Problem 1: (15 points)

Give the filtering probabilities for the water table's three values in the following cases:

(1.1) $P(X_1|e_1 = \text{“}notflooded\text{”})$

Observe:

$P(X_1|e_1 = \text{“}notflooded\text{”})$

$$
\begin{aligned}
&= \alpha P(X_1, e_1 = \text{“}notflooded\text{”}) \\
&= \alpha \sum_{x_0} P(X_0, X_1, e_1 = \text{“}notflooded\text{”}) \\
&= \alpha \sum_{x_0} P(X_0)P(X_1|X_0)P(e_1 = \text{“}notflooded\text{”}|X_1) \qquad \text{// chain rule} \\
&= \alpha\ P(e_1 = \text{“}notflooded\text{”}|X_1) \sum_{x_0} P(X_0)P(X_1|X_0) \\
&= \alpha\ P(e_1 = \text{“}notflooded\text{”}|X_1)[P(x_0 = low)P(X_1|x_0 = low) \\
&\quad + P(x_0 = med)P(X_1|x_0 = med) + P(x_0 = high)P(X_1|x_0 = high)]
\end{aligned}
$$

$$
\begin{aligned}
&= \alpha < \big(P(e_1 = \text{“}notflooded\text{”}|x_1 = low)\big[P(x_0 = low)P(x_1 = low|x_0 = low) \\
&\qquad + P(x_0 = med)P(x_1 = low|x_0 = med) + P(x_0 = high)P(x_1 = low|x_0 = high)\big]\big), \\
&\quad \big(P(e_1 = \text{“}notflooded\text{”}|x_1 = med)\big[P(x_0 = low)P(x_1 = med|x_0 = low) \\
&\qquad + P(x_0 = med)P(x_1 = med|x_0 = med) + P(x_0 = high)P(x_1 = med|x_0 = high)\big]\big), \\
&\quad \big(P(e_1 = \text{“}notflooded\text{”}|x_1 = high)[P(x_0 = low)P(x_1 = high|x_0 = low) \\
&\qquad + P(x_0 = med)P(x_1 = high|x_0 = med) + P(x_0 = high)P(x_1 = high|x_0 = high)]\big) >
\end{aligned}
$$

$$
\begin{aligned}
&= \alpha < (1)\left[\left(\frac{1}{3}\right)(0.6) + \left(\frac{1}{3}\right)(0.2) + \left(\frac{1}{3}\right)(0)\right], \\
&\quad (0.95)\left[\left(\frac{1}{3}\right)(0.35) + \left(\frac{1}{3}\right)(0.6) + \left(\frac{1}{3}\right)(0.5)\right], \\
&\quad (0.6)\left[\left(\frac{1}{3}\right)(0.05) + \left(\frac{1}{3}\right)(0.2) + \left(\frac{1}{3}\right)(0.5)\right] >
\end{aligned}
$$

$$
= \alpha < \frac{4}{15}, \frac{551}{1200}, \frac{3}{20} >
$$

$$
\begin{aligned}
&\approx < 0.304471931494, 0.524262607041, 0.171265461465 > \\
&\approx < f_{1:1} = low,\ f_{1:1} = med,\ f_{1:1} = high >
\end{aligned}
$$

(1.2) $P(X_2|e_1 = \text{“}notflooded\text{”}, e_2 = \text{“}notflooded\text{”})$

Let the filter probabilities $P(X_t|e_{1:t})$ be denoted by the filtering “message” $f_{1:t}$.

Next, recall that $P(X_{t+1}|e_{1:t+1}) = \alpha P(e_{t+1}|X_{t+1}) \sum_{x_t} P(X_{t+1}|x_t)P(x_t|e_{1:t})$.

It follows that $f_{1:t+1} = \alpha P(e_{t+1}|X_{t+1}) \sum_{x_t} P(X_{t+1}|x_t) f_{1:t}$

Using the filtering message $f_{1:1}$ found in the previous problem, then, we can determine the value of $f_{1:2} = P(X_2|e_1 = \text{“}notflooded\text{”}, e_2 = \text{“}notflooded\text{”})$ since matching evidence variables in the two problems are equal (i.e. $e_1 = \text{“}notflooded\text{”}$ in both problems).

Observe,

$$
\begin{aligned}
&P(X_2|e_1 = \text{“}notflooded\text{”}, e_2 = \text{“}notflooded\text{”}) \\
&\qquad = f_{1:2} \\
&\qquad = \alpha\ P(e_2 = \text{“}notflooded\text{”}|X_2) \sum_{x_1} P(X_2|x_1) f_{1:1} \\
&\qquad = \alpha\ P(e_2 = \text{“}notflooded\text{”}|X_2) \big[P(X_2|x_1 = low) f_{1:1=low} \\
&\qquad\qquad + P(X_2|x_1 = med) f_{1=med} + P(X_2|x_1 = high) f_{1:1=high}\big] \\
&\qquad = \alpha\ <\ P(e_2 = \text{“}notflooded\text{”}|X_2 = low)\big[P(X_2 = low|x_1 = low) f_{1:1=low} \\
&\qquad\qquad\qquad + P(X_2 = low|x_1 = med) f_{1:1=med} + P(X_2 = low|x_1 = high) f_{1:1=high}\big], \\
&\qquad\qquad P(e_2 = \text{“}notflooded\text{”}|X_2 = med)\big[P(X_2 = med|x_1 = low) f_{1:1=low} \\
&\qquad\qquad\qquad + P(X_2 = med|x_1 = med) f_{1:1=med} + P(X_2 = med|x_1 = high) f_{1:1=high}\big], \\
&\qquad\qquad P(e_2 = \text{“}notflooded\text{”}|X_2 = high)\big[P(X_2 = high|x_1 = low) f_{1:1=low} \\
&\qquad\qquad\qquad + P(X_2 = high|x_1 = med) f_{1:1=med} + P(X_2 = high|x_1 = high) f_{1:1=high}\big] > \\
&\qquad \approx\ \alpha\ < (1)\big[(0.6)(0.304471931494) + (0.2)(0.524262607041) + (0)(0.171265461465)\big], \\
&\qquad\qquad (0.95)\big[(0.35)(0.304471931494) + (0.6)(0.524262607041) + (0.5)(0.171265461465)\big], \\
&\qquad\qquad (0.6)\big[(0.05)(0.304471931494) + (0.2)(0.524262607041) + (0.5)(0.171265461465)\big] > \\
&\qquad \approx\ \alpha\ < 0.2875356803, 0.48141769743, 0.12342530922 > \\
&\qquad \approx < 0.3222126262,\ 0.5394769109,\ 0.1383104628 > \\
&\qquad \approx < f_{1:2} = low,\ f_{1:2} = med,\ f_{1:2} = high >
\end{aligned}
$$

(1.3) $P(X_3|e_1 = \text{"}notflooded\text{"}, e_2 = \text{"}notflooded\text{"}, e_3 = \text{"}flooded\text{"})$

Using the filtering message $f_{1:2}$ found in the previous problem, we can determine the value of $f_{1:3} = P(X_3|e_1 = \text{"}notflooded\text{"}, e_2 = \text{"}notflooded\text{"}, e_3 = \text{"}flodded\text{"})$ since matching evidence variables in the two problems are equal (i.e. $e_1 = \text{"}notflooded\text{"}$ in both problems).

Observe,

$$
\begin{aligned}
&P(X_3|e_1 = \text{"}notflooded\text{"}, e_2 = \text{"}notflooded\text{"}, e_3 = \text{"}flooded\text{"}) \\
&\qquad = f_{1:3} \\
&\qquad = \alpha\ P(e_3 = \text{"}flooded\text{"}|X_3) \sum_{x_2} P(X_3|x_2) f_{1:2} \\
&\qquad = \alpha\ P(e_3 = \text{"}flooded\text{"}|X_3) \big[P(X_3|x_2 = low) f_{1:2=low} \\
&\qquad\qquad + P(X_3|x_2 = med) f_{1:2=med} + P(X_3|x_2 = high) f_{1:2=high} \big] \\
&\qquad = \alpha\ <\ P(e_3 = \text{"}flooded\text{"}|X_3 = low) \big[P(X_3 = low|x_2 = low) f_{1:2=low} \\
&\qquad\qquad + P(X_3 = low|x_2 = med) f_{1:2=med} + P(X_3 = low|x_2 = high) f_{1:2=high} \big], \\
&\qquad\qquad P(e_3 = \text{"}flooded\text{"}|X_3 = med) \big[P(X_3 = med|x_2 = low) f_{1:2=low} \\
&\qquad\qquad + P(X_3 = med|x_2 = med) f_{1:2=med} + P(X_3 = med|x_2 = high) f_{1:2=high} \big], \\
&\qquad\qquad P(e_3 = \text{"}flooded\text{"}|X_3 = high) \big[P(X_3 = high|x_2 = low) f_{1:2=low} \\
&\qquad\qquad + P(X_3 = high|x_2 = med) f_{1:2=med} + P(X_3 = high|x_2 = high) f_{1:2=high} \big] > \\
&\qquad \approx\ \alpha\ < (0)\big[(0.6)(0.3222126262) + (0.2)(0.5394769109) + (0)(0.1383104628)\big], \\
&\qquad\qquad (0.05)\big[(0.35)(0.3222126262) + (0.6)(0.5394769109) + (0.5)(0.1383104628)\big], \\
&\qquad\qquad (0.4)\big[(0.05)(0.3222126262) + (0.2)(0.5394769109) + (0.5)(0.1383104628)\big] > \\
&\qquad \approx\ \alpha\ < 0, 0.02528078985, 0.07726449795 > \\
&\qquad \approx < 0,\ 0.246532926,\ 0.753467074 > \\
&\qquad \approx < f_{1:3} = low,\ f_{1:3} = med,\ f_{1:3} = high >
\end{aligned}
$$

Problem 2: (15 points)

Given the same sequence of evidence as problem 1, find the smoothed estimates for x_1, x_2 and x_3. (In other words, the evidence is still: $e_1 =$ "$notflooded$", $e_2 =$ "$notflooded$", $e_3 =$ "$flooded$")

Then plot the probabilities both filtering and smoothed estimates on the same graph. (Note: you will need two points/lines for a single probability, so overall you should have four lines.)

Begin by finding the three backwards messages $b_{4:3}, b_{3:3}, b_{2:3}$

Recall that, by definition (see Norvig p. 574), $b_{k+1:t} = P(e_{k+1:t}|X_k) = \sum_{x_{k+1}} P(e_{k+1}|x_{k+1})(b_{k+2:t})P(x_{k+1}|X_k)$

Observe:

$$\begin{aligned} b_{4:3} &= P(e_{4:3}|X_3) = P(\ |X_3)\mathbf{1} \qquad \text{// where } \mathbf{1} \text{ is the one's vector} \\ &= \mathbf{1} \end{aligned}$$

$$\begin{aligned} b_{3:3} &= \sum_{x_{2+1}} P(e_{2+1}|X_{2+1})(b_{2+2:3})P(X_{2+1}|X_2) \\ &= \sum_{x_3} P(e_3|X_3)(b_{4:3})P(X_3|X_2) \\ &= \sum_{x_3} P(e_3 = \text{"}flooded\text{"}|X_3)(\mathbf{1})P(X_3|X_2) \\ &= P(e_3 = \text{"}flooded\text{"}|x_3 = low)(\mathbf{1})P(x_3 = low|X_2) \\ &\quad + P(e_3 = \text{"}flooded\text{"}|x_3 = med)(\mathbf{1})P(x_3 = med|X_2) \\ &\quad + P(e_3 = \text{"}flooded\text{"}|x_3 = high)(\mathbf{1})P(x_3 = high|X_2) \\ \\ &= <[(0)(1)(0.6) + (0.05)(1)(0.35) + (0.4)(1)(0.05)], \\ &\qquad [(0)(1)(0.2) + (0.05)(1)(0.6) + (0.4)(1)(0.2)], \\ &\qquad [(0)(1)(0) + (0.05)(1)(0.5) + (0.4)(1)(0.5)]> \\ \\ &= <0.0375,\ 0.11,\ 0.225> \\ &= <b_{3:3} = low,\ b_{3:3} = med,\ b_{3:3} = high> \end{aligned}$$

$$\begin{aligned} b_{2:3} &= \sum_{X_{1+1}} P(e_{1+1}|X_{1+1})(b_{1+2:3})P(X_{1+1}|X_1) \\ &= \sum_{X_2} P(e_2|X_2)(b_{3:3})P(X_2|X_1) \\ &= \sum_{X_2} P(e_2 = \text{"}notflooded\text{"}|X_2)(b_{3:3})P(X_2|X_1) \\ &= P(e_2 = \text{"}notflooded\text{"}|x_2 = low)(b_{3:3})P(x_2 = low|X_1) \\ &\quad + P(e_2 = \text{"}notflooded\text{"}|x_2 = med)(b_{3:3})P(x_2 = med|X_1) \\ &\quad + P(e_2 = \text{"}notflooded\text{"}|x_2 = high)(b_{3:3})P(x_2 = high|X_1) \\ \\ &= <[(1)(0.0375)(0.6) + (0.95)(0.0375)(0.35) + (0.6)(0.0375)(0.05)], \\ &\qquad [(1)(.11)(0.2) + (0.95)(.11)(0.6) + (0.6)(.11)(0.2)], \\ &\qquad [(1)(0.225)(0) + (0.95)(0.225)(0.5) + (0.6)(0.225)(0.5)]> \\ \\ &= <0.03609375,\ 0.0979,\ 0.174375> \\ &= <b_{2:3} = low,\ b_{2:3} = med,\ b_{2:3} = high> \end{aligned}$$

Continued on next page ⟶

Next, find the smoothed estimates for X_1, X_2, and X_3:

By definition, the smoothed esimate for X_k is $P(X_k|e_{1:t}) = \alpha\ f_{1:k} \times b_{k+1:t}$.

The forward messages $f_{1:1}, f_{1:2}$, and $f_{1:3}$ were found in problem 1:

$$f_{1:1} \approx <0.304471931494,\ 0.524262607041,\ 0.171265461465>$$
$$f_{1:2} \approx <0.3222126262,\ 0.5394769109,\ 0.1383104628>$$
$$f_{1:3} \approx <0,\ 0.246532926,\ 0.753467074>$$

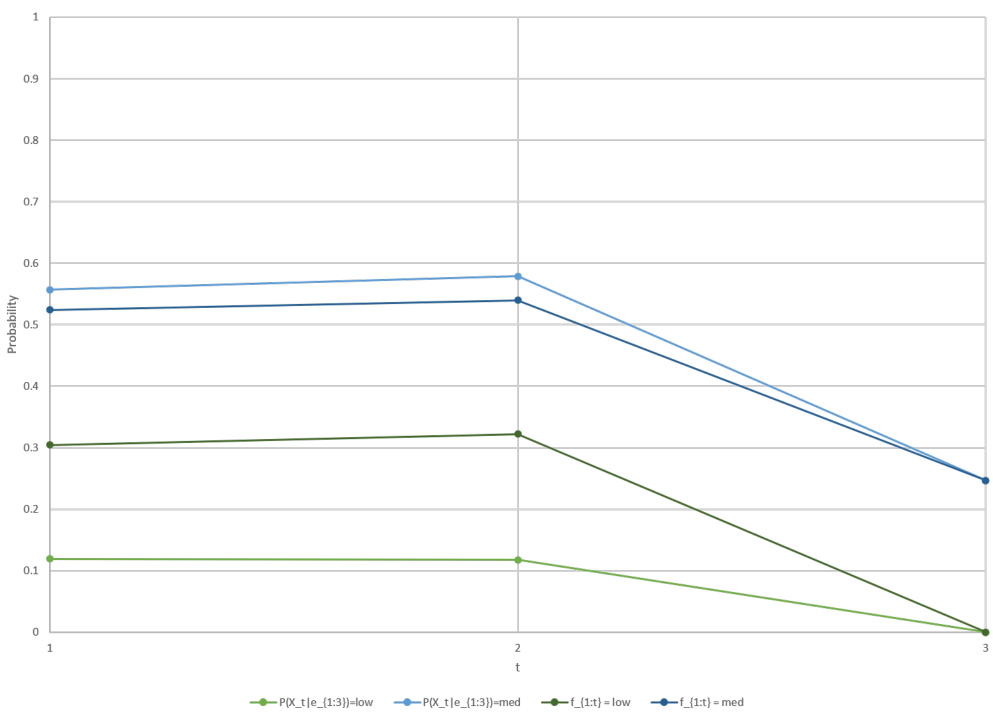
Therefore,

$$\begin{aligned}
P(X_1|e_{1:3}) &= \alpha\ f_{1:1} \times b_{2:3} \\
&\approx\ \alpha\ <0.304471931494,\ 0.524262607041,\ 0.171265461465> \ \times\ <0.03609375,\ 0.0979,\ 0.174375> \\
&\approx\ \alpha\ <(0.304471931494)(0.03609375),\ (0.524262607041)(0.0979),\ (0.171265461465)(0.174375)> \\
&\approx\ \alpha\ <0.01098953377,\ 0.05132530922,\ 0.02986441484> \\
&\approx <0.1192191609,\ 0.5567988985,\ 0.3239819407>
\end{aligned}$$

$$\begin{aligned}
P(X_2|e_{1:3}) &= \alpha\ f_{1:2} \times b_{3:3} \\
&\approx\ \alpha\ <0.3222126262,\ 0.5394769109,\ 0.1383104628> \ \times\ <0.0375,\ 0.11,\ 0.225> \\
&\approx\ \alpha\ <(0.3222126262)(0.0375),\ (0.5394769109)(0.11),\ (0.1383104628)(0.225)> \\
&\approx\ \alpha\ <0.01208297348,\ 0.0593424602,\ 0.03111985413> \\
&\approx <0.1178306068,\ 0.578695145,\ 0.3034742482>
\end{aligned}$$

$$\begin{aligned}
P(X_3|e_{1:3}) &= f_{1:3} \\
&\approx <0,\ 0.246532926,\ 0.753467074>
\end{aligned}$$

Note that $P(X_t|e_{1:3}) = high$ and $f_{1:t} = high$ are not included in the graph below, but can be still be found by using the fact that $\sum_{\omega \in \Omega} P(\omega) = 1$, where Ω is the set of all possible outcomes. This was done to keep the graph clearer.



Problem 3: (20 points)

Assume we are using the same HMM as problems 1 & 2, but we have different evidence:
$e_1 = \text{“}flooded\text{”}, e_2 = \text{“}notflooded\text{”}, e_3 = \text{“}notflooded\text{”}, e_4 = \text{“}flooded\text{”}, e_5 = \text{“}notflooded\text{”}$

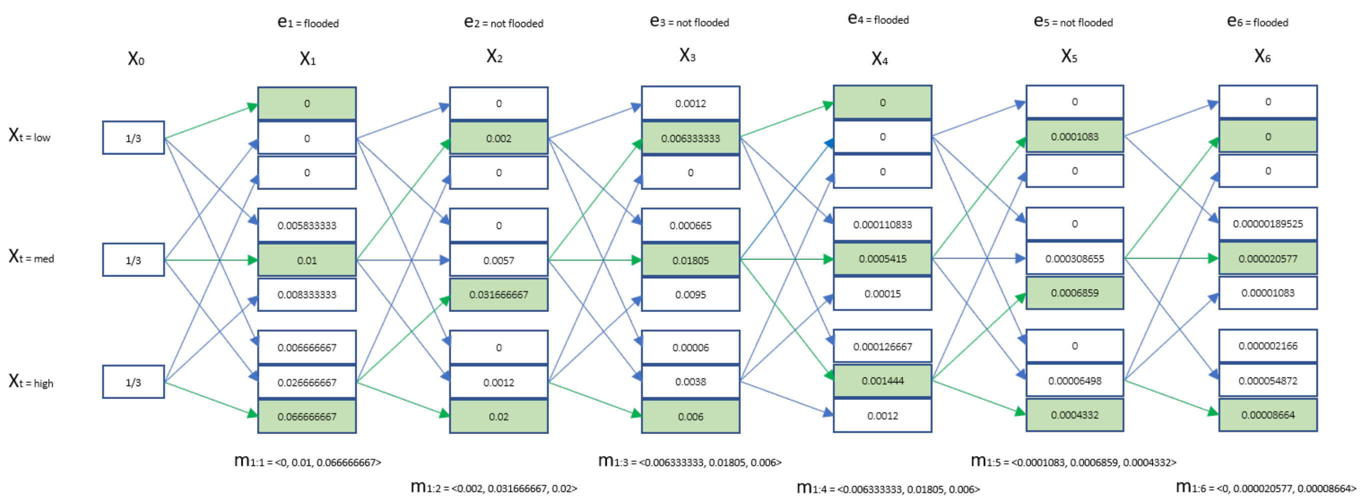
What is the most likely sequence of water table levels for these five days?
If you found that day 6 was “flooded” (i.e. $e_6 = \text{“}flooded\text{”}$), what is the most likely sequence now?

Let message $m_{1:t}$ denote:

$$\begin{aligned} m_{1:t} &= \max_{X_0 \ldots X_{t-1}} P(X_0, ..., X_t | e_{1:t}) \\ &= P(e_t|X_t) \max_{x_{t-1}} (P(X_t|X_{t-1})\ m_{1:t-1}) \end{aligned}$$

Additionally, let $m_{1:0} = P(X_0) = < \frac{1}{3}, \frac{1}{3}, \frac{1}{3} >$

Using this information, the following directed graph can be constucted from X_0 to X_6 with the green directed edges inidicating the best sequence leading to each of the states for any variable X_t. Note that each cell is found using the following product: $P(e_t|X_t)\ P(X_t|X_{t-1})\ m_{1:t-1}$



By tracing the green directed edges from the most likely state at variable X_5 back to X_1, we find that the most likely sequence of water table levels for days 1 to 5 given $e_1 = \text{“}flooded\text{”}, e_2 = \text{“}notflooded\text{”}, e_3 = \text{“}notflooded\text{”}, e_4 = \text{“}flooded\text{”}, e_5 = \text{“}notflooded\text{”}$ is:

$$[X_1 = high, X_2 = med, X_3 = med, X_4 = high, X_5 = high]$$

Similarily, the most likely sequence of water table levels for days 1 to 6 given $e_1 = \text{“}flooded\text{”}, e_2 = \text{“}notflooded\text{”}, e_3 = \text{“}notflooded\text{”}, e_4 = \text{“}flooded\text{”}, e_5 = \text{“}notflooded\text{”}, e_6 = \text{“}flooded\text{”}$ is:

$$[X_1 = high, X_2 = med, X_3 = med, X_4 = high, X_5 = high, X_6 = high]$$

Problem 4: (25 points)

Use particle filtering to estimate:
$P(x_{10}|e_1 = \text{``}notflooded\text{''}, e_2 = \text{``}notflooded\text{''}, e_3 = \text{``}flooded\text{''}, e_4 = \text{``}flooded\text{''}, e_5 = \text{``}notflooded\text{''}, e_6 = \text{``}notflooded\text{''}, e_7 = \text{``}notflooded\text{''}, e_8 = \text{``}notflooded\text{''}, e_9 = \text{``}flooded\text{''}, e_{10} = \text{``}notflooded\text{''})$

(i.e. the days flooded are 3,4 and 9. The rest are not flooded.) Give the number of particles used in your sampling, along with the probability for the water table values.

The code used to generate the following filtering approximation can be found in ParticleFiltering.py (submitted to HW3-code on Canvas):

Particle Filtering result with n = 10000000 samples: $P(X_{10}|e_{1:10}) \approx < 0.0608883, 0.6204157, 0.318696 >$

Therefore,

$$P(X_{10} = low \mid e_{1:10}) \approx 0.0608883$$
$$P(X_{10} = med \mid e_{1:10}) \approx 0.6204157$$
$$P(X_{10} = high \mid e_{1:10}) \approx 0.318696$$

Problem 5: (20 points)

Assume we are using the Frisbee example from class, where: $P(x_0) = N(0, 1)$ **and** $P(e_t|x_t) = N(x_t, 0.75)$ **How accurate do you need to be so that after 10 throws, the variance is not more than 10?**

Recall that $\sigma^2_{t+1} = \frac{(\sigma^2_t+\sigma^2_x)\sigma^2_z}{\sigma^2_t+\sigma^2_x+\sigma^2_z}$ (see Norvig p. 587)

We wish to find the values of σ^2_x for which $\sigma^2_{10} <= 10$. Given that $\sigma^2_z = 0.75$ and $\sigma^2_0 = 1$, however, there is no value of σ^2_x for which this expression is false. This statement can be verified using induction to prove that for any $t \geq 1$, $\sigma^2_t < 1$:

For the base case, let $t = 1$. Observe the following proof that $\sigma^2_1 < 1$:

Assume, to the contrary, that $\sigma^2_1 \geq 1$. It follows from the recursive defintion of σ^2_t that

$$\begin{aligned}
\sigma^2_1 = \frac{(\sigma^2_0+\sigma^2_x)\sigma^2_z}{\sigma^2_0+\sigma^2_x+\sigma^2_z} \geq 1 & \qquad \text{// note that all terms in } \sigma^2_0+\sigma^2_x+\sigma^2_z \text{ are positive since they are variances} \\
\iff (\sigma^2_0+\sigma^2_x)\sigma^2_z &\geq \sigma^2_0+\sigma^2_x+\sigma^2_z \\
\iff (\sigma^2_0\sigma^2_z)+(\sigma^2_x\sigma^2_z) &\geq \sigma^2_0+\sigma^2_x+\sigma^2_z \\
\iff (\sigma^2_0\sigma^2_z) &\geq \sigma^2_0+(\sigma^2_x-(\sigma^2_x\sigma^2_z))+\sigma^2_z \\
\iff (\sigma^2_0\sigma^2_z) &\geq \sigma^2_0+(1-\sigma^2_z)\sigma^2_x+\sigma^2_z \\
\iff (1)(0.75) &\geq (1)+(1-(0.75))\sigma^2_x+(0.75) \qquad \text{// substitute in given information} \\
\iff 0.75 &\geq 1.75+(0.25)\sigma^2_x \\
\iff 0 &\geq 1+(0.25)\sigma^2_x \\
\iff -1 &\geq (0.25)\sigma^2_x \\
\iff -4 &\geq \sigma^2_x
\end{aligned}$$

This, however, is a contradiction since σ^2_x must be non-negative as it represents the transition variance, and any variance must be non-negative. As such, the initial assumption that $\sigma^2_1 \geq 1$ must be false. Therefore, $\sigma^2_1 < 1$.

For the inductive case, we aim to prove that $\sigma^2_t < 1 \Rightarrow \ \sigma^2_{t+1} < 1$. Observe:

Assume, to the contrary, that $\sigma^2_{t+1} \geq 1$. It follows from the recursive definition of σ^2_{t+1} that

$$\begin{aligned}
\sigma^2_{t+1} = \frac{(\sigma^2_t+\sigma^2_x)\sigma^2_z}{\sigma^2_t+\sigma^2_x+\sigma^2_z} \geq 1 & \\
\iff (\sigma^2_t+\sigma^2_x)\sigma^2_z &\geq \sigma^2_t+\sigma^2_x+\sigma^2_z \\
\iff (\sigma^2_t)(\sigma^2_z)+(\sigma^2_x)(\sigma^2_z) &\geq \sigma^2_t+\sigma^2_x+\sigma^2_z \\
\iff 0 &\geq \sigma^2_t+\sigma^2_x+\sigma^2_z-(\sigma^2_t)(\sigma^2_z)-(\sigma^2_x)(\sigma^2_z) \\
\iff 0 &\geq \sigma^2_t-(\sigma^2_t)(\sigma^2_z)+\sigma^2_x-(\sigma^2_x)(\sigma^2_z)+\sigma^2_z \\
\iff 0 &\geq (1-\sigma^2_z)\sigma^2_t+(1-\sigma^2_z)\sigma^2_x+\sigma^2_z \\
\iff 0 &\geq (1-(0.75))\sigma^2_t+(1-(0.75))\sigma^2_x+(0.75) \qquad \text{// substitute in given info} \\
\iff 0 &\geq (0.25)\sigma^2_t+(0.25)\sigma^2_x+(0.75)
\end{aligned}$$

This, however, is a contradiction since both terms in the expression $(0.25)\sigma^2_t+(0.25)\sigma^2_x$ are nescessarily greater than or equal to 0 (any variance must be non-negative), and 0.75 is strictly greater than 0. As such, the initial assumption that $\sigma^2_{t+1} \geq 1$ must be false. Therefore, $\sigma^2_{t+1} < 1$.

Having proved the base case $\sigma^2_1 < 1$, and the inductive case $\sigma^2_t < 1 \Rightarrow \ \sigma^2_{t+1} < 1$ it follows, by the principle of induction, that for any $t \geq 1$, $\sigma^2_t < 1$.

Therefore, we have proven that the variance will be less than 10 after 10 throws, regardless of how accurate the Frisbee thrower is.

Problem 6: (15 points)

Suppose you have the simple HMM 3-node Bayes net shown below. Assume that $P(X_0)$ is uniformly distributed between [-1, 1]. The transition probability, $P(X_t+1|x_t)$, is $N(x_t, 1)$. The evidence probability is: $P(e_t|x_t) = N(x_t, 0.5)$
Use this description to approximate the distribution for $P(x_1|e_1 = 0)$ (using whatever method you prefer) and plot this distribution on a graph. Extend this network to 5 nodes to estimate the distribution for $P(x_2|e_1 = 0, e_2 = 0)$ and plot this distribution on a graph. As $t \to \infty$, describe what happens to the distribution of the filtering probabilities (assuming the evidence always says zero).

Since all of the distributions after $P(X_0)$ are gaussian, we can use the following recursive definitions to calculate the distribution mean and varance for any $t > 1$:

$\sigma_{t+1}^2 = \frac{(\sigma_t^2+\sigma_x^2)\sigma_z^2}{\sigma_t^2+\sigma_x^2+\sigma_z^2}$ and $\mu_{t+1} = \frac{(\sigma_t^2+\sigma_x^2)\ z_{t+1}+\sigma_z^2\ \mu_t}{\sigma_t^2+\sigma_x^2+\sigma_z^2}$ (see Norvig p. 587)

After substituting in the given values for this problem, we see that

$\sigma_{t+1}^2 = \frac{(\sigma_t^2+1)(0.5)}{\sigma_t^2+1+0.5} = \frac{\sigma_t^2+1}{2\sigma_t^2+3}$ and $\mu_{t+1} = \frac{(\sigma_t^2+1)(0)+(0.5)\ \mu_t}{\sigma_t^2+1+(0.5)} = \frac{(0.5)\ \mu_t}{\sigma_t^2+1.5} = \frac{\mu_t}{2\sigma_t^2+3}$

First consider the mean $\mu_{t+1} = \frac{\mu_t}{2\sigma_t^2+3}$. Notice that the denominator of the right-hand side of the expression is necessarily greater than or equal to 3. As such, each successive mean will be at most one-third the size of the previous mean. The limit of such an expression as $t \to \infty$ is clearly 0.

Next, consider the variance. For this expression to be in equilibrium, σ_{t+1}^2 must equal σ_t^2.

$$\begin{aligned}
& \sigma_{t+1}^2 = \sigma_t^2 \\
\iff\ & \frac{(0.5)\sigma_t^2 + 0.5}{\sigma_t^2 + 1.5} = \sigma_t^2 \\
\iff\ & (0.5)\sigma_t^2 + 0.5 = \sigma_t^2(\sigma_t^2 + 1.5) \\
\iff\ & 0 = \sigma_t^2(\sigma_t^2 + 1.5) - (0.5)\sigma_t^2 - 0.5 \\
\iff\ & 0 = (\sigma_t^2)^2 + \sigma_t^2 - 0.5 \\
\iff\ & \sigma_t^2 = \frac{-1 \pm \sqrt{1+2}}{2} \qquad \text{// using quadratic formula} \\
\iff\ & \sigma_t^2 = \frac{-1+\sqrt{3}}{2} \qquad \text{// variance must be positive} \\
& \approx 0.366025403784
\end{aligned}$$

We can see that the variance quickly converges to this value by using the approximation for the filtering distribution variance at time $t = 1$ found using particle filtering (see next page) to generate the next few variances (using the recursive definition above) for times t = 2 to 10:

t	1	2	3	4	5	6	7	8	9	10
σ_t^2	0.3674715	0.366129148	0.366032852	0.366025939	0.366025442	0.366025407	0.366025404	0.366025404	0.366025404	0.366025404

Therefore, as $t \to \infty$, the distribution of the filtering probabilites will converge to a mean of $\mu_t = 0$ and a variance of $\sigma_t^2 = \frac{\sqrt{3}-1}{2} \approx 0.366025403784$ assuming that the evidence is always zero.

Using particle filtering (n = 1000000), the two distibutions below–$P(X_1|e_1 = 0) \approx N(-0.0002219, 0.3674715)$, and $P(X_2|e_1 = 0, e_2 = 0) \approx N(-0.0001751, 0.3661570)$–were generated. Note that the code used to generate these distributions can be found in ParticleFiltering.py (submitted to HW3-code on canvas). The orange line is the normal distribution plotted with the same mean and variance as the sample distribution.