad \text{using (4.5.6)} \tag{4.5.7}$$

So

$$
\begin{aligned}
\rho_{a,b} &= \frac{1}{D}\left\{\mathbb{E}\left[(1-F)^{2(a-b)\bar{\mathbf{d}}\wedge t}\right] - \mathbb{P}\left[\bar{\mathbf{d}} \geq t\right](1-F)^{2(a-b)t}\right\} \text{ where} \\
D &= \left\{\mathbb{E}\left[(1-F)^{2a\bar{\mathbf{d}}\wedge t}\right] - \mathbb{P}\left[\bar{\mathbf{d}} \geq t\right](1-F)^{2at}\right\} \cdot \\
&\quad \cdot \left\{\mathbb{E}\left[(1-F)^{-2b\bar{\mathbf{d}}\wedge t}\right] - \mathbb{P}\left[\bar{\mathbf{d}} \geq t\right](1-F)^{-2bt}\right\}, \text{ by (4.5.2)-(4.5.5)} \\
&\quad \text{and (4.5.7), and} \\
\rho_{a_1,\ldots,a_k} &= \rho_{a_1,a_2+\ldots+a_k}\rho_{a_2,\ldots,a_k} \text{ recursively.} \qquad (4.5.8)
\end{aligned}
$$

. These $\rho_{a_1,\ldots,a_k}$ of lemma (2.9) can be computed using the formulae of section 3.3 so that they can be available to calculate (2.8) along the lines of (2.10)-(2.11).

5 Derivation of the Approximation Series

The only remaining loose end is to justify the series (2.2)-(2.4) suggested to approximate the moments of $\mathbf{r}_t$. These follow from an Edgeworth expansion of the probability density function for $\ln(\mathbf{r}_t)$. Section 5.1 will present the general Edgeworth expansion. Section 5.2 will apply it to the approximation of the moments of $\mathbf{r}_t$.

5.1 Edgeworth expansion

The Edgeworth usually gets presented in terms of Hermite polynomials, taking advantage of their orthogonality properties. For example, see the development in [8]. For this presentation, we work directly with Fourier transforms because the nature of the approximation will be more transparent and the polynomials will disappear almost as soon as they arise anyway as we move toward the moments we need.

We'll define our Fourier transforms as

$$
\begin{aligned}
\hat{f}(t) &= \int_{-\infty}^{\infty} e^{-itx} f(x)\,dx, \text{ so} \\
\hat{f}^{(n)}(t) &= (-1)^n i^n \int_{-\infty}^{\infty} x^n e^{-itx} f(x)\,dx \\
&= i^{-n}\widehat{(x^n f)}(t), \text{ and} \qquad (5.1.1) \\
\widehat{\left(f^{(n)}\right)}(t) &= i^n t^n \hat{f}(t) \qquad (5.1.2)
\end{aligned}
$$

When $f_{\mathbf{X}}(x)$ is the probability density function for a random variable $\mathbf{X}$

$$
\hat{f}_{\mathbf{X}}^{(n)}(0) = i^{-n}\mathbf{E}\left[\mathbf{X}^n\right] \text{ when the expectation exists.} \qquad (5.1.3)
$$

If $\phi(x)$ is the standard normal probability density function

$$\begin{aligned}
\phi(x) &= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}, \text{ then} && (5.1.4)\\
\widehat{\phi}(t) &= e^{-\frac{1}{2}t^2} \text{ is well-known (integrate by parts) and} && (5.1.5)\\
\widehat{\phi}^{(n)}(0) &= i^{-n} n? \text{ follows, where we define} && (5.1.6)\\
n? &= (n-1)(n-3)\cdots 1 \text{ for } n \text{ even and} && (5.1.7)\\
n? &= 0 \text{ for } n \text{ odd, and} &&\\
\left(\frac{1}{\widehat{\phi}(t)}\right)^{(n)}_{t=0} &= n? \text{ by (5.1.5) and (5.1.6) (it can only differ} && (5.1.8)\\
&\quad \text{from (5.1.6) by signs and the reciprocal of} &&\\
&\quad \text{(5.1.5) can produce only } + \text{ signs.)} &&
\end{aligned}$$

Let $\mathbf{W}$ be any random variable with mean 0 and variance 1; let $f_{\mathbf{W}}(x)$ be its probability density function. Then

$$\begin{aligned}
\widehat{f_{\mathbf{W}}}(t) &= \left[\widehat{f_{\mathbf{W}}}(t)\left(\frac{1}{\widehat{\phi}(t)}\right)\right]\widehat{\phi}(t)\\
&= \left\{\sum_{n=0}^{\infty}\frac{1}{n!}\left[\widehat{f_{\mathbf{W}}}(t)\left(\frac{1}{\widehat{\phi}(t)}\right)\right]^{(n)}_{t=0} t^n\right\}\widehat{\phi}(t), \text{ by Taylor's expansion}
\end{aligned}$$

so

$$f_{\mathbf{W}}(w) = \sum_{n=0}^{\infty}\frac{1}{n!}\left[\widehat{f_{\mathbf{W}}}(t)\left(\frac{1}{\widehat{\phi}(t)}\right)\right]^{(n)}_{t=0} i^{-n}\phi^{(n)}(w), \text{ by (5.1.2)} \qquad (5.1.9)$$

Use Leibniz's rule on the derivative, keeping the first term isolated to prepare for a trick

$$\left[\widehat{f_{\mathbf{W}}}(t)\left(\frac{1}{\widehat{\phi}(t)}\right)\right]^{(n)}_{t=0} = \left(\frac{1}{\widehat{\phi}(t)}\right)^{(n)}_{t=0} + \sum_{j=1}^{n}\frac{n!}{j!\,(n-j)!}\widehat{f_{\mathbf{W}}}^{(j)}(0)\left(\frac{1}{\widehat{\phi}(t)}\right)^{(n-j)}_{t=0}.$$

Here's the trick:

$$0 = \left[\widehat{\phi}(t)\left(\frac{1}{\widehat{\phi}(t)}\right)\right]^{(n)}_{t=0} = \left(\frac{1}{\widehat{\phi}(t)}\right)^{(n)}_{t=0} + \sum_{j=1}^{n}\frac{n!}{j!\,(n-j)!}\widehat{\phi}^{(j)}(0)\left(\frac{1}{\widehat{\phi}(t)}\right)^{(n-j)}_{t=0},$$

so subtracting

$$\left[\widehat{f_{\mathbf{W}}}(t)\left(\frac{1}{\widehat{\phi}(t)}\right)\right]^{(n)}_{t=0} = \sum_{j=1}^{n}\frac{n!}{j!\,(n-j)!}\left(\widehat{f_{\mathbf{W}}}^{(j)}(0) - \widehat{\phi}^{(j)}(0)\right)\left(\frac{1}{\widehat{\phi}(t)}\right)^{(n-j)}_{t=0}.$$

But $\mathbf{W}$ is mean 0 variance 1 so by (5.1.3)

$$\left[\widehat{f_{\mathbf{W}}}(t)\left(\frac{1}{\widehat{\phi}(t)}\right)\right]^{(n)}_{t=0}=\sum_{j=3}^{n}\frac{n!}{j!\,(n-j)!}\left(\widehat{f_{\mathbf{W}}}^{(j)}(0)-\widehat{\phi}^{(j)}(0)\right)\left(\frac{1}{\widehat{\phi}(t)}\right)^{(n-j)}_{t=0}.$$

Now use (5.1.3), (5.1.6) and (5.1.8) to provide expressions for the derivatives

$$\left[\widehat{f_{\mathbf{W}}}(t)\left(\frac{1}{\widehat{\phi}(t)}\right)\right]^{(n)}_{t=0}=\sum_{j=3}^{n}\frac{n!\,(n-j)?}{j!\,(n-j)!}i^{-j}\left(\mathbb{E}\left[\mathbf{W}^{j}\right]-j?\right).$$

Put this expression into (5.1.9)

$$f_{\mathbf{W}}(w)=\phi(w)+\sum_{n=0}^{\infty}\frac{1}{n!}\sum_{j=3}^{n}\frac{n!\,(n-j)?}{j!\,(n-j)!}i^{-n-j}\left(\mathbb{E}\left[\mathbf{W}^{j}\right]-j?\right)\phi^{(n)}(w)$$

and change the order of summation, emphasizing the limit so as to prepare for a change of variables

$$f_{\mathbf{W}}(w)=\phi(w)+\lim_{N\to\infty}\sum_{j=3}^{N}\frac{1}{j!}\left(\mathbb{E}\left[\mathbf{W}^{j}\right]-j?\right)\sum_{n=j}^{N}\frac{(n-j)?}{(n-j)!}i^{-n-j}\phi^{(n)}(w).$$

Now change $n+j$ for n

$$\begin{aligned}f_{\mathbf{W}}(w)&=\phi(w)+\lim_{N\to\infty}\sum_{j=3}^{N}\frac{1}{j!}\left(\mathbb{E}\left[\mathbf{W}^{j}\right]-j?\right)\sum_{n=0}^{N-j}\frac{n?}{n!}i^{-n-2j}\phi^{(n+j)}(w)\\&=\phi(w)+\lim_{N\to\infty}\sum_{j=3}^{N}\frac{1}{j!}\left(\mathbb{E}\left[\mathbf{W}^{j}\right]-j?\right)\sum_{n=0}^{\left\lfloor\frac{N-j}{2}\right\rfloor}\frac{(2n)?}{(2n)!}i^{-2n-2j}\phi^{(2n+j)}(w),\\&\qquad\text{by (5.1.7) where }\left\lfloor\frac{N-j}{2}\right\rfloor\text{ is the largest integer in }\frac{N-j}{2}.\\&=\phi(w)+\lim_{N\to\infty}\sum_{j=3}^{N}\frac{1}{j!}\left(\mathbb{E}\left[\mathbf{W}^{j}\right]-j?\right)\sum_{n=0}^{\left\lfloor\frac{N-j}{2}\right\rfloor}\frac{(2n)?}{(2n)!}(-1)^{n+j}\phi^{(2n+j)}(w).\end{aligned}\tag{5.1.10}$$

The derivative can be evaluated directly from the definition (5.1.4) of ϕ, either laboriously or more expeditiously using Faá di Bruno's formula (see [5]) for the chain rule of higher derivatives. Either way the result is

$$\phi^{(2n+j)}(w)=\left[\sum_{k=0}^{n+\left\lfloor\frac{j}{2}\right\rfloor}\frac{(2n+j)!\,(2k)?}{(2n+j-2k)!\,(2k)!}(-1)^{2n+j-k}w^{2n+j-2k}\right]\phi(w).$$

The polynomials in the brackets are the Hermite polynomials. Substitute into (5.1.10) and simplify the signs

$$f_{\mathbf{W}}(w) = \phi(w) + \lim_{N\to\infty} \sum_{j=3}^{N} \frac{1}{j!} \left(\mathbb{E}\left[\mathbf{W}^j\right] - j?\right) \sum_{n=0}^{\left\lfloor \frac{N-j}{2} \right\rfloor} \frac{(2n)?}{(2n)!} (-1)^n \cdot$$
$$\cdot \sum_{k=0}^{n+\left\lfloor \frac{j}{2} \right\rfloor} \frac{(2n+j)!\,(2k)?}{(2n+j-2k)!\,(2k)!} (-1)^k w^{2n+j-2k} \phi(w)$$

This is the Edgeworth expansion for $\mathbf{W}$. Remember that $\mathbf{W}$ is mean 0 variance 1. For a more general random variable $\mathbf{Y} = \sigma\mathbf{W}+\mu$ a simple change of variables gives

$$f_{\mathbf{Y}}(y) = \frac{1}{\sigma}\phi\left(\frac{y-\mu}{\sigma}\right) + \lim_{N\to\infty} \sum_{j=3}^{N} \frac{1}{j!} \left(\frac{\mu_j}{\sigma^j} - j?\right) \sum_{n=0}^{\left\lfloor \frac{N-j}{2} \right\rfloor} \frac{(2n)?}{(2n)!} (-1)^n \cdot$$
$$\cdot \sum_{k=0}^{n+\left\lfloor \frac{j}{2} \right\rfloor} \frac{(2n+j)!\,(2k)?}{(2n+j-2k)!\,(2k)!} (-1)^k \left(\frac{y-\mu}{\sigma}\right)^{2n+j-2k} \frac{1}{\sigma}\phi\left(\frac{y-\mu}{\sigma}\right),$$

where μ_j is the j-th central moment of $\mathbf{Y}$.

(5.1.11)

(5.1.11) is the Edgeworth expansion for the probability density function of a general random variable $\mathbf{Y}$. This is the nexus of a lot of possible alternatives so before moving on to use it in section 5.2 for our current application it is worth pointing out some of the sights. If we had chosen to make the Taylor's expansion at (5.1.9) around some point other than $t = 0$ we would have arrived at an expression that, upon a certain optimization of the point t, is known to statisticians as the saddlepoint approximation ([8]) and to actuaries as the Esscher approximation ([3] and [4]).

If we had chosen to set up the expression at (5.1.9) using the Fourier transform $\widehat{\psi}$ of some other probability density ψ rather than $\widehat{\phi}$ of the standard normal ϕ, then we would have arrived at an expression involving $\psi\left(\frac{y-\mu}{\sigma}\right)$, a different polynomial (or some other function) reflecting derivatives of ψ (Faá di Bruno's formula [5] could come in handy), and the difference between standardized central moments of $\mathbf{Y}$ and the central moments of ψ. Thus ψ would serve as a kind of model for $f_{\mathbf{Y}}$ with an approximation analogous to (5.1.11) building off departures of the moments of $f_{\mathbf{Y}}$ from those of ψ. Depending upon the situation, a judicious choice for ψ could affect dramatically either the convergence of the approximation or the expanatory power of the expansion, or both. For heavy-tail situations (not what we have in the present application) a logistic ψ might be suggested. In applications involving regime-switching, likely candidates for ψ include the gamma, the inverse Gaussian, and the inverse logistic (defined by analogy with the inverse Gaussian) all of which are implicated in various waiting times.

Considering the power that the Esscher has shown in actuarial and financial applications (see [6], for the primal example of the latter) this nexus of possible models ought to be a rich mine for actuaries and financial engineers looking for computable and/or explanatory models.

5.2 Approximation for the Moments of $\mathbf{r}_t$

Take $\mathbf{Y}$ in (5.1.11) to be $\ln(\mathbf{r}_t)$ and use it to create an expression for the l-th moment $\mathbb{E}\left[\mathbf{r}_t^l\right]$ of $\mathbf{r}_t$.

$$\begin{aligned}\mathbb{E}\left[\mathbf{r}_t^l\right] &= \mathbb{E}\left[\mathbf{e}^{l\ln(\mathbf{r}_t)}\right] \\ &= \mathbb{E}\left[\mathbf{e}^{l\mathbf{Y}}\right] \\ &= \int_{-\infty}^{\infty} e^{ly} f_{\mathbf{Y}}(y)\, dy\end{aligned}$$

$$= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} \left\{ 1 + \lim_{N\to\infty} \sum_{j=3}^{N} \frac{1}{j!} \left(\frac{\mu_j}{\sigma^j} - j? \right) \sum_{n=0}^{\left\lfloor \frac{N-j}{2} \right\rfloor} \frac{(2n)?}{(2n)!} (-1)^n \cdot \sum_{k=0}^{n+\left\lfloor \frac{j}{2} \right\rfloor} \frac{(2n+j)!\,(2k)?}{(2n+j-2k)!\,(2k)!} (-1)^k \left(\frac{y-\mu}{\sigma} \right)^{2n+j-2k} \right\} e^{ly-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2} dy$$

Now complete the square in the exponent, just as if you were finding the moments of the ordinary lognormal

$$= \frac{e^{l\mu+\frac{1}{2}(l\sigma)^2}}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} \left\{ 1 + \lim_{N\to\infty} \sum_{j=3}^{N} \frac{1}{j!} \left(\frac{\mu_j}{\sigma^j} - j? \right) \sum_{n=0}^{\left\lfloor \frac{N-j}{2} \right\rfloor} \frac{(2n)?}{(2n)!} (-1)^n \cdot \sum_{k=0}^{n+\left\lfloor \frac{j}{2} \right\rfloor} \frac{(2n+j)!\,(2k)?}{(2n+j-2k)!\,(2k)!} (-1)^k \left(\frac{y-\mu}{\sigma} \right)^{2n+j-2k} \right\} e^{-\frac{1}{2}\left(\frac{y-\mu}{\sigma}-l\sigma\right)^2} dy$$

and make the usual change of variables $z = \frac{y-\mu}{\sigma} - l\sigma$

$$= e^{l\mu+\frac{1}{2}(l\sigma)^2} \int_{-\infty}^{\infty} \left\{ 1 + \lim_{N\to\infty} \sum_{j=3}^{N} \frac{1}{j!} \left(\frac{\mu_j}{\sigma^j} - j? \right) \sum_{n=0}^{\left\lfloor \frac{N-j}{2} \right\rfloor} \frac{(2n)?}{(2n)!} (-1)^n \cdot \sum_{k=0}^{n+\left\lfloor \frac{j}{2} \right\rfloor} \frac{(2n+j)!\,(2k)?}{(2n+j-2k)!\,(2k)!} (-1)^k (z+l\sigma)^{2n+j-2k} \right\} \phi(z)\, dz$$

and expand the binomial, applying (5.1.3), (5.1.6), and (5.1.7) to evaluate the resulting integrals against $\phi(z)$

$$\begin{aligned}\mathbb{E}\left[\mathbf{r}_t^l\right] &= e^{l\mu+\frac{1}{2}(l\sigma)^2}\left\{1+\lim_{N\to\infty}\sum_{j=3}^{N}\frac{1}{j!}\left(\frac{\mu_j}{\sigma^j}-j?\right)\sum_{n=0}^{\left\lfloor\frac{N-j}{2}\right\rfloor}\frac{(2n)?}{(2n)!}(-1)^n\cdot\right.\\ &\quad\cdot\sum_{k=0}^{n+\left\lfloor\frac{j}{2}\right\rfloor}\frac{(2n+j)!\,(2k)?}{(2n+j-2k)!\,(2k)!}(-1)^k\cdot\\ &\quad\left.\cdot\sum_{m=0}^{n+\left\lfloor\frac{j}{2}\right\rfloor-k}\frac{(2n+j-2k)!\,(2m)?}{(2n+j-2k-2m)!\,(2m)!}(l\sigma)^{2n+j-2k-2m}\right\}.\end{aligned}$$

Cancel factors where possible and distribute powers of l and σ to logical terms across all the sums

$$\begin{aligned}&= e^{l\mu+\frac{1}{2}(l\sigma)^2}\left\{1+\lim_{N\to\infty}\sum_{j=3}^{N}\frac{l^j}{j!}\left(\mu_j-j?\sigma^j\right)\sum_{n=0}^{\left\lfloor\frac{N-j}{2}\right\rfloor}\frac{(2n)?}{(2n)!}(-1)^n(l\sigma)^{2n}\cdot\right.\\ &\quad\left.\cdot\sum_{k=0}^{n+\left\lfloor\frac{j}{2}\right\rfloor}\sum_{m=0}^{n+\left\lfloor\frac{j}{2}\right\rfloor-k}\frac{(2n+j)!\,(2k)?\,(2m)?}{(2n+j-2k-2m)!\,(2k)!\,(2m)!}(-1)^k(l\sigma)^{-2(k+m)}\right\},\end{aligned}$$

make the change of variables $m-k$ for m, and reverse the order of the summations

$$\begin{aligned}&= e^{l\mu+\frac{1}{2}(l\sigma)^2}\left\{1+\lim_{N\to\infty}\sum_{j=3}^{N}\frac{l^j}{j!}\left(\mu_j-j?\sigma^j\right)\sum_{n=0}^{\left\lfloor\frac{N-j}{2}\right\rfloor}\frac{(2n)?}{(2n)!}(-1)^n(l\sigma)^{2n}\cdot\right.\\ &\quad\left.\cdot\sum_{m=0}^{n+\left\lfloor\frac{j}{2}\right\rfloor}\frac{(2n+j)!}{(2n+j-2m)!}(l\sigma)^{-2m}\sum_{k=0}^{m}\frac{(2k)?\,(2(m-k))?}{(2k)!\,(2(m-k))!}(-1)^k\right\}.\end{aligned}\tag{5.2.1}$$

Remarkably, the sum over k is 0 for each $m>1$. This is by inspection for odd m but a true holiday venture to prove algebraically for even m. Luckily there's a trick to avoid the algebra. Use Leibniz's rule, (5.1.6) and (5.1.8) to evaluate

$$
\begin{aligned}
0 &= \left[\widehat{\phi}(t)\left(\frac{1}{\widehat{\phi}(t)}\right)\right]_{t=0}^{(2m)} \\
&= (2m)!\sum_{s=0}^{2m}\frac{(s)?\,(2m-s)?}{(s)!\,(2m-s)!}i^{-s} \\
&= (2m)!\sum_{k=0}^{m}\frac{(2k)?\,(2\,(m-k))?}{(2k)!\,(2\,(m-k))!}i^{-2k}, \text{ by (5.1.7).}
\end{aligned}
$$

Therefore (5.2.1) reduces to

$$
\mathbb{E}\left[\mathbf{r}_t^l\right] = e^{l\mu+\frac{1}{2}(l\sigma)^2}\left\{1+\lim_{N\to\infty}\sum_{j=3}^{N}\frac{l^j}{j!}\left(\mu_j - j?\sigma^j\right)\sum_{n=0}^{\left\lfloor\frac{N-j}{2}\right\rfloor}\frac{(2n)?}{(2n)!}\left(-1\right)^n\left(l\sigma\right)^{2n}\right\},
$$

which is the approximation series that we sought. For the application at (2.2)-(2.4) we had $\mu_j = 0$ for odd j so we were able to express both the limit and the sum in terms of even integers.

6 References

1. Bridgeman, J. G., "*Random switching times among randomly parameterized regimes of random interest rate scenarios*", Actuarial Research Clearing House (ARCH) 2007.1 (January 2007)

2. de la Peña, V. H. and Giné, E., Decoupling, From Dependence To Independence, Springer (1999)

3. Esscher, F., "*On the probability function in the collective theory of risk*", Skand. Act. Tidskr.(1932) 175-195

4. Esscher, F., "*On approximate computations when the corresponding characteristic functions are known*", Skand. Act. Tidskr. (1963) 78-86

5. Flanders, H., "*From Ford to Faá*", The American Mathematical Monthly 108 (June-July 2001) 559-561

6. Gerber, H. U. and Shiu, E. S. W., "*Option pricing by Esscher transforms*", Transactions of the Society of Actuaries 46 (1994) 99-140

7. Ho, T. S. and Lee, S. B., The Oxford Guide To Financial Modeling, Oxford University Press (2004)

8. Jensen, J. L., Saddlepoint Approximations, Oxford University Press (1995)