계량재무분석 I

Chapter 5 & 6
Probability &
Probability Distribution I

경영대학 재무금융학과
윤선중

Objectives

- 확률의 부여방법
  - 1st Stage: 완전하고 (Exhaustive) 상호배타적인 (Mutually Exclusive) 형태
  - 2nd Stage
    - 고전적 방법 (Classical Approach)
    - 상대도수 방법(Relative Frequency Approach)
    - 주관적 방법 (Subjective Approach)
- 확률의 계산
  - 확률법칙 (Probability Rules)
  - 확률 나무 (Probability Tree)
- Bayes' Law : 조건부 확률의 계산을 도움.

I. Probability

Definition

- 확률 (probability)
  - 어떤 사건이 일어날 가능성을 숫자로 표현한 척도
  - 불확실성(uncertainty) 혹은 랜덤성(randomness)를 연구하기 위한 방법
- Uncertainty vs. Probability
  - (1) 계량재무분석 I 성적의 예
    - A+ ~ F 까지 그 무엇도 받을 수 있는 ________
    - A+ ~F를 받을 수 있는 각각의 가능성: ________
  - (2) 내일 삼성전자 주가의 예
    - 오늘 종가가 50만원이라고 가정
    - 425,000 ~ 575,000원까지 얼마가 될지 정확히 알 수 없는 _______
    - 각각의 가격이 될 가능성: _______
  - 불확실한 상황에서의 의사결정을 위해서는 확률분석이 필요함

Terminology

- 확률실험 (random experiment)
  - 여러가지 가능한 결과들 중 하나의 결과를 발생시키는 활동/과정
    - 앞면(H)과 뒷면(T) 중 하나의 결과가 발생하는 동전 던지기
    - 1,2,…,6 중 하나의 결과가 발생하는 주사위 던지기
    - 0점에서 100점까지 중 하나의 결과가 발생하는 시험평가
    - 0 ~inf 까지 중 하나의 결과가 발생할 수 있는 1년 후의 KOSPI 지수 관찰
- 표본공간(sample space)
  - 확률실험으로부터 발생할 수 있는 모든 가능한 결과들의 집합
    - S={H,T} S={1,2,3,4,5,6}
    - S={1,2,3,…,100} S={x|-inf <x<inf}
  - 표본공간에 포함된 결과들은 반드시 _______(exclusive)하고 _______ (exhaustive) 해야함

Terminology

- 단순 사건 (simple event)
  - 표본공간 상의 개별 결과를 지칭함
    - {H}, {3}, {50}, {3.58%} 등
- 사건(Event)
  - 표본공간에 있는 하나의 단순사건 또는 두 개 이상의 단순사건들의 집합
    - 앞면이 나올 사건: {H}
    - 짝수가 나올 사건: {2,4,6}
    - 90점 이상이 나올 사건: {90,91,…,100}
    - 20% 이상(초과) 상승할 사건: {x|x>20%}

Mathematical Definition

- 확률

  $P{:}\Omega \rightarrow Y$

  - 단, $\Omega = \{X_1, X_2, \ldots, X_n\}, Y = \{y \mid y \in [0,1]$인 실수$\}$

  $(1)\ 0 \le P(X_i) \le 1$

  $(2) \sum P(X_i)$

  $(3) X_i \cap X_j = \phi, \Rightarrow P(X_i + X_j) = P(X_i) + P(X_j)$

- 예제: 주사위 던지기

  $\Omega = \{1, 2, \ldots, 6\}$

  $P(1) = 1/6, \ldots, P(6) = 1/6$

  $\sum P(\cdot) = 1/6$

  $P(\{1, 2\}) = 1/3 = P(1) + P(2)$

Calculation of Probability

- **(1) 수학적 방법 (고전적 방법)**
  - 표본공간을 구성하는 사건들을 모두 정의한 후, 확률의 정의에 따라 계산
- **(2) 통계적 방법 (상대도수 방법)**
  - 무한히 많은 실험/관찰을 반복한 후, 각각의 사건이 발생하는 상대 도수들을 계산
- **(3) 주관적 방법**
  - 해당 사건에 대해 가지고 있는 주관적인 확신에 따라 계산
- 예제
  - 평평한 동전 던지기에서 앞면이 나올 확률?
  - A씨는 동전의 앞면이 나올 확률이 11/20이라고 생각하지만, 실제 1000번을 실험해 보았더니 450번 앞면이 나왔음

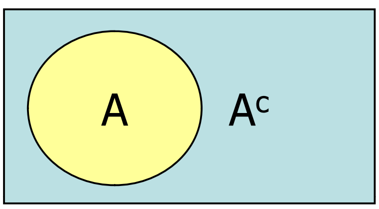
  => 각각의 방법에 의한 확률은?

Interpretation of Probability

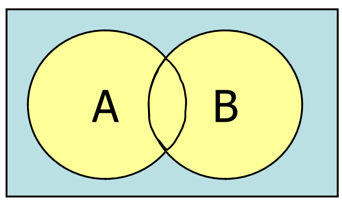
- 어떤 확률 방법을 선택하든 동일한 해석 가능
- 무한히 반복되는 상대도수 방법으로 해석 가능
- 이러한 개념을 이용하여,
  - 모집단과 표본을 연결 시키는 방법 !!!
  - 뒤 에서 배울 것임.

Basic Concept of Probability

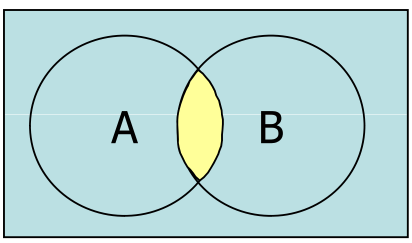
Complement of Event



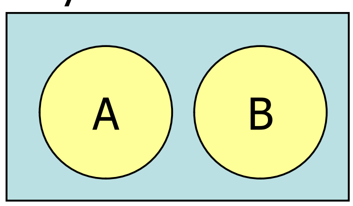
Union of Events



Intersection of Events



Mutually Exclusive Events



Joint Probability

- 정의
  - 결합확률
  - 두 개의 사건이 동시에 발생할 확률 : 교사건이 발생할 확률
  - 교사건 (intersection): $A \cap B$: A와 B가 동시에 발생할 사건
- 예제 5.1 뮤츄얼펀드매니저의 성공 요인 PART 1

	Mutual fund outperforms the market	Mutual fund doesn't outperform the market
Top 20 MBA program	.11	.29
Not top 20 MBA program	.06	.54

Marginal Probability

- 정의
  - 한계확률, 주변확률
  - 개별 사건들의 확률: P(A) 또는 P(M) 등
  - 결합확률 표에서 가장자리를 따라 더하면 한계확률이 됨
- 예제

	B_1	B_2	$P(A_i)$
A_1	.11	.29	.40
A_2	.06	.54	.60
$P(B_j)$	.17	.83	1.00

Conditional Probability

- 정의
  - 조건부 확률
  - 하나의 사건이 발생했다는 전제 하에서 다른 사건이 발생할 확률

$$P(A \mid B) = \frac{P(A\ and\ B)}{P(B)}$$

$$P(B \mid A) = \frac{P(A\ and\ B)}{P(A)}$$

- 예제
  - 앞서의 예제에서,
  - (1) 임의의 한 명을 뽑았을 때 남자라고 하면, 그가 승진하였을 확률은?
    - P(A|M)?
  - (2) 임의의 한 명을 뽑았을 때, 승진한 사람이라면, 그 사람이 남자일 확률은?
    - P(M|A)

Independence vs. Exclusiveness

- 정의
  - 독립사건: 한 사건의 발생확률이 다른 사건의 발생에 영향을 받지 않음
  - 배반사건: 두 사건이 동시에 발생할 가능성이 전혀 없음
  - 독립사건
    - P(A|B)=P(A) 또는 P(B|A)=P(B)
    - $P(A \cap B)=P(A)P(B)$
  - 배반사건
    - $P(A \cap B)=0$
- 예제
  - 주사위를 던져서 짝수의 눈이 나오는 사건과 3의 배수의 눈이 나오는 사건은 서로 독립? 배반 사건?
  - 주사위를 던져서 짝수의 눈이 나오는 사건과 홀수의 눈이 나오는 사건은 서로 독립사건? 배반사건?

Probability Rules – Multiplication Rule

- 정의
  - 두 사건의 결합확률을 계산하기 위하여 사용
  - (1) 만약 두사건 A, B가 독립이 아니면,
    - P(A and B) = P(A | B)•P(B)= P(B | A) • P(A)
  - (2) 만약 두 사건 A, B가 독립이면
    - P(A and B) = P(A)•P(B)
- 예제
  - 계량재무분석 I 수강생은 남학생 7, 여학생 3명
  - 2명의 학생을 임의로 선발하여 장학금 지급
  - 사건 A: 첫 번째 선발된 학생이 여학생일 사건
  - 사건 B: 두번째 선발된 학생이 여학생일 사건
  - 장학금 두번의 수혜자가 모두 여학생일 확률
  - Case 1: 장학금 중복 지급이 되지 않는 경우 Case 2: 장학금 중복 지급이 되지 않는 경우

Probability Rules – Rule of Complement

- 여사건의 정의
  - 사건 A의 여사건: A^c
  - 사건 A가 발생하지 않는 사건
- 여사건의 법칙
  - $P(A)+P(A^c)=1$; $P(A^c)=1-P(A)$
- 예제
  - 주사위를 3회 반복하여 던질 때, 짝수가 적어도 한 번 이상 나타날 확률은?

Probability Rules – Addition Rule

- 정의
  - 사건 A 또는 사건 B가 발생할 확률을 계산하기 위하여 사용
  - 합사건 (A and B): 사건 A 또는 사건 B가 발생할 사건
  - P(A or B)=P(A)+P(B)-P(A and B)
- 예제
  - 동아일도 vs. 경향신문의 두 가지 신문만 존재한다고 가정
  - 도시 가구의 20%는 동아일보 구독
  - 도시 가구의 12%는 경향신문 구독
  - 두 신문을 모두 구독하는 가구는 3%
  - 어느 신문이든 신문을 구독하는 도시가구의 비율은?

Exercise

- 예제 5.5 복원이 없는 경우 두 학생의 선택
  - 남자 7명 여자 3명
  - 두명의 학생을 선택했을 때, 모두 여학생일 확률은?
  - A: 첫 번째 선택이 여학생: P(A) = 3/10 = .30
  - B: 두번째 선택이 여학생: P(B | A) = 2/9 = .22

  - P(A and B) = P(A)•P(B|A) = (3/10)(2/9) = 6/90 = .067
- 복원이 있는 경우 두학생의 선택
  - 독립성 조사 : 0.09

Bayes' Law

- 아래의 경우에 적용

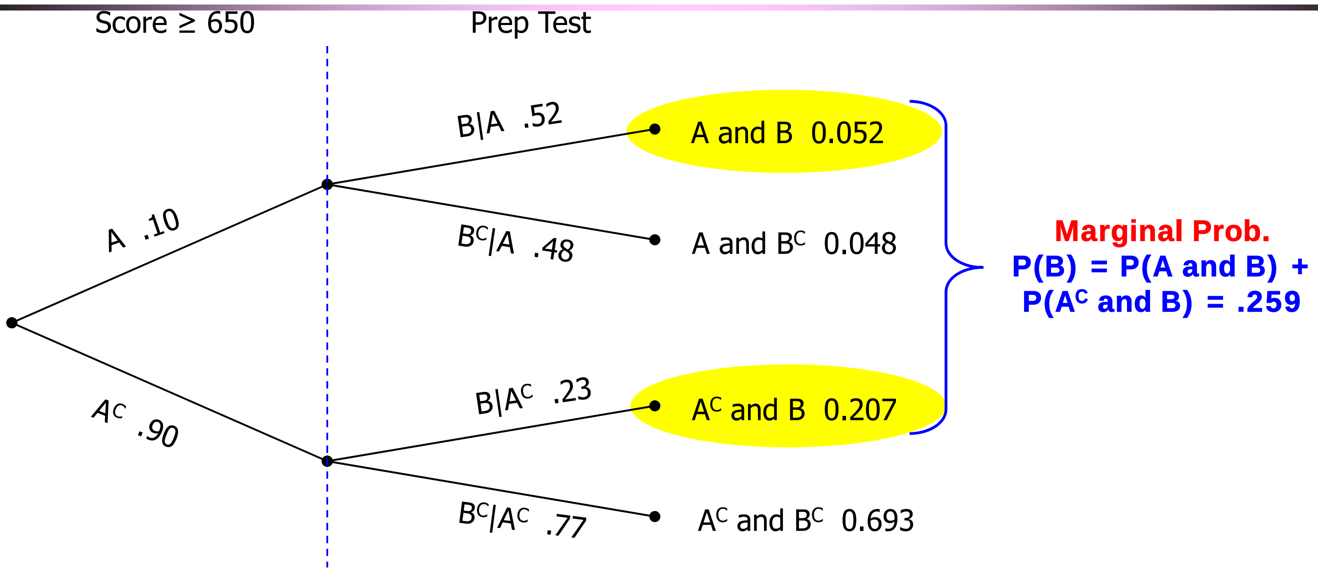
$$P(B \mid A) \Rightarrow P(A \mid B)$$

$$P(A \mid B) = \frac{P(A \text{ and } B)}{P(B)}$$

Example

- 예제 5.9 MBA 지원자는 GMAT 준비과목을 수강하여야 하는가?
  - **A: GMAT 점수가 650점 이상**
  - **A^C: GMAT 점수가 650점 미만**
  - **B: 준비과목을 수강**
  - **B^C: 준비과목을 수강하지 않는다**
  - $P(A) = .10$; $P(A^C) = 1 - .10 = .90$
  - $P(B \mid A) = .52$
  - $P(B \mid A^C) = .23$
  - $P(B^C \mid A) = 1 - .52 = .48$
  - $P(B^C \mid A^C) = 1 - .23 = .77$
  - $P(A \mid B) = ??$

Example



$$P(A \mid B) = \frac{P(A \text{ and } B)}{P(B)} = \frac{.052}{.259} = .201$$

II. Probability Distribution

Introduction to Random Variable

- **확률변수(random variable)의 정의**
  - **확률실험의 각 결과 (outcome)들에 하나의 실수를 부여하는 함수**
    - f: Ω -> R
  - **단정적인 하나의 값이 아닌 여러 개의 값을 확률적으로 가질 수 있음**
- **예제: 동전던지기**
  - **앞면 -> 1, 뒷면 ->0**

Introduction to Random Variable

- **예제**
  - **(1) 동전을 두 번 던졌을 때 앞면이 나온 개수를 확률 변수 X라고 하면,**
    - 정의역: {HH, HT, TH, TT}
    - 치역: {0, 1, 2}
  - **(2) 삼성전자에 1천만원, 현대차에 2천만원을 투자한 포트폴리오의 1년 후 가치를 Y라고 하면,**
    - 정의역: {x | x는 0보다 같거나 큰 실수}
    - 치역: {x | x는 0보다 같거나 큰 실수}
- **확률변수의 종류**
  - **이산확률변수 (discrete random variable) – 치역이 셀 수 있는 집합의 경우**
  - **연속확률변수 (continuous random variable) – 치역이 셀수 없는 실수의 집합**

Introduction to Random Variable

- 확률분포의 정의
  - 확률변수가 가지는 값과 그러한 값이 나타날 확률을 나타낸 표, 그래프 혹은 공식
  - 이산확률분포 vs. 연속확률분포
- 예제
  - (1) 동전을 두 번 던졌을 때 앞면이 나오는 개수의 분포
  - (2) 주사위를 두 번 던졌을 때 나온 눈의 합의 분포

Properties of Probability Distributions

- 이산확률분포

$$1.\ 0 \le \mathrm{P(x)} \le 1 \ for\ all\ x$$

$$2. \sum_{all\ x_i} P(x) = 1$$

  - 확률변수 X가 0~10사이의 정수 값을 가질 수 있다고 하면,
    - P(X=0), P(X=1), …, P(X=10)은 모두 ____과 ___ 사이
    - P(X=0), P(X=1), …, P(X=10)를 모두 더하면 ____이 되어야 함
- 연속확률분포

$$1. \forall x,\ f(x) \ge 0$$

$$2. \int_{\Omega} f(x)dx = 1$$

  - 확률변수 X가 음의 무한대에서 양의 무한대까지 실수값을 가진다면,
    - 임의의 실수 x에 대하여 f(x)는 0보다 같거나 커야함
    - -inf에서 inf까지 f(x)를 적분하면 1이 되어야 함

Example

- 예제 1
  - 세 명이 탈 수 있는 전세 비행기
  - 3명의 고객이 사전 예약
  - 출발 당일 실제로 비행기에 탑승하는 승객의 수를 X
  - 각 고객이 탑승할 확률이 50%라면, X=2일 확률은?
  - X의 확률 분포는?
- 예제 2
  - 휴대폰 충전지는 최소 20시간에서 최대 100시간까지 사용할 수 있음
  - 충전지의 예상 수명이 20시간~100시간 사이에서 얼마가 될지 확률은 모두 같다면
    - 충전지의 예상 수명이 40시간 이상 60시간 미만일 확률은?
    - 충전지의 예상 수명에 대한 확률분포는?

Expected Values & Variances

- 기대값 (expected value / 모평균)

$$E(X) = \mu = \begin{cases} \sum xP(X=x) \\ \int xf(x)dx \end{cases} \Rightarrow \frac{\sum X_i}{N} \text{과 비교}$$

- 분산 (Variance)

$$V(X) = \sigma^2 = \begin{cases} \sum (x-\mu)^2 P(X=x) \\ \int (x-\mu)^2 f(x)dx \end{cases} \Rightarrow \frac{\sum (X_i-\mu)^2}{N} \text{과 비교}$$

  - 참고:

$$V(X) = E(X^2) - E(X)^2 = \begin{cases} \sum x^2 P(X=x) - (\sum xP(X=x))^2 \\ \int x^2 f(x)dx - (\int xf(x)dx)^2 \end{cases}$$

Example

■ 예제 6.1: 가구당 컬러 TV 대수 조사

# of Televisions	# of Households	x	P(x)
0	1,218	0	0.012
1	32,379	1	0.319
2	37,961	2	0.374
3	19,387	3	0.191
4	7,714	4	0.076
5	2,842	5	0.028
	101,501		**1.000**

- 컬러 TV 대수를 X라고 하면 X의 확률 분포는?

Example

- X의 기대값과 분산은?

$$E(X) = \mu = \sum_{all\ x} xP(x) = 0 \bullet P(0) + 1 \bullet P(1) + \ldots + 5 \bullet P(5)$$

=0(.012) + 1(.319) + 2(.374) + 3(.191) + 4(.076) + 5(.028)

= **2.084**

$$V(X) = \sigma^2 = \sum_{all\ x} (x - \mu)^2 P(x)$$

$= (0 – 2.084)^2(.012) + (1 – 2.084)^2(.319)+\ldots+(5 – 2.084)^2(.028)$

= **1.107**

Rules of Expected Value and Variance

- 기대값의 법칙
  - X는 확률변수, c를 임의의 상수라고 할 때,
  - $E(c) = c$
  - $E(X + c) = E(X) + c$
  - $E(cX) = cE(X)$
- 분산의 법칙
  - X는 확률변수, c를 임의의상수라고 할 때,
  - $V(c)=0$
  - $V(X+c)=V(X)$
  - $V(cX) = c^2V(X)$

Bivariate Probability Distribution

- 이변량 확률분포의 정의
  - 두 확률변수의 결합확률(joint prob)을 나타내는 표, 그래프 혹은 공식
- 예제
  - X: 현대 부동산이 한 달에 중개하는 주택의 수
  - Y: 삼성 부동산이 한 달에 중개하는 주택의 수
  - X와 Y는 각각 0~2채
  - 이변량 확률 분포의 예:

		x			
		0	1	2	
y	0	0.12	0.42	0.06	0.6
	1	0.21	0.06	0.03	0.3
	2	0.07	0.02	0.01	0.1
		0.4	0.5	0.1	1.00

  - 각각의 한계확률(marginal probability)은?

Covariance & Correlation Coefficient

- 공분산 및 상관계수

$$COV(X,Y) = \sum_{all\ x} \sum_{all\ y} (x-\mu_X)(y-\mu_Y)P(x,y)$$

$$COV(X,Y) = \sum_{all\ x} \sum_{all\ y} xyP(x,y) - \mu_X \mu_Y$$

$$\rho = \frac{COV(X,Y)}{\sigma_x \sigma_y}$$

- 예제 6.6: 두 부동산 중개인에 의해 판매되는 주택수의 공분산과 상관계수

COV(X,Y) = (0 – .7)(0 – .5)(.12) + (1 – .7)(0 – .5)(.42) + …
… + (2 – .7)(2 – .5)(.01) = **–.15**

$$\rho = \frac{COV(X,Y)}{\sigma_x \sigma_y} = -0.15 \div [(.64)(.67)] = \mathbf{-.35}$$

Rules of Expected Value and Variance

- 두 변수의 합에 대한 기대값과 분산 – 앞 예제
  - P(X+Y=2) = P(0,2) + P(1,1) + P(2,0) = .07 + .06 + .06 = .19

x + y	0	1	2	3	4
P(x+y)	0.12	0.63	0.19	0.05	0.01

  - E(X + Y) = 0(.12) + 1(.63) + 2(.19) + 3(.05) + 4(.01) = 1.2
  - V(X + Y) = (0 – 1.2)2(.12) + … + (4 – 1.2)2(.01) = .56

$$\sigma_{x+y} = \sqrt{\mathrm{Var}(X+Y)} = \sqrt{.56} = .75$$

- **Rules**
  - E(X + Y) = E(X) + E(Y)
  - V(X + Y) = V(X) + V(Y) + 2COV(X, Y)
  - 만약 두 변수가 독립이라면, COV(X, Y) = 0 => V(X + Y) = V(X) + V(Y)

Covariance & Correlation Coefficient

■ 예제

- X:주사위를 던져서 나온 눈을 3으로 나눈 나머지
- Y: 2로 나눈 나머지
- X={0, 1, 2}   Y={0, 1}

		X		
		0	1	2
Y	0			
	1			

- 공분산 및 상관계수?

Application: Chapter 6.3

■ 예제

- 포트폴리오 수익률과 위험의 계산
- McDonald's 주식의 기대수익률은 8%, 표준편차는 12%
- Cisco Systems 주식의 기대수익률은 15%; 표준편차는 22%
- McDonald's 주식에 투자액의 25%, Cisco Systems 주식에 투자액의 75%를 투입한다고 하면
- (1) 포트폴리오의 기대 수익률은?
- (2) 다음과 같은 가정 하에서 포트폴리오 수익률의 표준편차를 계산하면
  - 두 주식의 수익률이 완전한 양(+)의 상관관계를 가짐
  - 두 주식의 수익률간 상관계수는 0.5
  - 두 주식의 수익률은 전혀 상관관계 없음