Instructions

1. 데이터 분석 시험은 3시간 동안 진행되며 총 150점 만점으로 점수가 매겨집니다.

2. 세부 과정 / 대략적인 과정을 수행하기 위한 **세부 워크시트(Detailed Worksheets)**가 있습니다. 각 **세부 워크시트에는** 다음을 기입하세요.

- 학생 코드
- 문제 번호
- 장 번호 및 총 페이지 수

3. 각 문제는 세부 워크시트의 새 페이지에서 시작하세요. 종이의 인쇄된 면에만 답안을 작성하십시오. 뒷면을 사용하지 마십시오. 채점되고 싶지 않은 내용을 종이에 적었다면, 줄로 그어 지우십시오.

4. 답안에는 그래프 용지가 필요합니다. 각 그래프 용지 시트에 다음을 기입하십시오.

- 학생 코드
- 문제 번호.
- 그래프 번호 및 사용된 그래프 용지의 총 개수.

5. 여러분의 최종 답안을 위해 학생 ID 코드가 포함된 요약 **답안지(summary Answer Sheet)가** 있습니다.

6. 채점자가 여러분의 언어를 이해하지 못할 수도 있습니다. 가능한 한 수학적 표현과 숫자만을 사용하여 답안을 작성하십시오. 글로 설명해야 하는 경우 짧은 문장(가능하면 영어로)을 사용하세요.

7. 허가 없이 시험 책상을 떠날 수 없습니다. 도움이 필요한 경우(계산기 오작동, 화장실 방문, 세부 워크시트가 더 필요한 경우 등) 손을 들어 감독관에게 신호를 보내주세요.

8. 시험 시작과 종료는 긴 소리 신호로 알려드립니다. 또한, 대회 종료 15분 전(마지막 긴 신호음이 울리기 전)에는 짧은 신호음이 울립니다.

9. 시험가 끝나면 즉시 답안 작성을 중단해야 합니다. 요약 답안지, 그래프 용지, 세부 워크시트를 순서대로 정리하고 한 스택에 넣습니다. 다른 모든 종이들은 다른 스택에 넣습니다. 시험장 밖으로 어떤 종이도 가져갈 수 없습니다.

10. 봉투가 수거될 때까지 여러분의 책상에서 기다립니다. 모든 봉투가 수거되면 학생 가이드가 시험장 밖으로 안내할 것입니다.

D2. 샤플리 가설 (75점)

구상 성단은 은하의 가장 오래된 구성 요소 중 하나이다. 약 한 세기 전, 할로우 샤플리는 태양에서 은하 중심까지의 거리를 결정하기 위해, 은하 중심 주위에 구상성단이 대칭적으로 분포한다는 가설을 세우고 우리 은하의 구상 성단 분포를 연구했다. 아래 표는 우리 은하에서 알려진 몇 개의 구상 성단의 위치와 거리 지수를 보여준다. 표의 처음 세 열은 성단 이름, 은하 경도*(은경, l)*, 은하 위도*(은위, b)*를 나타낸다. 네번째 열은 거리지수(즉, 겉보기 등급와 절대 등급의 차이)를 표시하며, 이 값은 소광 보정된 값이다. 표의 데이터:

이름	l (도)	b (도)	거리지수(등급)
NGC 6522	1.025	-3.926	14.3
NGC 6401	3.450	3.980	14.4
NGC 6342	4.898	9.725	14.5
NGC 6553	5.253	-3.029	13.6
NGC 6440	7.729	3.801	14.6
Ter 12	8.358	-2.101	13.6
VVV-CL160	10.151	0.302	14.2
2MASS-GC01	10.471	0.100	12.6
NGC 6517	19.225	6.762	14.8
NGC 6402	21.324	14.804	14.8
NGC 6712	25.354	-4.318	14.3
NGC 6426	28.087	16.234	16.6
NGC 5466	42.150	73.592	16.0
NGC 7089	53.371	-35.770	15.3
NGC 288	151.285	-89.380	14.8
NGC 2298	245.629	-16.006	15.0
NGC 4590	299.626	36.051	15.1
NGC 4372	300.993	-9.884	13.8
NGC 362	301.533	-46.247	14.7
BH 140	303.171	-4.307	13.4
NGC 5927	326.604	4.860	14.6
Patchick 126	340.381	-3.826	14.5
NGC 5897	342.946	30.294	15.5
NGC 6380	350.182	-3.422	14.9
Djor 1	356.675	-2.484	15.0

(a) (25점) 태양으로부터 각 구상 성단의 거리(파섹 단위)와 구상성단의 직교 좌표 (x, y, z)를 계산하시오. 이때 x-축은 은하 중심을 가리키고 y-축은 은하 자전 방향을 향한다. 이 좌표계는 오른손 법칙을 따른다. (right-handed)

(b) (15점) 주어진 자료를 이용하여, 태양으로부터 구상성단 분포의 중심까지의 거리를 추산하고 그 불확도(uncertainty)를 구하라.

(c) (30점) 구상 성단이 은하 중심 주위에 대칭적으로 분포한다는 샤플리의 가설의 타당성을 테스트하기 위해, x, y, z 방향으로의 각 분포 마다 5개의 구간을 가진 히스토그램들을 작성하시오(즉, 데이터를 정렬하고 데이터를 5개의 동일한 크기의 구간으로 나눈다는 것을 의미함). 세 개의 분포들의 사분위수 값(Q_1, Q_2, Q_3)을 각 히스토그램에 실선으로 표시하시오.

힌트: 세 개의 사분위수는 정렬된 표본을 네 개의 구간으로 나누며, 각 구간은 전체 데이터의 25%씩을 포함한다. 이때 두 번째와 세 번째 구간이 자료 중심분포를 포함하는 사분범위(Interquartile Range)를 나타낸다.

(d) (5점) 사분위수를 사용하여 주어진 세 분포의 대칭 척도 값을 다음과 같이 계산하시오:

$$\Phi_x = \frac{|Q_{1,x}+Q_{3,x}-2Q_{2,x}|}{Q_{3,x}-Q_{1,x}}, \Phi_y = \frac{|Q_{1,y}+Q_{3,y}-2Q_{2,y}|}{Q_{3,y}-Q_{1,y}}, \Phi_z = \frac{|Q_{1,z}+Q_{3,z}-2Q_{2,z}|}{Q_{3,z}-Q_{1,z}}$$

아래 표의 정보와 계산한 대칭 척도 값에 근거하여 x, y, 그리고 z 방향으로의 세 분포 유형을 분류하시오. 이에 따라서, 분석된 표본이 샤플리의 가설을 따르는 경우 답안지에 참**(T)**을, 그렇지 않은 경우 거짓**(F)**을 적으시오.

대칭 척도 값	대칭 유형
$0.0 \leq \Phi \leq 0.1$	대칭 (symmetrical)
$0.1 < \Phi \leq 0.2$	준대칭 (quasi-symmetrical)
$\Phi > 0.2$	비대칭 (asymmetrical)