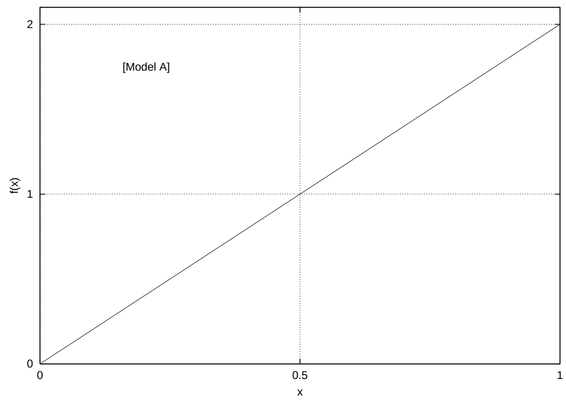
【演習問題】

ある確率事象についての観測データが〔表 1〕のように与えられている．この確率事象について以下の問に答えよ．

表 1: 確率変数 x の観測データ．x は 0 以上 1 以下の値域を有する

試行	1	2	3	4	5	6	7	8
x の観測値	0.8	0.75	0.6	0.25	0.8	0.4	0.2	0.9

【問１】
このデータを説明するため，A氏は〔図 1〕の右側で示される確率モデルの仮説を主張し，B氏は〔図 1〕左側の確率モデルによる仮説を主張している．モデルAとBのどちらが表 (1) のデータをより”もっともらしく”説明しているか？各モデルの尤度を計算して答えよ．
〔ヒント〕数値計算がやや煩雑だが，表の値を分数表記に直すなど計算を工夫すると良い．



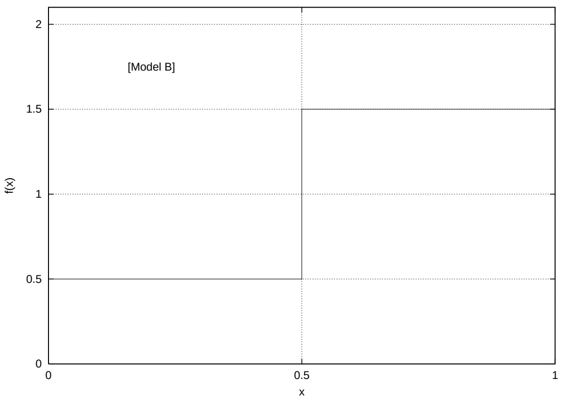


図 1: 確率密度関数モデルの図．縦軸は確率密度，横軸は確率変数を表す．左側のモデルAは線形関数による確率密度関数モデル，右側のモデルBは $x = [0, 0.5)$ の区間を $f(x) = 0.5$ の一様分布，$x = [0.5, 1]$ の区間を $f(x) = 1.5$ の一様分布とした確率密度関数モデル

【問２】
モデルBを改良し，$x = [0, 0.5)$ の区間は $f(x) = f_1$ の一様分布，$x = [0.5, 1)$ の区間は $f(x) = f_2$ の一様分布として，パラメータ f_1，f_2 を〔表 1〕のデータから最尤推定せよ．またそのときの尤度を求め，モデルAと修正したモデルBのどちらが良いモデルであるか判断せよ．
〔ヒント〕確率密度関数 $f(x)$ は，確率変数 x の値域が $0 \leq x \leq 1$ なので，$\int_0^1 f(x)dx = 1$ という制約条件がある．

【演習解答】

【問１】モデルAの尤度を L_A とすると,

$$L_A = (2\times 0.8)\times(2\times 0.75)\times(2\times 0.6)\times(2\times 0.25)\times(2\times 0.8)\times(2\times 0.4)\times(2\times 0.2)\times(2\times 0.9)$$

$$L_A = \frac{3^4\times 2^{15}}{10^6} = 1.327104$$

モデルBの尤度を L_B とすると，$x < 0.5$ のデータが 3 個で，このときの $f(x) = 0.5$ であり，また $0.5 \leq x$ のデータが 5 個で，このときの $f(x) = 1.5$ であるので,

$$L_B = (0.5)^3\times(1.5)^5 = \frac{3^5}{2^8} = \frac{243}{256} = 0.949$$

よって $L_A > L_B$ よりモデルAのほうがもっともらしい。

【問２】
$\int_0^1 f(x)dx = 1$ の制約条件より,

$$0.5f_1 + 0.5f_2 = 1$$

よって $f_2 = 2 - f_1$ である.
改良したモデルBの尤度を $L_{B'}$ とすると、

$$L_{B'} = f_1^3(2-f_1)^5 \quad ,$$

$$\frac{\partial}{\partial f_1}L_{B'} = 3f_1^2(2-f_1)^5 - 5f_1^3(2-f_1)^4 = f_1^2(2-f_1)^4\,(3(2-f_1)-5f_1) = f_1^2(2-f_1)^4\,(6-8f_1)$$

ここで，$L_{B'}$ を最大にする f_1 は $\frac{\partial}{\partial f_1}L_{B'} = 0$ を満たし，かつ $0 < f_1 < 1$ であるので，$f_1 = \frac{3}{4}$，$f_2 = \frac{5}{4}$.
このときの尤度は,

$$L_{B'} = \left(\frac{3}{4}\right)^3\left(2-\frac{3}{4}\right)^5 = 1.2875$$

である．よってモデルAの尤度のほうが大きいので、モデルAのほうが良い.