

### 이산&연속 확률 분포 1 - Uniform distribution(균등 분포)

어떤 두 점 사이의 구간에서 확률 밀도 함수가 일정하게 나타내는 분포를 균등 분포(Uniform distribution)라 하고,  $X \sim U(a, b)$  로 나타낸다.

**모양의 특성** - 모든 확률 변수에 대해 균일한 확률을 갖기 때문에 수평적으로 되어있다.

**사용처** : 동전 던지기, 주사위 던지기, 난수 생성

**연관된 분포**-이항분포

그래프로 만들 때 필요한 매개변수-dunif(x, min = 0, max = 1, log = FALSE)

### 연속 확률 분포 2 - Exponential distribution(지수 분포)

임의의 양수  $\lambda$ 에 대하여 다음과 같이 확률 밀도 함수  $f(x)$  를 갖는 확률 분포를 모수  $\lambda$ 인 지수 분포(Exponential distribution)

**모양의 특성** - 우측으로 긴 꼬리(tail): 확률값이 감소할수록 긴 꼬리 형태를 나타냄

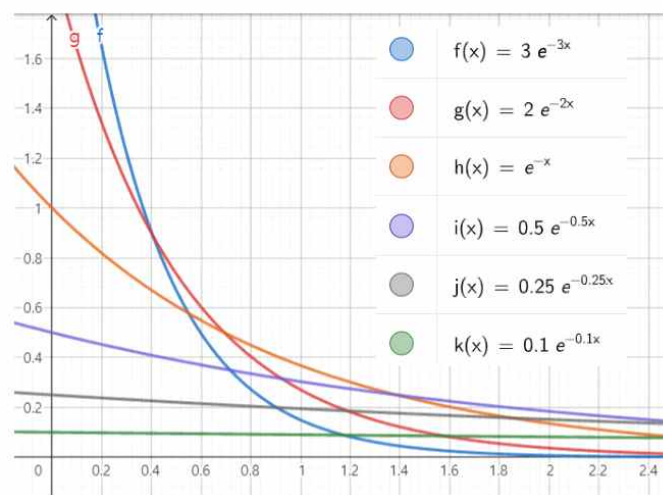
**사용처** - 서비스 대기 시간 분석: 은행에서 고객이 대기하는 시간, 콜센터에서 통화를 기다리는 시간

고장 날 확률 분석: 전자 제품의 수명, 자동차의 고장 확률

**연관된 분포** - 포아송 분포

지수 분포는 사건과 사건 사이 간격의 확률을 계산 그리고, 각 사건은 포아송 분포에 의해 발생

그래프로 만들 때 필요한 매개변수-dexp(x, rate = 1, log = FALSE)



### 연속 확률 분포 3 - Normal Distribution(정규분포)

모양의 특성 - 평균을 기준으로 대칭 (종모양)

사용처 - 자연과학 분야: 지진 발생 간격, 지진 규모  
사회과학 분야: 지역의 인구수, 교육 수준, 소득 수준

연관된 분포-카이제곱분포, 포아송분포, 이항분포, t분포

그래프로 만들 때 필요한 매개변수-평균: `dnorm(x, mean = 0, sd = 1, log = FALSE)`

### 연속 확률 분포 4 - Gamma Distribution(감마 분포)

모양의 특성 - 알파와 람다에 따라 모양이 바뀌지만 기본적으로는 비대칭성을 가지며, 0에 가까워질수록 밀도함수 값은 커지고 무한히 커질수록 0으로 수렴하게끔 오목한 함수형태

사용처

재고 모델링: 제품 생산 시간의 분포가 감마 분포를 따른다고 가정 -> 재고 모델링에 활용

신호 처리: 감마 분포는 신호 처리에서 사용되는 필터링 기법인 감마 필터(Gamma filter)에서 사용

연관된 분포- 지수분포는 첫번째 사건이 발생할 때 까지의 대기시간의 분포인 반면, 감마분포는  $\alpha$ 번째 사건이 발생할 때 까지의 대기시간의 분포, 카이제곱분포

그래프로 만들 때 필요한 매개변수-`dgamma(x, shape, rate = 1, scale = 1/rate, log = FALSE)`

### 연속 확률 분포 5- 카이 제곱 분포(Chi-squared distribution)

모양의 특성 - 왼쪽으로 치우쳐진 비대칭 분포: 카이제곱 분포는 일반적으로 왼쪽으로 치우쳐진 비대칭 분포를 가집니다. 이는 분산 값이 음수가 될 수 없다는 점 때문입니다.

사용처 - 통계학: 분산값의 신뢰구간을 구할 때나, 두 분산값의 비교를 위해 사용됩니다.  
머신 러닝

연관된 분포 - 정규분포, 감마분포, t분포

그래프로 만들 때 필요한 매개변수-`dchisq(x, df, ncp = 0, log = FALSE)`

### 연속 확률 분포 6-t 분포(t-distribution)

모양의 특성- 원점에 대해 좌우 대칭 + 자유도가 증가할수록 표준 정규분포에 가까워짐

사용처 - 어떤 치료법이 효과가 있는지를 검증하는 실험에서 작은 표본 크기를 가지고 가설을 검정할 때

연관된분포- 표준 정규분포, 카이제곱분포, F-분포

그래프로 만들 때 필요한 매개변수-dt(x, df, ncp, log = FALSE)

#### 연속 확률 분포 7- 로그 정규 분포 (Log-Normal)

모양의 특성 - 비대칭성 + 왼쪽으로 치우침

사용처-금융 분야에서의 수익률 분포 : 금융 분야에서는 자산의 수익률을 분석할 때

연관된 분포 - 정규분포

그래프로 만들 때 필요한 매개변수-lnorm(x, meanlog = 0, sdlog = 1, log = FALSE)

#### 연속 확률 분포 8- 베이불분포 (Weibull distribution)

모양의 특성 - 지수분포와 같은 다른 통계적인 분포를 따라간다

사용처 - 부품의 수명 추정 분석

산업 현장에서 어떤 제품의 제조와 배달에 걸리는 시간을 나타낸다.

날씨예보

신뢰성공학에서 실패분석

연관된 분포 -지수분포

그래프로 만들 때 필요한 매개변수-dweibull(x, shape, scale = 1, log = FALSE)

#### 이산확률 분포 1- 포아송분포(Poisson distributio)

모양의 특성 - 오른쪽으로 길게 늘어진 + 비대칭

사용처 - 일정 주어진 시간 동안에 도착한 고객의 수

1km 도로에 있는 흙집의 수

하룻 동안 발생하는 출생자 수

연관된 분포 - 이항분포, 정규분포, 지수분포

그래프로 만들 때 필요한 매개변수-dpois(x, lambda, log = FALSE)

## 이산확률 분포 2 - 이항분포(Binomial Distribution)

모양의 특성 -  $p$ 가 0 또는 1에 가까우면 분포가 왜곡되어 대부분의 확률 질량이 분포의 한쪽 끝에 집중+  $p$ 가 0.5에 가까울수록 분포가 대칭적으로 된다

사용처 - 타율의 0.247인 타자가 5번 타석에서 안타를 칠 확률  
농구골대에서 슛을 5번 넣어 골을 4번 할 확률

연관된 분포 - 베르누이분포, 정규분포, 포아송분포

그래프로 만들 때 필요한 매개변수- dbinom(x, size, prob, log = FALSE)

## 이산확률 분포 3- 초기하분포(hypergeometric distribution)

모양의 특성 -확률 질량의 대부분이 분포의 한쪽 끝에 집중되어 분포가 왜곡 +  $N$ (모집단 크기)이 커질수록 분포가 더 대칭적입니다.

사용처-비교적 작은 모집단에서 비복원으로 추출되는 표본에 사용

연관된분포 - 이항분포, 정규분포

그래프로 만들 때 필요한 매개변수-dhyper(x, m, n, k, log = FALSE)

$m$  : the number of white balls in the urn.

$n$  : the number of black balls in the urn

## 이산확률분포 4 - 베르누이분포((Bernoulli distribution))

모양의 특성- 0과 1에 대한 값밖에 존재하지 않는다.

사용처-성공과 실패의 두 가지 중 하나로만 나오는 실험을 반복적으로 수행할 때

연관된 분포- “이항 분포”는  $n$ 개의 독립적이고 동일하게 분포된 베르누이 확률 변수의 합  
+기하분포

그래프로 만들 때 필요한 매개변수-dbinom(x, size, prob, log = FALSE)

## 이산확률분포 5 - 기하분포

**모양의 특성**- 성공 확률이 클 수록 초반에 '성공'이 나타날 확률이 크고, 성공 확률이 작을수록 초반에 '성공'이 나타날 확률이 작다.

**사용처**- 베르누이 시행을 반복할 때 처음으로 알고자 하는 사건을 관찰에 성공하기 까지의 시도 횟수를 확률변수로 예를 들어 연애에서 결혼까지 이어질 확률이 10%라면 x번째 연애에 결혼하게 되는 것을 구하고 싶을 때 사용한다.

**연관된 분포**-음이항분포, 이항분포, 베르누이 분포

**그래프로 만들 때 필요한 매개변수**-`dgeom(x, prob, log = FALSE)`

## 이산확률분포 6 - 음이항분포

**모양의 특성**-왼쪽으로 치우쳐진 비대칭 분포

**사용처**-이항분포는 성공할 확률이  $p$ 인 베르누이 시행을  $r$ 번 성공할 때 까지 시행하는 경우 실패(시행)횟수를 구할 때 사용

**연관된분포**-포아송 분포 , 이항 분포, 기하분포

**그래프로 만들 때 필요한 매개변수**-`dnbinom(x, size, prob, mu, log = FALSE)`