**오차 연산을 통한 결과 검증**

회귀를 통하여 독립변수와 종속변수 간의 연관성을 나타내는 함수가 완성되었다면 그 결과값을 확인하여 함수의 정확도를 검증할 필요가 있다. 이번 프로젝트에서는 데이터셋에서 회귀에 사용된 독립변수만을 뽑아내고 함수에 대입하여 계산된 결과값과 데이터셋에서 종속변수로 사용된 값을 비교하여 이를 검증하였다. 검증에 사용되는 데이터의 개수는 사용자의 입력에 따라 최대 1880 개를 포함할 수 있다.

검증은 데이터마다 데이터셋의 종속변수 값과 함숫값의 절대오차를 logscale로 계산하여 이들의 평균을 구하는 방식으로 수행된다. 오차 계산 공식을 선택함에 있어서, 프로그램에 사용된 데이터셋 각 열의 값이 종류마다 단위가 달라 그 절댓값도 크게 차이가 나기 때문에 수치화 할 때 전체적인 평준화가 이루어지면서 직관성을 유지하도록 logscale을 적용하였다. 검증에 사용된 공식은 다음과 같다.

n은 사용자가 입력한 검증용 데이터의 크기이다. Yi는 i번째 검증용 데이터에서 종속변수의 값이다. F는 다중 회귀를 통하여 완성한 함수이고 xij는 i번째 검증용 데이터에서 독립변수 j 의 값이다. 수식을 계산함으로써 우리는 다중 회귀로 완성한 함수가 얼마나 실제 데이터의 경향성을 잘 나타냈는지 수치화하여 확인할 수 있었다.