COMPUTATIONAL FINANCE & RISK MANAGEMENT
UNIVERSITY *of* WASHINGTON
Department of Applied Mathematics

CFRM 410: Probability and Statistics for Computational Finance

Week 5 Probability Theory

Jake Price
Instructor, Computational Finance and Risk Management
University of Washington
Slides originally produced by Kjell Konis

Outline

Random Experiments

Probability theory allows us to develop models for random phenomenon

A *random experiment* is an experiment where it is impossible to predict the outcome

In principle, if a random experiment is repeated indefinitely under identical conditions, the outcome will vary from repetition to repetition

Examples:

- Flipping a coin
- Selecting a card from a well-shuffled deck

Probability Models

The set of all possible outcomes of a random experiment is called the sample space and is denoted by S

We think of S as a set where each element $s \in S$ represents one elementary outcome of the random experiment

Any subset of S is called an *event*. An event may be composed of one or more outcomes of the random experiment

Example *Rolling a 6-sided die:*

$$\begin{aligned} S &= \{1, 2, 3, 4, 5, 6\} \\ A &= \{6\} \text{ (i.e., rolling a 6) is both an elementary outcome and} \\ &\quad\ \text{an event} \\ B &= \{1, 2, 3\} \text{ is a (compound) event} \end{aligned}$$

Events

Let A and B be events (A and B subsets of the sample space S)

- The event A is said to occur if the outcome of the random experiment is in A

- Define containment

  $$A \subset B \iff \{x \in A \implies x \in B\}$$

  to order subsets

- Equality is then defined for sets as

  $$A = B \iff \{A \subset B \text{ and } B \subset A\}$$

Operations on Subsets: Intersection

A and $B \quad \longleftrightarrow \quad A \cap B \qquad$ (intersection of sets A and B)

- The *intersection* of two sets A and B consists of the elements that are in both A and B
- If there are no elements that are in both A and B then their intersection is the empty set $\emptyset$ (i.e., the set containing zero elements)
- The intersection of sets A and B is symmetric: $A \cap B = B \cap A$

Examples

- One throw of a die: $A = \textit{result even}$, $B = \textit{result prime}$

$$A \cap B \quad = \quad \{2, 4, 6\} \cap \{2, 3, 5\} \quad = \quad \{2\}$$

- One throw of a die: $A = \textit{result even}$, $B = \{3\}$

$$A \cap B \quad = \quad \{2, 4, 6\} \cap \{3\} \quad = \quad \emptyset$$

Operations on Subsets: Union

A or $B \quad \longleftrightarrow \quad A \cup B \qquad$ (union of sets A and B)

- The *union* of two sets A and B consists of the elements that are in A, in B, or in both
- The union of two sets A and B is empty only if $A = B = \emptyset$
- The union of sets A and B is symmetric: $A \cup B = B \cup A$

Example

- One throw of a die: $A =$ *result even*, $B =$ *result prime*

$$A \cup B \quad = \quad \{2, 4, 6\} \cup \{2, 3, 5\} \quad = \quad \{2, 3, 4, 5, 6\}$$

Operations on Subsets: Complement

Not $A \quad \longleftrightarrow \quad A^c \qquad$ (A complement)

- The *complement* of a set A is the set containing all of the elements of S that are not in A
- The complement of A is empty only if $A = S$
- Obviously $A \cup A^c = S$ and $A \cap A^c = \emptyset$

Example

- One throw of a die: $A = \textit{result even}$

$$A^c \quad = \quad \{2, 4, 6\}^c \quad = \quad \{1, 3, 5\}$$

Operations on Subsets: Set Difference

$$A \text{ but not } B \quad \longleftrightarrow \quad A \setminus B = A \cap B^c \qquad \text{(set difference)}$$

- The *difference* between a set A and a set B is the set of elements of A that are not in B
- *WARNING:* the set difference is not in general symmetric

$$A \setminus B = A \cap B^c \neq B \cap A^c = B \setminus A$$

- The difference between two sets A and B is empty only if $A \subset B$

Example

- One throw of a die: $A = \textit{result odd}$, $B = \textit{result prime}$

$$A \setminus B = \{1, 3, 5\} \setminus \{2, 3, 5\} = \{1\}$$
$$B \setminus A = \{2, 3, 5\} \setminus \{1, 3, 5\} = \{2\}$$

Operations on Subsets: Intervals

Subset operations also work for intervals on the real line

We use the notation [to denote closed and (to denote open

Examples

- $[1, 5] \cap (4, 7) = (4, 5]$
- $[1, 5] \cup (4, 7) = [1, 7)$
- $(-\infty, 0)^c = [0, \infty)$
- $[1, 5] \setminus (4, 7) = [1, 4]$

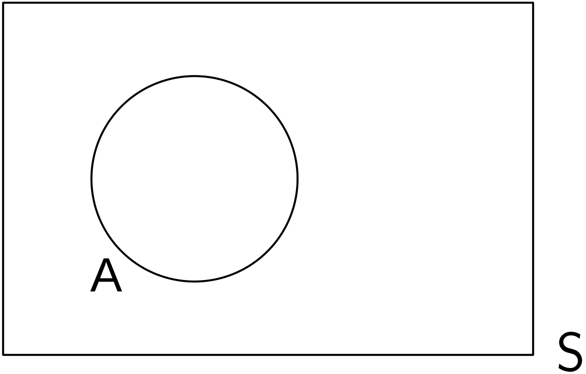
The intervals may also be disjoint

- $[1, 5] \setminus [2, 4] = \{ \ [1, 2), (4, 5] \ \}$
- $(1, 3) \cup (3, 5) = (1, 5) \setminus 3$

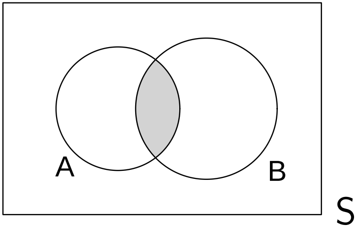
Venn Diagrams

A *Venn diagram* is a simple tool for visualizing sets and the operations between them

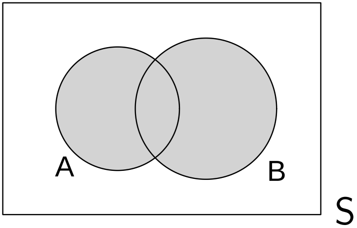
- The sample space S is represented by a rectangle
- Subsets are represented by discs contained in the rectangle



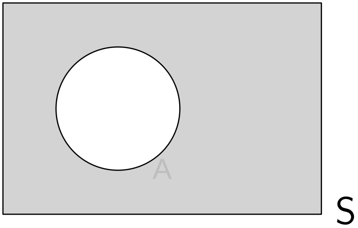
Venn Diagrams: Operations on Subsets



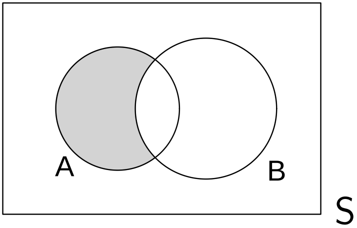
(a) $A \cap B$



(b) $A \cup B$



(c) A^c



(d) $A \setminus B$

Operations on Subsets: Properties

For two events A and B we have the following properties

Commutativity:	$A \cap B = B \cap A$ $A \cup B = B \cup A$
Associativity:	$A \cap (B \cap C) = (A \cap B) \cap C$ $A \cup (B \cup C) = (A \cup B) \cup C$
Distributive:	$A \cap (B \cup C) = (A \cap C) \cup (B \cap C)$ $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$
DeMorgan's Laws:	$(A \cup B)^c = A^c \cap B^c$ $(A \cap B)^c = A^c \cup B^c$

Outline

Set Theory

Probability Functions

Random Variables

Probability Functions

A *probability function* P satisfies the following properties:

- $P(\emptyset) = 0$
- $P(A^c) = 1 - P(A)$
- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
- $A \subset B \Rightarrow P(A) \leq P(B)$

Example Flipping a coin two times:

$$S = \{HH, HT, TH, TT\}$$

Let A be the event that there is at least one heads and let B be the event that there is at least one tails

What are the probabilities of the events $A \cap B$ and $A \cup B$ if each of the 4 outcomes in S is equally likely?

Equally Likely Outcomes

Often assume that each of the outcomes composing the sample space S is equally likely

For any event $A \subset S$

$$\begin{aligned} P(A) &= \frac{\text{number of outcomes in } A}{\text{number of outcomes in } S} \\ &= \frac{\text{number of favorable cases}}{\text{total number of cases}} \end{aligned}$$

Example Suppose each face of a 6-sided die is equally likely

$$P(\{1\}) = P(\{2\}) = \cdots = P(\{6\}) = \frac{1}{6}$$

then

$$P(\text{outcome odd}) = P(\{1,3,5\}) = \frac{\#\{1,3,5\}}{\#\{1,2,3,4,5,6\}} = \frac{3}{6} = \frac{1}{2}$$

Counting

When the elementary outcomes are all equally likely, computing probabilities is simply a matter of counting

Fundamental Theorem of Counting If a job consists of k separate tasks and task i can be done in n_i ways, then the entire job can be done in $n_1 \times n_2 \times \cdots \times n_k$ ways.

Combinations and Permutations

When counting ...

- if the order matters it is a *permutation*
- if the order does not matter it is a *combination*

In fact, a permutation is an ordered combination

Examples:

Combination The number of 5-card hands containing 2 kings

Permutation The number of *combinations* on a Master Lock dial *combination* lock (should be called a permutation lock!)

With and Without Replacement

When counting . . .

- if the objects being counted can be counted multiple times then we are counting *with* replacement
- if the objects being counted can be counted at most once, then we are counting *without* replacement

Examples:

- How many numbers we can write down using 3 digits?
- How many ways can 3 competitors finish a race?

Counting Permutations

Definition For a positive integer n, $n!$ (n *factorial*) is the product

$$n! = n \times (n-1) \times (n-2) \times \cdots \times 3 \times 2 \times 1$$

Further, 0! is defined to be equal to 1

- Fundamental Theorem of Counting: the number of permutations of n things is equal to $n!$
- The number of permutations of k things drawn from a population of size n is

$$\frac{n!}{(n-k)!} = \frac{n \times (n-1) \times \cdots \times (n-k+1) \times (n-k)!}{(n-k)!}$$

Counting Combinations

Recall that a permutation is an ordered combination

We just saw that there are

$$\frac{n!}{(n-k)!}$$

permutations of k things taken from a population of size n

For each combination of k things taken from a population of size n, there are $k!$ permutations, thus

$$k! \times \{\# \text{ of combinations}\} = \frac{n!}{(n-k)!}$$

The number of combinations of k things taken from a population of size n (read *n choose k*) is

$$\{\# \text{ of combinations}\} = \frac{n!}{k!(n-k)!} = \binom{n}{k}$$

Counting Summary

How many ways are there to choose k things from a population of n things ...

	Without Replacement	With Replacement
Unordered	$\binom{n}{k}$	$\binom{n+k-1}{k}$
Ordered	$\frac{n!}{(n-k)!}$	n^k

Conditional Probability and Independence

Knowledge that an event B has occurred may influence the probability of another event A

Definition The *conditional probability* of A given that an event B has occurred is defined to be

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \quad \text{with} \quad P(B) > 0$$

Definition Two events A and B are said to be *independent* if

$$P(A \cap B) = P(A) \times P(B)$$

Or equivalently: $P(A|B) = P(A)$

Independence (continued)

Example In two flips of a fair coin, find the conditional probability that the second flip is heads given that the first flip was heads

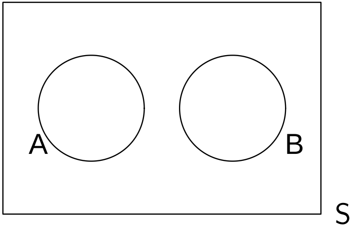
Let $A = \{HH, TH\}$ and $B = \{HH, HT\}$, then

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(\{HH\})}{P(\{HH, HT\})} = \frac{1/4}{1/2} = \frac{1}{2}$$

Example In one roll of a 6-sided die, are the events $A = \{2, 4\}$ and $B = \{2, 4, 6\}$ independent?

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(\{2, 4\})}{P(\{2, 4, 6\})} = \frac{1/3}{1/2} = \frac{2}{3} \neq \frac{1}{3} = P(A)$$

Independence (warning)



Do not confuse independent events with incompatible (disjoint) events

$$\begin{aligned} A \cap B = \emptyset \quad &\Rightarrow \quad A,\, B \text{ are disjoint} \\ &\Rightarrow \quad P(A \cup B) = P(A) + P(B) \end{aligned}$$

If A and B are disjoint with $P(A) > 0$ and $P(B) > 0$) then

$$P(A \cap B) = P(\emptyset) = 0, \quad \text{but} \quad P(A) \times P(B) \neq 0$$

thus A and B are dependent (i.e., not independent)

Total Probability

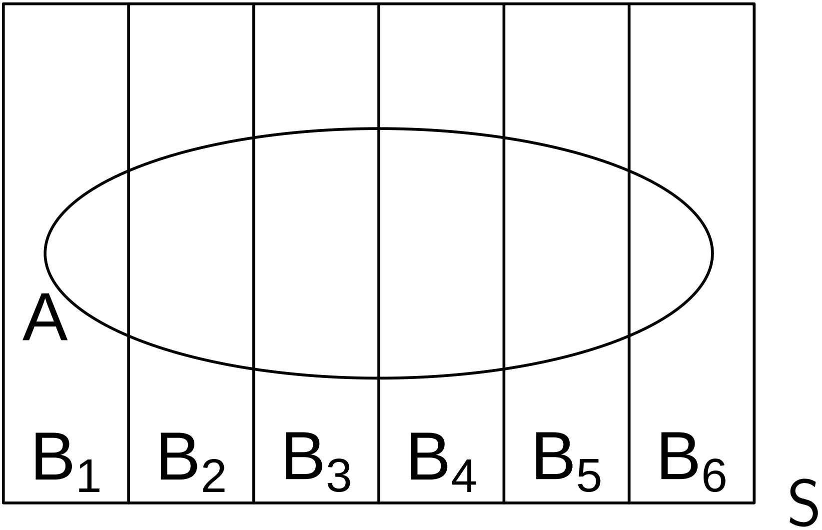
Definition Let A be an event defined on the sample space S, the *law of total probability* states that

$$P(A) = \sum_{i=1}^{n} P(A \cap B_i) = \sum_{i=1}^{n} P(A|B_i)P(B_i)$$

where the events $\{B_1, \ldots, B_n\}$ form a *partition* of S

$$B_i \cap B_j = \emptyset \text{ for } i \neq j \quad \text{and} \quad \bigcup_{i=1}^{n} B_i = S$$

Total Probability: Venn Diagram



Example: Bond Portfolio

Suppose that a bond portfolio is allocated

50% High credit-quality investment grade
30% Medium credit-quality investment grade
20% Low credit-quality or "junk bonds"

Further, suppose the probabilities of default for each rating are

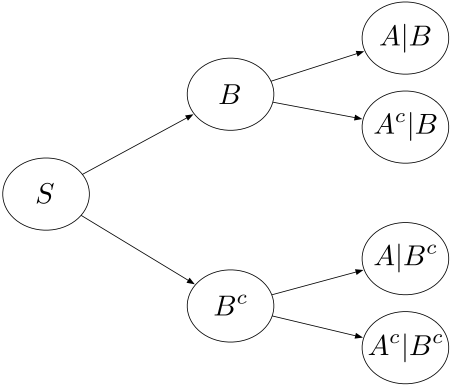
High credit-quality	0.01
Medium credit-quality	0.05
Low credit-quality	0.25

What is the probability that a randomly selected bond in the portfolio will default?

Bayes' Theorem

Theorem Let A be an event defined on a sample space S, and let $B_1, \ldots, B_n$ be a partition of S, then

$$P(B_i|A) = \frac{P(B_i \cap A)}{P(A)} = \frac{P(A|B_i)P(B_i)}{\sum_{j=1}^{n} P(A|B_j)P(B_j)}$$



Example: Bond Portfolio

Suppose that a bond portfolio is allocated

- 50% High credit-quality investment grade
- 30% Medium credit-quality investment grade
- 20% Low credit-quality or "junk bonds"

Further, suppose the probabilities of default for each rating are

High credit-quality	0.01
Medium credit-quality	0.05
Low credit-quality	0.25

What is the probability that a defaulted bond is medium credit-quality?

Outline

Set Theory

Probability Functions

Random Variables

Random Variables

Example Suppose we roll 2 6-sided dice but are interested in the sum of the values on the two faces rather than the specific outcome

Often more convenient to consider a function of the outcome of a random experiment rather than the outcome itself

Definition Let S be a sample space. A *random variable* X defined on S is a function from S to $\mathbb{R}$ (i.e., the real numbers):

$$\begin{aligned} X \quad : \quad & S \to \mathbb{R} \\ & s \to X(s) \end{aligned}$$

where $s \in S$ is an outcome of the random experiment

Examples

The set of possible values of a random variable X is called the support of X and is denoted by S_X

The support S_X can be discrete or continuous, and finite (bounded) or infinite (unbounded)

- The number of heads in ten flips of a coin
- The number of tails before the first head
- A randomly chosen compass direction
- The yearly rainfall in Seattle

Discrete Random Variables

Definition A random variable X is *discrete* if its support S_X is a finite or countably infinite set

The function

$$f_X(x) = P(X = x)$$

is called the *probability mass function*

A discrete random variable is characterized by

- the support $S_X = \{x_1, x_2, \dots\}$
- the table of probabilities $f_X(x_i) = P(X = x_i)$ for all $x_i \in S_X$

Probability Mass Function

The probability mass function $f_X(x)$ satisfies the following properties:

- $0 \leq f_X(x_i) \leq 1$ for $x_i \in S_X$
- $f_X(x) = 0$ for every other value of x
- $\sum_{x_i \in S_X} f_X(x_i) = 1$

Example Consider rolling 2 6-sided dice

(a) Let X be the sum of the two dice. Find $f_X(x)$

(b) Let Y be the maximum of the two dice. Find $f_Y(y)$

Cumulative Distribution Function

Definition The *cumulative distribution function* (cdf) $F_X(x)$ of a random variable X is

$$F_X(x) = P(X \leq x)$$

The cumulative distribution function has the following properties:

- $F_X(x)$ takes values in the interval $[0, 1]$
- $F_X(x)$ is nondecreasing
- if X is a discrete random variable then

  $$F_X(x) = \sum_{x_i \leq x} P(X = x_i)$$

  $F_X(x)$ is right-continuous at $x = x_i$
- $P(a < X \leq b) = F_X(b) - F_X(a)$

Quantile Function

Definition The quantile function Q of a distribution F_X specifies, for a given probability $p \in [0, 1]$, the smallest value of $x \in \mathbb{R}$ such that $F_X(x) \geq p$

$$Q_{F_X}(p) = \inf_{x \in \mathbb{R}} \{F_X(x) \geq p\}$$

If $F_X(x)$ is a strictly increasing function of x then the quantile function is its reciprocal, e.g.,

$$Q_{F_X}(p) = F_X^{-1}(p)$$

Examples:

- The *median* is the 0.5-quantile
- *Quartiles* are the $\{0.25, 0.5, 0.75\}$-quantiles
- *Percentiles* are the $\{0.01, 0.02, \ldots, 0.98, 0.99\}$-quantiles

Notation

Uppercase letters X, Y, Z, W, etc., denote random variables

Lowercase letters x, y, z, w, etc., are possible values of the random variable denoted by the corresponding uppercase letter

The cumulative distribution function of a random variable X is denoted by an upper case letter, e.g., $F_X(x)$

The probability mass function of a random variable X is denoted by a lowercase letter, e.g., $f_X(x)$

The notation $X \sim F$ is read 'X is distributed F' and means that the cumulative distribution function (or distribution) of X is F

Continuous Random Variables

Definition A random variable X that takes values in an interval (or collection of intervals) on the real line is called a *continuous random variable*

A continuous random variable X is characterized by a function $f_X(x)$ called the *probability density function* (or simply the *density*) of X

The function $f_X(x)$ is defined so that

$$P(X \in A) = \int_A f_X(x)dx$$

where A is a subset of S_X

Example Suppose $S_X = \mathbb{R}$ and let A be the interval $[a, b]$, then

$$P(X \in A) = P(a \leq X \leq b) = \int_a^b f_X(x)dx$$

Density Functions

The density function $f_X(x)$ has the following properties:

- Essential properties:

$$f_X(x) \geq 0 \quad \forall\, x \qquad \text{and} \qquad \int_{-\infty}^{\infty} f_X(x)\, dx = 1$$

- When $b = a$

$$P(X = a) = \int_a^a f_X(x)\, dx = 0$$

- The cdf $F_X(x) = P(X \leq x) = P(X < x)$ of a continuous random variable X is

$$F_X(x) = \int_{-\infty}^{x} f_X(u)\, du$$

Cumulative Distribution Function

- For a continuous random variable X

$$P(a < X \leq b) = F_X(b) - F_X(a)$$

- At every point x where $F_X(x)$ is differentiable

$$\frac{d}{dx} F_X(x) = F'_X(x) = f_X(x)$$

and since $f_X(x) \geq 0$, $F_X(x)$ is nondecreasing

Expected Value

The *expected value* (or *mean*) of a random variable X is

$$\mathrm{E}(X) = \begin{cases} \displaystyle\int_{-\infty}^{\infty} x\, f_X(x)\, dx & X \text{ continuous} \\[2em] \displaystyle\sum_{x_i \in S_X} x_i\, f_X(x_i) & X \text{ discrete} \end{cases}$$

The expected value can be interpreted as the center of gravity of the distribution

Properties of the Expected Value

For any function $g(x)$

$$\mathrm{E}(g(X)) = \begin{cases} \displaystyle\int_{-\infty}^{\infty} g(x) f_X(x)\, dx & \text{X} \quad \text{continuous} \\[2ex] \displaystyle\sum_{x_i \in S_X} g(x_i) f_X(x_i) & \text{X} \quad \text{discrete} \end{cases}$$

If $g(X) = aX + b$ where a and b are any two constants

$$\mathrm{E}(aX + b) = a\mathrm{E}(X) + b$$

Moments

An important class of expectations are the *moments* of a distribution

For positive integer n, the n^{th} *moment* μ'_n of X is

$$\mu'_n = \mathrm{E}(X^n)$$

The n^{th} *central moment* μ_n of X is

$$\mu_n = \mathrm{E}([X - \mathrm{E}X]^n)$$

Variance

The second central moment of a random variable X is called the *variance* of X

$$\mathrm{Var}(X) = \mathrm{E}(X - \mathrm{E}X)^2 = \mathrm{E}(X^2) - [\mathrm{E}(X)]^2$$

The *standard deviation* of X is the square root of the variance of X

$$\mathrm{SD}(X) = \sqrt{\mathrm{Var}(X)}$$

The variance satisfies the following properties:

- $\mathrm{Var}(X) \geq 0$
- $\mathrm{Var}(X) = 0$ implies that X is constant
- $\mathrm{Var}(aX + b) = a^2\,\mathrm{Var}(X)$ (where a and b are constants)

Skewness and Kurtosis

Recall that μ_n is the n^{th} central moment of a random variable X

- The coefficient of *skewness* of X is

  $$\alpha_3 = \frac{\mu_3}{(\mu_2)^{3/2}}$$

  and measures lack of symmetry of the probability density function

- The coefficient of *kurtosis* of X is

  $$\alpha_4 = \frac{\mu_4}{\mu_2^2}$$

  and measures the peakedness (leptokurtic) or flatness (platykurtic) of the probability density function

Moment Generating Function

Sometimes moments are difficult to compute

$$\mu'_n = \mathrm{E}(X^n) = \int_{-\infty}^{\infty} x^n\, f_X(x)\, dx$$

Alternative: *moment generating function* (mgf)

$$M_X(t) = \mathrm{E}(e^{tX}) = \int_{-\infty}^{\infty} e^{tx}\, f(x)\, dx$$

The moment generating function has the property that

$$\frac{d^n}{dt^n}\big[M_X(t)\big]\bigg|_{t=0} = \mathrm{E}(X^n)$$

Caveat: not all densities have a moment generating function

Moment Generating Function

Let $f_X(x) = \begin{cases} 1 & x \in [0,1] \\ 0 & \text{otherwise} \end{cases}$

Moment Generating Function

$$M_X(t) = \frac{e^t - 1}{t}$$

Change of Variables Formula

Let X be a random variable with probability density function $f_X(x)$ and define a new random variable

$$Y = g(X)$$

where g is a monotonic function

Then the pdf of the random variable Y is

$$f_Y(y) = f_X\left(g^{-1}(y)\right)\left|\frac{d}{dy}g^{-1}(y)\right|$$

where g^{-1} denotes the inverse function of g

Example: Change of Variables Formula: Example

Let X be a continuous random variable with support the positive real line [i.e., $S_X = (0, \infty)$] and define $Y = X^2$

The cumulative distribution function of Y is

$$F_Y(y) = P(Y \leq y)$$

Change of Variables Formula: Example (continued)

Differentiating $F_Y(y)$ yields the density of Y

$$f_Y(y) = \frac{d}{dy} F_Y(y) = \frac{d}{dy} F_X(\sqrt{y}) = \frac{1}{2\sqrt{y}} f_X(\sqrt{y})$$

Alternatively, use the change of variables formula

$$g(x) = x^2 \quad \Rightarrow \quad g^{-1}(y) = \sqrt{y}$$

$$f_Y(y) = f_X(\sqrt{y}) \left| \frac{d}{dy} \left(\sqrt{y}\right) \right| = \frac{1}{2\sqrt{y}} f_X(\sqrt{y})$$

COMPUTATIONAL FINANCE & RISK MANAGEMENT
UNIVERSITY of WASHINGTON
Department of Applied Mathematics

http://computational-finance.uw.edu