

乱数の生成

樋口さぶろお <https://hig3.net>

龍谷大学大学院 先端理工学研究科 数理・情報科学コース

計算科学特論 L09(2025-11-19 Wed)

最終更新: Time-stamp: "2025-11-19 Wed 13:46 JST hig"

今日の目標

- 確率変数の関数の確率密度関数を計算できる
- 逆関数法による乱数生成を説明できる



L08-Q1

Quiz 解答: 連続型の条件付き確率

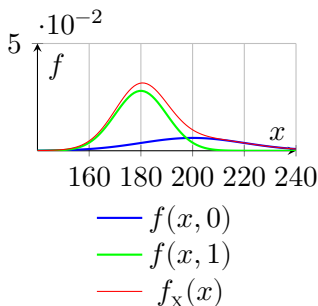
品種 A, B を $Y = 0, 1$, 実の質量を X とする.

周辺分布は $Y \sim B(1, \frac{7}{10})$.

条件付き分布は $X|Y = 0 \sim N(200, 20^2)$, $X|Y = 1 \sim N(180, 10^2)$.

①

$$\begin{aligned}
 f_X(x) &= \sum_{y=0,1} f_{X|Y}(x|y) f_Y(y) \\
 &= \frac{1}{(2\pi 20^2)^{1/2}} e^{-\frac{(x-200)^2}{2 \cdot 20^2}} \cdot \frac{3}{10} + \frac{1}{(2\pi 10^2)^{1/2}} e^{-\frac{(x-180)^2}{2 \cdot 10^2}} \cdot \frac{7}{10}.
 \end{aligned}$$



2

$$\begin{aligned}
 p' = f_{Y|X}(0|190) &= \frac{f_{X|Y}(190|0)f_Y(0)}{f_X(190)} \\
 &= \frac{\frac{1}{(2\pi 20^2)^{1/2}} e^{-\frac{(190-200)^2}{2 \cdot 20^2}} \cdot \frac{3}{10}}{\frac{1}{(2\pi 20^2)^{1/2}} e^{-\frac{(190-200)^2}{2 \cdot 20^2}} \cdot \frac{3}{10} + \frac{1}{(2\pi 10^2)^{1/2}} e^{-\frac{(190-180)^2}{2 \cdot 10^2}} \cdot \frac{7}{10}} = 0.238.
 \end{aligned}$$

すなわち、 Y の事前分布は $B(1, \frac{7}{10})$, 事後分布は $B(1, 1 - p')$.

L08-Q2

Quiz 解答: 混合ガウス分布のナীবベイズ

$$\begin{aligned}
 P(Y = 0|X = 1) &= \frac{\frac{1}{(2\pi 2^2)^{1/2}} e^{-\frac{(5-2)^2}{2 \cdot 2^2}} \cdot \frac{3}{10}}{\frac{1}{(2\pi 2^2)^{1/2}} e^{-\frac{(5-2)^2}{2 \cdot 2^2}} \cdot \frac{3}{10} + \frac{1}{(2\pi (1/2)^2)^{1/2}} e^{-\frac{(5-6)^2}{2 \cdot (1/2)^2}} \cdot \frac{7}{10}} \\
 &= \frac{\frac{3}{20} \phi\left(\frac{3}{2}\right)}{\frac{3}{20} \phi\left(\frac{3}{2}\right) + \frac{7}{5} \phi(2)}
 \end{aligned}$$

L08-Q3

Quiz 解答: 多項分布の確率

- ① 四面体サイコロを 5 回振ったときの目の出方は、各面の出方の確率に基づく多項分布に従います。具体的には、各目 $i = 1, 2, 3, 4$ の出る回数 x_i は次のように表されます。

$$(x_1, x_2, x_3, x_4) \sim M(5, 0.3, 0.2, 0.4, 0.1)$$

②

$$P(x_1 = 3, x_2 = 0, x_3 = 1, x_4 = 1) = \frac{5!}{3!0!1!1!} \cdot 0.3^3 \cdot 0.2^0 \cdot 0.4^1 \cdot 0.1^1 = 0.02$$

L08-Q4

Quiz 解答: ディリクレ分布の確率

- ① $(1, 0, 0), (0, 1, 0), (0, 0, 1)$ を 3 頂点とする三角形の周と内部。
 ② 三角形のうち $p_1 \geq 0$ の部分は、相似比 $\frac{1}{2}$ の三角形で、その面積は、もとの三角形の面積の $(1/2)^2 = \frac{1}{4}$ 。

乱数 random number とは

確率変数 X に対応する乱数 (列) random number (sequence) とは, X の実現値と見なせる (同じ分布にしたがう) 数 (列).

各項が独立で, 長さ n の列なら, X のサイズ n の標本.

乱数生成器 RNG=random number generator 乱数を生成する物理的または計算機を用いた仕組み

擬似乱数しかし, 計算機は決定論的な機械なので, 実は独立にはできない. ‘一見したところ’ 独立そうに見えるもので許す, という意味で**擬似乱数** pseudo-random number’ とよぶ.

多くの処理系で, 基本的な $Y \sim U(0, 1)$ に対応する擬似乱数 ($[0, 1)$ 一様乱数) を与える関数が準備されている. 以下では, C で `double getuniform(void)` がそれだとする.

$[0, 1)$ 一様乱数を与える関数は多くの処理系やライブラリにある.

NumPy `np.random.random_sample(1)`

SciPy Statsmodels `stat.uniform(loc=c, scale=d-c)`

PyMC `pm.Uniform('Y', lower=0.0, upper=1.0)`

以下では, $[0, 1)$ 一様乱数はあったとして, そこから, 他の分布にしたがう乱数を構成する方法 (がなかなかないこと) を知る.

NumPy, SciPy.statsmodels, PyMC などでは, これらの方法により, 他の有名な分布にしたがう擬似乱数を与える関数も用意されている.

やりたいことの例

ベイズ推定にでてきた，この確率密度関数で，乱数 $(\mu^{(1)}, \mu^{(2)}, \mu_0)$ を生成したい！

$$\propto e^{-\frac{(55-\mu^{(1)})^2}{2 \cdot 2^2}} e^{-\frac{(57-\mu^{(1)})^2}{2 \cdot 2^2}} e^{-\frac{(\mu^{(1)}-\mu_0)^2}{2 \cdot 3^2}} e^{-\frac{(61-\mu^{(1)})^2}{2 \cdot 2^2}} e^{-\frac{(65-\mu^{(1)})^2}{2 \cdot 2^2}} e^{-\frac{(\mu^{(1)}-\mu_0)^2}{2 \cdot 3^2}} e^{-\frac{(\mu_0-50)^2}{2 \cdot 10^2}}.$$

困難

- 複雑… $\leadsto$ 実は多次元正規分布にすぎない
- 3変数… $\leadsto$ 実は多次元正規分布にすぎない

多次元正規分布が共役事前分布になってればずっとハッピー，だけど一般にそうはいかない

任意の確率密度関数にしたがう乱数を生成するにはどうするの？

ここまで来たよ

8 潜在変数モデルと混合ガウス分布

9 乱数の生成

- 確率変数の関数
- 逆関数法
- 棄却法

L09-Q1

Quiz(確率変数の変換)

あるクッキーマシンの作る正方形のクッキーの面積 (生地 の 量) Q は, 次の確率密度関数にしたがう (単位省略).

$$f_Q(q) = \begin{cases} \frac{1}{36} & (64 \leq q < 100) \\ 0 & (\text{他}) \end{cases}$$

クッキーの一辺の長さは $R = g(Q) = \sqrt{Q}$ で与えられる (単位省略).

- ① Q の母平均値と母分散を求めよう.
- ② 確率 $P(Q > 82)$ を求めよう.
- ③ $f_R(r)$ を求めよう.
- ④ R の母平均値と母分散を求めよう (2つの方法で).
- ⑤ 確率 $P(R > 9)$ を求めよう (2つの方法で).

確率変数の関数 I

L09-Q2

Quiz(確率変数の変換)

一様分布 $U(0, 1)$ に従う連続型確率変数 Y と, $R = g(Y) = 2\sqrt{Y}$ で定まる連続型確率変数 R を考える.

- ① $E[R], V[R]$ を求めよう.
- ② 母比率 (確率) $P(0.2 < R < 0.8)$ を求めよう.
- ③ 母比率 (確率) $F_R(r) = P(R < r)$ を求めよう.
- ④ R の確率密度関数 $f_R(r)$ を求めよう.

R の乱数生成は簡単.

```
1 double getrandom(double y){  
2     double r;  
3     r=2*sqrt(y);  
4     return r;  
5 }  
6 r=getrandom(getuniform());
```

標本

y	$r = 2\sqrt{y}$
0.00	0.00
0.49	1.40
$\vdots$	$\vdots$
0.81	1.80

復習 (累積分布関数)

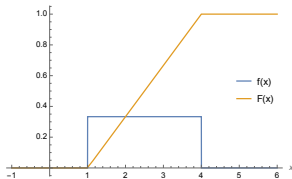
岩薩林 §4.1(4.4)(4.5)

確率密度関数 (累積) 分布関数

$$f(x) \xrightarrow{\text{積分}} F(x) = \int_{-\infty}^x f(s) ds = P(X < x)$$

$$f(x) = \frac{dF}{dx} \xleftarrow{\text{微分}} F(x)$$

U(1,4)



例

確率変数の関数

連続型確率変数 Y に対して, 単調増加な $g(y)$ で連続型確率変数 $X = g(Y)$ を定める. このとき, 次が成立するはず.

$$E[g(Y)] = E[X]$$

$$F_Y(b) - F_Y(a) = P(a \leq Y \leq b) = P(g(a) \leq X \leq g(b)) = F_X(g(b)) - F_X(g(a))$$

$$F_Y(y) = F_X(g(y))$$

これまで $E[aY + b]$ とか考えてたのは $X = g(Y) = aY + b$ を考えてたことに相当. 岩薩林 (3.6)

$f_Y(y)$ はわかってて, 右辺を求めろと言われたときの方針

- 左辺に直して計算する.
- X の確率密度関数 $f_X(x)$ を求めてから, 右辺で計算する.

Y	a	$\rightarrow$	b
$Z = g(Y)$	$g(a)$	$\rightarrow$	$g(b)$

$f_X(x)$ と $f_Y(y)$ との関係

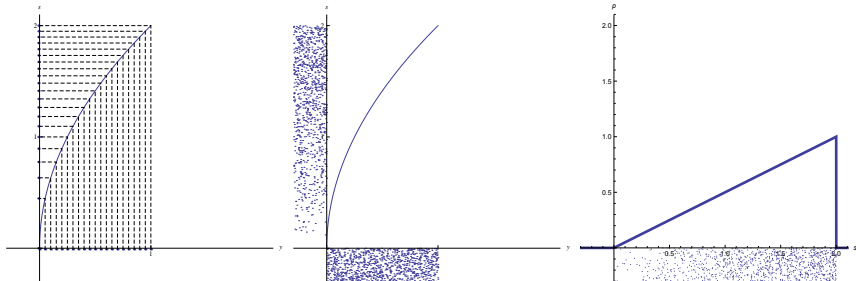
$x = g(y)$ のとき, F の関係式の両辺を y で微分して,

$$f_Y(y) = f_X(g(y)) \cdot \frac{dg}{dy}(y)$$

確率密度関数の変換のおぼえ方

$x = g(y)$ のとき,

$$f_Y(y) dy = f_X(x) dx$$



L09-Q3

Quiz(確率変数の変換)

一様分布 $U(0, 1)$ に従う連続型確率変数 Y と, $X = g(y) = (d - c)Y + c$ で定まる連続型確率変数 X を考える.

X の確率密度関数 $f_X(x)$ を求めよう.

ここまで来たよ

8 潜在変数モデルと混合ガウス分布

9 乱数の生成

- 確率変数の関数
- 逆関数法
- 棄却法

確率変数の関数

逆関数法

$Y \sim U(0, 1)$, $X \sim$ 与えられた確率密度関数 $f_x(x)$ のとき, $x = g(y)$ で関係づけたい.

$g(y)$ を求めるには, 確率密度関数の変換の式を解けばよい.

$g(y)$ が求まれば, 連続型一様分布 $U(0, 1)$ にしたがう Y を g で変換して, $f_x(x)$ にしたがう X を生成できる.

L09-Q4

Quiz(逆関数法)

確率密度関数

$$f_x(x) = \begin{cases} \frac{3\sqrt{2}}{8}\sqrt{x} & (0 \leq x < 2) \\ 0 & (\text{他}) \end{cases}$$

に従う確率変数 X に対応する乱数を, $[0, 1)$ 一様乱数 y から $x = g(y)$ で作りたい. $g(y)$ を求めよう.

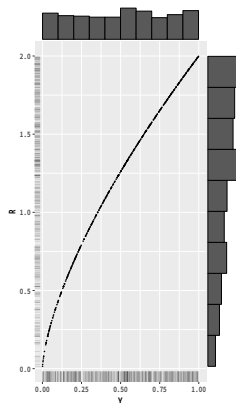
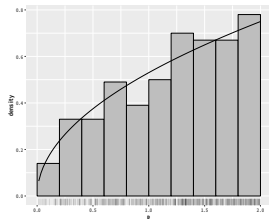
$[0, 1)$ 一様乱数=連続一様分布 $U(0, 1)$ にしたがう乱数


```

1 double getRandom(double y){
2     return 2*pow(y, 2.0/3);
3 }

```

出力



L09-Q5

Quiz(逆関数法)

確率密度関数

$$f_R(r) = \begin{cases} \frac{1}{3} & (2 \leq r < 5) \\ 0 & (\text{他}) \end{cases}$$

に従う確率変数 R に対応する乱数を, $[0, 1)$ 一様乱数 y から $r = g(y)$ で作りたい. $g(y)$ を求めよう.

ここまで来たよ

8 潜在変数モデルと混合ガウス分布

9 乱数の生成

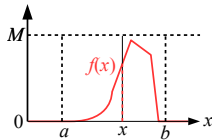
- 確率変数の関数
- 逆関数法
- 棄却法

棄却法による乱数生成

連続型確率変数 X の確率密度関数 $f(x)$ が次を満たすとする.

$$f(x) = 0 \quad (x < a, b \leq x)$$

$$0 \leq f(x) \leq M \quad (a \leq x < b)$$



このとき, X に対応する乱数を次で生成できる (効率はおもかく).

棄却法

```

1 double f(double x); /* プログラムが定義 */
2 double getRandom() {
3     double x, y; double a, b, M; /* プログラムが定義 */
4     while(1) {
5         x = a + (b - a) * getuniform(); /* X ~ U(a, b) */
6         z = M * getuniform();           /* Z ~ U(0, M), Z/M = R ~ U(0, 1) */
7         if (z < f(x)) {
8             break; /* 受理=accept */
9             /* 受理の確率 P(Z < f(x)) = P(R < f(x)/M) = f(x)/M */
10        } /* 棄却=reject, やりなおし */
11    }
12    return x;
13 }

```

理由 だって、 x を得る確率密度って、 $f(x)$ に比例してるでしょ。
直観的にもわかるけど、下のような計算方法もある。

- 棄却 0 回で x を得る確率密度 $\frac{1}{M(b-a)} f(x)$
- 棄却 1 回で x を得る確率密度 $(1 - \frac{1}{M(b-a)}) \frac{1}{M(b-a)} f(x)$
- $\vdots$
- 棄却 n 回で x を得る確率密度 $(1 - \frac{1}{M(b-a)})^n \frac{1}{M(b-a)} f(x)$
- $\vdots$

この排反事象列の和事象の確率密度は、無限等比級数の和を考えて、

$$\frac{1}{M(b-a)} f(x) \times \frac{1}{1 - (1 - \frac{1}{M(b-a)})} = f(x).$$

L09-Q6

Quiz(棄却法による乱数生成)

確率密度関数

$$f(x) = \begin{cases} \frac{1}{21}x^2 & (1 \leq x < 4) \\ 0 & (\text{他}) \end{cases}$$

に従う乱数を、棄却法を利用して返す関数 `double getrandom(void)` を C で書こう。

平均して、1 回の実行 (乱数 1 個の生成) に、`double getuniform()` の何回の呼び出しを要するか考えよう。

C でも Python でも．タイプでも手書きでも．