

## W~Y 엔트로피 조사

```
In [1]: import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
plt.style.use('fivethirtyeight')
import warnings
warnings.filterwarnings('ignore')
%matplotlib inline
```

```
In [2]: import os
# 운영체제별 한글 폰트 설정
if os.name == 'posix': # Mac 환경 폰트 설정
    plt.rc('font', family='AppleGothic')
elif os.name == 'nt': # Windows 환경 폰트 설정
    plt.rc('font', family='Malgun Gothic')

plt.rc('axes', unicode_minus=False) # 마이너스 폰트 설정

# 글씨 선명하게 출력하는 설정
%config InlineBackend.figure_format = 'retina'
```

```
In [3]: data2017 = pd.read_csv('NHIS_OPEN_GJ_2017_100.csv', encoding='euc-kr')
data2018 = pd.read_csv('NHIS_OPEN_GJ_2018_100.csv')
```

```
In [4]: data2017 = data2017[['성별코드', '식전혈당(공복혈당)', '(혈청지오티)AST', '(혈청지오티)ALT',
data2018 = data2018[['성별코드', '식전혈당(공복혈당)', '(혈청지오티)AST', '(혈청지오티)ALT',
```

## 부모 엔트로피

```
In [5]: data = pd.concat([data2017, data2018]) # 2017, 2018 데이터 합침
len(data)
```

Out[5]: 2000000

```
In [6]: data.dropna(axis=0) # null 값 제거
```

```
Out[6]:
```

	성별코드	식전혈당(공복혈당)	(혈청지오티)AST	(혈청지오티)ALT	감마지티피
0	1	99.0	21.0	35.0	40.0
1	1	106.0	20.0	36.0	27.0
2	1	98.0	47.0	32.0	68.0
3	1	95.0	29.0	34.0	18.0
4	1	101.0	19.0	12.0	25.0
...	...	...	...	...	...
999995	2	107.0	18.0	15.0	19.0

	성별코드	식전혈당(공복혈당)	(혈청지오티)AST	(혈청지오티)ALT	감마지티피
999996	1	114.0	25.0	36.0	90.0
999997	1	98.0	21.0	15.0	36.0
999998	1	94.0	20.0	18.0	14.0
999999	1	85.0	17.0	12.0	11.0

1994034 rows × 5 columns

```
In [7]: data = data[data['식전혈당(공복혈당)'] < 300] # 공복혈당 300 이상 제거
data
```

```
Out[7]:
```

	성별코드	식전혈당(공복혈당)	(혈청지오티)AST	(혈청지오티)ALT	감마지티피
0	1	99.0	21.0	35.0	40.0
1	1	106.0	20.0	36.0	27.0
2	1	98.0	47.0	32.0	68.0
3	1	95.0	29.0	34.0	18.0
4	1	101.0	19.0	12.0	25.0
...	...	...	...	...	...
999995	2	107.0	18.0	15.0	19.0
999996	1	114.0	25.0	36.0	90.0
999997	1	98.0	21.0	15.0	36.0
999998	1	94.0	20.0	18.0	14.0
999999	1	85.0	17.0	12.0	11.0

1990911 rows × 5 columns

```
In [72]: diabetes = data['식전혈당(공복혈당)']
diabetes_man = data['식전혈당(공복혈당)'][data['성별코드'] == 1]
diabetes_woman = data['식전혈당(공복혈당)'][data['성별코드'] == 2]
len(diabetes_man)
# len(diabetes_man) + len(diabetes_woman)
```

Out[72]: 1061762

```
In [76]: diabetes_get = data[data['식전혈당(공복혈당)'] >= 126 ]
diabetes_get = diabetes_get['식전혈당(공복혈당)']

diabetes_mget = data[data['식전혈당(공복혈당)'] >= 126]
diabetes_mget = diabetes_mget['식전혈당(공복혈당)'][diabetes_mget['성별코드'] == 1]

diabetes_wget = data[data['식전혈당(공복혈당)'] >= 126]
diabetes_wget = diabetes_wget['식전혈당(공복혈당)'][diabetes_wget['성별코드'] == 2]

diabetes_wget
```

```
Out[76]: 37      128.0
75      155.0
191     155.0
```

```

206      160.0
271      163.0
...
999919    138.0
999924    126.0
999940    143.0
999942    150.0
999991    217.0
Name: 식전혈당(공복혈당), Length: 52017, dtype: float64

```

```

In [10]: print("공복혈당 126 이상 : ", len(diabetes_get))
         print("공복혈당 126 이하 : ", len(diabetes)-len(diabetes_get))

```

```

공복혈당 126 이상 : 152768
공복혈당 126 이하 : 1838143

```

```

In [11]: print("당뇨 발병 확률 : ", len(diabetes_get)/1990911)
         print("정상 확률 : ", (len(diabetes)-len(diabetes_get))/1990911)

```

```

당뇨 발병 확률 : 0.07673271180881516
정상 확률 : 0.9232672881911849

```

```

In [81]: P = np.array([len(diabetes_get)/1990911, (len(diabetes)-len(diabetes_get))/1990911])
         mP = np.array([len(diabetes_mget)/1061762, (len(diabetes_man)-len(diabetes_mge))])
         wP = np.array([len(diabetes_wget)/929149, (len(diabetes_woman)-len(diabetes_wget))])

         def H(p):
             id_p = np.where(p != 0)
             return -np.sum(p[id_p]*np.log2(p[id_p]))

         parent = H(P)
         mparent = H(mP)
         wparent = H(wP)
         parent # 부모 엔트로피

```

```

Out[81]: 0.3905607434045131

```

```

In [ ]:

```

## (혈청지오티)AST 엔트로피

(혈청지오티)AST의 정상치는 0~40IU/L 이다.

수치가 120 IU/L 을 넘을 경우 지방간으로 인한 당뇨병 발병 확률을 0.1%로 보았다.

또한 2000 IU/L 를 넘어가는 수치는 이상치로 판단하여 제거하였다.

- 1) 정상인 0 ~ 40 IU/L
- 2) 약간 높음(지방간) 40 ~ 120 IU/L
- 3) 경도 120 ~ 200 IU/L
- 4) 중등도 200 ~ 400 IU/L
- 5) 중증 400 ~ 2000 IU/L

```

In [13]: ast = data[data['(혈청지오티)AST'] < 2000] # 이상치 제거
         ast

```

Out[13]:

	성별코드	식전혈당(공복혈당)	(혈청지오티)AST	(혈청지오티)ALT	감마지티피
0	1	99.0	21.0	35.0	40.0
1	1	106.0	20.0	36.0	27.0
2	1	98.0	47.0	32.0	68.0
3	1	95.0	29.0	34.0	18.0
4	1	101.0	19.0	12.0	25.0
...	...	...	...	...	...
999995	2	107.0	18.0	15.0	19.0
999996	1	114.0	25.0	36.0	90.0
999997	1	98.0	21.0	15.0	36.0
999998	1	94.0	20.0	18.0	14.0
999999	1	85.0	17.0	12.0	11.0

1990888 rows × 5 columns

In [14]:

```
ast1 = ast[(ast['(혈청지오티)AST'] >= 0) & (ast['(혈청지오티)AST'] <= 40)]
ast2 = ast[(ast['(혈청지오티)AST'] > 40) & (ast['(혈청지오티)AST'] <= 120)]
ast3 = ast[(ast['(혈청지오티)AST'] > 120) & (ast['(혈청지오티)AST'] <= 200)]
ast4 = ast[(ast['(혈청지오티)AST'] > 200) & (ast['(혈청지오티)AST'] <= 400)]
ast5 = ast[(ast['(혈청지오티)AST'] > 400)]
```

In [15]:

```
print(len(ast1))
print(len(ast2))
print(len(ast3))
print(len(ast4))
print(len(ast5))
len(ast1)+len(ast2)+len(ast3)+len(ast4)+len(ast5) # 총 갯수 일치하는지 확인, 이상없음
```

```
1840640
143426
4811
1579
432
```

Out[15]: 1990888

In [16]:

```
ast1 = ast1['식전혈당(공복혈당)']
ast2 = ast2['식전혈당(공복혈당)']
ast3 = ast3['식전혈당(공복혈당)']
ast4 = ast4['식전혈당(공복혈당)']
ast5 = ast5['식전혈당(공복혈당)']
```

In [17]:

```
ast1_diabetes_get = ast1>=126
ast2_diabetes_get = ast2>=126
ast3_diabetes_get = ast3>=126
ast4_diabetes_get = ast4>=126
ast5_diabetes_get = ast5>=126
```

In [18]:

```
print("ast1 공복혈당 126 이상 :", ast1_diabetes_get.sum())
print("ast1 공복혈당 126 이하 :", len(ast1)-ast1_diabetes_get.sum(), end="\n\n")
print("ast2 공복혈당 126 이상 :", ast2_diabetes_get.sum())
```

```
print("ast2 공복혈당 126 이하 : ",len(ast2)-ast2_diabetes_get.sum(), end="\n\n")
print("ast3 공복혈당 126 이상 : ", ast3_diabetes_get.sum())
print("ast3 공복혈당 126 이하 : ",len(ast3)-ast3_diabetes_get.sum(), end="\n\n")
print("ast4 공복혈당 126 이상 : ", ast4_diabetes_get.sum())
print("ast4 공복혈당 126 이하 : ",len(ast4)-ast4_diabetes_get.sum(), end="\n\n")
print("ast5 공복혈당 126 이상 : ", ast5_diabetes_get.sum())
print("ast5 공복혈당 126 이하 : ",len(ast5)-ast5_diabetes_get.sum())
```

```
ast1 공복혈당 126 이상 : 128015
ast1 공복혈당 126 이하 : 1712625
```

```
ast2 공복혈당 126 이상 : 23312
ast2 공복혈당 126 이하 : 120114
```

```
ast3 공복혈당 126 이상 : 1076
ast3 공복혈당 126 이하 : 3735
```

```
ast4 공복혈당 126 이상 : 302
ast4 공복혈당 126 이하 : 1277
```

```
ast5 공복혈당 126 이상 : 59
ast5 공복혈당 126 이하 : 373
```

In [19]:

```
Px = [128015/1840640, 23312/143426, 1076/4811, 302/1579, 59/432]
Py = [1712625/1840640, 120114/143426, 3735/4811, 1277/1579, 373/432]
print("Px :",Px, "\nPy :",Py)
```

```
Px : [0.06954917854659248, 0.16253677854782256, 0.2236541259613386, 0.19126029
132362254, 0.13657407407407407]
Py : [0.9304508214534075, 0.8374632214521774, 0.7763458740386614, 0.8087397086
763775, 0.8634259259259259]
```

In [20]:

```
ast_entropy = []
for x,y in zip(Px,Py):
    P = np.array([x,y])
    ast_entropy.append(H(P))
```

In [21]:

```
print("ast1 엔트로피 :",ast_entropy[0])
print("ast2 엔트로피 :",ast_entropy[1])
print("ast3 엔트로피 :",ast_entropy[2])
print("ast4 엔트로피 :",ast_entropy[3])
print("ast5 엔트로피 :",ast_entropy[4])
```

```
ast1 엔트로피 : 0.3642390171395185
ast2 엔트로피 : 0.6403439416641256
ast3 엔트로피 : 0.7667839207101522
ast4 엔트로피 : 0.7041004594912995
ast5 엔트로피 : 0.5751958149700604
```

## AST 정보증가량

In [22]:

```
ast_ig = parent - (1840640/1990888*ast_entropy[0]+ 143426/1990888*ast_entropy
print("AST 정보증가량 : ",ast_ig)
```

```
AST 정보증가량 : 0.0051427131863678865
```

## (혈청지오티)AIT 엔트로피

(혈청지오티)ALT의 정상치는 0~40IU/L 이다.

ALT 또한 2000 IU/L 를 넘어가는 수치는 이상치로 판단하여 제거하였다.

- 1) 정상인 0 ~ 40 IU/L
- 2) 약간 높음(지방간) 40 ~ 120 IU/L
- 3) 경도 120 ~ 200 IU/L
- 4) 중등도 200 ~ 400 IU/L
- 5) 중증 400 ~ 2000 IU/L

```
In [23]: alt = data[data['(혈청지오티)ALT'] < 2000] # 이상치 제거
alt
```

```
Out[23]:
```

	성별코드	식전혈당(공복혈당)	(혈청지오티)AST	(혈청지오티)ALT	감마지티피
0	1	99.0	21.0	35.0	40.0
1	1	106.0	20.0	36.0	27.0
2	1	98.0	47.0	32.0	68.0
3	1	95.0	29.0	34.0	18.0
4	1	101.0	19.0	12.0	25.0
...	...	...	...	...	...
999995	2	107.0	18.0	15.0	19.0
999996	1	114.0	25.0	36.0	90.0
999997	1	98.0	21.0	15.0	36.0
999998	1	94.0	20.0	18.0	14.0
999999	1	85.0	17.0	12.0	11.0

1990874 rows × 5 columns

```
In [24]: alt1 = alt[(alt['(혈청지오티)ALT'] >= 0) & (alt['(혈청지오티)ALT'] <= 40)]
alt2 = alt[(alt['(혈청지오티)ALT'] > 40) & (alt['(혈청지오티)ALT'] <= 120)]
alt3 = alt[(alt['(혈청지오티)ALT'] > 120) & (alt['(혈청지오티)ALT'] <= 200)]
alt4 = alt[(alt['(혈청지오티)ALT'] > 200) & (alt['(혈청지오티)ALT'] <= 400)]
alt5 = alt[(alt['(혈청지오티)ALT'] > 400)]
```

```
In [25]: print(len(alt1))
print(len(alt2))
print(len(alt3))
print(len(alt4))
print(len(alt5))
len(alt1)+len(alt2)+len(alt3)+len(alt4)+len(alt5) # 총 갯수 일치하는지 확인, 이상없음
```

```
1736206
240756
11014
2388
510
```

```
Out[25]: 1990874
```

```
In [26]: alt1 = alt1['식전혈당(공복혈당)']
alt2 = alt2['식전혈당(공복혈당)']
```

```
alt3 = alt3['식전혈당(공복혈당)']
alt4 = alt4['식전혈당(공복혈당)']
alt5 = alt5['식전혈당(공복혈당)']
```

```
In [27]: alt1_diabetes_get = alt1>=126
alt2_diabetes_get = alt2>=126
alt3_diabetes_get = alt3>=126
alt4_diabetes_get = alt4>=126
alt5_diabetes_get = alt5>=126
```

```
In [28]: print("alt1 공복혈당 126 이상 :", alt1_diabetes_get.sum())
print("alt1 공복혈당 126 이하 :", len(alt1)-alt1_diabetes_get.sum(), end="\n\n")
print("alt2 공복혈당 126 이상 :", alt2_diabetes_get.sum())
print("alt2 공복혈당 126 이하 :", len(alt2)-alt2_diabetes_get.sum(), end="\n\n")
print("alt3 공복혈당 126 이상 :", alt3_diabetes_get.sum())
print("alt3 공복혈당 126 이하 :", len(alt3)-alt3_diabetes_get.sum(), end="\n\n")
print("alt4 공복혈당 126 이상 :", alt4_diabetes_get.sum())
print("alt4 공복혈당 126 이하 :", len(alt4)-alt4_diabetes_get.sum(), end="\n\n")
print("alt5 공복혈당 126 이상 :", alt5_diabetes_get.sum())
print("alt5 공복혈당 126 이하 :", len(alt5)-alt5_diabetes_get.sum())
```

```
alt1 공복혈당 126 이상 : 115969
alt1 공복혈당 126 이하 : 1620237
```

```
alt2 공복혈당 126 이상 : 34383
alt2 공복혈당 126 이하 : 206373
```

```
alt3 공복혈당 126 이상 : 1958
alt3 공복혈당 126 이하 : 9056
```

```
alt4 공복혈당 126 이상 : 394
alt4 공복혈당 126 이하 : 1994
```

```
alt5 공복혈당 126 이상 : 59
alt5 공복혈당 126 이하 : 451
```

```
In [29]: Px = [115969/1736206, 34383/240756, 1958/11014, 394/2388, 59/510]
Py = [1620237/1736206, 206373/240756, 9056/11014, 1994/2388, 451/510]
print("Px :", Px, "\nPy :", Py)
```

```
Px : [0.06679449328017528, 0.1428126401834222, 0.17777374250953332, 0.16499162
479061977, 0.11568627450980393]
Py : [0.9332055067198247, 0.8571873598165778, 0.8222262574904666, 0.8350083752
093802, 0.884313725490196]
```

```
In [30]: alt_entropy = []
for x,y in zip(Px,Py):
    P = np.array([x,y])
    alt_entropy.append(H(P))
```

```
In [31]: print("alt1 엔트로피 :", alt_entropy[0])
print("alt2 엔트로피 :", alt_entropy[1])
print("alt3 엔트로피 :", alt_entropy[2])
print("alt4 엔트로피 :", alt_entropy[3])
print("alt5 엔트로피 :", alt_entropy[4])
```

```
alt1 엔트로피 : 0.3538458269929723
alt2 엔트로피 : 0.591557729171504
alt3 엔트로피 : 0.6751825243285104
```

alt4 엔트로피 : 0.6461184832849831  
alt5 엔트로피 : 0.5168327427401807

## ALT 정보증가량

In [32]:

```
alt_ig = parent - (1736206/1990874*alt_entropy[0]+ 240756/1990874*alt_entropy[1])
print("ALT 정보증가량 : ", alt_ig)
```

ALT 정보증가량 : 0.0057984246920222415

## 감마지티피 엔트로피

감마지티피의 정상치는 남성이 11 ~ 63IU/L, 여성이 8 ~ 35IU/L 이다.

따라서 범위를 남자/여자로 나누어 구했다

남성 1) 낮음 0 ~ 10 IU/L  
2) 정상인 11 ~ 63IU/L  
3) 경도 63 ~ 99 IU/L  
4) 중등도 100 ~ 199 IU/L  
5) 약 고도 200 ~ 599 IU/L  
6) 고도 600 ~ IU/L

여성 1) 낮음 0 ~ 7 IU/L  
2) 정상인 8 ~ 35IU/L  
3) 경도 35 ~ 99 IU/L  
4) 중등도 100 ~ 199 IU/L  
5) 약 고도 200 ~ 599 IU/L  
6) 고도 600 ~ IU/L

In [33]:

```
gtp = data
gtp
```

Out[33]:

	성별코드	식전혈당(공복혈당)	(혈청지오티)AST	(혈청지오티)ALT	감마지티피
0	1	99.0	21.0	35.0	40.0
1	1	106.0	20.0	36.0	27.0
2	1	98.0	47.0	32.0	68.0
3	1	95.0	29.0	34.0	18.0
4	1	101.0	19.0	12.0	25.0
...	...	...	...	...	...
999995	2	107.0	18.0	15.0	19.0
999996	1	114.0	25.0	36.0	90.0
999997	1	98.0	21.0	15.0	36.0
999998	1	94.0	20.0	18.0	14.0
999999	1	85.0	17.0	12.0	11.0

1990911 rows × 5 columns



```
In [34]: gtp_man = gtp [gtp['성별코드'] == 1]
gtp_woman = gtp [gtp['성별코드'] == 2]
len(gtp_man) + len(gtp_woman) # 1990911 row 확인
```

Out[34]: 1990911

## 감마지티피 남자

```
In [35]: gtp_man1 = gtp_man[ (gtp_man['감마지티피'] >= 0) & (gtp_man['감마지티피'] <= 10) ]
gtp_man2 = gtp_man[ (gtp_man['감마지티피'] > 10) & (gtp_man['감마지티피'] <= 63) ]
gtp_man3 = gtp_man[ (gtp_man['감마지티피'] > 63) & (gtp_man['감마지티피'] <= 99) ]
gtp_man4 = gtp_man[ (gtp_man['감마지티피'] > 99) & (gtp_man['감마지티피'] <= 199) ]
gtp_man5 = gtp_man[ (gtp_man['감마지티피'] > 199) & (gtp_man['감마지티피'] <= 599) ]
gtp_man6 = gtp_man[ (gtp_man['감마지티피'] > 599) ]
```

```
In [36]: print(len(gtp_man1))
print(len(gtp_man2))
print(len(gtp_man3))
print(len(gtp_man4))
print(len(gtp_man5))
print(len(gtp_man6))
len(gtp_man1)+len(gtp_man2)+len(gtp_man3)+len(gtp_man4)+len(gtp_man5)+len(gtp
```

9129  
844277  
110940  
73125  
22133  
2153

Out[36]: 1061757

```
In [37]: gtp_man1 = gtp_man1[ '식전혈당(공복혈당)' ]
gtp_man2 = gtp_man2[ '식전혈당(공복혈당)' ]
gtp_man3 = gtp_man3[ '식전혈당(공복혈당)' ]
gtp_man4 = gtp_man4[ '식전혈당(공복혈당)' ]
gtp_man5 = gtp_man5[ '식전혈당(공복혈당)' ]
gtp_man6 = gtp_man6[ '식전혈당(공복혈당)' ]
```

```
In [38]: gtp_man1_diabetes_get = gtp_man1>=126
gtp_man2_diabetes_get = gtp_man2>=126
gtp_man3_diabetes_get = gtp_man3>=126
gtp_man4_diabetes_get = gtp_man4>=126
gtp_man5_diabetes_get = gtp_man5>=126
gtp_man6_diabetes_get = gtp_man6>=126
```

```
In [39]: print("gtp_man1 공복혈당 126 이상 : ", gtp_man1_diabetes_get.sum())
print("gtp_man1 공복혈당 126 이하 : ", len(gtp_man1)-gtp_man1_diabetes_get.sum(),
print("gtp_man2 공복혈당 126 이상 : ", gtp_man2_diabetes_get.sum())
print("gtp_man2 공복혈당 126 이하 : ", len(gtp_man2)-gtp_man2_diabetes_get.sum(),
print("gtp_man3 공복혈당 126 이상 : ", gtp_man3_diabetes_get.sum())
print("gtp_man3 공복혈당 126 이하 : ", len(gtp_man3)-gtp_man3_diabetes_get.sum(),
print("gtp_man4 공복혈당 126 이상 : ", gtp_man4_diabetes_get.sum())
print("gtp_man4 공복혈당 126 이하 : ", len(gtp_man4)-gtp_man4_diabetes_get.sum(),
print("gtp_man5 공복혈당 126 이상 : ", gtp_man5_diabetes_get.sum())
print("gtp_man5 공복혈당 126 이하 : ", len(gtp_man5)-gtp_man5_diabetes_get.sum(),
print("gtp_man6 공복혈당 126 이상 : ", gtp_man6_diabetes_get.sum())
print("gtp_man6 공복혈당 126 이하 : ", len(gtp_man6)-gtp_man6_diabetes_get.sum())
```

```

gtp_man1 공복혈당 126 이상 : 521
gtp_man1 공복혈당 126 이하 : 8608

gtp_man2 공복혈당 126 이상 : 67161
gtp_man2 공복혈당 126 이하 : 777116

gtp_man3 공복혈당 126 이상 : 15245
gtp_man3 공복혈당 126 이하 : 95695

gtp_man4 공복혈당 126 이상 : 12298
gtp_man4 공복혈당 126 이하 : 60827

gtp_man5 공복혈당 126 이상 : 4828
gtp_man5 공복혈당 126 이하 : 17305

gtp_man6 공복혈당 126 이상 : 697
gtp_man6 공복혈당 126 이하 : 1456

```

In [40]:

```

Px = [521/9129, 67161/844277, 15245/110940, 12298/73125, 4828/22133, 697/2153]
Py = [8608/9129, 777116/844277, 95695/110940, 60827/73125, 17305/22133, 1456/2153]
print("Px :", Px, "\nPy :", Py)

```

```

Px : [0.05707087304195421, 0.07954853679538824, 0.13741662159725979, 0.1681777777777778, 0.21813581529842316, 0.3237343241987924]
Py : [0.9429291269580458, 0.9204514632046118, 0.8625833784027402, 0.8318222222222222, 0.7818641847015768, 0.6762656758012077]

```

In [41]:

```

gtp_man_entropy = []
for x,y in zip(Px,Py):
    P = np.array([x,y])
    gtp_man_entropy.append(H(P))

```

In [42]:

```

print("gtp_man1 엔트로피 :", gtp_man_entropy[0])
print("gtp_man2 엔트로피 :", gtp_man_entropy[1])
print("gtp_man3 엔트로피 :", gtp_man_entropy[2])
print("gtp_man4 엔트로피 :", gtp_man_entropy[3])
print("gtp_man5 엔트로피 :", gtp_man_entropy[4])
print("gtp_man6 엔트로피 :", gtp_man_entropy[5])

```

```

gtp_man1 엔트로피 : 0.3157059315005649
gtp_man2 엔트로피 : 0.4005864305757727
gtp_man3 엔트로피 : 0.577432983931476
gtp_man4 엔트로피 : 0.653519281935413
gtp_man5 엔트로피 : 0.7567489182557392
gtp_man6 엔트로피 : 0.9083962624581801

```

## 감마지티피 여자

In [43]:

```

gtp_woman1 = gtp_woman[(gtp_woman['감마지티피'] >= 0) & (gtp_woman['감마지티피'] <= 7)]
gtp_woman2 = gtp_woman[(gtp_woman['감마지티피'] > 7) & (gtp_woman['감마지티피'] <= 35)]
gtp_woman3 = gtp_woman[(gtp_woman['감마지티피'] > 35) & (gtp_woman['감마지티피'] <= 99)]
gtp_woman4 = gtp_woman[(gtp_woman['감마지티피'] > 99) & (gtp_woman['감마지티피'] <= 199)]
gtp_woman5 = gtp_woman[(gtp_woman['감마지티피'] > 199) & (gtp_woman['감마지티피'] <= 599)]
gtp_woman6 = gtp_woman[(gtp_woman['감마지티피'] > 599)]

```

In [44]:

```

print(len(gtp_woman1))
print(len(gtp_woman2))
print(len(gtp_woman3))
print(len(gtp_woman4))
print(len(gtp_woman5))

```

```
print(len(gtp_woman6))
len(gtp_woman1)+len(gtp_woman2)+len(gtp_woman3)+len(gtp_woman4)+len(gtp_woman5)+len(gtp_woman6)
```

```
10811
807079
96824
11122
2986
326
```

Out[44]: 929148

```
In [45]: gtp_woman1 = gtp_woman1['식전혈당 (공복혈당)']
gtp_woman2 = gtp_woman2['식전혈당 (공복혈당)']
gtp_woman3 = gtp_woman3['식전혈당 (공복혈당)']
gtp_woman4 = gtp_woman4['식전혈당 (공복혈당)']
gtp_woman5 = gtp_woman5['식전혈당 (공복혈당)']
gtp_woman6 = gtp_woman6['식전혈당 (공복혈당)']
```

```
In [46]: gtp_woman1_diabetes_get = gtp_woman1>=126
gtp_woman2_diabetes_get = gtp_woman2>=126
gtp_woman3_diabetes_get = gtp_woman3>=126
gtp_woman4_diabetes_get = gtp_woman4>=126
gtp_woman5_diabetes_get = gtp_woman5>=126
gtp_woman6_diabetes_get = gtp_woman6>=126
```

```
In [47]: print("gtp_woman1 공복혈당 126 이상 : ", gtp_woman1_diabetes_get.sum())
print("gtp_woman1 공복혈당 126 이하 : ", len(gtp_woman1)-gtp_woman1_diabetes_get.sum())
print("gtp_woman2 공복혈당 126 이상 : ", gtp_woman2_diabetes_get.sum())
print("gtp_woman2 공복혈당 126 이하 : ", len(gtp_woman2)-gtp_woman2_diabetes_get.sum())
print("gtp_woman3 공복혈당 126 이상 : ", gtp_woman3_diabetes_get.sum())
print("gtp_woman3 공복혈당 126 이하 : ", len(gtp_woman3)-gtp_woman3_diabetes_get.sum())
print("gtp_woman4 공복혈당 126 이상 : ", gtp_woman4_diabetes_get.sum())
print("gtp_woman4 공복혈당 126 이하 : ", len(gtp_woman4)-gtp_woman4_diabetes_get.sum())
print("gtp_woman5 공복혈당 126 이상 : ", gtp_woman5_diabetes_get.sum())
print("gtp_woman5 공복혈당 126 이하 : ", len(gtp_woman5)-gtp_woman5_diabetes_get.sum())
print("gtp_woman6 공복혈당 126 이상 : ", gtp_woman6_diabetes_get.sum())
print("gtp_woman6 공복혈당 126 이하 : ", len(gtp_woman6)-gtp_woman6_diabetes_get.sum())
```

```
gtp_woman1 공복혈당 126 이상 : 156
gtp_woman1 공복혈당 126 이하 : 10655
```

```
gtp_woman2 공복혈당 126 이상 : 36527
gtp_woman2 공복혈당 126 이하 : 770552
```

```
gtp_woman3 공복혈당 126 이상 : 12947
gtp_woman3 공복혈당 126 이하 : 83877
```

```
gtp_woman4 공복혈당 126 이상 : 1806
gtp_woman4 공복혈당 126 이하 : 9316
```

```
gtp_woman5 공복혈당 126 이상 : 511
gtp_woman5 공복혈당 126 이하 : 2475
```

```
gtp_woman6 공복혈당 126 이상 : 70
gtp_woman6 공복혈당 126 이하 : 256
```

```
In [48]: Px = [156/10811, 36527/807079, 12947/96824, 1806/11122, 511/22986, 70/326]
Py = [10655/10811, 770552/807079, 83877/96824, 9316/11122, 2475/2986, 256/326]
print("Px :", Px, "\nPy :", Py)
```

```
Px : [0.01442974747941911, 0.045258270875589625, 0.13371684706271172, 0.162380
```

```
86675058444, 0.022230923170625597, 0.2147239263803681]
Py : [0.9855702525205808, 0.9547417291244104, 0.8662831529372883, 0.8376191332
494156, 0.8288680509042197, 0.7852760736196319]
```

```
In [49]: gtp_woman_entropy = []
          for x,y in zip(Px,Py):
              P = np.array([x,y])
              gtp_woman_entropy.append(H(P))
```

```
In [50]: print("gtp_woman1 엔트로피 :",gtp_woman_entropy[0])
          print("gtp_woman2 엔트로피 :",gtp_woman_entropy[1])
          print("gtp_woman3 엔트로피 :",gtp_woman_entropy[2])
          print("gtp_woman4 엔트로피 :",gtp_woman_entropy[3])
          print("gtp_woman5 엔트로피 :",gtp_woman_entropy[4])
          print("gtp_woman6 엔트로피 :",gtp_woman_entropy[5])
```

```
gtp_woman1 엔트로피 : 0.1089019663373415
gtp_woman2 엔트로피 : 0.26590224620047215
gtp_woman3 엔트로피 : 0.5675442448499202
gtp_woman4 엔트로피 : 0.6399750409308562
gtp_woman5 엔트로피 : 0.34652197484850034
gtp_woman6 엔트로피 : 0.7504158499790908
```

## GTP 정보증가량

남자 / 여자로 나뉘는데 ,, 정보증가량이 원래 집합(부모 집합) - 분할된 집합(자식집합) 들의 합 이니까 ,, 아래 처럼 남녀 따로 계산하는게 맞나용?

```
In [82]: gtp_man_ig = mparent - (9129/1061757*gtp_man_entropy[0]+ 844277/1061757*gtp_m
          print("GTP 남자 정보증가량 :",gtp_man_ig)

          gtp_woman_ig = wparent - (10811/929148*gtp_woman_entropy[0]+ 807079/929148*gt
          print("GTP 여자 정보증가량 :",gtp_woman_ig)
```

```
GTP 남자 정보증가량 : 0.00837707881283617
GTP 여자 정보증가량 : 0.010874509481908656
```

```
In [ ]:
```