Question 3 solution

Part 1

1. False: influence can flow along the path Season → Flu → Chills, since Flu is unobserved.
2. True: influence cannot flow through Glu, since it is observed. There are no other paths linking Season and Chills.
3. False: influence can flow along the path Season → Dehydration → Headache, since Dehydration is unobserved.
4. True: since both Flu and Dehydration are observed, influence cannot flow along any path that links Season and Headache
5. False: influence can flow along the path formed by Season → Flue → Headache → Dizziness → Nausea, since Flu, Headache, and Dizziness are unobserved.
6. True: influence cannot flow along the path Season → Dehydration → Nausea, since Dehydration is observed; influence cannot flow along the path Season → Flu → Headache → Dizziness → Nausea, since Headache is observed; influence cannot flow along the path Season → Flu → Headache ← Dehydration → Nausea, even though there is an observed v-structure centered at Headache, because Dehydration is observed.
7. False: influence can flow along the path Flu ← Season → Dehydration, since Season is unobserved.
8. False: influence can flow along the path Flu → Headache ← Dehydration, since this is a v-structure and Headache is observed.
9. True: influence cannot flow through Season, which is observed, nor through Headache or Nausea, since both form v-structures and both are unobserved
10. False: influence can flow along the path Flu → Headache → Dizziness → Nausea ← Dehydration, since Headache and Dizziness are unobserved and there is a v-structure at Nausea, which is observed.
11. False: influence can flow along the path Chills ← Flu ← Season → Dehydration → Nausea, since Flu, Season, and Dehydration are all unobserved.
12. False: influence can flow along the path Chills ← Flu → Headache ← Dehydration → Nausea, since there is a v-structure at Headache, which is observed.

Part 2

1. $P(S, F, D, C, H, Z, N) = P(S)P(F \mid S)P(D \mid S)P(C \mid F)P(H \mid F, D)P(Z \mid H)P(N \mid D, Z)$

part 3

1. This translates to $p(Flu = true)$

$p(F = true) = \sum_s p(F = true, S = s) = \sum_s p(F = true \mid S = s)p(S = s) = p(F = true \mid S = wint)p(S = wint) + p(F = true \mid S = summ)p(S = summ) = 0.4 \times 0.5 + 0.1 \times 0.5 = 0.25$

2. This translates to $p(Flu = true \mid Season = winter)$

$p(F = true \mid S = wint) = 0.4$

3. This translates to $p(Flu = true \mid Season = winter, Headache = true)$

$p(F = true \mid S = wint, H = true) = \frac{p(F=true,S=wint,H=true)}{p(S=wint,H=true)} = \frac{\sum_d p(F=true,S=wint,H=true,D=d)}{\sum_{f,d} p(F=f,S=wint,H=true,D=d)} = \frac{\sum_d p(H=true|F=true,D=d)p(F=true|S=wint)p(D=d|S=wint)p(S=wint)}{\sum_{f,d} p(H=true|F=f,D=d)p(F=f|S=wint)p(D=d|S=wint)p(S=wint)} = \frac{0.9\times0.4\times0.1\times0.5+0.8\times0.4\times0.9\times0.5}{0.9\times0.4\times0.1\times0.5+0.8\times0.4\times0.9\times0.5+0.8\times0.6\times0.1\times0.5+0.3\times0.6\times0.9\times0.5} = 0.61$

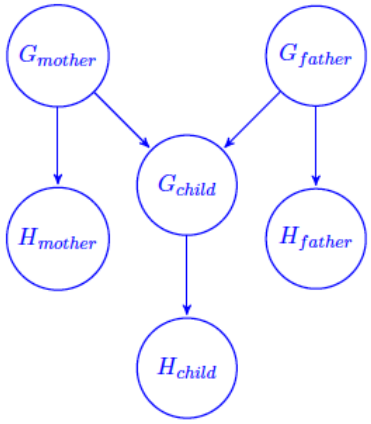
4. This translates to $p(Flu = true \mid Season = winter, Headache = true, Dehydration = true)$

$p(F = true \mid S = winter, H = true, D = true) = \frac{p(F=true,S=winter,H=true,D=true)}{p(S=winter,H=true,D=true)} = \frac{p(F=true,S=winter,H=true,D=true)}{\sum_f p(F=f,S=winter,H=true,D=true)} = \frac{p(H=true|F=true,D=true)p(F=true|S=wint)p(D=true|S=wint)p(S=wint)}{\sum_f p(H=true|F=f,D=true)p(F=f|S=wint)p(D=true|S=wint)p(S=wint)} = \frac{0.9\times0.4\times0.1\times0.5}{0.9\times0.4\times0.1\times0.5+0.8\times0.6\times0.1\times0.5} = 0.43$

5. Knowing that you are dehydrated decreases the likelihood that you have the flu. This makes sense because the headache symptom is "explained away" by the dehydration.

Bonus Question

(a)



(b)

- Yes, because on the path of interest, there are head-to-head edges going into G_{child}, which is not in the evidence set and neither is its descendent H_{child}.
- No, because on the path of interest, there are head-to-head edges going into G_{child}, and its descendent H_{child} is in the evidence set.
- Yes, because on the path of interest there are head-to-tail edges going into G_{child}, which is in the evidence set.
- No, because on the path of interest, there are no head-to-head edges and there are no nodes in the evidence set.

G_{mother}	G_{father}	$P(G_{child} = l \mid \cdots)$	$P(G_{child} = r \mid \cdots)$
l	l	$1 - m$	m
l	r	0.5	0.5
r	l	0.5	0.5
r	r	m	$1 - m$

(c)

(d)

$p(G_{child} = l) = \sum_{g_m, g_f} p(G_{child} = l \mid G_{mother} = g_m, G_{father} = g_f) p(G_{mother} = g_m, G_{father} = g_f) =$
$\sum_{g_m, g_f} p(G_{child} = l \mid G_{mother} = g_m, G_{father} = g_f) p(G_{mother} = g_m) p(G_{father} = g_f) =$
$(1 - m)x^2 + 0.5x(1 - x) + 0.5(1 - x)x + m(1 - x)^2 = x + m - 2mx$

(e)

$p(G_{child} = l) = p(G_{father} = l) \rightarrow x + m - 2mx = x \rightarrow 2mx = m \rightarrow x = 0.5$

$p(G_{child} = l) = 0.5$ implies $p(H_{child} = l) = 0.5p + 0.5(1 - p) = 0.5$, which means half of the population should be left-handed under genetic equilibrium. Since we know only a small fraction of the population is left-handed, the hypothesis for handedness described in this question must be incorrect (if we assume genetic equilibrium).