Samuel Kim (ssk1)
15-381
HW2

1.1 Hosting Tournament

Probability for just one match between one pair…
P(match) = x * (1 – x) = $x - x^2$

Let M_i be an indicator variable that is a 1 if a match between two players was successful, 0 otherwise. With i ranging from 1 to N for each possible match.

We want to maximize the expected number of total matches:

MAX: E(Total Matches)
= MAX: E (SUM (i=1) to N of Xi)

i.e. we want to maximize the expected number of matches, which is maximizing the sum of all the indicator variables signifying there has been a match between two players.

E (SUM (i=1 to N) of E(Xi))
= SUM (i=1 to N) of E(Xi), by linearity of expectation
= SUM (i=1 to N) of (1 * ($x - x^2$)), definition of expectation is value * probability
= SUM (i=1 to N) of ($x - x^2$)

So, we'd like to maximize $x - x^2$
We can find the derivative, 1 – 2x, and observe **x = 1/2** is the max.

1.2 Bayes Net Independence

1) **False.** A and F are not conditionally independent given C, H, K; because there is a path that connects them (A – D – B – E – G – J – K – I – F).
2) **False.** E and F are not conditionally independent given D, I, J; because there is a path that connects them (E – B – D – A – C – F)
3) **True.** E and F are conditionally independent given B, I, K; because there is no path that connects them, by d-separation.

Samuel Kim (ssk1)
15-381
HW2

1.3 The Burglar Scenario

1) Table computed using formula, P(B, E, A, J, M) = P(J|A)*P(M|A)*P(A|B,E)*P(B)*P(E)

B	E	A	J	M	P(B, E, A, J, M)
F	F	F	F	F	0.93674270062
F	F	F	F	T	0.00946204748
F	F	F	T	F	0.04930224740
F	F	F	T	T	0.00049800250
F	F	T	F	F	0.00002991006
F	F	T	F	T	0.00006979014
F	F	T	T	F	0.00026919054
F	F	T	T	T	0.00062811126
F	T	F	F	F	0.00133417449
F	T	F	F	T	0.00001347651
F	T	F	T	F	0.00007021971
F	T	F	T	T	0.00000070929
F	T	T	F	F	0.00001738260
F	T	T	F	T	0.00004055940
F	T	T	T	F	0.00015644340
F	T	T	T	T	0.00036503460
T	F	F	F	F	0.00005631714
T	F	F	F	T	0.00000056886
T	F	F	T	F	0.00000296406
T	F	F	T	T	0.00000002994
T	F	T	F	F	0.00002814360
T	F	T	F	T	0.00006566840
T	F	T	T	F	0.00025329240
T	F	T	T	T	0.00059101560
T	T	F	F	F	0.00000009405
T	T	F	F	T	0.00000000095
T	T	F	T	F	0.00000000495
T	T	F	T	T	0.00000000005
T	T	T	F	F	0.00000005700
T	T	T	F	T	0.00000013300
T	T	T	T	F	0.00000051300
T	T	T	T	T	0.00000119700

Samuel Kim (ssk1)

P(b,e|a)
= P(a|b,e) * P(b,e) / P(a)
= 0.95 * (0.001 * 0.002) / (P(a|b,e)*P(b,e) + P(a|b, -e)*P(b, -e) + P(a|-b,e)*P(-b,e) + P(a|-b, -e)*P(-b, -e))
= 0.95 *(0.001 * 0.002) / (0.95*0.001*0.002 + 0.94*0.001*0.998 + 0.29*0.999*0.002 + 0.001*0.999*0.998)
= **0.0007550342905**

P(b,e|-a)
= P(-a|b,e) * P(b,e) / P(-a)
= 0.05 * (0.001 * 0.002) / (P(-a|b,e)*P(b,e) + P(-a|b, -e)*P(b, -e) + P(-a|-b,e)*P(-b,e) + P(-a|-b, -e)*P(-b, -e))
= 0.05 * (0.001 * 0.002) / (0.05*0.001*0.002 + 0.06*0.001*0.998 + 0.71*0.999*0.002 + 0.999*0.999*0.998)
= **0.000000100252279**

P(b|a)
= P(a|b) * P(b) / P(a)
=(P(a|b,e)*P(e) + P(a|b, -e)*P(-e)) * P(b) / P(a)
= (P(a|b,e)*P(e) + P(a|b, -e)*P(-e)) * P(b) / P(a|b,e)*P(b,e) + P(a|b, -e)*P(b, -e) + P(a|-b,e)*P(-b,e) + P(a|-b, -e)*P(-b, -e)
= ((0.95*0.002 + 0.94*0.998) * 0.001) / (0.95*0.001*0.002 + 0.94*0.001*0.998 + 0.29*0.999*0.002 + 0.001*0.999*0.998)
= **0.3735512283**

P(e|a)
= P(a|e) * P(e) / P(a)
= (P(a|e,b)*P(b) + P(a|e, -b)*P(-b)) * P(e) / P(a)
= (P(a|e,b)*P(b) + P(a|e, -b)*P(-b)) * P(e) / P(a|b,e)*P(b,e) + P(a|b, -e)*P(b, -e) + P(a|-b,e)*P(-b,e) + P(a|-b, -e)*P(-b, -e)
= ((0.95*0.001 + 0.29*0.999) * 0.002) / (0.95*0.001*0.002 + 0.94*0.001*0.998 + 0.29*0.999*0.002 + 0.001*0.999*0.998)
= **0.231008702**

P(a)
= P(a|b,e)*P(b,e) + P(a|b, -e)*P(b, -e) + P(a|-b,e)*P(-b,e) + P(a|-b, -e)*P(-b, -e)
= (0.95*0.001*0.002 + 0.94*0.001*0.998 + 0.29*0.999*0.002 + 0.001*0.999*0.998)
= **0.002516442**

P(-a)
= 1 – P(a)
= **0.997483558**

P(m)
= P(m|a)*P(a) + P(m|-a)*P(-a)
= P(m|a) * (P(a|b,e)*P(b,e) + P(a|b, -e)*P(b, -e) + P(a|-b,e)*P(-b,e) + P(a|-b, -e)*P(-b, -e)) + P(m|-a) * (P(-a|b,e)*P(b,e) + P(-a|b, -e)*P(b, -e) + P(-a|-b,e)*P(-b,e) + P(-a|-b, -e)*P(-b, -e))
= 0.70 * (0.95*0.001*0.002 + 0.94*0.001*0.998 + 0.29*0.999*0.002 + 0.001*0.999*0.998) + 0.01 * (0.05*0.001*0.002 + 0.06*0.001*0.998 + 0.71*0.999*0.002 + 0.999*0.999*0.998)
= **0.01173634498**

P(-m)
= 1 – P(m)
= **0.988263655**

P(j)
= P(j|a)*P(a) + P(j|-a)*P(-a)
= P(j|a) * (P(a|b,e)*P(b,e) + P(a|b, -e)*P(b, -e) + P(a|-b,e)*P(-b,e) + P(a|-b, -e)*P(-b, -e)) + P(j|-a) * (P(-a|b,e)*P(b,e) + P(-a|b, -e)*P(b, -e) + P(-a|-b,e)*P(-b,e) + P(-a|-b, -e)*P(-b, -e))
= 0.90 * (0.95*0.001*0.002 + 0.94*0.001*0.998 + 0.29*0.999*0.002 + 0.001*0.999*0.998) + 0.05 * (0.05*0.001*0.002 + 0.06*0.001*0.998 + 0.71*0.999*0.002 + 0.999*0.999*0.998)
= **0.0521389757**

P(-j)
= 1 – P(j)
= **0.9478610243**

2) B and E are no longer conditionally independent if A is given, because B and E are both parents of A, so B and E become d-separated when A is given.

   We can verify this by showing P(+B,+E|+A) =/= P(+B|+A) * P(+E|+A):

   P(+B,+E|+A)
   = 0.0007550342905

   P(+B|+A) * P(+E|+A)
   = 0.3735512283 * 0.231008702
   = 0.08629358438

Samuel Kim (ssk1)

1.4 Approximate Inference of the Bayes Net

1) By enumeration:
   - P(+D|+C)
     = P(+D,+C) / P(+C)
     = (P (+D,+C,-A,-B) + P(+D,+C,-A,+B)) / P(+C)
     = (0.0 +P(+D,+C,-A,+B)) / P(+C)
     = P(+D|+B,+C)*P(+C|-A,+B)*P(-A)*P(+B) / P(+C)
     = P(+D|+B,+C)*P(+C|-A,+B)*P(-A)*P(+B) / P(+C|-A,+B)*P(-A)*P(+B)
     = 0.75*0.5*1.0*0.9 / 0.5*1.0*0.9
     = **0.75**
   - P(B|+C)
     = P(+B,+C) / P(+C)
     = (P(+C|+B,-A)*P(+B)*P(-A)) / P(+C)
     = (P(+C|+B,-A)*P(+B)*P(-A)) / P(+C|-A,+B)*P(-A)*P(+B)
     = 0.5*1.0*0.9 / 0.5*1.0*0.9
     = **1.0**
   - P(D|-A,+B)
     = (P(+D,-A,+B,+C) + P(+D,-A,+B,-C)) / P(-A,+B)
     = (P(+D|+B,+C)*P(+C|-A,+B)*P(-A)*P(+B)
       + P(+D|+B,-C)*P(-C|-A,+B)*P(-A)*P(+B)) / P(-A,+B)
     = (P(+D|+B,+C)*P(+C|-A,+B)*P(-A)*P(+B)
       + P(+D|+B,-C)*P(-C|-A,+B)*P(-A)*P(+B)) / (P(-A)*P(+B))
     = (0.75*0.5*1.0*0.9
       + 0.1*0.5*1.0*0.9) / (1.0*0.9)
     = **0.425**

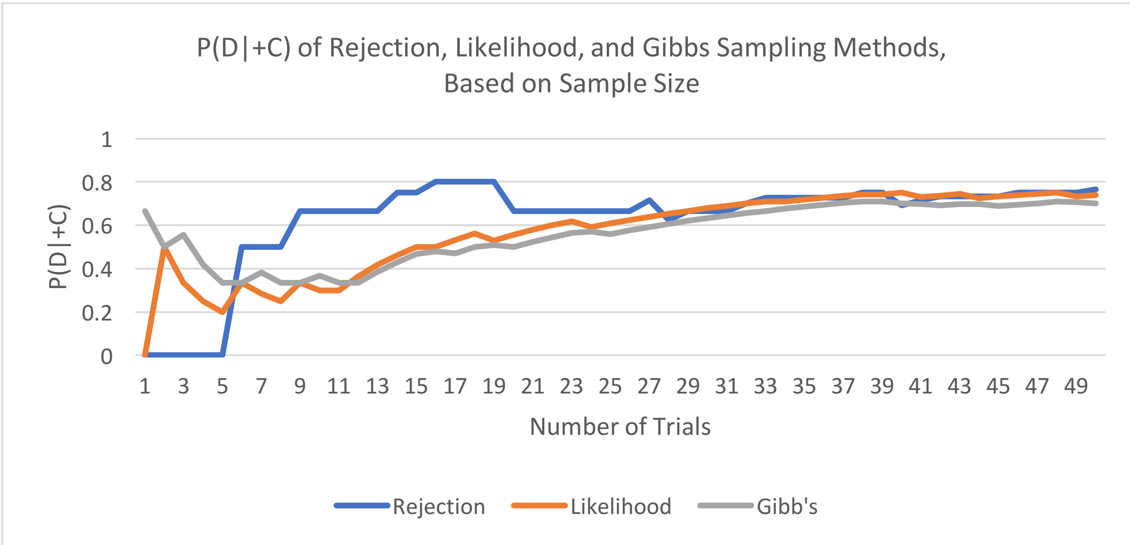
   By approximation, with 1000 samples:

	Rejection	**Likelihood**	**Gibb's**
P(D|+C)	0.765486725664	0.739618406285	0.734
P(B|+C)	1.0	1.0	1.0
P(D|-A,+B)	0.426339285714	0.405	0.3995

   The enumerations by hand are very close to the samples. This is expected, because we chose to do 1,000 trials so the samples should be very close to the actual query.

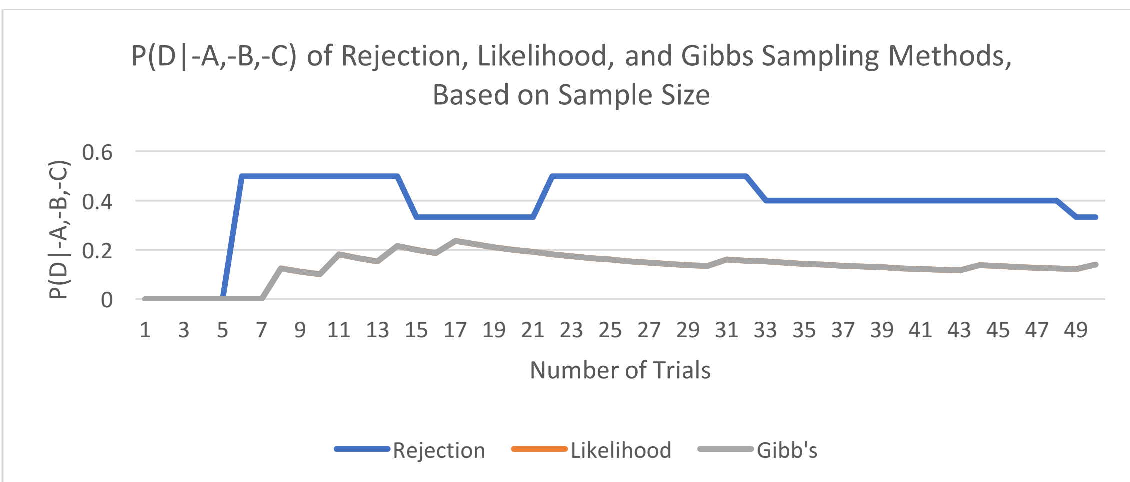
Samuel Kim (ssk1)
15-381
HW2

2)



There is no appreciable difference in the convergence rates for the 3 sampling methods. While they each take different ways of getting there, they all eventually begin to converge to the final value of P(D|+C) around 40 samples.

3)



Knowing how rejection sampling works, we want to choose a query that is very restrictive (has many evidence variables) so that rejection sampling will throw away most of its trials. So, I chose to query P(D|-A,-B,-C). We can see in the graph that rejection sampling has many flat lines, because it's rejecting a lot of the trials that don't fit the evidence variables.

Note that likelihood and gibb's sampling are identical in this case, but this makes sense because both methods fix the evidence variables so all variables except D will be fixed and D will be the only variable continually sampled.

Burglary Net

1) By enumeration, some results copied and pasted from 1.3/#1
P(E|+A) =
P(e|a)
= P(a|e) * P(e) / P(a)
= (P(a|e,b)*P(b) + P(a|e, -b)*P(-b)) * P(e) / P(a)
= (P(a|e,b)*P(b) + P(a|e, -b)*P(-b)) * P(e) / P(a|b,e)*P(b,e) + P(a|b, -e)*P(b, -e) + P(a|-b,e)*P(-b,e) + P(a|-b, -e)*P(-b, -e)
= ((0.95*0.001 + 0.29*0.999) * 0.002) / (0.95*0.001*0.002 + 0.94*0.001*0.998 + 0.29*0.999*0.002 + 0.001*0.999*0.998)
= **0.231008702**

P(E|-A)
= P(-A|E) * P(E) / P(-A)
= (P(-A|E,B)*P(B) + P(-A|E,-B)*P(-B)) * P(E) / P(-A)
= (P(-A|E,B)*P(B) + P(-A|E,-B)*P(-B)) * P(E) / (P(-A|+B,+E)*P(+B,+E) + P(-A|+B,-E)*P(+B,-E) + P(-A|-B,+E)*P(-B,+E) + P(-A|-B,-E)*P(-B,-E))
= ((0.05*0.001 + 0.71*0.999) * 0.002) / (0.05*0.001*0.002 + 0.06*0.001*0.998 + 0.71*0.999*0.002 + 0.999*0.999*0.998)
= **0.001422259032**

P(b|a)
= P(a|b) * P(b) / P(a)
=(P(a|b,e)*P(e) + P(a|b, -e)*P(-e)) * P(b) / P(a)
= (P(a|b,e)*P(e) + P(a|b, -e)*P(-e)) * P(b) / P(a|b,e)*P(b,e) + P(a|b, -e)*P(b, -e) + P(a|-b,e)*P(-b,e) + P(a|-b, -e)*P(-b, -e)
= ((0.95*0.002 + 0.94*0.998) * 0.001) / (0.95*0.001*0.002 + 0.94*0.001*0.998 + 0.29*0.999*0.002 + 0.001*0.999*0.998)
= **0.3735512283**

P(+A) =
= P(A|+B,+E)*P(+B,+E) + P(A|+B, -E)*P(+B,-E) + P(A|-B,+E)*P(-B,+E) + P(A|-B,-E)*P(-B,-E)
= (0.95*0.001*0.002 + 0.94*0.001*0.998 + 0.29*0.999*0.002 + 0.001*0.999*0.998)
= **0.002516442**

P(-A)
= 1 – P(+A)
= **0.997483558**

Samuel Kim (ssk1)
15-381
HW2

By approximation, with 1000 samples:

	Rejection	**Likelihood**	**Gibb's**
P(E|+A)	0.0	0.428486997636	0.214
P(E|-A)	0.00100401606426	0.0035620539907	0.00425
P(B|+A)	0.666666666667	0.277777777778	0.399
P(+A)	0.004	0.004	0.0074
P(-A)*	0.996	0.996	0.9926

* - calculated using 1 – P(A), because the program only calculates probabilities of True

Again, the enumerations by hand are very close to the samples. This is expected, because we chose to do 1,000 trials so the samples should be very close to the actual query.