

6. R

6.1 apply 함수

- 계열 함수를 이용하면 반복문 사용보다 코드 최소화, 수행 속도가 빠르다.

1) apply : 행렬이나 배열 차원별로 지정함 함수에 적용한다.

2) lapply : list apply (함수를 적용한 결과를 list로 반환)

3) sapply : lapply와 유사하나 결과를 리스트 대신 행렬, 벡터로 반환한다.

4) vapply : sapply와 유사하나 FUN의 모든 값이 FUN.VALUE와 호환되는지 확인한다.

5) mapply : sapply와 유사하나, 여러 개 인자를 함수에 전달할 수 있다.

6) tapply : 그룹별 처리를 위한 apply 함수다. (대상은 list)

7) by : 데이터프레임 대상으로 사용한다. (tapply함수)

6.2 doBy 패키지 함수

1) summaryBy : 그룹별로 그룹을 특정 짓는 통계적 요약 값에 사용한다.

2) orderBy : 데이터 프레임의 특정 변수로 데이터 프레임 행을 정렬한다.

3) sampleBy : 데이터 프레임은 포물러 변수에 따라 분할되어 특정 비율의 샘플을 추출한다.

6.3 계통추출

- 계통추출은 체계적 표집이라고 하며 첫번째 요소를 선정 후 그 샘플로부터 동일한 간격의 데이터를 샘플로 추출하는 방법이다.