ECS7005P - Risk and Decision Making for Big Data and AI

In [ ]:
```
pip install pyAgrum
```

In [ ]:
```
conda install python-graphviz
```

In [ ]:
```
pip install pyAgrum
```

In [4]:
```
from pylab import *
import matplotlib.pyplot as plt
import pyAgrum as gum
import pyAgrum.lib.notebook as gnb
import numpy as np
from scipy.stats import norm
```

Question 1

In [12]:
```
#probability diagnosed with an STD
(0.2 * 0.35) + (0.7* 0.65)

(0.2*0.35 / 0.5249999999999999)
# probability that the new patient has a UTI is 0.13
```

Out[12]:
```
0.13333333333333333
```

Question 2

In [13]:
```
#create network and discrete nodes
bn = gum.BayesNet('DiseaseDiagnosis')
bn.add(gum.LabelizedVariable('Disease', '', 2))
bn.add(gum.LabelizedVariable('Diagnosis', '', 2))
```

Out[13]:
```
1
```

In [14]:
```
bn.addArc('Disease', 'Diagnosis')
#gnb.showBN(bn)
```

In [24]:
```
#Conditional Probability Table
bn.cpt('Disease').fillWith([0.65, 0.35])

bn.cpt('Diagnosis').fillWith([0.70, 0.30,   #STD 0.70 Diagnosis
                                            #UTI 0.30 incorrectly diagnosed
                              0.20, 0.80])  #STD 0.20 incorrectly diagnosed
                                            #UTI Diagnosis
```

Out[24]:

	Diagnosis	
Disease	**0**	**1**
0	0.7000	0.3000
1	0.2000	0.8000

In [38]:
```
ie = gum.LazyPropagation(bn)
ie.makeInference()

ie.setEvidence({'Diagnosis': 0})
ie.makeInference()
p_UTI_given_STD = ie.posterior('Disease')[1]

print('Probability that a new patient diagnosed with an STD actually has a UTI: \n', p
```

```
Probability that a new patient diagnosed with an STD actually has a UTI:
 0.13333333333333333
```

Question 3

a) Incorrect the relationship being described is a positive correlation both factors ice cream and weather decrease together

b) Incorrect the interpretation finds a causal link where there is no significance and wrongly states the claim correlation does not mean causation and another confounding variable might be influencing the factors.

c) Incorrect p value is greater than the significance level ($\alpha = 0.05$), meaning it is not significant and unable to reject the null hypothesis. There is not enough statistical evidence to show the observed effects are not due to chance.

d) Incorrect the evidence shows no statistical evidence but that does not mean that there is no relationship. This could be due to the model used, sample size or confounding variables not accounted for in study. The study's power may be too low to detect a real effect if it exists.

e) Correct the interpretation reflects 95% confidence interval that if the study were repeated many times, 95% of those confidence intervals would contain the true average weight loss.

In [48]:
```
#3ii

mean = -3
variance = 10
std_dev = variance**0.5

#a) between 1 and 3
p_a = norm.cdf(3, mean, std_dev) - norm.cdf(1, mean, std_dev)
#0.07406181980423554

#b) greater than -5
p_b = 1 - norm.cdf(-5, mean, std_dev)
#0.736455371567231
```

```
print(f'a)  The probability that the growth rate will be between 1 and 3 \n is approxi
print(f'b) The probability that the growth rate will be greater than -5 \n is approxim
```

```
a)  The probability that the growth rate will be between 1 and 3
 is approximately: 0.074
b) The probability that the growth rate will be greater than -5
 is approximately: 0.736
```

Question 4

i) Treatment B = 68/100

ii) Treatment A = 16/20

iii) Treatment A = 30/80

iv) Simpson's paradox is when a trend in different groups of data reverses when the groups are combined. Treatment A shows a higher success rate for acute and chronic disease types, but when combining the success rates across both disease types, the overall effectiveness shows Treatment B is more effective. Simpson's paradox can indicate the presence of a confounding variable that isn't accounted for in the data.

Question 5

In [2]:
```
#i
((20-5) * 1/3) + (-5 *2/3)
```

Out[2]:
```
1.6666666666666665
```

In [56]:
```
#ii
(((20-5) * 1/3) + (-5 *2/3))+7
```

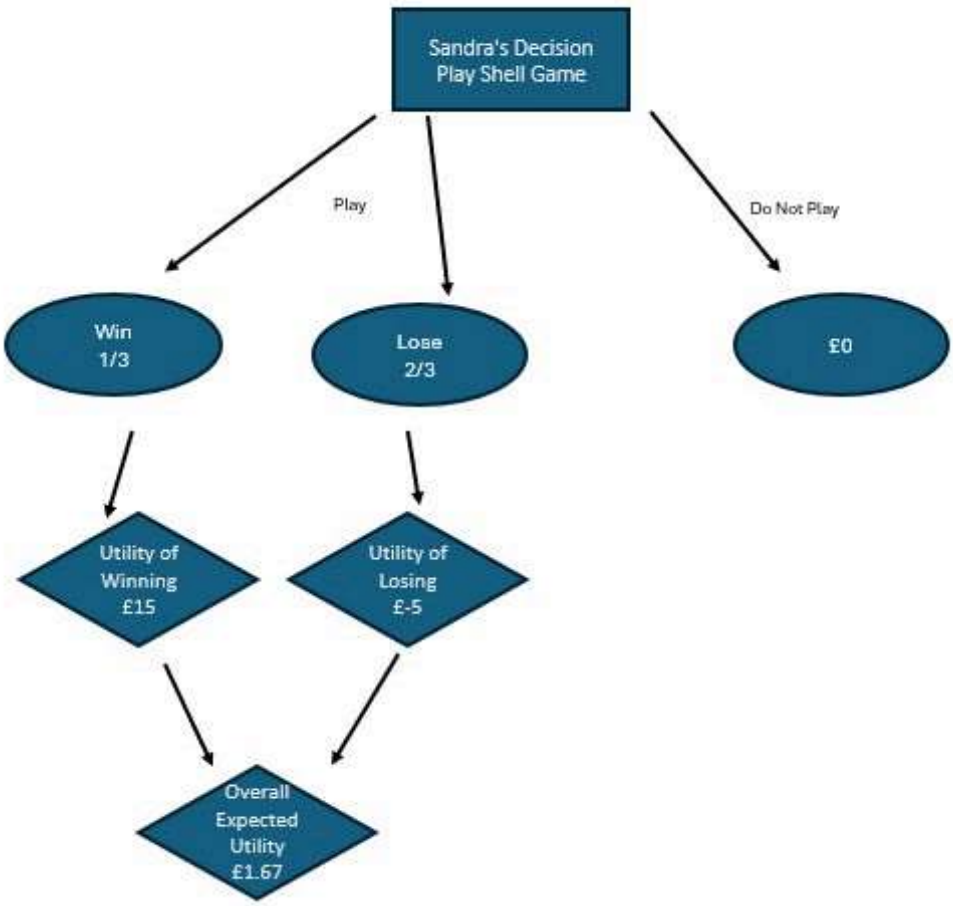
Out[56]:
```
8.666666666666666
```

iii) Sandra should play the game the expected utility is positive

In [3]:
```
#iv)
from IPython.display import Image

Image(filename='Influence diagram.jpg')
```

Out[3]:


Sandra's Decision
Play Shell Game
Play
Do Not Play
Win
1/3
Lose
2/3
£0
Utility of
Winning
£15
Utility of
Losing
£-5
Overall
Expected
Utility
£1.67