Normal Random Variables and Probability

An Undergraduate Introduction to Financial Mathematics

J. Robert Buchanan

2015

Discrete *vs.* Continuous Random Variables

Think about the probability of selecting X from the interval $[0, 1]$ when

- $X \in \{0, 1\}$

Discrete *vs.* Continuous Random Variables

Think about the probability of selecting X from the interval $[0, 1]$ when

- $X \in \{0, 1\}$
- $X \in \{k/10 \, : \, k = 0, 1, \ldots, 10\}$

Think about the probability of selecting X from the interval $[0, 1]$ when

- $X \in \{0, 1\}$
- $X \in \{k/10 \,:\, k = 0, 1, \ldots, 10\}$
- $X \in \{k/n \,:\, k = 0, 1, \ldots, n\}$ and $n \in \mathbb{N}$

Think about the probability of selecting X from the interval $[0, 1]$ when

- $X \in \{0, 1\}$
- $X \in \{k/10 \, : \, k = 0, 1, \dots, 10\}$
- $X \in \{k/n \, : \, k = 0, 1, \dots, n\}$ and $n \in \mathbb{N}$

Question: what happens to the last probability as $n \to \infty$?

Continuous Random Variables

Definition

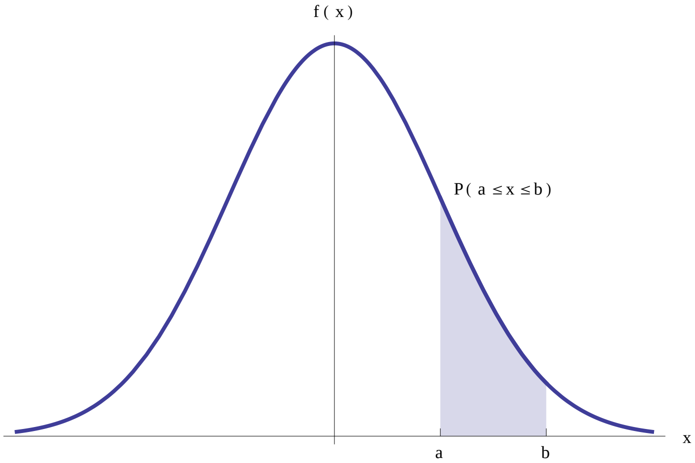
A random variable X has a **continuous distribution** (or **probability distribution function** or **probability density function**) if there exists a non-negative function $f : \mathbb{R} \to \mathbb{R}$ such that for an interval $[a, b]$ the

$$\mathbb{P}\left(a \leq X \leq b\right) = \int_a^b f(x)\,dx.$$

The function f must, in addition to satisfying $f(x) \geq 0$, have the following property,

$$\int_{-\infty}^{\infty} f(x)\,dx = 1.$$

Area Under the PDF



Remark: the area under the curve may be interpreted as probability.

Uniformly Distributed Continuous Random Variables

Definition

A continuous random variable X is **uniformly distributed** in the interval $[a, b]$ (with $b > a$) if the probability that X belongs to any subinterval of $[a, b]$ is equal to the length of the subinterval divided by $b - a$.

Uniformly Distributed Continuous Random Variables

Definition

A continuous random variable X is **uniformly distributed** in the interval $[a, b]$ (with $b > a$) if the probability that X belongs to any subinterval of $[a, b]$ is equal to the length of the subinterval divided by $b - a$.

Question: Assuming the PDF vanishes outside of $[a, b]$ and is constant on $[a, b]$, what is the PDF?

Definition

A continuous random variable X is **uniformly distributed** in the interval $[a, b]$ (with $b > a$) if the probability that X belongs to any subinterval of $[a, b]$ is equal to the length of the subinterval divided by $b - a$.

Question: Assuming the PDF vanishes outside of $[a, b]$ and is constant on $[a, b]$, what is the PDF?

Answer: $f(x) = \begin{cases} \frac{1}{b-a} & \text{if } a \leq x \leq b, \\ 0 & \text{otherwise.} \end{cases}$

Example (1 of 2)

Random variable X is continuously and uniformly randomly distributed in the interval $[-5, 5]$. Find the probability that $-1 \leq X \leq 2$.

Example (1 of 2)

Random variable X is continuously and uniformly randomly distributed in the interval $[-5, 5]$. Find the probability that $-1 \leq X \leq 2$.

$$\mathbb{P}\left(-1 \leq X \leq 2\right) = \frac{2 - (-1)}{5 - (-5)} = \frac{3}{10}$$

Random variable X is continuously and uniformly randomly distributed in the interval $[-10, 10]$. Find the probability that $-3 \leq X \leq 1$ or $X > 7$.

Example (2 of 2)

Random variable X is continuously and uniformly randomly distributed in the interval $[-10, 10]$. Find the probability that $-3 \leq X \leq 1$ or $X > 7$.

$$\begin{aligned}
\mathbb{P}((-3 \leq X \leq 1) \vee (X > 7)) &= \mathbb{P}(-3 \leq X \leq 1) + \mathbb{P}(X > 7) \\
&= \frac{1-(-3)}{10-(-10)} + \frac{10-7}{10-(-10)} \\
&= \frac{7}{20}
\end{aligned}$$

Expected Value

Definition

The **expected value** or **mean** of a continuous random variable X with probability density function $f(x)$ is

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} x\, f(x)\, dx.$$

Example

Find the expected value of X, if X is a continuously uniformly distributed random variable on the interval $[-10, 80]$.

Example

Find the expected value of X, if X is a continuously uniformly distributed random variable on the interval $[-10, 80]$.

$$
\begin{aligned}
\mathbb{E}[X] &= \int_{-\infty}^{\infty} x\, f(x)\, dx \\
&= \int_{-10}^{80} \frac{x}{90}\, dx \\
&= \left. \frac{x^2}{180} \right|_{-10}^{80} \\
&= \frac{6400}{180} - \frac{100}{180} \\
&= 35
\end{aligned}
$$

Example

Find the expected value of X, if X is a continuously uniformly distributed random variable on the interval $[-10, 80]$.

$$\begin{aligned}
\mathbb{E}[X] &= \int_{-\infty}^{\infty} x\, f(x)\, dx \\
&= \int_{-10}^{80} \frac{x}{90}\, dx \\
&= \left.\frac{x^2}{180}\right|_{-10}^{80} \\
&= \frac{6400}{180} - \frac{100}{180} \\
&= 35
\end{aligned}$$

Question: if X is a uniformly distributed but integer-valued RV, what is its expected value?

Expected Value of a Function

Definition

The expected value of a function g of a continuously distributed random variable X which has probability density function f is defined as

$$\mathbb{E}\left[g(X)\right] = \int_{-\infty}^{\infty} g(x) f(x)\, dx,$$

provided the improper integral converges absolutely, *i.e.*, $\mathbb{E}\left[g(X)\right]$ is defined if and only if

$$\int_{-\infty}^{\infty} |g(x)| f(x)\, dx < \infty.$$

Example

Find the expected value of X^2 if X is continuously distributed on $[0, \infty)$ with probability density function $f(x) = e^{-x}$.

Example

Find the expected value of X^2 if X is continuously distributed on $[0, \infty)$ with probability density function $f(x) = e^{-x}$.

$$\begin{aligned}
\mathbb{E}\left[X^2\right] &= \int_0^\infty x^2 e^{-x}\,dx \\
&= \lim_{M\to\infty} \int_0^M x^2 e^{-x}\,dx \\
&= \lim_{M\to\infty} \left[-(x^2+2x+2)e^{-x}\right]\Big|_0^M \\
&= \lim_{M\to\infty} \left[2-(M^2+2M+2)e^{-M}\right] \\
&= 2
\end{aligned}$$

Joint and Marginal Distributions

Definition

A **joint probability density** for a pair of random variables, X and Y, is a non-negative function $f(x, y)$ for which

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y)\, dx\, dy = 1.$$

Joint and Marginal Distributions

Definition

A **joint probability density** for a pair of random variables, X and Y, is a non-negative function $f(x, y)$ for which

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y)\, dx\, dy = 1.$$

Definition

If X and Y are continuous random variables with joint distribution $f(x, y)$ then the **marginal density** for X is defined as the function

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y)\, dy.$$

Joint and Marginal Distributions

Definition

A **joint probability density** for a pair of random variables, X and Y, is a non-negative function $f(x, y)$ for which

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y)\, dx\, dy = 1.$$

Definition

If X and Y are continuous random variables with joint distribution $f(x, y)$ then the **marginal density** for X is defined as the function

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y)\, dy.$$

Remark: a similar definition may be stated for the marginal density for Y.

If the joint probability density of X and Y is given by

$$f(x,y) = \begin{cases} 1/\pi & \text{if } x^2 + y^2 \le 1, \\ 0 & \text{otherwise} \end{cases}$$

find the marginal probability density of X.

Example

If the joint probability density of X and Y is given by

$$f(x,y) = \begin{cases} 1/\pi & \text{if } x^2+y^2 \leq 1, \\ 0 & \text{otherwise} \end{cases}$$

find the marginal probability density of X.

$$f_X(x) = \int_{-\infty}^{\infty} f(x,y)\,dy = \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} \frac{1}{\pi}\,dy = \begin{cases} \frac{2}{\pi}\sqrt{1-x^2} & \text{if } |x| \leq 1, \\ 0 & \text{otherwise} \end{cases}$$

Independence of Jointly Distributed RVs

Definition

Two continuous random variables are **independent** if and only if the joint probability density function factors into the product of the marginal densities of X and Y. In other words X and Y are independent if and only if

$$f(x, y) = f_X(x) f_Y(y)$$

for all real numbers x and y.

Example

The joint probability density function of X and Y is

$$f(x,y) = \begin{cases} \frac{1}{2}x\,y & \text{if } 0 \leq x \leq y \text{ and } 0 \leq y \leq 2 \\ 0 & \text{otherwise.} \end{cases}$$

Are X and Y independent?

Example

The joint probability density function of X and Y is

$$f(x,y) = \begin{cases} \frac{1}{2}x\,y & \text{if } 0 \le x \le y \text{ and } 0 \le y \le 2 \\ 0 & \text{otherwise.} \end{cases}$$

Are X and Y independent?

No, since $f_X(x) = x - x^3/4$ if $0 \le x \le 2$ and $f_Y(y) = y^3/4$ if $0 \le y \le 2$.

Example (1 of 2)

Consider the jointly distributed random variables $(X, Y) \in [0, \infty) \times [-2, 2]$ whose density is the function $f(x, y) = \dfrac{1}{4e^x}$. Find the mean of $X + Y$.

Example (2 of 2)

$$
\begin{aligned}
\mathbb{E}\left[X+Y\right] &= \int_0^\infty \int_{-2}^2 (x+y)\left(\frac{1}{4e^x}\right) dy\, dx \\
&= \int_0^\infty \frac{1}{4} e^{-x} \int_{-2}^2 (x+y)\, dy\, dx \\
&= \int_0^\infty \frac{1}{4} e^{-x} (4x)\, dx \\
&= \int_0^\infty x e^{-x}\, dx \\
&= \lim_{M\to\infty} \int_0^M x e^{-x}\, dx \\
&= \lim_{M\to\infty} \left(1 - M e^{-M} - e^{-M}\right) \\
&= 1
\end{aligned}
$$

Properties of the Expected Value

Theorem

If $X_1, X_2, \ldots, X_k$ are continuous random variables with joint probability density $f(x_1, x_2, \ldots, x_k)$ then
$\mathbb{E}[X_1 + X_2 + \cdots + X_k] = \mathbb{E}[X_1] + \mathbb{E}[X_2] + \cdots + \mathbb{E}[X_k]$.

Theorem

If $X_1, X_2, \ldots, X_k$ are continuous random variables with joint probability density $f(x_1, x_2, \ldots, x_k)$ then $\mathbb{E}[X_1 + X_2 + \cdots + X_k] = \mathbb{E}[X_1] + \mathbb{E}[X_2] + \cdots + \mathbb{E}[X_k]$.

Theorem

Let X_1, X_2, ..., X_k be pairwise independent random variables with joint density $f(x_1, x_2, \ldots, x_k)$, then

$$\mathbb{E}[X_1 X_2 \cdots X_k] = \mathbb{E}[X_1]\,\mathbb{E}[X_2] \cdots \mathbb{E}[X_k].$$

Variance and Standard Deviation

Definition

If X is a continuously distributed random variable with probability density function $f(x)$, the **variance** of X is defined as

$$\mathbb{V}(X) = \mathbb{E}\left[(X-\mu)^2\right] = \int_{-\infty}^{\infty} (x-\mu)^2 f(x)\,dx,$$

where $\mu = \mathbb{E}[X]$. The **standard deviation** of X is $\sigma(X) = \sqrt{\mathbb{V}(X)}$.

Variance and Standard Deviation

Definition

If X is a continuously distributed random variable with probability density function $f(x)$, the **variance** of X is defined as

$$\mathbb{V}(X) = \mathbb{E}\left[(X-\mu)^2\right] = \int_{-\infty}^{\infty} (x-\mu)^2 f(x)\, dx,$$

where $\mu = \mathbb{E}[X]$. The **standard deviation** of X is $\sigma(X) = \sqrt{\mathbb{V}(X)}$.

Theorem

Let X be a random variable with probability density f and mean μ, then $\mathbb{V}(X) = \mathbb{E}\left[X^2\right] - \mu^2$.

Example

Suppose X is continuously distributed on $[0, \infty)$ with probability density function $f(x) = e^{-x}$. Find $\mathbb{V}(X)$.

Example

Suppose X is continuously distributed on $[0, \infty)$ with probability density function $f(x) = e^{-x}$. Find $\mathbb{V}(X)$.

$$\begin{aligned}
\mathbb{V}(X) &= \mathbb{E}\left[X^2\right] - (\mathbb{E}[X])^2 \\
&= \int_0^\infty x^2 e^{-x}\,dx - \left(\int_0^\infty x e^{-x}\,dx\right)^2 \\
&= 2 - \left(\int_0^\infty x e^{-x}\,dx\right)^2 \\
&= 2 - (1)^2 \\
&= 1
\end{aligned}$$

Properties of Variance

Theorem

Let X be a continuous random variable with probability density $f(x)$ and let $a, b \in \mathbb{R}$, then

$$\mathbb{V}(aX + b) = a^2 \, \mathbb{V}(X).$$

Theorem

Let X be a continuous random variable with probability density $f(x)$ and let $a, b \in \mathbb{R}$, then

$$\mathbb{V}(aX + b) = a^2\,\mathbb{V}(X).$$

Theorem

Let $X_1, X_2, \ldots, X_k$ be pairwise independent continuous random variables with joint probability density $f(x_1, x_2, \ldots, x_k)$, then

$$\mathbb{V}(X_1 + X_2 + \cdots + X_k) = \mathbb{V}(X_1) + \mathbb{V}(X_2) + \cdots + \mathbb{V}(X_k).$$

Assumption: any characteristic of an object subject to a large number of independently acting forces typically takes on a **normal distribution**.

Assumption: any characteristic of an object subject to a large number of independently acting forces typically takes on a **normal distribution**.

- We will develop the normal probability density function from the probability function for the binomial random variable.

Assumption: any characteristic of an object subject to a large number of independently acting forces typically takes on a **normal distribution.**

- We will develop the normal probability density function from the probability function for the binomial random variable.
- Recall that if X is a binomial random variable of n trials and probability of success on a single trial of p, then for $x \in \{0, 1, \ldots, n\}$:

$$\mathbb{P}(X = x) = \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x}$$

Thought experiment: Imagine standing at the origin of the number line and for each tick of a clock taking a step to the left or the right. In the long run where will you stand?

Overview of Derivation

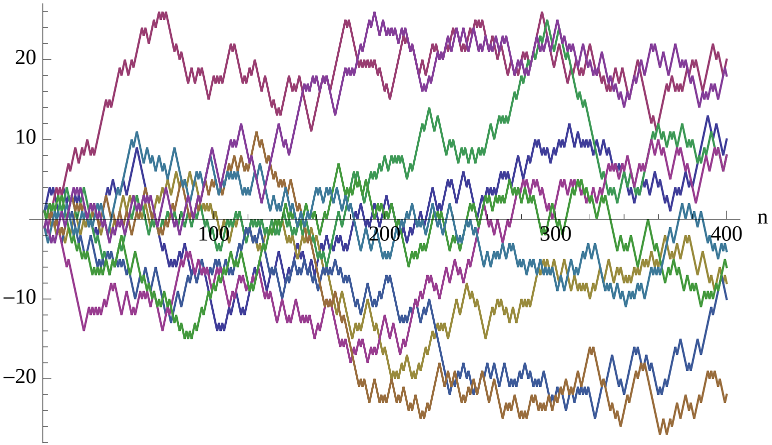
Thought experiment: Imagine standing at the origin of the number line and for each tick of a clock taking a step to the left or the right. In the long run where will you stand?

Assumptions:

1. n steps/ticks,
2. random walk takes place during time interval $[0, t]$, which implies a "tick" lasts $\Delta t = t/n$,
3. on each tick move a distance $\Delta x > 0$,
4. $n(\Delta x)^2 = 2kt$ or equivalently $(\Delta x)^2 = 2k(\Delta t)$, for some positive constant k,
5. probability of moving left/right is $1/2$,
6. all steps are independent.

Simulation

Simulate taking 400 steps, 10 different trials.



Average position on last step: $\overline{x} = 1.4$ with $\sigma(x) \approx 13.5$.

Take a Few Steps

Suppose r out of n steps $(0 \leq r \leq n)$ have been to the right.

Question: Where are you?

Take a Few Steps

Suppose r out of n steps ($0 \le r \le n$) have been to the right.

Question: Where are you?

$$(r - (n - r))\Delta x = (2r - n)\Delta x = m\Delta x$$

Question: What is the probability of standing there?

Take a Few Steps

Suppose r out of n steps ($0 \leq r \leq n$) have been to the right.

Question: Where are you?

$$(r - (n - r))\Delta x = (2r - n)\Delta x = m\Delta x$$

Question: What is the probability of standing there?

$$\begin{aligned}
\mathbb{P}\left(X = m\Delta x\right) &= \mathbb{P}\left(X = (2r - n)\Delta x\right) \\
&= \binom{n}{r}\left(\frac{1}{2}\right)^r \left(\frac{1}{2}\right)^{n-r} \\
&= \frac{n!}{r!(n-r)!}\left(\frac{1}{2}\right)^n \\
&= \frac{n!\left(\frac{1}{2}\right)^n}{\left(\frac{1}{2}(n+m)\right)!\left(\frac{1}{2}(n-m)\right)!}
\end{aligned}$$

Bernoulli Steps

Each step is a Bernoulli experiment with outcomes Δx and $-\Delta x$.

Questions:

- What is the expected value of a single step?
- What is the variance in the outcomes?

Each step is a Bernoulli experiment with outcomes Δx and $-\Delta x$.

Questions:

- What is the expected value of a single step?

$$\mathbb{E}[X] = 0$$

- What is the variance in the outcomes?

Bernoulli Steps

Each step is a Bernoulli experiment with outcomes Δx and $-\Delta x$.

Questions:

- What is the expected value of a single step?

$$\mathbb{E}[X] = 0$$

- What is the variance in the outcomes?

$$\mathbb{V}(X) = (\Delta x)^2$$

The Sum of Bernoulli Steps

Questions: after n steps,

- What is the expected value of where you stand?

- What is the variance in final position?

The Sum of Bernoulli Steps

Questions: after n steps,

- What is the expected value of where you stand?

$$\mathbb{E}\left[\sum_{i=1}^{n} X\right] = n\,\mathbb{E}[X] = 0$$

- What is the variance in final position?

The Sum of Bernoulli Steps

Questions: after *n* steps,

- What is the expected value of where you stand?

$$\mathbb{E}\left[\sum_{i=1}^{n} X\right] = n\,\mathbb{E}\left[X\right] = 0$$

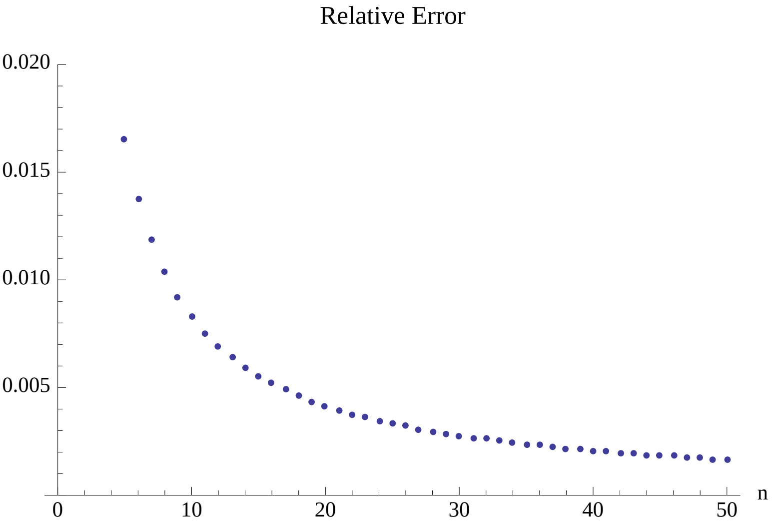
- What is the variance in final position?

$$\mathbb{V}\left(\sum_{i=1}^{n} X\right) = n\,\mathbb{V}\left(X\right) = n(\Delta x)^2 = 2k\,t$$

Stirling's formula approximates $n!$ for large n.

$$n! \approx \sqrt{2\pi}e^{-n}n^{n+1/2}$$

n	$n!$	$\sqrt{2\pi}e^{-n}n^{n+1/2}$
5	120	118.019
10	3.6288×10^6	3.5987×10^6
15	1.30767×10^{12}	1.30043×10^{12}
20	2.4329×10^{18}	2.42279×10^{18}
30	2.65253×10^{32}	2.64517×10^{32}


Relative Error
0.020
0.015
0.010
0.005
0
10
20
30
40
50
n

$$n! \approx \sqrt{2\pi}e^{-n}n^{n+1/2}$$

Replace all the factorials with Stirling's Formula.

$$n! \approx \sqrt{2\pi} e^{-n} n^{n+1/2}$$

Replace all the factorials with Stirling's Formula.

$$\begin{aligned}
\mathbb{P}(X = m\Delta x) \\
&= \frac{n! \left(\frac{1}{2}\right)^n}{\left(\frac{1}{2}(n+m)\right)! \left(\frac{1}{2}(n-m)\right)!} \\
&= \frac{\sqrt{2\pi} e^{-n} n^{n+\frac{1}{2}} \left(\frac{1}{2}\right)^n}{\sqrt{2\pi} e^{-\frac{n+m}{2}} \left(\frac{1}{2}(n+m)\right)^{\frac{n+m+1}{2}} \sqrt{2\pi} e^{-\frac{n-m}{2}} \left(\frac{1}{2}(n-m)\right)^{\frac{n-m+1}{2}}} \\
&= \frac{2}{\sqrt{2n\pi}} \left(1 + \frac{m}{n}\right)^{-\frac{m}{2}} \left(1 - \frac{m}{n}\right)^{\frac{m}{2}} \left(1 - \frac{m^2}{n^2}\right)^{-\frac{n+1}{2}}
\end{aligned}$$

Since $m = x/\Delta x$ and $n = t/\Delta t$,

$$\begin{aligned}
&\mathbb{P}\left(X = m\Delta x\right) \\
&= \frac{2}{\sqrt{2n\pi}}\left(1+\frac{m}{n}\right)^{-m/2}\left(1-\frac{m}{n}\right)^{m/2}\left(1-\frac{m^2}{n^2}\right)^{-(n+1)/2} \\
&= \frac{2\sqrt{\Delta t}}{\sqrt{2\pi t}}\left(1+\frac{x\Delta t}{t\Delta x}\right)^{-\frac{x}{2\Delta x}}\left(1-\frac{x\Delta t}{t\Delta x}\right)^{\frac{x}{2\Delta x}}\left(1-\left[\frac{x\Delta t}{t\Delta x}\right]^2\right)^{-\frac{1+t/\Delta t}{2}} \\
&= \frac{\Delta x}{\sqrt{k\pi t}}\left[1+\frac{x\,\Delta x}{2kt}\right]^{-\frac{x}{2\Delta x}}\left[1-\frac{x\,\Delta x}{2kt}\right]^{\frac{x}{2\Delta x}}\left[1-\left(\frac{x\,\Delta x}{2kt}\right)^2\right]^{-\frac{kt}{(\Delta x)^2}-\frac{1}{2}}
\end{aligned}$$

since $(\Delta x)^2 = 2k\Delta t$.

Passing to the Limit

As $\Delta x \to 0$, the probability of standing at exactly one, specific location becomes 0.

Instead we must change our thinking and ask for

$$\mathbb{P}\left((m-1)\Delta x < X < (m+1)\Delta x\right) \approx 2(\Delta x) f(x,t).$$

Passing to the Limit

As $\Delta x \to 0$, the probability of standing at exactly one, specific location becomes 0.

Instead we must change our thinking and ask for

$$\mathbb{P}\left((m-1)\Delta x < X < (m+1)\Delta x\right) \approx 2(\Delta x) f(x,t).$$

$$\begin{aligned}
f(x,t) &= \frac{1}{2\sqrt{k\pi t}} \lim_{\Delta x \to 0} \left[1 + \frac{x\,\Delta x}{2kt}\right]^{\frac{-x}{2\Delta x}} \left[1 - \frac{x\,\Delta x}{2kt}\right]^{\frac{x}{2\Delta x}} \left[1 - \left(\frac{x\,\Delta x}{2kt}\right)^2\right]^{-\frac{kt}{(\Delta x)^2} - \frac{1}{2}} \\
&= \frac{1}{2\sqrt{k\pi t}} \left(e^{\frac{x}{2kt}}\right)^{-\frac{x}{2}} \left(e^{-\frac{x}{2kt}}\right)^{\frac{x}{2}} \left(e^{-\frac{x^2}{4k^2t^2}}\right)^{-kt} \\
&= \frac{1}{2\sqrt{k\pi t}} e^{-\frac{x^2}{4kt}}
\end{aligned}$$

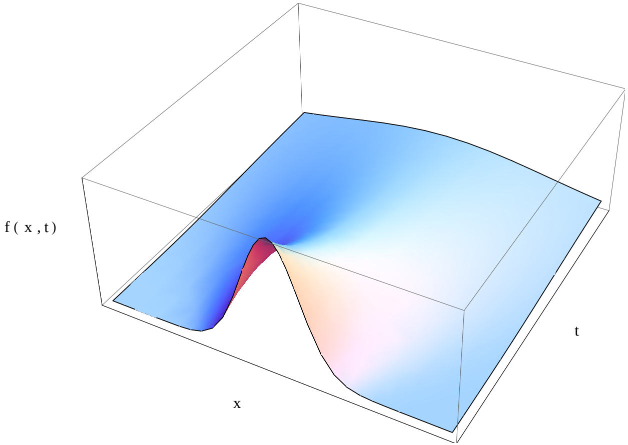
Is $f(x,t)$ a PDF?

Suppose $\displaystyle\int_{-\infty}^{\infty} \frac{1}{2\sqrt{k\pi t}} e^{-\frac{x^2}{4kt}}\,dx = S$, then

$$
\begin{aligned}
S^2 &= \int_{-\infty}^{\infty} \frac{1}{2\sqrt{k\pi t}} e^{-\frac{x^2}{4kt}}\,dx \int_{-\infty}^{\infty} \frac{1}{2\sqrt{k\pi t}} e^{-\frac{y^2}{4kt}}\,dy \\
&= \frac{1}{4k\pi t} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-(x^2+y^2)/4kt}\,dx\,dy \\
&= \frac{1}{4k\pi t} \int_{0}^{2\pi} \int_{0}^{\infty} r\,e^{-r^2/4kt}\,dr\,d\theta \\
&= 1
\end{aligned}
$$

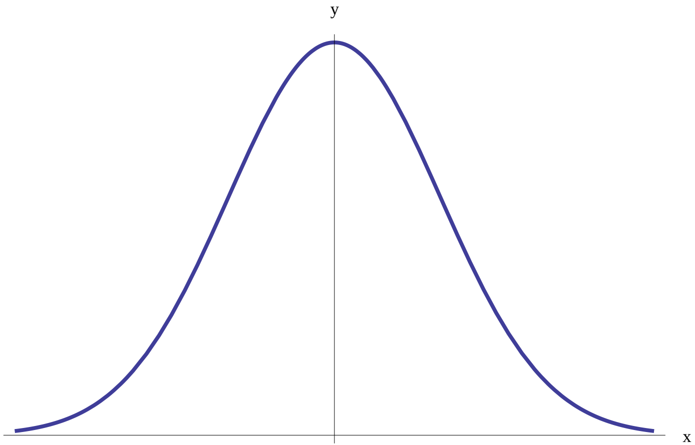
Surface Plot

The graph of the PDF resembles:



The Bell Curve

For a fixed value of t, the graph of the PDF resembles:



If X is a continuously distributed random variable with PDF:

$$f(x,t) = \frac{1}{2\sqrt{k\pi t}} e^{-\frac{x^2}{4kt}}$$

then

Expected Value and Variance

If X is a continuously distributed random variable with PDF:

$$f(x,t) = \frac{1}{2\sqrt{k\pi t}} e^{-\frac{x^2}{4kt}}$$

then

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} \frac{x}{2\sqrt{k\pi t}} e^{-\frac{x^2}{4kt}}\, dx = 0$$

and

Expected Value and Variance

If X is a continuously distributed random variable with PDF:

$$f(x,t) = \frac{1}{2\sqrt{k\pi t}} e^{-\frac{x^2}{4kt}}$$

then

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} \frac{x}{2\sqrt{k\pi t}} e^{-\frac{x^2}{4kt}}\, dx = 0$$

and

$$\mathbb{V}(X) = \int_{-\infty}^{\infty} \frac{x^2}{2\sqrt{k\pi t}} e^{-\frac{x^2}{4kt}}\, dx - (\mathbb{E}[X])^2 = 2k\,t$$

and thus $2k\,t = \sigma^2$ and we express the PDF for a **normally distributed random variable** with mean μ and variance σ^2 as

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

Standard Normal Distribution

When $\mu = 0$ and $\sigma = 1$, the PDF $\phi(x) = \dfrac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$ is called the **standard normal density.**

Standard Normal Distribution

When $\mu = 0$ and $\sigma = 1$, the PDF $\phi(x) = \dfrac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$ is called the **standard normal density.**

The **cumulative distribution function** (CDF) $\Phi(x)$ is defined as

$$\Phi(x) = \mathbb{P}(X < x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} \, du.$$

When $\mu = 0$ and $\sigma = 1$, the PDF $\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$ is called the **standard normal density**.

The **cumulative distribution function** (CDF) $\Phi(x)$ is defined as

$$\Phi(x) = \mathbb{P}(X < x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}}\, du.$$

Remarks:

- Pay close attention of the case of the symbol used for the probability density and cumulative distribution of X.
- The values of $\Phi(x)$ can be produced by many scientific calculators or by looking them up in printed tables.

Change of Variable

Theorem

If X is a normally distributed random variable with expected value μ and variance σ^2, then $Z = \dfrac{X - \mu}{\sigma}$ is normally distributed with an expected value of zero and a variance of one.

Central Limit Theorem (1 of 2)

Suppose the random variables $X_1, X_2, \ldots, X_n$:

1. are pairwise independent but not necessarily identically distributed,
2. have means $\mu_1, \mu_2, \ldots, \mu_n$ and variances $\sigma_1^2, \sigma_2^2, \ldots, \sigma_n^2$,

and we define a new random variable Y_n as

$$Y_n = \frac{\sum_{i=1}^{n} (X_i - \mu_i)}{\sqrt{\sum_{i=1}^{n} \sigma_i^2}}.$$

Central Limit Theorem (1 of 2)

Suppose the random variables $X_1, X_2, \ldots, X_n$:

1. are pairwise independent but not necessarily identically distributed,
2. have means $\mu_1, \mu_2, \ldots, \mu_n$ and variances $\sigma_1^2, \sigma_2^2, \ldots, \sigma_n^2$,

and we define a new random variable Y_n as

$$Y_n = \frac{\sum_{i=1}^n (X_i - \mu_i)}{\sqrt{\sum_{i=1}^n \sigma_i^2}}.$$

A **Central Limit Theorem** due to Liapounov implies that Y_n has the standard normal distribution.

Central Limit Theorem (2 of 2)

Theorem

Suppose that the infinite collection $\{X_i\}_{i=1}^{\infty}$ *of random variables are pairwise independent and that for each* $i \in \mathbb{N}$ *we have* $\mathbb{E}\left[|X_i - \mu_i|^3\right] < \infty$. *If in addition,*

$$\lim_{n\to\infty} \frac{\sum_{i=1}^{n} \mathbb{E}\left[|X_i - \mu_i|^3\right]}{\left(\sum_{i=1}^{n} \sigma_i^2\right)^{3/2}} = 0$$

then for any $x \in \mathbb{R}$

$$\lim_{n\to\infty} \mathbb{P}\left(Y_n \leq x\right) = \Phi\left(x\right)$$

where random variable Y_n *is defined as above.*

Suppose the annual snowfall in Millersville, PA is 14.6 inches with a standard deviation of 3.2 inches and is normally distributed. Snowfall amounts in different years are independent. What is the probability that the sum of the snowfall amounts in the next two years will exceed 30 inches?

Solution: If X represents the random variable standing for the snowfall received in Millersville, PA for one year then $X + X$ is the random variable representing the snowfall of two years. The random variable $X + X$ has mean $\mu = 2(14.6) = 29.2$ inches and variance $\sigma^2 = (3.2)^2 + (3.2)^2 = 20.48$.

$$\begin{aligned} \mathbb{P}(X + X > 30) &= \mathbb{P}\left(Z > \frac{30 - 29.2}{\sqrt{20.48}}\right) \\ &= 1 - \mathbb{P}(Z \leq 0.176777) \\ &= 1 - \Phi(0.176777) \\ &= 0.429842 \end{aligned}$$

Lognormal Random Variables

Definition

A random variable X is a **lognormal random variable** with parameters μ and σ if $\ln X$ is a normally distributed random variable with mean μ and variance σ^2.

Lognormal Random Variables

Definition

A random variable X is a **lognormal random variable** with parameters μ and σ if $\ln X$ is a normally distributed random variable with mean μ and variance σ^2.

Remarks:

- The parameter μ is sometimes called the **drift.**
- The parameter σ is sometimes called the **volatility**.

Lognormal PDF (1 of 2)

Suppose X is lognormal, then $Y = \ln X$ is normal and

$$\begin{aligned} \mathbb{P}\,(X < x) &= \mathbb{P}\,(Y < \ln x) \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\ln x} e^{-(t-\mu)^2/2\sigma^2}\,dt. \end{aligned}$$

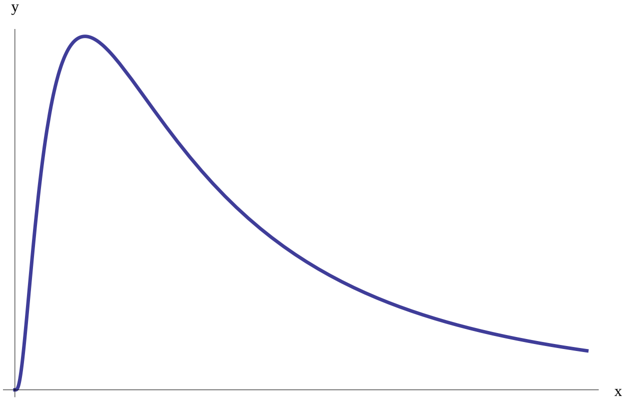
If we let $u = e^t$ and $du = e^t\,dt$, then

$$\mathbb{P}\,(X < x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \frac{1}{u} e^{-(\ln u-\mu)^2/2\sigma^2}\,du,$$

the cumulative distribution function for the lognormally distributed random variable X.

The probability density function for lognormal X is

$$f(x) = \frac{1}{(\sigma\sqrt{2\pi})x} e^{-(\ln x - \mu)^2/2\sigma^2}.$$



Mean and Variance of a Lognormal RV

Lemma

If X is a lognormal random variable with parameters μ and σ then

$$\begin{aligned}\mathbb{E}[X] &= e^{\mu+\sigma^2/2},\\ \mathbb{V}(X) &= e^{2\mu+\sigma^2}\left(e^{\sigma^2}-1\right).\end{aligned}$$

Let X be lognormally distributed with parameters μ and σ, then

$$\begin{aligned}
\mathbb{E}[X] &= \frac{1}{\sigma\sqrt{2\pi}} \int_0^\infty x \left(\frac{1}{x} e^{-(\ln x - \mu)^2/2\sigma^2} \right) dx \\
&= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^\infty e^t e^{-(t-\mu)^2/2\sigma^2}\, dt \\
&= e^{\mu+\sigma^2/2} \left[\frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^\infty e^{-(t-(\mu+\sigma^2))^2/2\sigma^2}\, dt \right] \\
&= e^{\mu+\sigma^2/2}.
\end{aligned}$$

Let X be lognormally distributed with parameters μ and σ, then

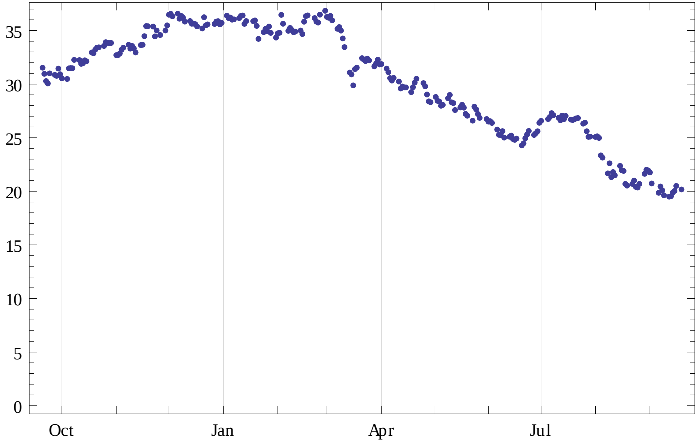
$$\begin{aligned}
\mathbb{V}(X) &= \mathbb{E}\left[X^2\right] - (\mathbb{E}[X])^2 \\
&= \frac{1}{\sigma\sqrt{2\pi}} \int_0^\infty x^2 \left(\frac{1}{x} e^{-(\ln x-\mu)^2/2\sigma^2}\right) dx - \left(e^{\mu+\sigma^2/2}\right)^2 \\
&= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^\infty e^{2t} e^{-(t-\mu)^2/2\sigma^2}\, dt - e^{2\mu+\sigma^2} \\
&= e^{2(\mu+\sigma^2)} \left[\frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^\infty e^{-(t-(\mu+2\sigma))^2/2\sigma^2}\, dt\right] - e^{2\mu+\sigma^2} \\
&= e^{2(\mu+\sigma^2)} - e^{2\mu+\sigma^2} \\
&= e^{2\mu+\sigma^2}\left(e^{\sigma^2} - 1\right).
\end{aligned}$$

Observation:

- Let $S(0)$ denote the price of a security at some starting time arbitrarily chosen to be $t = 0$.
- For $n \geq 1$, let $S(n)$ denote the price of the security on day n.
- The random variable $X(n) = \dfrac{S(n)}{S(n-1)}$ for $n \geq 1$ is lognormally distributed, *i.e.*, $\ln X(n) = \ln S(n) - \ln S(n-1)$ is normally distributed.

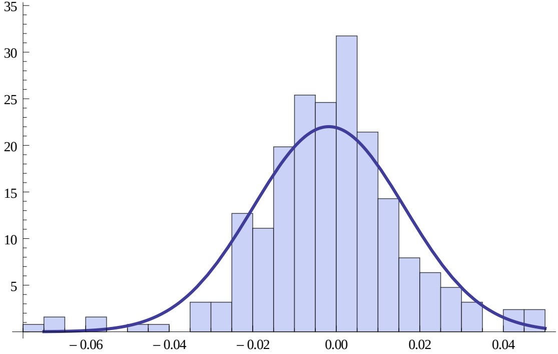
Closing Prices of Sony (SNE) Stock

Closing prices of Sony Corporation stock (09/20/2010–09/19/2011):



Lognormal Behavior of Sony (SNE) Stock

Lognormal behavior of closing prices:



$$\mu = -0.00177279 \quad \sigma = 0.0181285$$

Example (1 of 2)

What is the probability that the closing price of Sony Corporation stock will be higher today than yesterday?

What is the probability that the closing price of Sony Corporation stock will be higher today than yesterday?

$$\begin{aligned}
\mathbb{P}\left(\underbrace{\frac{S(n)}{S(n-1)}}_{\text{lognormal}} > 1\right) &= \mathbb{P}\left(\underbrace{\ln\frac{S(n)}{S(n-1)}}_{\text{normal}} > \ln 1\right) \\
&= \mathbb{P}\left(X > 0\right) \\
&= \mathbb{P}\left(Z > \frac{0-(-0.00177279)}{0.01811285}\right) \\
&= 1 - \mathbb{P}\left(Z \leq 0.09779\right) \\
&= 1 - \Phi\left(0.09779\right) \\
&= 0.46105
\end{aligned}$$

Example (2 of 2)

What is the probability that tomorrow's closing price will be higher than yesterday's closing price?

What is the probability that tomorrow's closing price will be higher than yesterday's closing price?

$$\begin{aligned}
\mathbb{P}\left(\frac{S(n+1)}{S(n-1)} > 1\right) &= \mathbb{P}\left(\frac{S(n+1)}{S(n)}\,\frac{S(n)}{S(n-1)} > 1\right) \\
&= \mathbb{P}\left(\ln\frac{S(n+1)}{S(n)} + \ln\frac{S(n)}{S(n-1)} > 0\right) \\
&= \mathbb{P}\left(X + X > 0\right) \\
&= \mathbb{P}\left(Z > \frac{0 - 2(-0.00177279)}{\sqrt{2(0.01811285)^2}}\right) \\
&= 1 - \mathbb{P}\left(Z \leq 0.138296\right) \\
&= 1 - \Phi\left(0.138296\right) \\
&= 0.445003
\end{aligned}$$

If an item is worth K but can only be sold for X, a rational investor would sell only if $X \geq K$.

The net **payoff** of the sale can be expressed as

$$(X - K)^+ = \begin{cases} X - K & \text{if } X \geq K, \\ 0 & \text{if } X < K. \end{cases}$$

Payoff When X is Normal

Corollary

If X is normal random variable with mean μ and variance σ^2 and K is a constant, then

$$\mathbb{E}\left[(X-K)^+\right] = \frac{\sigma}{\sqrt{2\pi}} e^{-(\mu-K)^2/2\sigma^2} + (\mu-K)\Phi\left(\frac{\mu-K}{\sigma}\right),$$

$$\begin{aligned}
&\mathbb{V}\left((X-K)^+\right) \\
&= \left((\mu-K)^2+\sigma^2\right)\Phi\left(\frac{\mu-K}{\sigma}\right) + \frac{(\mu-K)\sigma}{\sqrt{2\pi}} e^{-(\mu-K)^2/2\sigma^2} \\
&\quad - \left(\frac{\sigma}{\sqrt{2\pi}} e^{-(\mu-K)^2/2\sigma^2} + (\mu-K)\Phi\left(\frac{\mu-K}{\sigma}\right)\right)^2.
\end{aligned}$$

Payoff When X is Lognormal

Corollary

If X is a lognormally distributed random variable with parameters μ and σ^2 and $K > 0$ is a constant then

$$\begin{aligned}
\mathbb{E}\left[(X-K)^+\right] &= e^{\mu+\sigma^2/2}\Phi\left(\frac{\mu-\ln K}{\sigma}+\sigma\right) - K\Phi\left(\frac{\mu-\ln K}{\sigma}\right), \\
\mathbb{V}\left((X-K)^+\right) &= e^{2(\mu+\sigma^2)}\Phi\left(w+2\sigma\right) + K^2\Phi\left(w\right) \\
&\quad - 2Ke^{\mu+\sigma^2/2}\Phi\left(w+\sigma\right) \\
&\quad - \left(e^{\mu+\sigma^2/2}\Phi\left(w+\sigma\right) - K\Phi\left(w\right)\right)^2
\end{aligned}$$

where $w = (\mu - \ln K)/\sigma$.

Credits

These slides are adapted from the textbook,

An Undergraduate Introduction to Financial Mathematics, 3rd edition, (2012).

author: J. Robert Buchanan

publisher: World Scientific Publishing Co. Pte. Ltd.

address: 27 Warren St., Suite 401–402, Hackensack, NJ 07601

ISBN: 978-9814407441