M11 Lognormal Models for Investment Rates

Topics in Insurance, Risk, and Finance

Yuyu Chen

2026

Learning outcomes

- Know and derive the properties of lognormal distributions.
- Derive explicit formulas for the distributions and related quantities of S_n and V_n in the lognormal model.
- Understand the unique role of the lognormal distribution for modelling rates of return in varying rate models.
- Define, calculate, and interpret Value-at-Risk.

Review: normal distribution

Lognormal model

- We have studied the moments of cash series.
- To have a more detailed analysis of investment activities (e.g., probability of default), we need to derive distribution functions.
- In particular, distributions can be used to determine the capital requirements for financial institutions (Value-at-Risk for insurance companies and Expected Shortfall for banks).
- The task of finding distributions is however very challenging.
- We will study the case when the accumulation factor $(1 + i_t)$ follows a lognormal distribution, in which case we can derive the exact distribution of S_n.

Review: normal distribution I

We say a random variable X follows a normal distribution with parameters μ and σ^2, denoted by $X \sim N(\mu, \sigma^2)$, if the density function of X can be written as

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right).$$

- μ and σ^2 are the mean and variance of X.
- If $\mu = 0$ and $\sigma = 1$, X is usually denoted by Z and is called a standard normal random variable.
- The distribution function and the density function of Z are usually denoted by Φ and ϕ, and are called the standard normal distribution function/density function.

Review: normal distribution II

- Symmetry: $\Phi(x) + \Phi(-x) = 1$ for $x \in \mathbb{R}$.
- Standardization: For $X \sim N(\mu, \sigma^2)$, we have

$$\frac{X-\mu}{\sigma} \sim N(0, 1).$$

- Moment generating function: Let $X \sim N(\mu, \sigma^2)$. For $t \in \mathbb{R}$,

$$\mathbb{E}[\exp(tX)] = \exp(\mu t + t^2\sigma^2/2).$$

- If $X_1 \sim N(\mu_1, \sigma_1^2)$ and $X_2 \sim N(\mu_2, \sigma_2^2)$ are independent,

$$c_1X_1 + c_2X_2 \sim N(c_1\mu_1 + c_2\mu_2, c_1^2\sigma_1^2 + c_2^2\sigma_2^2).$$

Question: If $X_1 \sim N(\mu_1, \sigma_1^2)$ and $X_2 \sim N(\mu_2, \sigma_2^2)$, does $X_1 + X_2$ always follow the normal distribution?

Review: normal distribution III

STANDARD NORMAL DISTRIBUTION: Table Values Represent AREA to the LEFT of the Z score.

Z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
-3.9	.00005	.00005	.00004	.00004	.00004	.00004	.00004	.00004	.00003	.00003
-3.8	.00007	.00007	.00007	.00006	.00006	.00006	.00006	.00005	.00005	.00005
-3.7	.00011	.00010	.00010	.00010	.00009	.00009	.00008	.00008	.00008	.00008
-3.6	.00016	.00015	.00015	.00014	.00014	.00013	.00013	.00012	.00012	.00011
-3.5	.00023	.00022	.00022	.00021	.00020	.00019	.00019	.00018	.00017	.00017
-3.4	.00034	.00032	.00031	.00030	.00029	.00028	.00027	.00026	.00025	.00024
-3.3	.00048	.00047	.00045	.00043	.00042	.00040	.00039	.00038	.00036	.00035
-3.2	.00069	.00066	.00064	.00062	.00060	.00058	.00056	.00054	.00052	.00050
-3.1	.00097	.00094	.00090	.00087	.00084	.00082	.00079	.00076	.00074	.00071
-3.0	.00135	.00131	.00126	.00122	.00118	.00114	.00111	.00107	.00104	.00100
-2.9	.00187	.00181	.00175	.00169	.00164	.00159	.00154	.00149	.00144	.00139
-2.8	.00256	.00248	.00240	.00233	.00226	.00219	.00212	.00205	.00199	.00193
-2.7	.00347	.00336	.00326	.00317	.00307	.00298	.00289	.00280	.00272	.00264
-2.6	.00466	.00453	.00440	.00427	.00415	.00402	.00391	.00379	.00368	.00357
-2.5	.00621	.00604	.00587	.00570	.00554	.00539	.00523	.00508	.00494	.00480
-2.4	.00820	.00798	.00776	.00755	.00734	.00714	.00695	.00676	.00657	.00639
-2.3	.01072	.01044	.01017	.00990	.00964	.00939	.00914	.00889	.00866	.00842
-2.2	.01390	.01355	.01321	.01287	.01255	.01222	.01191	.01160	.01130	.01101
-2.1	.01786	.01743	.01700	.01659	.01618	.01578	.01539	.01500	.01463	.01426
-2.0	.02275	.02222	.02169	.02118	.02068	.02018	.01970	.01923	.01876	.01831
-1.9	.02872	.02807	.02743	.02680	.02619	.02559	.02500	.02442	.02385	.02330
-1.8	.03593	.03515	.03438	.03362	.03288	.03216	.03144	.03074	.03005	.02938
-1.7	.04457	.04363	.04272	.04182	.04093	.04006	.03920	.03836	.03754	.03673
-1.6	.05480	.05370	.05262	.05155	.05050	.04947	.04846	.04746	.04648	.04551
-1.5	.06681	.06552	.06426	.06301	.06178	.06057	.05938	.05821	.05705	.05592
-1.4	.08076	.07927	.07780	.07636	.07493	.07353	.07215	.07078	.06944	.06811
-1.3	.09680	.09510	.09342	.09176	.09012	.08851	.08691	.08534	.08379	.08226
-1.2	.11507	.11314	.11123	.10935	.10749	.10565	.10383	.10204	.10027	.09853
-1.1	.13567	.13350	.13136	.12924	.12714	.12507	.12302	.12100	.11900	.11702
-1.0	.15866	.15625	.15386	.15151	.14917	.14686	.14457	.14231	.14007	.13786
-0.9	.18406	.18141	.17879	.17619	.17361	.17106	.16853	.16602	.16354	.16109
-0.8	.21186	.20897	.20611	.20327	.20045	.19766	.19489	.19215	.18943	.18673
-0.7	.24196	.23885	.23576	.23270	.22965	.22663	.22363	.22065	.21770	.21476
-0.6	.27425	.27093	.26763	.26435	.26109	.25785	.25463	.25143	.24825	.24510
-0.5	.30854	.30503	.30153	.29806	.29460	.29116	.28774	.28434	.28096	.27760
-0.4	.34458	.34090	.33724	.33360	.32997	.32636	.32276	.31918	.31561	.31207
-0.3	.38209	.37828	.37448	.37070	.36693	.36317	.35942	.35569	.35197	.34827
-0.2	.42074	.41683	.41294	.40905	.40517	.40129	.39743	.39358	.38974	.38591
-0.1	.46017	.45620	.45224	.44828	.44433	.44038	.43644	.43251	.42858	.42465
-0.0	.50000	.49601	.49202	.48803	.48405	.48006	.47608	.47210	.46812	.46414

Review: normal distribution example

For $X \sim N(5, 5^2)$, what is $\mathbb{P}(X \leqslant 6)$?

Lognormal distribution

Lognormal distribution: definition

Let Z be a standard normal random variable, i.e., $Z \sim N(0,1)$. Let $\mu \in \mathbb{R}$ and $\sigma > 0$. A lognormal random variable can be written as

$$X = \exp(\mu + \sigma Z).$$

- We write $X \sim LN(\mu, \sigma^2)$.
- Note that μ and σ^2 here are not the mean and variance of X.
- The logarithm of X is normally distributed (with mean μ and variance σ^2), and hence the name.
- Equivalently, we can write

  $$X = \exp(Y)$$

  where $Y \sim N(\mu, \sigma^2)$.

Lognormal distribution: density

Property 1. For $X \sim LN(\mu, \sigma^2)$, its density function is

$$f(x) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left(-\frac{(\log x - \mu)^2}{2\sigma^2}\right), \quad \text{for} \quad x > 0.$$

Proof:

Lognormal distribution: moments

Property 2. For $X \sim LN(\mu, \sigma^2)$, we have

- $\mathbb{E}[X] = \exp(\mu + 0.5\sigma^2)$;
- $\mathrm{Var}(X) = (\exp(\sigma^2) - 1)\exp(2\mu + \sigma^2) = (\exp(\sigma^2) - 1)\,\mathbb{E}[X]^2$;
- The rth moment of X is

  $$\mathbb{E}[X^r] = \mathbb{E}[\exp(r \log X)] = \exp(r\mu + r^2\sigma^2/2).$$

Proof:

Lognormal distribution: product

Property 3. If $X_1 \sim LN(\mu_1, \sigma_1^2)$ and $X_2 \sim LN(\mu_2, \sigma_2^2)$ are independent, then

$$X_1 X_2 \sim LN(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2).$$

Proof:

Lognormal distribution: reciprocal

Property 4. If $X \sim LN(\mu, \sigma^2)$, then $X^{-1} \sim LN(-\mu, \sigma^2)$.

Proof:

Lognormal models: S_n and V_n

Lognormal model

- Due the properties of lognormal distributions, it is convenient to model the accumulation factors (i.e., $1 + i_t$) as lognormal distributions.
- We can derive distributions of S_n and V_n using lognormal models, which turn out to be lognormal distributions.
- Even if the accumulation factor is not lognormally distributed, we can still use lognormal distributions to approximate the distribution of S_n and V_n.
- However, distributions of A_n and P_n are still not available.

Lognormal model: S_n and V_n in a varying rate model

Proposition 1. *Suppose that in a varying rate model* $1+i_t \sim LN(\mu, \sigma^2)$ *for* $t = 1, \dots, n$*. Then for* $s > 0$*,*

$$\mathbb{P}(S_n \leqslant s) = \Phi\left(\frac{\log s - n\mu}{\sqrt{n}\sigma}\right) \quad \textit{and} \quad \mathbb{P}(V_n \leqslant s) = \Phi\left(\frac{\log s + n\mu}{\sqrt{n}\sigma}\right).$$

Question: What if the rates are not identically distributed? Any criticism on the lognormal assumption?

Proof:

Lognormal model: S_n example

In a varying rate model, suppose that the iid returns, i_k, are such that $1+i_k \sim LN(0.12, 0.04^2)$. What is the probability that the accumulated value of 10000 at time 5 is greater than 21000?

Lognormal model: V_n example

In a varying rate model, suppose the iid returns, i_k, are such that $1+i_k \sim LN(0.08, 0.04^2)$. What is the probability that the present value of a benefit of 100 at time 10 is less than 40?

Lognormal model: approximation I

In a varying model, even if accumulation factors do not follow a lognormal distribution, one can still use lognormal distributions to approximate S_n and V_n.

Theorem 1. Let $X_1, \dots, X_n$ be iid with mean μ_X and variance $\sigma_X^2 < \infty$. Then

$$\frac{\sum_{i=1}^n X_i - n\mu_X}{\sqrt{n}\sigma_X} \to_d N(0,1),$$

where $\to_d$ means convergence in distribution.

Lognormal model: approximation II

- Since $\log S_n = \sum_{t=1}^n \log(1+i_t)$, the summation of iid random variables, by the central limit theorem, $\log S_n$ has a normal distribution as its limiting distribution, after normalization.
- Therefore, S_n can be approximated by a lognormal distribution.
- To apply the central limit theorem, technically, you need to verify the moment constraints/assumptions (why).

Lognormal model: S_n and V_n in a fixed rate model

Proposition 2. *Suppose that in a fixed rate model* $1+i \sim LN(\mu, \sigma^2)$ *where* i *is the rate of return for each period of time. Then for* $s > 0$*,*

$$\mathbb{P}(S_n \leqslant s) = \Phi\left(\frac{\log s - n\mu}{n\sigma}\right) \quad \textit{and} \quad \mathbb{P}(V_n \leqslant s) = \Phi\left(\frac{\log s + n\mu}{n\sigma}\right).$$

Proof:

Lognormal model: a comparison

Let the accumulation factor for one period follow $LN(\mu, \sigma^2)$. Calculate the coefficient of variation of S_n in a varying rate model and fixed rate model respectively.

We will see that S_n is more spread-out in the fixed rate model, which is intuitive (why?).

Value-at-Risk

With the distribution of S_n, more risk metrics of S_n can be calculated. For a random variable X, Value-at-Risk (VaR) at level $p \in (0,1)$ is defined as

$$\mathrm{VaR}_p(X) = F^{-1}(p) = \inf\{x : \mathbb{P}(X \leqslant x) \geqslant p\}.$$

- Here, X is regarded as losses.
- In general, p is close to 1 (e.g., $p = 0.99$).

- VaR is used as a regulatory risk measure in the realm of bank and insurance.
- VaR: the event that the loss is greater than this level has a probability less than $1-p$.
- For a continuous random variable $X \sim F$, $\mathrm{VaR}_p(X)$ is the inverse of F.

Value-at-Risk : example A

Suppose that a loss X has the distribution that $P(X=-7)=0.04$ and $P(X=5.5)=0.96$. What is $\mathrm{VaR}_{0.95}(X)$? Ans: $\mathrm{VaR}_{0.95}(X)=5.5$.

Value-at-Risk : example B I

In a varying rate model, the annual accumulation factor follows a lognormal distribution with parameters $\mu=0.075$ and $\sigma^2=0.025^2$. Let the initial investment be 1000 and the accumulated value at the end of 5 years be X. What are $\mathrm{VaR}_{0.25}(X)$ and $\mathrm{VaR}_{0.75}(X)$?

Value-at-Risk : example B II