Problem 1:

a). The solution here is to use a slice sampler. In more detail, this means generate samples $(x_i, y_i)_{i=1,\cdots,n}$ by first sampling each $x_i \overset{iid}{\sim} p$ followed by sampling $y_i \overset{iid}{\sim}$ Unif$[0, p(x_i)]$.

b). Overfitting can occur when the training data is perfectly separable (as is always the case when there are more features than observations). This can manifest in many ways. One of which is that the norm of the weight vector grows without bound and the decision function approaches a step function at the boundary. Another is that the decision boundary may be poorly defined and an unfavorable (in terms of generalization out of set) one may be selected. Since there are many ways to answer this question, we will accept any of these explanations.

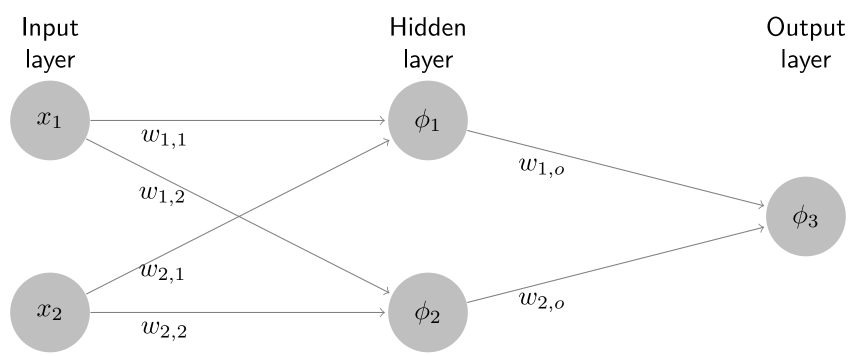
c). X_n and X_2 are clearly dependent. Any explanation which notes that X_2 depends on Z_2 and X_n depends on Z_n where Z_2 and Z_n are clearly dependent is acceptable.

Problem 2:

There are many valid answers to this problem, so you will need to spot check the student's solutions to see if they compute the correct output. One such solution is:

$$\{\phi_1(x) = \phi_2(x) = \phi_3(x) = I\{x > 0\}, w_{11} = w_{21} = 1, w_{12} = w_{22} = -1, w_{1o} = w_{2o} = 1\}$$

mapped onto the network:



Problem 3:

The observed data consists of one point in category 3. We can tell since the mass of the posterior is shifted towards category 3 as compared to the prior.

Problem 4: Neural network classifier

$x : \mathbb{I}(-3\frac{1}{\sqrt{2}} - \frac{1}{2\sqrt{2}}) = 0$

$x' : \mathbb{I}(\frac{1}{2\sqrt{2}} + \frac{1}{2\sqrt{2}} - \frac{1}{2\sqrt{2}}) = 1$

Problem 5: Rejection and importance sampling

a) M = 2

b) Acceptance Probability = 0.5. n = 2m

c) $\mathbb{E}_{x\sim p}[X] \approx \frac{1}{n}\sum x_i \frac{p(x_i)}{q(x_i)} = \frac{1}{n}\sum x_i(2x_i) = \frac{2}{n}\sum x_i^2$

$var_{x\sim q}(\frac{2}{n}\sum x_i^2) = \frac{4}{n^2}\sum var_{x\sim q}(x_i^2) = \frac{16}{45n^2}$

because $var_{x\sim q}(x_i^2) = \mathbb{E}_{x\sim q}(x_i^4) - \mathbb{E}_{x\sim q}(x_i^2)^2 = \frac{1}{5} - \frac{1}{9} = \frac{4}{45}$

Problem 6: Rejection sampling without rejection

a)

$$\begin{aligned}\mathbb{P}(X_i \leq x \,|\, R) = \mathbb{P}\left(X_i \leq x \,\middle|\, U_i \geq \tfrac{p(X_i)}{kr(X_i)}\right) &= \frac{\mathbb{P}\left(X_i \leq x\,,\, U_i \geq \frac{p(X_i)}{kr(X_i)}\right)}{\mathbb{P}\left(U_i \geq \frac{p(X_i)}{kr(X_i)}\right)} \\ &= \frac{\int_{-\infty}^{x}\left[\int_{p(x_i)/kr(x_i)}^{1} dy\right] r(x_i)\,dx_i}{\int_{-\infty}^{\infty}\left[\int_{p(x_i)/kr(x_i)}^{1} dy\right] r(x_i)\,dx_i} = \frac{\int_{-\infty}^{x} r(x_i) - \frac{p(x_i)}{k}\,dx_i}{\int_{-\infty}^{\infty} r(x_i) - \frac{p(x_i)}{k}\,dx_i} \\ &= \frac{k(\int_{-\infty}^{x} r(x_i) - \frac{p(x_i)}{k}\,dx_i)}{k-1}\end{aligned}$$

Therefore,

$$p(x|R) = \frac{kr(x) - p(x)}{k-1}$$

b)

$$\begin{aligned}\delta &:= \frac{1}{N}\sum_{i=1}^{N} h(X_i)\frac{p(X_i)}{q(X_i)} = \frac{1}{N}\left(\sum_{i\,:\,Z_i=A} h(X_i)\frac{p(X_i)}{p(X_i|A)} + \sum_{i\,:\,Z_i=R} h(X_i)\frac{p(X_i)}{p(X_i|R)}\right) \\ &= \frac{1}{N}\left(\sum_{i\,:\,Z_i=A} h(X_i) + \sum_{i\,:\,Z_i=R} h(X_i)\frac{(k-1)p(X_i)}{kr(X_i)-p(X_i)}\right)\end{aligned}$$