주요 연속형 확률분포함수

Z_α $N[0,1]$ — χ^2 분포, t 분포, F 분포 — 정의, E, V, 성질

- 카이제곱(χ^2) 분포 (chi-squared distribution)
  - 카이제곱 확률변수와 확률밀도함수
    - $Z_1, Z_2, \dots, Z_k$가 k개의 서로 독립인 표준정규 확률변수 $(Z_i \overset{iid}{\sim} N(0,1),\ i = 1,2,\dots,k)$일 때, (1-α)분위수
      - ► 확률변수 $X = Z_1^{\ 2} + Z_2^{\ 2} + \cdots + Z_k^{\ 2}$는 자유도가 k인 카이제곱 분포를 따르는 것으로 정의.

indep.

    - 자유도 k인 카이제곱분포의 확률밀도함수는 다음과 같고, 이 경우, $X \sim \chi^2(k)$라고 함.

$$\int_0^\infty x^2 f(x)\,dx$$

$$f(x) = \frac{1}{\Gamma\left(\frac{1}{2}\right) 2^{\frac{k}{2}}} x^{\frac{k}{2}-1} e^{-\frac{x}{2}}, \quad 0 \le x < \infty$$

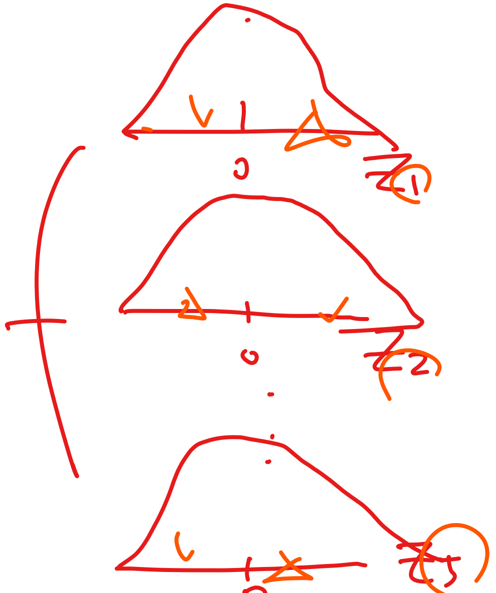
$\chi^2[5]$

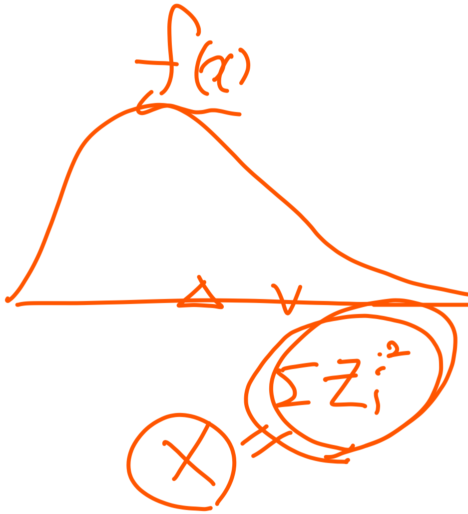
  - 카이제곱 분포의 평균과 분산

    $X \sim \chi^2(k)$인 경우,
    - $E[X] = k$
    - $V[X] = 2k$

$E(X^2) - k^2$

자유도



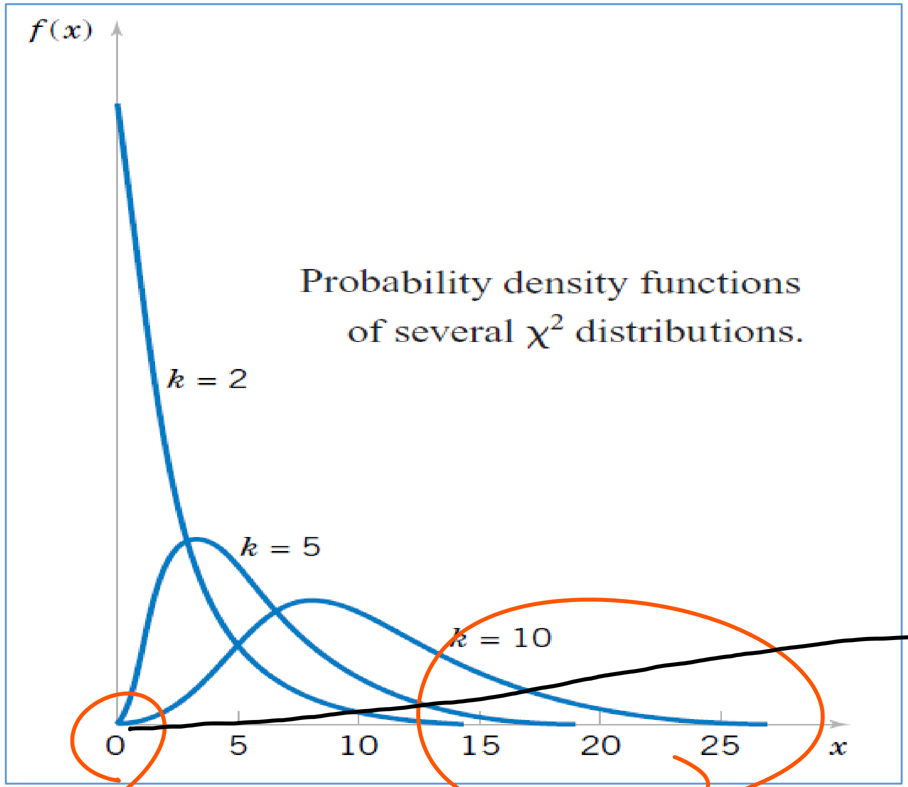


주요 연속형 확률분포함수

- **카이제곱 분포 확률밀도함수 개형**
  - 오른쪽으로 치우친 비대칭 구조이지만 자유도 k가 커질수록 정규분포와 비슷하게 평균에 대칭인 모양이 됨.

Positive (right) skewed.

정규분포의 근사.



주요 연속형 확률분포함수

- **카이제곱 분포의 가법성**
  - 확률변수 U는 자유도가 k_1인 카이제곱 분포를, 확률변수 V는 자유도가 k_2인 카이제곱 분포를 따르며, U와 V는 서로 독립이라고 할 때,
  - ► $U+V$는 자유도가 k_1+k_2인 카이제곱 분포를 따른다.

$U \sim \chi^2[k_1]$ $\quad Z_1^2+\cdots+Z_{k_1}^2$ 독립

$V \sim \chi^2[k_2]$ $\quad Y_1^2+\cdots+Y_{k_2}^2$ 독립

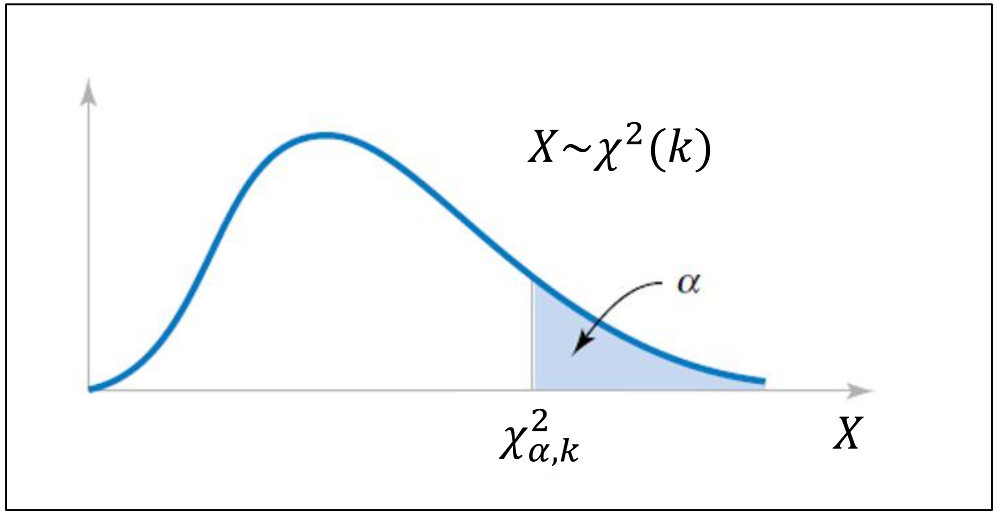
U, V 독립.

$\Rightarrow X(=U+V) \sim \chi^2[k_1+k_2]$

$Z_1^2+\cdots+Z_{k_1}^2+Y_1^2+\cdots+Y_{k_2}^2$

주요 연속형 확률분포함수

- **카이제곱 확률변수의 $(\mathbf{1}-\boldsymbol{\alpha})$분위수 : $\boldsymbol{\chi^2_{\alpha,k}}$**
  - $X \sim \chi^2(k)$일 때, $P[X > c] = \alpha$를 만족하는 X의 $(1-\alpha)$분위수 c를 $\chi^2_{\alpha,k}$으로 표기함.



주요 연속형 확률분포함수

- 예제
  - 다음의 $\chi^2_{\alpha,k}$를 구하여라.
    - $\chi^2_{0.05,8}$
    - $\chi^2_{0.95,8}$

  - $X \sim \chi^2(16)$이라고 할 때, 다음을 만족하는 b, c를 구하여라..
    - $P(X < b) = 0.10.$
    - $P(X < c) = 0.95.$

주요 연속형 확률분포함수

```
> qchisq(0.95, df=8 )
[1] 15.50731
> qchisq(0.05, df=8 )
[1] 2.732637
>
> qchisq(0.10, df=16)
[1] 9.312236
> qchisq(0.95, df=16)
[1] 26.29623
```

주요 연속형 확률분포함수

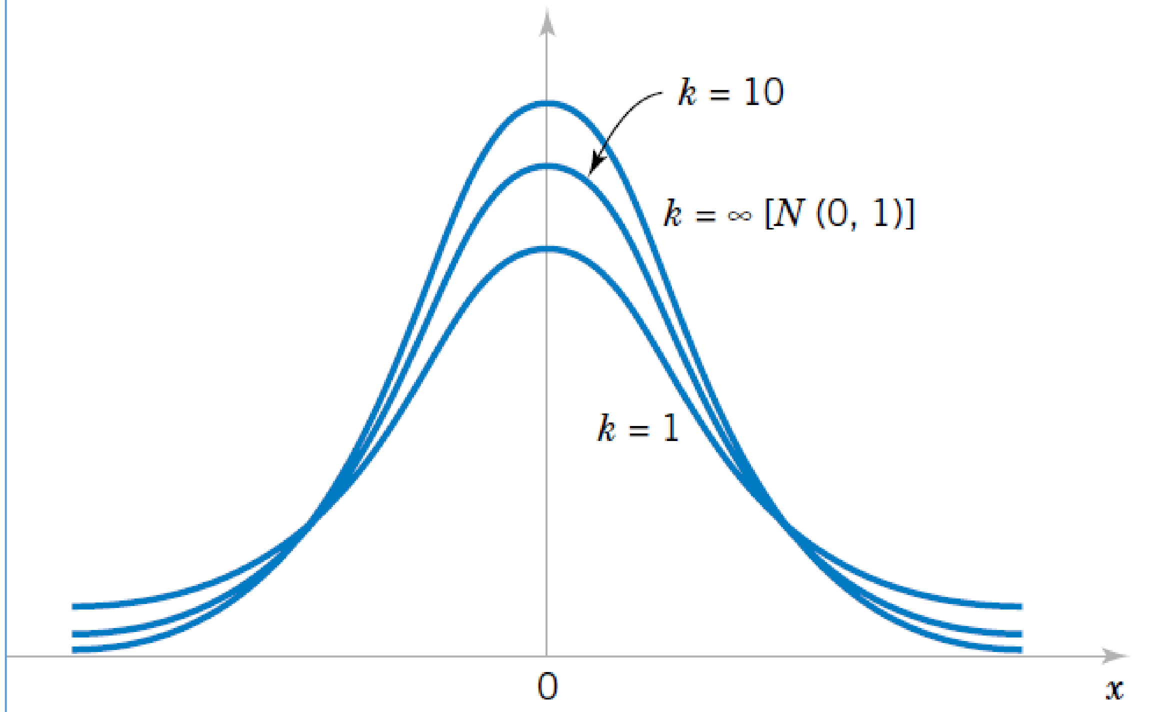
- **t 분포 (t-distribution)**
  - **t 확률변수와 확률밀도함수**
    - Z가 표준정규 확률변수 $Z \sim N(0,1)$이며,
      X가 자유도가 k인 카이제곱 확률변수 $X \sim \chi^2(k)$ 이며,
      Z와 X는 서로 독립이라고 할 때,
      ► 확률변수 $T = \frac{Z}{\sqrt{X/k}}$는 자유도가 k인 t 분포를 따르는 것으로 정의.
    - 자유도가 k인 t 분포의 확률밀도함수는 다음과 같고, 이 경우 $T \sim t\ (k)$라고 함.

$$f(t) = \frac{\Gamma\left(\frac{k+1}{2}\right)}{\Gamma\left(\frac{k}{2}\right)} \frac{1}{\sqrt{k\pi}} \frac{1}{\left(1 + \frac{t^2}{k}\right)^{\frac{k+1}{2}}}, \quad -\infty < t < \infty$$

  - **t 분포의 평균과 분산**
    $X \sim t\ (k)$인 경우
    - $E[X] = 0$
    - $V[X] = \frac{k}{k-2}$ (단, $k > 2$)

주요 연속형 확률분포함수

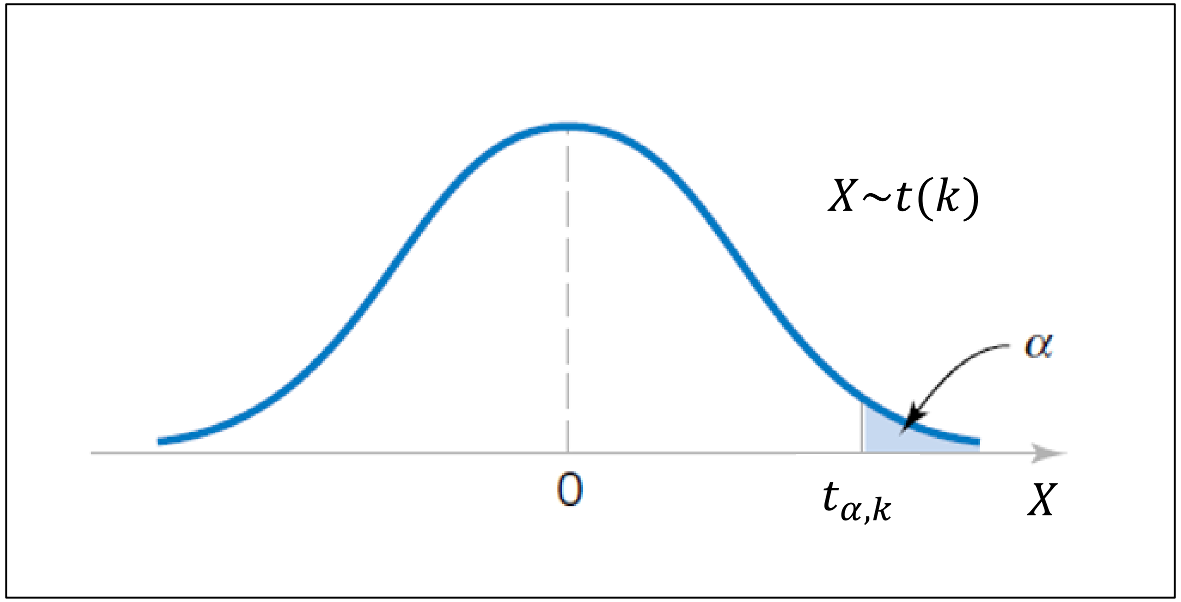
- **t 분포 확률밀도함수 개형**
  - 표준정규분포처럼 0을 중심으로 대칭이지만 표준정규분포보다 꼬리가 더 두꺼움. 자유도 k 가 커질수록 표준정규분포의 밀도함수에 근사하게 됨.



Probability density functions of several t distributions.

주요 연속형 확률분포함수

- **t 확률변수의 $(1-\boldsymbol{\alpha})$분위수 : $\boldsymbol{t_{\alpha,k}}$**
  - $X \sim t\ (k)$일 때, $\mathrm{P}[X > c] = \alpha$를 만족하는 X의 $(1-\alpha)$분위수 c를 $t_{\alpha,k}$으로 표기함.



주요 연속형 확률분포함수

- 예제
  - 다음의 $t_{\alpha,k}$를 구하여라.
    - $t_{0.05,25}$
    - $t_{0.025,15}$
    - $t_{0.95,10}$
    - $t_{0.975,8}$

주요 연속형 확률분포함수

```
> qt(0.95, df=25 )
[1] 1.708141
> qt(0.975, df=15 )
[1] 2.13145
> qt(0.05, df=10)
[1] -1.812461
> qt(0.025, df=8)
[1] -2.306004
```

주요 연속형 확률분포함수

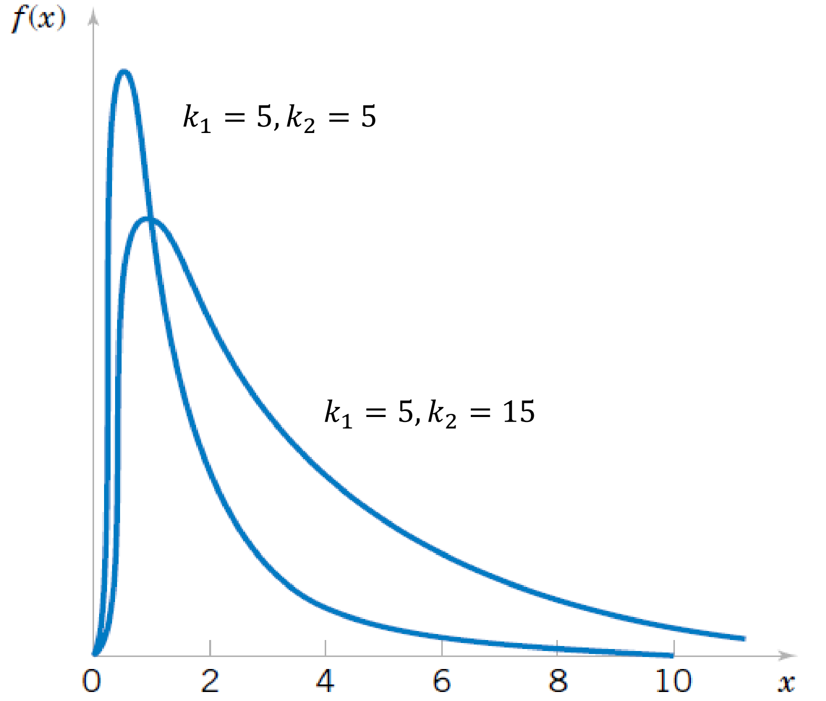
- **F 분포 (F distribution)**
  - **F 확률변수와 확률밀도함수**
    - U가 자유도가 k_1인 카이제곱 확률변수($U \sim \chi^2(k_1)$)이며,
      V가 자유도가 k_2 인 카이제곱 확률변수($V \sim \chi^2(k_2)$) 이고,
      U와 V는 서로 독립이라고 할 때,
      ► 확률변수 $X = \frac{U/k_1}{V/k_2}$ 는 자유도 k_1, k_2 인 F 분포를 따르는 것으로 정의.

    - 자유도 k_1, k_2 인 F 분포의 확률밀도함수는 다음과 같고, 이 경우, $X \sim F(k_1, k_2)$ 라고 함.

$$f(x) = \frac{\Gamma\left(\frac{k_1+k_2}{2}\right)}{\Gamma\left(\frac{k_1}{2}\right)\Gamma\left(\frac{k_2}{2}\right)}\left(\frac{k_1}{k_2}\right)^{\frac{k_1}{2}} x^{\frac{k_1}{2}-1}\left(1+\frac{k_1}{k_2}x\right)^{-\frac{1}{2}(k_1+k_2)}, \quad 0 \le x < \infty$$

주요 연속형 확률분포함수

- **F분포 확률밀도함수 개형**
  - 오른쪽으로 치우친 비대칭 구조임.



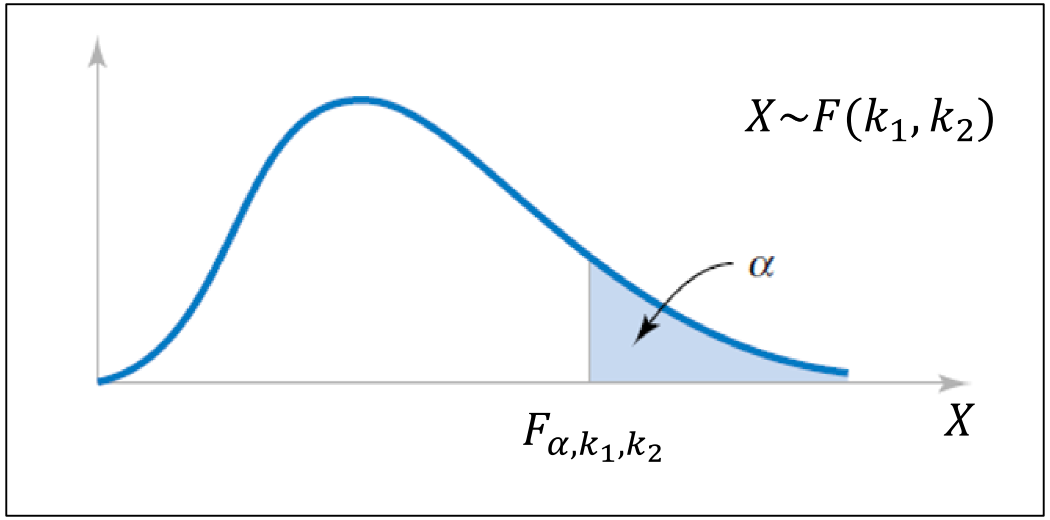
Probability density functions of two F distributions.

- **F분포의 평균과 분산**

  $X \sim F(k_1, k_2)$ 인 경우
  - $E[X] = \frac{k_2}{k_2 - 2}$
  - $V[X] = \frac{2k_2^2(k_1 + k_2 - 2)}{k_1(k_2 - 2)^2(k_2 - 4)}$

주요 연속형 확률분포함수

- **F 확률변수의 $(1-\alpha)$분위수 : F_{α,k_1,k_2}**
  - $X \sim F(k_1, k_2)$ 일 때, $\mathrm{P}[X > c] = \alpha$를 만족하는 X의 $(1-\alpha)$분위수 c를 F_{α,k_1,k_2}으로 표기한다.



주요 연속형 확률분포함수

- 예제
  - 다음의 F_{α,k_1,k_2}를 구하여라.
    - $F_{0.05,5,7}$
    - $F_{0.95,4,8}$

주요 연속형 확률분포함수

```
> qf(0.95, df1=5, df2=7 )
[1] 3.971523
> qf(0.05, df1=4, df2=8 )
[1] 0.1655343
```

주요 연속형 확률분포함수

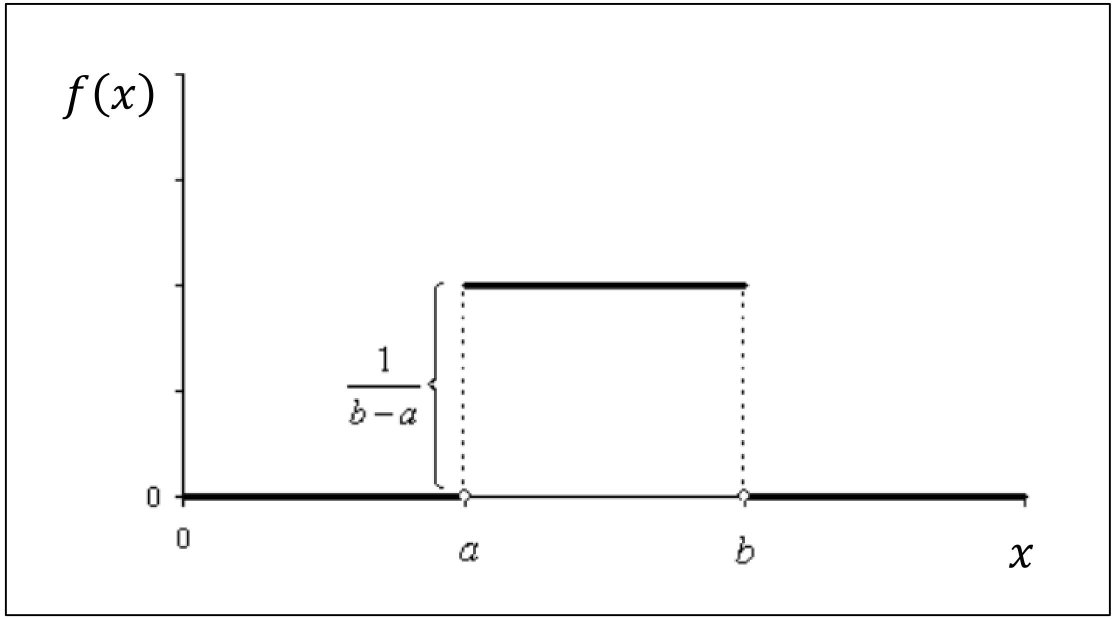
- **균일분포 (uniform distribution)**
  - 균일확률변수와 확률밀도함수
    - 확률변수 X가 실구간 (a, b)에서 균일하게 분포되어 있을 때, 그 확률밀도함수와 분포함수는 아래와 같이 주어짐.

$$f(x) = \begin{cases} \dfrac{1}{b-a}, & a \le x \le b \text{ 인경우} \\ 0, & \text{그외의 경우} \end{cases}$$

    - 이 경우 $X \sim U(a, b)$라고 함.

주요 연속형 확률분포함수

- **균일분포의 확률밀도함수 개형**



- 균일분포의 특성치

  $X \sim U(a,b)$인 경우

  - $E[X] = \frac{a+b}{2}$
  - $V[X] = \frac{(b-a)^2}{12}$

주요 연속형 확률분포함수

- 예제
  - 어느 기업의 주식의 수익률이 하루 동안 -1%~9%까지 변할 수 있고 이 수익률의 변화는 연속형 균일분포를 따른다고 한다. 하루동안 이 주식의 수익률의 변화가 -1%~1% 사이에 있을 확률은 얼마인가?