**[2014년 회귀분석실습 중간고사] 2014.10.15**

**\*\*\*\*모든 답은 답안지에 써서 낼 것. (번호 명시 확실히)**

**\*\*\*\*본 시험은 실습시험이므로, 본인이 답안지에 쓴 코드가 실제로 R에서 실행되지 않을 시 점수를 드릴 수 없습니다.**

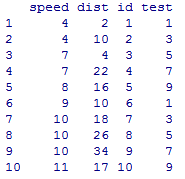
**\*\*\*\*중간에 못 푼 문제가 있더라도, 데이터 명은 문제에 맞게 설정하셔야 합니다.**

**1. “data(cars)”를 실행하고 다음 문항들에 답변하시오.**

1) 아래의 두 벡터를 생성하여, 새로운 컬럼으로 cars데이터에 붙여 cars2를 생성하시오

- cars의 row수를 확인하고, row수가 있는 만큼 1부터 연속된 수를 같는 벡터 (변수명=id)

- 10이하의 홀수가 반복되는 벡터 (변수명 test)



2) cars2에서 dist의 평균을 계산하고, dist의 평균보다 크면 1, 작으면 0을 나타내는 변수 z를 생성하여 새로운 컬럼으로 cars2데이터에 붙여 cars3를 생성하시오.

3) cars3와 제공된 csv파일 “id.csv”를 병합하여 cars4를 생성하시오. (기준변수= id, no)

4) cars4를 다음 순으로 동시에 정렬하시오.

1. test : 오름차순 2.x : 내림차순

5) cars4에서 z값에 따른 speed의 평균값을 구하여라. (코드와 결과값 모두 제시)

6) cars4에서 z와 test 간의 이원분할표를 구하시오. (코드와 결과값 모두 제시)

7) cars4 데이터를 z를 기준으로 두 개의 데이터 셋 part1(z=1)과 part2(z=0)으로 분할 하시오.

8) 분할 한 두 데이터의 speed의 분포를 비교하려 한다. 각 데이터의 speed에 대한 상대도수 히스토그램을 한 화면 안에 그리고 비교하시오. (density라인 추가)

9) 이번에는 z를 기준으로 x의 분포를 비교하려 한다. 상자 그림을 그려 비교하시오.

(제목 : “Do they have the same distribution?”)

**2. 주어진 DATA ex2.txt(신혼부부의 키)를 사용하여 다음에 답하시오.**

**(Y : 부인의 키, X : 남편의 키)**

**\*\*\*\*\*\*\* 모든 계산 과정에 필요한 요소들은 R의 output을 불러와 사용할 것!**

1) 회귀분석을 하기 전 확인해야 할 2가지를 R code를 사용하여 보이시오. (그림, 수치)

2) 과 에 대하여 99% 신뢰구간을 구하여라.

3) 반응변수를 이용하여 기울기가 0.5보다 크다는 귀무가설을 95% 유의수준에서 검정하여라.

4) 어떤 신혼부부중 남편의 키가 187이라고 하자. 이 신혼부부의 예측된 아내의 키에 대하여 구간추정치를 95% 신뢰수준 하에서 제시하여라.