まえがき

○○○○○○○○○○○○○○○○○○○○○○○○○○○○○○○
○○○○○○○○○○○○○○○○○○○○○○○○○○○○○○○
○○○○○○○○○○○○○○○○○○○○○○○○○○○○○○○
○○○○○○○○○○○○○○○○○○○○○○○○○○○○○○○
○○○○○○○○○○○○○○○○○○○○○○○○○○○○○○○
○○○○○○○○○○○○○○○○○○○○○○○○○○○○○○○
○○○○○○○○○○○○○○○○○○○○○○○○○○○○○○○
○○○○○○○○○○○○○○○○○○○○○○○○○○○○○○○
○○○○○○○○○○○○○○○○○○○○○○○○○○○○○○○
○○○○○○○○○○○○○○○○○○○○○○○○○○○○○○○
○○○○○○○○○○○○○○○○○○○○○○○○○○○○○○○
○○○○○○○○○○○○○○○○○○○○○○○○○○○○○○○
○○○○○○○○○○○○○○○○○○○○○○○○○○○○○○○

20XX

○○

○○

目次

★問題の頁数のあとのマス目は，自分の理解の度合いを記入しておくのにご利

Chapter 1

概要

□□□□□□□□□□□□□□□□□□□□□□□□□
□□□□□□□□□□□□□□□□□□□□□□□□□
□□□□□□□□□□

問題 01 ベイズの定理（Bayes' Theorem）

minipage のテスト

有望と思えるガン発見のある検査法が開発された。大病院のガン患者の 97% がこの検査に陽性反応を示し，ガンでない患者の 5% が同じ陽性反応を示したとしよう。病院の患者の 2% がガンにかかっているとするとき，無作為に選んだある患者がこの検査に陽性反応を示したとして，以下の問いに答えよ。

表 1.1 table のテスト

A	B	C
A	B	C
A	B	C

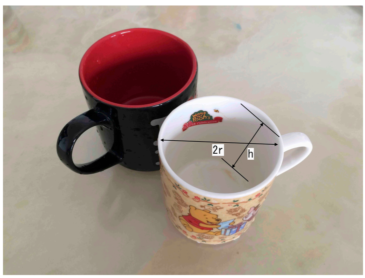

図 1.1 figure のテスト

(1) 抽出された患者が本当にガンにかかっている確率を求めよ。

(2) (1) で得られた結果について吟味せよ。

Ω は標本空間で，$H_1, H_2, \cdots, H_k$ は Ω の部分集合である。標本空
$H_1, H_2, \cdots, H_k$ によって素分割されている。すなわち、次が成り立つ。

(1) $\Omega = H_1 \cup H_2 \cup \cdots \cup H_k$ であり，
(2) $H_1, H_2, \cdots, H_k$ は互いに排反である。すなわち，$H_i \cap H_j = \phi$
ある。

このとき，任意の事象 A と任意 $i = 1, \cdots, k$ に対して，次が成り立つ。

$$P(H_i|A) = \frac{P(H_i)\,P(A|H_i)}{\displaystyle\sum_{j=1}^{k} P(H_j)\,P(A|H_j)}$$

ベイズの定理は次のように証明される。まず, 事象 A は Ω の部分集
集合の吸収則により，$A = A \cap \Omega$ となる。条件により，$\Omega = H_1 \cup H_2$
なので，集合の分配則により，

$$A = (A \cap H_1) \cup (A \cap H_2) \cup \cdots \cup (A \cap H_k)$$

が成り立つ。確率の加法性により，A の確率は

$$P(A) = \sum_{j=1}^{k} P(A \cap H_j) = \sum_{j=1}^{k} P(H_j)\,P(A|H_j)$$

と分解できる。2 番目の等号は乗法定理 $P(A \cap B) = P(B)\,P(A|B)$ に
は**全確率の定理** (law of total probability) として知られる。条件付確
より，

$$P(H_i|A) = \frac{P(A \cap H_i)}{P(A)} = \frac{P(H_i)\,P(A|H_i)}{P(A)}$$

と書ける。上の式の分母に全確率の定理を適用すれば，ベイズの定理
られる。

解答

したがって，全確率は

$$P(A) = P(H_1)\,P(A|H_1) + P(H_2)\,P(A|H_2) = 0.02 \times 0.97 + 0.98 \times 0.0$$

となる。ベイズの定理より

$$P(H_1|A) = \frac{P(A|H_1)\,P(H_1)}{P(A)} = \frac{0.02 \times 0.97}{0.0684} \approx 0.284 = 28.4\%$$

(2)　ベイズの定理における $P(H_i)$ は事前確率（prior probability）と呼ばれ，これ
どに基づいて，病気などにおける確率に関する先験的（prior）知識を表す。新たな
得られたとき，病気になる事後確率（posterior probability）$P(H_i|A)$ を計算
ベイズの定理である。より広い文脈において，H_i を原因，A を結果とすると，ベイ
結果が得られたときの原因の確率 $P(H_i|A)$ を更新するための理論として非常に重
上の例において，$P(H_1|A) = 28.4\%$ がそれほど高くないことに意外と感じ
い。これがむしろ一般的な現象である。もし，$\Omega = H_1 \cup H_2$ で，$P(H_1)$ が
れば，$P(H_2) = 1 - P(H_1) \approx 1$ となる。$P(A) \approx P(A|H_2)$ となることに
$P(H_1|A) \approx 0/P(A|H_2) = 0$ となることが理解できよう。

問題 *02* **正規分布 $N(\mu, 1)$ の 2 乗の分布**

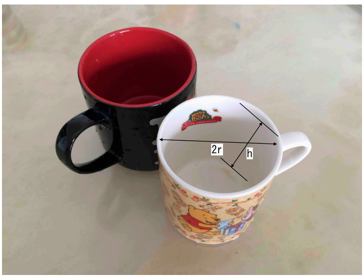

図 **1.2**　wrapfigure のテスト

湖に野生してる蓮の花がそ
れ、時折不規則的に水面に触
面から 25 センチ突き出てい
をじっと見ていたら、53 セン
れた水面に触れることに段々
た。しかし、離れる距離は目測
誤差が伴うことも承知する。
は、平均が 0 センチ, 標準偏
ンチの正規分布に従うとしよ

(1)　水の深さの期待値は何センチか。

(2)　水の深さが 40 センチ以上となる確率はいくらか。

確率変数の関数の分布や積率母関数の求め方を理解することが前提

$$\begin{aligned} F(y) &= \Pr[Y \le y] = \Pr[X^2 \le y] \\ &= \Pr[-\sqrt{y} \le X \le \sqrt{y}] \\ &= \Pr[-\sqrt{y} - \mu \le X - \mu \le \sqrt{y} - \mu] \\ &= 1 - \Phi(-\sqrt{y} - \mu) + \Phi(\sqrt{y} + \mu) \end{aligned}$$

となり、$Y = X^2$ の密度関数 $f(y) = F'(y)$ は次のように書ける。

$$f(y) = \frac{1}{\sqrt{2\pi y}} e^{-\frac{1}{2}(y+\mu^2)} \cosh(\mu\sqrt{y}) \quad (y > 0)$$

ただし、$\cosh x = (e^x + e^{-x})/2$ は双曲線余弦関数である。一方、積率
のように計算できる。

$$\begin{aligned} M(t) &= \mathbb{E}[e^{tY}] = \mathbb{E}[e^{tX^2}] \\ &= \int_{-\infty}^{\infty} e^{tx^2} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x-\mu)^2}\, dx \\ &= \frac{1}{\sqrt{1-2t}} e^{\frac{t\mu^2}{1-2t}} \qquad (t < 1/2) \end{aligned}$$

$M(t)$ を t について微分することにより、Y の 1 次と 2 の積率を求め
きる。

$$\begin{aligned} \mathbb{E}(Y) &= M'(0) = \mu^2 + 1 \\ \mathbb{E}(Y^2) &= M''(0) = \mu^4 + 6\mu^2 + 3 \end{aligned}$$

解 答

(1) 蓮の花の長さは、水面に出ている部分 $a = 25$(cm) と、水の深さ Y(cm) の和か
により、蓮の花が水面に触れるときの距離を W(cm) とすると、

$$W = b + \text{誤差} = b + N(0, \sigma^2) = N(b, \sigma^2)$$

となる。ただし、$b = 53$(cm), $\sigma = 5$(cm) である。三平方の定理

となる。(1.5) により、水の深さ Y の期待値は

$$\mathbb{E}(Y) = \frac{\sigma^2}{2a}\mathbb{E}(X^2) = \frac{\sigma^2}{2a}\left(\frac{b^2}{\sigma^2}+1\right) = \frac{b^2+\sigma^2}{2a}$$

となる。具体的数値を代入すると、$\mathbb{E}(Y) = 56.68(\mathrm{cm})$ となる。測定誤差がなけ
となり、$\mathbb{E}(Y) = b^2/(2a)$ となる。これが「ハスの問題」[1] として知られている。

(2) 次に、$Y \geq c = 40(\mathrm{cm})$ の確率を求める。$\lambda = 2ac/\sigma^2$ とすると、(1.7), (1.3)

$$\begin{aligned}
\mathrm{P}[Y \geq c] &= \mathrm{P}[X^2 \geq \lambda] \\
&= \int_\lambda^\infty \frac{1}{\sqrt{2\pi y}} e^{-\frac{1}{2}(y+\mu^2)} \cosh(\mu\sqrt{y})\, dy \\
&= \int_\lambda^\infty \frac{1}{\sqrt{2\pi y}} e^{-\frac{1}{2}(y+\mu^2)} \frac{e^{\mu\sqrt{y}}+e^{-\mu\sqrt{y}}}{2}\, dy \\
&= 2 - \Phi\left(\sqrt{\lambda}+\mu\right) - \Phi\left(\sqrt{\lambda}-\mu\right) \\
&= 2 - \Phi\left(\frac{\sqrt{2ac}}{\sigma}+\frac{b}{\sigma}\right) - \Phi\left(\frac{\sqrt{2ac}}{\sigma}-\frac{b}{\sigma}\right) \\
&= 2 - \Phi\left(\frac{\sqrt{2ac}+b}{\sigma}\right) - \Phi\left(\frac{\sqrt{2ac}-b}{\sigma}\right)
\end{aligned}$$

ここで、$\Phi\left[(\sqrt{2ac}+b)/\sigma\right] \approx 1$ に注意すると、

$$\mathrm{P}[Y \geq c] \approx 1 - \Phi\left(\frac{\sqrt{2ac}-b}{\sigma}\right)$$

となる。具体的数値を代入すると、

$$\mathrm{P}[Y \geq 40] \approx 1 - \Phi(-1.656) \approx 0.95 = 95\%$$

となる。

Tea Time ‥‥‥ 実験の計画が重要である：円周率の

データを要約したら、データを支配する法則を推量したりする方法論の構築が
命である。しかし、統計的方法論の選択よりも、まず質の高いデータの取得が
ることを肝に銘じてほしい。ここで円周率 π を未知の定数として、π を推定する
て実験計画の重要性を考えよう。

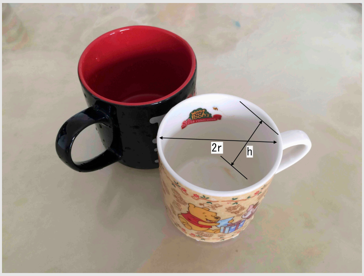

図 1.3　同じ高さのマグカップ

半径 r、高さ h の円柱状のマグカ

$$V_c = 底面積 \times 高さ = (\pi r^2$$

となるので、円周率は

$$\pi = \frac{V_c}{hr^2}$$

となる。しかし、容積 V_c が分から
のように間接的に V_c を測る[2]。

図 1.4 は、横幅 a = 95mm、
110mm、一丁 400g 入れの豆腐容器である。マグカップにある満杯の水を豆
れ、水の深さ X(mm) を測る。豆腐容器内の水の体積は

$$V_t = a \times b \times X \qquad (\mathrm{mm}^3)$$

であり、これがマグカップの容積 V_c の間接的な測定値と見なせる。

上の議論により、円周率の推定値として

$$\hat{\pi} = \frac{V_t}{hr^2} = \frac{ab}{hr^2}X \qquad (1.11)$$

を得る。X 以外を定数として、$\hat{\pi}$ の精度を表す標準偏差は

$$\sqrt{\mathrm{V}(\hat{\pi})} = \sqrt{\frac{a^2b^2}{h^2r^4}\mathrm{V}(X)} \propto 1/r^2 \qquad (1.12)$$

となり、マグカップの半径の 2 乗に応じて減少する。従って、円周率のよい推定量を得るため、

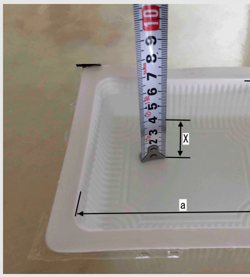

図 1.4　一丁 400g 入れ

できるだけ大口のマグカップを使うべきであることが分かる。表 1.2 は図 1.3 の
を使った実験結果を纏めたものである。大口のカップを使った場合、一桁の精度

TEST shuffle 20

TEST 01 年

$\boxed{1}$ □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

$\boxed{2}$ s □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
□□□□□□□□□□□□□□□□□□□□□□□□

TEST 02

年

$\boxed{1}$ □□□□□□□□□□

$\boxed{2}$ □□□□□□□□□□