

# 確率密度

## 確率・統計 第3回

村田 昇

2020年8月23日

## 確率測度と確率密度

実数上に値を取る一般の確率測度を具体的に表現するためには、確率密度による積分表現を用いることになる。

### 歪んだ目盛りのルーレットの例

- 問題

起点 0 から周に沿った長さ  $l$  の位置の目盛りの値を

$$x = l^2$$

としてルーレットを作る。

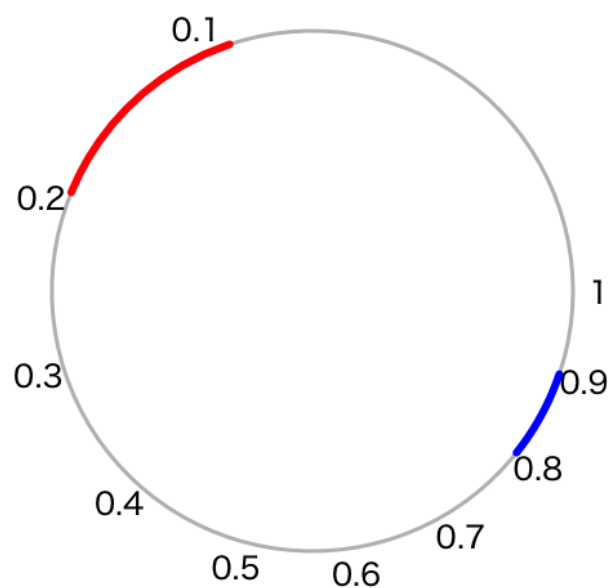
事象として同じ長さの区間  $[0.1, 0.2]$ ,  $[0.8, 0.9]$  を考えたとき、その確率

$$P([0.1, 0.2]), \quad P([0.8, 0.9])$$

の大小関係はどのような関係になるか?

- 解答

$$P([0.1, 0.2]) > P([0.8, 0.9])$$



- 考え方

これまで扱ってきた単純なルーレットに対応づけて考えるればよい.

- 歪んだ目盛りの指す区間:  $[a, b]$
- これまでのルーレットでの区間:  $[\sqrt{a}, \sqrt{b}]$

したがって確率測度は以下で計算される.

$$P([a, b]) = \sqrt{b} - \sqrt{a}$$

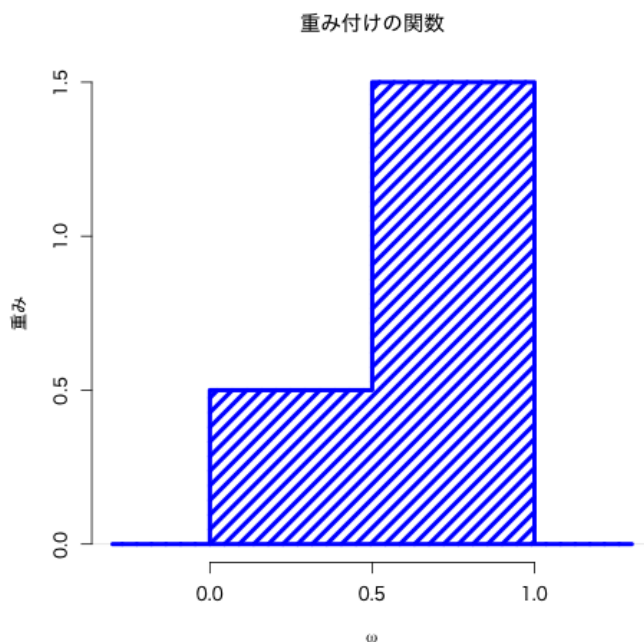
## 一様でない抜き出しの例

- 問題

「 $\Omega = [0, 1]$  の点を異なる確率で抜き出す」という試行において, 区間  $[0, 0.5)$  から点が抜き出される事象と, 区間  $[0.5, 1]$  から点が抜き出される事象を比較すると, 前者より後者が3倍起こりやすい状況を考える.

このとき確率測度をどのように定義すれば良いか?

- 解答



$$f(x) = \begin{cases} 0.5, & (0 \leq x < 0.5) \\ 1.5, & (0.5 \leq x \leq 1) \end{cases}$$

として, 確率測度を

$$P(A) = \int_A f(\omega) d\omega$$

で定義する.

- 確率の偏りの確認

例えば区間  $[0.1, 0.2]$  と  $[0.8, 0.9]$  を比べてみる

$$P([0.1, 0.2]) = \int_{0.1}^{0.2} f(\omega) d\omega = [0.5\omega]_{0.1}^{0.2} = 0.05$$

$$P([0.8, 0.9]) = \int_{0.8}^{0.9} f(\omega) d\omega = [1.5\omega]_{0.8}^{0.9} = 0.15$$

## 確率測度の条件 (再掲)

- 以下の 3 つの条件が成立する

$$(P.1) \quad P(A) \geq 0, \quad A \in \mathcal{F}$$

$$(P.2) \quad P\left(\sum_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n), \quad A_n \in \mathcal{F}$$

$$(P.3) \quad P(\Omega) = 1$$

## 積分による表現の正当性

- (P.1)  $f(x)$  の正値性から明らか
- (P.2) 積分区間を分割しても積分値が変わらないという積分の性質から明らか
- (P.3) 区間  $[0, 1]$  上の積分

$$P([0, 1]) = \int_0^1 f(\omega) d\omega = 0.5 \times 0.5 + 0.5 \times 1.5 = 1$$

より全確率が 1 となっている

## 確率密度

- 定義

確率測度がある関数の積分

$$P(A) = \int_A f(\omega) d\omega$$

で書かれているとき, その被積分関数を**確率密度 (関数)** (probability density function) という.

また確率密度を持つ確率測度を**絶対連続** (absolutely continuous) な分布と呼ぶ.

## 確率密度の例

- 歪んだ目盛りのルーレット

歪んだ目盛をもつルーレットの確率密度は, 任意の  $a, b$  ( $0 < a < b \leq 1$ ) に対して

$$P([a, b]) = \int_a^b f(\omega) d\omega = \sqrt{b} - \sqrt{a}$$

となる以下の関数  $f$  である.

$$f(\omega) = \begin{cases} \frac{1}{2\sqrt{\omega}}, & (0 < \omega \leq 1) \\ 0, & (\text{それ以外}) \end{cases}$$

## 演習

- 起点 0 から周に沿った長さ  $l$  の位置に対して, 以下のような歪んだ目盛り  $x$  を付けたルーレットの確率密度を求めよ.

- $x = \sqrt{l}$
- $x = -\log(l)$
- $x = \sin^{-1}(l)$  ( $0 < x \leq \pi/2$ )

## ヒント

- 歪んでいないルーレットと対応づければ良い.

$$\begin{aligned}P(a < x < b) &= P(a < g(l) < b) = P(g^{-1}(a) < l < g^{-1}(b)) \\&= g^{-1}(b) - g^{-1}(a) \\&= \int_a^b f(\omega) d\omega\end{aligned}$$

( $g$  が単調増加の場合, 減少する場合は符号に注意)

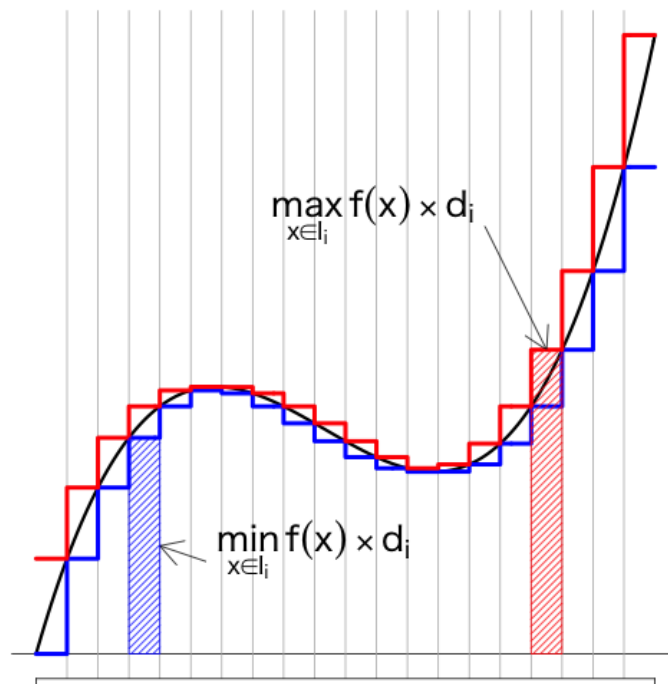
## 異なる2つの積分

### 積分の考え方

- Riemann 積分 (区分求積法に基づく)
- Lebesgue 積分 (確率測度に基づく)

### Riemann 積分

- 定義



有界な積分領域  $K$  を  $n$  個の区間  $I_i$  (幅は  $d_i$ ) に分割する. 各区間における関数  $f(x)$  の最小値と最大値を用いて有限和  $s_n, S_n$  を次のように定義する.

$$\begin{aligned}s_n &= \sum_{i=1}^n \min_{x \in I_i} f(x) \times d_i \\S_n &= \sum_{i=1}^n \max_{x \in I_i} f(x) \times d_i\end{aligned}$$

各区間の大小関係から

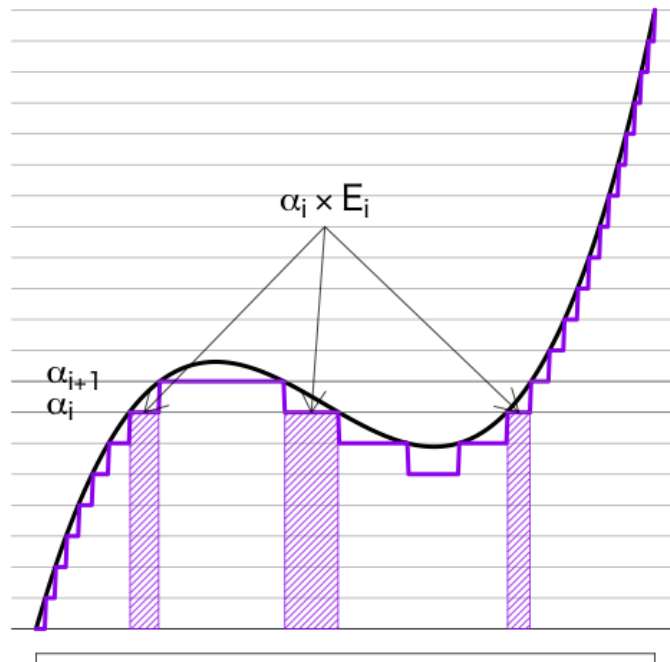
$$s_n \leq S_n$$

が成り立っているが、分割を細かくしたとき、それぞれの極限が存在して、なおかつ一致するならば、これを  $f(x)$  の **Riemann 積分** (Riemann integral) という。

$$\lim_{n \rightarrow \infty} s_n = \lim_{n \rightarrow \infty} S_n = \int_K f(x) dx$$

## Lebesgue 積分

- 定義



区間  $K$  上で定義された関数  $f(x)$  の値域を  $n$  個の区間に分割する。区間の端点を  $\alpha_i$  とし、 $f(x)$  が区間  $[\alpha_i, \alpha_{i+1})$  で値を取る  $x$  の領域を  $E_i \subset K$  で表し、 $E_i$  上で値が  $\alpha_i$  となる階段関数  $f_n(x)$  (単関数) を考える。

このとき  $\mu$  を Lebesgue 測度とし、単関数  $f_n(x)$  の **Lebesgue 積分** を

$$\int_K f_n(x) dx = \sum_{i=1}^n \alpha_i \mu(E_i)$$

で定義する。分割を細かくしたとき  $f_n(x)$  は  $f(x)$  に近づいていくが、この極限が存在するならば、これを  $f(x)$  の **Lebesgue 積分** (Lebesgue integral) という。

$$\lim_{n \rightarrow \infty} \int_K f_n(x) dx = \int_K f(x) dx$$

## 表記に関する注意

- Lebesgue 積分特有の記法

2つの積分を明示的に区別するために、Lebesgue 積分を

$$\int_K f(x) \mu(dx), \quad (\mu \text{ は Lebesgue 測度})$$

と書くことがある。

- この講義での以降の積分は、特に断わらないかぎり Lebesgue 積分を考える

## 2つの積分の違い

- 特異な例

$[0, 1]$  上で次のような関数

$$f(x) = \begin{cases} 1, & x \text{ が有理数} \\ 0, & x \text{ が無理数} \end{cases}$$

を考える.

- この関数は Dirichlet の関数と呼ばれ

$$f(x) = \lim_{n \rightarrow \infty} \lim_{k \rightarrow \infty} \cos^{2k}(n! \pi x)$$

と書けることが知られている.

- Riemann 積分の場合

どんな短い区間  $I_i$  にも有理数と無理数は混在するので

$$\min_{x \in I_i} f(x) = 0, \quad \max_{x \in I_i} f(x) = 1$$

$$s_n = \sum_{i=0}^n 0 \times d_i = 0$$

$$S_n = \sum_{i=0}^n 1 \times d_i = 1$$

であり,  $n$  をいくら大きくしても極限は一致しない. したがってこの関数の Riemann 積分は存在しない.

- Lebesgue 積分の場合

適当な値域の分割において

$$\alpha_i \leq 0 < \alpha_{i+1}, \quad \alpha_j \leq 1 < \alpha_{j+1}$$

となるのであれば, Lebesgue 測度の性質より

$$\int f_n(x) dx = \alpha_i \times \mu(\text{無理数}) + \alpha_j \times \mu(\text{有理数}) = \alpha_i \cdot 1 + \alpha_j \cdot 0 = \alpha_i$$

となるが,  $n \rightarrow \infty$  のとき  $\alpha_i \rightarrow 0, \alpha_j \rightarrow 1$  となるので

$$\int f(x) dx = 0 \times \mu(\text{無理数}) + 1 \times \mu(\text{有理数}) = 0$$

となり Lebesgue 積分可能である.

## 積分の一致する条件

- 定理

有界閉区間  $K$  で  $f$  が連続なら,  $K$  上の  $f$  の Lebesgue 積分の値は Riemann 積分の値と一致する.

- 応用上重要な問題に現れる「あまり特殊でない関数」ではほとんど一致すると考えてよい

## 演習

- 以下の関数が  $\mathbb{R}$  上の確率密度関数となるように  $a$  の値を定めよ.

1.

$$f(x) = \begin{cases} 0 & (x < 0) \\ \exp(ax) & (x \geq 0) \end{cases}$$

2.

$$g(x) = \begin{cases} 0 & (|x| > 1) \\ a \cos(\pi x/2) & (|x| \leq 1) \end{cases}$$

## ヒント

- 見本空間で関数が正値で、積分が 1 となればよい.

1.

$$\int_0^{\infty} \exp(ax) dx = \left[ \frac{\exp(ax)}{a} \right]_0^{\infty} = \frac{-1}{a} = 1$$

2.

$$\int_{-1}^1 a \cos\left(\frac{\pi x}{2}\right) dx = \left[ a \frac{2}{\pi} \sin\left(\frac{\pi x}{2}\right) \right]_{-1}^1 = a \frac{4}{\pi} = 1$$

## 確率密度関数の例

### 正規分布 (normal/Gaussian distribution)

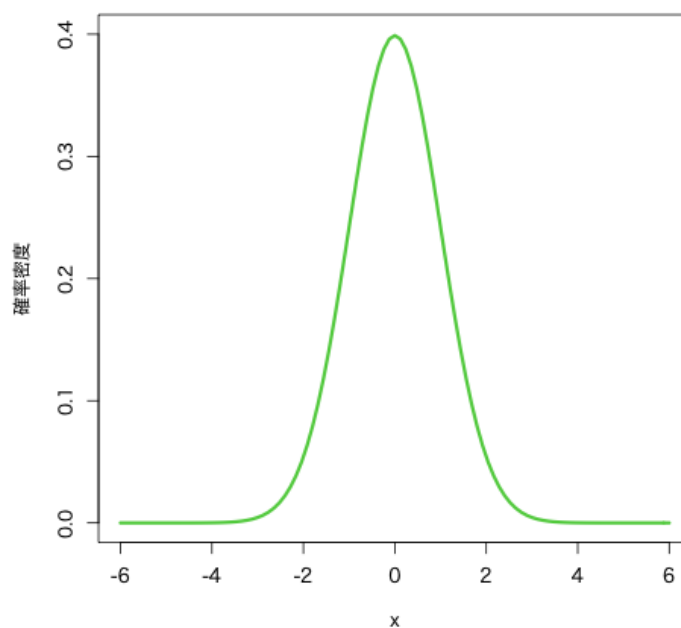


図 1: 正規分布 (平均 0, 分散 1)

- 見本空間:  $(-\infty, \infty)$
- 母数: 平均  $\mu$ , 分散  $\sigma^2$
- 密度関数:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

- 備考:  $\mu = 0, \sigma = 1$  のとき **標準正規分布** と呼ぶ.

## 一様分布 (uniform distribution)

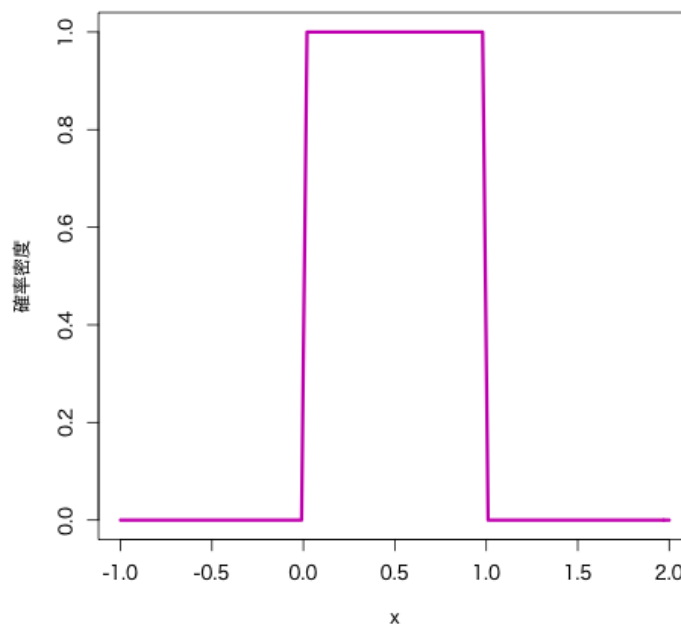


図 2: 一様分布 (区間  $[0, 1]$ )

- 見本空間: 区間  $[a, b]$
- 母数: 区間の端点  $a, b$
- 密度関数:

$$\frac{1}{b-a} \quad (b \leq x \leq a)$$

- 備考: ルーレット回しの確率密度関数に相当する.

## Cauchy 分布 (Cauchy distribution)

- 見本空間:  $(-\infty, \infty)$
- 母数: 位置 (location)  $\mu$ , 尺度 (scale)  $\sigma$
- 密度関数:

$$f(x) = \frac{1}{\pi\sigma} \left( 1 + \frac{(x-\mu)^2}{\sigma^2} \right)^{-1}$$

- 備考: 裾の重い分布の典型として用いられる. 正規分布の比から導かれる.



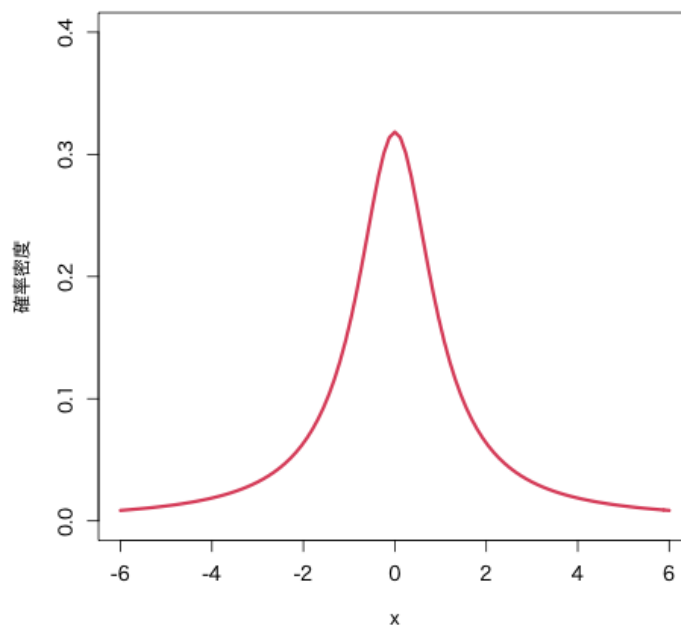


図 3: Cauchy 分布 (位置 0, 尺度 1)

### $\chi^2$ -分布 ( $\chi^2$ distribution)

- 見本空間:  $[0, \infty)$
- 母数: 自由度  $\nu$
- 密度関数:

$$f(x) = \frac{1}{2^{\nu/2}\Gamma(\frac{\nu}{2})} x^{\nu/2-1} e^{-x/2}$$

$$\Gamma(z) = \int_0^\infty e^{-t} t^{z-1} dt$$

- 備考:  $\nu$  個の標準正規分布の 2 乗和の分布で, 検定に利用される.

### $t$ -分布 (Student's t distribution)

- 見本空間:  $(-\infty, \infty)$
- 母数: 自由度  $\nu$
- 密度関数:

$$f(x) = \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\nu\pi}\Gamma(\frac{\nu}{2})} \left(1 + \frac{x^2}{\nu}\right)^{-\frac{1}{2}(\nu+1)}$$

- 備考: 標準正規分布と自由度  $\nu$  の  $\chi^2$  -分布の比の分布で, 検定に利用される.

### 指数分布 (exponential distribution)

- 見本空間:  $[0, \infty)$
- 母数: 比率 (rate)  $\lambda > 0$

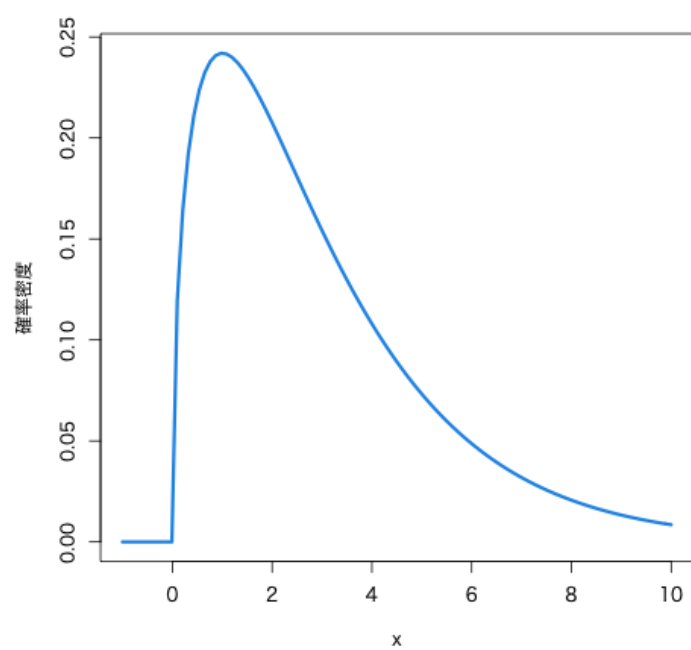


图 4:  $\chi^2$  -分布 (自由度 3)

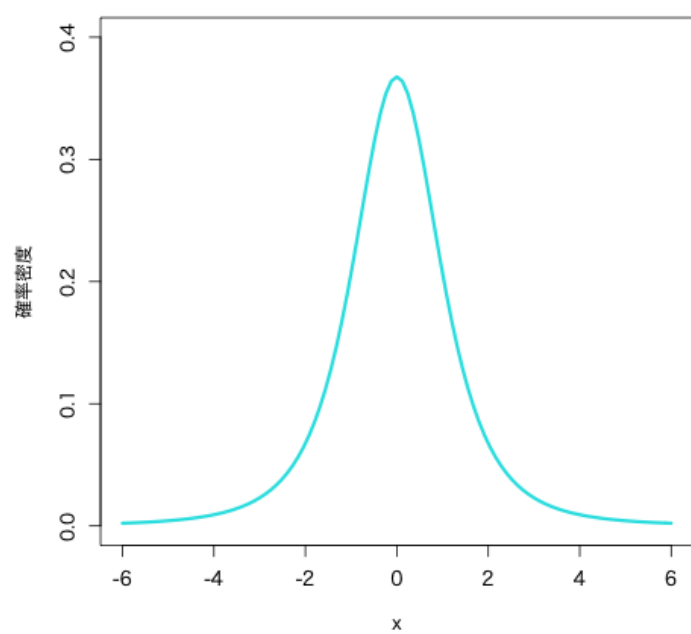


图 5:  $t$  -分布 (自由度 3)

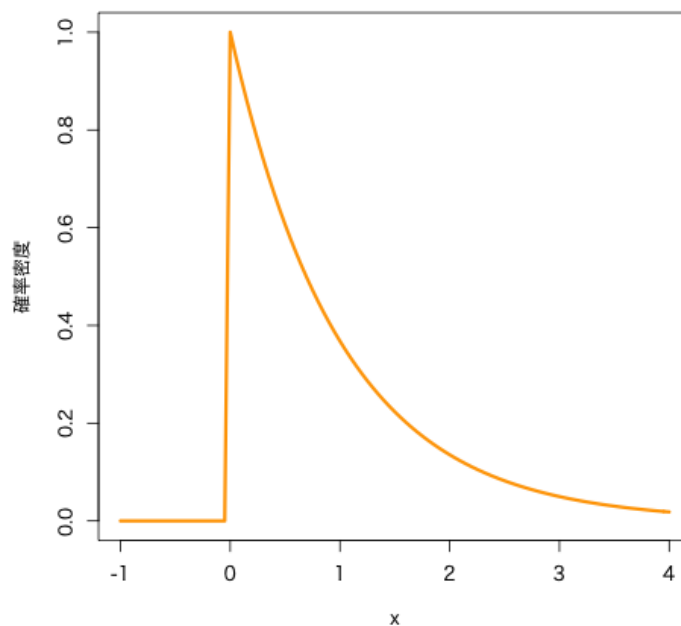


図 6: 指数分布 (比率 1)

- 密度関数:

$$f(x) = \frac{1}{\lambda} e^{-\frac{x}{\lambda}}$$

- 特徴: タクシーなどの待ち時間のモデルとして利用される。

## 確率密度がない分布

### 確率密度が定義されない例 (1/2)

- $\delta$  分布

確率測度を

$$P(A) = \delta_a(A) = \begin{cases} 1, & a \in A \\ 0, & a \notin A \end{cases}$$

で定義する。これは

$$(a \text{ が抜き出される確率}) = 1$$

$$(\Omega - a \text{ の点が抜き出される確率}) = 0$$

であるから、 $\Omega$  から 1 点抜き出すとそれは **ほとんど確実に**  $a$  であるような抜き出しを考えていることになる。

### 確率密度が定義されない例 (2/2)

- 有理数の抜き出し

有理数  $\mathbb{Q}_{[0,1]}$  全体に番号をふり, これを  $q_1, q_2, \dots$  で表す.

$$P(A) = \sum_{i: q_i \in A} 2^{-i}$$

という確率測度を考える. 例えば

$$P(\mathbb{Q}_{[0,1]}) = 1, \quad P(\Omega - \mathbb{Q}_{[0,1]}) = 0$$

である.

この確率測度においては確率密度を考えることができない. なぜなら 1 点  $q_i$  の起こる確率は

$$P\{q_i\} = 2^{-i}$$

であるので, 密度関数  $f$  は

$$P\{q_i\} = \int_{\{q_i\}} f(x) \mu(dx) = f(q_i) \times \mu(q_i) = 2^{-i}$$

とならなくてははいけませんが, 1 点の Lebesgue 測度は 0 なので

$$f(q_i) \times \mu(q_i) = f(q_i) \times 0 = 0$$

となり,  $f$  を構成することは不可能である.

## 演習

- 正規分布の確率密度関数

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

が  $(-\infty, \infty)$  で積分すると 1 となることを確かめよ.

- 次の関数が確率密度となるように  $a$  を定めよ.

$$h(x) = a \exp(-4x^2) \quad (-\infty < x < \infty)$$

## ヒント

- 2 つの正規分布の積を考えて, 重積分を極座標に変換すればよい. 簡単のため変数変換した

$$g(x) = e^{-x^2}$$

を考えると以下のように計算できる.

$$\begin{aligned} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x)g(y)dx dy &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-x^2-y^2} dx dy \\ &= \int_0^{2\pi} \int_0^{\infty} e^{-r^2} r dr d\theta \end{aligned}$$

## まとめ

- 確率密度
  - 確率密度の積分による確率測度の表現
  - 確率密度の満たすべき条件
  - Riemann 積分と Lebesgue 積分
  - いろいろな確率分布