7CCMFM02 Risk Neutral valuation: Pricing and Hedging Derivatives

Welcome to FM02. In this course you will learn the basic building blocks and mathematical techniques for pricing and hedging financial derivatives. In the first part of this course, you will learn about **discrete-time models**, in particular the **binomial model** where the stock price can only go up or down by a fixed amount at each discrete time step. At first glance this seems like a crude model, but many insights comes out of this model when we start pricing so-called **European options**, in particular the important notion of **risk-neutral pricing**, which is the backbone of the course and essentially says that the fair (or **no-arbitrage** price of an option does not depend on the real world **drift** of the stock price process. You will also learn how to price exotic **path-dependent options**, such as **American options** and **barrier options**, and this can easily be implemented in e.g. Python, Matlab, or Excel. After this we will learn about **Brownian motion** and **continuous time models** which are more interesting but more mathematically sophisticated, and the financial models and instruments you will learn about are used on real life trading desks, where trillions of dollars worth of derivatives trade hands every day. I use to work in the city on Foreign Exchange and interest rate derivatives desks, so let me know if you have any careers or technical questions about that.

The course textbook is *Risk-Neutral Valuation: Pricing and Hedging of Financial Derivatives*, by N.Bingham and R.Kiesel but I do not actively use the book these days. Assessment for the module is a two hour exam in January. There will be **weekly Homework questions** with solutions in the tutorials.

In semester 1, I strongly recommend the FM06 Numerical Methods module because you will need it for your summer project, and in Semester 2 I recommend the FM20 Stochastic Control and applications to algorithmic trading and FM18 Machine Learning modules

Prerequisites for this course

- Some background in basic **probability**, e.g. calculating **expectations** of random variables, the **Normal distribution** etc.
- You should also attend FM01 lectures on **measure theory**, since we will sometimes be talking about σ**-algebras**, and if you're keen the undergraduate courses 388 (discrete time, Semester 1) and 338 (continuous time, Semester 2)

Summary of course

On this course you will learn about:

- **The binomial model-** Pricing/hedging of European, American and other types of exotic options.
- **Brownian motion** (see first graph below), which is a continuous time random process which has independent, identically distributed increments with the property that $W_t - W_s \sim N(0, t-s)$ for $s < t$, i.e. Normally distributed with mean zero and variance $t-s$ Brownian motion has a continuous sample paths (i.e. it does not jump) but is not differentiable with respect to t, so we cannot apply standard calculus techniques to W_t, so we have to develop a new calculus called **stochastic calculus**.
- **Ito's lemma** - how to do calculus on $f(W_t)$, where f is a twice differentiable function- we will see the rules are subtly different to standard 1st year calculus.
- The **Black-Scholes model/formula and PDE** for pricing a European call option with payoff $\max(S_T - K, 0)$ on a stock price process $S_t = S_0 e^{(\mu - \frac{1}{2}\sigma^2)t + \sigma W_t}$ where W_t is Brownian motion. μ is the **drift** of the process, which describes the overall upward/downward trend, and σ is the **volatility**, which describes its variability.
- TBA

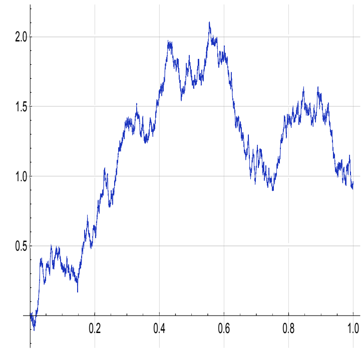

Figure 1: Euler scheme approximation for Brownian motion in Mathematica with large (left) and small (right) step size

Applied Probability Revision

- Recall that a **continuous random variable** X is a random quantity which has the property that for any set A, the probability that the realized value of X lies in the set A is given by

$$\mathbb{P}(X \in A) \quad = \quad \int_A f_X(u)du \tag{1}$$

for some function $f_X(x)$ know as the **density** of X. Note that X is a random variable here and u is just a **dummy variable** of integration, and we are assuming that X cannot be $+\infty$ or $-\infty$. In particular, if we set $A = (-\infty, x]$, the **distribution function** of X is defined as

$$F_X(x) \quad = \quad \mathbb{P}(X \in (-\infty, x]) \quad = \quad \mathbb{P}(X \le x) \quad = \quad \int_A f_X(u)du \quad = \quad \int_{-\infty}^x f_X(u)du\,. \tag{2}$$

Then differentiating and using the **fundamental theorem of calculus** we see that

$$F_X'(x) \quad = \quad \frac{d}{dx}F_X(x) \quad = \quad f_X(x)\,.$$

Setting $x = \infty$ in (2) we see that

$$\mathbb{P}(X \le \infty) \quad = \quad \mathbb{P}(X < \infty) \quad = \quad 1 \quad = \quad \int_{-\infty}^{\infty} f_X(x)dx$$

i.e. the density of any continuous random variable has to **integrate to** 1. If we set A equal to the event $\{X \le x\}$, then its **complement** $A^c = \{X > x\}$, and $\mathbb{P}(A \cup A^c) = \mathbb{P}(A) + \mathbb{P}(A^c) = 1$, since A and A^c are two **disjoint** events such that $\mathbb{P}(A \cup A^c) = 1$. The **complementary cdf** of X is defined as $\mathbb{P}(X > x)$, which is equal to $1 - \mathbb{P}(A) = 1 - \mathbb{P}(X \le x)$, and differentiating this expression with respect to x we see that

$$\frac{d}{dx}\mathbb{P}(X > x) \quad = \quad \frac{d}{dx}(1 - \mathbb{P}(X \le x)) \quad = \quad -f_X(x)$$

which will be used many times on the course. Note also that

$$\mathbb{P}(X = x) \quad = \quad \int_x^x f_X(u)du \quad = \quad 0$$

i.e. the probability that a continuous random variable X takes a particular value x is zero.

- **Example**: a standard **uniform random variable** U on $[0, 1]$ has density $f_X(x) = 1$ for $x \in [0, 1]$ and zero otherwise. The distribution function of U is obtained by integrating this density from 0 to x:

$$\mathbb{P}(U \le x) \quad = \quad \int_0^x 1\,du \quad = \quad u|_{u=x} - u|_{u=0} \quad = \quad x \tag{3}$$

for $x \in [0, 1]$. Note that the lower limit of integration is 0 not $-\infty$ here since the density of U is zero outside $[0, 1]$. Fact of the day: the sum of two iid uniform random variables is not uniform (EFR).

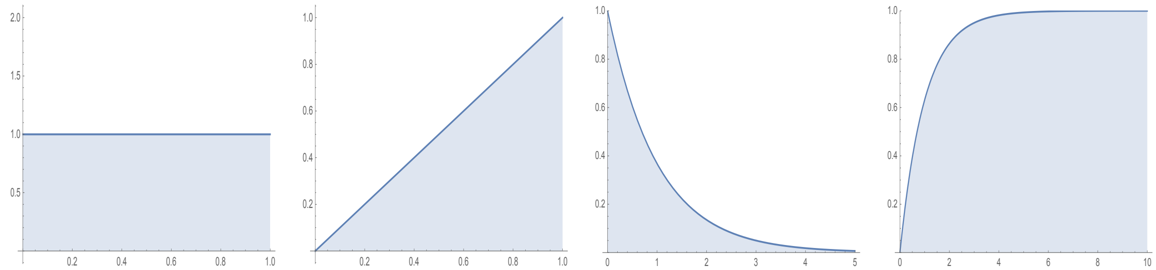

Figure 2: On the left we see the density of a standard $U[0,1]$ uniform random variable, and in the second from left panel we have plotted its distribution function $F_U(x) = x$. In the third panel we have plotted the density of an exponential random variable with parameter $\lambda = 1$, and on the right we have plotted its distribution function.

- **Example**: an **exponential random variable** X has density $f_X(x) = \lambda e^{-\lambda x}$ for $x \in (0, \infty)$ and zero otherwise, for some parameter $\lambda > 0$ (we say that X is an $\mathrm{Exp}(\lambda)$ random variable). The distribution function of X is again obtained by integrating this density from 0 to x as follows:

$$\mathbb{P}(X \leq x) \;=\; \int_0^x \lambda e^{-\lambda u} du \;=\; -e^{-\lambda u}|_{u=x} - (-e^{-\lambda u}|_{u=0}) \;=\; 1 - e^{-\lambda x}.$$

As a sanity check, we see that $\lim_{x \to 0} \mathbb{P}(X \leq x) = \mathbb{P}(X \leq 0) = 0$, and $\lim_{x \to +\infty} \mathbb{P}(X \leq x) = \mathbb{P}(X < \infty) = 1$, as we would expect since $X \geq 0$ and $X < \infty$.

- The **expectation** of a continuous random variable X with density $f_X(x)$ is

$$\mathbb{E}(X) \;=\; \int_{-\infty}^{\infty} x f_X(x) dx$$

and using this definition and the fact that $\int_{-\infty}^{\infty} f_X(x) dx = 1$ we see that

$$\mathbb{E}(X + a) \;=\; \mathbb{E}(X) + a \tag{4}$$

for any constant a. Similarly the expectation of $g(X)$ for some function g is given by

$$\mathbb{E}(g(X)) \;=\; \int_{-\infty}^{\infty} g(x) f_X(x) dx \tag{5}$$

so in particular for any constant $a \in \mathbb{R}$, setting $g(x) = ax$, we have

$$\mathbb{E}(aX) \;=\; \int_{-\infty}^{\infty} a x f_X(x) dx \;=\; a\mathbb{E}(X). \tag{6}$$

- Two random variables X and Y are **independent** if $\mathbb{P}(X \leq x, Y \leq y) = \mathbb{P}(X \leq x \cap Y \leq y) = \mathbb{P}(X \leq x)\,\mathbb{P}(Y \leq y)$. If X and Y are independent, and have density $f_X(x)$ and $f_Y(y)$ respectively, then

$$\mathbb{E}(g(X)h(Y)) \;=\; \int_{-\infty}^{\infty} g(x) f_X(x) dx \cdot \int_{-\infty}^{\infty} h(y) f_Y(y) dy. \tag{7}$$

- Recall that if X is a random variable with expectation $\mathbb{E}(X) = \mu$, the **variance** of X is

$$\mathrm{Var}(X) \;=\; \mathbb{E}[(X-\mu)^2] \;=\; \mathbb{E}[X^2 - 2X\mu + \mu^2] \;=\; \mathbb{E}(X^2) - 2\mathbb{E}(X)\mu + \mu^2 \;=\; \mathbb{E}(X^2) \,-\, \mu^2.$$

From this we see that for any constant $a \in \mathbb{R}$

$$\mathrm{Var}(aX) \;=\; \mathbb{E}((aX)^2) \,-\, (\mathbb{E}(aX))^2 \;=\; a^2\mathbb{E}(X^2) \,-\, a^2(\mathbb{E}(X))^2 \;=\; a^2\mathrm{Var}(X)$$

where we are using (6) to obtain the middle equality here. Using (4), we can also easily see that $\mathrm{Var}(X + a) = \mathbb{E}((X + a - (\mu + a))^2) = \mathrm{Var}(X)$, i.e. the constant a has no effect on variance.

- For two random variables X and Y we have

$$\begin{aligned}
\mathrm{Var}(X+Y) &= \mathbb{E}((X+Y)^2) - (\mathbb{E}(X+Y))^2 \\
&= \mathbb{E}(X^2) + 2\mathbb{E}(XY) + \mathbb{E}(Y^2) \,-\, (\mathbb{E}(X)^2 + \mathbb{E}(Y)^2 - 2\mathbb{E}(X)\mathbb{E}(Y)) \\
&= \mathrm{Var}(X) \,+\, \mathrm{Var}(Y) \,+\, 2\mathrm{Cov}(X,Y)
\end{aligned}$$

where $\mathrm{Cov}(X,Y) := \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y) = \mathbb{E}((X-\mu_X)(Y-\mu_Y))$ is known as the **covariance** of X and Y, where $\mu_X := \mathbb{E}(X)$ and $\mu_Y := \mathbb{E}(Y)$.

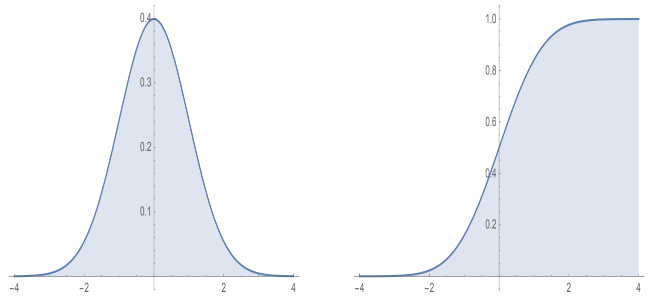

Figure 3: Here we have plotted the density $n(x)$ (left) and the distribution function $\Phi(x)$ (right) for the standard Normal distribution.

- If X and Y are **independent** random variables, then $\mathbb{E}((X-\mu_X)(Y-\mu_Y)) = \mathbb{E}(X-\mu_X)\mathbb{E}(Y-\mu_Y) = 0 \times 0 = 0$, so $\mathrm{Var}(X+Y)$ simplifies to

$$\mathrm{Var}(X+Y) \quad = \quad \mathrm{Var}(X) + \mathrm{Var}(Y)\,.$$

- If Z has a standard $N(0,1)$ **Normal distribution**, $\mathbb{E}(Z) = 0$, $\mathrm{Var}(Z) = 1$ and the density of Z is the "bell-shaped" function:

$$n(z) \quad = \quad \frac{1}{\sqrt{2\pi}} e^{-z^2/2}$$

for $z \in \mathbb{R}$ (see first plot in Figure 2 above), and the distribution function of Z is the "s-shaped" function:

$$\mathbb{P}(Z \le x) \quad = \quad \Phi(x) \quad := \quad \int_{-\infty}^{x} n(z)dz$$

and hence

$$\Phi'(x) \quad = \quad n(x)$$

(see second plot in Figure 2 above).

- Proving that $\int_{-\infty}^{\infty} n(z)dz = 1$ is not trivial, and requires working in **polar coordinates** with 2 i.i.d. (independent and identically distributed) standard Normal random variables. $\Phi(x)$ cannot be computed exactly, but there are useful **asymptotic formula** for $\Phi(x)$ when x is small or x is large, or we can look up Φ in **tables**.

- For a general Normal random variable $X \sim N(\mu, \sigma^2)$ with mean μ and variance σ^2, it can be shown that $Z = (X-\mu)/\sigma$ is a standard $N(0,1)$ random variable. Then

$$\mathbb{P}(X > x) \quad = \quad \mathbb{P}(\frac{X-\mu}{\sigma} > \frac{x-\mu}{\sigma}) \quad = \quad \mathbb{P}(Z > \frac{x-\mu}{\sigma}) \quad = \quad \Phi^c(z)$$

where $z = \frac{x-\mu}{\sigma}$ and $\Phi^c(x) = \int_x^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz = 1 - \Phi(z)$. We can then look up $\Phi^c(z)$ in a Normal table or compute it in e.g. Matlab, Excel, Python or Mathematica.

- The **moment generating function** (mgf) of a general random variable X is defined as

$$M(p) \quad := \quad \mathbb{E}(e^{pX}) \quad = \quad \int_{-\infty}^{\infty} e^{px} f_X(x)dx$$

for $p \in \mathbb{R}$, where I am using (5) for the final equality.

- If $X \ge 0$ and $\mathbb{E}(e^{pX}) < \infty$ for p in some open interval $I = (p_*, p^*)$ which contains zero, then from a result called **Fubini's theorem**, we can interchange the expectation and the integral if the integrand is non-negative, so

$$\int_0^p \mathbb{E}(Xe^{qX})dq \quad = \quad \mathbb{E}(\int_0^p Xe^{qX}dq) \quad = \quad \mathbb{E}(e^{qX}|_{q=0}^{q=p}) \quad = \quad \mathbb{E}(e^{pX} - 1)\,.$$

Hence differentiating the left hand and right hand expressions and again using the fundamental theorem of calculus for the left expression, we see that

$$\frac{d}{dp}\int_0^p \mathbb{E}(Xe^{qX})dq \quad = \quad \mathbb{E}(Xe^{pX}) \quad = \quad \frac{d}{dp}\mathbb{E}(e^{pX}-1) \quad = \quad \frac{d}{dp}\mathbb{E}(e^{pX})\,.$$

Then setting $p=0$ we see that

$$\frac{d}{dp}\mathbb{E}(e^{pX})|_{p=0} \quad = \quad \mathbb{E}(X)\,.$$

By repeating this procedure, we can show that

$$M^{(n)}(p) \quad := \quad \frac{d^n}{dp^n}\mathbb{E}(e^{pX}) \quad = \quad \mathbb{E}(X^n e^{pX})$$

so $M^{(n)}(0) = \mathbb{E}(X^n)$, which is the nth moment of X. This is why M is called the moment generating function.

Note: To extend to a real-valued random variable X which may be negative, we have to check that $\int_0^p \mathbb{E}(|Xe^{qX}|)dq < \infty$ to justify use of Fubini (we will not do this here).

- If $X_1, ..., X_n$ is a sequence of i.i.d. random variables, then

$$\mathbb{E}(e^{p(X_1+...+X_n)}) \quad = \quad \mathbb{E}(e^{pX_1})...\mathbb{E}(e^{pX_n}) \quad = \quad \mathbb{E}(e^{pX_1})^n\,.$$

- If $X \sim N(\mu,\sigma^2)$, then using (5) we can show that the mgf of X is given by

$$\mathbb{E}(e^{pX}) \quad = \quad \mathbb{E}(e^{p(\sigma Z+\mu)}) \quad = \quad \int_{-\infty}^{\infty} e^{p(\sigma z+\mu)}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}z^2}dz \quad = \quad ... \quad = \quad e^{\mu p+\frac{1}{2}\sigma^2 p^2}$$

where $Z \sim N(0,1)$, and we have used that $X \sim \sigma Z+\mu$, since we already know that $(X-\mu)/\sigma = Z$, and I have skipped some tedious details to get to the final expression.

Example questions

1. For a **continuous random variable** X with density $f_X(x)$, prove that $\mathbb{E}(1_A(X)) = \mathbb{P}(X \in A)$ for any set A, where $1_A(X) = 1$ if $X \in A$ and zero otherwise. $1_A(X)$ is known as the **indicator function**, and will be used many times in the course, and we sometimes abbreviate this to just 1_A.

Solution.

$$\mathbb{E}(1_A(X)) \quad = \quad \int_{-\infty}^{\infty} 1_A(x)f_X(x)dx \quad = \quad \int_A f_X(x)dx \quad = \quad \mathbb{P}(X\in A)\,.$$

2. Let W_t be a random function such that $W_t \sim N(0,t)$ (**Brownian motion** has this nice property as we shall see next week). Compute $\mathbb{P}(W_t > x)$.

Solution.

$$\mathbb{P}(W_t > x) \quad = \quad \mathbb{P}(\frac{W_t-0}{\sqrt{t}} > \frac{x-0}{\sqrt{t}}) \quad = \quad \mathbb{P}(Z > \frac{x}{\sqrt{t}}) \quad = \quad \Phi^c(\frac{x}{\sqrt{t}})$$

since $Z = (W_t - 0)/\sqrt{t}$ is a standard $N(0,1)$ random variable. The general rule here is: **do to one side what you do to the other side**.

3. **Simulating random variables with a given distribution**. Let X be a random variable with a continuous strictly increasing distribution function $F_X(x)$. What is the distribution of $F_X^{-1}(U)$, where U is a standard Uniform random variable on $[0,1]$?

Solution.

$$\mathbb{P}(F_X^{-1}(U) \le x) \quad = \quad \mathbb{P}(U \le F_X(x)) \quad = \quad F_X(x) \tag{8}$$

since $F_X(F_X^{-1}(x)) = x$, i.e. $F_X^{-1}(U) \sim X$, where we have used that $\mathbb{P}(U \le x) = x$ from (3). So the conclusion here is that $F_X^{-1}(U)$ has the same distribution as X. This is how we typically generate a random variable with a given distribution in practice on a computer.

Similarly, we can compute the distribution function of $F_X(X)$ as

$$\mathbb{P}(F_X(X) \le x) \quad = \quad \mathbb{P}(X \le F_X^{-1}(x)) \quad = \quad F_X(F_X^{-1}(x)) \quad = \quad x \tag{9}$$

so we see that $F_X(X)$ has the same distribution function as a $U[0,1]$ random variable (see Eq (3) above), so $F_X(X) \sim U[0,1]$.

More advanced concepts

- **Strong law of large numbers** (SLLN). Let $X_1, X_2, ...$ denote an infinite sequence of independent identically distributed (i.i.d) random variables with $\mathbb{E}(X_i) = \mu$ and $\mathbb{E}(|X_i|) < \infty$ for all $i = 1, 2, ...$ Let $S_n = \sum_{i=1}^{n} X_i$. Then the SLLN says that

$$\mathbb{P}(\lim_{n\to\infty} \frac{S_n}{n} = \mu) \quad = \quad 1$$

i.e. the **sample average** $\frac{S_n}{n}$ tends to the **true expectation** μ of X_i as $n \to \infty$, as we would intuitively. This result is the cornerstone of **Monte Carlo simulation** (and why bad gamblers lose in the long run), namely that if μ is unknown, we can estimate μ with $\hat{\mu}_n = \frac{S_n}{n}$ for n large (see e.g. chapter 7 of David Williams book "Probability with Martingales" for proof).

Transformations of a 1-d random variable. Let X be a continuous random variable with density $f_X(x)$, and let $Y = g(X)$, where g is differentiable and strictly increasing, which implies that $g'(x) > 0$ for all $x \in \mathbb{R}$, which further implies that g has a unique inverse $h = g^{-1}$ such that $g^{-1}(g(x)) = x$. Then the distribution function of Y is

$$\mathbb{P}(Y \le y) \quad = \quad \mathbb{P}(X \le g^{-1}(y)) \quad = \quad F_X(g^{-1}(y)) \quad = \quad F_X(h(y))\,.$$

Differentiating with respect to y, we obtain the density of Y:

$$f_Y(y) \quad = \quad \frac{d}{dy}\mathbb{P}(Y \le y) \quad = \quad h'(y)F'_X(h(y)) \quad = \quad h'(y)f_X(h(y))$$

using the chain rule.

- **Generating correlated Normal random variables**. Let Z_1, Z_2 be two i.i.d $N(0,1)$ random variables and

$$\begin{aligned} X &= Z_1 \\ Y &= \rho Z_1 + \sqrt{1-\rho^2}\,Z_2 \end{aligned} \tag{10}$$

for some ρ with $-1 \le \rho \le 1$. This definition implies that $\mathbb{E}(Y|X) = \rho Z_1 = \rho X$ since $\mathbb{E}(Z_2) = 0$. Then it turns out that Y is also a standard $N(0,1)$ random variable, and X and Y have (in general non-zero) **correlation** ρ, i.e.

$$\mathrm{Corr}(X,Y) \quad := \quad \frac{\mathrm{Cov}(X,Y)}{\sigma_X \sigma_Y} \quad = \quad \frac{\mathbb{E}((X-\mu_X)(Y-\mu_Y))}{\sigma_X \sigma_Y} \quad = \quad \mathbb{E}(XY) \quad = \quad \mathbb{E}(Z_1(\rho Z_1) + \sqrt{1-\rho^2}\,Z_2) \quad = \quad \rho \tag{11}$$

where $\mu_X = \mathbb{E}(X) = 0$, $\mu_Y = \mathbb{E}(Y) = 0$, and σ_X, σ_Y denote the **standard deviations** of X and Y (which in this case are also both 1 since Z_1 and Z_2 are standard Normal RVs).

Clearly $\sigma_X = 1$. To prove that $\sigma_Y = 1$, we note that since Z_1 and Z_2 are independent

$$\begin{aligned} \mathrm{Var}(Y) \quad = \quad \mathrm{Var}(\rho Z_1 + \sqrt{1-\rho^2}\,Z_2) \quad = \quad \mathrm{Var}(\rho Z_1) + \mathrm{Var}(\sqrt{1-\rho^2}\,Z_2) \quad &= \quad \rho^2 \mathrm{Var}(Z_1) + (1-\rho^2)\mathrm{Var}(Z_2) \\ &= \quad \rho^2 + 1 - \rho^2 \quad = \quad 1\,. \end{aligned}$$

To verify the correlation of X and Y, we see that

$$\mathbb{E}(XY) \quad = \quad \mathbb{E}(Z_1(\rho Z_1 + \sqrt{1-\rho^2}\,Z_2)) \quad = \quad \rho\mathbb{E}(Z_1^2) + \sqrt{1-\rho^2}\,\mathbb{E}(Z_1)\mathbb{E}(Z_2) \quad = \quad \rho\mathbb{E}(Z_1^2) \quad = \quad \rho$$

since $\mathrm{Var}(Z_1) = \mathbb{E}(Z_1^2) - \mathbb{E}(Z_1)^2 = \mathbb{E}(Z_1^2)$ as $\mathbb{E}(Z_1) = 0$, which verifies (11).

To show that Y is a normal random variable, we appeal to the general result that any linear combination of two independent Normal random variables $X_1 \sim N(\mu_1, \sigma_1^2)$ and $X_2 \sim N(\mu_2, \sigma_2^2)$ is Normal, since

$$\mathbb{E}(e^{p(X_1+X_2)}) \quad = \quad \mathbb{E}(e^{pX_1})\mathbb{E}(e^{pX_2}) \quad = \quad e^{\mu_1 p + \frac{1}{2}\sigma_1^2 p^2} e^{\mu_2 p + \frac{1}{2}\sigma_2^2 p^2} \quad = \quad e^{(\mu_1+\mu_2)p + \frac{1}{2}(\sigma_1^2+\sigma_2^2)p^2}$$

which is the the m.g.f. of a $N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$ random variable.