#79 Tokyo.R

2019.06.29

BeginneR Session
- 確率の基礎 -

@kilometer00

Who！？

Who！？

名前： 三村 @kilometer

職業： ポスドク (こうがくはくし)

専門： 行動神経科学(霊長類)

脳イメージング

医療システム工学

R歴： ~ 10年ぐらい

流行: Netflix

2018.06.09 Tokyo.R #70
BeginneR Session – Bayesian Modeling

2018.07.15 Tokyo.R #71
Landscape with R – the Japanese R community

2018.10.20 Tokyo.R #73
BeginneR Session – Visualization & Plot

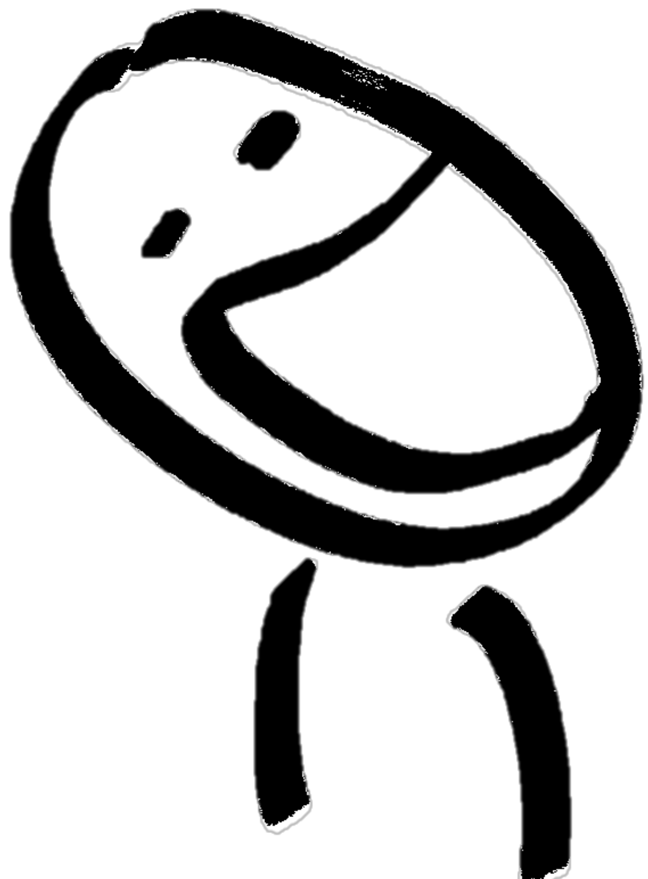
2019.01.19 Tokyo.R #75
BeginneR Session – Data pipeline

2019.03.02 Tokyo.R #76
BeginneR Session – Data pipeline

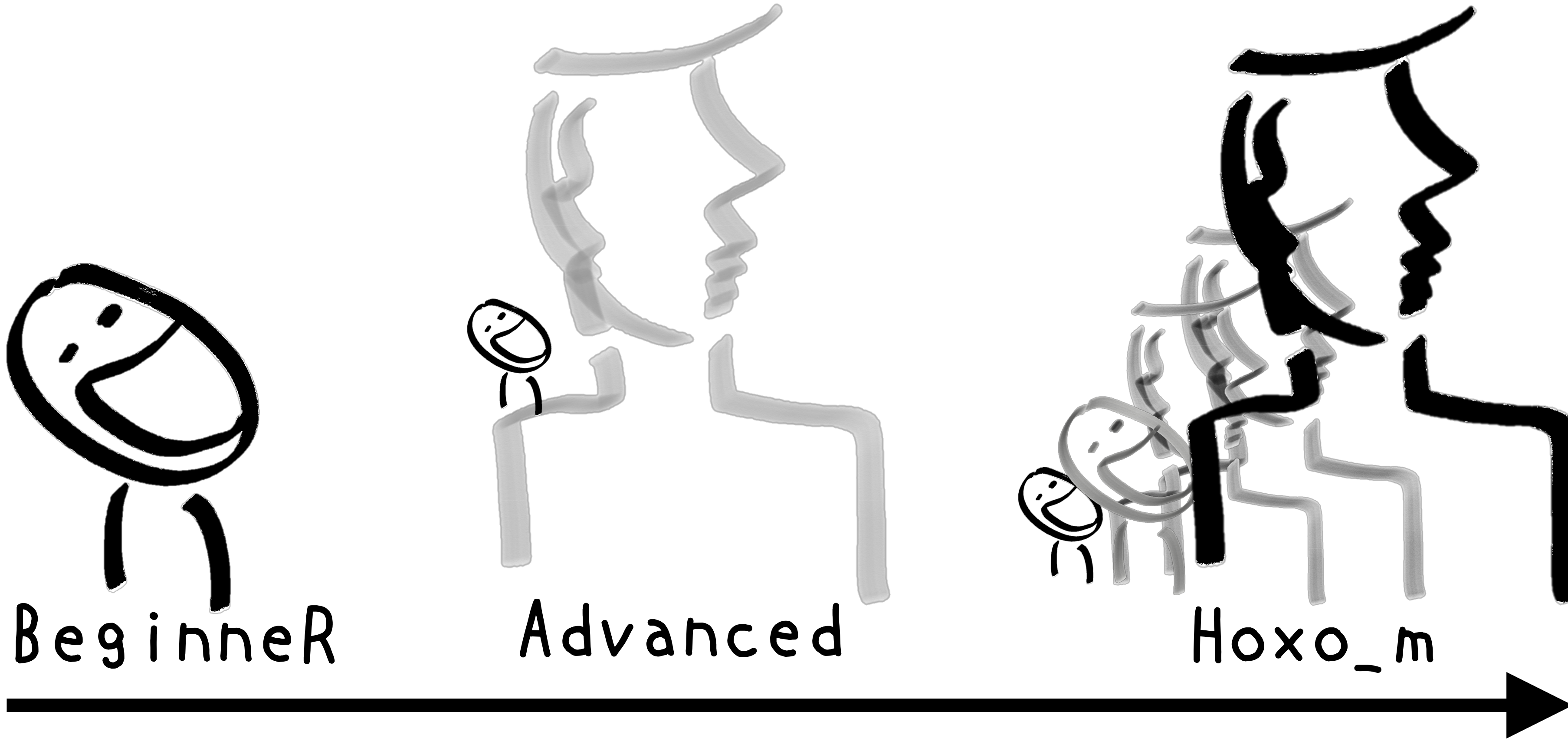
2019.04.13 Tokyo.R #77
BeginneR Session – Data analysis

2019.05.25 Tokyo.R #78
BeginneR Session – Data analysis

BeginneR Session

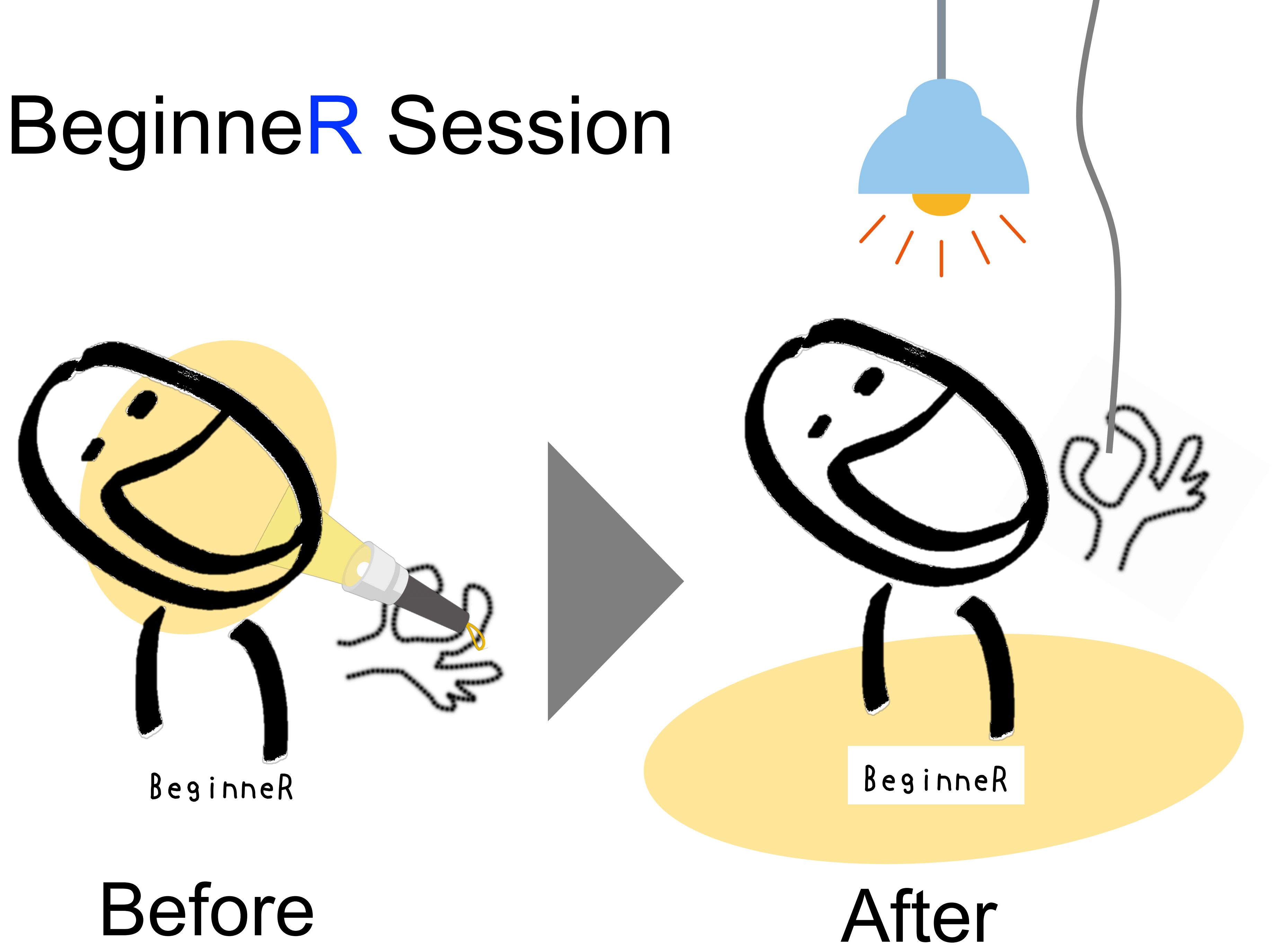


BeginneR



If I have seen further it is by standing on the shoulders of Giants.

-- Sir Isaac Newton, 1676


BeginneR Session
BeginneR
Before
BeginneR
After

BeginneR Session
- 確率の基礎 -

今回は

新作です！

- スライドが日本語です
- 難しいかも
- 分かりやすさを優先するため単純化しました。

- スライドが日本語です
- 難しいかも
- 分かりやすさを優先するため単純化し~~ました~~ていません(ちょっとだけ)。

で、

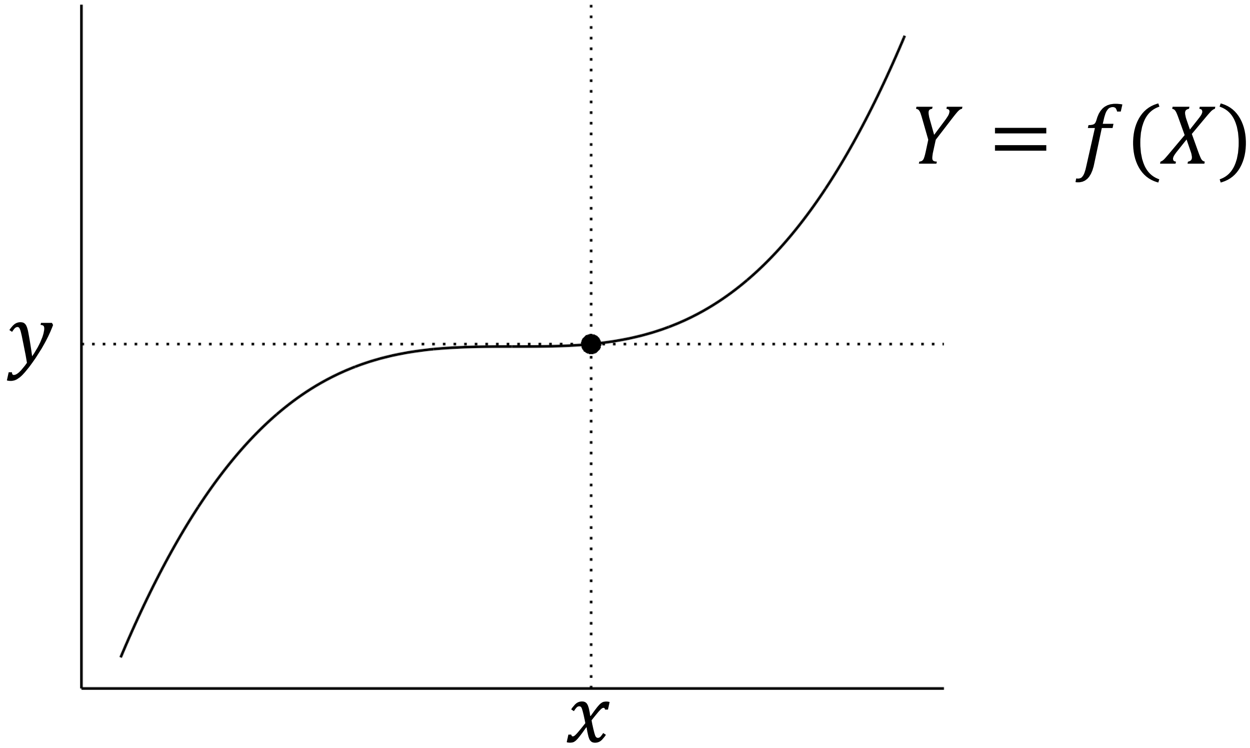
確率変数Xは正規分布$\mathcal{N}$に従う

【今日の目標】

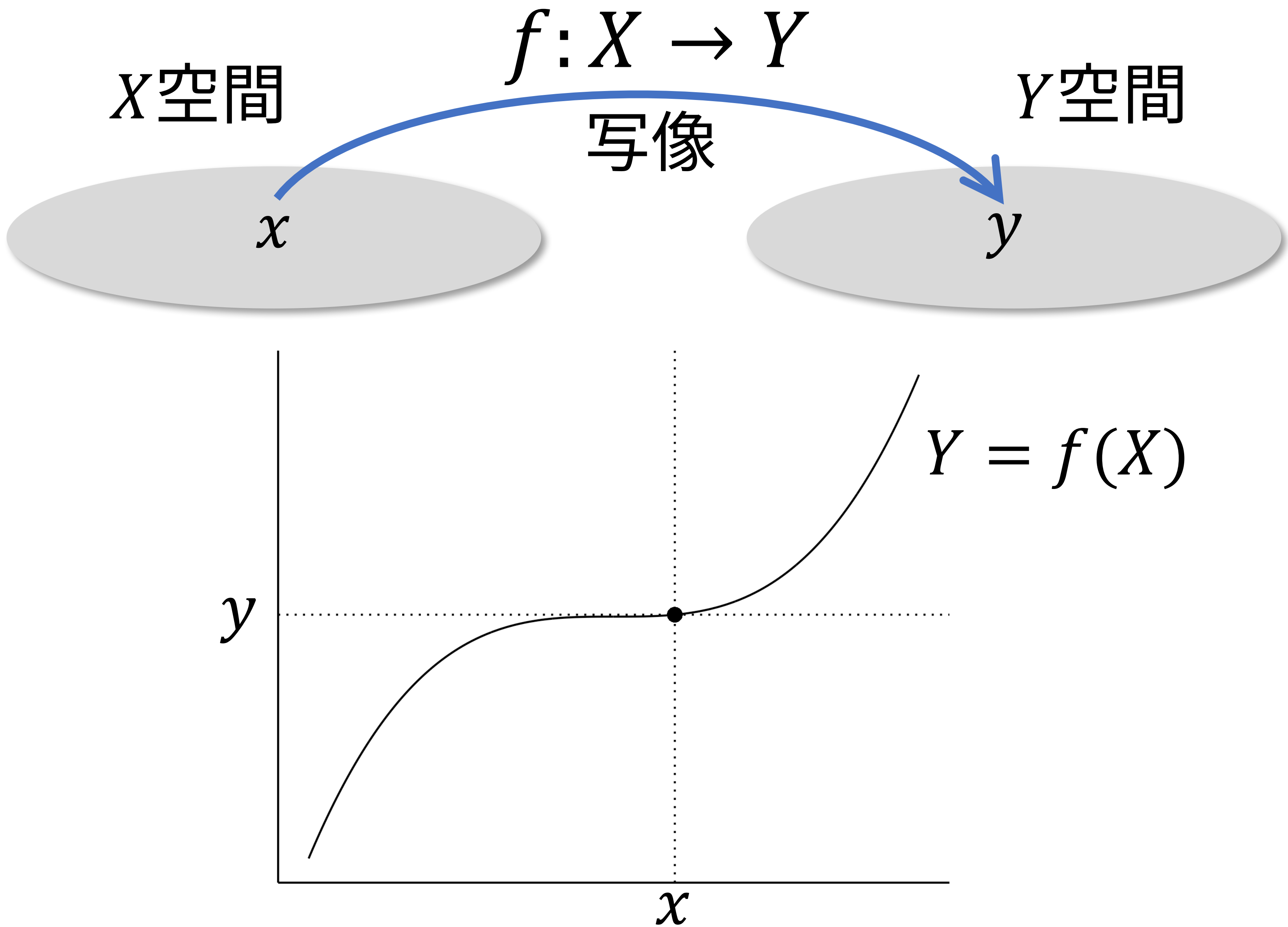
確率変数Xは正規分布$\mathcal{N}$に従う

を完全に理解する。

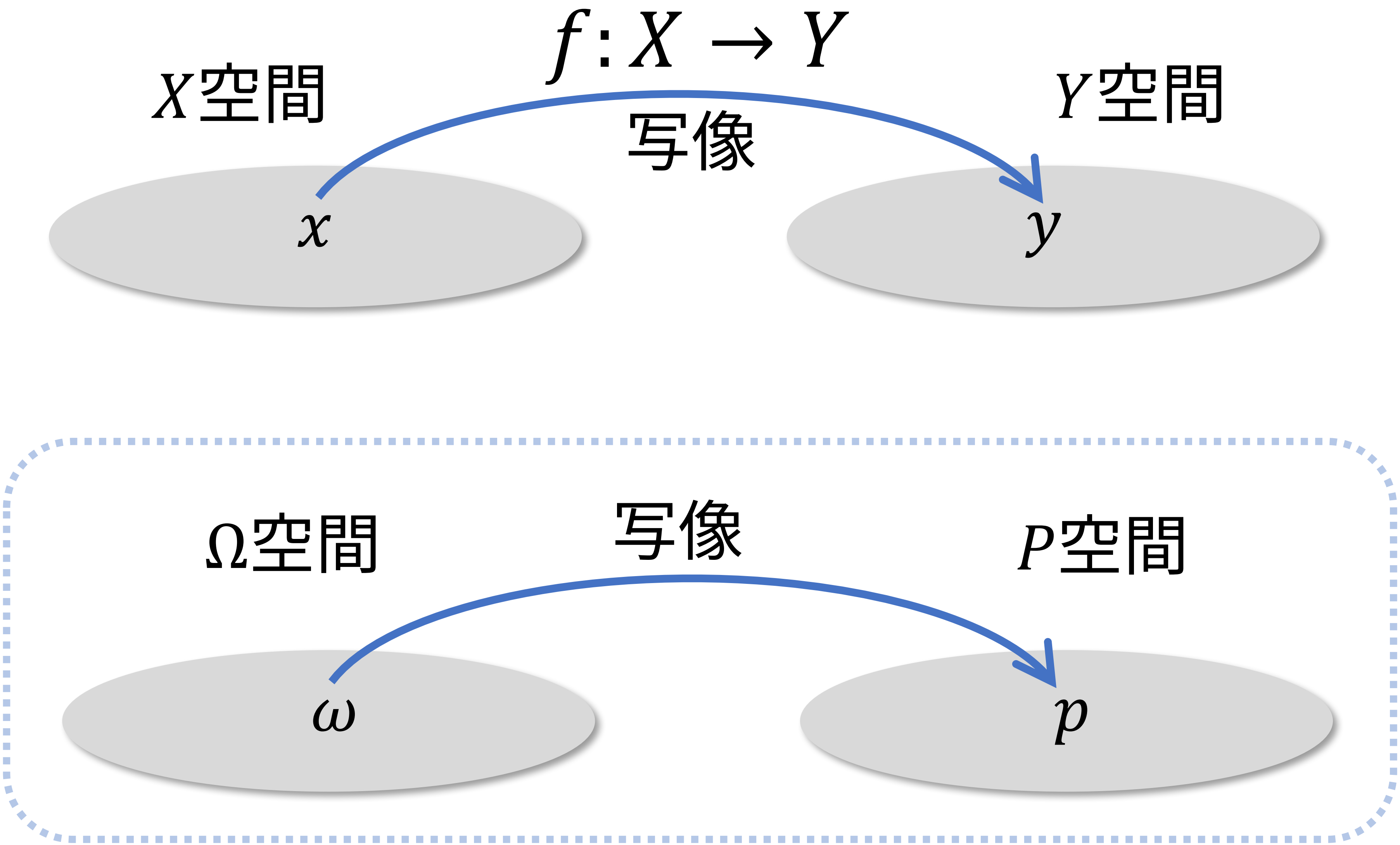
関数



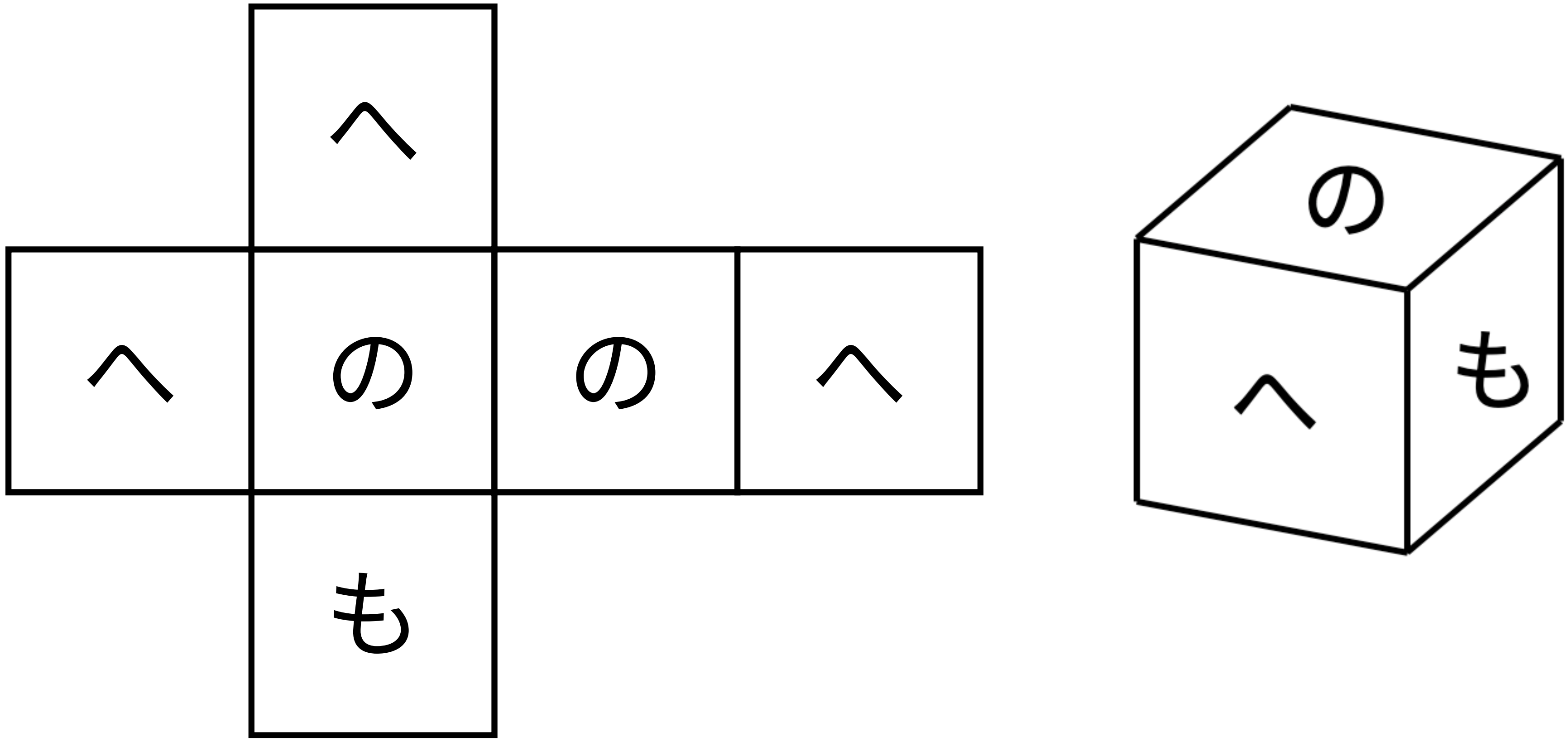
関数は写像。



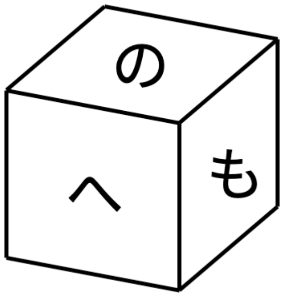
確率を求める関数も写像。



へのへのもへサイコロ



へのへのもへサイコロ



標本空間Ω

面$_1$: へ

面$_2$: の

面$_3$: へ

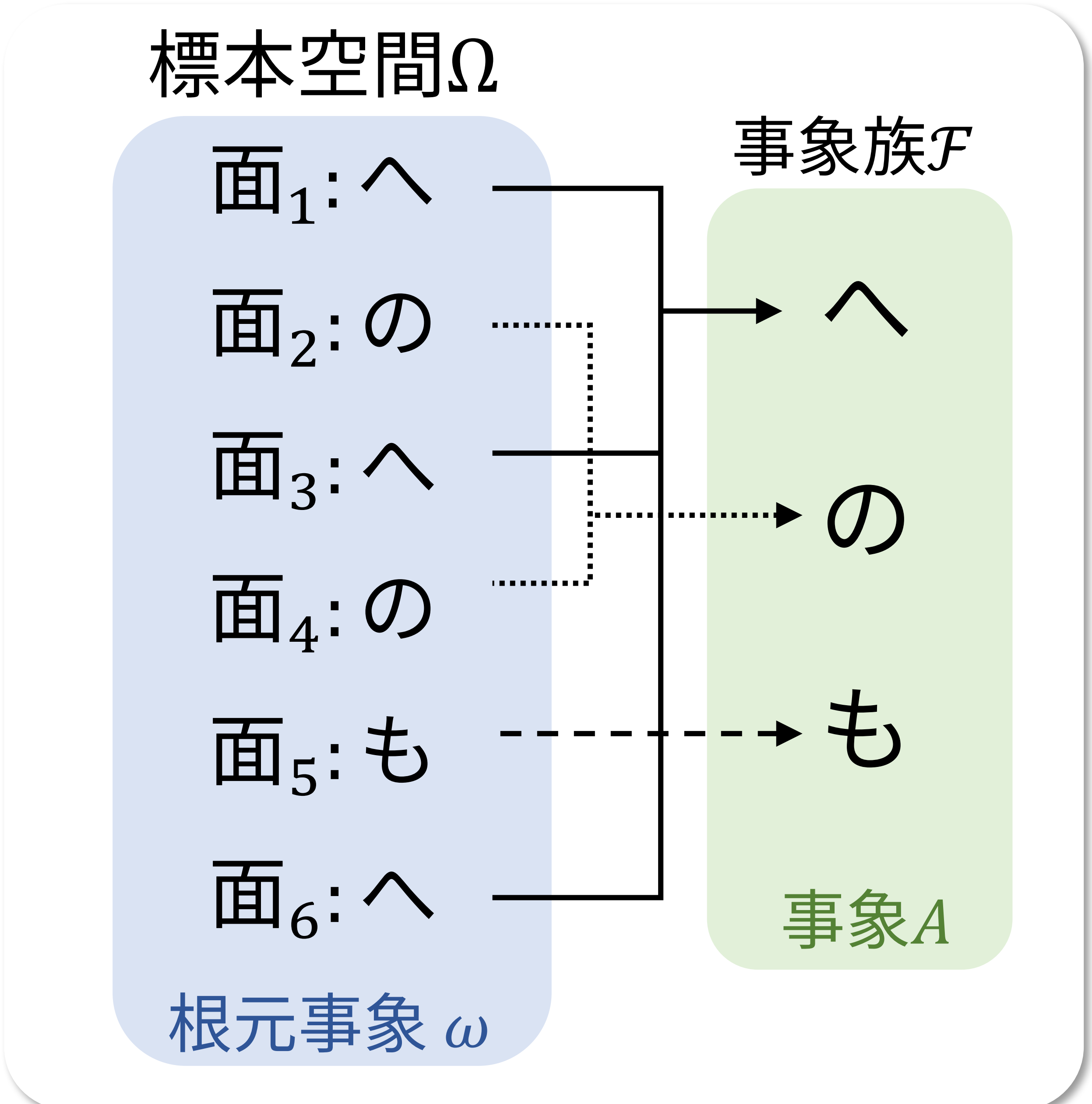
面$_4$: の

面$_5$: も

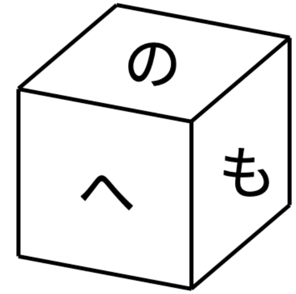
面$_6$: へ

根元事象 ω

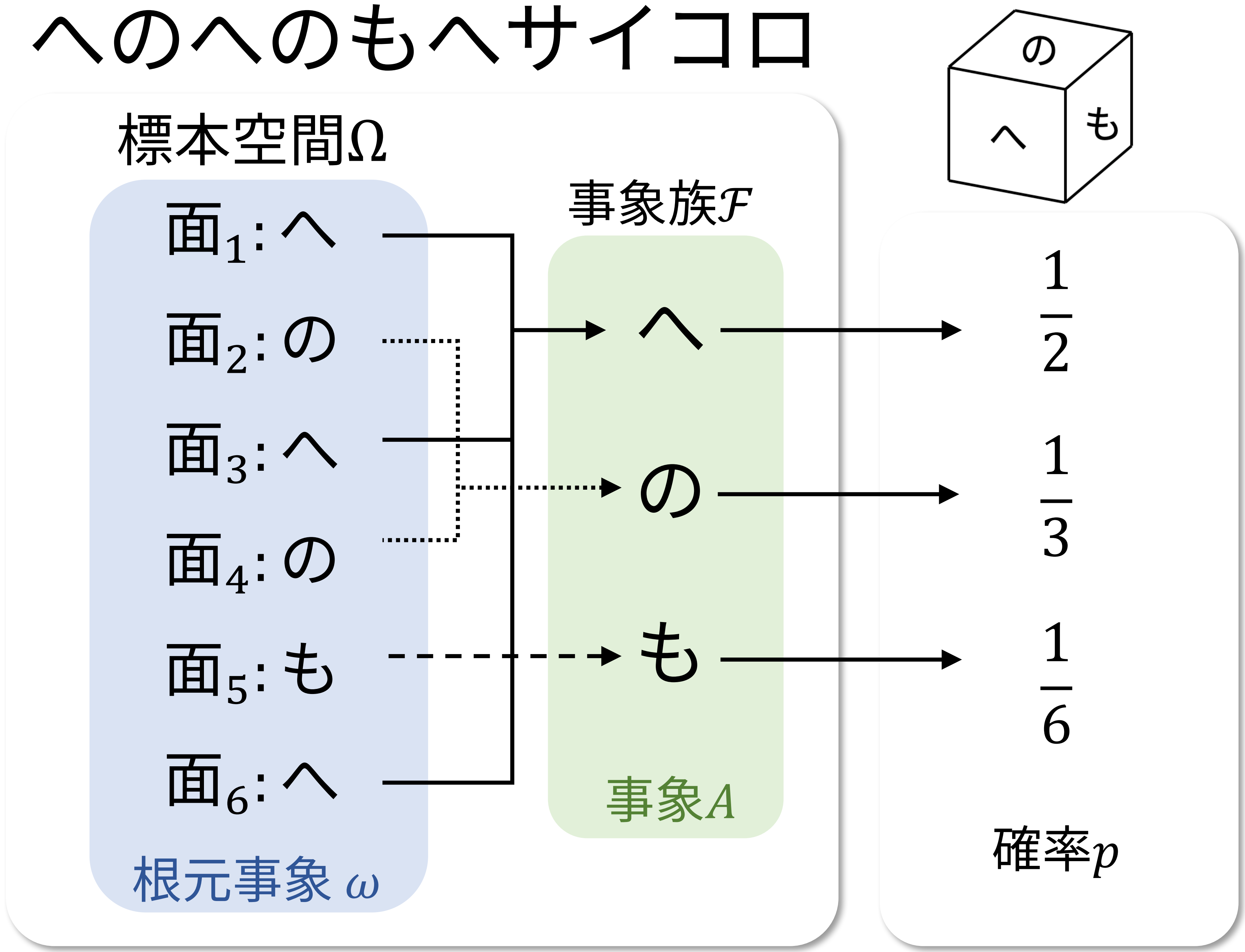
へのへのもへサイコロ



標本空間Ω
事象族$\mathcal{F}$
面$_1$: へ
面$_2$: の
面$_3$: へ
面$_4$: の
面$_5$: も
面$_6$: へ
へ
の
も
事象A
根元事象 ω

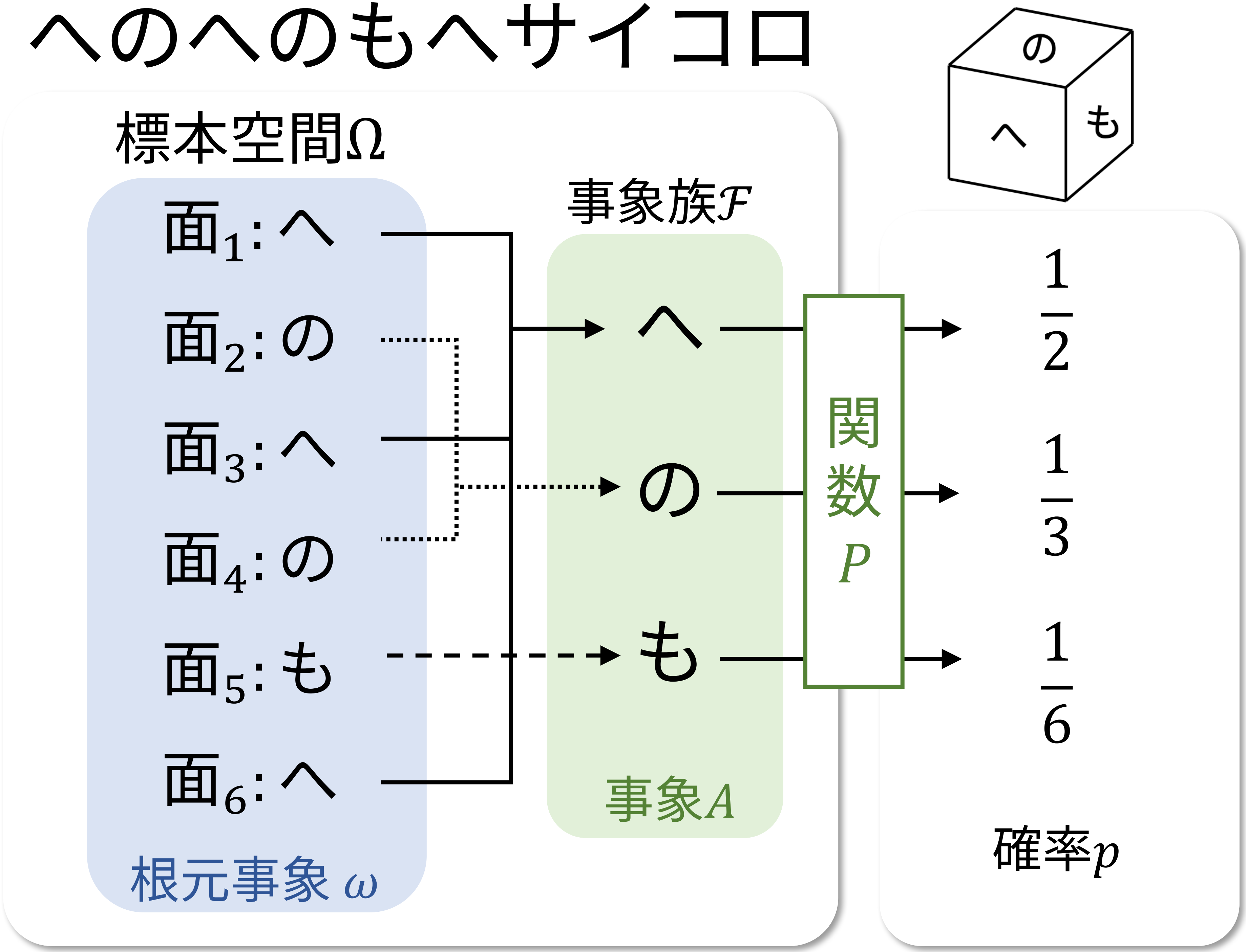


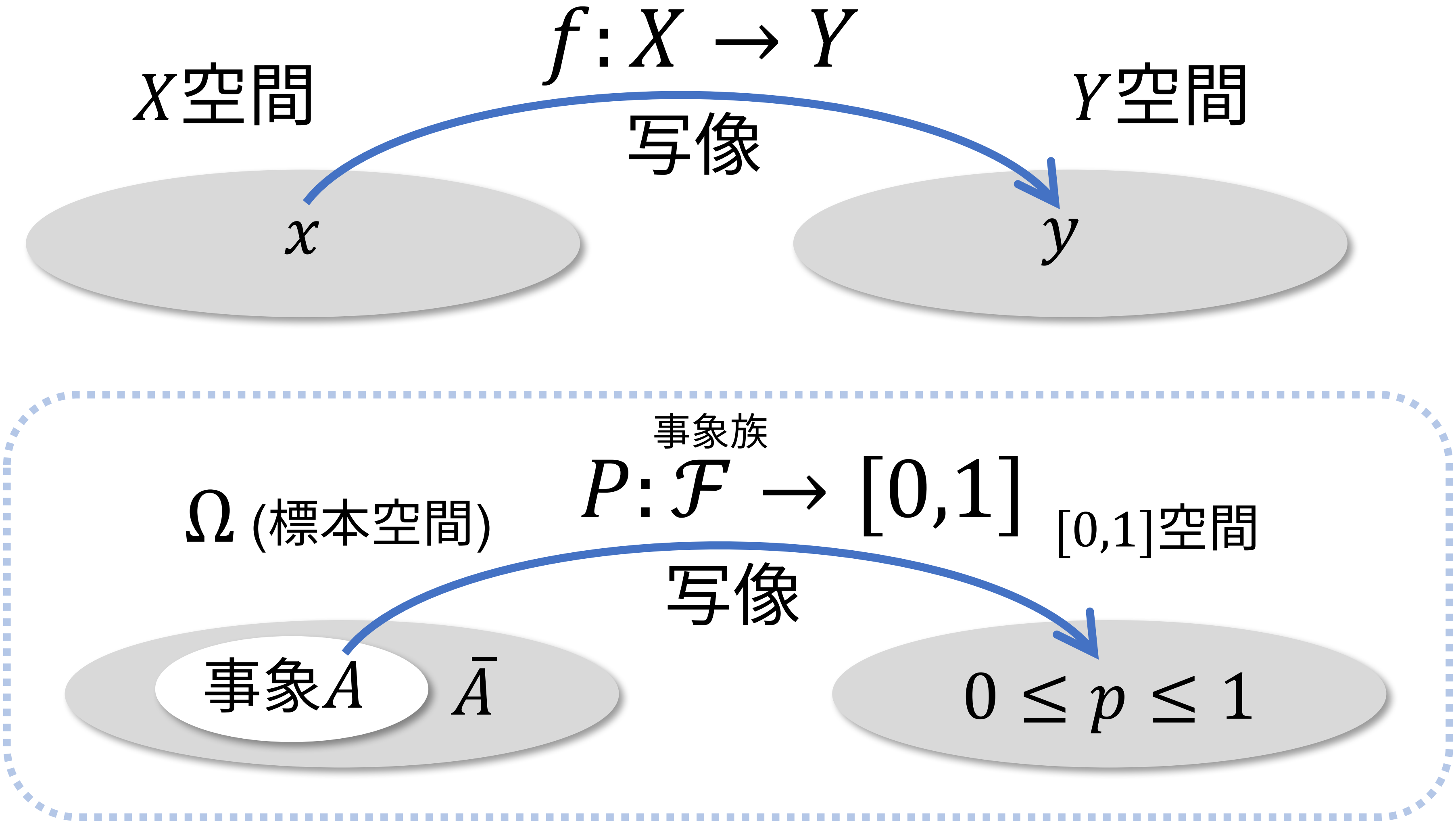
の
へ
も



へのへのもへサイコロ
の
へ
も
標本空間Ω
面1: へ
面2: の
面3: へ
面4: の
面5: も
面6: へ
根元事象 ω
事象族ℱ
へ
の
も
事象A
1/2
1/3
1/6
確率p

へのへのもへサイコロ


の
へ
も
標本空間Ω
面$_1$: へ
面$_2$: の
面$_3$: へ
面$_4$: の
面$_5$: も
面$_6$: へ
根元事象 ω
事象族$\mathcal{F}$
へ
の
も
事象A
関数P
$\frac{1}{2}$
$\frac{1}{3}$
$\frac{1}{6}$
確率p



$f: X \to Y$
X空間
Y空間
写像
x
y
事象族
$P: \mathcal{F} \to [0,1]$
Ω (標本空間)
[0,1]空間
写像
事象A
$\bar{A}$
$0 \le p \le 1$

事象族$\mathcal{F}$の性質

$$\mathcal{F} \subseteq \Omega$$

$$\emptyset, \Omega \in \mathcal{F}$$

$$\mathcal{F} = \{A_1, A_2, \dots\} \neq \emptyset$$

$$A \in \mathcal{F} \;\Rightarrow\; \bar{A} \in \mathcal{F}$$

$$A_i \in \mathcal{F} \Rightarrow \bigcup_i A_i \in \mathcal{F}$$

→ $\mathcal{F}$はσ-加法族である

写像Pの性質

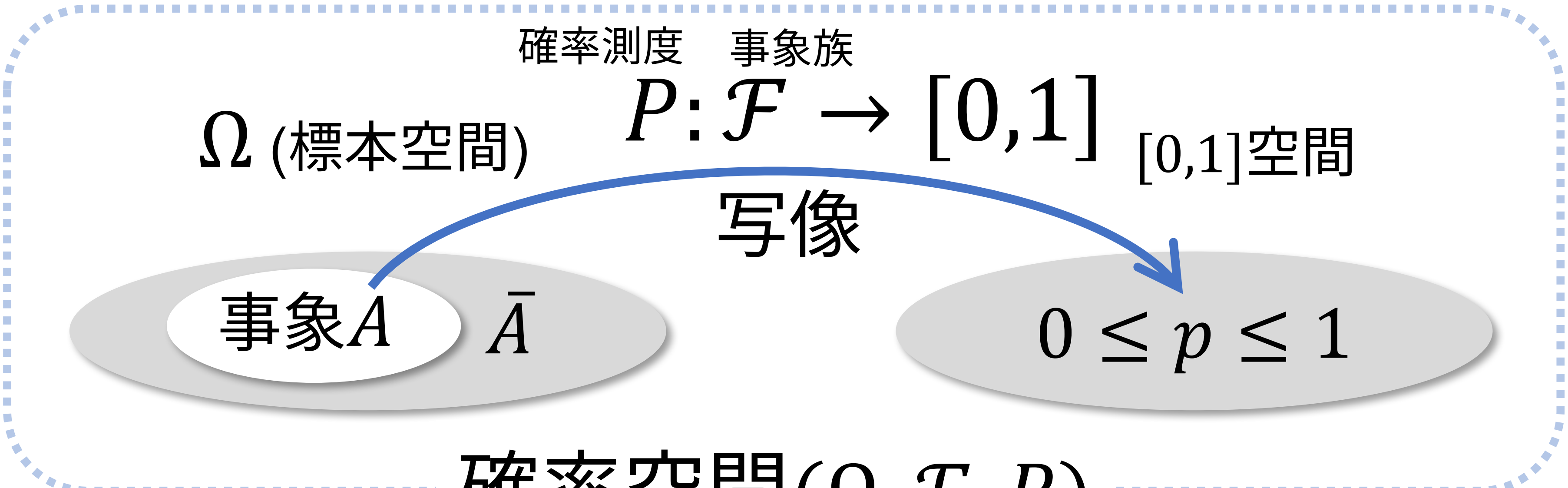
$$P: \mathcal{F} \to [0,1]$$

$$P(\Omega) = 1$$

$$\forall \left(A_i \cap A_j \right) = \emptyset \quad (i \neq j)$$

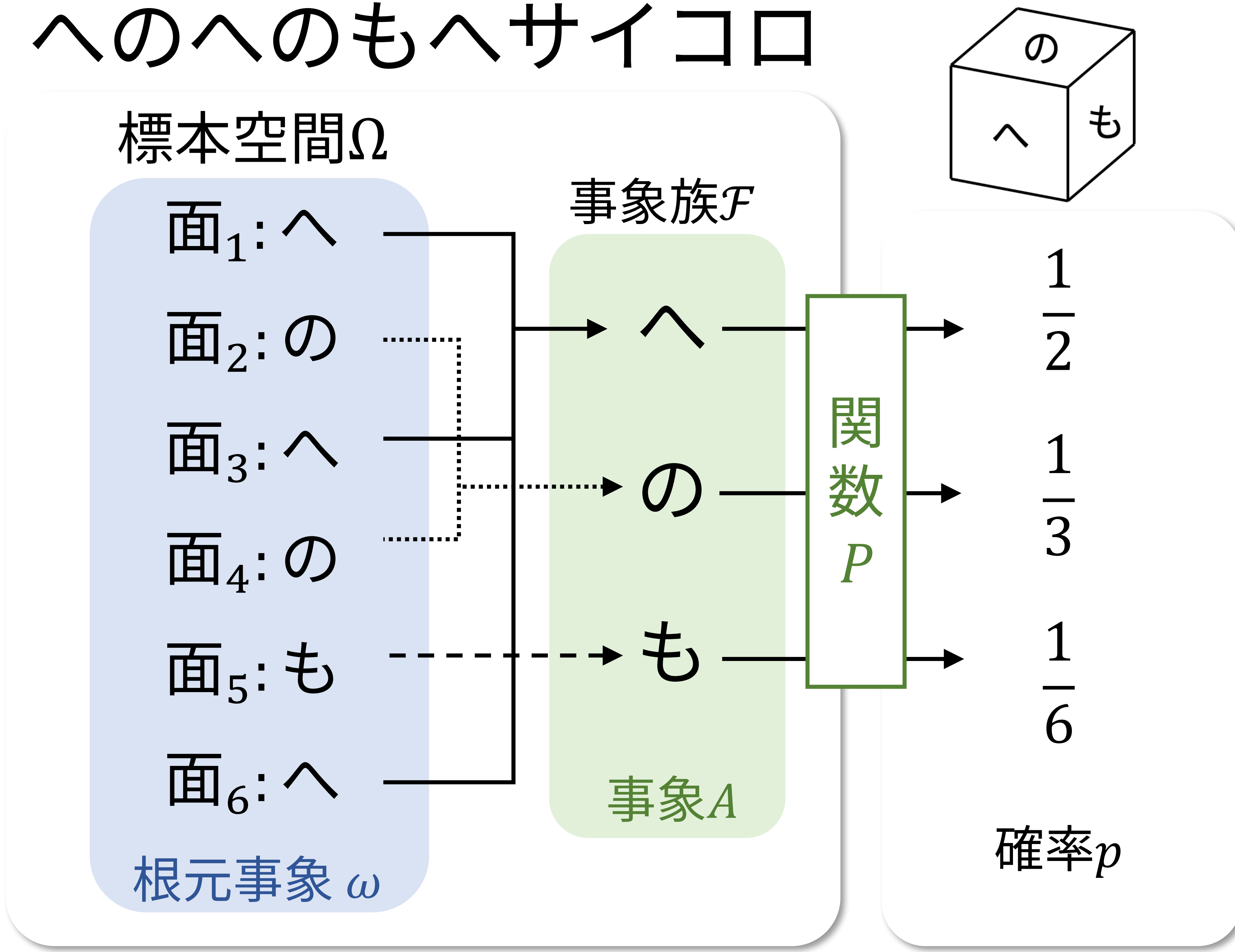
$$\Rightarrow \quad P\left(\bigcup_i A_i\right) = \sum_i P(A_i)$$

$\rightarrow P$が確率測度である。



確率空間$(\Omega, \mathcal{F}, P)$
probability space

へのへのもへサイコロ

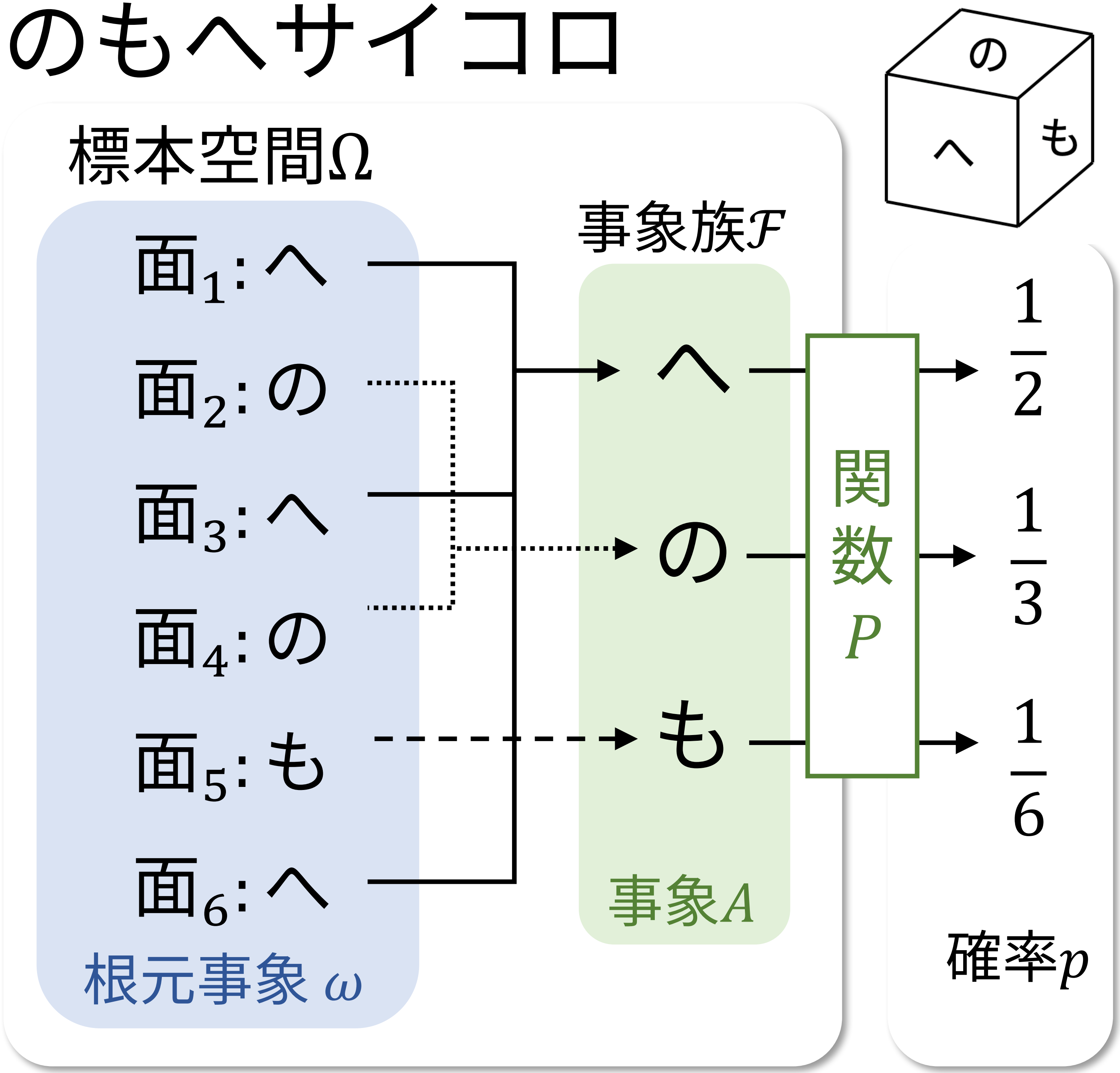


の
へ
も
標本空間Ω
面$_1$: へ
面$_2$: の
面$_3$: へ
面$_4$: の
面$_5$: も
面$_6$: へ
根元事象 ω
事象族$\mathcal{F}$
へ
の
も
事象A
関数P
$\frac{1}{2}$
$\frac{1}{3}$
$\frac{1}{6}$
確率p

「で、君、それは何点なんだね？」

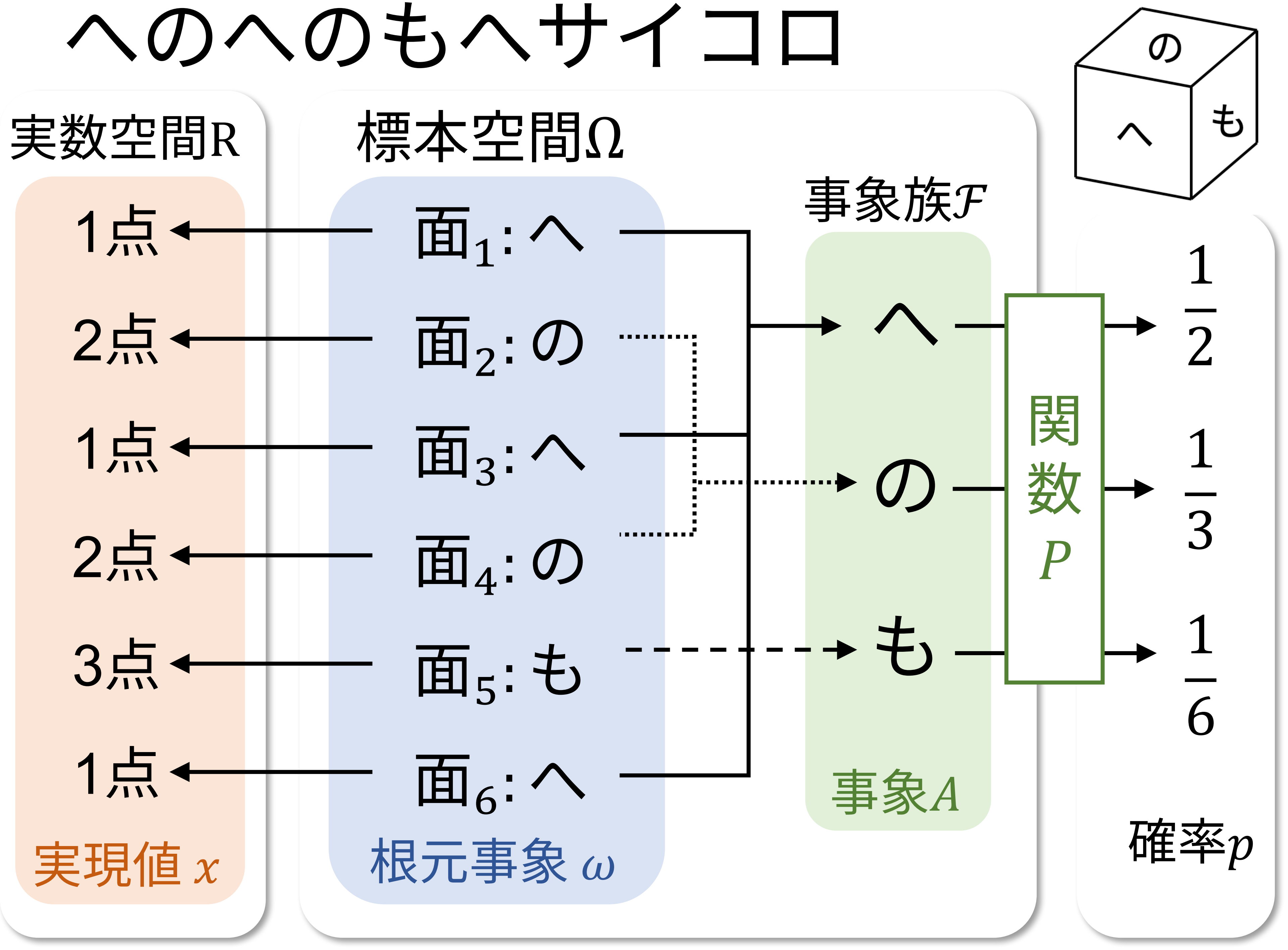
Mr. IF

へのへのもへサイコロ

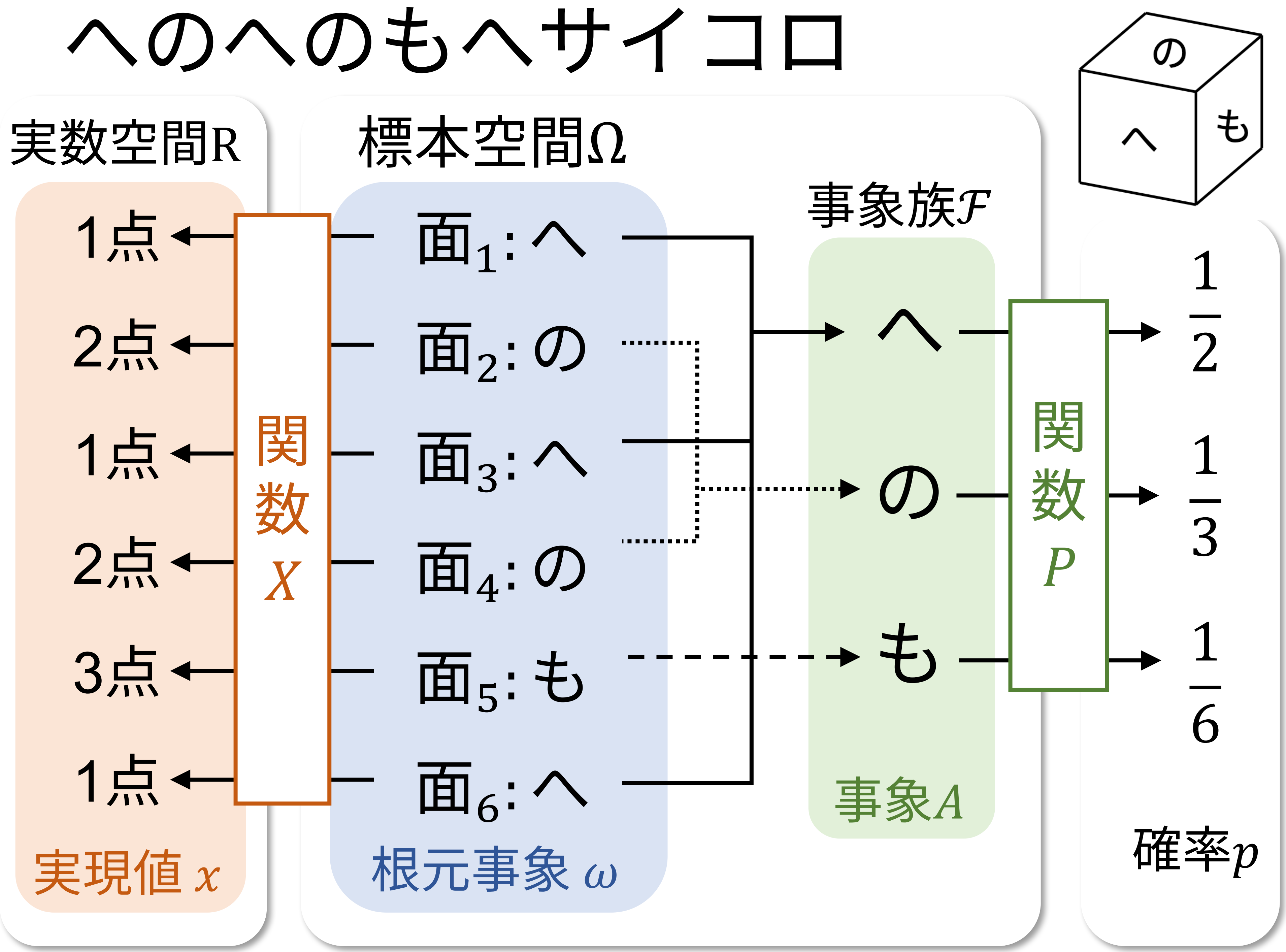


の
へ
も
標本空間Ω
面$_1$: へ
面$_2$: の
面$_3$: へ
面$_4$: の
面$_5$: も
面$_6$: へ
根元事象 ω
事象族$\mathcal{F}$
へ
の
も
事象A
関数P
$\frac{1}{2}$
$\frac{1}{3}$
$\frac{1}{6}$
確率p

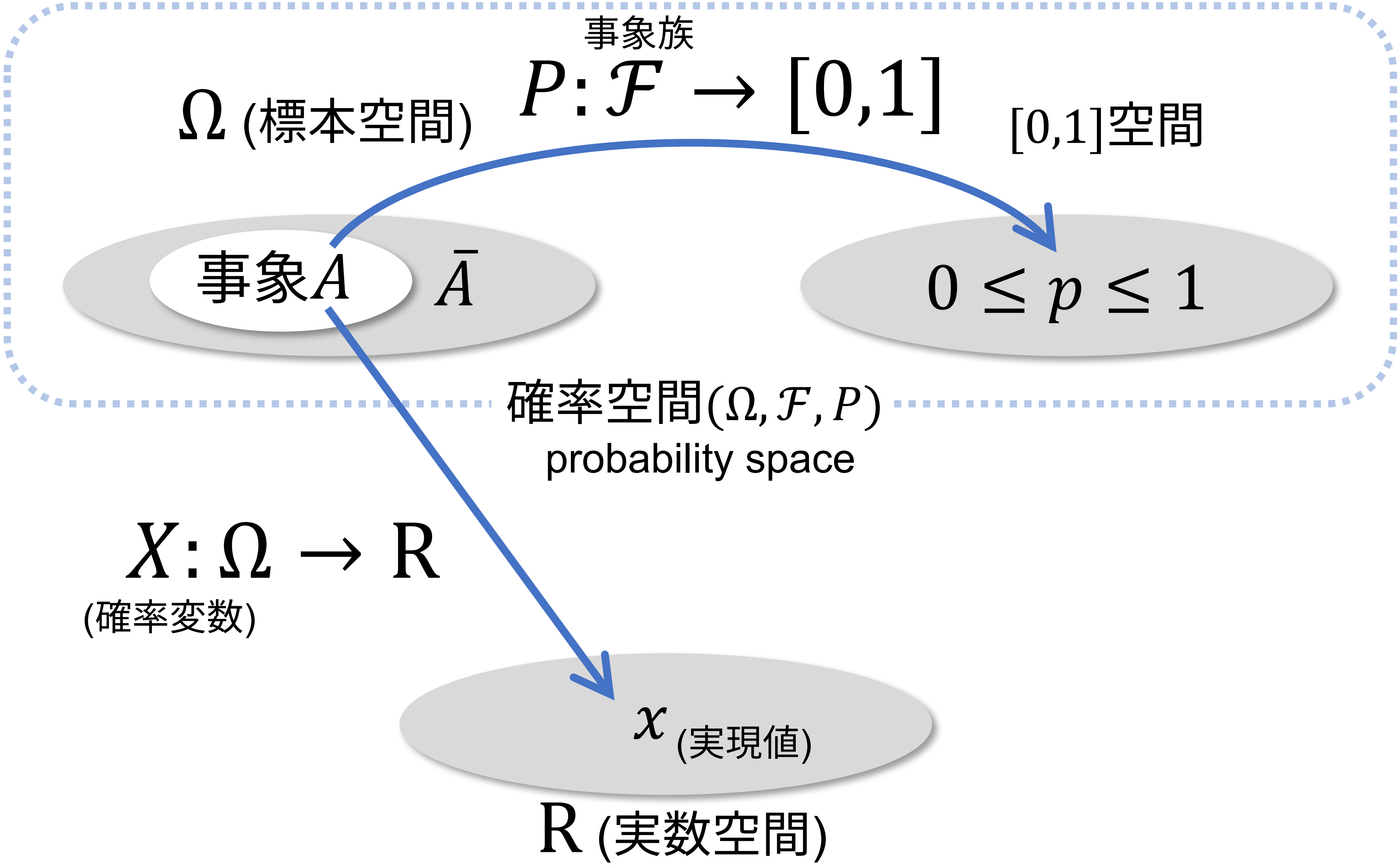
へのへのもへサイコロ

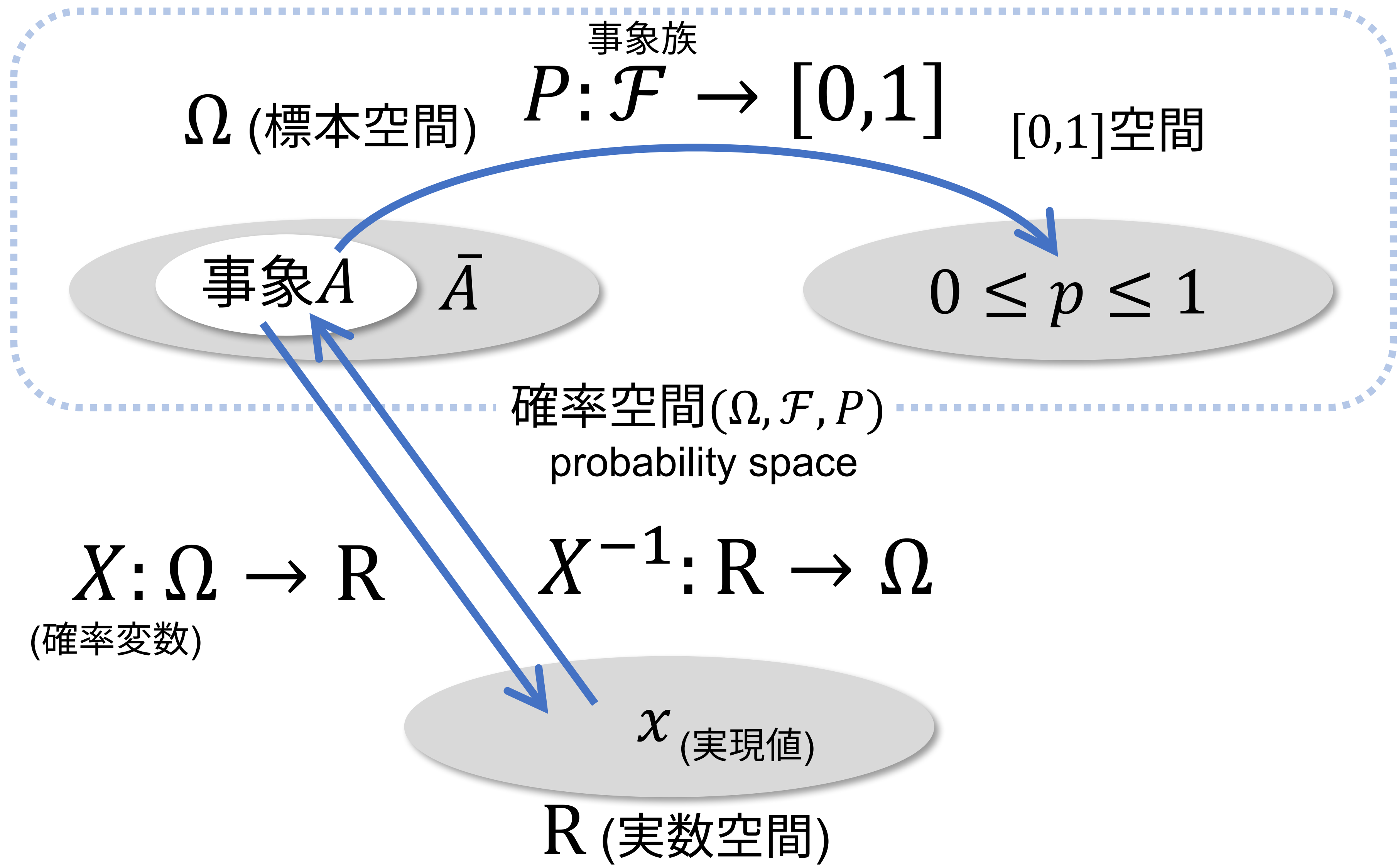


の
へ
も
実数空間R
標本空間Ω
事象族$\mathcal{F}$
1点
2点
1点
2点
3点
1点
面$_1$:へ
面$_2$:の
面$_3$:へ
面$_4$:の
面$_5$:も
面$_6$:へ
へ
の
も
関数P
$\frac{1}{2}$
$\frac{1}{3}$
$\frac{1}{6}$
実現値 x
根元事象 ω
事象A
確率p



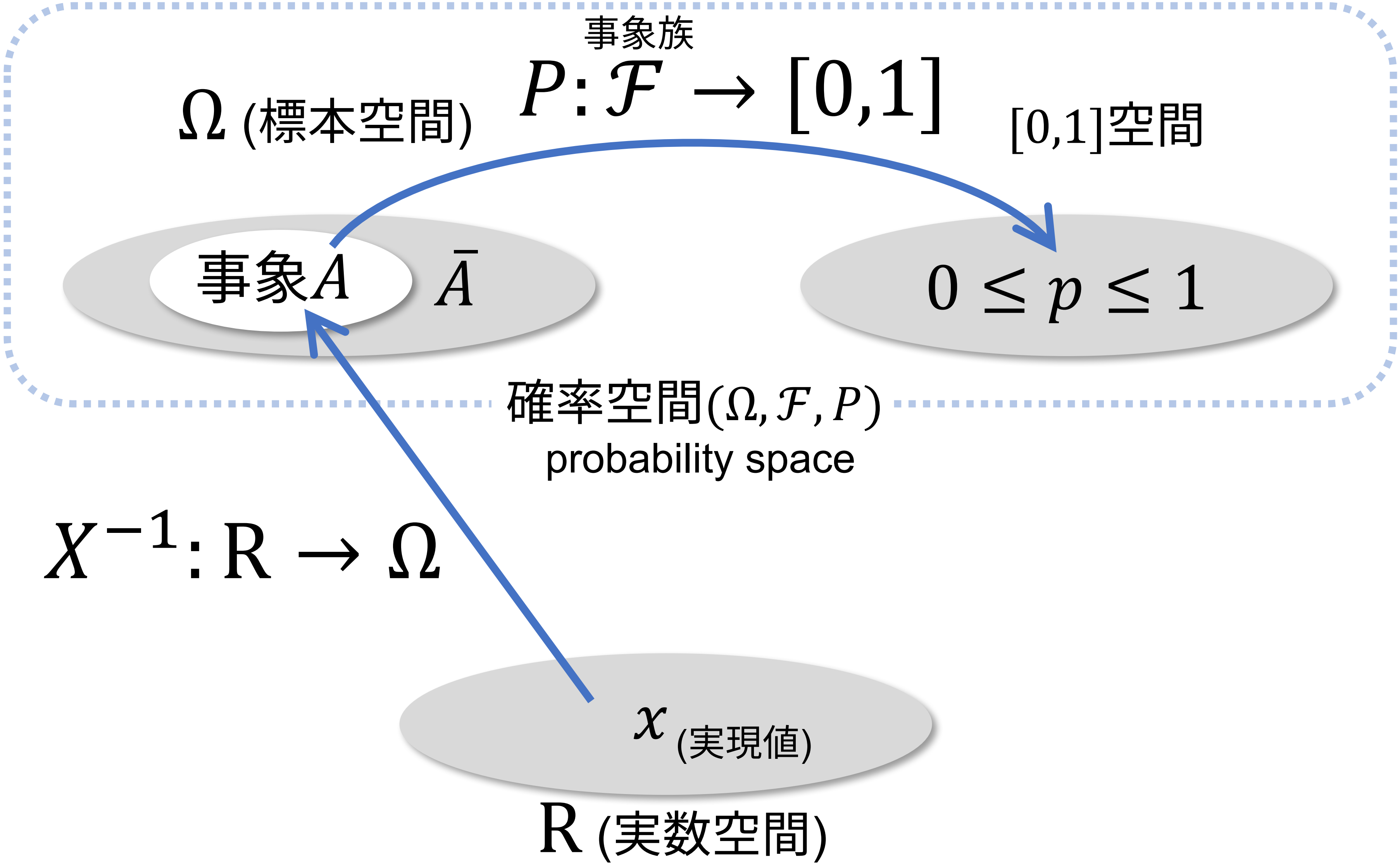
へのへのもへサイコロ
の
へ
も
実数空間R
標本空間Ω
事象族$\mathcal{F}$
1点
2点
1点
2点
3点
1点
関数X
面$_1$: へ
面$_2$: の
面$_3$: へ
面$_4$: の
面$_5$: も
面$_6$: へ
へ
の
も
関数P
$\frac{1}{2}$
$\frac{1}{3}$
$\frac{1}{6}$
実現値 x
根元事象 ω
事象A
確率p

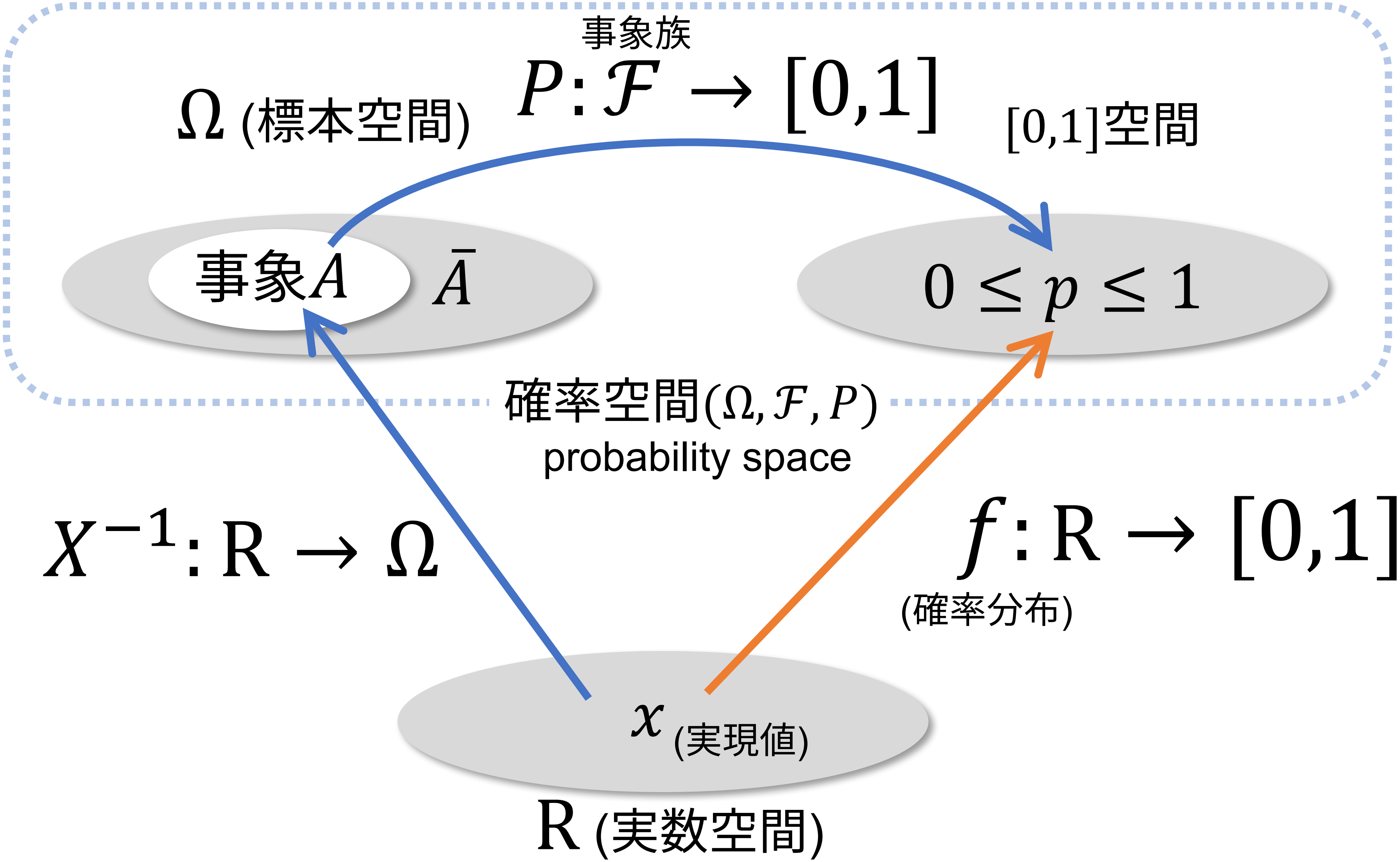

事象族
$P: \mathcal{F} \to [0,1]$
Ω (標本空間)
[0,1]空間
事象A
$\bar{A}$
$0 \leq p \leq 1$
確率空間$(\Omega, \mathcal{F}, P)$
probability space
$X: \Omega \to \mathrm{R}$
(確率変数)
x (実現値)
R (実数空間)

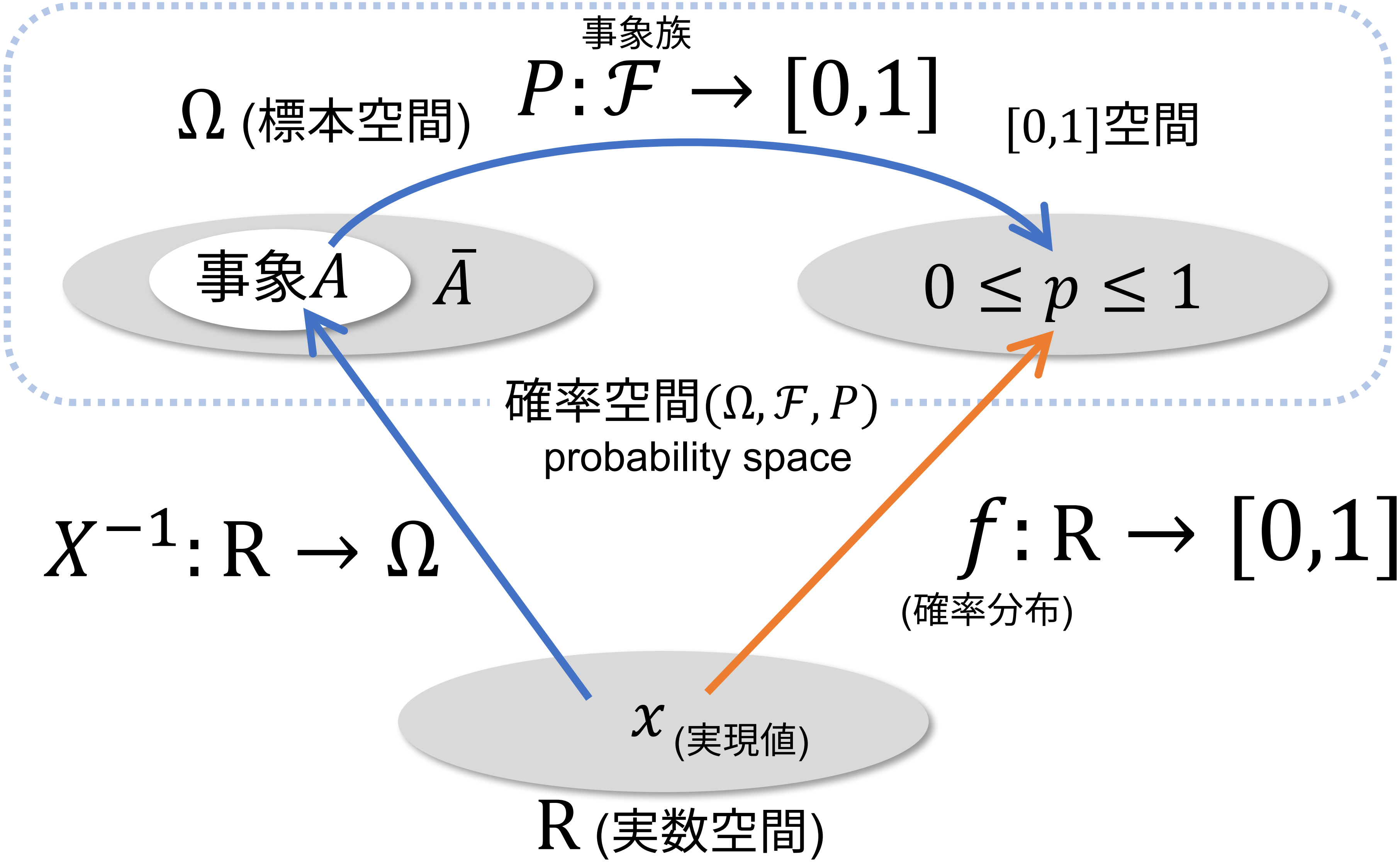


$$X(A) = x \quad \Leftrightarrow \quad X^{-1}(x) = A$$

より厳密には $X^{-1}(x) := \{\omega \in \Omega | X(\omega) \in x\} \in \mathcal{F}$

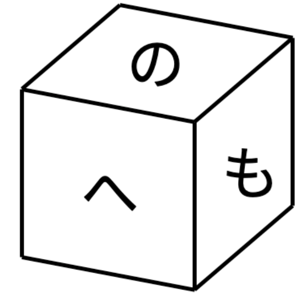

事象族
$P\colon \mathcal{F} \to [0,1]$
Ω (標本空間)
[0,1]空間
事象A
$\bar{A}$
$0 \le p \le 1$
確率空間$(\Omega, \mathcal{F}, P)$
probability space
$X^{-1}\colon \mathrm{R} \to \Omega$
x (実現値)
R (実数空間)


事象族
$P\colon \mathcal{F} \to [0,1]$
Ω (標本空間)
[0,1]空間
事象A
$\bar{A}$
$0 \leq p \leq 1$
確率空間$(\Omega, \mathcal{F}, P)$
probability space
$X^{-1}\colon \mathrm{R} \to \Omega$
$f\colon \mathrm{R} \to [0,1]$
(確率分布)
x (実現値)
R (実数空間)



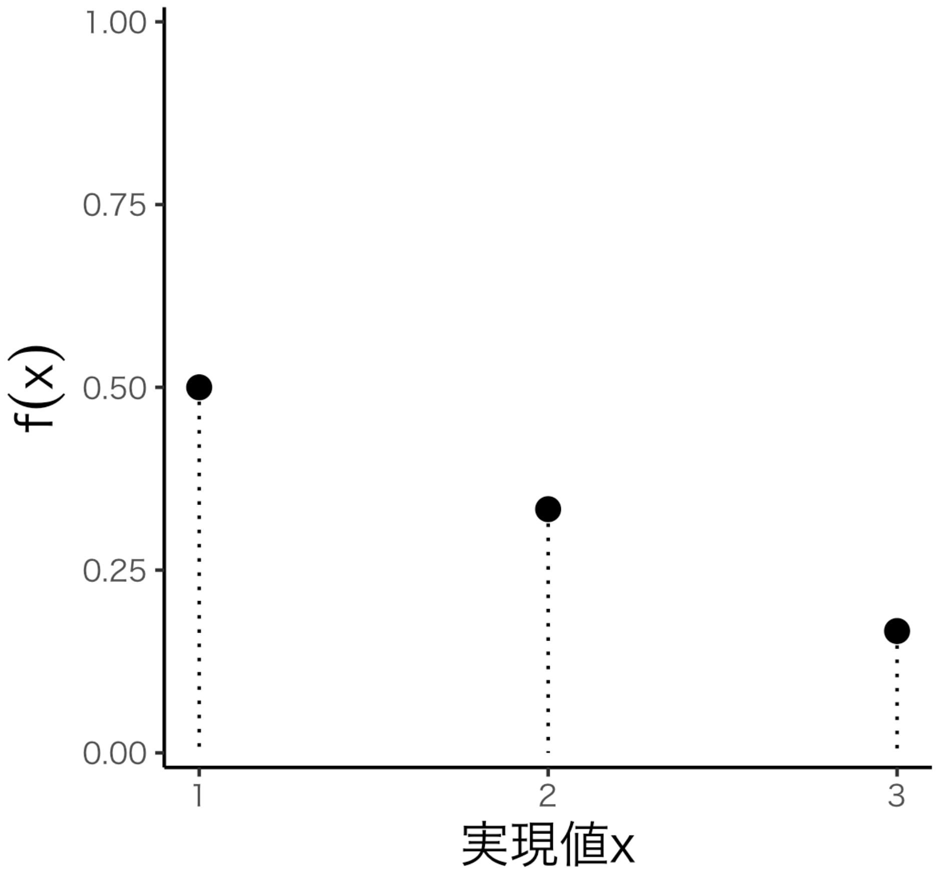
$$\overset{\text{確率分布}}{f(x)} = \overset{\text{確率測度}}{P}\left(\underset{\text{実現値}}{X^{-1}(x)}\right) = P(\underset{\text{事象}}{A} \ni \underset{\text{根元事象}}{\omega} | \overset{\text{確率変数}}{X}(\omega) = x) = \overset{\text{確率}}{p}$$

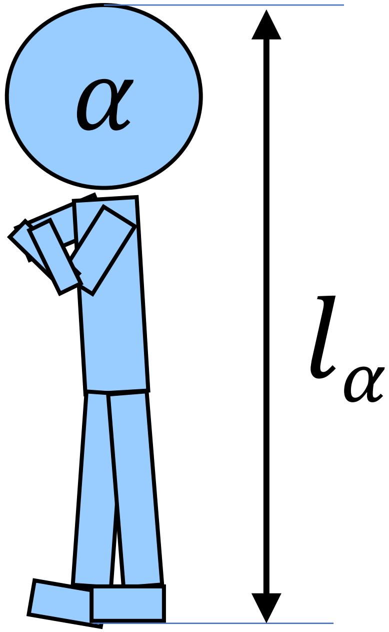
へのへのもへサイコロ



確率分布

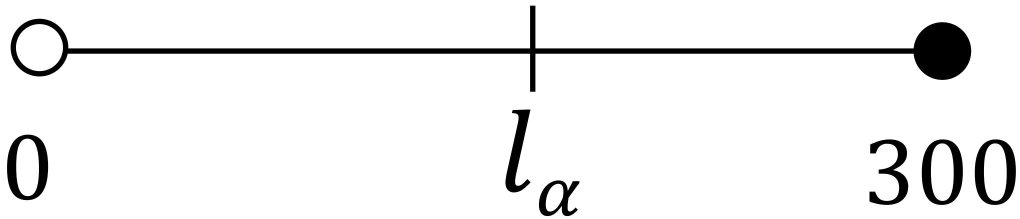
$$f(x) = P(X = x) = p$$

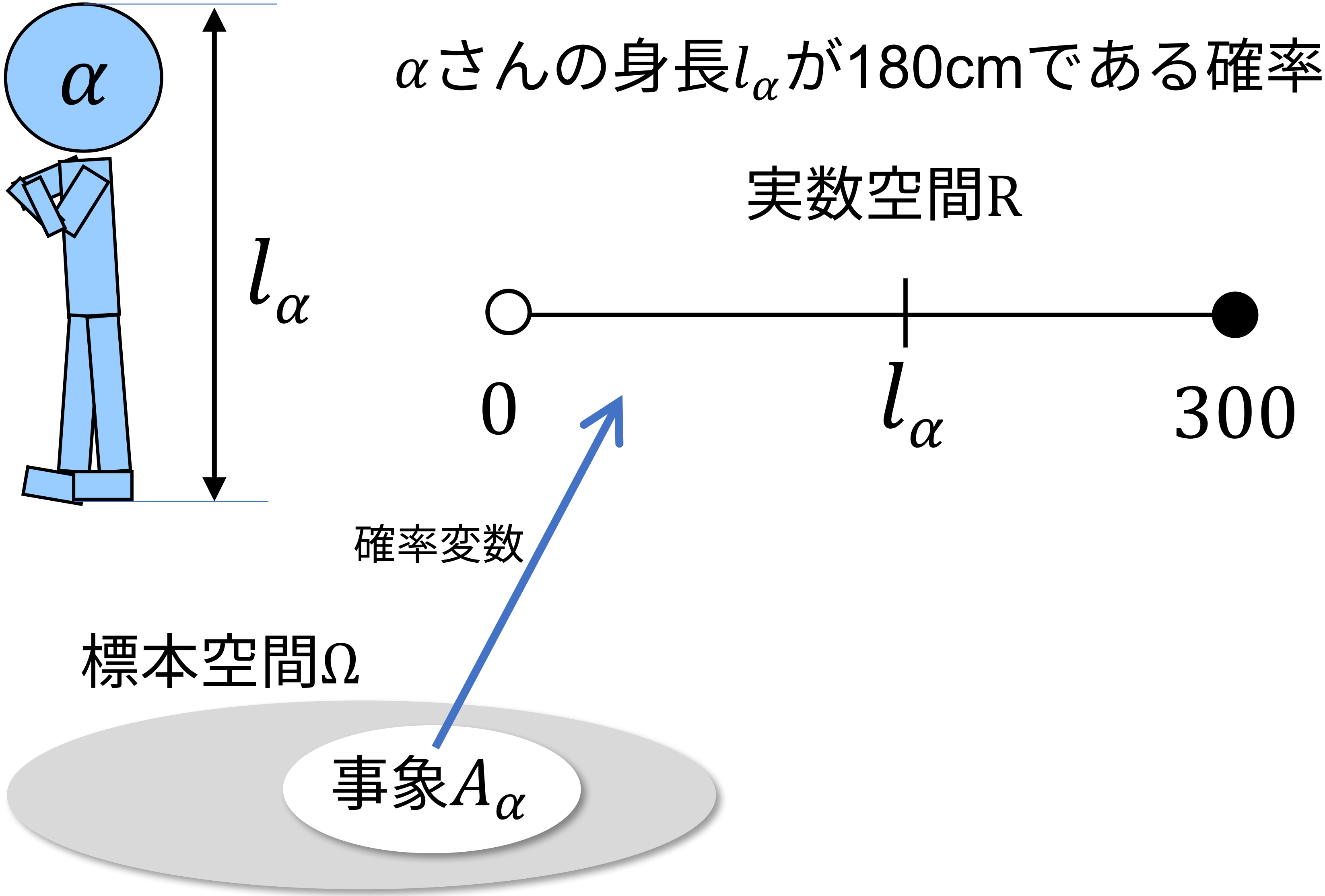


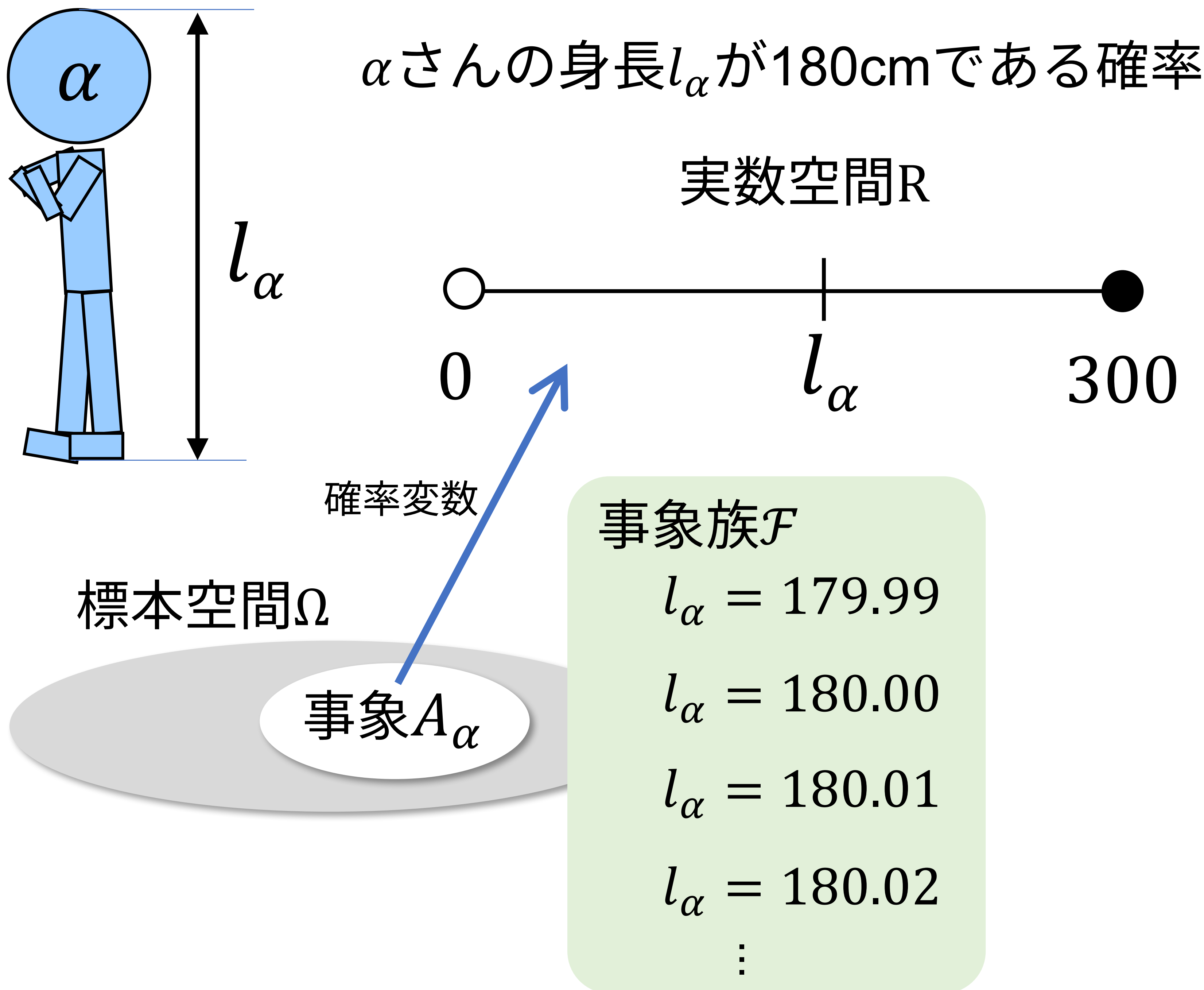


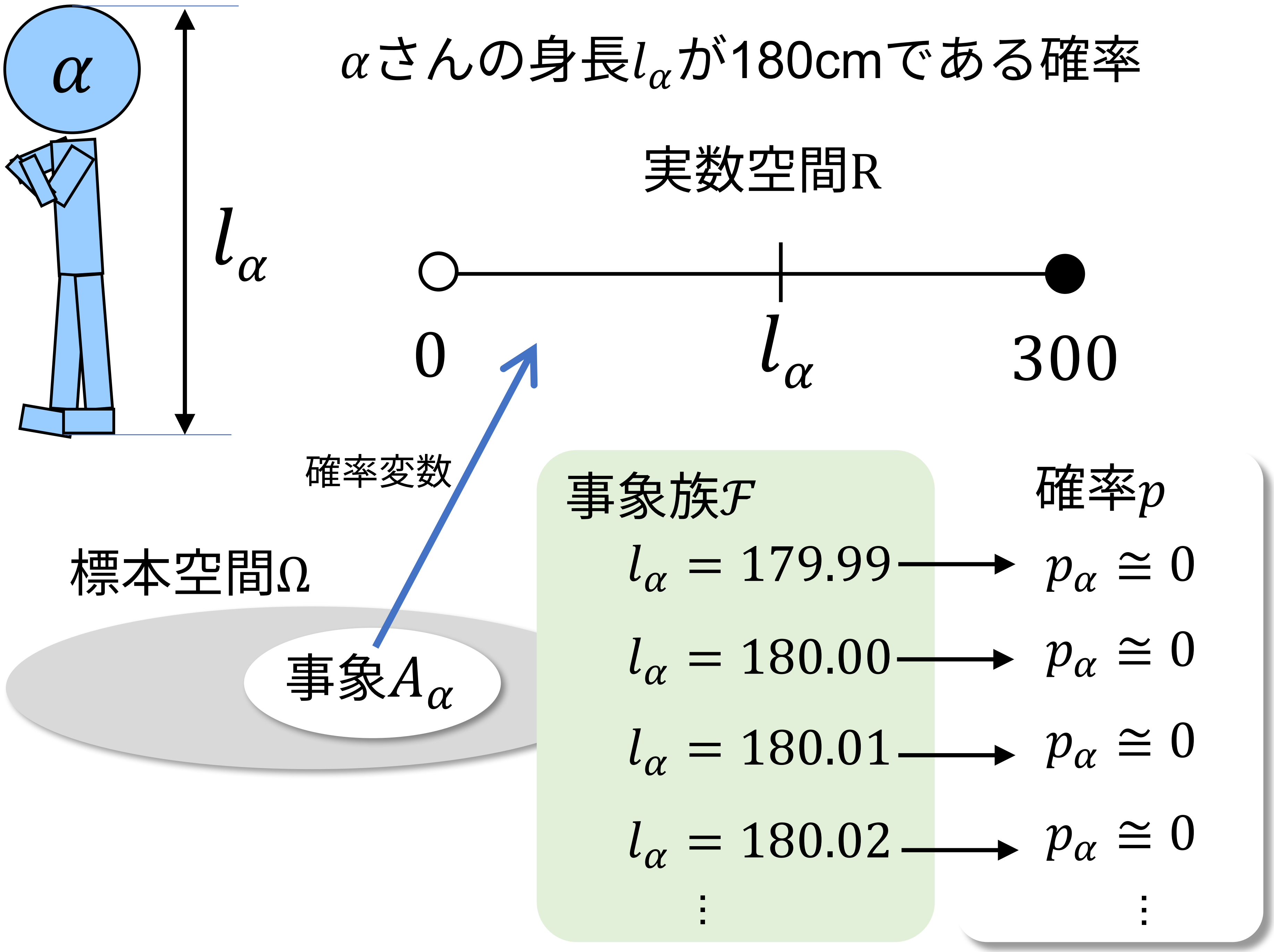
αさんの身長l_αが180cmである確率

実数空間R



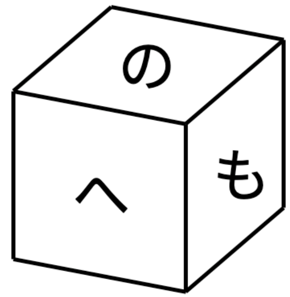

α
l_α
αさんの身長l_αが180cmである確率
実数空間R
0
l_α
300
確率変数
標本空間Ω
事象A_α


αさんの身長lαが180cmである確率
α
lα
実数空間R
0
lα
300
確率変数
標本空間Ω
事象Aα
事象族F
lα = 179.99
lα = 180.00
lα = 180.01
lα = 180.02
⋮

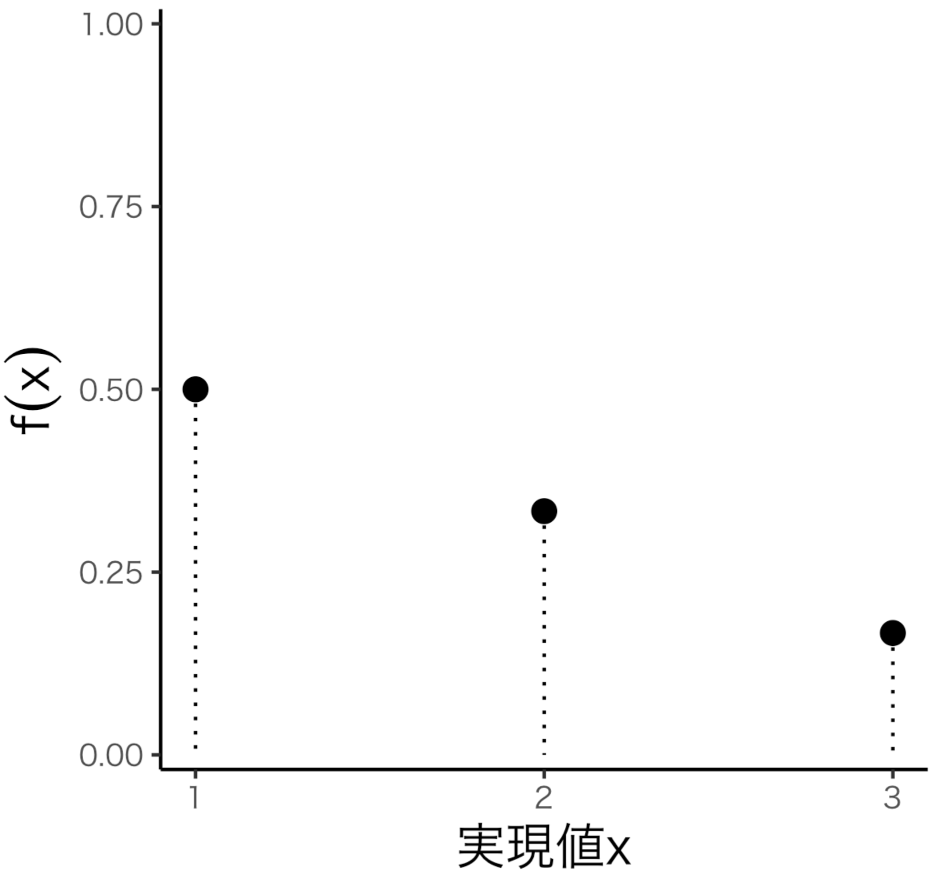


αさんの身長l_αが180cmである確率
α
l_α
実数空間R
0
l_α
300
確率変数
標本空間Ω
事象A_α
事象族$\mathcal{F}$
$l_\alpha = 179.99$
$l_\alpha = 180.00$
$l_\alpha = 180.01$
$l_\alpha = 180.02$
⋮
確率p
$p_\alpha \cong 0$
$p_\alpha \cong 0$
$p_\alpha \cong 0$
$p_\alpha \cong 0$
⋮

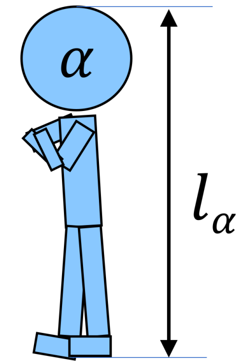
確率分布



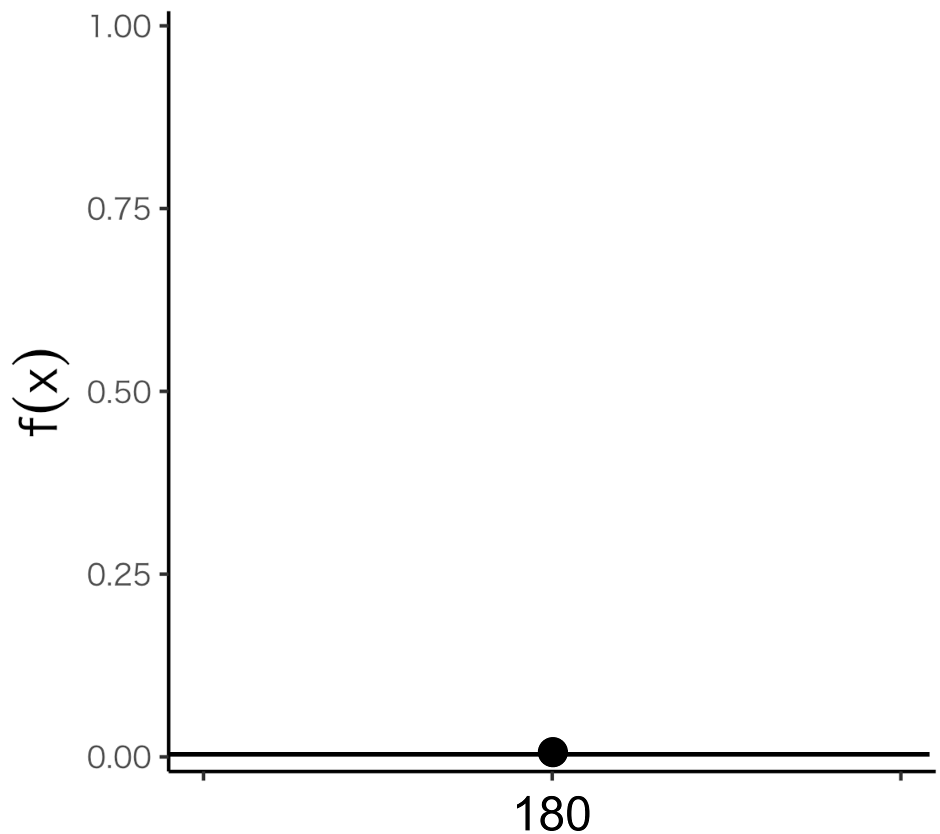
$$f(x) = P(A|X = x) = p$$

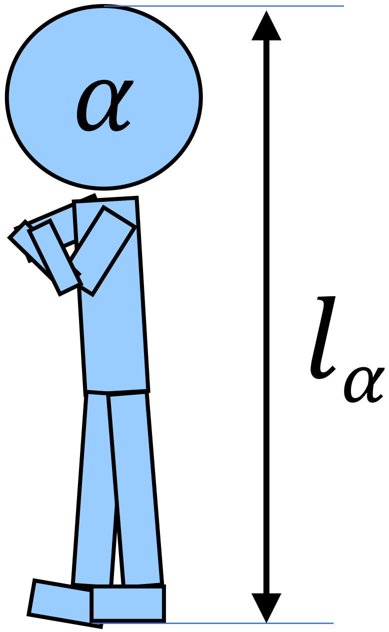


確率分布



$$P(A|X = 180) \cong 0\ !?$$





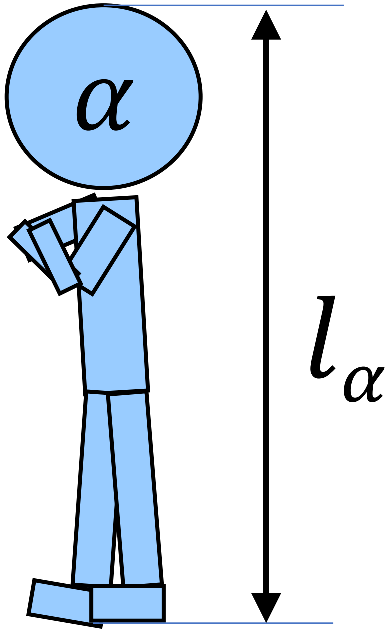
αさんの身長l_αが180cmである確率

だいたい180cmである確率

事象族$\mathcal{F}$の性質より、

$$A_i \in \mathcal{F} \Rightarrow \bigcup_i A_i \in \mathcal{F}$$

（足したものも$\mathcal{F}$ に含まれるよ）



αさんの身長l_αが180cmである確率

だいたい180cmである確率

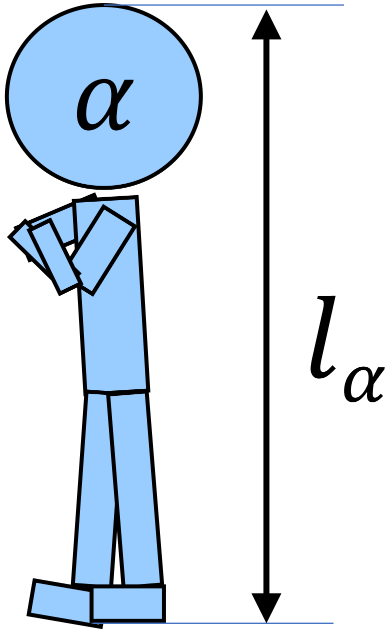
事象族$\mathcal{F}$の性質より、

$$A_i \in \mathcal{F} \Rightarrow \bigcup_i A_i \in \mathcal{F}$$

確率測度Pの性質より、

$$\forall \left(A_i \cap A_j\right)_{i \neq j} = \emptyset \Rightarrow P\left(\bigcup_i A_i\right) = \sum_i P(A_i)$$

（事象が互いに素なら、和事象の確率は、事象の確率の和だよ）



αさんの身長l_αが180cmである確率

だいたい180cmである確率

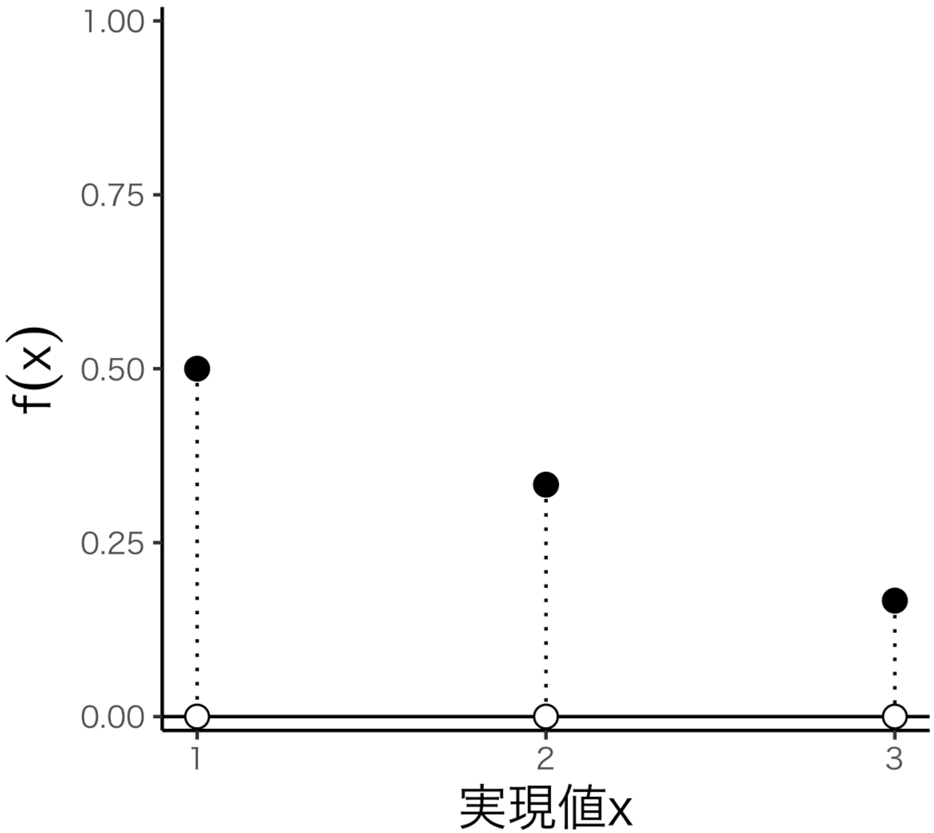
$$P(A|X \cong 180) = P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$$

$$= \int_{180-\Delta}^{180+\Delta} P(A|X = x)dx = \int_{180-\Delta}^{180+\Delta} f(x)dx$$

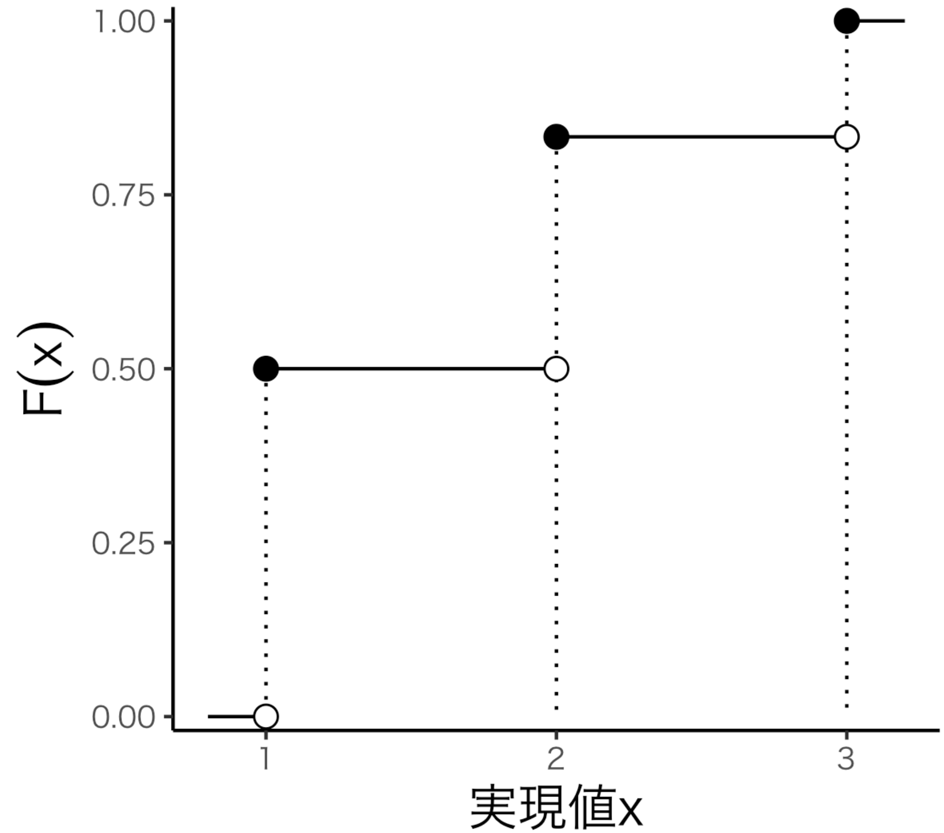
$$f(x) = P(A|X = x) = p$$

$$F(x) = \int_{-\infty}^{x} f(x)dx$$

確率分布



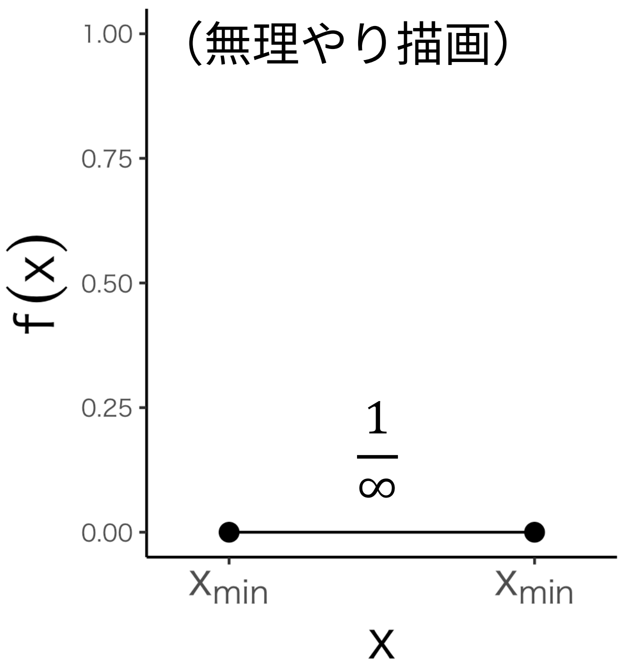
累積分布



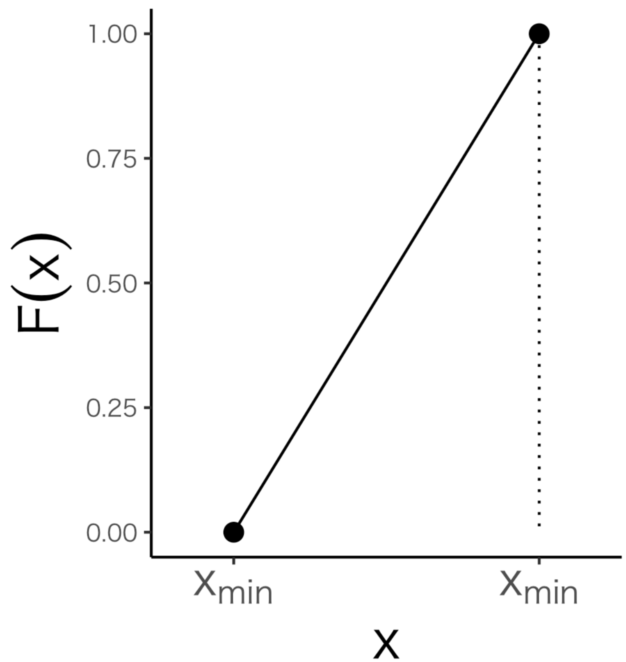
確率分布$f(x)$: 定義域$[x_{min}, x_{max}]$において均一

$$p = f(x) \qquad F(x) = \int_{-\infty}^{x} f(x)dx$$

確率分布



累積分布

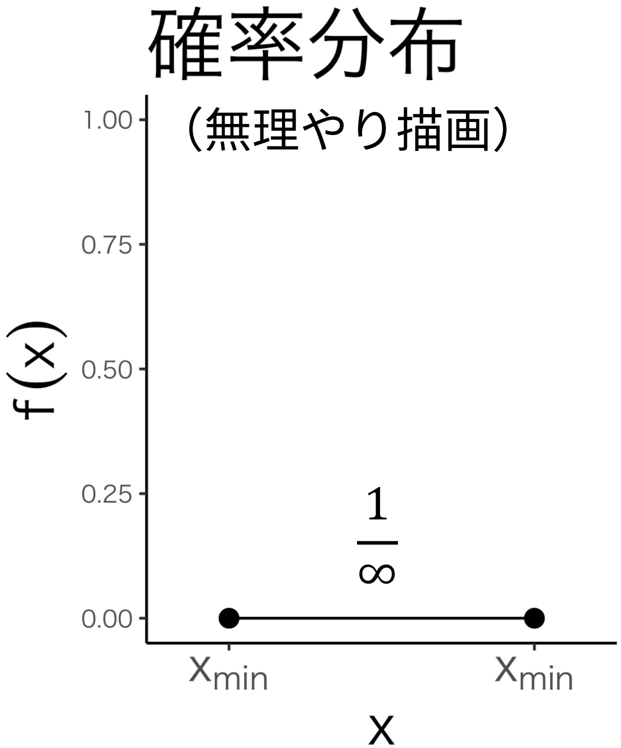


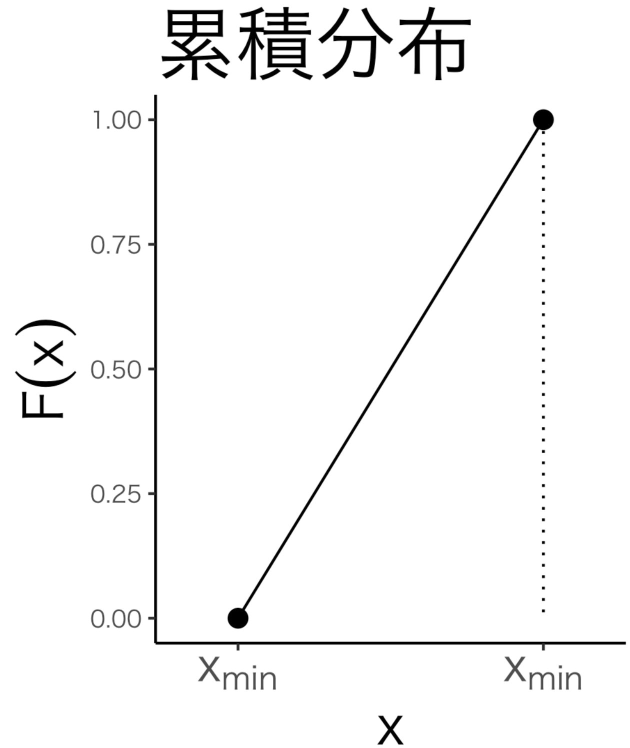
確率分布$f(x)$: 定義域$[x_{min}, x_{max}]$において均一

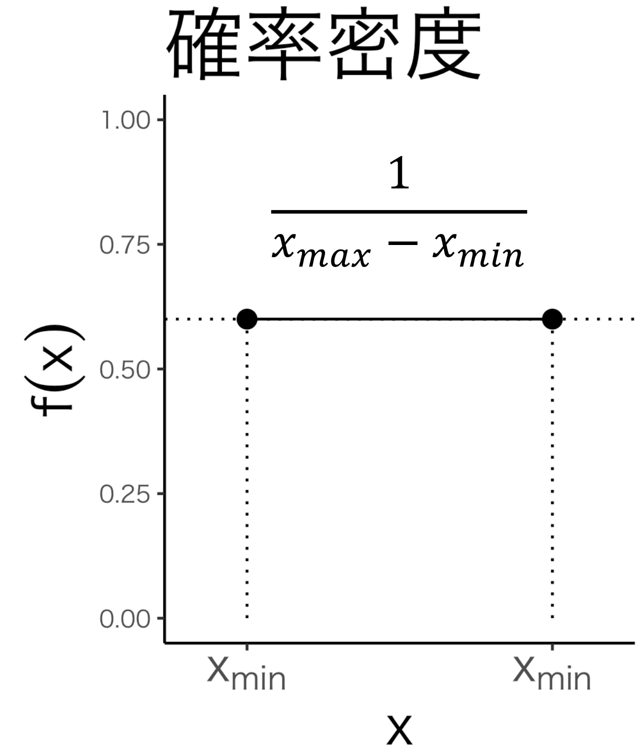
$$p = f(x)$$

$$F(x) = \int_{-\infty}^{x} f(x)dx$$

$$\mathrm{f}(x) = \frac{1}{dx}F(x)$$







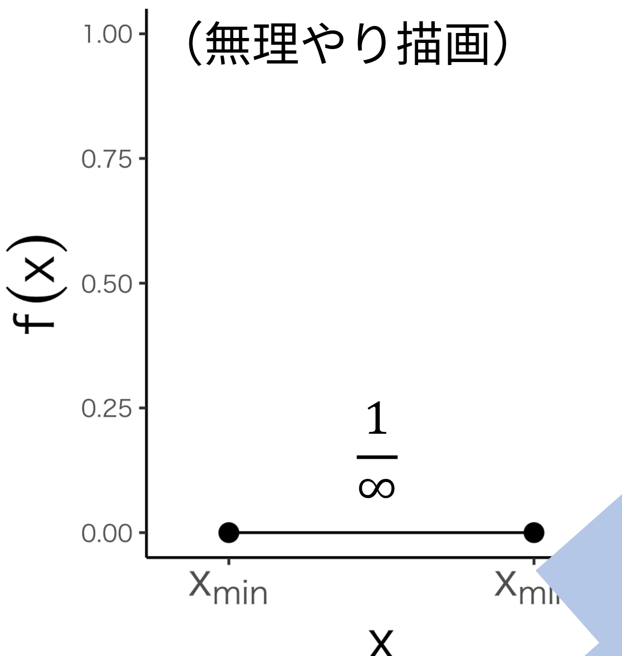
確率分布$f(x)$: 定義域$[x_{min}, x_{max}]$において均一

$p = f(x)$

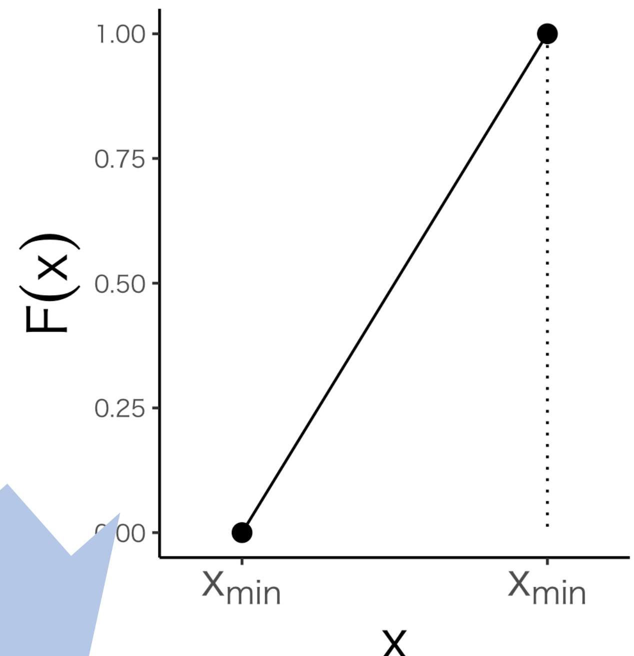
$F(x) = \int_{-\infty}^{x} f(x)dx$

$\mathrm{f}(x) = \frac{1}{dx}F(x)$

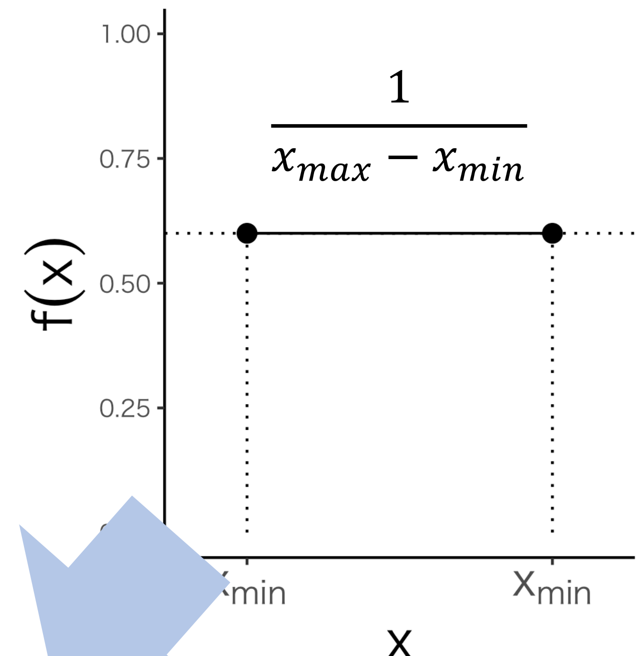
確率分布



累積分布



確率密度



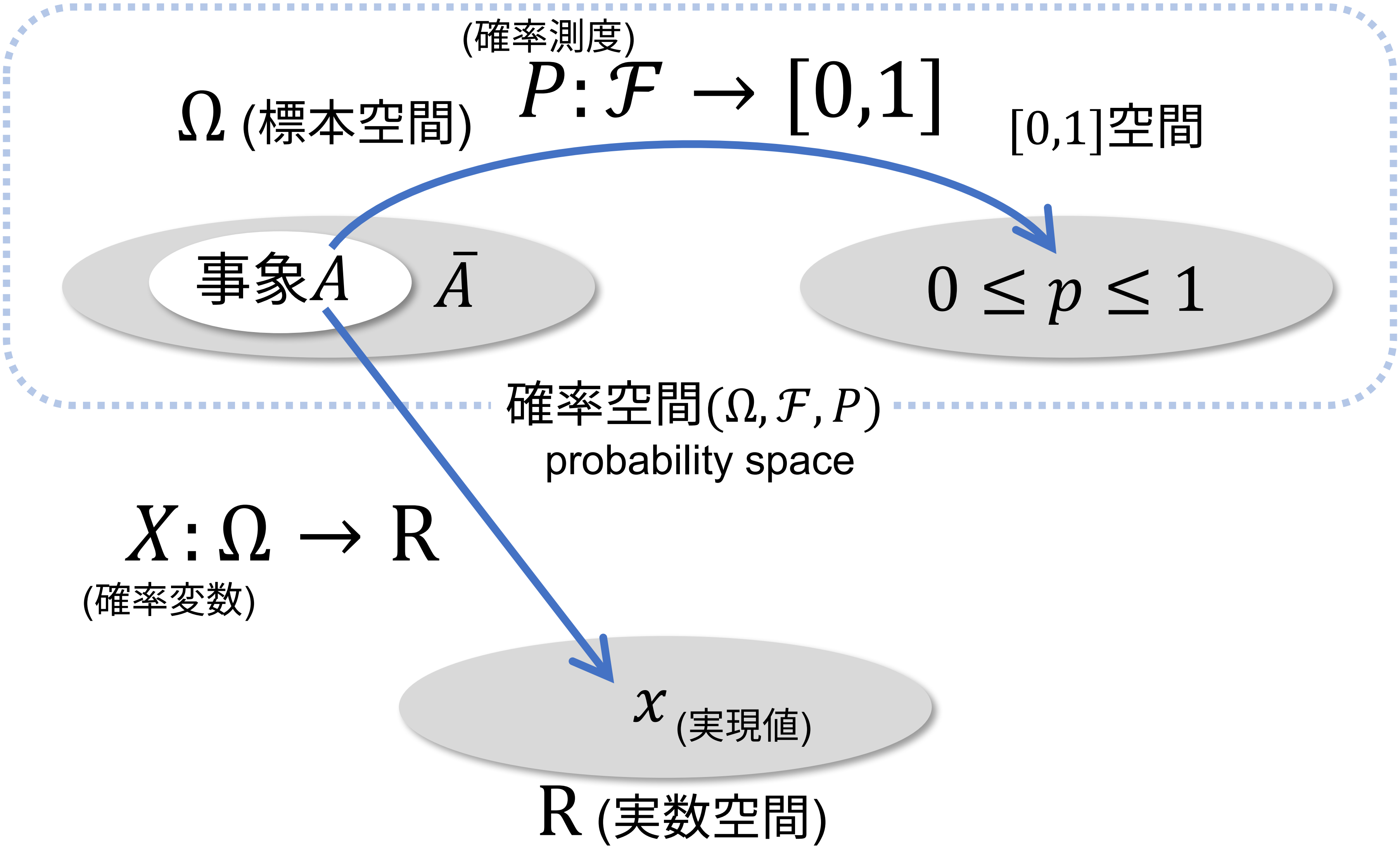
$$\int_a^b f(x)dx = \int_a^b \mathrm{f}(x)dx = p_{|a \le x \le b}$$

確率分布　$f(x) = P(A|X = x) = p$

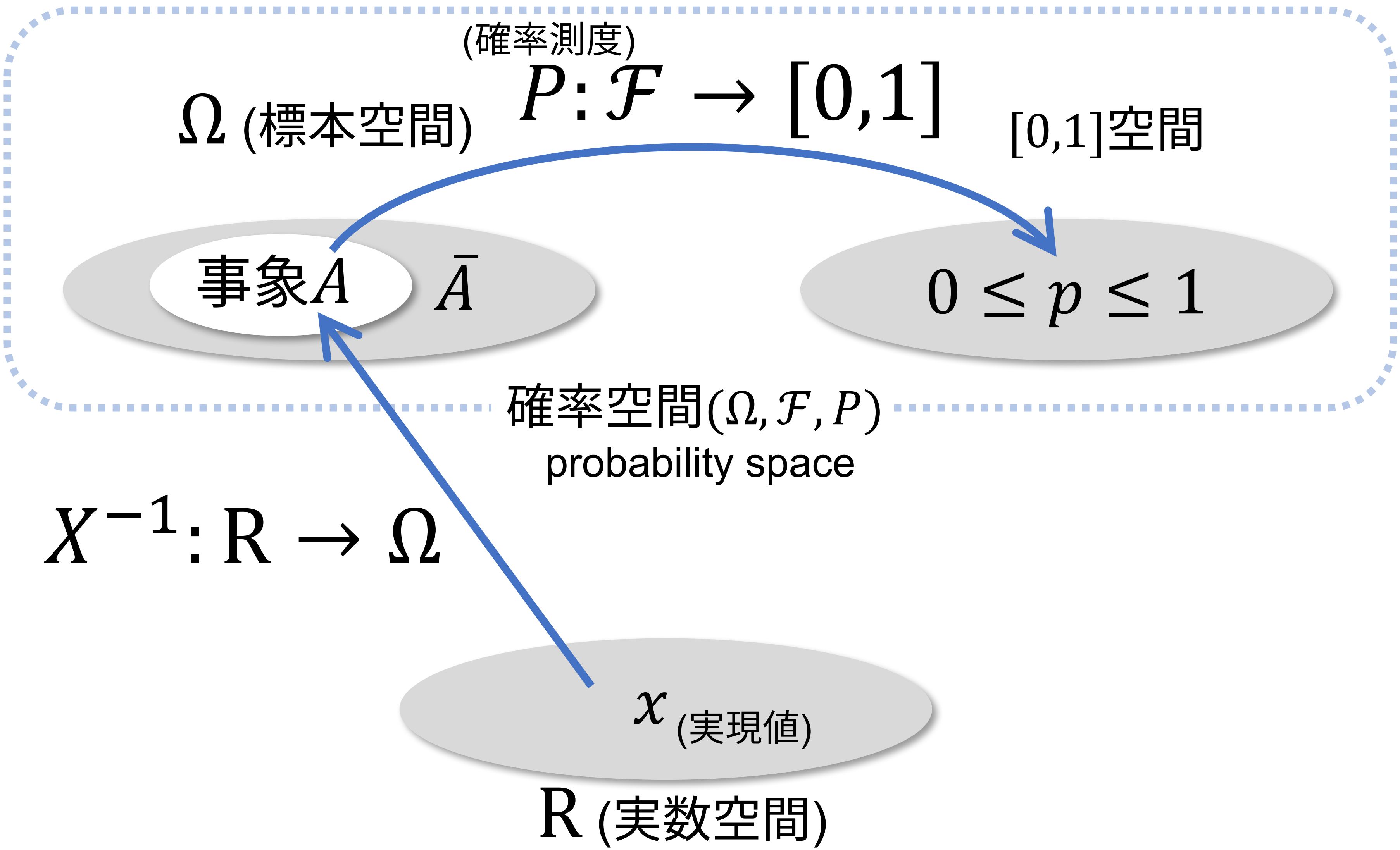
累積分布　$F(x) = \int_{-\infty}^{x} f(x)dx$

確率密度　$\mathrm{f}(x) = \frac{1}{dx}F(x)$

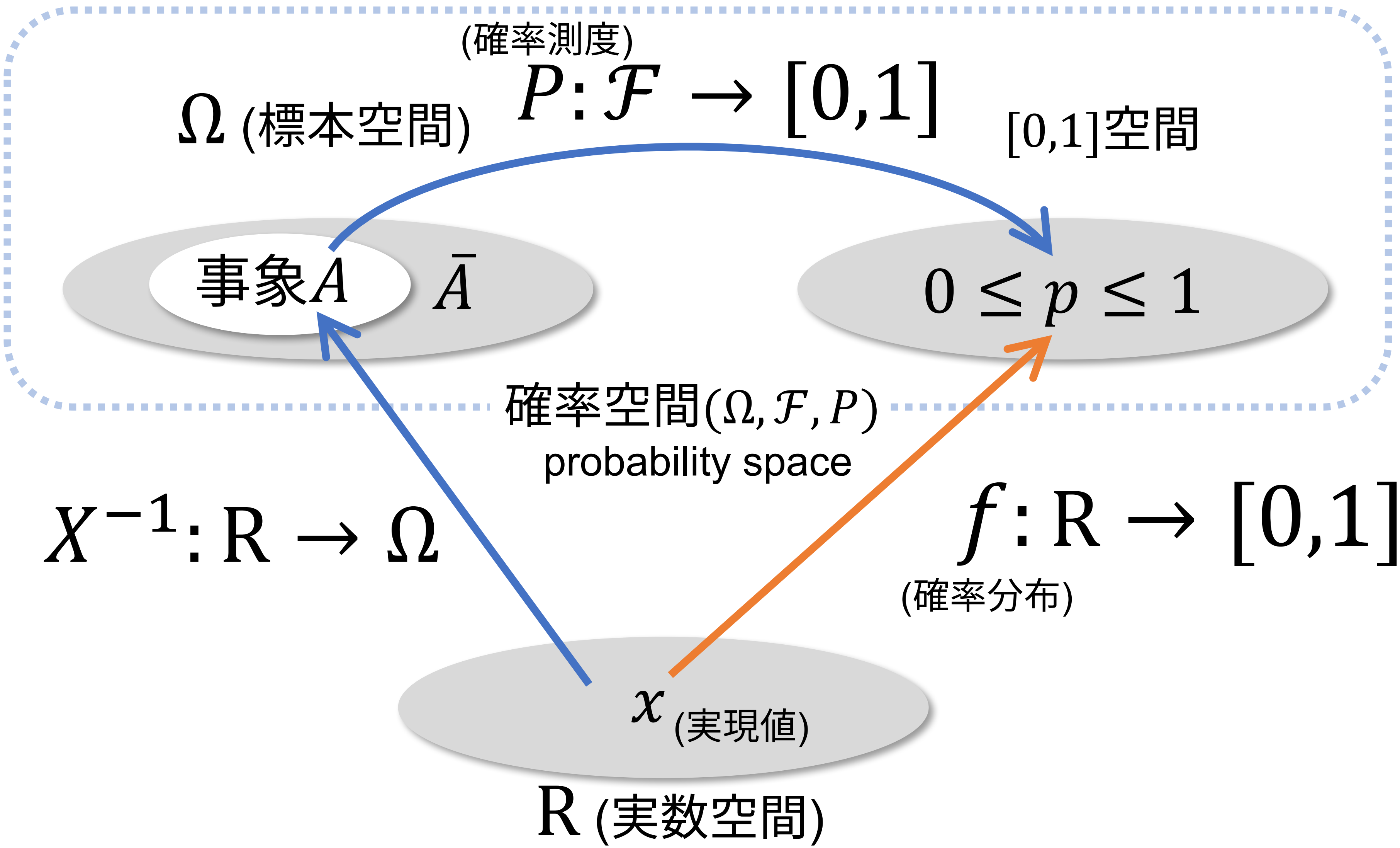
確率　$p_{|a \leq x \leq b} = \int_{a}^{b} \mathrm{f}(x)dx$

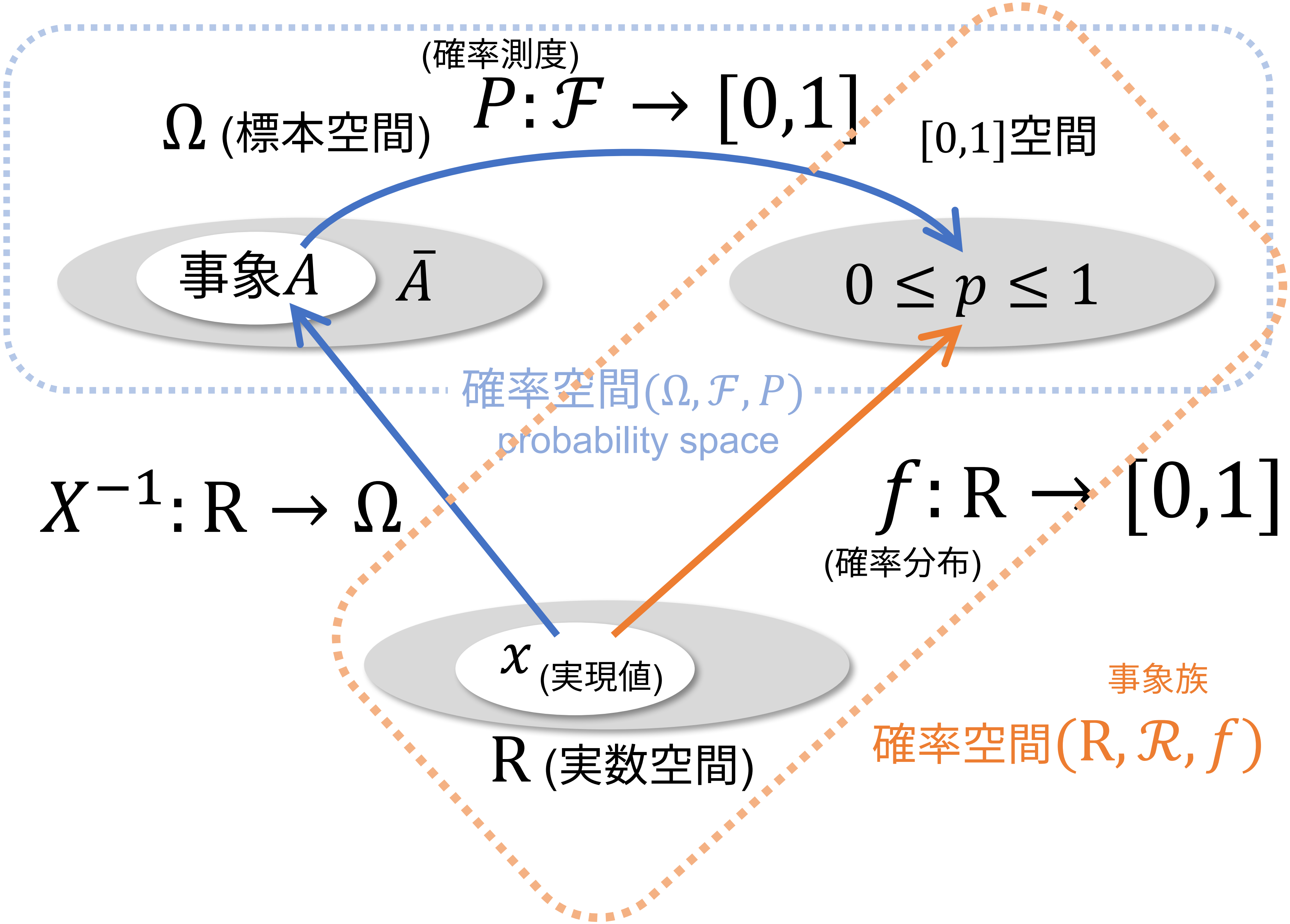


(確率測度)
$P: \mathcal{F} \to [0,1]$
Ω (標本空間)
[0,1]空間
事象A
$\bar{A}$
$0 \le p \le 1$
確率空間$(\Omega, \mathcal{F}, P)$
probability space
$X: \Omega \to \mathrm{R}$
(確率変数)
x (実現値)
R (実数空間)



(確率測度)
$P: \mathcal{F} \to [0,1]$
Ω (標本空間)
[0,1]空間
事象A
$\bar{A}$
$0 \leq p \leq 1$
確率空間$(\Omega, \mathcal{F}, P)$
probability space
$X^{-1}: \mathrm{R} \to \Omega$
x (実現値)
R (実数空間)


(確率測度)
$P: \mathcal{F} \to [0,1]$
Ω (標本空間)
[0,1]空間
事象A
$\bar{A}$
$0 \le p \le 1$
確率空間$(\Omega, \mathcal{F}, P)$
probability space
$X^{-1}: \mathrm{R} \to \Omega$
$f: \mathrm{R} \to [0,1]$
(確率分布)
x (実現値)
R (実数空間)



(確率測度)
$P: \mathcal{F} \to [0,1]$
Ω (標本空間)
[0,1]空間
事象A
$\bar{A}$
$0 \le p \le 1$
確率空間$(\Omega, \mathcal{F}, P)$
probability space
$X^{-1}: \mathrm{R} \to \Omega$
$f: \mathrm{R} \to [0,1]$
(確率分布)
x (実現値)
事象族
確率空間$(\mathrm{R}, \mathcal{R}, f)$
R (実数空間)

職業、イケメン。テラモナギ
@teramonagi

$x=X(\omega)$は $\{\omega\in\Omega; X(\omega)=x\}$の略ですね。なのでこれは$\Omega$の部分集合を表すための略記です。$P(x=X(\omega)) = P(\{\omega\in\Omega; X(\omega)=x\})$ということです。なので確率測度は$\Omega$の部分集合を引数にとる関数です！

15:26 - 2015年11月17日

もし、わからない事があったら
完全に理解しているヒトにドンドン聞いてみよう！

確率変数Xは正規分布に従う

確率変数Xは正規分布に従う

確率密度関数f(x)が

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

確率密度関数f(x)が

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

理解

「お、要するにy $= e^{-x^2}$を考えて
$x=\mu$でy_{max}, $x=\pm\infty$で$y_{min}=0$。
そのなだらかさの度合いがδかな。」

確率密度関数f(x)が

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

🤔
理解

「お、要するにy $= e^{-x^2}$を考えて
$x = \mu$でy_{max}, $x = \pm\infty$で$y_{min} = 0$。
そのなだらかさの度合いがδかな。」

🐱
完全理解
(論理)

「にゃーん😸😸😸😸」

確率密度関数f(x)が

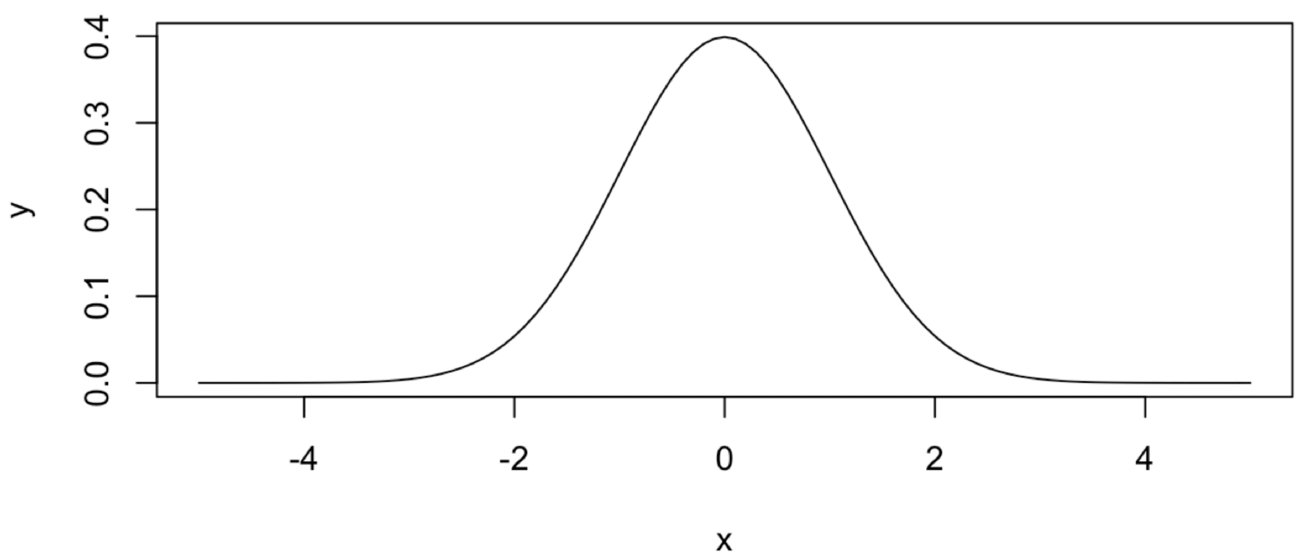
$$f(x|\mu,\delta^2) = \frac{1}{\sqrt{2\pi\delta^2}} e^{-\frac{(x-\mu)^2}{2\delta^2}}$$

```
y <- function(x, mu = 0, s = 1){
  (2 * pi * s^2)^(-1/2) * exp(-(x - mu)^2/(2 * s^2))
}

plot(y, -5, 5)
```

確率密度関数f(x)が

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

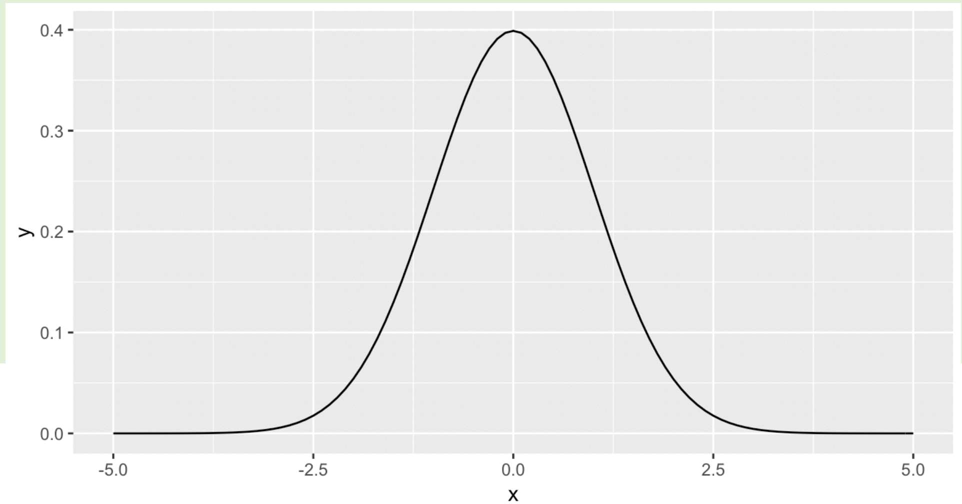
```
y <- function(x, mu = 0, s = 1){
  (2 * pi * s^2)^(-1/2) * exp(-(x - mu)^2/(2 * s^2))
}

plot(y, -5, 5)
```



確率密度関数f(x)が

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

```
library(tidyverse)

dat_g <- data.frame(x = seq(-5, 5, by = 0.1)) %>%
  mutate(y = dnorm(x, 0, 1))

dat_g %>%
  ggplot(aes(x, y))+
  geom_path()
```



確率分布に関わるRの関数群

確率密度関数	d** (第1引数はx座標)
累積確率分布	p** (第1引数はx座標)
乱数	r** (第1引数は個数)
確率点	q** (第1引数は割合)

確率分布の略号

正規分布：norm

ポアソン分布：pois

一様分布：unif

対数正規分布：lnorm

ベータ分布：beta

などなど

確率分布に関わるRの関数群

済	確率密度関数	d** (第1引数はx座標)
済	累積確率分布	p** (第1引数はx座標)
	乱数	r** (第1引数は個数)
	確率点	q** (第1引数は割合)

確率分布の略号

正規分布：norm

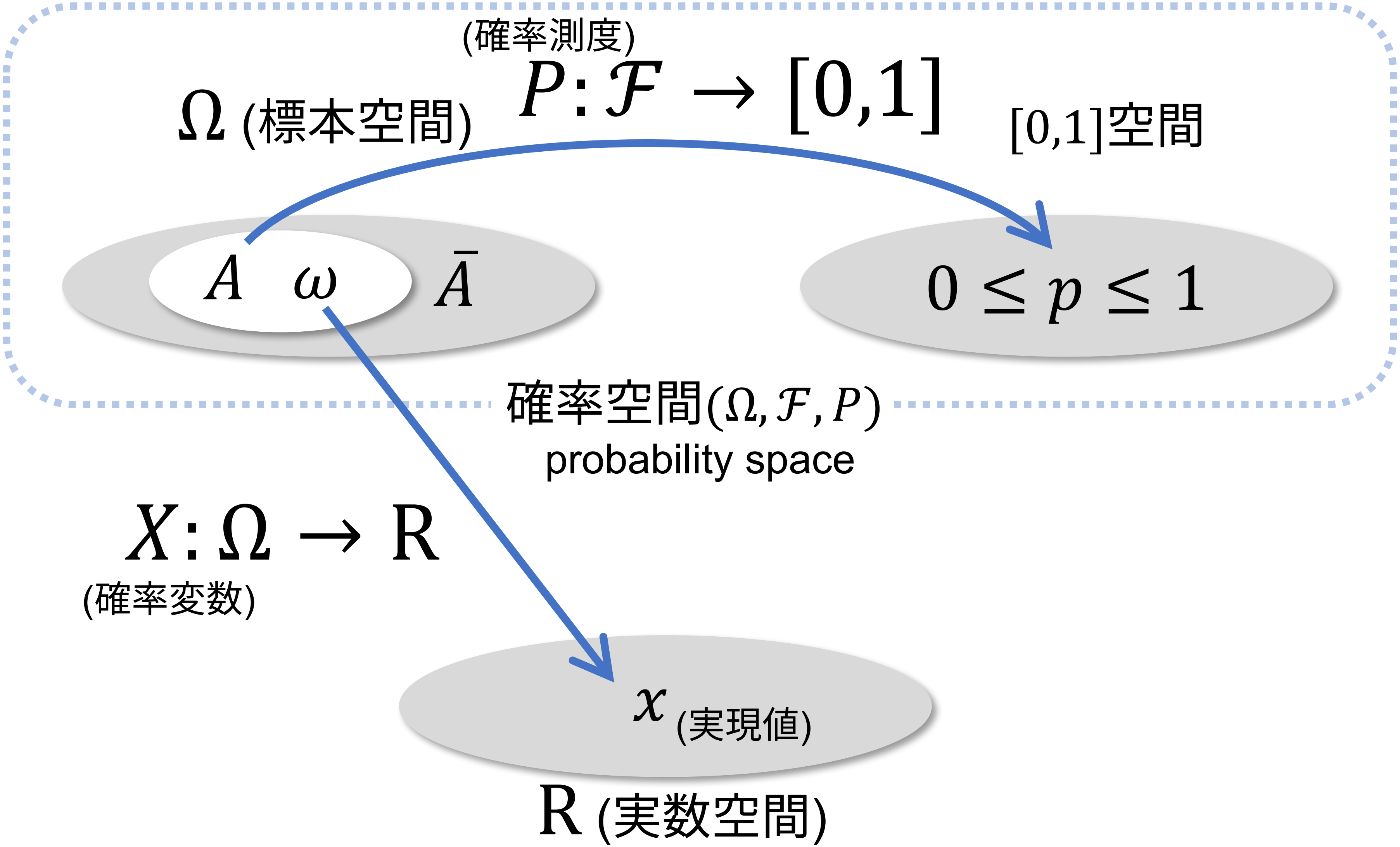
ポアソン分布：pois

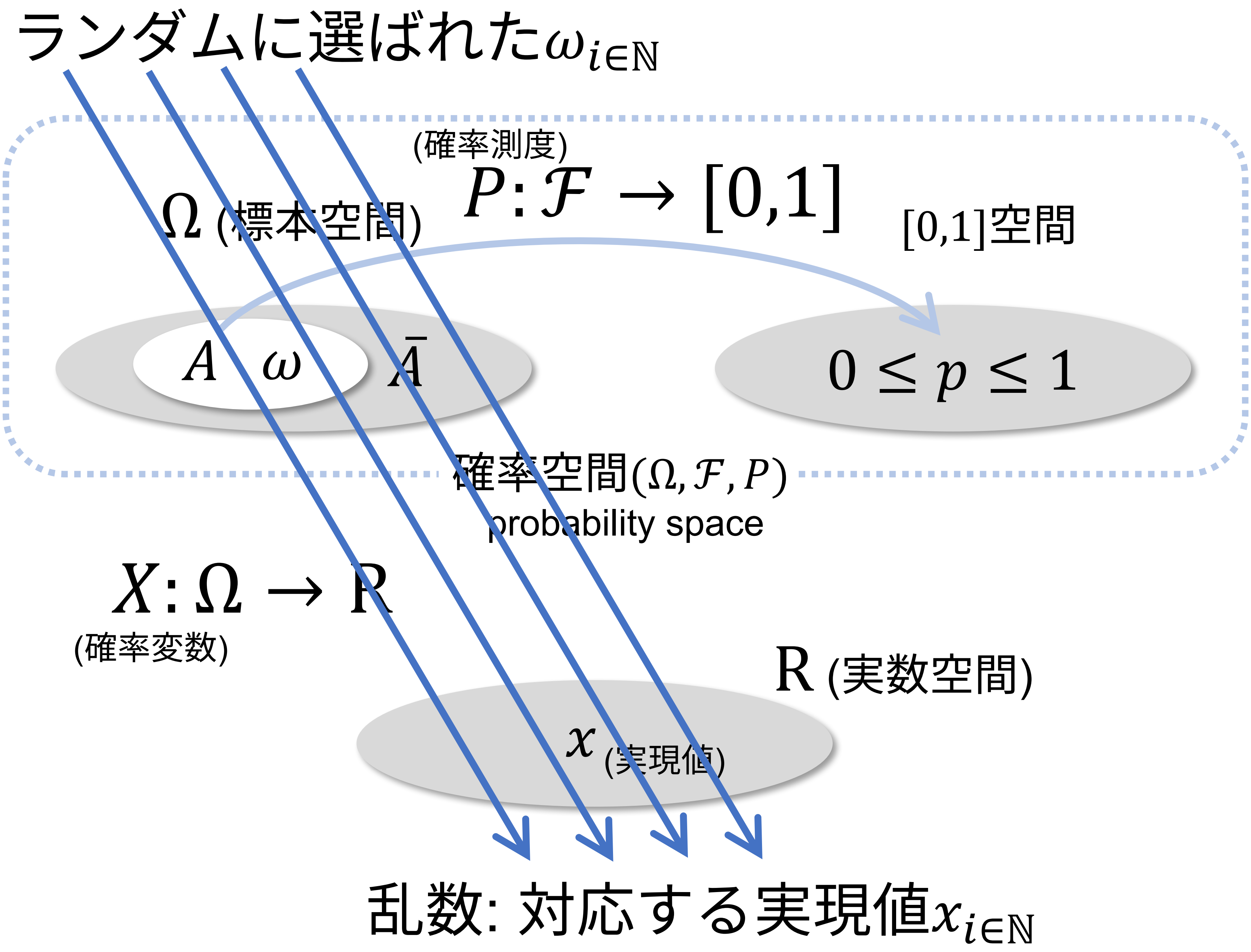
一様分布：unif

対数正規分布：lnorm

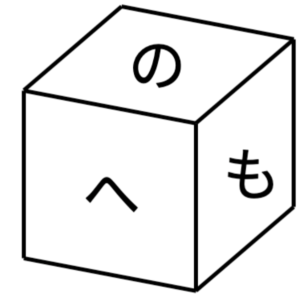
ベータ分布：beta

などなど

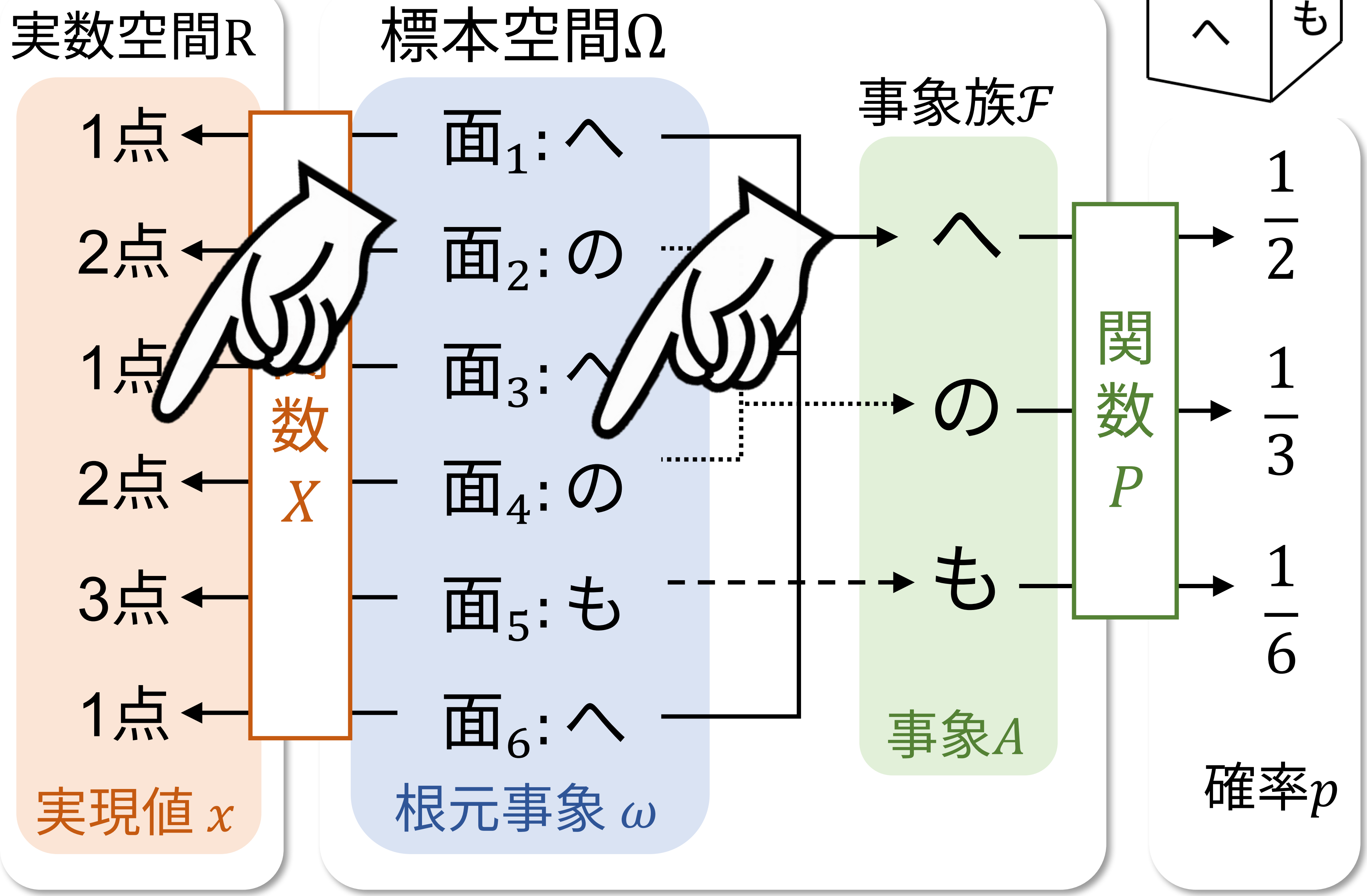

(確率測度)
$P: \mathcal{F} \to [0,1]$
Ω (標本空間)
[0,1]空間
A ω $\bar{A}$
$0 \le p \le 1$
確率空間$(\Omega, \mathcal{F}, P)$
probability space
$X: \Omega \to \mathrm{R}$
(確率変数)
x (実現値)
R (実数空間)


ランダムに選ばれた$\omega_{i\in\mathbb{N}}$
(確率測度)
$P: \mathcal{F} \to [0,1]$
Ω (標本空間)
[0,1]空間
A
ω
$\bar{A}$
$0 \le p \le 1$
確率空間$(\Omega, \mathcal{F}, P)$
probability space
$X: \Omega \to \mathrm{R}$
(確率変数)
R (実数空間)
x (実現値)
乱数: 対応する実現値$x_{i\in\mathbb{N}}$

へのへのもへサイコロ

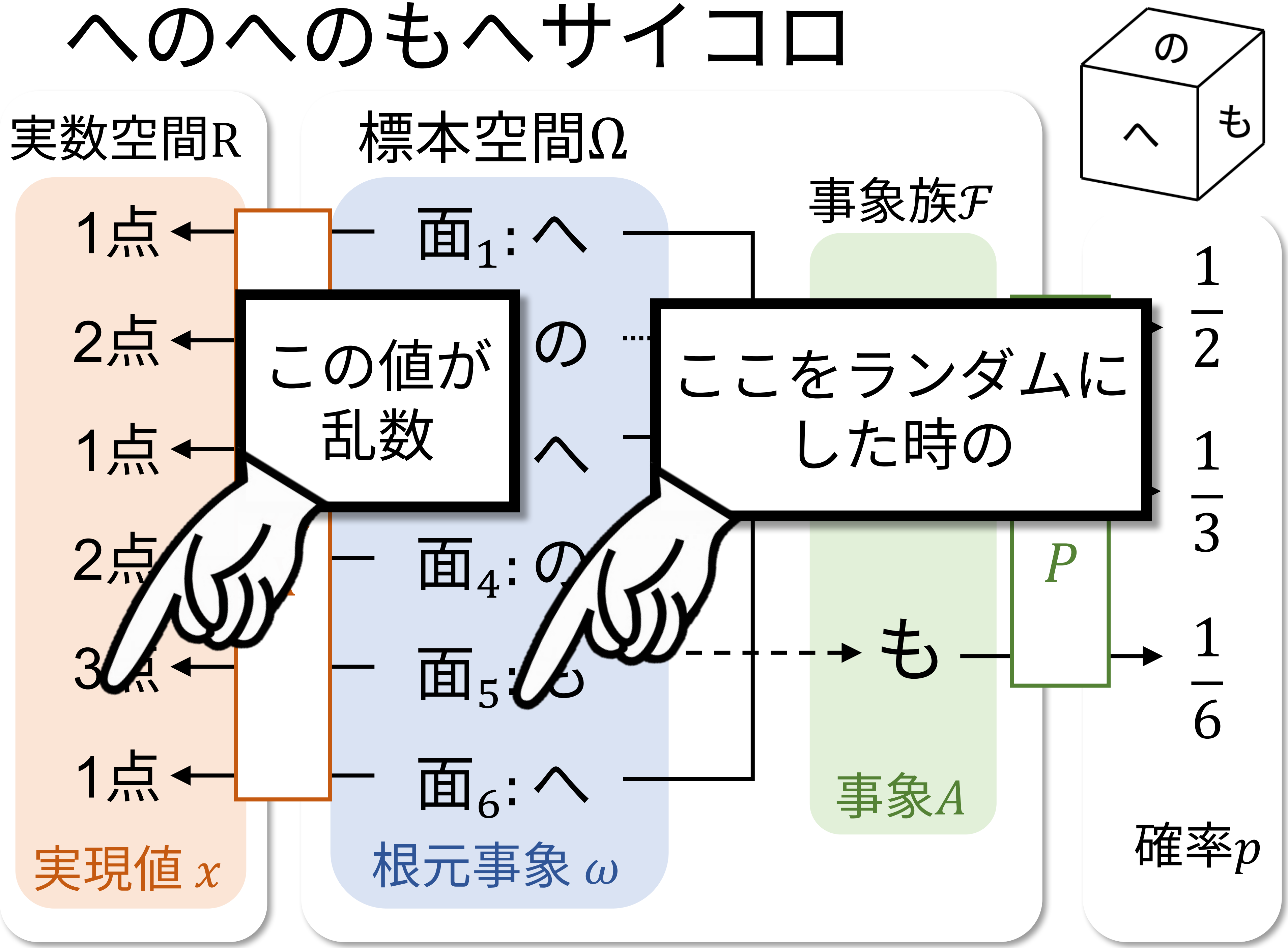


の
へ
も

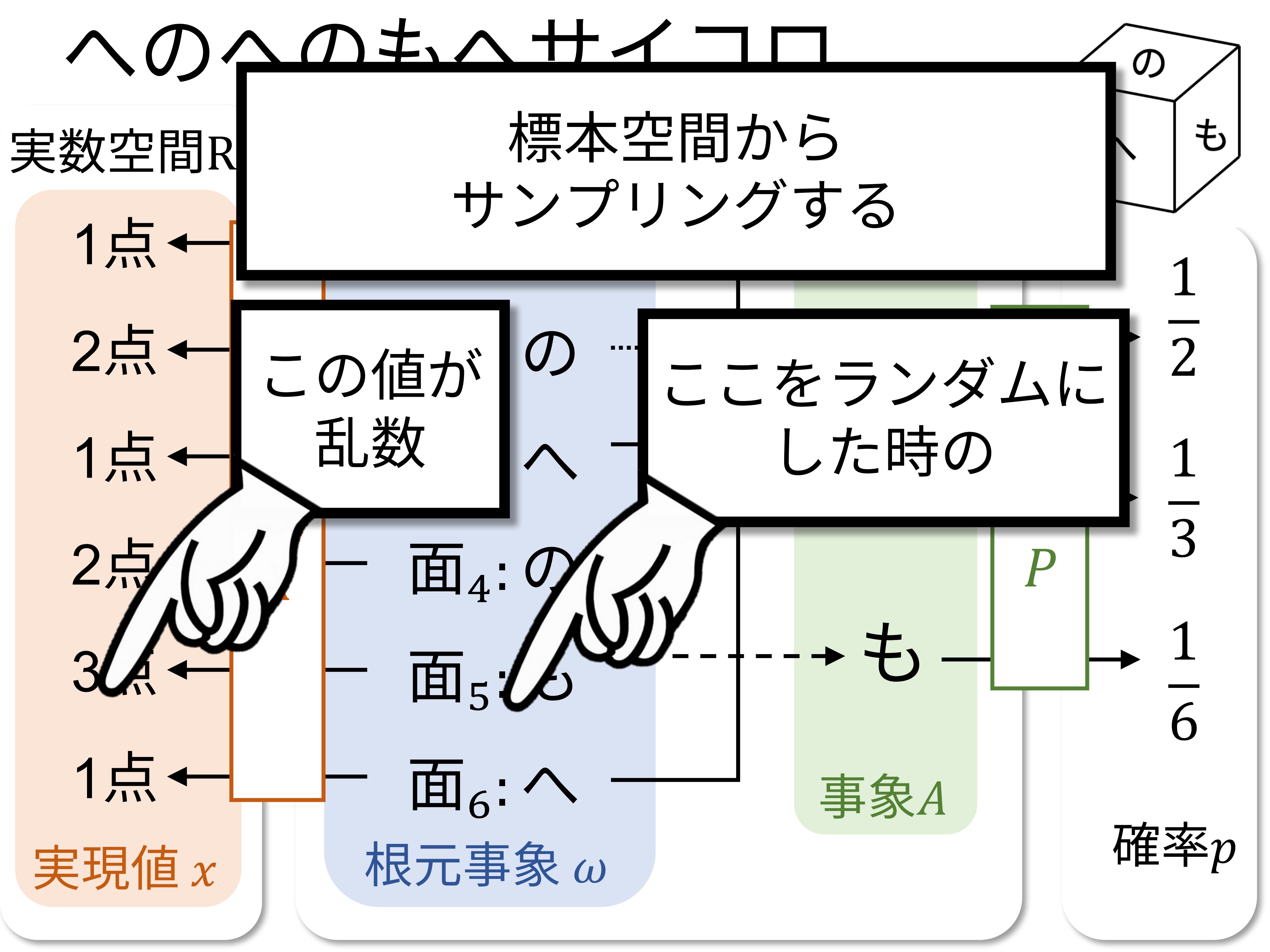


実数空間R
標本空間Ω
事象族$\mathcal{F}$
1点
2点
1点
2点
3点
1点
数
X
面$_1$:へ
面$_2$:の
面$_3$:へ
面$_4$:の
面$_5$:も
面$_6$:へ
へ
の
も
関数
P
$\frac{1}{2}$
$\frac{1}{3}$
$\frac{1}{6}$
実現値 x
根元事象 ω
事象A
確率p

へのへのもへサイコロ



の
へ
も
実数空間R
標本空間Ω
事象族𝓕
1点
2点
1点
2点
3点
1点
実現値 x
面1: へ
の
へ
面4: の
面5: も
面6: へ
根元事象 ω
も
事象A
P
1/2
1/3
1/6
確率p
この値が
乱数
ここをランダムに
した時の

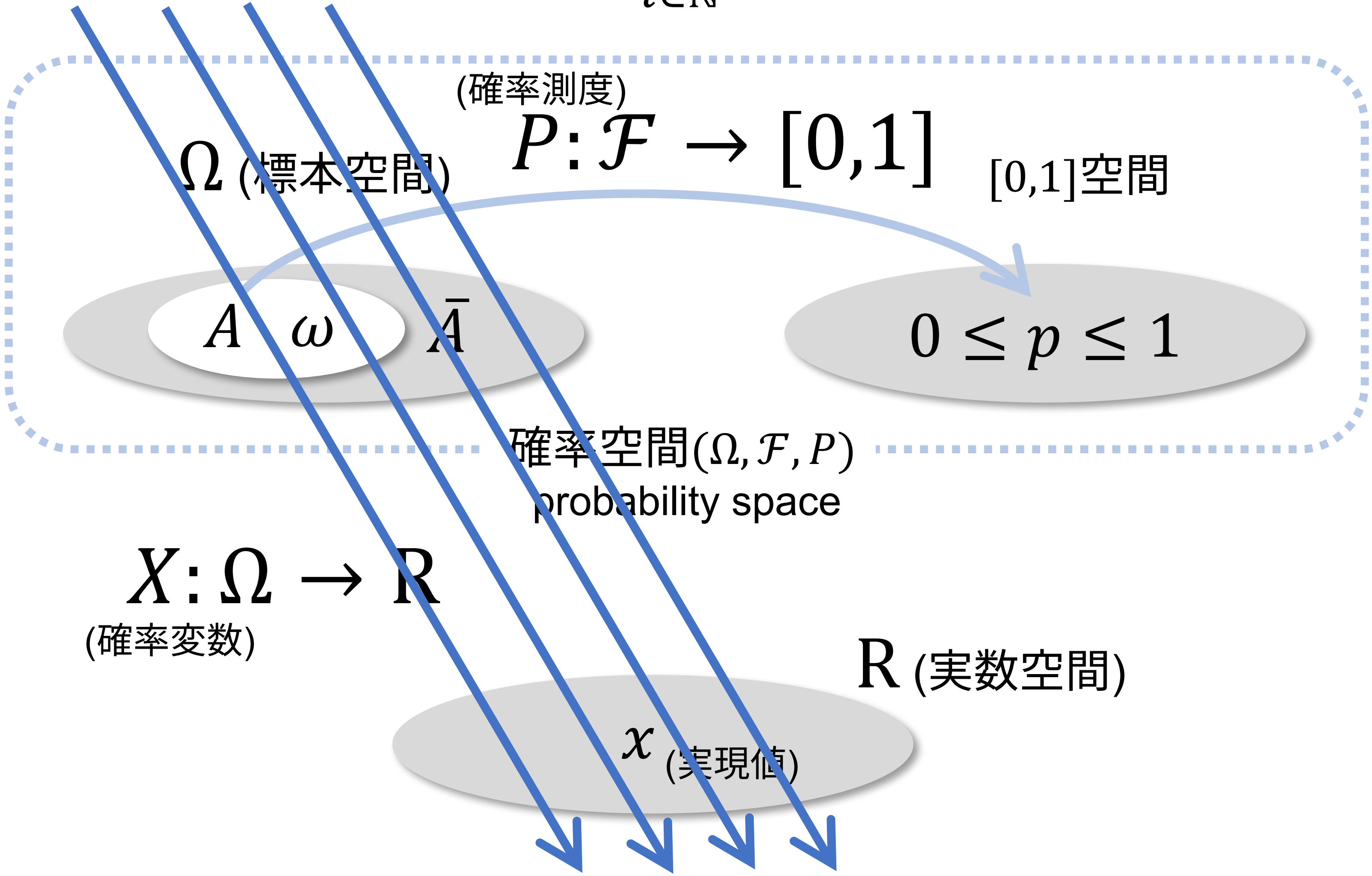


へのへのもへサイコロ
の
も
実数空間R
標本空間から
サンプリングする
1点
2点
1点
2点
3点
1点
この値が
乱数
の
へ
面$_4$: の
面$_5$: も
面$_6$: へ
ここをランダムに
した時の
も
P
事象A
$\frac{1}{2}$
$\frac{1}{3}$
$\frac{1}{6}$
実現値 x
根元事象 ω
確率p

正規分布の乱数

Ωにおけるωの分布は
一般には取り扱えない
(ピッタリ180cmの確率はゼロ)

ランダムに選ばれた$\omega_{i\in\mathbb{N}}$



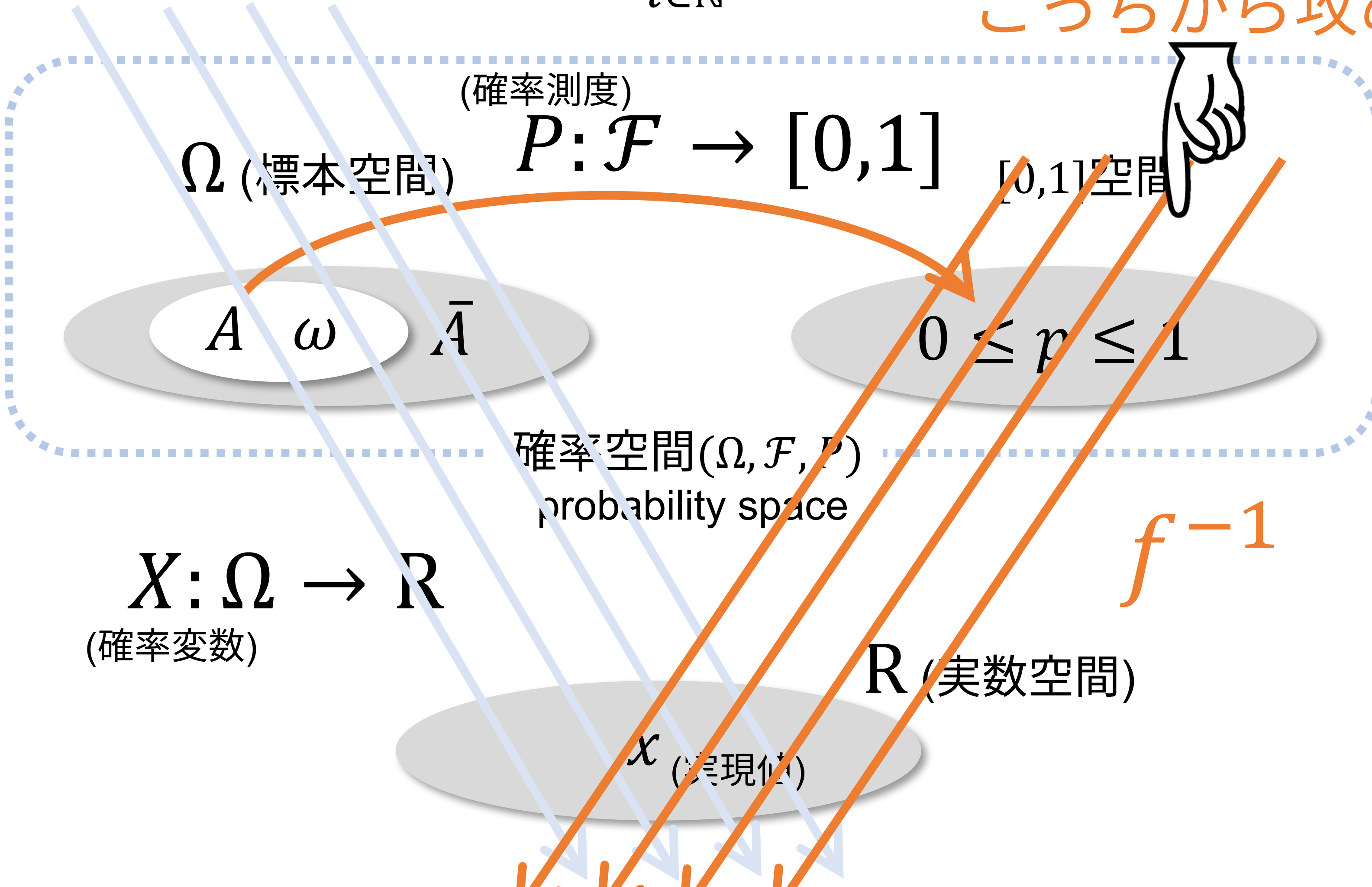
乱数: 対応する実現値$x_{i\in\mathbb{N}}$

正規分布の乱数

Ωにおけるωの分布は
一般には取り扱えない
(ピッタリ180cmの確率はゼロ)

ランダムに選ばれた$\omega_{i \in \mathbb{N}}$

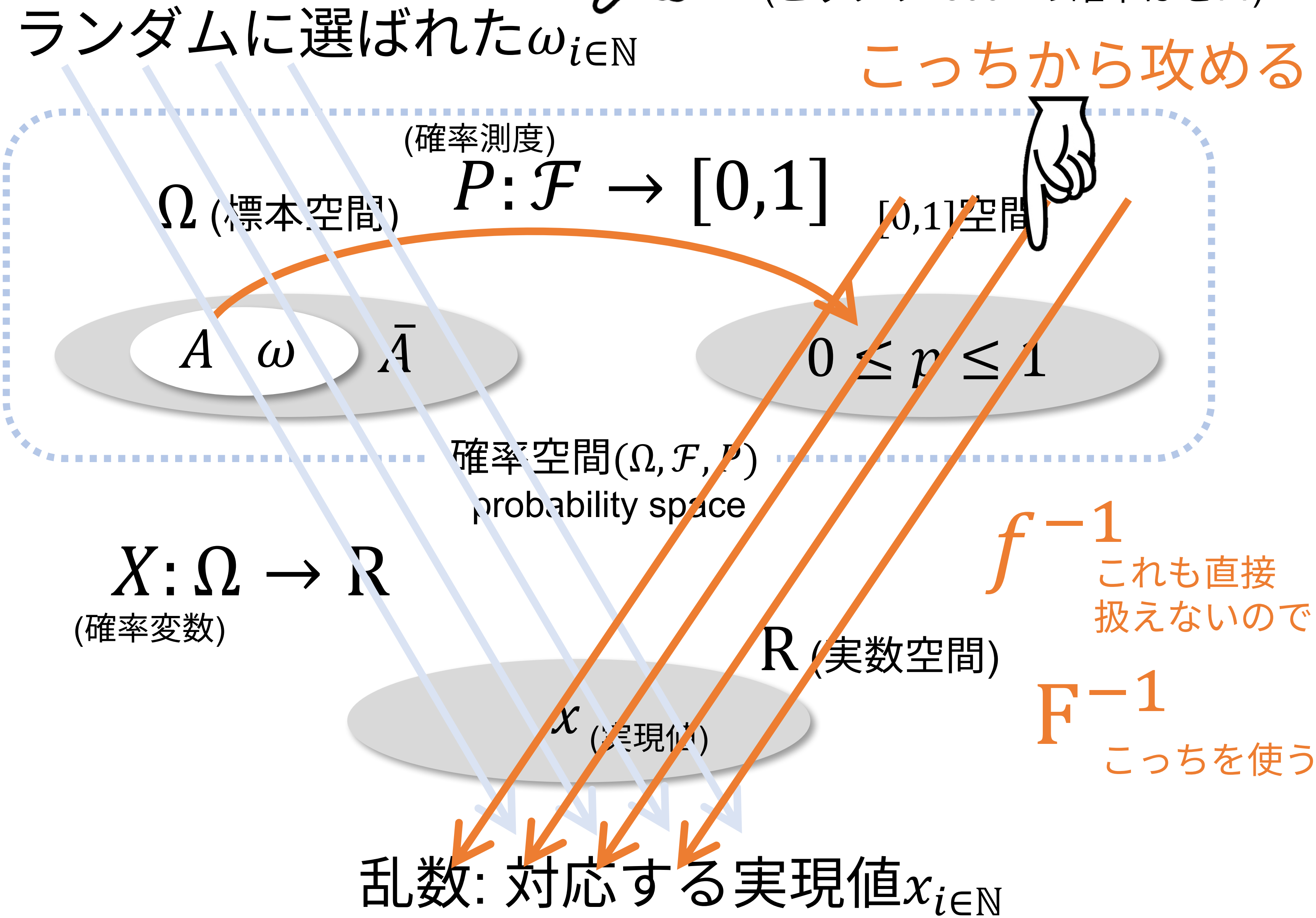
こっちから攻める



乱数: 対応する実現値$x_{i \in \mathbb{N}}$

正規分布の乱数

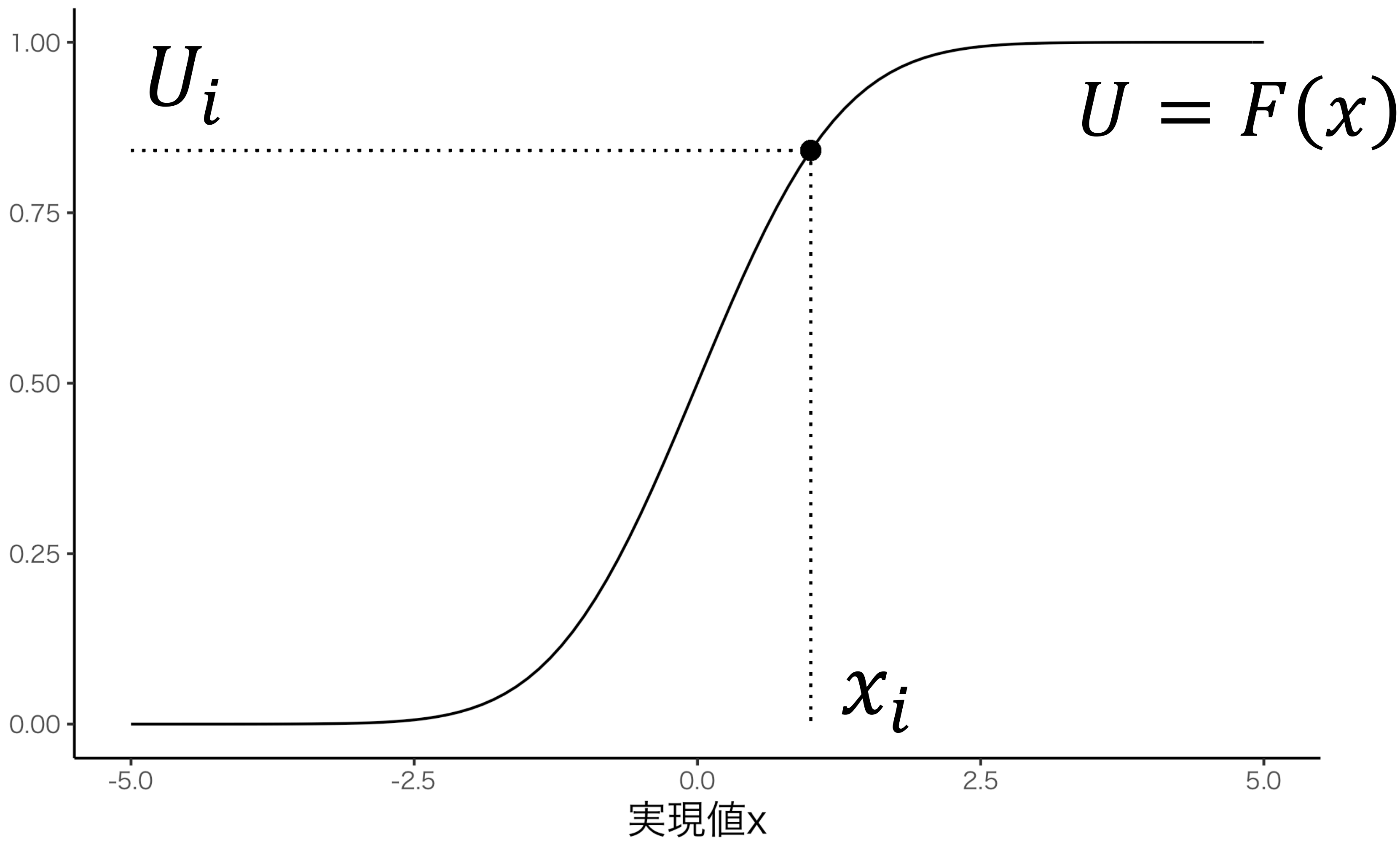
Ωにおけるωの分布は
一般には取り扱えない
(ピッタリ180cmの確率はゼロ)



正規分布の乱数

累積分布$F(x)$は必ず$0 \leq F(x) \leq 1$で、かつ、単調増加関数。

$F(x) = U$を$[0,1]$において均一にサンプリングすれば良い。

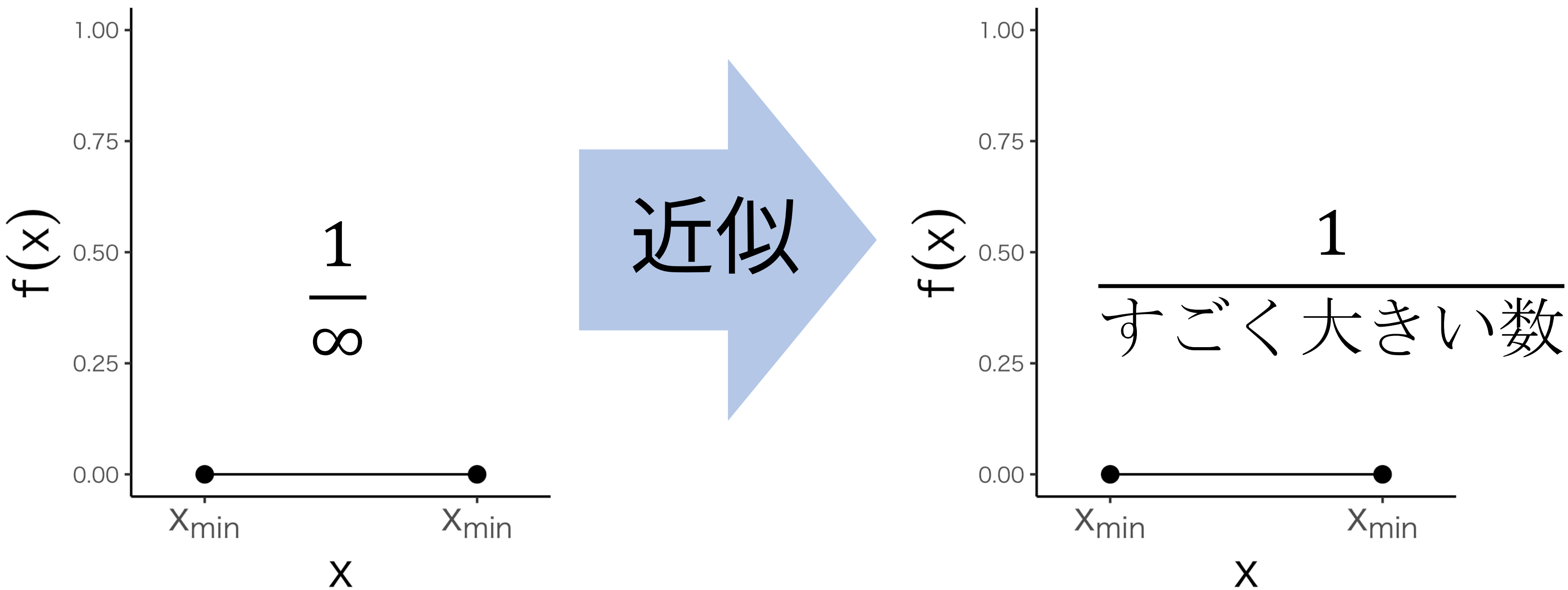


正規分布の乱数

累積分布$F(x)$は必ず$0 \le F(x) \le 1$で、かつ、単調増加関数。

$F(x) = U$を$[0, 1]$において均一にサンプリングすれば良い。

一様分布の確率分布$f(x)$

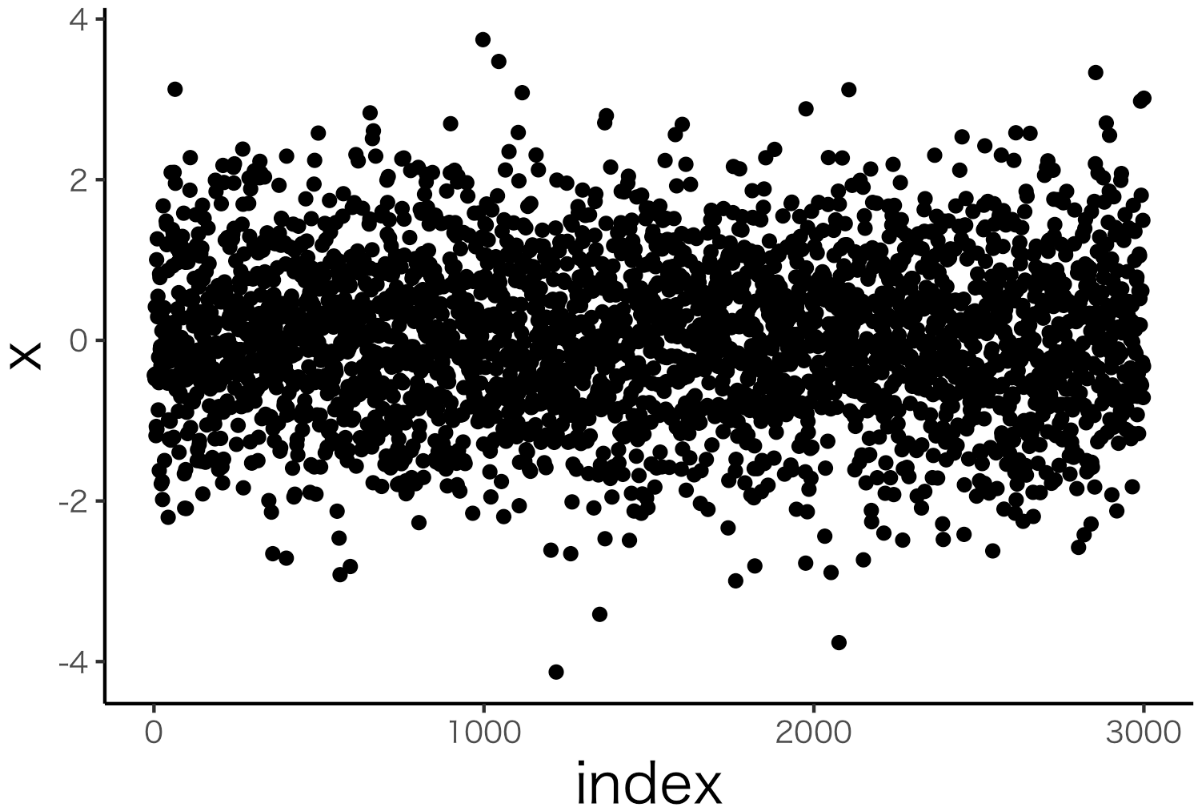


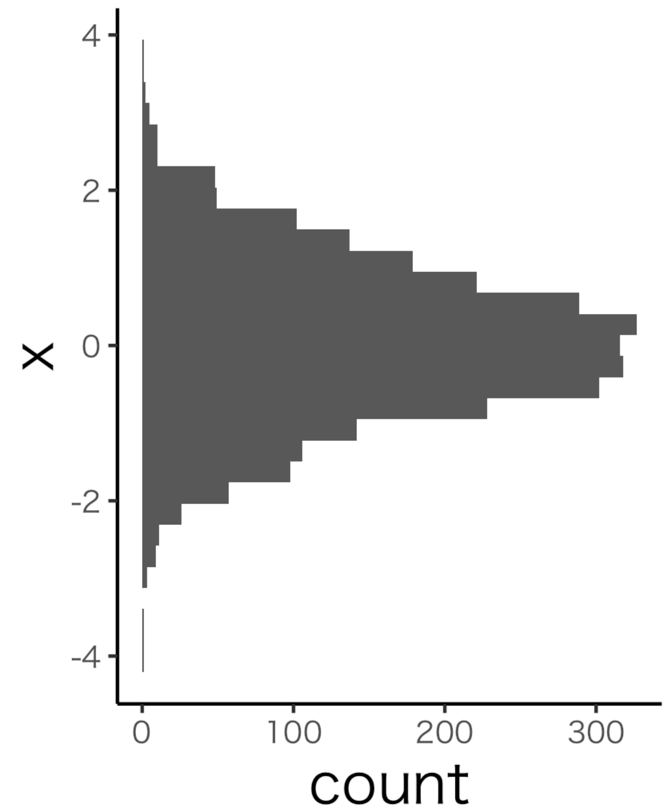
Rで取り扱える最大数は10^{300}ぐらい？

```r
set.seed(71)
dat <- data.frame(x = rnorm(3000)) %>%
  rowid_to_column("index")
```

```r
ggplot(dat, aes(index, x))+
  geom_point()
```

```r
ggplot(dat, aes(x))+
  geom_histogram()+
  coord_flip()
```





確率変数Xは正規分布に従う

⇩

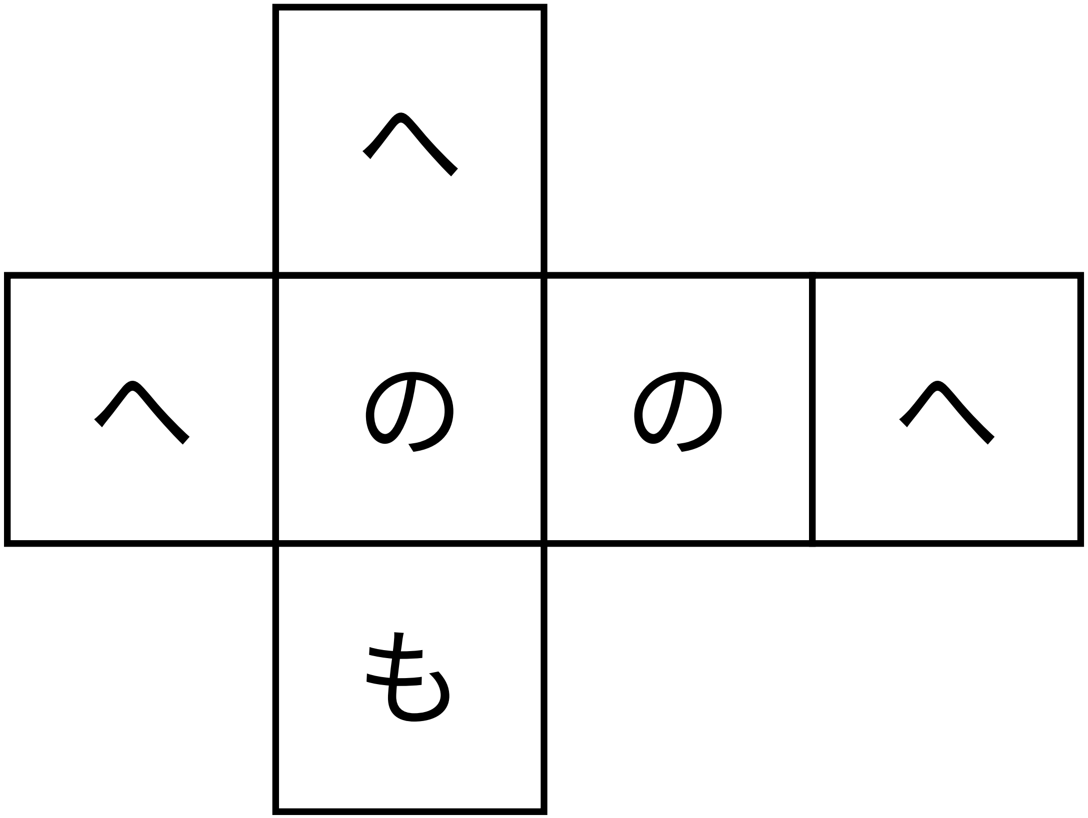
確率密度関数f(x)が

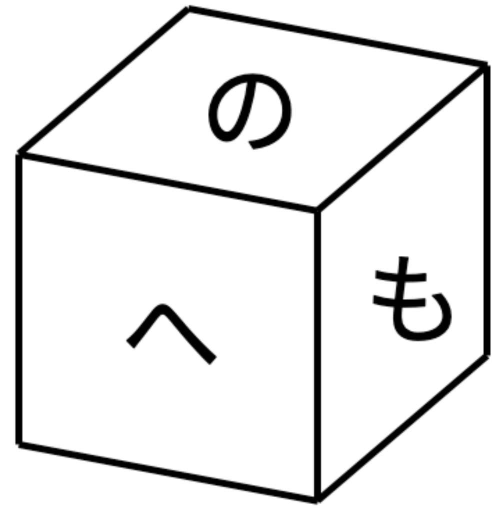
$$f(x|\mu,\delta^2)=\frac{1}{\sqrt{2\pi\delta^2}}e^{-\frac{(x-\mu)^2}{2\delta^2}}$$

となる確率分布から生じるサンプルxを実現値に持つ確率変数Xを考える。

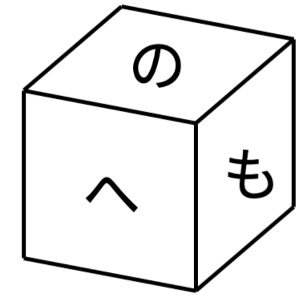
まとめ

へのへのもへサイコロ

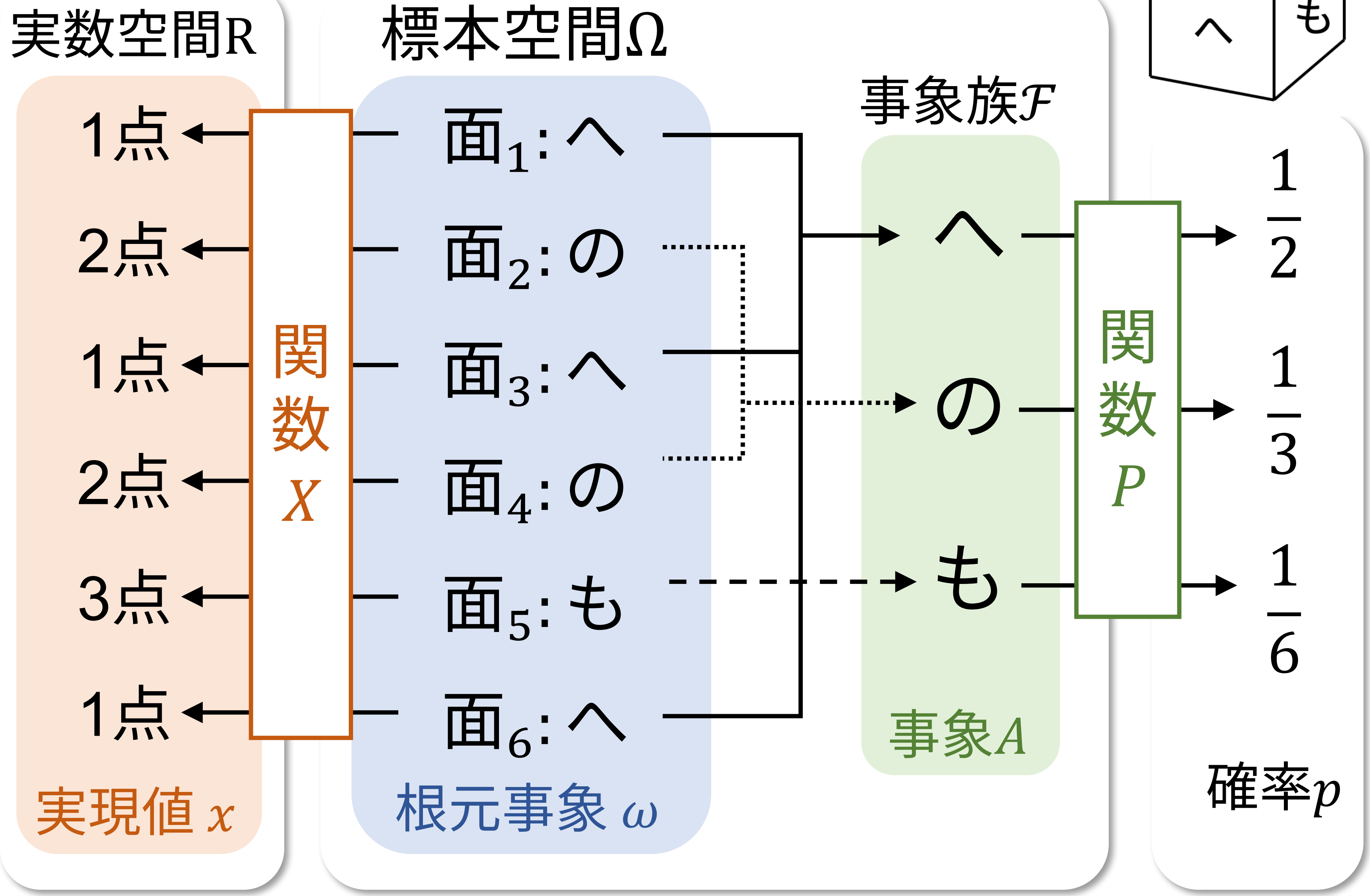




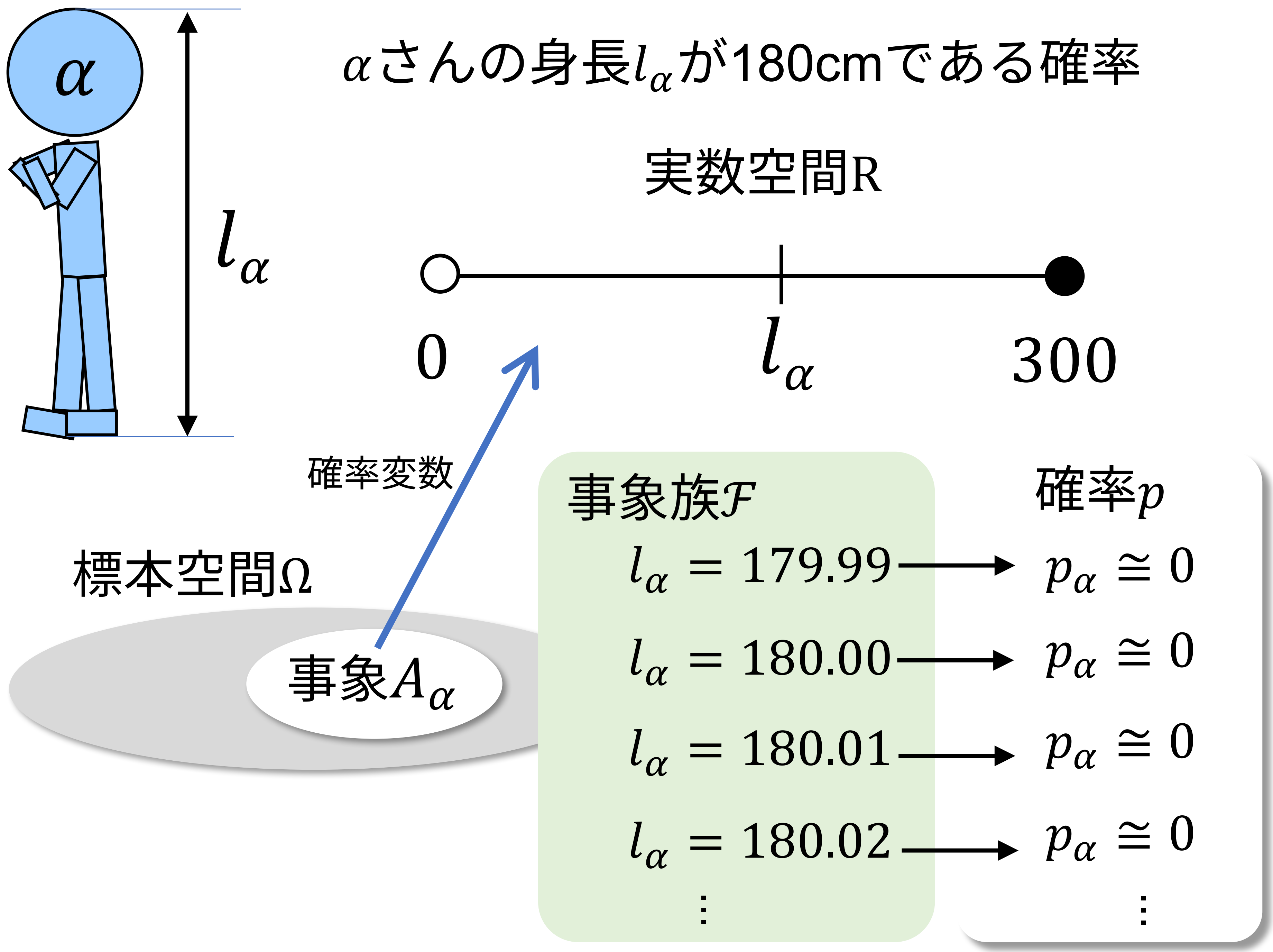
へのへのもへサイコロ



の
へ
も



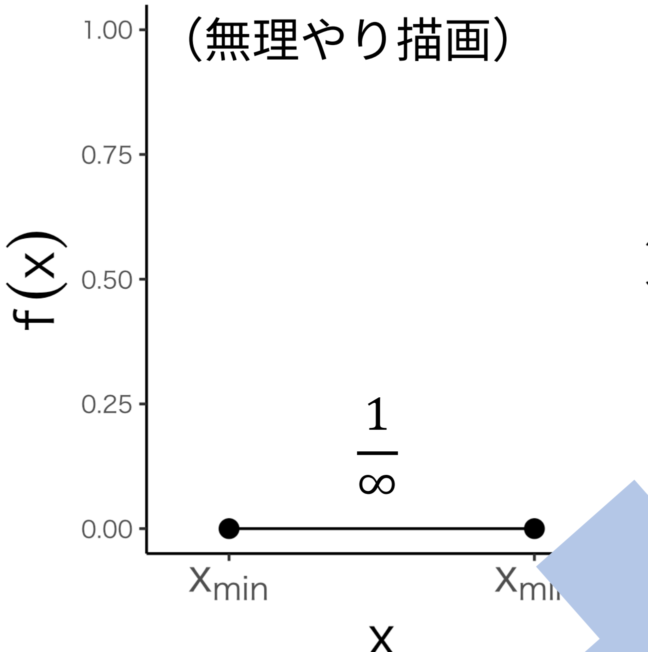
実数空間R
1点
2点
1点
2点
3点
1点
実現値 x
関数 X
標本空間Ω
面$_1$: へ
面$_2$: の
面$_3$: へ
面$_4$: の
面$_5$: も
面$_6$: へ
根元事象 ω
事象族$\mathcal{F}$
へ
の
も
事象A
関数 P
$\frac{1}{2}$
$\frac{1}{3}$
$\frac{1}{6}$
確率p


αさんの身長l_αが180cmである確率
α
l_α
実数空間R
0
l_α
300
確率変数
事象族$\mathcal{F}$
確率p
標本空間Ω
事象A_α
$l_\alpha = 179.99$
$p_\alpha \cong 0$
$l_\alpha = 180.00$
$p_\alpha \cong 0$
$l_\alpha = 180.01$
$p_\alpha \cong 0$
$l_\alpha = 180.02$
$p_\alpha \cong 0$
⋮
⋮

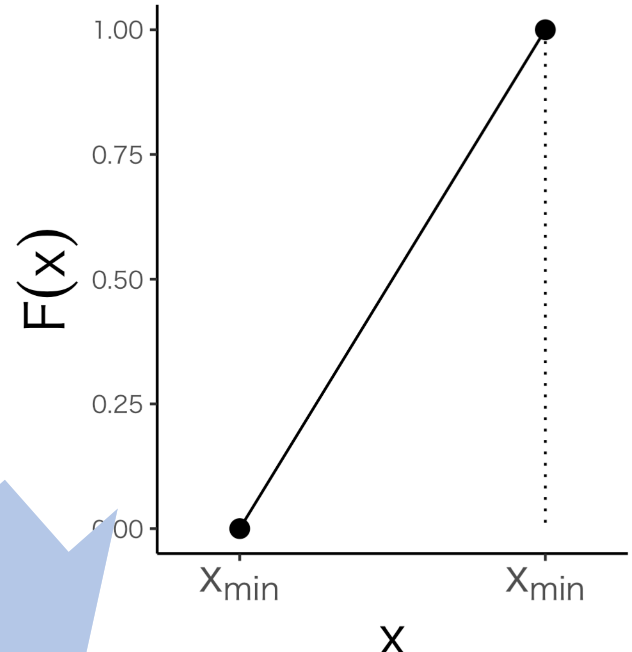
確率分布$f(x)$: 定義域$[x_{min}, x_{max}]$において均一

$$p = f(x) \qquad F(x) = \int_{-\infty}^{x} f(x)dx \qquad \mathrm{f}(x) = \frac{1}{dx}F(x)$$

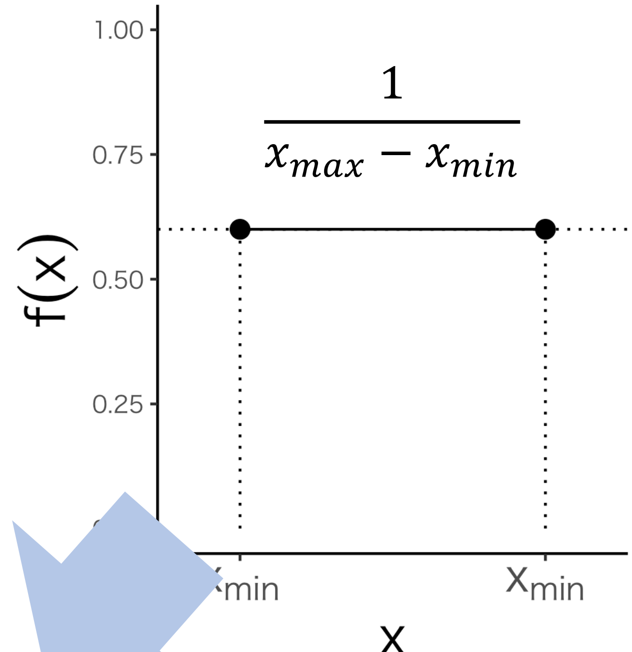
確率分布



累積分布



確率密度



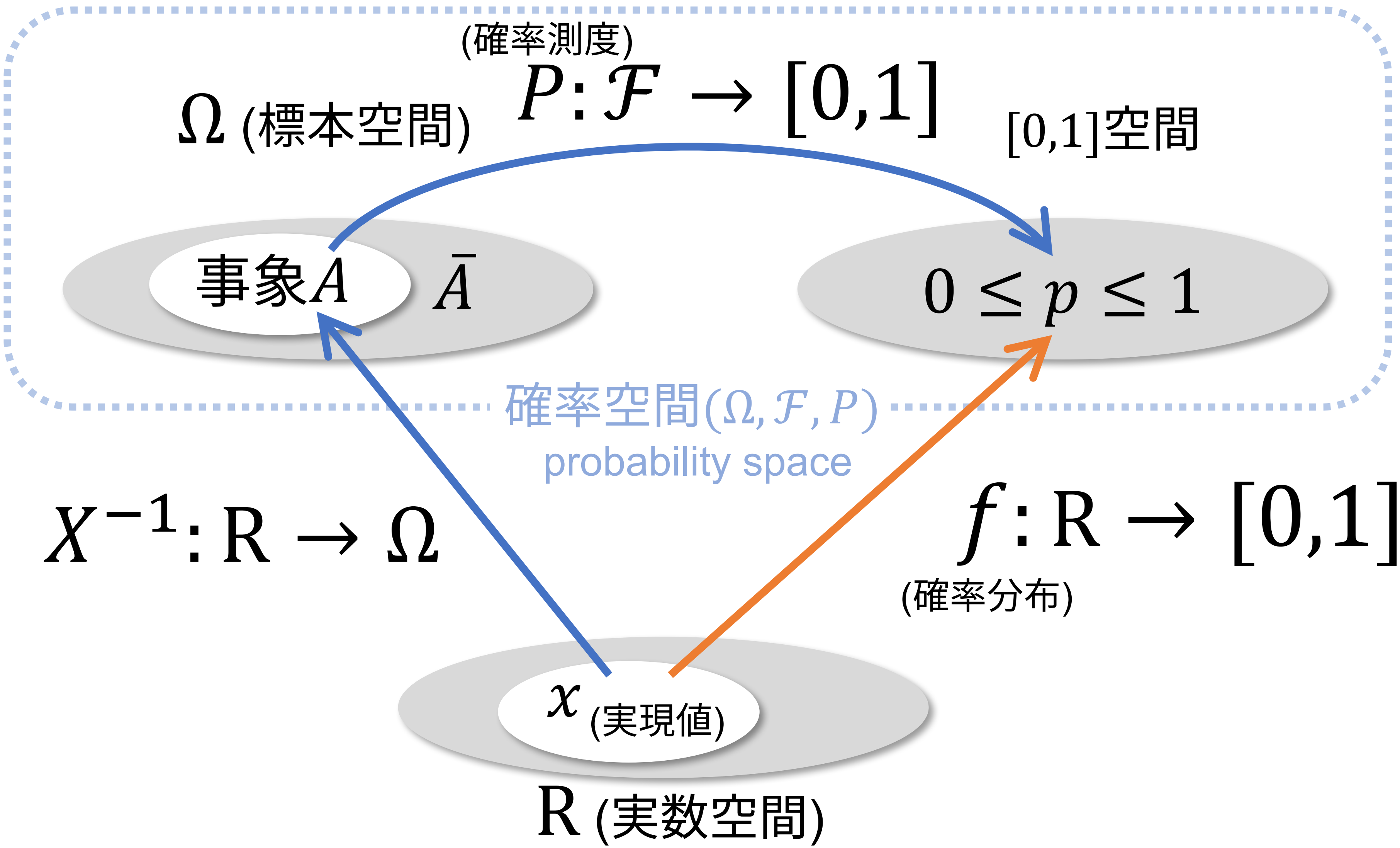
$$\int_a^b f(x)dx = \int_a^b \mathrm{f}(x)dx = p_{|a \le x \le b}$$

確率分布 $f(x) = P(A|X = x) = p$

累積分布 $F(x) = \int_{-\infty}^{x} f(x)dx$

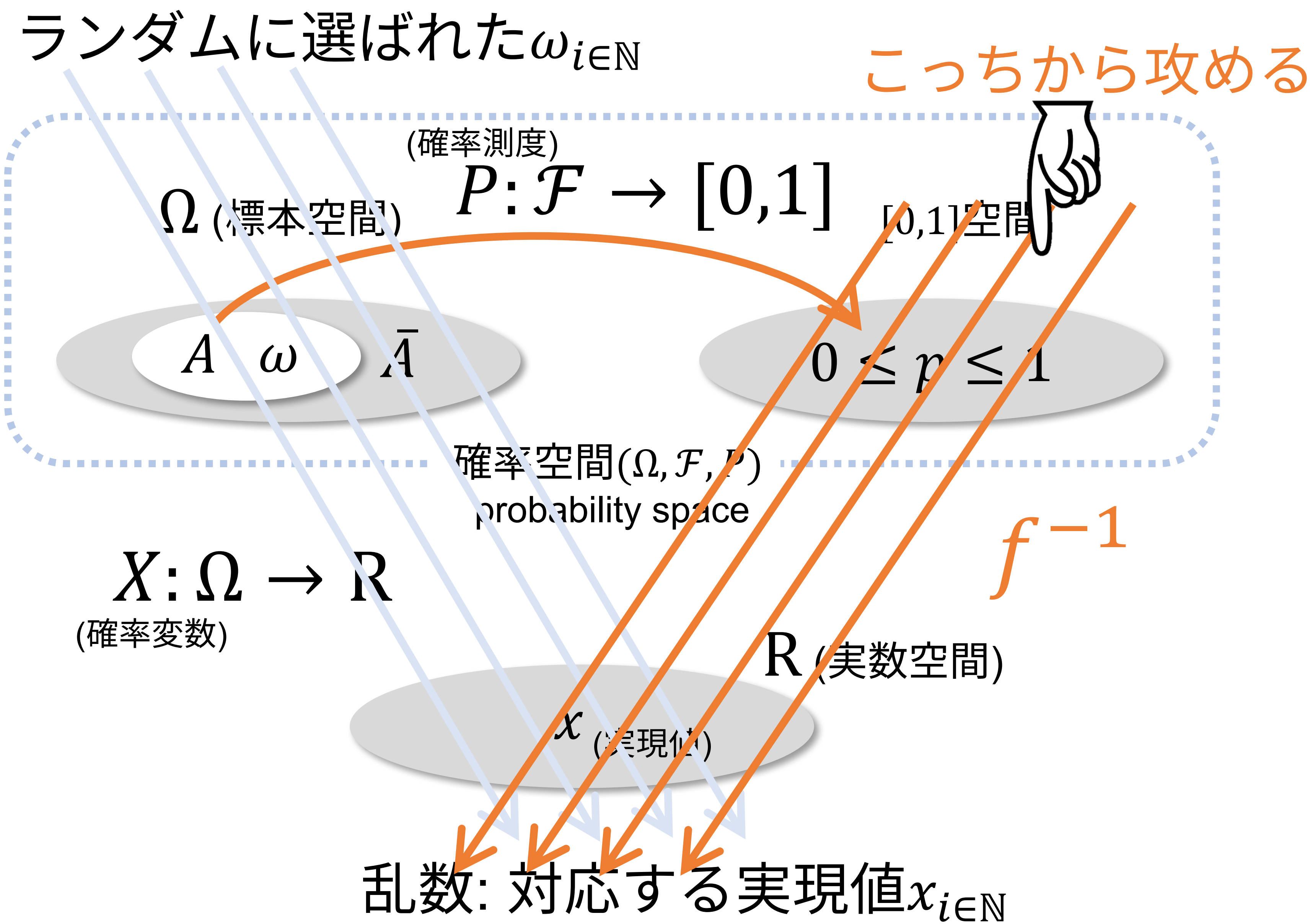
確率密度 $\mathrm{f}(x) = \frac{1}{dx} F(x)$

確率 $p_{|a \leq x \leq b} = \int_{a}^{b} \mathrm{f}(x)dx$



(確率測度)
$P: \mathcal{F} \to [0,1]$
Ω (標本空間)
[0,1]空間
事象A
$\bar{A}$
$0 \le p \le 1$
確率空間$(\Omega, \mathcal{F}, P)$
probability space
$X^{-1}: \mathrm{R} \to \Omega$
$f: \mathrm{R} \to [0,1]$
(確率分布)
x (実現値)
R (実数空間)

正規分布に従う乱数

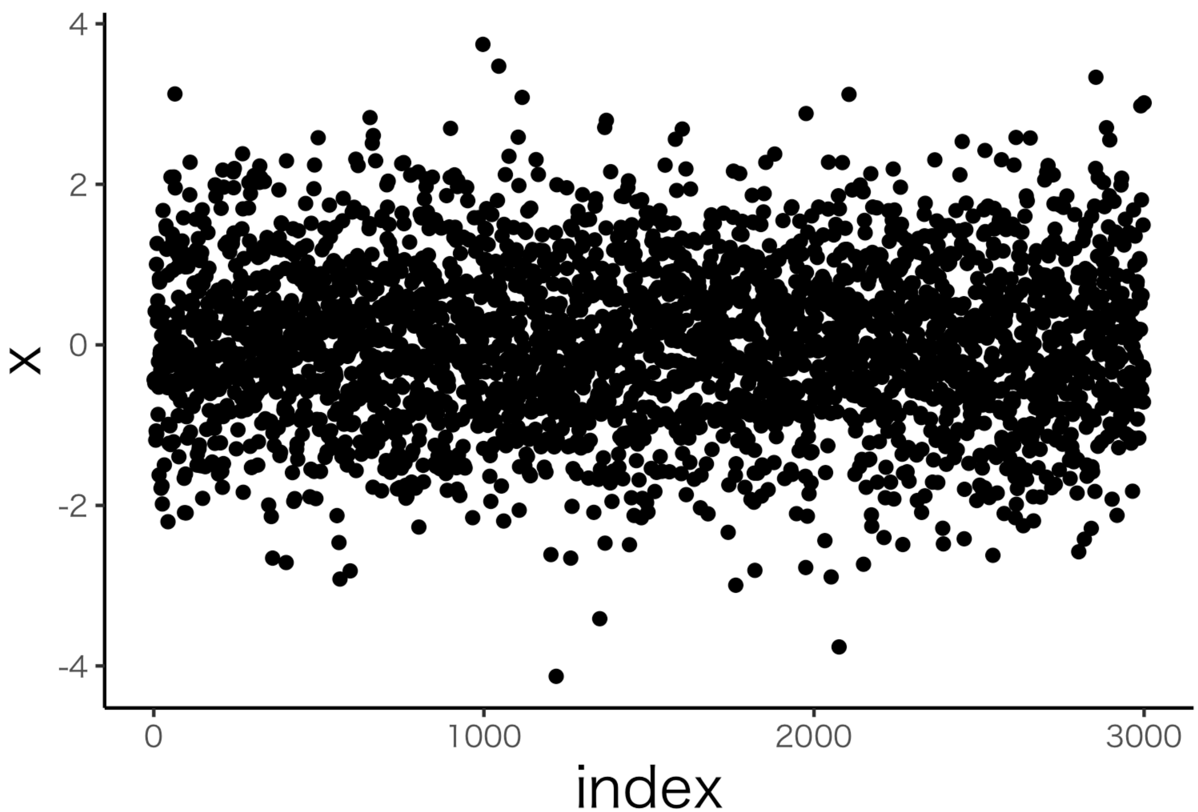


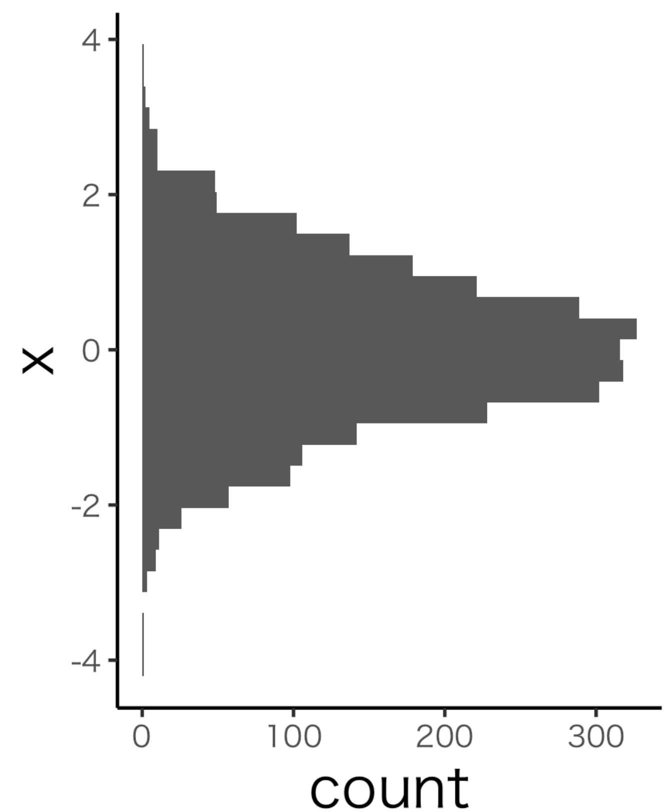
ランダムに選ばれた$\omega_{i\in\mathbb{N}}$
こっちから攻める
(確率測度)
$P: \mathcal{F} \to [0,1]$
Ω (標本空間)
[0,1]空間
A ω $\bar{A}$
$0 \leq p \leq 1$
確率空間$(\Omega, \mathcal{F}, P)$
probability space
f^{-1}
$X: \Omega \to \mathrm{R}$
(確率変数)
R (実数空間)
x (実現値)
乱数: 対応する実現値$x_{i\in\mathbb{N}}$

```
set.seed(71)
dat <- data.frame(x = rnorm(3000)) %>%
  rowid_to_column("index")
```

```
ggplot(dat, aes(index, x))+
  geom_point()
```

```
ggplot(dat, aes(x))+
  geom_histogram()+
  coord_flip()
```





確率分布に関わるRの関数群

済	確率密度関数	d** (第1引数はx座標)
済	累積確率分布	p** (第1引数はx座標)
済	乱数	r** (第1引数は個数)
	確率点	q** (第1引数は割合)

確率分布の略号

正規分布：norm

ポアソン分布：pois

一様分布：unif

対数正規分布：lnorm

ベータ分布：beta

などなど

【今日の目標】

確率変数Xは正規分布に従う

を完全に理解する。

職業、イケメン。テラモナギ

@teramonagi

$x=X(\omega)$は $\{\omega\in\Omega; X(\omega)=x\}$の略ですね。なのでこれは$\Omega$の部分集合を表すための略記です。$P(x=X(\omega)) = P(\{\omega\in\Omega; X(\omega)=x\})$ということです。なので確率測度は$\Omega$の部分集合を引数にとる関数です！

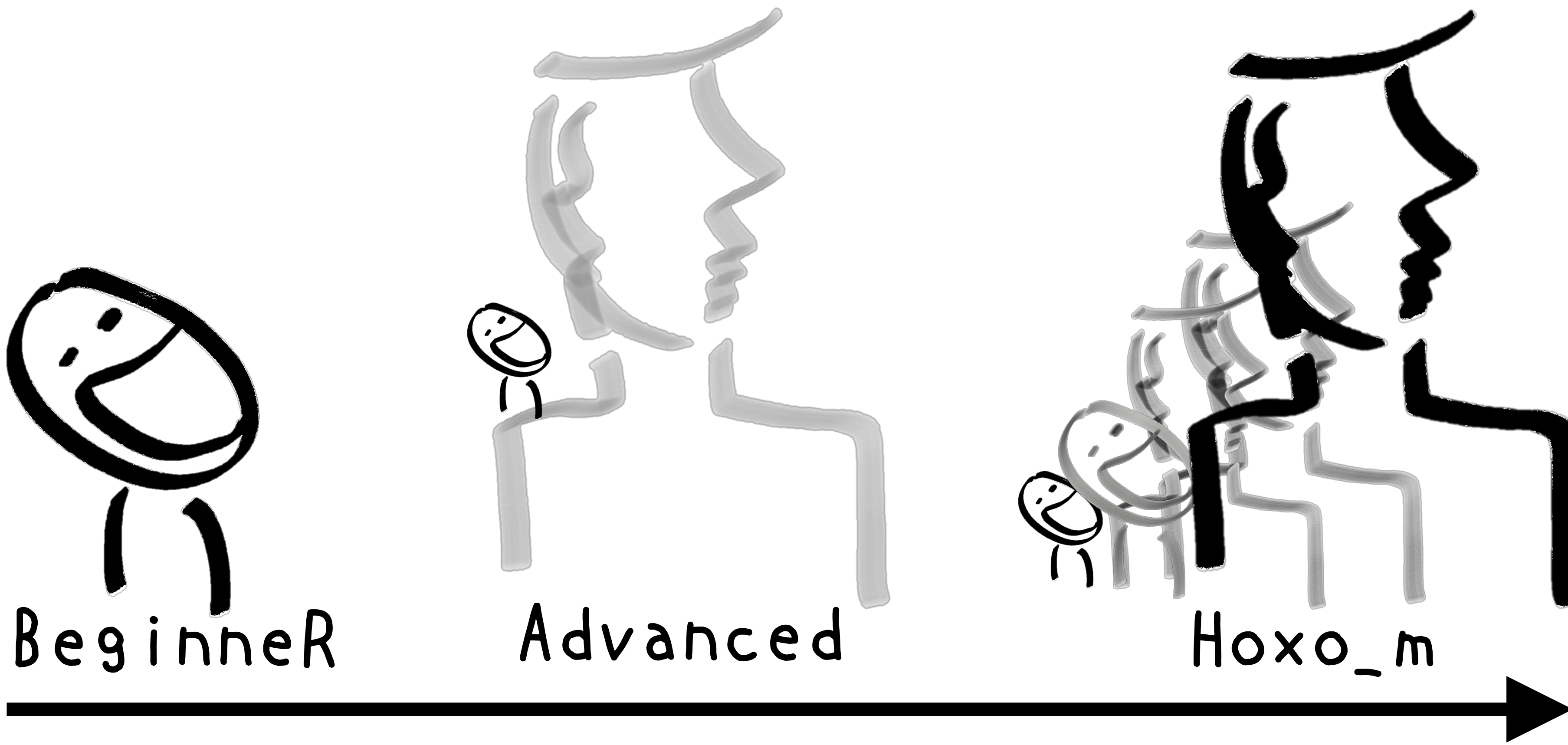
15:26 - 2015年11月17日

もし、わからない事があったら
完全に理解しているヒトにドンドン聞いてみよう！

BeginneR Session

Before

After



If I have seen further it is by standing on the shoulders of Giants.

-- Sir Isaac Newton, 1676

enjoy!


bar dradra
THE PENINSULA

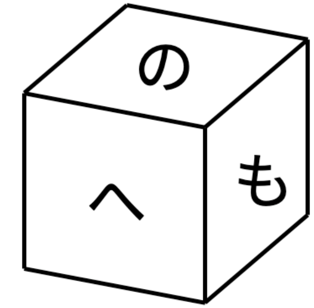
[1] 標本空間Ω上の事象族$\mathcal{F}$から区間[0,1]への写像として確率測度Pを定義することで、Ω上の事象Aに対応する確率pを得る。

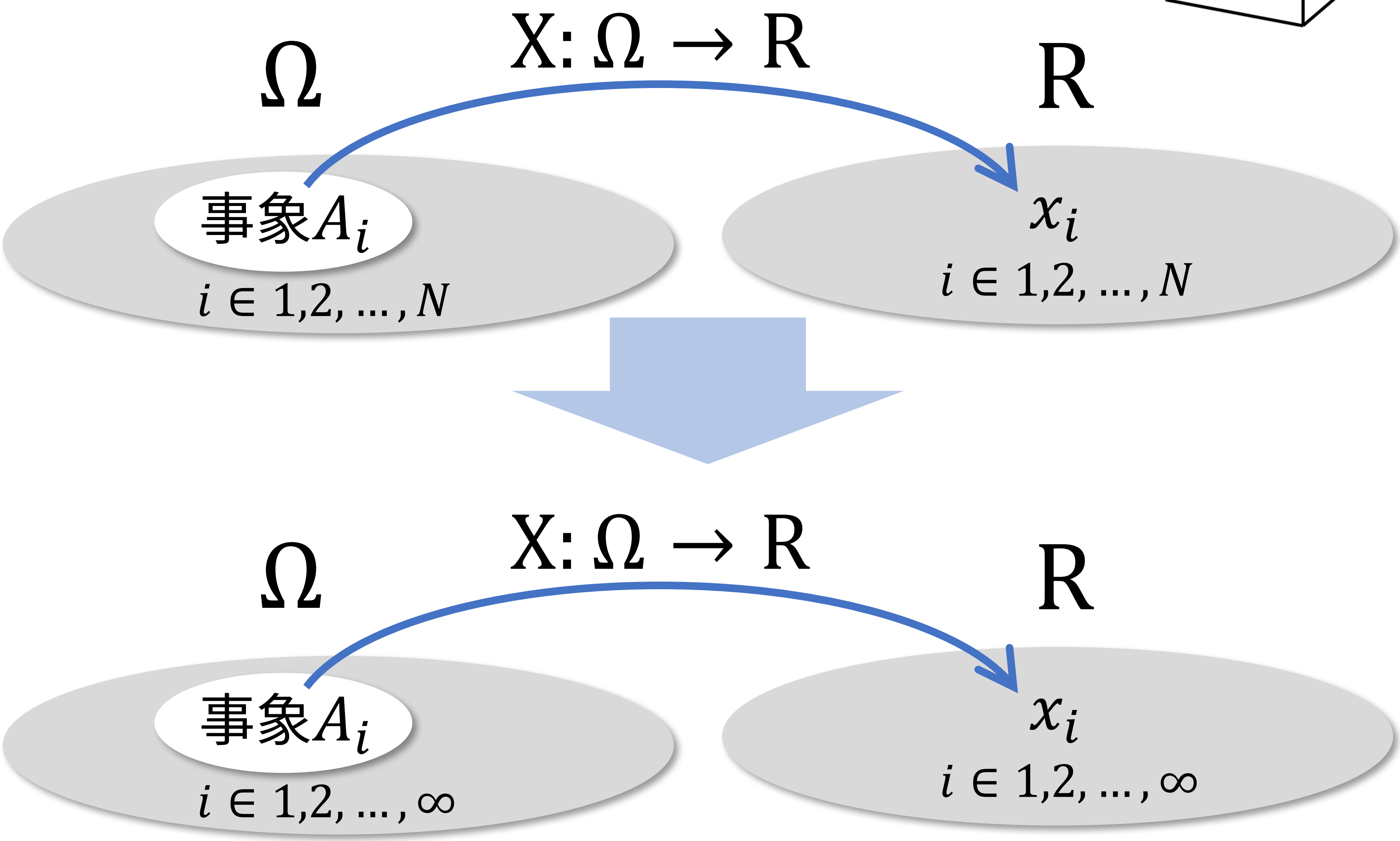
[2] Ωから実数空間Rへの写像を、確率変数Xとして定義することで、その逆によりR上に適切な事象族$\mathcal{R}$を得る。

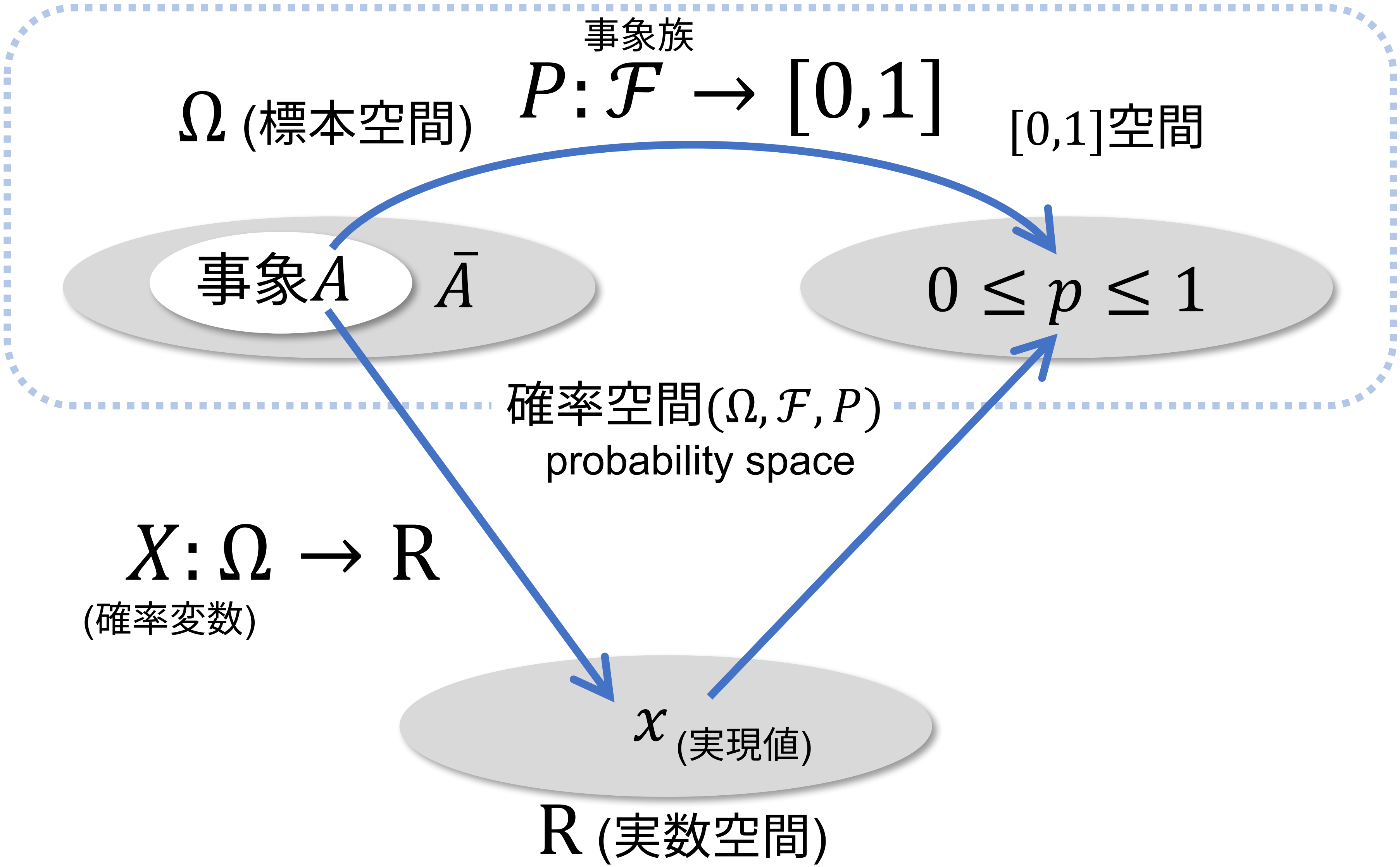
[3] Rから区間[0,1]への写像として、確率分布f(x)を定義することで、空間 $(R, \mathcal{R}, f)$ が確率空間として振る舞える。

[4] これによりXの実現値xに対応するpを、確率空間$(R, \mathcal{R}, f)$上で求める事が可能となる。

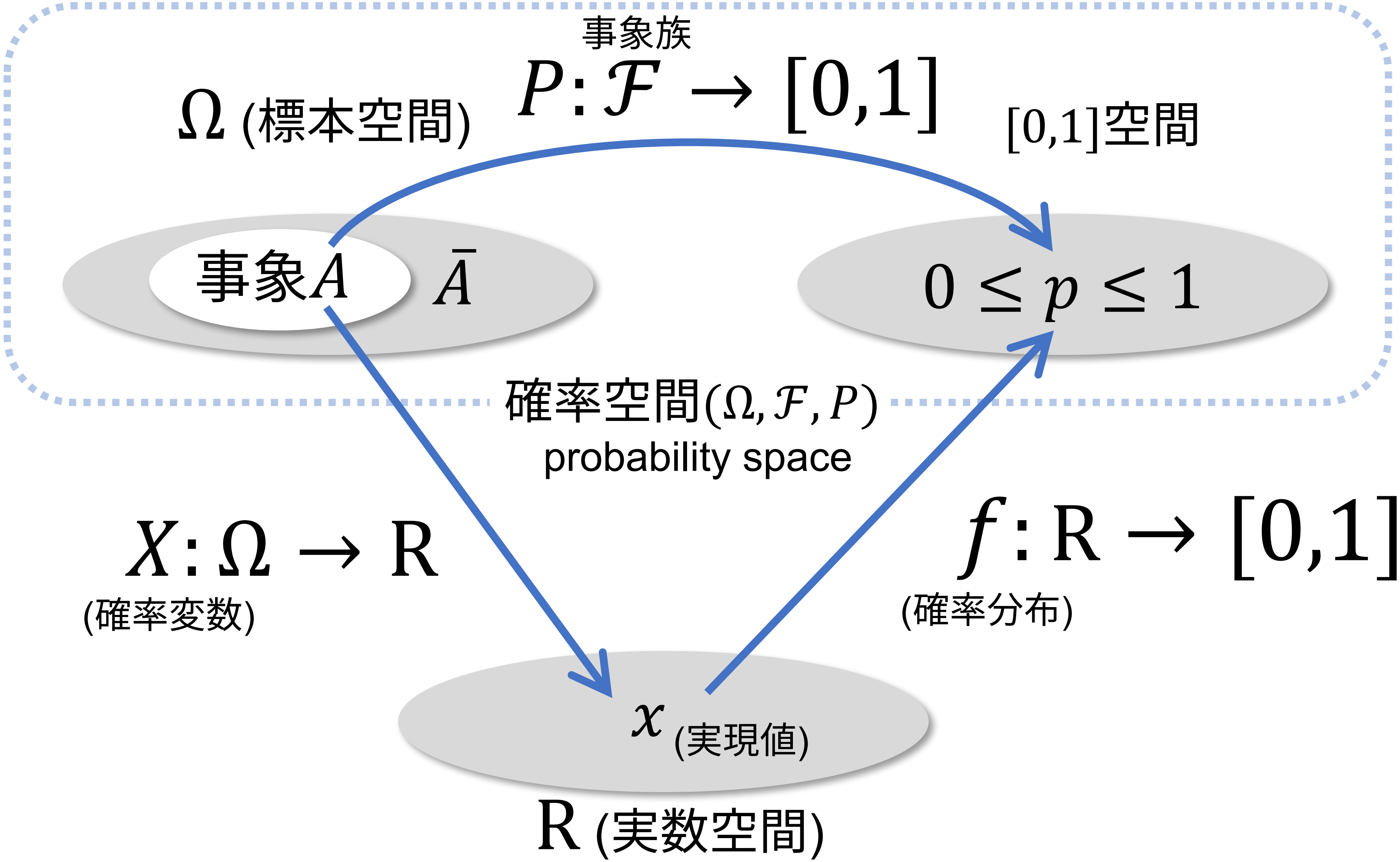
へのへのもへサイコロ


の
へ
も

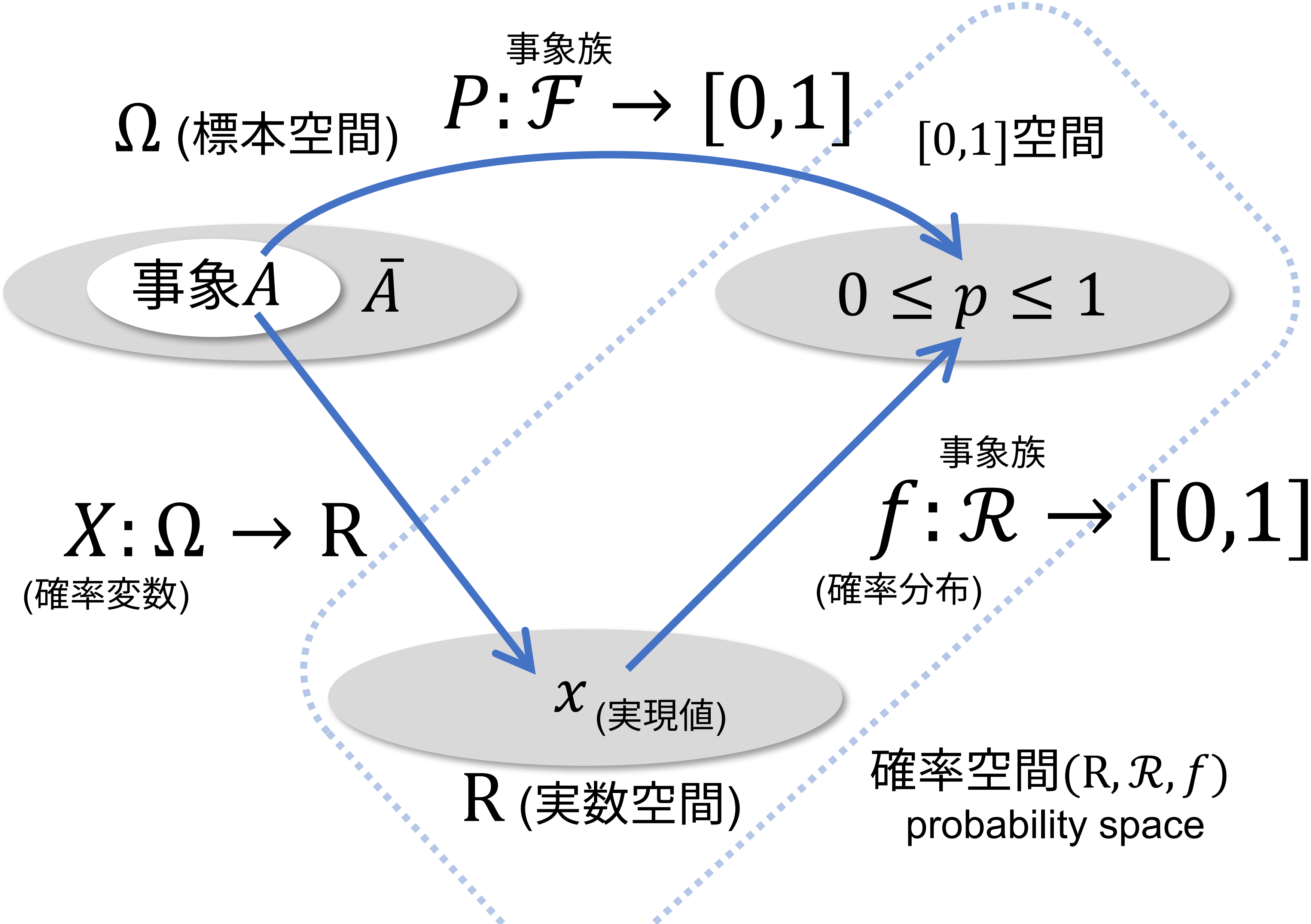

$X: \Omega \to R$
Ω
R
事象A_i
$i \in 1,2,\ldots,N$
x_i
$i \in 1,2,\ldots,N$
$X: \Omega \to R$
Ω
R
事象A_i
$i \in 1,2,\ldots,\infty$
x_i
$i \in 1,2,\ldots,\infty$



事象族
$P: \mathcal{F} \to [0,1]$
Ω (標本空間)
[0,1]空間
事象A
$\bar{A}$
$0 \le p \le 1$
確率空間$(\Omega, \mathcal{F}, P)$
probability space
$X: \Omega \to R$
(確率変数)
x (実現値)
R (実数空間)



$$\overset{\text{確率分布}}{f}(\underset{\text{実現値}}{x}) = P(\underset{\text{事象}}{A} \ni \omega | \overset{\text{確率変数}}{X}(\omega) = x) = \underset{\text{確率}}{p}$$



確率変数Xを考える事で、実現値xが存在する実数空間から[1,0]空間への写像である
確率分布fを定義することができ、$(\mathrm{R},\mathcal{S},f)$ を確率空間として捉えられるようになった。

事象族$\mathcal{F}$の性質より、

$$\mathcal{F} \subseteq \Omega$$

$$A_i \in \mathcal{F} \Rightarrow \bigcup_{i \in \mathbb{N}} A_i \in \mathcal{F}$$

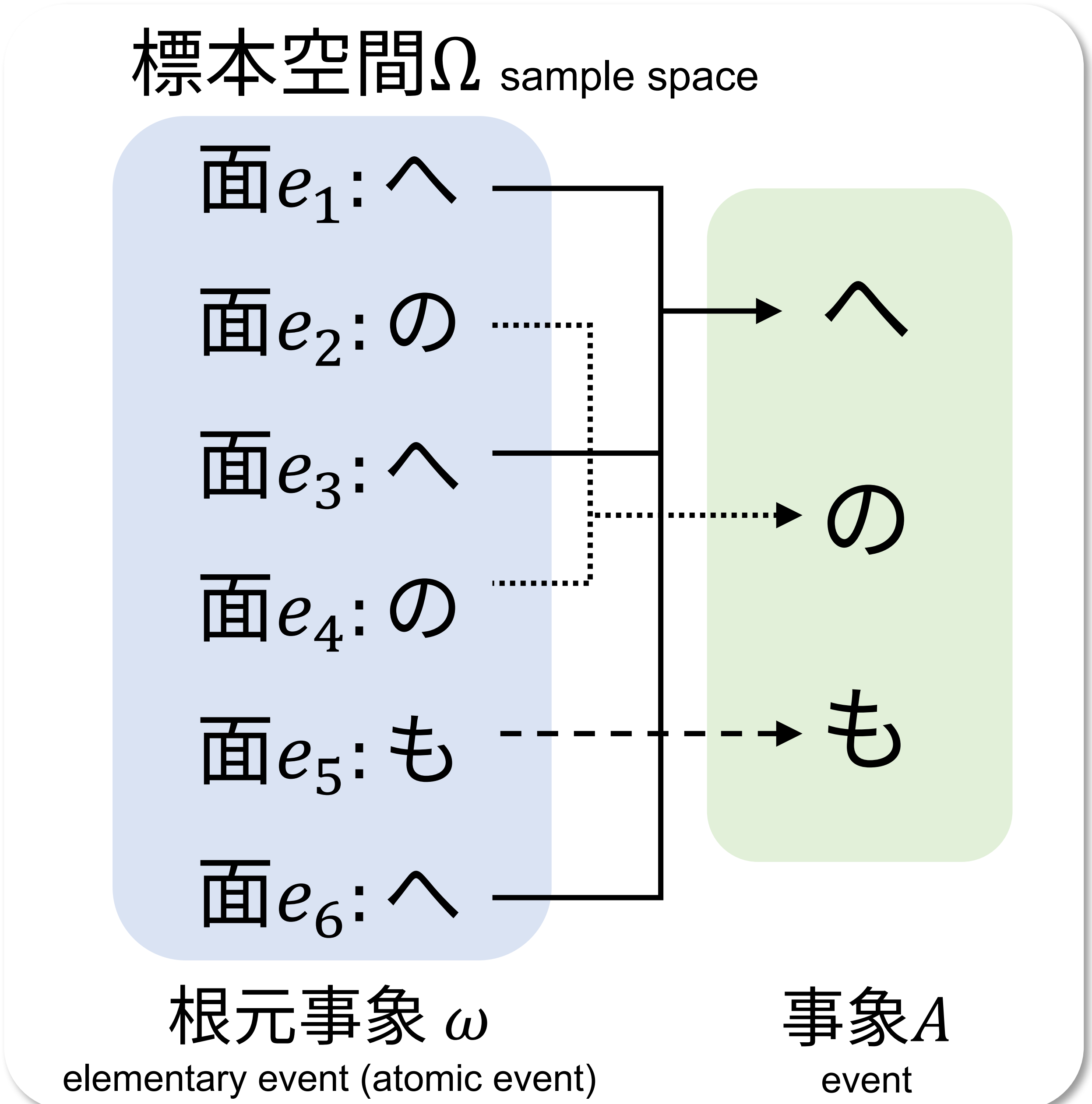
確率測度Pの性質より、

$$P(\Omega) = 1$$

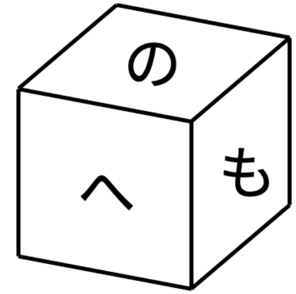
$$\forall \left(A_i \cap A_j \right) = \emptyset$$

$$\Rightarrow \ P(\cup_i A_i) = \sum_i P(A_i)$$

へのへのもへサイコロ

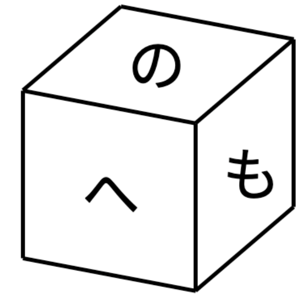


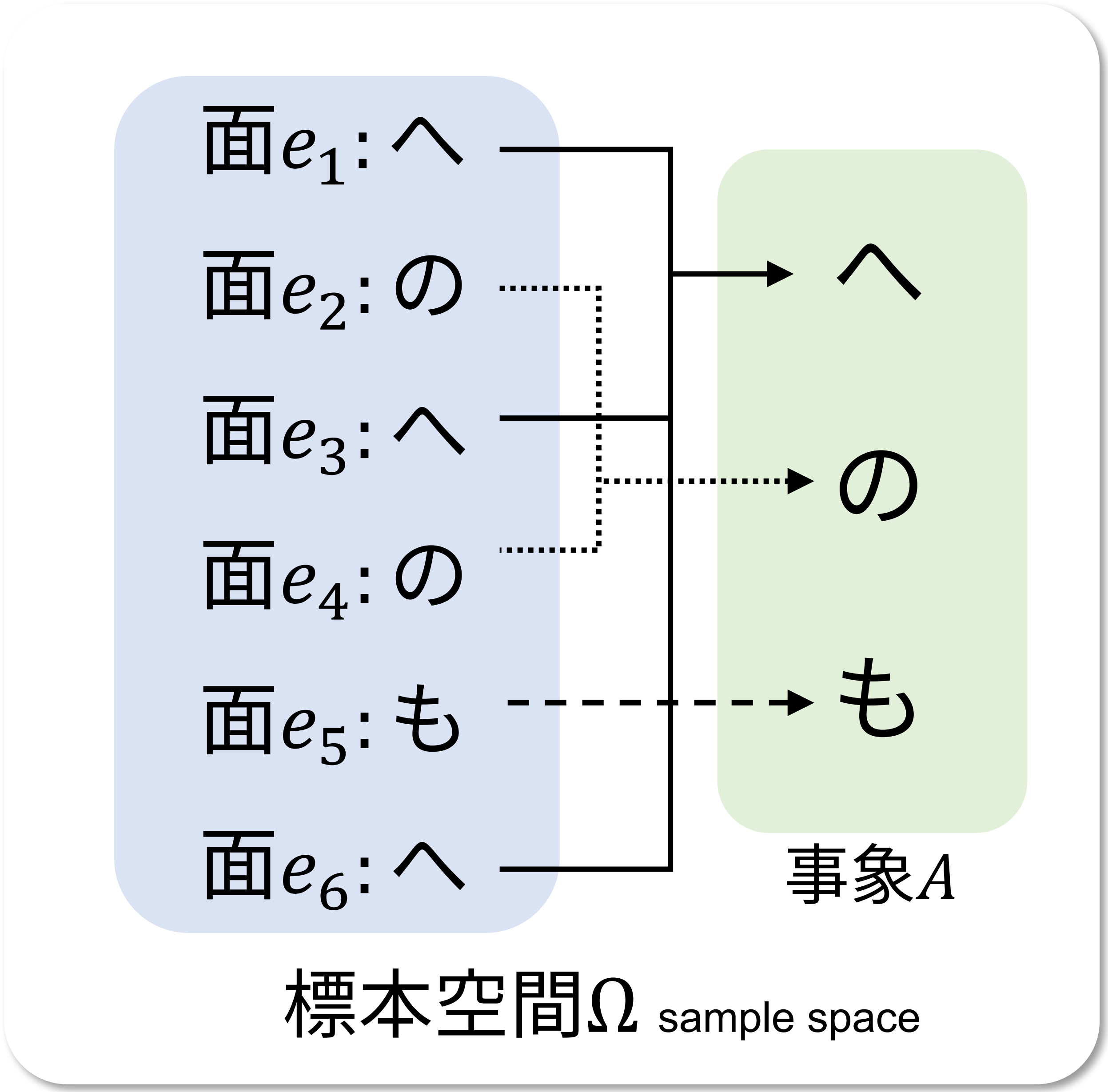
標本空間Ω sample space
面e_1: へ
面e_2: の
面e_3: へ
面e_4: の
面e_5: も
面e_6: へ
へ
の
も
根元事象 ω
elementary event (atomic event)
事象A
event

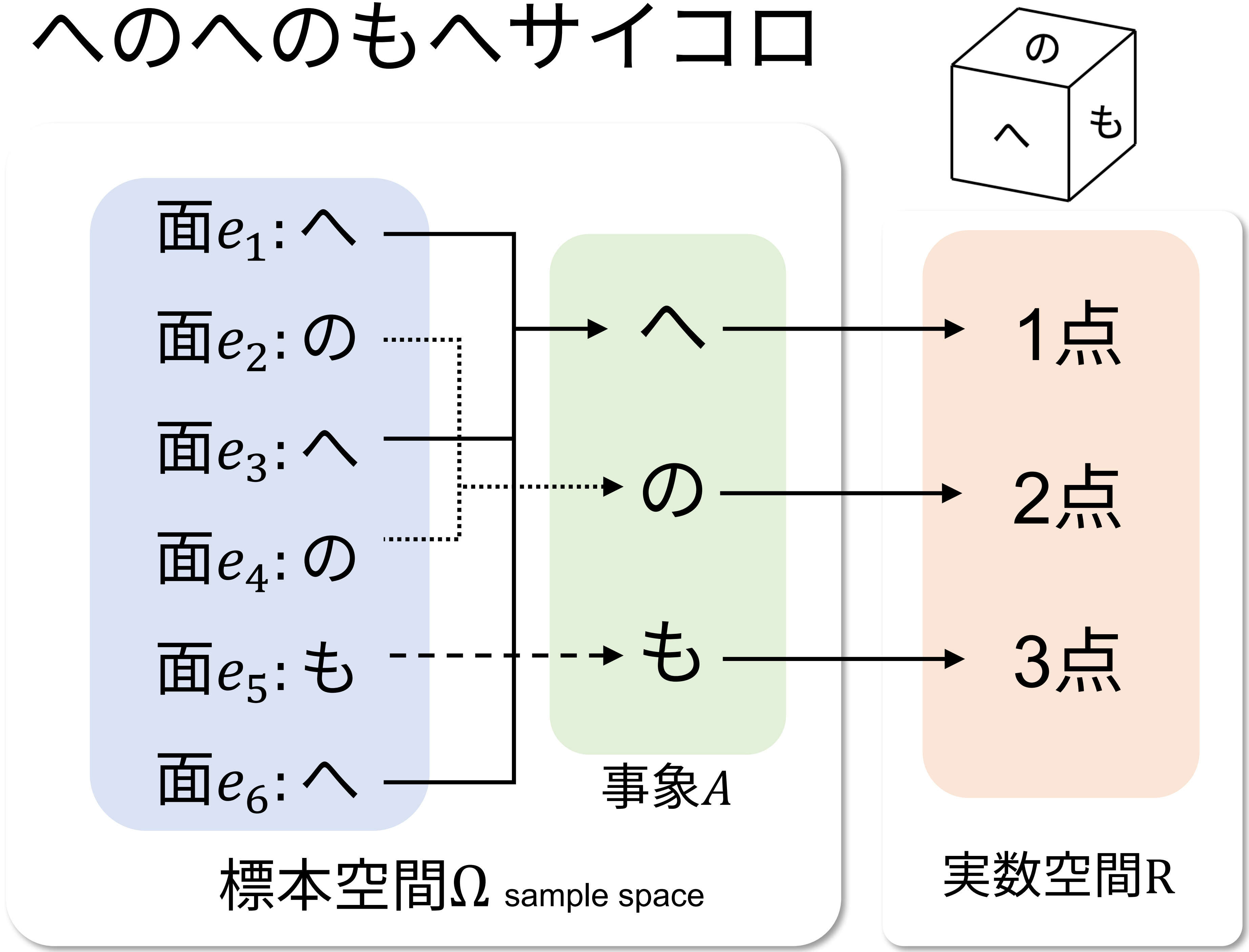


の
へ
も

へのへのもへサイコロ

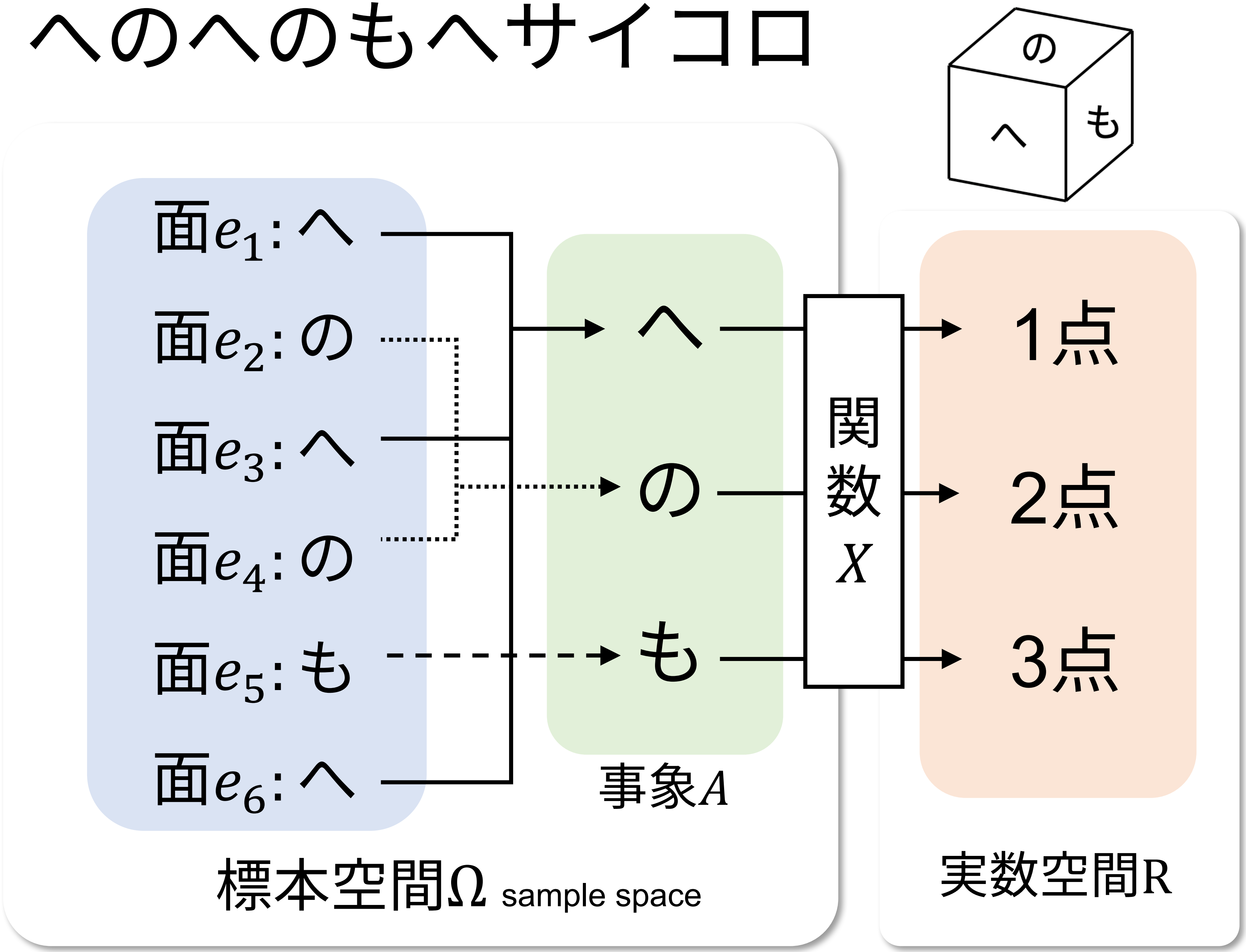






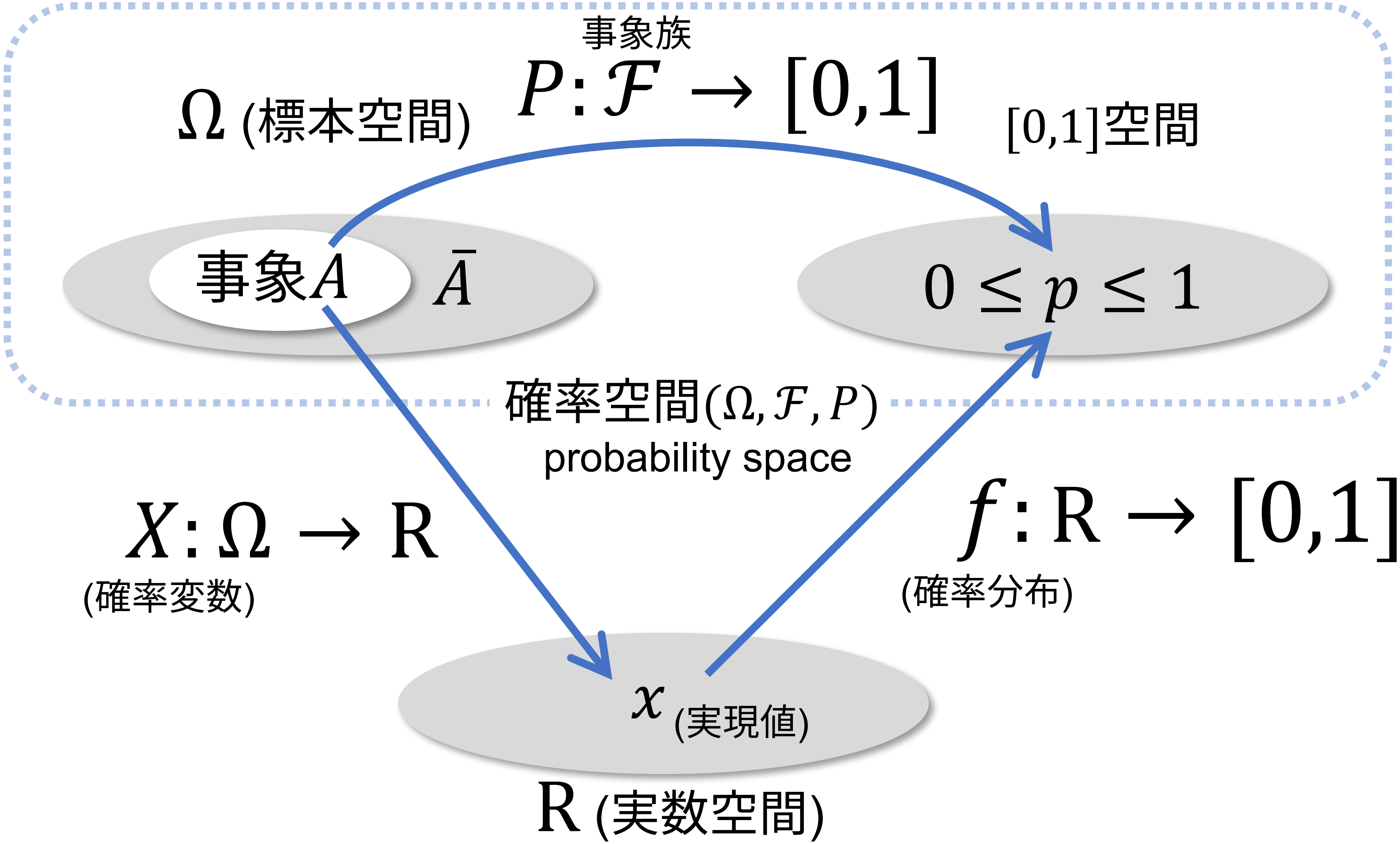
へのへのもへサイコロ
の
へ
も
面e_1: へ
面e_2: の
面e_3: へ
面e_4: の
面e_5: も
面e_6: へ
へ
の
も
事象A
1点
2点
3点
標本空間Ω sample space
実数空間R

へのへのもへサイコロ

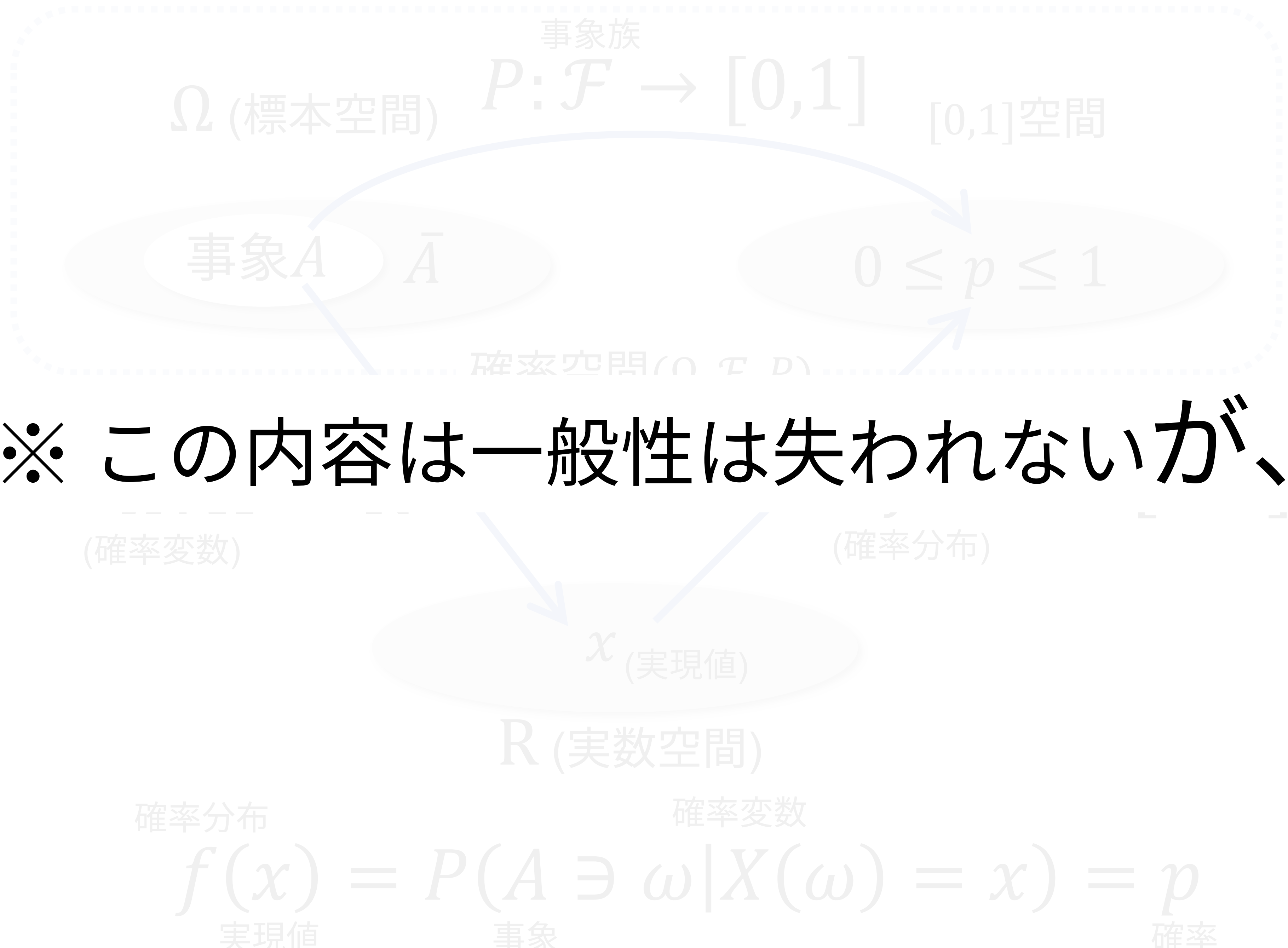


の
へ
も
面e_1: へ
面e_2: の
面e_3: へ
面e_4: の
面e_5: も
面e_6: へ
へ
の
も
事象A
関数X
1点
2点
3点
標本空間Ω sample space
実数空間R

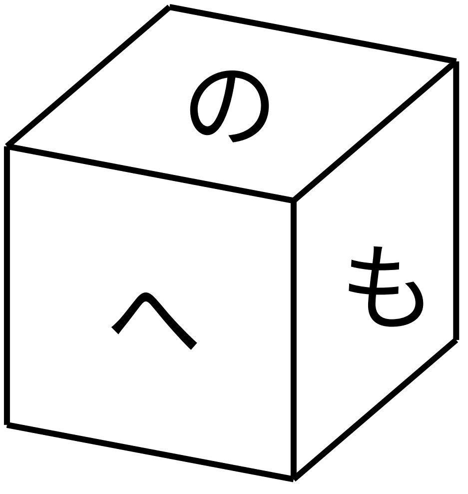
※ この内容は一般性は失われない



$$\overset{\text{確率分布}}{f}(\underset{\text{実現値}}{x}) = P(\underset{\text{事象}}{A} \ni \omega | \overset{\text{確率変数}}{X(\omega)} = x) = \underset{\text{確率}}{p}$$



※ この内容は一般性は失われないが、


の
へ
も