Introduction to Mathematical Finance

George Voutsadakis[1]

[1]Mathematics and Computer Science
Lake Superior State University

LSSU Math 500

Subsection 1

Brownian Motion

Brownian Motion

- Consider a collection of random variables $X(t)$, $t \geq 0$.
- Imagine we are observing some process as it evolves over time.
- The index parameter t represents time.
- $X(t)$ is interpreted as the state of the process at time t.

Definition

The collection of random variables $X(t)$, $t \geq 0$ is said to be a **Brownian motion** with drift parameter μ and variance parameter σ^2 if the following hold:

(a) $X(0)$ is a given constant.

(b) For $y, t > 0$, the random variable $X(y+t) - X(y)$:

- Is independent of the process values up to time y;
- Has a normal distribution with mean μt and variance $t\sigma^2$.

Consequence

- Assumption (b) says that, for any history of the process up to the present time y, the change in the value of the process over the next t time units is a normal random variable with mean μt and variance $t\sigma^2$.
- Note that any future value $X(y+t)$ is equal to the present value $X(y)$ plus the change in value $X(y+t)-X(y)$.
- Thus, the assumption implies that it is only the present value of the process, and not any past values, that determines probabilities about future values.

Continuity Property

- An important property of Brownian motion is that $X(t)$ will, with probability 1, be a continuous function of t.
- Althought this is a mathematically deep result, it is not difficult to see why it might be true.
- To prove that $X(t)$ is continuous, we must show that

$$\lim_{h\to 0}(X(t+h)-X(t))=0.$$

- Bu the random variable $X(t+h)-X(t)$ has mean μh and variance $h\sigma^2$.
- So it converges as $h\to 0$ to a random variable with mean 0 and variance 0.
- That is, it converges to the constant 0, thus arguing for continuity.

Nowhere Differentiability

- We saw that $X(t)$ is, with probability 1, a continuous function of t.
- However, it possesses the property of being nowhere differentiable.
- To see why this might be the case, note that

$$\frac{X(t+h)-X(t)}{h}$$

  has mean μ and variance $\frac{\sigma^2}{h}$.
- The variance of this ratio is converging to infinity as $h \to 0$.
- So it is not surprising that the ratio does not converge.

Subsection 2

Brownian Motion as a Limit of Simpler Models

Brownian Motion as a Limit of Simpler Models

- Let Δ be a small increment of time.
- Set $p = \frac{1}{2}(1 + \frac{\mu}{\sigma}\sqrt{\Delta})$.
- Consider a process such that, every Δ time units, the value of the process behaves in either of two ways:
  - It increases by the amount $\sigma\sqrt{\Delta}$ with probability p;
  - It decreases by the amount $\sigma\sqrt{\Delta}$ with probability $1 - p$.

  Successive changes in value are independent.
- Take Δ smaller and smaller.
  - The changes occur more and more frequently;
  - The change amounts become smaller and smaller.
- The process becomes a Brownian motion with drift parameter μ and variance parameter σ^2.
- Consequently, Brownian motion can be approximated by a relatively simple process that either increases or decreases by a fixed amount at regularly specified times.

Verification

- Let

$$X_i = \begin{cases} 1, & \text{if the change at time } i\Delta \text{ is an increase} \\ -1, & \text{if the change at time } i\Delta \text{ is a decrease} \end{cases}.$$

- Let $X(0)$ be the process value at time 0.
- Then its value after n changes is

$$X(n\Delta) = X(0) + \sigma\sqrt{\Delta}(X_1 + \cdots + X_n).$$

- By time t, there would have been $n = \frac{t}{\Delta}$ changes.
- This gives

$$X(t) - X(0) = \sigma\sqrt{\Delta}\sum_{i=1}^{t/\Delta} X_i.$$

Verification (Cont'd)

- Note that:
  - The X_i, $i = 1, \ldots, \frac{t}{\Delta}$, are independent;
  - As Δ goes to 0 there are more and more terms in $\sum_{i=1}^{t/\Delta} X_i$.
- Thus, the Central Limit Theorem suggests that this sum converges to a normal random variable.
- Consequently, as Δ goes to 0, the process value at time t becomes a normal random variable.
- To compute its mean and variance, note that

$$\begin{aligned} E[X_i] &= 1(p) - 1(1-p) = 2p - 1 = \frac{\mu}{\sigma}\sqrt{\Delta}; \\ \mathrm{Var}(X_i) &= E[X_i^2] - (E[X_i])^2 = 1 - (2p-1)^2. \end{aligned}$$

Verification (Cont'd)

- Hence,

$$\begin{aligned} E[X(t)-X(0)] &= E\left[\sigma\sqrt{\Delta}\sum_{i=1}^{t/\Delta}X_i\right] \\ &= \sigma\sqrt{\Delta}\sum_{i=1}^{t/\Delta}E[X_i] \\ &= \sigma\sqrt{\Delta}\frac{t}{\Delta}\frac{\mu}{\sigma}\sqrt{\Delta} \\ &= \mu t. \end{aligned}$$

- Furthermore,

$$\begin{aligned} \text{Var}(X(t)-X(0)) &= \text{Var}\left(\sigma\sqrt{\Delta}\sum_{i=1}^{t/\Delta}X_i\right) \\ &= \sigma^2\Delta\sum_{i=1}^{t/\Delta}\text{Var}(X_i) \\ &= \sigma^2 t[1-(2p-1)^2]. \end{aligned}$$

- We have $p \to \frac{1}{2}$ as $\Delta \to 0$.
- So $\text{Var}(X(t)-X(0)) \to t\sigma^2$ as $\Delta \to 0$.

Verification (Cont'd)

- Consequently, as Δ gets smaller and smaller, $X(t) - X(0)$ converges to a normal random variable with mean μ and variance σ^2.
- In addition:
  - Successive process changes are independent;
  - Each has the same probability of being an increase.
- Hence, $X(y+t) - X(y)$ has the same distribution as does $X(t) - X(0)$.
- Moreover, it is independent of earlier process changes before time y.
- Hence, as Δ goes to 0, the collection of process values over time becomes a Brownian motion process with drift parameter μ and variance parameter σ^2.

Independence of the Drift Parameter

Theorem

Given that $X(t) = x$, the conditional probability law of the collection of prices $X(y)$, $0 \leq y \leq t$, is the same for all values of μ.

- Let $s = X(0)$ be the price at time 0.

  Consider the approximating model where the price changes every Δ time units by an amount equal, in absolute value, to $c \equiv \sigma\sqrt{\Delta}$.

  Note that c does not depend on μ.

  By time t, there would have been $\frac{t}{\Delta}$ changes.

  Suppose the price has increased from time 0 to time t by $x - s$.

  It follows that, of the $\frac{t}{\Delta}$ changes, there have been:

  - A total of $\frac{t}{2\Delta} + \frac{x-s}{2c}$ positive changes;
  - A total of $\frac{t}{2\Delta} - \frac{x-s}{2c}$ negative changes.

  In fact $(\frac{t}{2\Delta} + \frac{x-s}{2c})c - (\frac{t}{2\Delta} - \frac{x-s}{2c})c = \frac{x-s}{c}c = x - s$.

Independence of the Drift Parameter (Cont'd)

- Each change is, independently, a positive change with the same probability p.

  So, conditional on there being a total of $\frac{t}{2\Delta} + \frac{x-s}{2c}$ positive changes out of the first $\frac{t}{\Delta}$ changes, all possible choices of the changes that were positive are equally likely.

  [That is, if a coin having probability p is flipped m times, then, given that k heads resulted, the subset of trials that resulted in heads is equally likely to be any of the $\binom{m}{k}$ subsets of size k.]

  Although p depends on μ, the conditional distribution of the history of prices up to time t, given that $X(t) = x$, does not depend on μ.

  It depends on σ, because c, the size of a change, depends on σ.

  So, if σ changed, then so would the number of the $\frac{t}{\Delta}$ changes that would have had to be positive for $S(t)$ to equal x.

  Letting Δ go to 0 now completes the proof.

Subsection 3

Geometric Brownian Motion

Geometric Brownian Motion

Definition

Let $X(t)$, $t \geq 0$ be a Brownian motion process with drift parameter μ and variance parameter σ^2, and let

$$S(t) = e^{X(t)}, \quad t \geq 0.$$

The process $S(t)$, $t \geq 0$, is said to be be a **geometric Brownian motion** process with drift parameter μ and variance parameter σ^2.

Geometric Brownian Motion Features

- Let $S(t)$, $t \geq 0$ be a geometric Brownian motion process with drift parameter μ and variance parameter σ^2.
- We have, by definition, that $\log(S(t))$, $t \geq 0$, is a Brownian motion.
- Moreover,

$$\log(S(t+y)) - \log(S(y)) = \log\left(\frac{S(t+y)}{S(y)}\right).$$

- Thus, by definition, for all $y, t > 0$, the quantity $\log(\frac{S(t+y)}{S(y)})$:
  - Is independent of the process values up to time y;
  - Has a normal distribution with mean μt and variance $t\sigma^2$.

Advantages for Modeling Prices of Securities

- When used to model the price of a security over time, the geometric Brownian motion process has some advantages over the Brownian motion process:
  - First, it is the logarithm of the stock's price, assumed to be a normal random variable.
    So the model does not allow for negative stock prices.
  - Second, it consists of ratios, rather than differences, of prices separated by a fixed amount of time that have the same distribution.
    So it makes what many feel is the more reasonable assumption of a percentage, rather than absolute, change in price whose probabilities do not depend on the current price.

Remarks

- When geometric Brownian motion is used to model the price of a security over time, it is common to call σ the **volatility parameter**.
- If $S(0) = s$, then we can write

$$S(t) = se^{X(t)}, \quad t \geq 0,$$

where $X(t)$, $t \geq 0$, is a Brownian motion process with $X(0) = 0$.

- If X is a normal random variable, then it can be shown that

$$E[e^X] = \exp\left\{E[X] + \frac{\mathrm{Var}(X)}{2}\right\}.$$

Remarks (Cont'd)

- Assume, now, that $S(t)$, $t \geq 0$, is a geometric Brownian motion process with:
  - Drift μ;
  - Volatility σ;
  - $S(0) = s$.
- Then
$$E[S(t)] = se^{\mu t + \frac{t\sigma^2}{2}} = se^{(\mu + \frac{\sigma^2}{2})t}.$$
- Thus, under geometric Brownian motion, the expected price of a security grows at rate $\mu + \frac{\sigma^2}{2}$.
- $\mu + \frac{\sigma^2}{2}$ is often called the **rate** of the geometric Brownian motion.
- Consequently, a geometric Brownian motion with rate parameter μ_r and volatility σ would have drift parameter $\mu_r - \frac{\sigma^2}{2}$.

Geometric Brownian Motion as a Limit

- Let $S(t)$, $t \geq 0$ be a geometric Brownian motion process with drift parameter μ and volatility parameter σ.
- Because $X(t) = \log(S(t))$, $t \geq 0$, is Brownian motion, we can use its approximating process to obtain an approximating process for geometric Brownian motion.
- We have
$$\frac{S(y+\Delta)}{S(y)} = e^{X(y+\Delta)-X(y)}.$$
- It follows that
$$S(y+\Delta) = S(y)e^{X(y+\Delta)-X(y)}.$$

Geometric Brownian Motion as a Limit (Cont'd)

- Set
$$u = e^{\sigma\sqrt{\Delta}}, \quad d = e^{-\sigma\sqrt{\Delta}}, \quad p = \frac{1}{2}\left(1 + \frac{\mu}{\sigma}\sqrt{\Delta}\right).$$
- We can approximate geometric Brownian motion by a model for the price of a security in which:
  - Price changes occur only at times that are integral multiples of Δ;
  - Price changes occur in one of two possible ways:
    - The price is multiplied by the factor u with probability p;
    - The price is multiplied by the factor d with probability $1 - p$.
- As Δ goes to 0, this model becomes geometric Brownian motion.
- Consequently, geometric Brownian motion can be approximated by a relatively simple process that goes either up or down by fixed factors at regularly spaced times.

Subsection 4

The Maximum Variable

The Maximum Variable

- Let $X(v)$, $v \geq 0$, be a Brownian motion process with drift parameter μ and variance parameter σ^2.
- Suppose that $X(0) = 0$, so that the process starts at state 0.
- Now, define
$$M(t) = \max_{0 \leq v \leq t} X(v)$$
to be the maximal value of the Brownian motion up to time t.
- We derive the conditional distribution of $M(t)$ given the value of $X(t)$.
- We then use this to derive the unconditional distribution of $M(t)$.

Conditional Distribution

Theorem

For $y > x$,

$$P(M(t) \geq y | X(t) = x) = e^{-2y(y-x)/t\sigma^2}, \quad y \geq 0.$$

- Because $X(0) = 0$, it follows that $M(t) \geq 0$.
  So the result is true when $y = 0$ (both sides are equal to 1).
  Suppose that $y > 0$.
  By a previous theorem, $P(M(t) \geq y | X(t) = x)$ does not depend on μ.
  So let us take $\mu = 0$.
  Let T_y denote the first time the Brownian motion reaches y.
  Brownian motion is continuous.
  So before the process can exceed y it must pass through y.
  So the event $M(t) \geq y$ is equivalent to $T_y \leq t$.

Conditional Distribution (Cont'd)

- Let h be a small positive number for which $y > x + h$. Then

$$\begin{aligned} & P(M(t) \geq y, x \leq X(t) \leq x + h) \\ &= P(T_y \leq t, x \leq X(t) \leq x + h) \\ &= P(x \leq X(t) \leq x + h | T_y \leq t) P(T_y \leq t). \end{aligned}$$

Now, given $T_y \leq t$, the event $x \leq X(t) \leq x + h$ will occur if, after hitting y, the additional amount $X(t) - X(T_y) = X(t) - y$ by which the process changes by time t is between $x - y$ and $x + h - y$.

The distribution of this change is symmetric about 0 ($\mu = 0$).

The distribution of a normal variable is symmetric about its mean.

So the additional change is just as likely to be between $-(x + h - y)$ and $-(x - y)$ as it is to be between $x - y$ and $x + h - y$.

Conditional Distribution (Cont'd)

- Consequently,

$$\begin{aligned}&P(x \leq X(t) \leq x+h | T_y \leq t)\\&= P(x-y \leq X(t)-y \leq x+h-y | T_y \leq t)\\&= P(-(x+h-y) \leq X(t)-y \leq -(x-y) | T_y \leq t).\end{aligned}$$

Combining the preceding equalities gives

$$\begin{aligned}&P(M(t) \geq y, x \leq X(t) \leq x+h)\\&= P(2y-x-h \leq X(t) \leq 2y-x | T_y \leq t) P(T_y \leq t)\\&= P(2y-x-h \leq X(t) \leq 2y-x, T_y \leq t)\\&= P(2y-x-h \leq X(t) \leq 2y-x).\end{aligned}$$

By hypothesis, $y > x+h$. This implies that $2y-x-h > y$.
So, by continuity, $2y-x-h \leq X(t)$ implies $T_y \leq t$.

Conditional Distribution (Cont'd)

- Now we have

$$P(M(t) \geq y | x \leq X(t) \leq x+h) = \frac{P(2y-x-h \leq X(t) \leq 2y-x)}{P(x \leq X(t) \leq x+h)}$$

$$\approx \frac{f_{X(t)}(2y-x)h}{f_{X(t)}(x)h} \text{ (for } h \text{ small)},$$

where $f_{X(t)}$, the density function of $X(t)$, is the density of a normal random variable with mean 0 and variance $t\sigma^2$.

On letting $h \to 0$ in the preceding, we obtain that

$$\begin{aligned} P(M(t) \geq y | X(t) = x) &= \frac{f_{X(t)}(2y-x)}{f_{X(t)}(x)} \\ &= \frac{e^{-(2y-x)^2/2t\sigma^2}}{e^{-x^2/2t\sigma^2}} \\ &= e^{-2y(y-x)/t\sigma^2}. \end{aligned}$$

Distribution

- With Z being a standard normal distribution function, let
$$\overline{\Phi}(x) = 1 - \Phi(x) = P(Z > x).$$

Corollary

For $y \geq 0$
$$P(M(t) \geq y) = e^{2y\mu/\sigma^2}\overline{\Phi}\left(\frac{\mu t + y}{\sigma\sqrt{t}}\right) + \overline{\Phi}\left(\frac{y - \mu t}{\sigma\sqrt{t}}\right).$$

- Conditioning on $X(t)$, and using the theorem gives
$$\begin{aligned} P(M(t) \geq y) &= \int_{-\infty}^{\infty} P(M(t) \geq y | X(t) = x) f_{X(t)}(x) dx \\ &= \int_{-\infty}^{y} P(M(t) \geq y | X(t) = x) f_{X(t)}(x) dx \\ &\quad + \int_{y}^{\infty} P(M(t) \geq y | X(t) = x) f_{X(t)}(x) dx \\ &= \int_{-\infty}^{y} e^{-2y(y-x)/t\sigma^2} f_{X(t)}(x) dx + \int_{y}^{\infty} f_{X(t)}(x) dx. \end{aligned}$$

Distribution (Cont'd)

- $f_{X(t)}$ is the density function of a normal random variable with mean μt and variance $t\sigma^2$:

$$\begin{aligned} & P(M(t) \geq y) \\ & = \int_{-\infty}^{y} e^{-2y(y-x)/t\sigma^2} \frac{1}{\sqrt{2\pi t\sigma^2}} e^{-(x-\mu t)^2/2t\sigma^2} dx + P(X(t) > y) \\ & = \frac{1}{\sqrt{2\pi t}\sigma} e^{-2y^2/t\sigma^2} e^{-\mu^2 t^2/2t\sigma^2} \\ & \times \int_{-\infty}^{y} \exp\{-\frac{1}{2t\sigma^2}(x^2 - 2\mu t x - 4yx)\} dx + P(X(t) > y) \\ & = \frac{1}{\sqrt{2\pi t}\sigma} e^{-(4y^2+\mu^2 t^2)/2t\sigma^2} \\ & \times \int_{-\infty}^{y} \exp\{-\frac{1}{2t\sigma^2}(x^2 - 2x(\mu t + 2y))\} dx + P(X(t) > y). \end{aligned}$$

Now, $x^2 - 2x(\mu t + 2y) = (x - (\mu t + 2y))^2 - (\mu t + 2y)^2$. So

$$\begin{aligned} P(M(t) \geq y) &= e^{-(4y^2+\mu^2 t^2-(\mu t+2y)^2)/2t\sigma^2} \frac{1}{\sqrt{2\pi t}\sigma} \\ & \times \int_{-\infty}^{y} e^{-(x-\mu t-2y)^2/2t\sigma^2} dx + P(X(t) > y). \end{aligned}$$

Distribution (Cont'd)

- We got

$$P(M(t) \geq y) = e^{-(4y^2+\mu^2t^2-(\mu t+2y)^2)/2t\sigma^2} \frac{1}{\sqrt{2\pi t}\sigma} \times \int_{-\infty}^{y} e^{-(x-\mu t-2y)^2/2t\sigma^2} dx + P(X(t) > y).$$

Let Z be a standard normal random variable.
Change variables $w = \frac{x-\mu t-2y}{\sigma\sqrt{t}}$.
Then $dx = \sigma\sqrt{t}dw$ and

$$\begin{aligned} P(M(t) \geq y) &= e^{2y\mu/\sigma^2} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{-\mu t-y}{\sigma\sqrt{t}}} e^{-w^2/2} dw \\ &\quad + P\left(\frac{X(t)-\mu t}{\sigma\sqrt{t}} > \frac{y-\mu t}{\sigma\sqrt{t}}\right) \\ &= e^{2y\mu/\sigma^2} P\left(Z < \frac{-\mu-y}{\sigma\sqrt{t}}\right) + P\left(Z > \frac{y-\mu t}{\sigma\sqrt{t}}\right) \\ &= e^{2y\mu/\sigma^2} P\left(Z > \frac{\mu t+y}{\sigma\sqrt{t}}\right) + P\left(Z > \frac{y-\mu t}{\sigma\sqrt{t}}\right). \end{aligned}$$

Distribution of Hitting Time

- In the proof of the theorem we let T_y denote the first time the Brownian motion is equal to y.
- That is,

$$T_y = \begin{cases} \infty, & \text{if } X(t) \neq y \text{ for all } t \geq 0 \\ \min(t : X(t) = y), & \text{otherwise} \end{cases}$$

- In addition, it follows from the continuity of Brownian motion paths that, for $y > 0$, the process would have hit y by time t if and only if the maximum of the process by time t is at least y. That is,

$$T_y \leq t \quad \Leftrightarrow \quad M(t) \geq y.$$

- Hence, the corollary yields that

$$P(T_y \leq t) = e^{2y\mu/\sigma^2}\overline{\Phi}\left(\frac{y+\mu t}{\sigma\sqrt{t}}\right) + \overline{\Phi}\left(\frac{y-\mu t}{\sigma\sqrt{t}}\right).$$

The Minimum Variable

- Let $M_{\mu,\sigma}(t)$ denote a random variable having the distribution of the maximum value up to time t of a Brownian motion process that starts at 0 and has drift parameter μ and variance parameter σ^2.
- The distribution of $M_{\mu,\sigma}(t)$ is given by the corollary.
- Suppose we want the distribution of

$$M^*(t) = \min_{0 \le v \le t} X(v).$$

- The process $-X(v)$, $v \geq 0$, is a Brownian motion with drift parameter $-\mu$ and variance parameter σ^2. So, for $y > 0$,

$$\begin{aligned} P(M^*(t) \le -y) &= P(\min_{0 \le v \le t} X(v) \le -y) \\ &= P(-\max_{0 \le v \le t} -X(v) \le -y) \\ &= P(\max_{0 \le v \le t} -X(v) \ge y) \\ &= P(M_{-\mu,\sigma}(t) \ge y) \\ &= e^{-2y\mu/\sigma^2}\overline{\Phi}(\tfrac{-\mu t + y}{\sigma\sqrt{t}}) + \overline{\Phi}(\tfrac{y+\mu t}{\sigma\sqrt{t}}). \end{aligned}$$

Subsection 5

The Cameron-Martin Theorem

Notation

- Consider a Brownian motion process with variance parameter σ^2.
- We use the notation

$$E_\mu$$

  to denote taking expectations under the assumption that the drift parameter is μ.
- E.g.,

$$E_0$$

  signifies that the expectation is taken under the assumption that the drift parameter is 0.

The Cameron-Martin Theorem

Theorem

Let W be a random variable whose value is determined by the history of the Brownian motion up to time t. That is, the value of W is determined by a knowledge of the values of $X(s)$, $0 \leq s \leq t$. Then,

$$E_\mu[W] = e^{-\mu^2 t/2\sigma^2} E_0[W e^{\mu X(t)/\sigma^2}].$$

- Condition on $X(t)$, which is normal with mean μt and variance $t\sigma^2$. Take into account that, given $X(t) = x$, the conditional distribution of the process W up to time t is the same for all values μ.

The Cameron-Martin Theorem (Cont'd)

- We obtain

$$\begin{aligned} E_\mu[W] &= \int_{-\infty}^{\infty} E_\mu[W|X(t)=x]\frac{1}{\sqrt{2\pi t\sigma^2}}e^{-(x-\mu t)^2/2t\sigma^2}dx \\ &= \int_{-\infty}^{\infty} E_0[W|X(t)=x]\frac{1}{\sqrt{2\pi t\sigma^2}}e^{-(x-\mu t)^2/2t\sigma^2}dx \\ &= \int_{-\infty}^{\infty} E_0[W|X(t)=x]\frac{1}{\sqrt{2\pi t\sigma^2}}e^{-x^2/2t\sigma^2}e^{(2\mu x-\mu^2 t)/2\sigma^2}dx. \end{aligned}$$

Define

$$Y = e^{-\mu^2 t/2\sigma^2}e^{\mu X(t)/\sigma^2} = e^{(2\mu X(t)-\mu^2 t)/2\sigma^2}.$$

Then

$$E_0[WY] = \int_{-\infty}^{\infty} E_0[WY|X(t)=x]\frac{1}{\sqrt{2\pi t\sigma^2}}e^{-x^2/2t\sigma^2}dx.$$

The Cameron-Martin Theorem (Cont'd)

- We have

$$E_0[WY] = \int_{-\infty}^{\infty} E_0[WY|X(t)=x]\frac{1}{\sqrt{2\pi t\sigma^2}}e^{-x^2/2t\sigma^2}dx.$$

But, given that $X(t)=x$, the random variable Y is equal to the constant $e^{(2\mu x-\mu^2 t)/2\sigma^2}$.

So the preceding yields

$$\begin{aligned} E_0[WY] &= \int_{-\infty}^{\infty} e^{(2\mu x-\mu^2 t)/2\sigma^2}E_0[W|X(t)=x]\frac{1}{\sqrt{2\pi t\sigma^2}}e^{-x^2/2t\sigma^2}dx \\ &= E_\mu[W], \end{aligned}$$

where the final equality used the equality of the preceding slide.