

## 제 3 강. 난수 생성하기

### Uniform 난수 생성

이제까지 우리는 변수를 선언하는 방법을 알아보았다. 이번 강의는 살짝 VBA에서 벗어나서 난수(Random Number)를 생성하는 방법에 대하여 알아보려고 한다. 독자 중 일부는 이미 엑셀을 이용하여 난수를 만들어낼 수 있다는 사실을 알고 있을 것이다. 바로 rand() 함수를 사용하는 방법인데 rand() 함수는 임의의 0에서 1사이의 값을 우리에게 되돌려 주는 함수이다. 여기에서 '임의'라는 말에는 0에서 1사이의 모든 값에 대하여 균등한 확률로 되돌려 준다는 말이 함축되어 있는 것이므로, rand() 함수는 우리에게 확률변수 Uniform(0, 1)을 따르는 Random Number를 제공해주는 셈이다.

VBA에서도 rand()함수와 똑 같은 기능을 제공하는 함수를 가지고 있는데, 바로 rnd() 함수이다. 자, 한번 앞 강을 복습할 겸 rnd()함수 사용 방법을 알아보자.

#### [Example 3.1]

```
Sub Use_Rnd()
    Dim A As Double
    Randomize
    A = Rnd()

    Sheets(1).Range("A1")(1, 1) = A

End Sub
```

위 코드를 입력하고 F5를 연속해서 누르면서 엑셀 시트를 확인해보면, "A1"셀에 =rand()를 써놓은 것과 같은 효과를 볼 수 있다. Rnd() 함수를 사용하기에 앞서 Randomize라는 구문이 있는데, 이것은 리셋의 효과<sup>1</sup>라고 생각하면 된다. 매번 난수를 발생시키기 전에 Randomize를 써주는 습관을 기르자.

자, 이번에는 조금 더 응용을 해서 "A1"셀에 반환되는 숫자들이 Uniform(3, 5)를 따르게 하기 위해서는 어떻게 해야 할까? 우리는 어떤 확률 변수에 상수를 곱하거나 더하여도 원래의 분포를 변화시키지는 않는다는 사실을 알고 있다.

따라서  $X \sim U(0, 1)$  일 때,  $Y = 2X + 3 \sim U(3, 5)$ 라는 것을 이용하면 Uniform(3, 5)를 따르는 난수를 쉽게 뽑아낼 수 있다.

---

<sup>1</sup> 엑셀에서 Uniform 난수를 발생시키는 방법이 나머지를 이용하는 방법인데, Randomize는 이 때 Seed를 다시 설정해주는 역할을 한다.

**[Example 3.2]**

```

Sub Use_Rnd()
    Dim A As Double
    Randomize
    A = 2*Rnd() + 3

    Sheets(1).Range("A1")(1, 1) = A

End Sub

```

**역 변환법(Inversion Method)**

수리통계 시간에 한번쯤은 배웠으나 난수 발생 방법을 배우기 전까지는 아무도 쳐다보지 않는 정리가 하나 있는데 바로 다음과 같다.

**정리. 확률적분변환**

$X$  는 분포함수와 밀도함수가 각각  $F_X(x)$ ,  $f_X(x)$  인 연속 확률 변수이고,  $F^{-1}(x)$  는 미분 가능할 때,

$$Y = F_X(X)$$

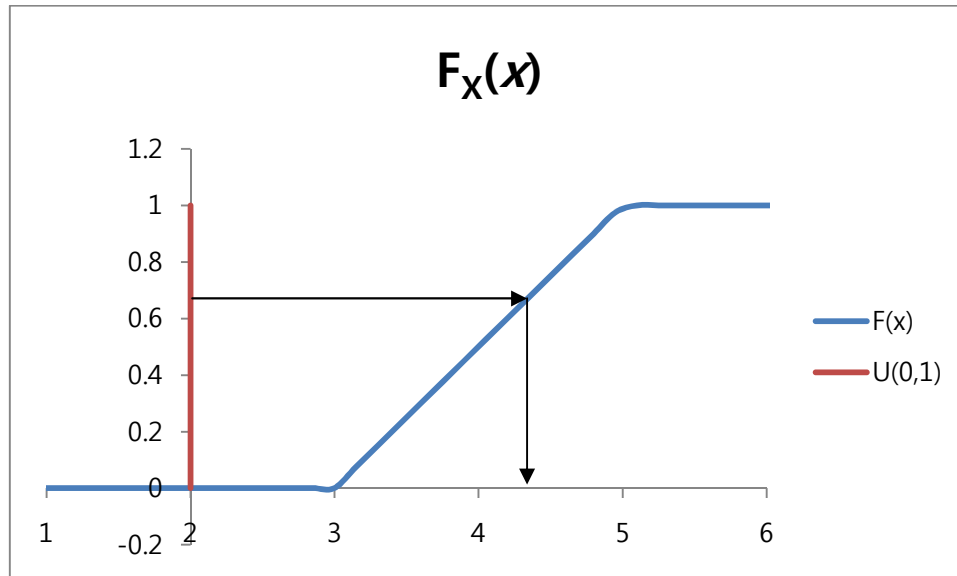
로 정의된 확률변수  $Y$  는 Uniform(0, 1)을 따른다.

뭔가 거창하다!! 하지만 알고 보면 별거 아니라는 사실. 이 별거 아닌 사실로 인해서 우리는 각종 분포를 따르는 난수를 생성해 낼 수 있다! 위의 정리를 수리적으로 증명한다거나 하는 것은 수리통계 시간에 한번씩 해보도록 하고, 무엇을 의미하는지를 생각해 보도록 하자.

정리를 잘 살펴보면  $Y$ 가 Uniform을 따른다는 사실인데, 가만히 생각해 보면 우리는 Uniform(0, 1)을 따르는 난수를 생성하는 방법을 알고 있다. 더군다나  $X$ 의 분포함수의 역함수가 미분가능이니 마니 하는 것을 보니 존재하는 것 같다. 그렇다면, 식 양변에  $F$ 의 역함수를 씌워서 살펴 보도록 하자.

$$F_X^{-1}(Y) = X$$

즉, 우리가 발생시키고 싶어하는 분포함수의 역함수에 Uniform 확률변수를 대입한 결과값이 우리가 원하는 분포를 따르는 난수가 된다는 것을 의미한다. 역시나 말로만 하는 것은 잘 이해가 되지 않으므로! 그림을 그려서 설명해 보겠다. 앞에서의 예제, Uniform(3, 5)를 따르는 난수를 발생시키는 방법을 역 변환법을 사용해 발생하는 방법을 생각해 보자.



파란색으로 그려진 함수가 Uniform(3, 5)의 누적 확률 분포함수를 나타내고 있다. 따라서 F(x)는 0에서 1사이의 값을 갖는다. **여기서 뽀인뜨는 바로!** 왼쪽 빨간 영역에서 임의의 값을 뽑은 후(이게 바로 Uniform(0, 1) 따르는 영역인 것!), 역함수를 통하여 찾은 X축의 값이 Uniform(3, 5) 따르게 되는 것이다.

$X \sim U(3, 5)$  일 때, 누적분포 함수를 구해보면 다음과 같이 나타낼 수 있다.

$$F_X(x) = \begin{cases} 0 & x < 3 \\ 0.5(x-3) & 3 \leq x < 5 \\ 1 & 5 \leq x \end{cases}$$

역함수를 구하는 과정은 다음과 같다.

$$U = F_X(X) = 0.5(X-3)$$

위의 식을 X에 관하여 정리하면, Uniform(0, 1)을 따르는 확률변수 U를 사용하여 X의 난수를 다음 식을 통하여 발생 시킬 수 있다는 결론이 난다.

$$\therefore X = 2U + 3$$

앗, 아까의 예제와 방법이 똑같이 않은가?! 즉, 역변환법은 분포 함수의 역함수를 알고 있을 때 rand() 함수를 통하여 아주 쉽게 발생시켜준다. '아... 각 분포 별 분포함수의 역함수를 어떻게 구하냐 말이다.' 하고 생각하는 독자가 있을지 모르겠다. 하지만 똑똑한 엑셀에서는 감동적이게도 각 분포함수의 역함수 값을 구해주는 함수를 제공한다는 사실. 엑셀 함수 중에서 =NormsInv(), =BetaInv(), =GammaInv() 같은 함수들이 바로 그것이다.

**엑셀 함수를 VBA에서!(WorksheetFunction)**

눈치가 빠른 사람들은 이미 알고 있었을 지도 모르겠지만, 엑셀 함수에서의 난수발생 함수는 =Rand()이고, VB에서는 Rnd()를 사용해야 한다는 사실에서 엑셀이 가지고 있는 함수와 VB가 가지고 있는 함수를 구분해야 한다는 것을 유추할 수 있다. 하지만 앞에서 보았듯, 엑셀이 가지고 있는 저 유용한 역함수들을 VB상에서 못쓰다고 놔두는 것은 너무나 비효율적인 일이다. 이러한 상황을 방지하기 위하여 WorksheetFunction 기능을 이용하여 엑셀의 함수를 VB로 끌어당겨 사용할 수 있다.

**[Example 3.3]**

```
Sub S_Normal_rand()  
    Dim Z As Double  
    Randomize  
    Z = WorksheetFunction.NormSInv(Rnd())  
  
    Sheets(1).Range("A1")(1, 1) = 3 * Z + 2  
  
End Sub
```

위의 예제는 Z에 표준정규분포를 따르는 난수를 받아, 엑셀에는 Normal(2, 3^2)를 따르는 난수를 발생하는 방법이다. 이렇게 엑셀 함수를 쓸 때에는 앞에 WorksheetFunction을 써주어야 한다.

**[HomeWork]**

1. VBA를 사용하여  $X \sim \text{Normal}(3, 4)$ 를 따르는 난수 10개를 발생시킨 후  $E[X]$ 와  $\text{Var}[X]$ 를 추정하여 보자.