確率論 II

2007/07/04, 西岡[1]
http://c-faculty.chuo-u.ac. jp/∼nishioka/

確率論とは, 「偶然事象の法則性」をテーマとする学問である. ただし, どんな偶然事象でも良いわけではなく, 同じような状況が幾度も繰り返された時に起こる偶然事象を対象とする. この場合, 幾度も繰り返すことにより, 何らかの法則性 (極限定理) が現れることが多いからである.

1 確率空間ー復習

「確率論 I」の復習から始める. 近代的な確率論では,

(i) まず適当な確率空間 $(\Omega, \mathbf{P})$ を考え,

(ii) その上で定義された確率変数 $X(w)$ のいろいろな性質を考える

という方法をとる.

1.1 定義

確率空間 $(\Omega, \mathbf{P})$ で Ω は空でない集合であり, **標本空間**[2] sample space と呼ばれ, その元 $w \in \Omega$ を標本 sample という. $\mathbf{P}$ は **確率測度** probability measure である.

[興味を持つ人のために]　厳密に言うと, 確率 (確率測度) $\mathbf{P}$ は,

$$\mathbf{P} : \Omega \text{ の任意の部分集合 } \to \text{ 区間 } [0,1] \text{ の数}$$

という対応をあたえる関数 (集合関数という) で, 次の性質を満たすものである:

$$\begin{aligned} &A \subset \Omega \;\Rightarrow\; 0 \le \mathbf{P}[A] \le 1, \quad \mathbf{P}[\Omega] = 1, \quad \mathbf{P}[\emptyset] = 0, \\ &A_k \subset \Omega, k = 1, 2, \cdots \text{ が } j \ne k \text{ なら } A_j \cap A_k = \emptyset \;\Rightarrow\; \mathbf{P}[\bigcup_{k=1}^{\infty} A_k] = \sum_{k=1}^{\infty} \mathbf{P}[A_k] \end{aligned} \tag{1.1}$$

1.2 確率空間の例

例 1.1. サイコロを 1 回投げる. 起こり得る事象は

$$w_1 \equiv \{1 \text{ の目がでる }\}, \quad w_2 \equiv \{2 \text{ の目がでる }\}, \cdots, w_6 \equiv \{6 \text{ の目がでる }\}$$

の 6 通りであり, 標本空間は $\Omega = \{w_1, \cdots, w_6\}$ である.

また, このサイコロが公平なものとすると, 確率測度 $\mathbf{P}$ は

$$\mathbf{P}[w_1] = \frac{1}{6}, \quad \mathbf{P}[w_2] = \frac{1}{6}, \cdots, \mathbf{P}[w_6] = \frac{1}{6}$$

となる.　◇

[1] nishioka@tamacc.chuo-u.ac.jp　2 号館 11 階 38 号室

[2] しばしば, “ 事象空間 ”とも呼ばれる.

例 1.2. コインを 2 回投げる. 起こり得る事象は

$$w_1 \equiv \{1\text{ 回目}=\text{表}, 2\text{ 回目}=\text{表}\}, \quad w_2 \equiv \{1\text{ 回目}=\text{裏}, 2\text{ 回目}=\text{表}\},$$
$$w_3 \equiv \{1\text{ 回目}=\text{表}, 2\text{ 回目}=\text{裏}\}, \quad w_4 \equiv \{1\text{ 回目}=\text{裏}, 2\text{ 回目}=\text{裏}\},$$

の 4 通りであり, 標本空間は $\Omega = \{w_1, \cdots, w_4\}$ である.
　また

(1.2)　　投げたコインは必ずしも公平ではなく, 表がでる確率が p, $(0 < p < 1)$

なら, 確率測度 $\mathbf{P}$ は

$$\mathbf{P}[w_1] = p^2, \quad \mathbf{P}[w_2] = \mathbf{P}[w_3] = p(1-p), \quad \mathbf{P}[w_4] = (1-p)^2$$

となる.

例 1.3. (i) 今度は (1.2) のコインを n 回投げる. 起こり得る事象は

$$\begin{aligned} w_1 &\equiv \{1\text{ 回目}=\text{表}, 2\text{ 回目}=\text{表}, 3\text{ 回目}=\text{表}, \cdots, n\text{ 回目}=\text{表}\}, \\ w_2 &\equiv \{1\text{ 回目}=\text{裏}, 2\text{ 回目}=\text{表}, 3\text{ 回目}=\text{表}, \cdots, n\text{ 回目}=\text{表}\}, \\ w_3 &\equiv \{1\text{ 回目}=\text{表}, 2\text{ 回目}=\text{裏}, 3\text{ 回目}=\text{表}, \cdots, n\text{ 回目}=\text{表}\}, \\ &\vdots \\ w_N &\equiv \{1\text{ 回目}=\text{裏}, 2\text{ 回目}=\text{裏}, 3\text{ 回目}=\text{裏}, \cdots, n\text{ 回目}=\text{裏}\}, \end{aligned} \tag{1.3}$$

の $N = 2^n$ 通りであり, 標本空間は $\Omega = \{w_1, \cdots, w_N\}$ となる.

(ii) [興味を持つ人のために]　このように極めて多数の元からなる Ω を扱う場合には, (1.3) を次のように書き直した方が分かり易くなる.

(1.4)　**別の表記法**： 1 が表, 0 が裏を表すとして
$w \in \Omega$ にたいし $w = (w^{(1)}, w^{(2)}, \cdots, w^{(2)})$　ここで $w^{(k)} = 1$ もしくは 0.

この表記法 (1.4) を使うと, (1.3) の $w_1, w_2, \cdots, w_n$ は

$$w_1 = (1, 1, \cdots, 1), \quad w_2 = (0, 1, \cdots 1),$$
$$w_3 = (1, 0, 1, \cdots, 1), \ \cdots, \ w_N = (0, 0, \cdots, 0)$$

となる. すると確率測度 $\mathbf{P}$ は

$$\begin{aligned} &w \in \Omega \text{ にたいし } \quad \mathbf{P}[w] = p^{|w|}(1-p)^{n-|w|}, \\ &\text{ここで } |w| \equiv w^{(1)} + w^{(2)} + \cdots + w^{(n)} = \text{表がでた回数} \end{aligned} \tag{1.5}$$

と比較的簡単に定義できる.　◇

2　確率変数–復習

2.1　定義と例

関数 $X(w): w \in \Omega \to \mathbf{R}$ を**確率変数** random variable と呼ぶ.

2 つの確率変数 $X(w)$ と $Y(w)$ で,

$$\begin{aligned} &X(w) \text{ は } x_1, x_2, \cdots, x_m \text{ の値をとり}, \\ &Y(w) \text{ は } y_1, y_2, \cdots, y_n \text{ の値をとる} \end{aligned} \tag{2.1}$$

とする. このとき $x = x_1, x_2, \cdots, x_m$ にたいして決まる関数

$$X(w) \text{ が } x_i \text{ の値をとる確率} = \mathbf{P}[X(w) = x_i], \quad i = 1, 2, \cdots, n$$

を " 確率変数 $X(w)$ の**分布** distribution " とよぶ.

" $X(w)$ と $Y(w)$ が**独立** independent " とは,

$$\begin{aligned} &\text{任意の } 1 \le i \le m,\ 1 \le j \le n \text{ にたいし} \\ &\mathbf{P}[X(w) = x_i,\ Y(w) = y_j] = \mathbf{P}[X(w) = x_i] \cdot \mathbf{P}[Y(w) = y_j] \end{aligned} \tag{2.2}$$

の等式が成立することである.

例 2.1. 例 1.2 の確率空間 $(\Omega, \mathbf{P})$ で 確率変数 $X_k(w), w \in \Omega, k = 1, 2,$ を次のように定義する:

$$X_k(w) \equiv \begin{cases} 1 & k \text{ 回目のコイントスが表} \\ -1 & k \text{ 回目のコイントスが裏} \end{cases} \tag{2.3}$$

すると,

$$\begin{aligned} \mathbf{P}[X_1(w) = 1] = \mathbf{P}[w_1 \cup w_3] &= \mathbf{P}[w_1] + \mathbf{P}[w_3] \\ &= p^2 + p(1-p) = p, \\ \mathbf{P}[X_1(w) = -1] = \mathbf{P}[w_2 \cup w_4] &= \mathbf{P}[w_2] + \mathbf{P}[w_4] \\ &= p(1-p) + (1-p)^2 = 1-p, \\ \mathbf{P}[X_2(w) = 1] = \mathbf{P}[w_1 \cup w_2] &= \mathbf{P}[w_1] + \mathbf{P}[w_2] \\ &= p^2 + p(1-p) = p, \\ \mathbf{P}[X_2(w) = -1] = \mathbf{P}[w_3 \cup w_4] &= \mathbf{P}[w_3] + \mathbf{P}[w_4] \\ &= p(1-p) + (1-p)^2 = 1-p. \quad \diamond \end{aligned}$$

問題 2.2. 上の 例 1.2 (ii) で述べた確率変数 $X_1(w), X_2(w)$ は独立である. 実際に

$$\mathbf{P}[X_1(w) = 1, X_2(w) = 1], \quad \mathbf{P}[X_1(w) = 1, X_2(w) = -1]$$

などを計算し, 独立であることを確かめよ. ♠

解答. 例 1.2 を見ると $\mathbf{P}[X_1(w) = 1, X_2(w) = 1] = \mathbf{P}[w_1] = p^2$. 上の 例 2.1 から $\mathbf{P}[X_1(w) = 1] = p, \mathbf{P}[X_2(w) = 1] = p$ となるので,

$$\mathbf{P}[X_1(w) = 1, X_2(w) = -1] = p^2 = \mathbf{P}[X_1(w) = 1] \cdot \mathbf{P}[X_2(w) = -1]$$

となる. 同様に

$$
\begin{aligned}
\mathbf{P}[X_1(w)=1, X_2(w)=-1] &= p(1-p) \\
&= \mathbf{P}[X_1(w)=1]\cdot\mathbf{P}[X_2(w)=-1] \\
\mathbf{P}[X_1(w)=-1, X_2(w)=1] &= p(1-p) \\
&= \mathbf{P}[X_1(w)=-1]\cdot\mathbf{P}[X_2(w)=1] \\
\mathbf{P}[X_1(w)=-1, X_2(w)=-1] &= (1-p)^2 \\
&= \mathbf{P}[X_1(w)=-1]\cdot\mathbf{P}[X_2(w)=-1]
\end{aligned}
$$

が成立する.

つまり $X_1(w)$ がとる値 $x_1=1, x_2=-1$, $X_2(w)$ のとる値 $y_1=1, y_2=-1$ にたいし

$$\mathbf{P}[X_1(w)=x_i, X_2(w)=y_j] = \mathbf{P}[X_1(w)=x_i]\cdot\mathbf{P}[X_2(w)=y_j], \quad i,j=1,2$$

が成立する事を示したので, $X_1(w)$ と $X_2(w)$ は独立. $\square$

——

例 2.3. つぎに例 1.3 の確率空間 $(\Omega, \mathbf{P})$ で, 確率変数 $Y_k(w), w\in\Omega$, $k=1,\cdots,n$, を

(2.4)
$$Y_k(w) \equiv \begin{cases} 1 & k \text{ 回のコイントスが表} \\ 0 & k \text{ 回のコイントスが裏} \end{cases}$$

と定義する. すると $Y_k(w), k=1,\cdots,n$, は互いに独立で, それらの分布は

(2.5)
$$
\begin{aligned}
&\mathbf{P}[Y_k(w)=1]=p \\
&\mathbf{P}[Y_k(w)=0]=1-p. \quad \diamond
\end{aligned}
$$

例 2.4 (重要)**.** やはり, 例 1.3 の確率空間 $(\Omega, \mathbf{P})$ で, (1.4) の表記法を使い, 新しい確率変数 $Z_n(w)$ を

(2.6)
$$Z_n(w) \equiv |w| = n \text{ 回コインを投げて表が出た回数}$$

とおく. この $Z_n(w)$ の分布は

(2.7)
$$\mathbf{P}[Z_n(w)=k] = {}_nC_k\ p^k\ (1-p)^{n-k}, \quad k=0,1,\cdots,n.$$

((2.7) の右辺は **2 項分布** binomial distribution と呼ばれ, 応用上重要である.) $\diamond$

問題 2.5. ある池に生息している魚の総数を推定するため, つぎの作業を行った:

(i) 魚を 50 匹捕まえ, 尻尾にマークを付けたのち放流する,

(ii) 数日後, 60 匹の魚を捕まえたら, そのうちの 10 匹にマークが付いていた.

この池に生息する魚の総数 N を何匹と推定するか? ♠

解答. 魚の総数を N とする.

(2.8)
$$
\begin{aligned}
& A \equiv \text{数日後に捕まえた 60 匹のなかで 10 匹にマークがあった} \\
& \Leftrightarrow \text{“ 表が出る確率 } p = 50/N \text{ のコイン ” を 60 回投げ, 10 回表が出た}
\end{aligned}
$$

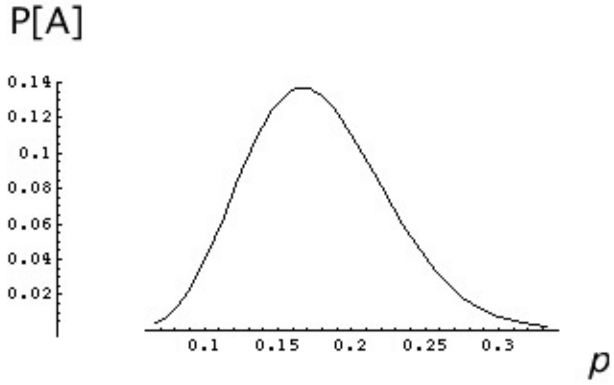
となるので, (2.8) の確率 $\mathbf{P}[A]$ は 2 項分布から計算できる:

$$\mathbf{P}[A] = {}_{60}C_{10}\ p^{10}\ (1-p)^{60-10}.$$

実際に計算を行うと

p	6/60	7/60	8/60	9/60	10/60	11/60	12/60	13/60	14/60
$\mathbf{P}[A]$	0.039	0.071	0.105	0.129	0.137	0.129	0.110	0.086	0.061

これより $p \leq 7/60$ では $\mathbf{P}[A] \leq 0.071$, $p \geq 13/60$ では $\mathbf{P}[A] \leq 0.086$ となり, (2.8) は "起こる確率が低い事が起きた" ことになる. 一方 $p = 10/60$ のとき (2.8) が起こる確率は最大で, $N = 50/p = 300$ 匹と推定できる.



さらに "$\mathbf{P}[A] \geq 0.10$ 以上の事だけが起こる" とすると, $8/60 \leq p \leq 12/60$ の場合が該当するので,

$$375 \text{ 匹} = 50 \cdot \frac{60}{8} \geq N \geq 50 \cdot \frac{60}{12} = 250 \text{ 匹}$$

と幅を持たせた推定となる. □

2.2 平均と分散

確率変数を特徴づける数値として, **平均**と **分散** が重要である.

定義 2.6. (i) $x_1, x_2, \cdots, x_m$ の値をとる確率変数 $X(w)$ にたいして,

$$\mathbf{E}[X(w)] \equiv \sum_{k=1}^{m} x_k\ \mathbf{P}[X(w) = x_k]\ \Big(= \mu \text{ とおく}\Big)$$

を "$X(w)$ の**平均**[3] mean" と呼ぶ. 一方,

$$\mathbf{V}[X] \equiv \mathbf{E}[\big(X(w) - \mu\big)^2] = \sum_{k=1}^{m} \big(x_k - \mu\big)^2\ \mathbf{P}[X(w) = x_k]$$

を "$X(w)$ の**分散** variance" と呼ぶ. ◇

[3] しばしば, **期待値** expectation とも呼ばれる.

例 2.7. 次の 2 種類の確率変数 $X(w)$ と $Y(w)$ を考えてみよう.

$$\mathbf{P}[X(w)=1]=\frac{1}{2}=\mathbf{P}[X(w)=-1],$$
$$\mathbf{P}[Y(w)=100]=\frac{1}{2}=\mathbf{P}[Y(w)=-100].$$

どちらの確率変数も

$$\mathbf{E}[X(w)]=1\cdot\frac{1}{2}-1\cdot\frac{1}{2}=0,\quad \mathbf{E}[Y(w)]=100\cdot\frac{1}{2}-100\cdot\frac{1}{2}=0,$$

と平均は 0 である. ところが分散を比べると

$$\mathbf{V}[X]=\mathbf{E}[\big(X(w)\big)^2]=1^2\cdot\frac{1}{2}+(-1)^2\cdot\frac{1}{2}=1$$
$$\mathbf{V}[Y]=\mathbf{E}[\big(Y(w)\big)^2]=(100)^2\cdot\frac{1}{2}+(-100)^2\cdot\frac{1}{2}=10000$$

となり大きく異なる. $X(w)$ は出入り 1 円, $Y(w)$ は出入り 100 円の公平な賭と考えると, “分散が大きい賭は, ハイリスク・ハイリターン” と解釈できる[4]. $\diamond$

一般に, 平均や分散の計算は易しくない. その計算を少しでも容易にするために, 次の 2 つの補題がある.

補題 2.8 (重要). $X(w), Y(w)$ を確率変数, a, b を定数とする.

(i) $\mathbf{E}\big[a\,X(w)+b\,Y(w)]=a\,\mathbf{E}[X(w)]+b\,\mathbf{E}[Y(w)]$.

(ii) 定数 a にたいしては, $\mathbf{E}[a]=a$.

(iii) $X(w)$ と $Y(w)$ が独立なら $\mathbf{E}[X(w)\cdot Y(w)]=\mathbf{E}[X(w)]\cdot\mathbf{E}[Y(w)]$. $\diamond$

補題 2.9 (重要). $X(w), Y(w)$ を確率変数, a, b を定数とする.

(i) $\mathbf{V}[aX(w)+b]=a^2\ \mathbf{V}[X(w)]$.

(ii) $\mathbf{V}[a]=0$. 逆に分散 $\mathbf{V}[X(w)]=0$ となる 確率変数 $X(w)$ は定数である.

(iii) $\mathbf{V}[X(w)]=\mathbf{E}[\big(\,X(w)\,\big)^2]-\big(\,\mathbf{E}[X(w)]\,\big)^2$.

(iv) $X(w)$ と $Y(w)$ が独立なら

$$\mathbf{V}[X(w)+Y(w)]=\mathbf{V}[X(w)]+\mathbf{V}[Y(w)].\quad\diamond$$

問題 2.10. 会社 A , B の株式収益率[5]をそれぞれ $X(w), Y(w)$ とする. $X(w), Y(w)$ は独立な確率変数で

$$\mathbf{E}[X(w)]=a,\quad \mathbf{E}[Y(w)]=a,\qquad \mathbf{V}[X(w)]=c^2,\quad \mathbf{V}[Y(w)]=2c^2$$

である. 資金 K を A と B に $t, 1-t$ と分配するとき, t をどう決めればリスクは最小になるか?

[4] 株式市場では, 分散は “ボラタリティー” と呼ばれ, 相場変動の大きさをボラタリティーの大きさで判定している.
[5] “収益率= (期末の株価 − 期首の株価)/期首の株価” で, 株の相対的な利益を表す.

解答. $X(w)$ と $Y(w)$ は独立なので, 補題 2.9 より

$$\begin{aligned}\mathbf{V}[K\,t\,X(w) + K\,(1-t)\,Y(w)] &= K^2\,t^2\,\mathbf{V}[X(w)] + K^2\,(1-t)^2\,\mathbf{V}[Y(w)]\\ &= K^2\big\{t^2\,c^2 + (1-t)^2\,2c^2\big\} = K^2\,c^2\big(3t^2 - 4t + 2\big)\\ &= K^2\,c^2\big\{3(t - \frac{2}{3})^2 + \frac{2}{3}\big\}.\end{aligned}$$

これより $\mathbf{V}[K\,t\,X(w) + K\,(1-t)\,Y(w)]$ は $t = 2/3$ のとき最小になるので, A 株に $\dfrac{2}{3}\,K$, B 株に $\dfrac{1}{3}\,K$ 投資するとき, リスクが最小になる. □

成功と失敗の 2 つの結果しかない試行を **ベルヌイ試行** Bernoulli trial と呼ぶ.

問題 2.11. 成功確率 p のベルヌイ試行を成功するまで繰り返す. “成功するまでに要する試行の回数 $T(w)$” の分布を求めよ. ◇

解答. $q \equiv 1 - p$ とおくと,

$$\mathbf{P}[T(w) = 1] = p, \quad \mathbf{P}[T(w) = 2] = q\,p, \quad \mathbf{P}[T(w) = 3] = q^2\,p, \cdots$$

となるので,

$$\mathbf{P}[T(w) = k] = q^{k-1}\,p, \quad k = 1, 2, \cdots \tag{2.9}$$

ここで,

$$W(p, \ell) \equiv \sum_{k=1}^{\ell} \mathbf{P}[T(w) = k] = \sum_{k=1}^{\ell} q^{k-1} p$$

とおくと, $W(p, \ell)$ は“成功確率 p の試行を繰り返すとき ℓ 回以内に成功する確率”となる. 実際に $W(p, \ell)$ の値を計算してみよう:

p	0.05	0.1	0.2	0.3	0.4	0.5	0.6	0.7
$W(p,3)$	0.14	0.27	0.49	0.66	0.78	0.88	0.93	0.97
$W(p,4)$	0.19	0.34	0.59	0.76	0.87	0.93	0.97	0.99
$W(p,5)$	0.22	0.41	0.67	0.83	0.92	0.97	0.99	~ 1
$W(p,6)$	0.26	0.47	0.74	0.88	0.95	0.98	~ 1	~ 1
$W(p,7)$	0.30	0.52	0.79	0.91	0.97	0.99	~ 1	~ 1

つまり

(i) 3 回試行して一度も成功しない (= 確率 1/2 以上のことが起こった). $\Rightarrow$ 成功確率 20% 以下の試行,

(ii) 7 回試行して一度も成功しない (= 確率 1/2 以上のことが起こった). $\Rightarrow$ 成功確率 10% 以下の試行,

といえるので, “ハードルを下げて成功確率を上げる” 方がよい. ちなみに, 成功確率の低い試行では,

k	10	13	16	19	22	25
$W(0.1,k)$	0.65	0.75	0.81	0.86	0.90	0.93
$W(0.05,k)$	0.40	0.49	0.56	0.62	0.68	0.72

となり, 成功を確信するためには 20 回程度もの試行が必要となる. □

3　極限定理

3.1　ポアッソンの少数法則

例 2.4 で述べた 2 項分布は応用上重要であるが, n が大きい場合, 実際の計算は易しくない. ところが,

試行の回数 n が大きく, 成功する確率 p が小さい

場合には, 2 項分布の計算が容易になる. つまり

(3.1)
[ポアッソンの小数法則での仮定] :

$n \to \infty,\ p \to 0$ であるが, ある定数 $\lambda > 0$ があり, $n\,p \to \lambda$

としよう. 例 2.4 より

$$\mathbf{P}[Z_n(w) = 0] = {}_nC_0\, p^0\,(1-p)^n = (1 - \frac{\lambda}{n})^n \to e^{-\lambda}, \quad n \to \infty. \tag{3.2}$$

注: ネイピア数 $e = 2.718\cdots$ (「基礎数学 I」で講義する) :

$$(1 + 1/n)^n \to e = 2.718\cdots$$
$$(1 + x/n)^n \to e^x.$$

また

$$\begin{aligned}\frac{\mathbf{P}[Z_n(w) = k+1]}{\mathbf{P}[Z_n(w) = k]} &= \frac{n!}{(k+1)!\,(n-k-1)!} \cdot \frac{k!\,(n-k)!}{n!} \cdot \frac{p}{1-p} \\ &= \frac{(n-k)\,p}{(k+1)\,(1-p)} = \frac{(1-k/n)\,\lambda}{(k+1)(1-\lambda/n)} \to \frac{\lambda}{k+1}, \qquad n \to \infty\end{aligned}$$

だから $\lim_{n\to\infty} \mathbf{P}[Z_n(w) = k+1] = \frac{\lambda}{k+1} \lim_{n\to\infty} \mathbf{P}[Z_n(w) = k+1]$ となる. よって (3.2) から次々と計算が進行する:

$$\begin{aligned}\lim_{n\to\infty} \mathbf{P}[Z_n(w) = 1] &= \frac{\lambda}{1} \cdot e^{-\lambda} = \lambda\, e^{-\lambda} \\ \lim_{n\to\infty} \mathbf{P}[Z_n(w) = 2] &= \frac{\lambda}{2} \cdot \lambda\, e^{-\lambda} = \frac{\lambda^2}{2!}\, e^{-\lambda} \\ &\vdots \\ \lim_{n\to\infty} \mathbf{P}[Z_n(w) = j+1] &= \frac{\lambda}{j+1} \cdot \frac{\lambda^j}{j!}\, e^{-\lambda} = \frac{\lambda^{j+1}}{(j+1)!}\, e^{-\lambda}.\end{aligned}$$

この極限の計算から得られた確率分布をポアッソン分布と呼ぶ:

定義 3.1 (ポアッソン分布)**.** 無限個の値をとる確率変数 $Z(w) : \Omega \to \{0, 1, \cdots\}$ で

$$\mathbf{P}[Z(w) = k] \equiv \frac{\lambda^k}{k!}\, e^{-\lambda}, \quad k = 0, 1, 2, \cdots \tag{3.3}$$

ここで $\lambda > 0$ はある定数 ($0! \equiv 1$ である.)

となるものを **ポアッソン分布** Poisson distribution に従う確率変数とよぶ.

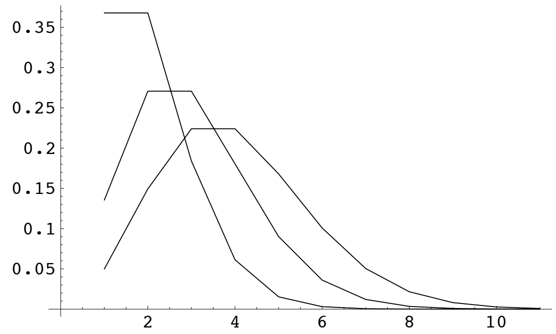


図 3.1: $k=2$ の位置で上から順に $\lambda=1,2,3$ の場合の, $\mathbf{P}[Z(w)=k]$ のグラフ

注意 3.2. ポアッソン分布は 2 項分布の良い近似になっている.
$n=50, p=1/50$, つまり $\lambda=1$ として

$$2\text{ 項分布}\quad Q(k) \equiv= {}_{50}C_k \,\Big(\frac{1}{50}\Big)^k \,\Big(1-\frac{1}{50}\Big)^{50-k}, \quad k=0,1,\cdots$$
$$\text{ポアッソン分布}\quad R(k) \equiv \frac{1}{k!}\,e^{-1}, \quad k=0,1,\cdots$$

の値を比較してみる.

k	0	1	2	3	4	5	6
$Q(k)$	0.364	0.372	0.186	0.061	0.015	0.003	~ 0
$R(k)$	0.368	0.368	0.184	0.061	0.015	0.003	~ 0

命題 3.3. $\mathbf{P}[Z(w)=0]+\mathbf{P}[Z(w)=1]+\cdots+\mathbf{P}[Z(w)=k]+\cdots=1.\quad\diamond$

証明. 指数関数のテイラー展開 を使うと (「解析学」で講義する):

$$\begin{aligned}&\mathbf{P}[Z(w)=0]+\mathbf{P}[Z(w)=1]+\cdots+\mathbf{P}[Z(w)=k]+\cdots\\&=e^{-\lambda}+\lambda\,e^{-\lambda}+\cdots+\frac{\lambda^k}{k!}\,e^{-\lambda}+\cdots\\&=e^{-\lambda}\Big(1+\lambda+\cdots+\frac{\lambda^k}{k!}+\cdots\Big)=e^{-\lambda}\cdot e^{\lambda}=1.\quad\square\end{aligned}$$

命題 3.4 (重要)**.** 確率変数 $Z(w)$ は (3.3) のポアッソン分布に従う. このとき, $Z(w)$ の平均と分散は以下の通り:

$$\mathbf{E}[Z(w)]=\lambda, \qquad \mathbf{V}[Z(w)]=\lambda.\quad\diamond$$

証明 平均を計算する:

$$\begin{aligned}\mathbf{E}[Z(w)]&=\sum_{k=0}^{\infty}k\;\mathbf{P}[Z(w)=k]=\sum_{k=0}^{\infty}k\;\frac{\lambda^k}{k!}\;e^{-\lambda}\\&=\lambda\,e^{-\lambda}\sum_{k=1}^{\infty}\frac{\lambda^{k-1}}{(k-1)!}=\lambda\,e^{-\lambda}\;e^{\lambda}=\lambda.\end{aligned}$$

この計算結果と 命題 3.3 より

$$\begin{aligned}\mathbf{E}[\big(Z(w)\big)^2] &= \sum_{k=0}^{\infty} k^2\ \mathbf{P}[Z(w)=k] \\ &= \sum_{k=1}^{\infty}\Big(k(k-1)+k\Big)\ \frac{\lambda^k}{k!}\ e^{-\lambda} = \sum_{k=2}^{\infty} k(k-1)\ \frac{\lambda^k}{k!}\ e^{-\lambda} + \lambda \\ &= \sum_{k=2}^{\infty}\lambda^2\ \frac{\lambda^{k-2}}{(k-2)!}\ e^{-\lambda} + \lambda = \lambda^2 \sum_{k=0}^{\infty}\frac{\lambda^k}{k!}\ e^{-\lambda} + \lambda = \lambda^2+\lambda.\end{aligned}$$

ここで 補題 2.9 を使うと,

$$\mathbf{V}[Z(w)] = \mathbf{E}[\big(Z(w)\big)^2] - \Big(\mathbf{E}[Z(w)]\Big)^2 = \lambda^2+\lambda-\lambda^2 = \lambda. \quad \square$$

例 3.5. ポアッソン分布は, 事故発生確率, 死亡確率などに応用されている.
1996 年から 2005 年にかけて総数 104 件の航空機事故が発生し, “ 一ヶ月間で何件の事故があったか ” の統計は以下の通りだった:

(3.4)

一ヶ月間の事故発生件数	0	1	2	3	4	5	
月数	52	44	15	6	3	0	計 120ヶ月

この 表 (3.4) を ポアッソン分布で検証してみよう. 毎月の事故発生件数を平均すると $104/120 \sim 0.867$ となるので, 命題 3.4 より λ の値が決まり, (3.4) の統計データとポアッソン分布による計算値が比較できる:

$$\lambda = 0.867, \qquad R(k) \equiv \frac{\lambda^k}{k!}\ e^{-\lambda}, \quad k=0,1,\cdots$$

一ヶ月間の事故発生件数	0	1	2	3	4	5
(3.4) に基づく確率	0.43	0.37	0.18	0.07	0.04	0
$R(k)$	0.42	0.36	0.16	0.05	0.01	0.002

3.2 大数の法則

数理統計への応用を考えるとき, 極限定理で特に重要なものは, “ 大数の法則 ” law of large numbers と “ 中心極限定理 ” central limit theorem である.

“ 大数の法則 ” のイメージ: 公平なサイコロを n 回振ったとき, 1 の目がでる回数を n_1 としよう. すると, n が十分大きいとき, $n_1 \sim n/6$ となることが経験/直感で知られている.

この直感を数学として正当化することが, 大数の法則である.

定理 3.6 (大数の法則)**.** $X_1(w), X_2(w), \cdots, X_n(w)$ は独立で同じ確率分布に従う確率変数であり, それぞれの平均と分散が存在する:

$$\mathbf{E}[X_1(w)] \equiv m, \quad V(X_1) = \mathbf{E}[\big(X_1-m\big)^2] \equiv \sigma^2. \tag{3.5}$$

このとき, 任意の $\varepsilon > 0$ にたいし,

$$\mathbf{P}[\Big|\frac{X_1(w)+\cdots+X_n(w)}{n} - m\Big| > \varepsilon] \to 0, \quad n\to\infty. \quad \diamondsuit$$

問題 3.7. 公平なサイコロを n 回振る. 確率変数 $X_k(w), k=1,2,\cdots,n,$ を

$$X_k(w) \equiv \begin{cases} 1 & k \text{ 回目に振ったサイコロの目が 1 か 2} \\ 0 & k \text{ 回目に振ったサイコロの目が 3 以上} \end{cases}$$

とする.

(i) このとき $\mathbf{P}[X_k(w)=1]$ を計算せよ.

(ii) n が十分大きいとき,

$$\overline{X_n}(w) \equiv \frac{X_1(w)+\cdots+X_n(w)}{n} \tag{3.6}$$

の値を予測せよ. ♠

注意 3.8. 大数の法則は, 古くから研究されており, 定理 3.6 で述べたことよりもっと強い結果[6]が知られている.

> Kolmogorov の大数の強法則 $X_1(w), X_2(w), \cdots, X_n(w)$ は独立な確率変数で, ある定数 m, c にたいし
>
> $$m = \mathbf{E}[X_1(w)] = \mathbf{E}[X_2(w)] = \cdots = \mathbf{E}[X_n(w)], \quad \sup_k \mathbf{E}[(X_k)^2] \le c$$
>
> が成立している. このとき確率 1 で,
>
> $$\frac{X_1(w)+\cdots+X_n(w)}{n} \to m \quad \big(n\to\infty\big). \quad \diamond$$

3.3 中心極限定理

中心極限定理は, §4 で述べた " 大数の法則 " の精密化であり, 数理統計学の数学的な基礎となる重要な定理である.

3.3.1 正規化

$X_1(w)$ を (3.5) である確率変数とする. " 確率変数 $X_1(w)$ の**正規化** $\widehat{X_1}(w)$ " を

$$\widehat{X_1}(w) \equiv \frac{X_1(w)-\mathbf{E}[X_1(w)]}{\sqrt{V(X_1)}} = \frac{X_1(w)-m}{\sqrt{\sigma^2}} \tag{3.7}$$

で定義する. 正規化された確率変数 $\widehat{X_1}(w)$ は

$$\text{平均} \ \ \mathbf{E}[\widehat{X_1}(w)]=0, \quad \text{分散} \ \ V[\widehat{X_1}]=1$$

である.

問題 3.9. 問題 3.7 で与えた確率変数 $X_k(w), k=1,2,\cdots,n,$ を正規化した確率変数 $\widehat{X_k}(w)$ をもとめよ. ♠

[6] こちらの方が直感に合致する.

3.3.2 中心極限定理

$X_1(w), X_2(w), \cdots, X_n(w)$ を 定理 3.6 で述べた確率変数とする. そこで取り扱った確率変数

$$X_1(w) + \cdots + X_n(w)$$

にたいし

$$S_n(w) \equiv X_1(w) + \cdots + X_n(w), \quad \overline{X_n}(w) \equiv \frac{S_n(w)}{n} \tag{3.8}$$

とおく. 確率変数 $S_n(w)$ の正規化 $\widehat{S_n}(w)$ は

$$\widehat{S_n}(w) \equiv \frac{S_n(w) - n\,m}{\sqrt{n\,\sigma^2}} = \frac{\overline{X_n} - m}{\sqrt{\sigma^2}/\sqrt{n}}, \quad n = 1, 2, \cdots$$

となる.

定理 3.10 (中心極限定理)**.** $\widehat{S_n}(w)$ の確率分布は $n \to \infty$ で **標準正規分布** に収束する. つまり, 任意の実数 $a < b$ にたいし

$$\mathbf{P}[a \le \widehat{S_n}(w) < b] \to \int_a^b g(t)\ dt \quad \text{as } n \to \infty,$$

$$\text{ただし}\quad \textbf{正規分布曲線}\quad g(t) \equiv \frac{1}{\sqrt{2\pi}}\ \exp\{-\frac{t^2}{2}\}. \quad \diamondsuit$$

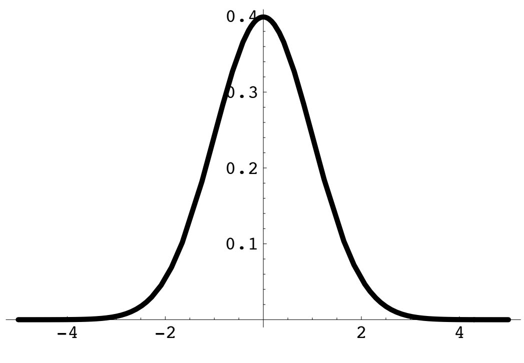


図 3.2: $g(t)$ のグラフ

注意 3.11 (数理統計の基本的な考え方)**.** 中心極限定理を大雑把に述べると, 次の主張に要約できる:

(i) n が大きいとき, (3.8) の $S_n(w)$ の確率はつぎの通り: 任意の $a < b$ にたいし

$$\mathbf{P}[a \le S_n(w) < b] = \int_a^b H(t)\ dt$$

$$\text{ただし}\quad H(t) = \frac{1}{\sqrt{2\pi\,n\,\sigma^2}}\ \exp\{-\frac{(t - n\,m)^2}{2n\,\sigma^2}\}.$$

(ii) n が大きいとき, (3.8) の $\overline{X_n}(w)$ の確率はつぎの通り: 任意の $a < b$ にたいし

$$\mathbf{P}[a \leq \overline{X_n}(w) < b] = \int_a^b h(t)\ dt$$

$$\text{ただし}\quad h(t) = \frac{1}{\sqrt{2\pi\,(\sigma^2/n)}}\ \exp\{-\frac{(t-m)^2}{2(\sigma^2/n)}\}. \quad \diamond$$

例 3.12. 確率変数 $X_k(w)$ を

$$X_k(w) \equiv \begin{cases} 1 & k \text{ 回目にサイコロを振って } 6 \text{ の目が出たとき} \\ 0 & k \text{ 回目にサイコロを振って } 6 \text{ 以外の目が出たとき} \end{cases}$$

とすると

$$\mathbf{P}[X_k(w) = 1] = \frac{1}{6} \quad \mathbf{P}[X_k(w) = 0] = \frac{5}{6}$$

となる. いま $S_{50}(w)$ を“サイコロを 50 回振ったとき 6 の目が出る回数”とおくと

$$S_{50}(w) = X_1(w) + \cdots + X_{50}(w)$$

である. $S_{50}(w)$ の確率分布は直ぐに計算でき,

$$\mathbf{P}[S_{50}(w) = k] = {}_{50}C_k\ \left(\frac{1}{6}\right)^k\ \left(\frac{5}{6}\right)^{50-k}.$$

これより

$$\mathbf{P}[6 \leq S_{50}(w) \leq 10] = \sum_{k=6}^{10} \mathbf{P}[S_{50} = k] \sim 0.66$$

と計算できる. 次に 中心極限定理 3.10 を使って, この確率を計算してみよう. まず

$$m = \mathbf{E}[X_1(w)] = \frac{1}{6}, \quad V(X_1) = \mathbf{E}[\big(X_1(w) - m\big)^2] = \frac{5}{36}$$

だから $S_{50}(w)$ の正規化 $\widehat{S_{50}}(w)$ は

$$\widehat{S_{50}}(w) = \frac{S_{50}(w) - 50 \times (1/6)}{\sqrt{50 \times (5/36)}}.$$

ここで 中心極限定理 3.10 を使って

$$\begin{aligned} &\mathbf{P}[6 \leq S_{50}(w) \leq 10] = \mathbf{P}[5.5 \leq S_{50}(w) \leq 10.5] \\ &= \mathbf{P}[-10.7 \leq \widehat{S_{50}}(w) \leq 0.82] = \int_{-10.7}^{0.82} \frac{1}{\sqrt{2\pi}}\ \exp\{-\frac{t^2}{2}\}\ dt \sim 0.65. \end{aligned}$$

ここで, 最後の計算は手では実行できないから, コンピューターか数表の助けを借りる必要がある. $\diamond$

4 標本分布

4.1 数理統計の目的

性質を知りたい集合[7] がある. その集合は,“数が多すぎて, 全数を調べることが出来ない”場合に, 集合の一部を調査し, その結果から母集団の性質を知ることである.

調査のために取り出した母集団の一部を **標本** sample, 取り出す操作を **標本調査** sampling, 標本の個数を **標本数** size と呼ぶ.

[7] **母集団** population という.

4.2 数理統計の考え方

母集団は無限個の要素からなっており, 母集団の知りたい性質はある確率分布に従っている[8]. この確率分布を **母集団分布** population distribution と呼ぶ.

まず個数 n の標本調査を行うが, 調査する標本 $X_1, \cdots, X_n$ は

母集団分布に従う 独立同分布の確率変数

と仮定する. 標本調査の結果 $X_1, \cdots, X_n$ の値が確定するが, それらは確率変数 $X_1, \cdots, X_n$ の実現値と考える.

例 4.1. 身長の統計調査を例にとる. 国民全体 (N 人) を U とおき, その上の確率 $\mathbf{Q}$ を

$$\mathbf{Q}[u] = \frac{1}{N}, \quad u \in U$$

とする. ここで確率空間 $(U, \mathbf{Q})$ が得られたので,

$$Y(u) = u \text{ 個人の身長}, \quad u \in U$$

として $(U, \mathbf{Q})$ 上の確率変数を定義する. この $Y(u)$ の分布が母集団分布である.

つぎに, U から $u_1, u_2, \cdots, u_n$ の n 人を無作為に選び, 標本調査を行う. その標本を

$$w \equiv (u_1, u_2, \cdots, u_n)$$

として, 次のように新しい確率空間 $(\Omega, \mathcal{F}, \mathbf{P})$ を与える:

$$\begin{aligned} &\Omega \equiv \{w = (u_1, \cdots, u_n) : u_1 \in U, \cdots, u_n \in U\} \\ &\mathbf{P}[w] = \left(\frac{1}{N}\right)^n \end{aligned}$$

$w \in \Omega$ にたいし. 確率変数 $X_k(w)$, $k = 1, \cdots, n$ を

$$X_k(w) \equiv Y(u_k) \ \Big(= u_k \text{ 個人の身長}\Big), \quad w = (u_1, \cdots, u_n),\ u_k \in U$$

と定義すると, $X_1(w), \cdots, X_n(w)$ は母集団分布に従う独立同分布の確率変数となる. $\diamond$

4.3 母数と統計量

母集団分布はどんなものかは判っていない. しかし中心極限定理から示唆されるように, また経験的にも " 母集団分布は 正規分布 $N(m, \sigma^2)$ に従う " と考えてよい場合が多い.

正規分布 $N(m, \sigma^2)$ は m と σ^2 で確定するので, 標本調査から m と σ^2 を決定すればよい. このように, 母集団分布の特徴を表す量を, **母数** parameter とよぶ. とくに 母集団分布の平均と分散は, **母平均** population mean, **母分散** population variance と呼ばれ重要な母数である.

一方, n 個の標本調査 $X_1, \cdots, X_n$ から得られる

$$\overline{X_n} \equiv \frac{1}{n}\Big(X_1 + \cdots + X_n\Big) \tag{4.1}$$

を **標本平均** sample mean,

$$s^2 \equiv \frac{1}{n}\Big((X_1 - \overline{X_n})^2 + \cdots + (X_n - \overline{X_n})^2\Big) \tag{4.2}$$

を **標本分散** sample variance と呼ぶ. このように標本調査から決定される量を **統計量** statistics と呼ぶ.

[8] 例えば, 日本人の身長はある確率分布に従っている.

注意 4.2. 標本平均, 標本分散などの統計量は確率変数である. ◇

5 推定

5.1 推定量の分類

推定とは, 標本調査で得られた統計量から母集団の母数を決めることである. この決めた値を **推定量** estimator と呼ぶ.

しかし, "どの統計量を採用するのが良いか" という決定的な結論はない. そこで, 推定量をその特徴に従って以下のように分類する.

5.1.1 不偏推定量

標本調査 $X_1, \cdots, X_n$ から得られた統計量 $T = T(X_1, \cdots, X_n)$ を母数 θ の推定量とする. この統計量 $T = T(X_1, \cdots, X_n)$ が

$$\mathbf{E}[T] = \theta$$

の性質を持っているとき, "θ の **不偏推定量** unbiased estimator: " と呼ぶ.

例 5.1. (i) (5.4) で与えられる標本平均 $\overline{X_n}$ は不偏推定量である.

(ii) 実数 $a_1, \cdots, a_n$ が

$$a_1 + \cdots + a_n = 1$$

を満たしている. このとき

$$\tilde{X_n} \equiv a_1 X_1 + a_2 X_2 + \cdots + a_n X_n$$

も不偏推定量である.

(iii) (4.2) で与えられる 標本分散 s^2 は不偏推定量ではない. 実際 $V(X_k) = \sigma^2$ でも

$$\mathbf{E}[s^2] = \frac{n-1}{n}\,\sigma^2$$

となる. そこで

$$u^2 \equiv \frac{1}{n-1}\left((X_1 - \overline{X_n})^2 + \cdots + (X_n - \overline{X_n})^2\right) = \frac{n}{n-1} s^2 \tag{5.1}$$

を **不偏標本分散** と定義する. 不偏標本分散は勿論, 不偏推定量である. ◇

5.1.2 一致推定量

母数 θ の推定量 $T = T(X_1, \cdots, X_n)$ が, 任意の $\varepsilon > 0$ にたいし

$$\mathbf{P}\left[\left|T(X_1, \cdots, X_n) - \theta\right| > \varepsilon\right] \to 0, \quad n \to \infty \tag{5.2}$$

を満たすとき **一致推定量** consistent estimator と呼ぶ.

例 5.2. 標本平均 $\overline{X_n}$, (5.4), は母分散の一致推定量である．また 不偏標本分散 u^2, (5.1), は母分散の一致推定量である． $\diamond$

問題 5.3. ある工場で生産されているボルトを 16 個, 無作為に抽出し, その長さ

$$X_1, X_2, \cdots, X_{16} \tag{5.3}$$

を調べた．“ この工場で生産されるボルトの長さは, 正規分布 $N(m,1)$ に従う確率変数 ”と仮定し, 以下の設問に答えよ．

(i) いま, 標本 (5.3) から得られる統計量として

$$T = \frac{1}{n}\Big(X_1 + X_2 + \cdots + X_{16}\Big) \equiv \overline{X_{16}} \tag{5.4}$$

を選んだ (この統計量は, 特に“ 標本平均 ”と呼ばれる)． $\overline{X_{16}}$ は母数 m の不偏推定量であることを説明せよ．

(ii) つぎに, 標本 (5.3) から得られる統計量として

$$T = X_2 \tag{5.5}$$

を選んだ．この統計量も母数 m の不偏推定量となることを説明せよ．

(iii) (5.4) で与えられる 統計量 $T = \overline{X_n}$ は母数 m の一致推定量であることを説明せよ．

(iv) (5.5) で与えられる 統計量 $T = X_2$ は母数 m の一致推定量ではないことを説明せよ． ♠

5.2 区間推定

母集団分布が正規分布 $N(m, \sigma^2)$ に従うことが判っているとする．さらに, 過去のデータが蓄積され, 母分散も大きく変化しないとする．つまり次を仮定する．

仮定 5.4. 母集団分布は 正規分布 $N(m, \sigma^2)$ に従い, その母分散 σ^2 は既知． $\diamond$

n 個の標本 $X_1, \cdots, X_n$ はそれぞれ正規分布 $N(m, \sigma^2)$ に従っている．すると標本平均 $\overline{X_n}$ は正規分布 $N(m, \sigma^2/n)$ に従うので, $\overline{X_n}$ の正規化 Z は

$$Z = \frac{\overline{X_n} - m}{\sqrt{\sigma^2/n}}$$

となり, この Z は正規分布 $N(0,1)$ に従う確率変数である．つまり $a > 0$ にたいし

$$\mathbf{P}[|Z| > G(a)] = a \tag{5.6}$$

となる関数 $G(a)$ が計算できる．(5.6) は

$$\overline{X_n} - G(a)\sqrt{\sigma^2/n} \le m \le \overline{X_n} + G(a)\sqrt{\sigma^2/n}$$

が確率 $1-a$ で成立することを主張している．これを統計学の言葉で言うと

$[\overline{X_n} - G(a)\ \sqrt{\sigma^2/n},\ \ \overline{X_n} + G(a)\ \sqrt{\sigma^2/n}]$ が母平均 m の信頼度 $1-a$ での信頼区間

である.

問題 5.5. 仮定 5.4 が成立しているとし, 前述の 問題 5.3 と同じ設定で以下の設問に答えよ.
(5.3) の標本として

長さ	9.9cm	10 cm	10.3 cm
本数	2 本	8 本	6 本

という結果を得た. この工場で製造されるボルトの平均長 m について, 信頼度 95% の信頼区間を求めよ. ただし,

$$\int_{-1.96}^{1.96} \frac{1}{\sqrt{2\pi}} \exp\{-\frac{t^2}{2}\}\, dt = 0.95$$

とせよ. ♠