

문제 1 파이썬 통계 기본 (10점)

A = [9, 17, 11, 22, 12, 26, 13, 23, 4, 8, 8, 15]

1) 위 파이썬 코드의 변수 A에 정수들이 저장되어 있다. 이 숫자들의 사분위수범위를 구하는 파이썬 코드를 작성하시오. (5)

```
import statistics
import scipy
import numpy
A = [9, 17, 11, 22, 12, 26, 13, 23, 4, 8, 8, 15]
print("사분위수범위 = ", numpy.quantile(a, .75) - numpy.quantile(a, .25))
```

2) A의 중간값을 구하는 파이썬 코드를 작성하시오. (5)

```
import statistics
import scipy
import numpy
A = [9, 17, 11, 22, 12, 26, 13, 23, 4, 8, 8, 15]
print("중간값 = ", statistics.median(a))
```

문제 2 (25점)

1) 주사위를 던지는 실험에서 A는 짝수가 나오는 사건, B는 2의 배수가 나오는 사건, C는 3보다 작은 수가 나오는 사건이라고 할 때, 아래 확률을 구하시오. (15)

$P(A \cap B) = \{2, 4, 6\}$

$P(A \cup C) = \{1, 2, 4, 6\}$

$P(A \cap B \cap C) = \{2\}$

$A=\{2,4,6\}, B=\{2,4,6\}, C=\{1,2\}$ 임

(2) 어떤 사람이 검은색과 흰색의 셔츠를 가지고 있는데, 매일 아침 $4/5$ 정도는 검은색 셔츠를 입고, $1/5$ 정도는 흰색 셔츠를 입는다. 이 사람이 검은색 셔츠를 입었을 때는 $4/5$ 정도 넥타이를 매고, 흰색 셔츠를 입었을 때는 $1/4$ 정도 넥타이를 맨다고 하자. 어느날 이 사람이 넥타이를 매다면, 이 사람이 검은색 셔츠를 입었을 확률을 구하시오. (10)

사건A

아침에 검은색셔츠를 입는 사건

$$P(A) = 4/5$$

사건 B

넥타이를 맨 사건

$$P(B | A) = 4/5, P(B | A^c) = 1/4$$

$$P(A | B) = 69/80$$

문제 3 (15점)

1) 어느 공정에서 생산되는 제품의 불량률이 10%라 할 때, 100개의 제품을 검사했을 때, 불량 개수의 평균과 표준편차를 구하는 파이썬 코드를 작성하시오. (5)

```
import math
import scipy
p=0.1
n=100
np= n * p
print("Average = ", np)
sd = math.sqrt(n * p * (1-p))
print("Stddev = ", sd)
[avg, v] = stats.binom.stats(n=100, p=0.1)
print(avg, math.sqrt(v))
```

3) 어느 식당에서 손님 한명이 서비스를 받기 위해 대기하는 시간의 길이가 평균 1분인 지수분포를 따른다고 한다. 어떤 손님이 그 식당에 갔을 때 2분 이내로 대기할 확률을 구하는 파이썬 코드를 작성하시오. (10)

```
import scipy.stats as stats
```

```
print("P[X<2] = ", stats.poisson.cdf(5, mu=1))
```

문제 4 (20점)

1) 어느 타이어 공장에서 새로운 공정으로 생산되는 타이어의 평균 수명을 추정하기 위해 36 개의 표본을 추출하여 수명을 조사한 결과 아래와 같다.

```
A=[31, 33, 29, 28, 25, 32, 32, 34, 26, 30, 29, 29, 32, 26, 27, 27,  
25, 26, 33, 29, 25, 33, 32, 26, 28, 34, 32, 29, 33, 30, 30, 31, 26,  
28, 28, 32]
```

평균 수명에 대한 95% 신뢰 구간을 구하는 파이썬 코드를 작성하시오. (10)

```
A=[31, 33, 29, 28, 25, 32, 32, 34, 26, 30, 29, 29, 32, 26, 27, 27,  
25, 26, 33, 29, 25, 33, 32, 26, 28, 34, 32, 29, 33, 30, 30, 31, 26,  
28, 28, 32]
```

```
import numpy as np  
import scipy.stats  
xbar = np.mean(w)  
sd = np.std(w, ddof=1)  
n = len(w)  
print("타이어수 n = ", n)  
alpha = 0.05  
zalpha = scipy.stats.norm.ppf(1 - alpha/2)  
se = zalpha * sd / np.sqrt(n)  
ci = [xbar - se, xbar + se]
```

```
print("평균 = %.3f, 표준편차 = %.3f, 신뢰구간 = [%.3f, %.3f]"%(xbar, sd, ci[0], ci[1]))
```

2) 위 문제에서 타이어의 평균 수명은 29.5 개월로 알려져 있다. 새로운 공정을 도입하여 생산된 타이어가 평균적으로 더 수명이 길다고 할 수 있는지 검정하는 파이썬 코드를 작성하시오. 단, 유의수준 5%로 검정하시오. (10)

```
mu=29.5  
xbar=np.mean(w)  
sd=np.std(w, ddof=1)  
print("평균 %.3f, 표준편차: %.3f" %(xbar, sd))  
z=(xbar-mu)/(sd/np.sqrt(len(w)))  
print("검정통계량: ", z)  
alpha=0.05
```

```
import scipy.stats
cri = scipy.stats.norm.ppf(1-alpha)
print("임계값: ", cri)
```

문제 5 (10점)

1) 어떤 확률 변수에 대한 확률분포가 아래 변수 P 와 같이 주어져 있다. 이 확률분포의 엔트로피 (entropy)를 비트로 계산하는 파이썬 코드를 작성하시오.(5)

P=[0.23, 0.22, 0.25, 0.19, 0.11]

```
import numpy as np
A=-(np.log(0.23))
B=-(np.log(0.22))
C=-(np.log(0.25))
D=-(np.log(0.19))
E=-(np.log(0.11))
print("A",A)
print("B",B)
print("C",C)
print("D",D)
print("E",E)
```

```
EB=(1*0.23)+(1*0.22)+(1*0.25)+(1*0.19)+(2*0.19)+(1*0.11)
print("EB",EB)
```

) 아래는 2 개의 확률 분포를 각각 변수 P 와 Q 에 저장하였다. 확률분포 P 에 대한 확률분포 Q 의 교차 엔트로피를 계산하는 파이썬 코드를 작성하시오. (5)

P=[0.28, 0.22, 0.15, 0.20, 0.15]
Q=[0.31, 0.23, 0.21, 0.1, 0.15]

```
import numpy as np
def crossentropy(P,Q):
    return sum([-P[i]*np.log2(Q[i]) for i in range(len(P))])

print(crossentropy(P,Q))
```