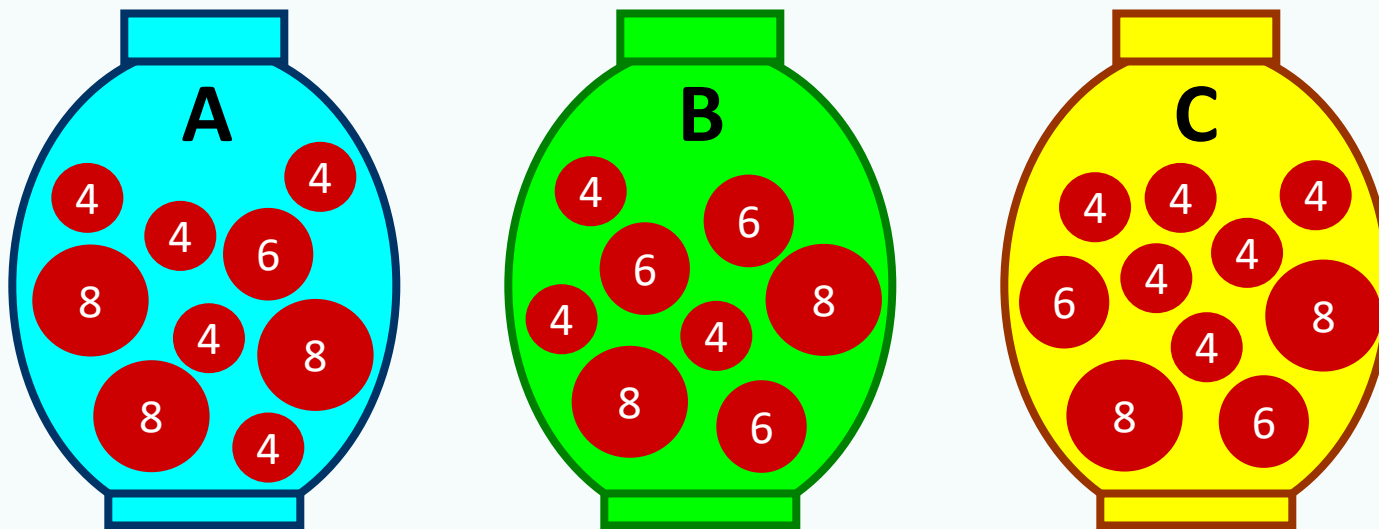


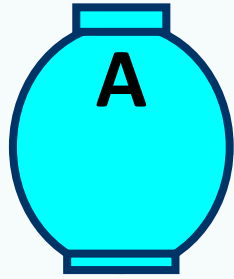
演習問題 2 (混合離散分布とEMアルゴリズム)

混合離散分布を推定する処理について以下の問いに答えよ.

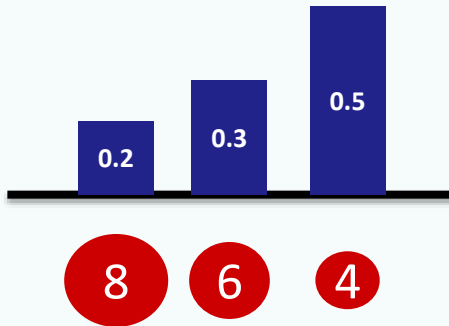
3つの壺 A, B, C があり, 各々の壺には重さが 8, 6, 4 のボールが入っている. いま, 壺 k に含まれるボールの重さは離散分布 ($P(8|k)$, $P(6|k)$, $P(4|k)$) に従うものとする. 適当な確率で壺を選んで, その壺からボールを取り出し, そのボールの重さを報告してまた壺に戻すという操作を繰り返す.



いま、各壺が選択される確率と各壺内のボールの重さの分布の初期値が以下のように与えられる。



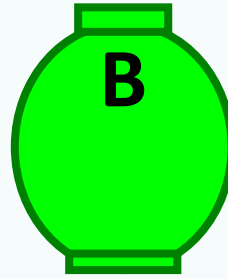
$$P(A) = 0.333$$



$$P(8 \mid A) = 0.2$$

$$P(6 \mid A) = 0.3$$

$$P(4 \mid A) = 0.5$$



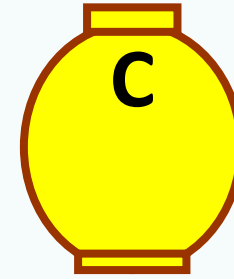
$$P(B) = 0.333$$



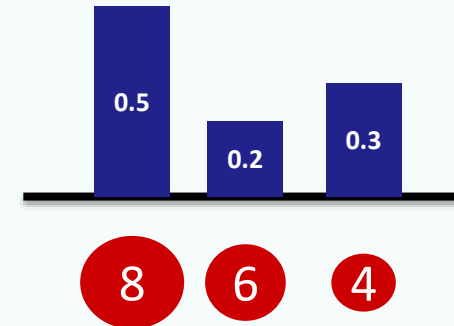
$$P(8 \mid B) = 0.4$$

$$P(6 \mid B) = 0.4$$

$$P(4 \mid B) = 0.2$$



$$P(C) = 0.333$$



$$P(8 \mid C) = 0.5$$

$$P(6 \mid C) = 0.2$$

$$P(4 \mid C) = 0.3$$

また、取り出されたボールの重さは以下のように報告された.



このとき、混合離散分布のモデルパラメータに相当する 各壺が選択される確率 および 各壺内のボールの重さの分布 をEMアルゴリズムにより再推定せよ.

解答) Eステップでは, 各壺が選択される事後確率 (EMカウント) を得るために, 完全データ尤度 (ボールの重さのみならず壺の名前が与えられた場合の尤度) を計算する.

$$P(8, A) = P(8 \mid A) P(A) = 0.2 \times 0.333 = 0.0666$$

$$P(8, B) = P(8 \mid B) P(B) = 0.4 \times 0.333 = 0.1332$$

$$P(8, C) = P(8 \mid C) P(C) = 0.5 \times 0.333 = 0.1665$$

$$P(6, A) = P(6 \mid A) P(A) = 0.3 \times 0.333 = 0.0999$$

$$P(6, B) = P(6 \mid B) P(B) = 0.4 \times 0.333 = 0.1332$$

$$P(6, C) = P(6 \mid C) P(C) = 0.2 \times 0.333 = 0.0666$$

$$P(4, A) = P(4 \mid A) P(A) = 0.5 \times 0.333 = 0.1665$$

$$P(4, B) = P(4 \mid B) P(B) = 0.2 \times 0.333 = 0.0666$$

$$P(4, C) = P(4 \mid C) P(C) = 0.3 \times 0.333 = 0.0999$$

各壺が選択される事後確率は以下のように計算される.

$$P(A|8) = P(A,8) / P(8) = P(A,8) / (P(A,8) + P(B,8) + P(C,8)) = 0.067 / (0.067 + 0.13 + 0.17) = 0.18$$

$$P(B|8) = P(B,8) / P(8) = P(B,8) / (P(A,8) + P(B,8) + P(C,8)) = 0.13 / (0.067 + 0.13 + 0.17) = 0.35$$

$$P(C|8) = P(C,8) / P(8) = P(C,8) / (P(A,8) + P(B,8) + P(C,8)) = 0.17 / (0.067 + 0.13 + 0.17) = 0.46$$

$$P(A|6) = P(A,6) / P(6) = P(A,6) / (P(A,6) + P(B,6) + P(C,6)) = 0.10 / (0.10 + 0.13 + 0.067) = 0.33$$

$$P(B|6) = P(B,6) / P(6) = P(B,6) / (P(A,6) + P(B,6) + P(C,6)) = 0.13 / (0.10 + 0.13 + 0.067) = 0.43$$













$$P(C|6) = P(C,6) / P(6) = P(C,6) / (P(A,6) + P(B,6) + P(C,6)) = 0.067 / (0.10 + 0.13 + 0.067) = 0.22$$

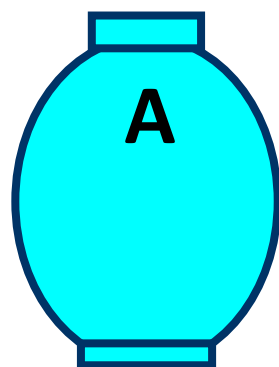
$$P(A|4) = P(A,4) / P(4) = P(A,4) / (P(A,4) + P(B,4) + P(C,4)) = 0.17 / (0.17 + 0.067 + 0.10) = 0.50$$

$$P(B|4) = P(B,4) / P(4) = P(B,4) / (P(A,4) + P(B,4) + P(C,4)) = 0.067 / (0.17 + 0.067 + 0.10) = 0.20$$

$$P(C|4) = P(C,4) / P(4) = P(C,4) / (P(A,4) + P(B,4) + P(C,4)) = 0.10 / (0.17 + 0.067 + 0.10) = 0.29$$

Mステップでは、既に得られている各壺が選択される事後確率（EMカウント）を用いて、壺Aに関する混合要素のパラメータ（壺Aが選択される確率および壺A内のボールの重さの分布）を以下のように更新できる.

												
A	.18	.33	.18	.50	.33	.33	.18	.50	.33	.18	.18	.33
B	.35	.43	.35	.20	.43	.43	.35	.20	.43	.35	.35	.43
C	.46	.22	.46	.29	.22	.22	.46	.29	.22	.46	.46	.22



Total responsibility : 3.6













$$P(A) = 3.6 / 12 = 0.30$$

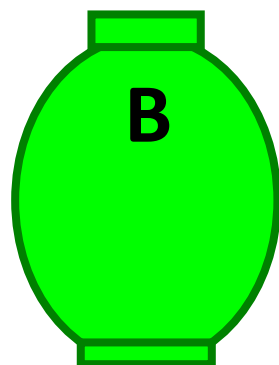
$$P(8 | A) = (0.18 * 5) / 3.6 = 0.25$$

$$P(6 | A) = (0.33 * 5) / 3.6 = 0.46$$

$$P(4 | A) = (0.50 * 2) / 3.6 = 0.28$$

Mステップでは、既に得られている各壺が選択される事後確率（EMカウント）を用いて、壺Bに関する混合要素のパラメータ（壺Bが選択される確率および壺B内のボールの重さの分布）を以下のように更新できる.

												
A	.18	.33	.18	.50	.33	.33	.18	.50	.33	.18	.18	.33
B	.35	.43	.35	.20	.43	.43	.35	.20	.43	.35	.35	.43
C	.46	.22	.46	.29	.22	.22	.46	.29	.22	.46	.46	.22



Total responsibility : 4.3

$$P(B) = 4.3 / 12 = 0.36$$

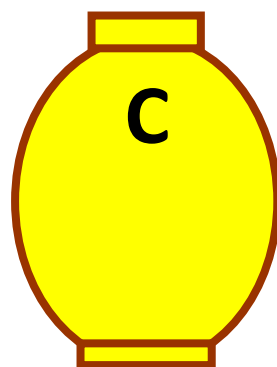
$$P(8 | B) = (0.35 * 5) / 4.3 = 0.41$$

$$P(6 | B) = (0.43 * 5) / 4.3 = 0.50$$

$$P(4 | B) = (0.20 * 2) / 4.3 = 0.093$$

Mステップでは、既に得られている各壺が選択される事後確率（EMカウント）を用いて、壺Cに関する混合要素のパラメータ（壺Cが選択される確率および壺C内のボールの重さの分布）を以下のように更新できる.

	8	6	8	4	6	6	8	4	6	8	8	6
A	.18	.33	.18	.50	.33	.33	.18	.50	.33	.18	.18	.33
B	.35	.43	.35	.20	.43	.43	.35	.20	.43	.35	.35	.43
C	.46	.22	.46	.29	.22	.22	.46	.29	.22	.46	.46	.22



Total responsibility : 4.0

$$P(C) = 4.0 / 12 = 0.30$$

$$P(8 \mid C) = (0.46 * 5) / 4.0 = 0.58$$

$$P(6 \mid C) = (0.22 * 5) / 4.0 = 0.28$$

$$P(4 \mid C) = (0.29 * 2) / 4.0 = 0.15$$