0 予備知識

0.1 和・2 項定理

(1) $\displaystyle\sum_{k=1}^{\infty} a_k = \lim_{n\to\infty}\sum_{k=1}^{n} a_k.$

(2) $\displaystyle\sum_{k=1}^{\infty}(ca_k + db_k) = c\sum_{k=1}^{\infty} a_k + d\sum_{k=1}^{\infty} b_k$ ($\sum$ **の線形性**).

(3) $\displaystyle\sum_{k=1}^{n} k = \frac{n(n+1)}{2},\ \ \sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}.$

(4) $\displaystyle\sum_{k=1}^{n} r^{k-1} = \frac{1-r^n}{1-r}.$ $|r|<1$ のとき，$\displaystyle\sum_{k=1}^{\infty} r^{k-1} = \frac{1}{1-r}.$

(5) $n! := n\cdot(n-1)\cdot\cdots\cdot 1,\ 0! := 1,\ {}_nC_k = \dfrac{n!}{(n-k)!k!}.$

(6) $\displaystyle(a+b)^n = \sum_{k=0}^{n} {}_nC_k a^k b^{n-k}$ (**二項定理**).
$(a+b)^2 = a^2 + 2ab + b^2,\ (a+b)^3 = a^3 + 3a^2b + 3ab^2 + b^3, \ldots$

(7) $\displaystyle\left(\sum_{i=1}^{n} a_k\right)\left(\sum_{i=1}^{m} b_k\right) = \sum_{i=1}^{n}\sum_{j=1}^{m} a_i b_j.$

$$(a_1+a_2)(b_1+b_2+b_3) = a_1b_1 + a_1b_2 + a_1b_3 + a_2b_1 + a_2b_2 + a_2b_3.$$

特に，$\displaystyle\left(\sum_{i=1}^{n} a_k\right)^2 = \sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j = \sum_{i=1}^{n} a_i^2 + \sum_{i\neq j} a_i a_j = \sum_{i=1}^{n} a_i^2 + 2\sum_{i<j} a_i a_j.$

$$\begin{aligned}(a_1+a_2+a_3)^2 &= a_1^2 + a_2^2 + a_3^2 + a_1a_2 + a_1a_3 + a_2a_1 + a_2a_3 + a_3a_1 + a_3a_2\\ &= a_1^2 + a_2^2 + a_3^2 + 2(a_1a_2 + a_1a_3 + a_2a_3).\end{aligned}$$

0.2 極限・微分

(1) $\displaystyle\lim_{n\to\infty}\frac{1}{n} = 0,\quad e := \lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n.$

(2) $\displaystyle f'(x) := \lim_{h\to 0}\frac{f(x+h)-f(x)}{h},\ \ \frac{df}{dx}(a) := f'(a),\ \ \frac{d}{dx}f(x) := f'(x).$

(3) $n = 0, 1, 2, \dots$ のとき，$(x^n)' = nx^{n-1}$. 特に定数 c に対して，$(c)' = 0$.

(4) $(e^x)' = e^x$, $(\log x)' = \frac{1}{x}$, $(\sin x)' = \cos x$, $(\cos x)' = -\sin x$.

(5) $\{af(x) + bg(x)\}' = af'(x) + bg'(x)$ (**微分の線形性**).
例. $(3x^2 + 5x - 3)' = 6x + 5$, $(2e^x + 3\log x)' = 2e^x + \frac{3}{x}$,
$(4\sin x + 5\cos x)' = 4\cos x - 5\sin x$

(6) $\{f(x)g(x)\}' = f'(x)g(x) + f(x)g'(x)$ (**積の微分**).

(7) $\{f(g(x))\}' = f'(g(x))g'(x)$ (**合成関数の微分**).

(a) $\{f(ax + b)\}' = af'(ax + b)$.
例. $\{(ax + b)^n\}' = n(ax + b)^{n-1}$, $\{e^{ax+b}\}' = ae^{ax+b}$.

(b) $\{(f(x))^n\}' = n(f(x))^{n-1}f'(x)$.
例.$\{(ae^x + b)^n\}' = n(ae^x + b)^{n-1} \cdot ae^x$.

(c) $\{e^{f(x)}\}' = e^{f(x)}f'(x)$.
例.$\{e^{ae^x+b}\}' = e^{ae^x+b} \cdot ae^x$, $\{e^{ax^2+bx}\}' = (2ax + b)e^{ax^2+bx}$.

(d) $\{\log f(x)\}' = \dfrac{f'(x)}{f(x)}$.
例. $f(x) = x^x$ とすると，$\log f(x) = x\log x$. 両辺微分すると $\frac{f'(x)}{f(x)} = \log x + 1$. これより，$f'(x) = (x^x)' = (\log x + 1)x^x$.

(8) $$\frac{d^m f}{dx^m}(x) := \frac{d}{dx}\left(\frac{d^{m-1}f}{dx^{m-1}}(x)\right), \quad f^{(m)}(x) := \frac{d^m f}{dx^m}(x),\ f^{(0)}(x) := f(x).$$

(9) $$f(x) = f(a) + f'(a)(x - a) + \cdots + \frac{1}{n!}f^{(n)}(a)(x - a)^n + \cdots$$ (**テイラー展開**).

(10) $$e^x = 1 + x + \frac{1}{2}x^2 + \frac{1}{6}x^3 + \cdots = \sum_{k=0}^{\infty}\frac{1}{k!}x^k$$ (**指数関数のテイラー展開**).

(11) 任意の $a > 0$, $b > 0$ に対して，$\lim_{x\to\infty} x^a e^{-bx} = 0$. (**指数関数の収束速度**).

(12) $$\frac{\partial f}{\partial x}(a, b) := \lim_{h\to 0}\frac{f(a + h, b) - f(a, b)}{h}, \quad \frac{\partial f}{\partial y}(a, b) := \lim_{k\to 0}\frac{f(a, b + k) - f(a, b)}{k}.$$
例. $$\frac{\partial(px + y)^2}{\partial x} = 2p(px + y), \quad \frac{\partial(px + y)^2}{\partial y} = 2(px + y), \ \frac{\partial e^{xy}}{\partial x} = ye^{xy}.$$

(13) $$\frac{df(g(t), h(t))}{dt} = \frac{\partial f}{\partial x}(g(t), h(t))g'(t) + \frac{\partial f}{\partial y}(g(t), h(t))h'(t)$$

0.3 積分

(1) $F'(x) = f(x)$ のとき，$\displaystyle\int f(x)dx := F(x) + C$($C$ は任意の定数),

$\displaystyle\int_a^b f(x)dx := F(b) - F(a)$．ここで $[F(x)]_a^b := F(b) - F(a)$ と書くことが多い．

(2) $$\int_a^b x^n dx = \left[\frac{1}{n+1}x^{n+1}\right]_a^b = \frac{1}{n+1}(b^{n+1} - a^{n+1}).$$

(3) $$\int_a^b \{cf(x) + dg(x)\}dx = c\int_a^b f(x)dx + d\int_a^b g(x)dx$$ (**積分の線形性**).

(4) $$\int_a^b f(g(x))dx = \int_{a'}^{b'} f(t)\left|\frac{dx}{dt}\right|dt,\ \ t = g(x),\ a' = g(a),\ b' = g(b)$$ (**置換積分**).

例. $\displaystyle\int_a^b (cx+d)^n dx = \left[\frac{1}{c}\frac{1}{n+1}(cx+d)^{n+1}\right]_a^b,\quad \int_a^b e^{cx}dx = \left[\frac{1}{c}e^{cx}\right]_a^b.$

(5) $$\int_a^b f(x)g'(x)dx = [f(x)g(x)]_a^b - \int_a^b f'(x)g(x)$$ (**部分積分の公式**).

(6) $$\int_a^b x^n e^{cx}dx = \left[\frac{1}{c}x^n e^{cx}\right]_a^b - \int_a^b \frac{n}{c}x^{n-1}e^{cx}dx.$$

特に，$\alpha > 0$ のとき，$\displaystyle\int_0^\infty x^n e^{-\alpha x}dx = \frac{n}{\alpha}\int_0^\infty x^{n-1}e^{-\alpha x}dx$

(7) $$\int\int_A f(x,y)dxdy = \int\int_B f(g(u,v), h(u,v))\mathrm{abs}(J(u,v))dudv.$$

ただし，$J(u,v) = \begin{vmatrix} \frac{\partial g}{\partial u} & \frac{\partial g}{\partial v} \\ \frac{\partial h}{\partial u} & \frac{\partial h}{\partial v} \end{vmatrix}$, $A = \{(x,y) = (g(u,v), h(u,v)) \mid (u,v) \in B\}$.

(a) $$\int\int_{\substack{0\leqq x\leqq 1\\ 0\leqq y\leqq 1}} f(x,y)dxdy = \int\int_{\substack{0\leqq au+bv\leqq 1\\ 0\leqq cu+dv\leqq 1}} f(au+bv, cu+dv)\mathrm{abs}(|C|)dudv$$

ただし，$C = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$.

(b) $$\int\int_{x^2+y^2\leqq 1} f(x^2+y^2)dxdy = 2\pi\int_0^1 uf(u)du$$

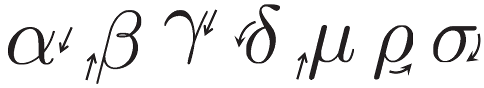
ここで，$x = u\cos v$, $y = u\sin v$ とした．そのとき，$J(u,v) = u$ である．

0.4　ギリシャ文字

数学や統計学では色々なギリシャ文字が利用される．

小	大	読み		小	大	読み	
α	A	アルファ	alpha	ν	N	ニュー	nu
β	B	ベータ	beta	ξ	Ξ	グザイ	xi
γ	Γ	ガンマ	gamma	o	O	オミクロン	omicron
δ	Δ	デルタ	delta	π	Π	パイ	pi
ϵ	E	イプシロン	epsilon	ρ	P	ロー	rho
ζ	Z	ゼータ	zeta	σ	Σ	シグマ	sigma
η	H	イータ	eta	τ	T	タウ	tau
θ	Θ	シータ	theta	υ	Υ	ウプシロン	upsilon
ι	I	イオタ	iota	ϕ	Φ	ファイ	phi
κ	K	カッパ	kappa	χ	X	カイ	chi
λ	Λ	ラムダ	lambda	ψ	Ψ	プサイ	psi
μ	M	ミュー	mu	ω	Ω	オメガ	omega

書き順を間違えると，違う文字に見えるので，正しい筆順で書かなければならない．よく使う文字の筆順は以下の通りである．



0.5　集合

$\mathbb{N}$	自然数の集合 $\{1, 2, 3, \ldots\}$
$\mathbb{Z}$	整数の集合 $\{0, \pm 1, \pm 2, \ldots\}$
$\mathbb{Q}$	有理数の集合
$\mathbb{R}$	実数の集合
$a \in A$, $A \ni a$	a は集合 A の要素である
$A \subset B$, $B \supset A$	集合 A が集合 B に含まれる
$A \cap B$	集合 A と集合 B の交わり (共通集合)
$A \cup B$	集合 A と集合 B の結び (和集合)
A^c	集合 A の補集合
ϕ	空集合

1　統計学とは

1.1　記述統計

統計学とは，対象とする集団の**全体的**特徴や傾向をとらえるために，いかなる方法が有効かを議論する理論体系である．

対象とする集団のすべての情報が得られている時は，全体的な特徴をとらえるために情報全体をいかに集約するかが問題となる．よく知られているように，全体的な大きさを表すには**平均**を求めればよいし，集団の中に平均からかけ離れた個体がどの程度あるのか (バラツキ) を表すには**標準偏差**を計算すればよい．あるいは，身長と体重の平均的関係を表すボディマス指標のように，集団の各個体に付随する 2 種類の特徴量の**平均的な関係**を求めることも必要であろう．

このような集団全体の情報から平均的な特徴を捉える方法を**記述統計**と呼ぶ．記述統計の基本的な手法は比較的よく知られているが，それらの性質や注意点を第 2 節 (p.7) で簡潔に説明する．

1.2　推測統計

実験や調査結果に基づき何らかの結論を下す時，再度実験や調査をしたら異なる結論になるのではないかという疑念が生じることがある．つまり，結論の**信頼性**が時として問題となる．

たとえば，ある農産物に含まれる有害物質が基準値より少ないかどうかを調べるために，生産された作物を適当に抜き出して検査をしたとしよう．この時，抜き取られた作物に含まれる有害物質の平均値が基準値を下回っていたならば，その農産物は安全であると考えていいだろうか？

この検査に関する疑問点として，

(1)　有害物質の少ないものばかりを抜き取ったのではないか？

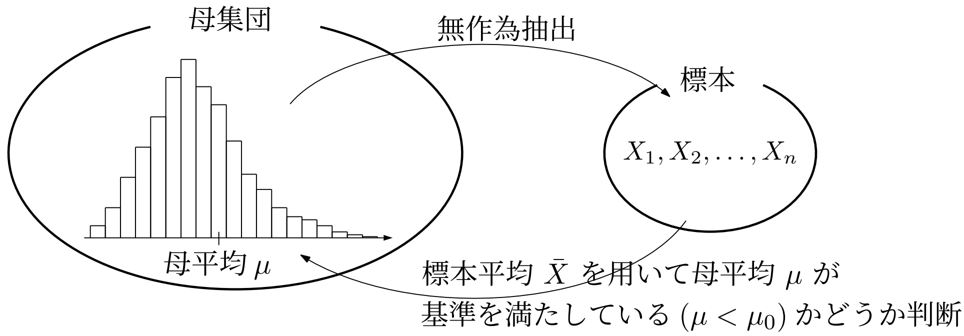
(2)　抜き取りではなくて，全数検査できないのか？

(3)　再度抜き取り検査したら平均値が基準値を上回ることはないのだろうか？

(4)　1 回の抜き取り検査で誰もが納得する判断ができるのだろうか？

などが挙げられよう．(1) に関しては，偏りのない抜き取り方法を考えれば解決できる．たとえば，生産環境が良いものばかりでなく，検査する側の恣意が入らないように，乱数などを使って無作為に抜き取ればよい．(2) に関しては，安全性を最優先するなら全数調査すべきであるが，農産物を粉々にするなどして食べられない状態にしないと検査できないなら，全数検査は不可能である．もちろん，全数調査には時間とコストがかかるという欠点もある．(3) と (4) については，以下で説明するような数学的なモデルを導入することで解決することができる．

実験や調査を何度も何度も繰り返し行った時に得られるだろうデータの仮想的な集合を考え，それを**母集団**と呼ぶことにする．上の農産物の例では，生産される農産物すべてに対して検査をした時に得られるデータの集合 (全数検査の結果) が母集団である．そこから無作為に抽出した数値の集合を**標本**と呼ぶことにしよう．実際に実験や調査を行ってデータを得ることは，母集団から標本を無作為に抽出することであると考える．ちょうど，数字が書いてある球がたくさん入っている壺から無作為に球を選ぶように，実験や調査の結果が得られると考えるのである．



どのような標本 (データ) が出やすいのかは，どのような数値が母集団の中に多く含まれるのかで決まる．つまり，標本の出現確率 (**確率分布**) は母集団の相対度数分布 (**母集団分布**) で決まる．標本の出現確率がわかれば，(原理的には) 標本の平均の出現確率もわかることになる．したがって，再検査したら平均値がどのような値をとるかは母集団分布で決まると考えることができる．

上の抜き取り検査では，母集団の平均値 (**母平均**)μ が基準より小さいかどうかを標本の平均値 (**標本平均**) $\bar{X}$ を見て判断することになる．母平均が安全基準を満たしていないにもかかわらず，標本平均が安全基準を下回る可能性が小さくなければ，標本平均には母平均よりも厳しい基準値を設ける必要がある．どの程度厳しくするのかは，安全でないものを安全であると判断ミスをした時の悪影響の度合いにもよるが，その可能性はできるだけ小さい方が望ましいのは言うまでもない．

このように，母集団を仮想してデータの変化を確率的にとらえ，判断ミスの可能性 (確率) を評価することで判断方法の性能を評価できるようになり，1 回の実験や調査から判定する基準を設定することができるようになる．母集団分布に応じた標本の確率的な変化 (確率分布) とそれに応じた標本平均の確率分布，あるいは，それから判断する基準の設定方法を数学的にどのように構成するのかを第 3 節 (p.13) 以降で解説する．データそのものの全体的特徴量を抽出する各種手法を記述統計と言うのに対して，データを標本ととらえて，その背後に潜む母集団の全体的特徴量を推し量る手法を**推測統計**と呼ぶ．調査対象の全データを知ることが困難であることが多いため，推測統計の各種手法は有効なアプローチとして多くの問題に利用されている．

2　データの要約

本節では，初等的な記述統計，つまり，データの全体的な特徴を捉える基本的な方法について概説する．

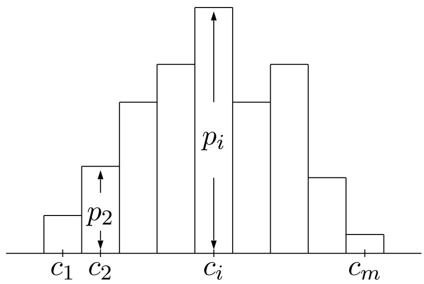
2.1　度数分布表・ヒストグラム

n 個のデータ全体を同じ幅の m 個の**階級** (小区間) に分割して，各階級の**度数** f_k (階級に含まれるデータの個数, $k=1,\dots,m$)，または**相対度数** $p_k = f_k/n$(データ全体の個数に対する度数の割合) を表 **2.1** のようにまとめたものを**度数分布表**といい，それを図 2.1 のように棒グラフで表したものを**ヒストグラム**と呼ぶ．また，各階級の中央の値 c_k を**階級値**と呼ぶ．

表 **2.1** 度数分布表

階級値	度数	相対度数
c_1	f_1	p_1
c_2	f_2	p_2
$\vdots$	$\vdots$	$\vdots$
c_m	f_m	p_m
計	n	1

図 **2.1** ヒストグラム



データの間隔が短く連続的に分布する時は，階級幅を小さくして分布を詳細に見ることが望ましい．ただし，そうすると各階級の相対度数が小さくなり，ヒストグラムの山が全体的に小さくなるため，分布の形状が分かりづらくなる．そんな時は，ヒストグラムの棒の面積が度数になるように描くと，階級幅が小さくなっても棒の高さは低くならないので，分布の概形を保ったまま分布が詳細になる．

階級の個数 m は，**スタージェス (Sturges) の公式**

$$m = \{(1+\log_2 n) \text{ を四捨五入した整数 }\}$$

から得られる個数を参考に，ヒストグラムの形状を見ながら調整することが多い．また，データが階級の境界上に存在する時は，測定単位より小さい目盛に境界を引くなどして，境界上にデータが存在しないように工夫するべきである．

2.2　代表値

広がりを持つデータ $x_1,\dots,x_n$ の大きさを 1 つの数値 (**代表値**という) で表すために，**平均**や**中央値**を用いる．平均を $\bar{x}$，中央値を Me と表すと，それぞれ，

$$\bar{x} := \frac{1}{n}\sum_{i=1}^{n} x_i, \quad Me := \begin{cases} x_{(j)} & (n \text{ が奇数 } 2j-1 \text{ の時}) \\ (x_{(j)} + x_{(j+1)})/2 & (n \text{ が偶数 } 2j \text{ の時}) \end{cases}$$

と定義される．ただし，$x_{(i)}$ はデータを小さい順に $x_{(1)} \leqq x_{(2)} \leqq \cdots \leqq x_{(n)}$ と並べた時の i 番目のデータ．また，$A := B$ あるいは $B =: A$ は A を B と定義することを意味する．

表 **2.1**(p.7) の度数分布表から平均 (の近似値) を求めるには，

$$\bar{x} \doteqdot \frac{1}{n}\sum_{k=1}^{m} c_k f_k = \sum_{k=1}^{m} c_k p_k$$

とすればよい．ここで，$\doteqdot$ は近似的に等しいことを表す記号である．

平均は，飛びぬけて大きいデータ，または飛びぬけて小さいデータ (**外れ値**) の影響を受けやすく，中央値は外れ値の影響を受けないことに注意しよう．

2.3 散らばりの指標と標準化

データ $x_1, \ldots, x_n$ に対して，$x_i - \bar{x}$ を x_i の**偏差**と呼び，その 2 乗和を**偏差平方和**とよび S_{xx} と表す．データ 1 個あたりの偏差平方和を**分散**と呼び，s_{xx} と表すと

$$S_{xx} := \sum_{i=1}^{n}(x_i - \bar{x})^2, \quad s_{xx} := \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2 = \frac{1}{n}S_{xx}$$

である．「平均 $\bar{x}$ から離れているデータ x_i が多ければ多いほど，s_{xx} が大きくなる」ので，分散 s_{xx} はデータの散らばりの指標であると考えられる．ただし，x_i の単位を d とすると，s_{xx} の単位は d^2 となることから，データ x_i や平均 $\bar{x}$ と単位を揃えるために，**標準偏差**と呼ばれる $s_x := \sqrt{s_{xx}}$ を散らばりの指標として用いることも多い．表 **2.1**(p.7) の度数分布表から分散 (の近似値) を求めるには，

$$s_{xx} \doteqdot \frac{1}{n}\sum_{k=1}^{m}(c_k - \bar{x})^2 f_k = \sum_{k=1}^{m}(c_k - \bar{x})^2 p_j$$

とすればよい．散らばりの指標には，偏差の絶対値の平均である**平均偏差** $\displaystyle\frac{1}{n}\sum_{i=1}^{n}|x_i - \bar{x}|$，最大値と最小値の差である**範囲** $\max\{x_1, \ldots, x_n\} - \min\{x_1, \ldots, x_n\}$ や**四分位範囲** $x_{(3n/4)} - x_{(n/4)}$ などもある．

偏差や分散には以下の公式が成り立つ．

【公式 1】 $\displaystyle\sum_{i=1}^{n}(x_i - \bar{x}) = 0$ (**偏差の和はゼロ!**)

〔証明〕 $\displaystyle\sum_{i=1}^{n}(x_i - \bar{x}) = \sum_{i=1}^{n} x_i - \sum_{i=1}^{n} \bar{x} = \sum_{i=1}^{n} x_i - n\bar{x} = 0.$

【公式 2】 $y_i = ax_i + b, i = 1, \ldots, n$ の平均 $\bar{y}$ と分散 s_{yy} は $\bar{y} = a\bar{x} + b, \quad s_{yy} = a^2 s_{xx}$.

〔証明〕$\bar{y} = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n}(ax_i + b) = \dfrac{1}{n}\left(a\sum_{i=1}^{n} x_i + \sum_{i=1}^{n} b\right) = \dfrac{1}{n}(an\bar{x} + nb) = a\bar{x} + b.$

$$s_{yy} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \bar{y})^2 = \frac{1}{n}\sum_{i=1}^{n}(ax_i + b - a\bar{x} - b)^2 = \frac{1}{n}\sum_{i=1}^{n} a^2(x_i - \bar{x})^2 = a^2 s_{xx}.$$

【公式 3】 $\dfrac{1}{n}\displaystyle\sum_{i=1}^{n}(x_i - a)^2 = s_{xx} + (\bar{x} - a)^2$, a は任意. $a = 0$ とすると次が成り立つ:

$$s_{xx} = \frac{1}{n}\sum_{i=1}^{n} x_i^2 - \bar{x}^2 \qquad \text{分散} = (2\text{ 乗の平均}) - (\text{平均の }2\text{ 乗})$$ **【分散公式】**

〔証明〕演習 **2.2**(p.11).

データ $x_1, \ldots, x_n$ を $z_i = (x_i - \bar{x})/s_x$ に変換することを**標準化**と呼ぶ. さらに, $w_i = z_i \times 10 + 50$ と変換された w_i を**偏差値**と呼ぶ. どのようなデータでも標準化すると平均が 0, 標準偏差が 1 となり, 偏差値は平均は 50, 標準偏差が 10 となる.

【公式 4】 $\bar{z} = 0$, $s_z = 1$. $\bar{w} = 50$, $s_w = 10$.

〔証明〕$z_i = \dfrac{1}{s_x}x_i - \dfrac{\bar{x}}{s_x}$ だから, $a = \dfrac{1}{s_x}, b = -\dfrac{\bar{x}}{s_x}$ として, 公式 2 を用いると

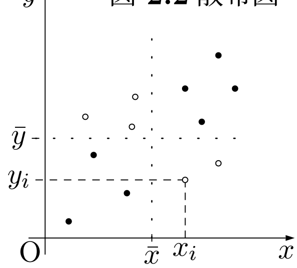
$$\bar{z} = \frac{1}{s_x}\bar{x} - \frac{\bar{x}}{s_x} = 0, \quad s_z^2 = \frac{1}{s_x^2}s_{xx} = 1.$$

$\bar{w} = 50$, $s_w^2 = 100$ についても同様にできる.

2.4 散布図・共分散・相関係数

2 次元データ $(x_1, y_1), \ldots, (x_n, y_n)$ の分布を視覚的にとらえるために, x 座標が x_i, y 座標が y_i である n 個の点 (x_i, y_i) を図 **2.2** のように表したものを**散布図**という. また, x と y の直線的関係の強さを表す指標として,

図 **2.2** 散布図



$$s_{xy} := \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}), \qquad r_{xy} := \frac{s_{xy}}{s_x\, s_y}$$

により定義される**共分散** s_{xy} と**相関係数** r_{xy} を用いる. (x_i, y_i) が $(\bar{x}, \bar{y})$ より右上, または左下にあると, $(x_i - \bar{x})(y_i - \bar{y})$ は正の値をとるため, 散布図において右上がりの傾向があると, s_{xy} は正の大きな値をとる. 逆に右下がりの傾向があると, s_{xy} は負

の小さな値をとる．相関係数 r_{xy} は標準偏差の大きさで s_{xy} で調整したものである．これらは以下の性質を持つ．

【公式 5】 $v_i = ax_i + b,\ w_i = cy_i + d$ とおくと，$s_{vw} = ac\, s_{xy}, \quad r_{vw} = \pm r_{xy}$
〔証明〕公式 2(p.9) より $\bar{v} = a\bar{x} + b,\ s_{vv} = a^2 s_{xx},\ \bar{w} = c\bar{y} + d,\ s_{ww} = c^2 s_{yy}$. また，公式 2(p.9) と同じようにすると，$s_{vw} = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} ac(x_i - \bar{x})(y_i - \bar{y}) = acs_{xy}$.

【公式 6】 $\dfrac{1}{n}\displaystyle\sum_{i=1}^{n}(x_i - a)(y_i - b) = s_{xy} + (\bar{x} - a)(\bar{y} - b),\ a, b$ 任意．$a = b = 0$ とすると，次がなりたつ:

$$s_{xy} = \frac{1}{n}\sum_{i=1}^{n} x_i y_i - \bar{x}\,\bar{y}, \quad \text{共分散} = (\text{積の平均}) - (\text{平均の積}) \quad \textbf{【共分散公式】}$$

〔証明〕演習 **2.6**(p.12)．

【公式 7】 2 次元データ $(x_1, y_1), \ldots, (x_n, y_n)$ の値がどんな値でも，$-1 \leqq r_{xy} \leqq 1$ が成り立つ．等号が成立するのは，すべてのデータが同一直線上にあるときのみ．
〔証明〕演習 **2.7**(p.12)．

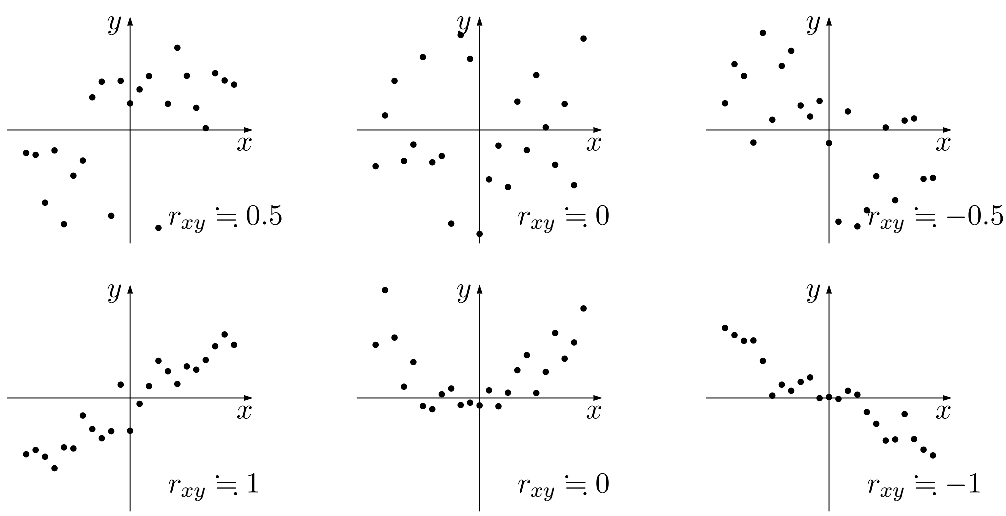
散布図の典型的な例を図 **2.3** に挙げた．それらをまとめると，

(1) 散布図が右肩上がりなら相関は正，右肩下がりなら相関は負．

(2) 相関係数は直線的関係の強さ．例えば 2 次関数的な関係は捉えられない (下中)．

となる．

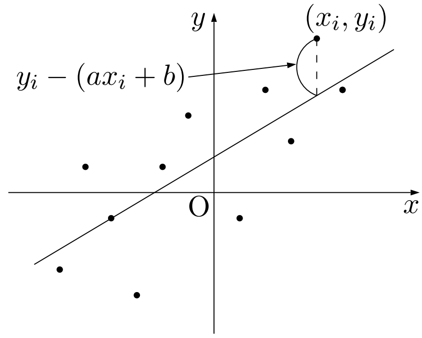
図 **2.3** 散布図と相関係数



2.5 最小２乗法

相関係数が強い 2 次元データ (x_1, y_1), $\ldots$, (x_n, y_n) に対して，背後に想定される直線 $y = ax + b$ の傾き a, b は，最小 2 乗推定量

$$\hat{a} := \frac{s_{xy}}{s_{xx}}, \quad \hat{b} := \bar{y} - \bar{x}\frac{s_{xy}}{s_{xx}}$$



により推定できる (演習 **2.7**(p.12))．これは，直線 $y = ax + b$ とデータのずれを

$$S(a, b) = \sum_{i=1}^{n} (y_i - (ax_i + b))^2$$

により測り，これを最小化する a, b を求めた結果得られる．直線 $y = \hat{a}x + \hat{b}$ を**回帰直線**と呼ぶ.

演習問題

演習 2.1 $4, 2, 1, 4, -9, -8, -4, 4, 6, -3, -1, -8$ の平均と分散，標準偏差を求めよ.

演習 2.2

(1) 公式 3(p.9) の $\frac{1}{n}\sum_{i=1}^{n}(x_i - a)^2 = s_{xx} + (\bar{x} - a)^2$ を示せ．(ヒント) $(x_i - a)^2 = (x_i - \bar{x} + \bar{x} - a)^2 = (x_i - \bar{x})^2 + 2(\bar{x} - a)(x_i - \bar{x}) + \cdots$ と変形して，公式 1(p.8) と s_{xx} の定義を使う.

(2) $\displaystyle\sum_{i=1}^{n}(x_i - 5)^2 = 273$, $s_{xx} = 12$, $\bar{x} = 8$ のとき，n の値を求めよ.

演習 2.3 英語のテストを 5 点刻みで採点したので，得点の集計をするために得点 x_1, $\ldots$, x_n を $y_i = (x_i - m)/5$ と変換したところ，$\bar{y} = 8$, $s_{yy} = 9$ となった．ただし，m は最低点とする．このとき，平均点 $\bar{x}$ は最低点 m より何点高いか．また，得点の標準偏差 s_x は何点か.

演習 2.4 次の度数分布表から平均と分散の近似値を求めよ.

階級値	1	3	5	7	9
度数	3	8	6	2	1

演習 2.5　次の 2 次元データの散布図に描き，共分散，相関係数を求めよ．さらに，回帰直線を求め，散布図に書き加えよ．ただし，x は演習 **2.1** と同一のものである．

x	4	2	1	4	-9	-8	-4	4	6	-3	-1	-8
y	-2	5	-9	4	-3	-7	-9	8	5	-1	-7	-8

演習 2.6　公式 6(p.10) の $\frac{1}{n}\sum_{i=1}^{n}(x_i-a)(y_i-b)=s_{xy}+(\bar{x}-a)(\bar{y}-b)$ を示せ．

演習 2.7　2 次元データ $(x_1,y_1),\ldots,(x_n,y_n)$ に対して，$X_i=x_i-\bar{x}, Y_i=y_i-\bar{y}$ とおく．以下の問いに答えよ．

(1) $\displaystyle\sum_{i=1}^{n}X_i^2=ns_{xx},\ \sum_{i=1}^{n}Y_i^2=ns_{yy},\ \sum_{i=1}^{n}X_iY_i=ns_{xy}$ を示せ．

(2) $\displaystyle f(t)=\frac{1}{n}\sum_{i=1}^{n}(X_it-Y_i)^2$ と定義する．このとき，$f(t)=s_{xx}t^2-2s_{xy}t+s_{yy}$ を示せ．さらに，$f(t)$ を最小にする t の値 t_0 を s_{xx}, s_{yy}, s_{xy} を用いて表せ．

(3) 任意の実数 t と任意の $i=1,\ldots,n$ に対して，$(X_it-Y_i)^2\geqq 0$ だから，任意の実数 t に対して $f(t)\geqq 0$ である．したがって，$f(t_0)\geqq 0$ である．このことから，$s_{yy}-\frac{s_{xy}^2}{s_{xx}}\geqq 0$ を示せ．さらに，$|r_{xy}|\leqq 1$ を示せ．

(4) $r_{xy}=\pm 1$ であることと，$f(t_0)=0$ が同値であることを示せ．さらに，それらと 2 次元データ $(x_1,y_1),\ldots,(x_n,y_n)$ が全て直線 $y=t_0(x-\bar{x})+\bar{y}$ 上にあることが同値であることを示せ．

演習 2.8　定数 x_1, …, x_n に対して，2 つの a の関数

$$F(a)=\sum_{i=1}^{n}(x_i-a)^2,\quad G(a)=\sum_{i=1}^{n}|x_i-a|$$

を考える．$F(a)$ と $G(a)$ それぞれの最小とその時の a の値を求めよ．

演習 2.9　演習 **2.5** の 2 次元データの散布図において，各点 (x_i,y_i) との距離が d_i である直線を $ax+by+c=0$ とする．このとき，

$$T(a,b,c)=\sum_{i=1}^{n}d_i^2$$

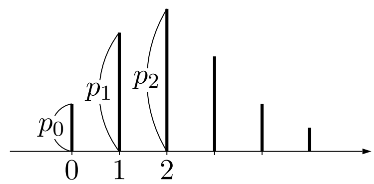
を a, b, c の式で表せ．さらに，$T(a,b,c)$ の値が最小になる a, b, c の値について検討せよ．

3　確率変数とその期待値

推測統計では，データは実験する前は確定しない量であるが，実験によって実現する値がどのような値をとりやすいかは，母集団分布によって決まる量であると考えた．本節では，ランダムに変化する変数 (確率変数という) とその期待値について概説する．確率変数やその期待値は確率論の概念であるが，推測統計を考える時，観測するまでどんな値になるかわからない標本は確率変数であり，母集団分布がその確率分布である．実際に観測されたデータは確率変数の実現値と考える．

3.1　確率変数

変数 X が非負整数 $\{0, 1, 2, \ldots,\}$ のいずれかの値をとり，任意の非負整数 k に対して $X = k$ となる確率 $P(X = k)$ が定義されているとき，X を**離散確率変数**と呼ぶ．$P(X = k)$ を p_k と表し，**確率関数**と呼ぶ．確率変数が実験や観測の結果を表すことがあるので，とりうる値の非負整数一つ一つを**実現値**と呼ぶ．離散確率変数は X, Y などの大文字を使って表し，それらの実現値は k, l などの小文字を使って表すことにする．確率の性質から，確率関数 p_k は



確率関数 p_k

$$\sum_{k=0}^{\infty} p_k = P(X = 0, 1, 2, \ldots) = 1$$

を満たす．確率変数 X の確率関数が p_k であることを簡単に $X \sim p_k$ と表すことにする．X の値が集合 A に含まれる確率を $P(X \in A)$，または，$P^X(A)$ と表し，

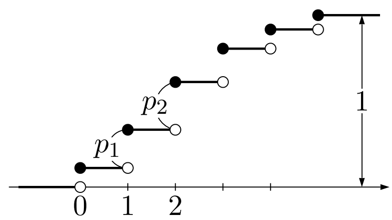
$$P(X \in A) = P^X(A) := \sum_{k \in A} p_k$$

と定義する．ここで，$\sum_{k \in A} p_k$ は A に含まれるすべての k についての p_k の合計．

さらに，

$$F(x) = P(X \leqq x) = \sum_{i \leqq x} p_i$$

を分布関数と呼ぶ．このとき，$F(x)$ は，右のような単調増加な階段関数であり，$F(-\infty) = 0$，$F(\infty) = 1$ を満たし右連続である．また，$P(a \leqq X \leqq b) = F(b) - F(a-1)$ が成り立つ．



分布関数 $F(x)$

《例 **3.1**》
右の表はある地域の家族数の相対度数分布である．その地域から無作為に 1 つ選んだ世帯の家族数を X とすると，$P(X=1)=0.35$, $P(X=2)=0.25$ のように相対度数と一致する．

人数	1	2	3	4	5
割合	0.35	0.25	0.2	0.15	0.05

《例 **3.2**》　大量に生産されるすべての製品 A に対して，不良品には 1，良品には 0 を対応させた集合 Ω を考える．Ω には 0 と 1 が多数含まれるが，そこにおける 1 の相対度数 p が不良品率を表す．Ω から無作為に 1 つの数を取り出した結果を X とすると，X の取りうる値は 0 と 1 であり，$p_1=P(X=1)=p$, $p_0=P(X=0)=1-p$ である．X の実現値が 1 となることは，無作為に 1 つの製品を抜き出した時それが不良品であることに対応している．推測統計では，Ω を母集団と呼び，X を標本と呼ぶ．確率変数 X の実現値である 0 や 1 は実際に観測されるデータである．

Ω から無作為に数を 1 つ取り出しては戻すことを n 回繰り返した時，取り出された 1 の個数を Y とする．Y の取りうる値は 0 から n までの整数であり，$Y=k$ となる確率 $P(Y=k)$ は ${}_nC_k\,p^k(1-p)^{n-k}$ である．たとえば，$n=3$ の時 $Y=2$ となるのは，取り出した結果が 011, 101, 110 の 3 通りであり，いずれの結果も $p^2(1-p)$ の確率で起こるので，$P(Y=2)=3p^2(1-p)$ となる．一般に，$Y=k$ となる結果は，何回目が 1 となるのかを考えると ${}_nC_k$ 通りあり，それぞれの結果が起こる確率はいずれも $p^k(1-p)^{n-k}$ なので $P(Y=k)={}_nC_k\,p^k(1-p)^{n-k}$ となることがわかる．この確率関数の表す確率分布を **2 項分布**と呼び，$B(n,p)$ と表す (4.1 節参照)．また，離散確率変数 Y が $B(n,p)$ の確率関数を持つことを Y は $B(n,p)$ に従うと言い，$Y\sim B(n,p)$ と表す．

$Y=k$ は製品から n 個抜き取った時，不良品が k 個になることに対応しており，$P(Y=k)$ は抜取検査の結果不良品が k 個になる確率を表している．

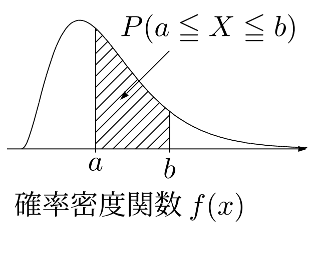
《例 **3.3**》　サイコロの出る目 X のとりうる値は $\{1,2,3,4,5,6\}$ であり，$k=1,2,\ldots,6$ に対して，$p_k=P(X=k)=\frac{1}{6}$ である．それ以外の k では $p_k=0$.

$$P(X=1,2)=\frac{1}{6}+\frac{1}{6}=\frac{1}{3},\ P(X\neq 3)=1-P(X=3)=\frac{5}{6}.$$

一方，変数 X のとりうる値が実数全体 $(-\infty,\infty)$ であり，$f(x)\geqq 0$ である関数を用いて，X が a 以上 b 以下である確率 $P(a\leqq X\leqq b)$ が

$$P(a\leqq X\leqq b)=\int_a^b f(x)dx$$

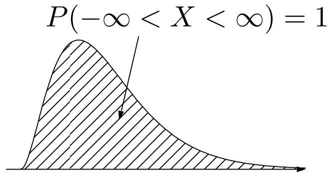
と定義されるとき，X を**連続確率変数**と呼び，$f(x)$



確率密度関数 $f(x)$

を**確率密度関数** (probability density function)，あるいは，略して **p.d.f.** と呼ぶ．このとき，すべての実数一つ一つが**実現値**である．連続確率変数は X, Y などの大文字を使って表し，実現値は x, y などの小文字を使って表すことにする．確率の性質より，$f(x)$ は

$$\int_{-\infty}^{\infty} f(x)dx = P(-\infty < X < \infty) = 1$$

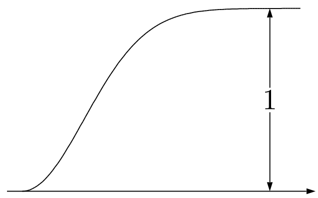


を満たす．X が p.d.f. が $f(x)$ である連続確率変数であることを $X \sim f(x)$ と表す．

さらに，

$$F(x) = P(X \leqq x) = \int_{-\infty}^{x} f(u)du$$

を**分布関数**と呼ぶ．このとき，$F(x)$ は，右のような単調増加関数であり，$F(-\infty) = 0, F(\infty) = 1$ を満たし，密度関数と $F'(x) = f(x)$ という関係がある．



分布関数 $F(x)$

注意 3.1. 任意の実数 c に対して，$P(X = c) = P(c \leqq X \leqq c) = \int_c^c f(x)dx = 0$ だから，

$$P(a \leqq X \leqq b) = P(a < X < b) = P(a \leqq X < b).$$

つまり，連続確率変数 X に関しては，不等号で示される範囲の確率を求める時，不等号の両端が含まれるかどうかで値は異ならない．

注意 3.2. X が離散でも連続でも，$P(X \in A) = P^X(A)$ は次の性質を満たす:

(1) $P^X(A^c) = 1 - P^X(A),\ P^X(\phi) = 0,\ P^X(\Omega) = 1$. ただし Ω は実現値の全体.

(2) $A \cap B = \phi \Rightarrow P^X(A \cup B) = P^X(A) + P^X(B)$.

《例 **3.4**》 連続確率変数 X の p.d.f. $f(x)$ は $f(x) = \begin{cases} (b-a)^{-1} & (a < x < b) \\ 0 & (\text{それ以外}) \end{cases}$ とする．ただし，$a < b$．この確率密度関数で表される確率分布を**一様分布**と呼び，$U(a,b)$ と表す (4.4 節参照)．X がこの確率密度関数を持つことを，X は $U(a,b)$ に従うと言い，$X \sim U(a,b)$ と表す．この $f(x)$ は，$\displaystyle\int_{-\infty}^{\infty} f(x)dx = \int_a^b (b-a)^{-1}dx = 1$ であり，$f(x) \geqq 0$ なので，確率密度関数であることが確認できる．また，

$$P\left(\frac{2a+b}{3} < X < \frac{a+2b}{3}\right) = \int_{\frac{2a+b}{3}}^{\frac{a+2b}{3}} (b-a)^{-1}dx = \left[(b-a)^{-1}x\right]_{\frac{2a+b}{3}}^{\frac{a+2b}{3}} = \frac{1}{3}.$$

《例 **3.5**》 耐久性の優れた蛍光灯の寿命について考える．未使用の蛍光灯が点灯しなくなるまでの累積点灯時間，つまり寿命を X とする．この X の取りうる値は $(0,\infty)$ であり，点灯しなくなるまで使用し続けないとわからないのである確率変数と考えられる．一般に，耐久性の優れた製品の寿命 X は $f(x)=\begin{cases}\lambda e^{-\lambda x} & (x>0)\\ 0 & (x\le 0)\end{cases}$ という p.d.f. を持つ連続確率変数であると考えられている．ただし，λ は耐久性を表す正の定数である．蛍光灯の寿命 X もこの確率密度関数を持つと仮定すると，

$$P(a<X<b)=\int_a^b \lambda e^{-\lambda x}dx=\Big[-e^{-\lambda x}\Big]_a^b=-e^{-b\lambda}+e^{-a\lambda}$$

である．上のような確率密度関数で表される確率分布を**指数分布**といい，$Ex(\lambda)$ と表す (4.6 節参照)．X の確率密度関数が上のように与えられることを，X が $Ex(\lambda)$ に従うと言い，$X\sim Ex(\lambda)$ と表す．

3.2 期待値と分散

$X\sim p_k$，または，$X\sim f(x)$ のとき，任意の関数 $h(x)$ に対して，$E[h(X)]$ を

$$E[h(X)]:=\begin{cases}\displaystyle\sum_{k=0}^{\infty}h(k)p_k & (X\sim p_k\text{の時})\\ \displaystyle\int_{-\infty}^{\infty}h(x)f(x)dx & (X\sim f(x)\text{ の時})\end{cases}$$

と定義し，$h(X)$ の**期待値**，または**平均**と呼び，μ_h と表す．また，

$$V[h(X)]:=E[\{h(X)-\mu_h\}^2]=E[\{h(X)-E[h(X)]\}^2]$$

と定義し，$h(X)$ の**分散**と呼び，σ_h^2 と表す．特に，$h(x)=x$ の時，$E[X]$ を X の期待値 (または平均) と呼び，μ_X と表す．また，$V[X]=E[(X-E[X])^2]=E[(X-\mu_X)^2]$ であるが，$V[X]$ を X の分散と呼び，σ_X^2 と表すことにする．さらに，$\sigma_X=\sqrt{V[X]}$ を X の**標準偏差**と呼ぶ．期待値や分散は確率変数ではない，定数であることに注意しよう．したがって，$E[\{h(X)-E[h(X)]\}^2]$ の中では，$E[h(X)]$ は定数であり，外の期待値 $E[\{\cdots\}^2]$ を計算する時には，中の $E[h(X)]$ は定数扱いすることに注意する必要がある．

《例 **3.6**》 例 **3.1**(p.14) の X に関して，

$$\begin{aligned}E[X]=&1\times 0.35+2\times 0.25+\cdots+5\times 0.05=2.3,\\ V[X]=&(1-2.3)^2\times 0.35+\cdots+(5-2.3)^2\times 0.05=1.51\end{aligned}$$

となり，この地域の家族数の平均と分散に一致する．

《例 **3.7**》 不良品の抜取検査の 例 **3.2**(p.14) における X に関しては，

$$
\begin{aligned}
E[X] =& 0 \times (1-p) + 1 \times p = p, \quad E[X^2] = 0^2 \times (1-p) + 1^2 \times p = p, \\
V[X] =& (0-p)^2 \times (1-p) + (1-p)^2 \times p = p(1-p).
\end{aligned}
$$

不良品の合計数 Y に関しては例 **3.14**(p.20) で説明する．

《例 **3.8**》 X がサイコロの目である時，例 **3.3**(p.14) のように確率は計算されるが，

$$
\begin{aligned}
E[X] =& 1 \times \frac{1}{6} + 2 \times \frac{1}{6} + \cdots + 6 \times \frac{1}{6} = \frac{7}{2}, \\
V[X] =& (1-\frac{7}{2})^2 \times \frac{1}{6} + (2-\frac{7}{2})^2 \times \frac{1}{6} + \cdots + (6-\frac{7}{2})^2 \times \frac{1}{6} = \frac{35}{12}
\end{aligned}
$$

である．また，

$$
\begin{aligned}
E[2X+3] =& (2 \times 1 + 3) \times \frac{1}{6} + \cdots + (2 \times 6 + 3) \times \frac{1}{6} = 10 \\
V[2X+3] =& (2 \times 1 + 3 - 10)^2 \times \frac{1}{6} + \cdots + (2 \times 6 + 3 - 10)^2 \times \frac{1}{6} = \frac{35}{3}.
\end{aligned}
$$

《例 **3.9**》 例 **3.4**(p.15) の X，つまり，一様分布 $U(a,b)$ に従う X に対して，

$$
\begin{aligned}
E[X] &= \int_a^b x(b-a)^{-1}dx = \left[(b-a)^{-1}\frac{1}{2}x^2\right]_a^b = \frac{1}{2(b-a)}(b^2-a^2) = \frac{a+b}{2}, \\
E[X^2] &= \int_a^b x^2(b-a)^{-1}dx = \left[\frac{1}{3(b-a)}x^3\right]_a^b = \frac{1}{3(b-a)}(b^3-a^3) = \frac{a^2+ab+b^2}{3}.
\end{aligned}
$$

つまり，X が一様分布 $U(a,b)$ に従う時，$E[X]$, $E[X^2]$ は上のようになることが分かった．

《例 **3.10**》 例 **3.5**(p.16) の蛍光灯の寿命 X に対して，$E[X^k] = \int_0^\infty x^k \lambda e^{-\lambda x} dx$ である．予備知識 0.3(6) より $E[X^k] = \lambda \left\{ \left[-\frac{1}{\lambda} x^k e^{-\lambda x} \right]_0^\infty + \frac{k}{\lambda} \int_0^\infty x^{k-1} e^{-\lambda x} dx \right\}$ であり，さらに予備知識 0.2(11) より，$\left[-\frac{1}{\lambda} x^k e^{-\lambda x} \right]_0^\infty = 0$ である．したがって，

$$
E[X^k] = \frac{k}{\lambda} \int_0^\infty x^{k-1} \lambda e^{-\lambda x} dx = \frac{k}{\lambda} E[X^{k-1}].
$$

いま，$E[X^0] = \int_0^\infty \lambda e^{-\lambda x} dx = \left[-e^{-\lambda x} \right]_0^\infty = 1$ だから，

$$
E[X^k] = \frac{k}{\lambda} \frac{k-1}{\lambda} \cdots \frac{1}{\lambda} E[X^0] = \frac{k!}{\lambda^k}
$$

である．特に $k=1,2$ とすると，$E[X]=\frac{1}{\lambda}$, $E[X^2]=\frac{2}{\lambda^2}$ がわかる．蛍光灯の寿命 X は指数分布 $Ex(\lambda)$ に従うので，一般に $X\sim Ex(\lambda)$ ならば $E[X]=\frac{1}{\lambda}$, $E[X^2]=\frac{2}{\lambda^2}$ であることが分かった．

3.3 期待値と分散の性質

確率変数の関数に対する期待値や分散には以下の公式が成り立つ．これらの公式を上手に利用すると，定義にかえらなくても，様々な計算が可能となる．

【公式 8】 関数 $h(x)$, $g(x)$ と定数 a, b に対して，

$$E[a\,h(X)+b\,g(X)]=a\,E[h(X)]+b\,E[g(X)] \ (\textbf{線型性})$$

〔証明〕X が連続確率変数で $X\sim f(x)$ のとき，

$$E[a\,h(X)+b\,g(X)]=\int_{-\infty}^{\infty}\{ah(x)+bg(x)\}f(x)dx$$
$$=a\int_{-\infty}^{\infty}h(x)f(x)dx+b\int_{-\infty}^{\infty}g(x)f(x)dx=a\,E[h(X)]+b\,E[g(X)].$$

X が離散確率変数の時も同様に示せる (演習 **3.4**(p.23))．

【公式 9】 $h(X)=a(a$ は定数$)$ のとき，$E[a]=a, V[a]=0$.
〔証明〕$X\sim p_k$ とする．

$$E[a]=\sum_{k=0}^{\infty}ap_k=a\sum_{k=0}^{\infty}p_k=a\times 1=a,\quad V[a]=E[\{a-E[a]\}^2]=E[0]=0$$

$X\sim f(x)$ の時も同様に示せる．

【公式 10】 $E[h(X)-\mu_h]=0$. 特に，$E[X-\mu_X]=0$. (**偏差の平均は 0**)
〔証明〕$E[h(X)-\mu_h]=E[h(X)]-E[\mu_h]=\mu_h-\mu_h=0$.

【公式 11】 任意の定数 a に対して，$E[(X-a)^2]=V[X]+\{E[X]-a\}^2$.
特に，$V[X]=E[X^2]-\{E[X]\}^2$(**分散公式**)
〔証明〕$(X-a)^2=(X-\mu_X+\mu_X-a)^2=(X-\mu_X)^2+2(\mu_X-a)(X-\mu_X)+(\mu_X-a)^2$ だから，期待値の線形性 (公式 8) を使うと，

$$E[(X-a)^2]=E[(X-\mu_X)^2]+2(\mu_X-a)E[X-\mu_X]+E[(\mu_X-a)^2].$$

この第 2 項は公式 10 より 0 であり，第 3 項 の $(\mu_X-a)^2$ は定数だから公式 9 を使うと

$$E[(X-a)^2]=E[(X-\mu_X)^2]+(\mu_X-a)^2=V[X]+\{E[X]-a\}^2.$$

この式は任意の a に対して成り立つが，特に，$a=0$ として，$V[X]$ について解くと分散公式が得られる．

【公式 12】 関数 $h(x)$ と定数 a, b に対して，$V[ah(X)+b]=a^2V[h(X)]$.
〔証明〕期待値の線形性 (公式 8(p.18)) より，$E[a\,h(X)+b]=a\,E[h(X)]+b$．したがって，

$$\begin{aligned}V[ah(X)+b] =& E[(ah(X)+b-E[ah(X)+b])^2]=E[(ah(X)-aE[h(X)])^2]\\ =& a^2E[(h(X)-E[h(X)])^2]=a^2V[h(X)].\end{aligned}$$

【公式 13】 $Z=\dfrac{X-\mu_X}{\sigma_X}$ とすると，$E[Z]=0, V[Z]=1$. X から期待値 μ_X を引いて，標準偏差 σ_X で割って，Z に変形することを**標準化**という．
〔証明〕$Z=\frac{1}{\sigma_X}X-\frac{\mu_X}{\sigma_X}$ だから，期待値の線形性 (公式 8(p.18)) より，

$$E[Z]=\frac{1}{\sigma_X}E[X]-\frac{\mu_X}{\sigma_X}=\frac{1}{\sigma_X}\mu_X-\frac{\mu_X}{\sigma_X}=0.$$

また，公式 12 より，$V[Z]=\left(\dfrac{1}{\sigma_X}\right)^2V[X]=\dfrac{1}{\sigma_X^2}\sigma_X^2=1.$

《例 **3.11**》 サイコロの出る目 X に対して，

$$E[X^2]=1^2\times\frac{1}{6}+2^2\times\frac{1}{6}+\cdots+6^2\times\frac{1}{6}=\frac{91}{6}$$

であり，例 **3.8**(p.17) で $E[X]=\frac{7}{2}$ だったので，分散公式 (公式 11(p.18)) を用いると

$$V[X]=\frac{91}{6}-\left(\frac{7}{2}\right)^2=\frac{35}{12}$$

となり，例 **3.8**(p.17) で求めた $V[X]$ と一致することがわかる．

《例 **3.12**》 例 **3.4**(p.15) で考えたように，一様分布 $U(a,b)$ に従う X に対して，例 **3.9**(p.17) で $E[X]=\frac{a+b}{2}$, $E[X^2]=\frac{a^2+ab+b^2}{3}$ がわかった．これらを確率変数の分散公式 (公式 11(p.18)) に代入すると，

$$V[X]=\frac{a^2+ab+b^2}{3}-\left(\frac{a+b}{2}\right)^2=\frac{(b-a)^2}{12}$$

となる．もちろん，$V[X]$ の定義から直接

$$V[X]=\int_a^b\left(x-\frac{a+b}{2}\right)^2\times(b-a)^{-1}dx=\left[\frac{1}{3(b-a)}\left(x-\frac{a+b}{2}\right)^3\right]_a^b=\frac{(b-a)^2}{12}$$

と求めることもできる．つまり，$X \sim U(a,b)$ の時，$E[X] = \frac{a+b}{2}$, $V[X] = \frac{(b-a)^2}{12}$ がわかった．

《例 **3.13**》 例 **3.5**(p.16) で定義した蛍光灯の寿命 X の p.d.f. に対して，例 **3.10**(p.17) では $E[X] = 1/\lambda$, $E[X^2] = 2/\lambda^2$ がわかった．確率変数の分散公式 (公式 11(p.18)) を用いると，$V[X] = \dfrac{2}{\lambda^2} - \left(\dfrac{1}{\lambda}\right)^2 = \dfrac{1}{\lambda^2}$ となる．つまり，$X \sim Ex(\lambda)$ ならば，$V[X] = \dfrac{1}{\lambda^2}$.

3.4 積率母関数

$M_X(t) := E[e^{tX}]$ により定義される t の関数 $M_X(t)$ を，X の**積率母関数** (moment generating function)，あるいは, 略して **m.g.f.** と呼ぶ．積率母関数の定義では，まず t を与えて，その t に対して，確率変数 X の関数 e^{tX} 期待値を求める．求めた期待値は確率的には定数であり，最初に決めた t の値ごとに色々な値をとる．したがって，$M_X(t)$ は t の値に依存して変化する t の関数である．ただし，期待値を求める時に確率変数 X の確率関数，あるいは，確率密度関数を用いるので，X の確率分布に依存している決まる関数である．以下の公式では，積率母関数が確率変数 X の期待値を求めるために利用できることや積率母関数と確率関数，あるいは，確率密度関数が一対一に対応することを説明する．

【公式 14】 $M'_X(0) = E[X]$, $M''_X(0) = E[X^2]$, $\dfrac{d^m}{dt^m} M_X(0) = E[X^m]$.

〔証明〕$X \sim p_k$ とする．$M_X(t) = \sum_{k=0}^{\infty} e^{tk} p_k$ だから，

$$M'_X(t) = \sum_{k=0}^{\infty} k e^{tk} p_k,\ M''_X(t) = \sum_{k=0}^{\infty} k^2 e^{tk} p_k,\ \ldots\ ,\ \frac{d^m}{dt^m} M_X(t) = \sum_{k=0}^{\infty} k^m e^{tk} p_k.$$

これらに $t = 0$ を代入すると，$e^{tk} = 1$ となるので，$M'_X(0) = \displaystyle\sum_{k=0}^{\infty} k p_k = E[X]$, $M''_X(0) = \displaystyle\sum_{k=0}^{\infty} k^2 p_k = E[X^2]$, $\ldots$, $\dfrac{d^m}{dt^m} M_X(0) = \displaystyle\sum_{k=0}^{\infty} k^m p_k = E[X^m]$. $X \sim f(x)$ の時も同様に証明できる (演習 **3.7**(p.23)).

この公式から，$E[X^m]$ が積率母関数 $M_X(t)$ を微分することで求めることができることが分かった．$E[X^m]$ のことを X の m 次の**積率 (モーメント)** と呼ぶが，$M_X(t)$ を微分することで積率が次々に産まれるため，このような名前がついている．

《例 **3.14**》 例 **3.2**(p.14) において，n 個の製品を抜取検査した時，不良品の個数 Y が k 個である確率は $P(Y = k) = {}_nC_k p^k (1-p)^{n-k}$ で与えれることがわかった．つま

り，$Y \sim B(n,p)$．したがって，Y の積率母関数 $M_Y(t)$ は

$$M_Y(t) = E[e^{tY}] = \sum_{k=0}^{n} e^{tk}{}_nC_k p^k(1-p)^{n-k} = \sum_{k=0}^{n} {}_nC_k(pe^t)^k(1-p)^{n-k}$$

となるが，予備知識 0.1(6) の二項定理において，$a = pe^t$, $b = 1-p$ とすると，$M_Y(t) = (1-p+pe^t)^n$ と表されることがわかる．これを t で微分すると，

$$\begin{aligned}
M_Y'(t) =& n(1-p+pe^t)^{n-1}\{1-p+pe^t\}' = n(1-p+pe^t)^{n-1}pe^t,\\
M_Y''(t) =& n\left\{(n-1)(1-p+pe^t)^{n-2}\{1-p+pe^t\}'pe^t + (1-p+pe^t)^{n-1}\{pe^t\}'\right\}\\
=& n\left\{(n-1)(1-p+pe^t)^{n-2}(pe^t)^2 + (1-p+pe^t)^{n-1}pe^t\right\}
\end{aligned}$$

となる．したがって，

$$\begin{aligned}
E[Y] =& M_Y'(0) = n(1-p+p)^{n-1}p = np,\\
E[Y^2] =& M_Y''(0) = n\left\{(n-1)(1-p+p)^{n-2}(p)^2 + (1-p+p)^{n-1}p\right\}\\
=& np\{(n-1)p+1\}
\end{aligned}$$

が導ける．これらより，$V[Y] = E[Y^2] - \{E[Y]\}^2 = np(1-p)$ が導ける．以上より，$Y \sim B(n,p)$ の時，$M_Y(t) = (1-p+pe^t)^n$, $E[Y] = np$, $V[Y] = np(1-p)$.

《例 **3.15**》 例 **3.5**(p.16) で考えた蛍光灯の寿命 X について，その積率母関数 $M_X(t)$ は

$$M_X(t) = \int_0^\infty e^{tx}\lambda e^{-\lambda x}dx = \left[\frac{\lambda}{t-\lambda}e^{(t-\lambda)x}\right]_0^\infty = \frac{\lambda}{\lambda - t}$$

と求められる．ただし，$t-\lambda<0$，すなわち，$t<\lambda$ である t について最後の等号は成り立つ．$\lambda>0$ なので，原点を含む区間 $I = (-\infty, \lambda)$ に含まれる t に対して，上の計算が正当化される．この $M_X(t)$ を微分すると，

$$\begin{aligned}
M_X'(t) =& \{\lambda(\lambda-t)^{-1}\}' = \lambda \times (-1)(\lambda-t)^{-2} \times (-1) = \frac{\lambda}{(\lambda-t)^2},\\
M_X''(t) =& \{\lambda(\lambda-t)^{-2}\}' = \lambda \times (-2)(\lambda-t)^{-3} \times (-1) = \frac{2\lambda}{(\lambda-t)^3}.
\end{aligned}$$

したがって，$E[X] = M_X'(0) = \frac{\lambda}{(\lambda-0)^2} = \frac{1}{\lambda}$, $E[X^2] = M_X''(0) = \frac{2\lambda}{(\lambda-0)^3} = \frac{2}{\lambda^2}$ となり，期待値の定義から直接計算した例 **3.10**(p.17) の結果と一致することがわかる．

以上より，$X \sim Ex(\lambda)$ の時，$E[X] = 1/\lambda$, $V[X] = 1/\lambda^2$, $M_X(t) = \lambda/(\lambda-t)$.

【公式 15】

(1) 確率変数 X と Y に対して，ある原点近傍 (原点を含む開区間) I に含まれる任意の $t \in I$ に対して $M_X(t) = M_Y(t)$ が成り立つことと，任意の集合 $A \subset \mathbb{R}$ に対して，$P(X \in A) = P(Y \in A)$ が成り立つことは同値である．

(2) 確率変数列 X_n と確率変数 X に対して，ある原点近傍 I の任意の t に対して $\lim_{n\to\infty} M_{X_n}(t) = M_X(t)$ であることと，任意の $A \subset \mathbb{R}$ に対して $\lim_{n\to\infty} P(X_n \in A) = P(X \in A)$ は同値.

〔証明〕$X \sim f(x), Y \sim g(y)$ であり,，それらの差が，$f(x) - g(x) = \sum_{k=0}^{\infty} a_k x^k$ のようにテイラー展開できる場合だけ証明する．まず，$M_X(t) = M_Y(t)$ なので，任意の非負の整数 k に対して，$E[X^k] = M_X^{(k)}(0) = M_Y^{(k)}(0) = E[Y^k]$ に注意しよう.

$$\begin{aligned}
\int_{-\infty}^{\infty} (f(x) - g(x))^2 dx &= \int_{-\infty}^{\infty} \sum_{k=0}^{\infty} a_k x^k \left(f(x) - g(x)\right) dx \\
&= \sum_{k=0}^{\infty} a_k \left(\int_{-\infty}^{\infty} x^k f(x) dx - \int_{-\infty}^{\infty} x^k g(x) dx \right) \\
&= \sum_{k=0}^{\infty} a_k (E[X^k] - E[Y^k]) = 0.
\end{aligned}$$

$f(x) \neq g(x)$ のとき上の積分は 0 とならないので，$f(x) = g(x)$．後半の証明は省略する.

この公式から，積率母関数と確率分布が一対一に対応していることがわかる.

《例 **3.16**》 例 **3.15**(p.21) では，指数分布 $Ex(\lambda)$ に従う蛍光灯の寿命 X の積率母関数 $M_X(t)$ は $M_X(t) = \dfrac{\lambda}{\lambda - t}$ となることが分かった．いま，ある確率変数 Y の積率母関数 $M_Y(t)$ が $M_Y(t) = \frac{1}{1-3t}$ で与えられるとしよう．このとき，$M_Y(t) = \frac{1/3}{1/3-t}$ なので，$\lambda = \frac{1}{3}$ のときの $M_X(t)$ と一致することがわかる．公式 15(p.21) (1) から，Y の p.d.f. は 例 **3.5**(p.16) における X の p.d.f. に $\lambda = \frac{1}{3}$ を代入したものに等しいことがわかる．つまり，$Y \sim Ex(1/3)$ がわかる.

《例 **3.17**》 確率関数が $p_k = \dfrac{\lambda^k}{k!} e^{-\lambda}, \lambda > 0$ で与えれれる離散確率変数 X について考える．この確率関数が表す確率分布を**ポアソン分布**と呼び，$Po(\lambda)$ と表す (4.2 節参照)．また，この確率関数を持つ X はポアソン分布 $Po(\lambda)$ に従うと言い，$X \sim Po(\lambda)$ と表す．さて，X の積率母関数 $M_X(t)$ は

$$M_X(t) = \sum_{k=0}^{\infty} e^{tk} \frac{\lambda^k}{k!} e^{-\lambda} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{(\lambda e^t)^k}{k!}$$

であるが，予備知識 0.2 (10) の指数関数のテイラー展開において，$x = \lambda e^t$ とすると，上の和が計算できて，$M_X(t) = e^{-\lambda} e^{\lambda e^t} = e^{\lambda(e^t - 1)}$ となる．一方，確率関数が

$q_k^{(n)} = {}_nC_k r_n^k (1-r_n)^{n-k}$, $r_n = \lambda/n$ である離散確率変数 X_n を考える．$X_n \sim B(n, r_n)$ なので，例 **3.14**(p.20) の Y と同様にすると，X_n の積率母関数は

$$M_{X_n}(t) = (1 - r_n + r_n e^t)^n = \left(1 + \frac{\lambda(e^t-1)}{n}\right)^n$$

であり，$m = \frac{n}{\lambda(e^t-1)}$ とおくと，$n = m\lambda(e^t-1)$ なので，$M_{X_n}(t) = \left\{\left(1+\dfrac{1}{m}\right)^m\right\}^{\lambda(e^t-1)}$ となる．予備知識 0.2(1) の e の定義から，$n \to \infty$ の時，$M_{X_n}(t) \to e^{\lambda(e^t-1)} = M_X(t)$ となることがわかり，公式 15(p.21) の (2) から，$q_k^{(n)} \to p_k$ がわかる．つまり，2 項分布 $B(n, r_n)$，$r_n = \lambda/n$ の確率関数 $q_k^{(n)}$ はポアソン分布 $Po(\lambda)$ の確率関数 p_k に近づくことがわかった．このことは，試行回数 n が多いが，対象の事象が起こる確率 $r_n = \lambda/n$ が小さい時，2 項分布はポアソン分布で近似できることを意味している．

演習問題

演習 3.1 全有権者における内閣支持率が 50% の時，有権者を無作為に選ぶことを 10 回繰り返して，その中で内閣支持した人の人数を X とする．$P(X=3)$ の値を求めよ．また，$P(X \geqq 4)$ の値を求めよ．

演習 3.2 関数 $f(x) = \begin{cases} 2x^2 + a & (-1 \leqq x \leqq 1) \\ 0 & (|x| > 1) \end{cases}$ において，定数 a を適当に決めると，$f(x)$ がある連続確率変数の確率密度関数になるかどうかを答えよ．もし，確率密度関数になる場合は，a の値を求めよ．

演習 3.3 $h(a) = E[(X-a)^2]$ を最小にする a の値と最小値を求めよ．

演習 3.4 期待値の線形性 (公式 8(p.18)) を離散確率変数 $X \sim p_k$ に対して示せ.

演習 3.5 確率変数 X に対して，$V[X] = E[X(X-1)] + E[X] - \{E[X]\}^2$ が成り立つことを示せ.

演習 3.6 サイコロの出た目を 4 で割った余りを X とする．$\mu_X = E[X]$, $E[X^2]$, $\sigma_X^2 = V[X]$ を求めよ．さらに，$E[4X+1]$, $V[4X+1]$ を求めよ.

演習 3.7 積率母関数の性質 (公式 14(p.20)) を連続確率変数 $X \sim f(x)$ に対して示せ.

演習 3.8 確率変数 X の積率母関数 $M_X(t)$ が $M_X(t) = (1-\beta t)^{-\alpha}$ である時，$E[X]$, $E[X^2]$, $V[X]$ を求めよ．ただし，$\alpha > 0$, $\beta > 0$.

演習 3.9　確率変数 X の積率母関数 $M_X(t)$ が $M_X(t) = e^{\mu t + \frac{\sigma^2}{2} t^2}$ である時，$E[X]$, $E[X^2]$, $V[X]$ を求めよ．ただし，μ は任意の実数であり，$\sigma > 0$.

演習 3.10　離散確率変数 X の確率関数 p_k が $p_k = qp^k$, $k = 0, 1, 2, \ldots$, $0 < p < 1$, $q = 1 - p$ で与えられる時，$t < -\log p$ を満たす t に対して，X の積率母関数 $M_X(t)$ を求めよ．それを利用して，$E[X]$, $E[X^2]$, $V[X]$ を求めよ．

演習 3.11　離散確率変数 X の確率関数が $p_k = \dfrac{18}{(k+1)(k+2)(k+3)(k+4)}$, $k = 0, 1, 2, \ldots$ で与えられる時，$E[X+1]$, $E[(X+1)(X+2)]$ を求めよ．その結果から，$E[X]$, $V[X]$ を求めよ．

演習 3.12　連続確率変数 X の p.d.f. が $f(x) = \frac{a}{2} e^{-a|x|}$, $a > 0$ である時，$E[X]$, $V[X]$ を求めよ．また，$M_X(t)$ を求めよ．

演習 3.13　連続確率変数 $X \sim f(x)$ に対して，$h(a) = E[|X - a|]$ とおく．また，$g(a) = \int_{-\infty}^{a} f(x)$, $m(a) = \int_{-\infty}^{a} x f(x) dx$ とする．

(1)　$h(a) = 2ag(a) - 2m(a) - a + \mu_X$ を示せ．ただし，$\mu_X = E[X]$.

(2)　$h'(a)$, $h''(a)$ を求め，$h''(a) \geqq 0$ を示せ．また，$h'(a) = 0$ を満たす a が満たすべき条件を求めよ．

(3)　$h(a)$ を最小値を求めよ．

演習 3.14　一辺の長さが 1 である正方形の内部から 1 点を無作為に選び，その点から最も近い辺までの距離を X とする．このとき，$P(a < X < b)$ を求めよ．また，X の p.d.f. $f(x)$ を求めよ．さらに，$E[X]$ と $V[X]$ を求めよ．

演習 3.15　$a > 0$ に対して，区間 $I = (-a, a)$ とおく．

(1)　任意の $t \in I$ に対して，$r_0 + r_1 e^t = 0$ が成り立つならば，$r_0 = r_1 = 0$ が成り立つことを示せ．

(2)　任意の $t \in I$ に対して，$r_0 + r_1 e^t + \cdots + r_n e^{nt} = 0$ が成り立つならば，$r_0 = r_1 = \cdots = r_n = 0$ が成り立つことを示せ．

(3)　二つの離散確率変数 X，Y の確率関数 p_k, q_k がある自然数 n に対して $p_{n+1} = p_{n+2} = \cdots = 0$, $q_{n+1} = q_{n+2} = \cdots = 0$ であると仮定する．このとき，任意の $t \in I$ に対して $M_X(t) = M_Y(t)$ が成り立つならば，$p_0 = q_0, \ldots, p_n = q_n$ が成り立つことを示せ．

4　代表的な確率変数

4.1　2 項分布

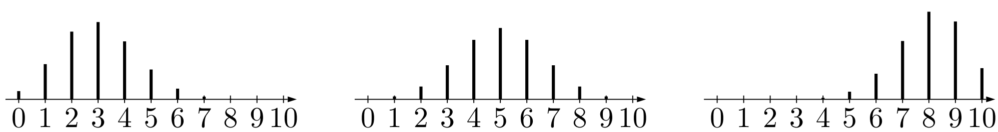
離散確率変数 X の確率関数 $p_k = P(X = k)$ が

$$p_k = \begin{cases} {}_nC_k p^k (1-p)^{n-k} & k = 0, 1, \ldots, n \\ 0 & k = n+1, n+2, \ldots \end{cases}$$

であるとき，X は **2 項分布**に従うといい，$X \sim B(n, p)$ と表す．ただし，$0 < p < 1$.

例 **3.2**(p.14) において説明した製品の抜取検査における不良品の個数 Y の確率関数は 2 項分布の確率関数であり，つまり，$Y \sim B(n, p)$ である．2 項分布に従う確率変数の例として，表が出る確率が p であるコインを n 回振った時の表の回数，内閣支持率が p である有権者の集団から無作為に選んだ n 人の有権者の中の支持者の数など，1 回の観測や測定で起こる確率が p である事象を繰り返し n 回観測した時にその事象が起こる回数 X が 2 項分布に従う．

確率関数 p_k の形状は，下に例示するように，$p = 0.5$ の時は左右対称であるが，そうでない時は歪んでいる．



$B(10, 0.3)$　　$B(10, 0.5)$　　$B(10, 0.8)$

確率関数として満たすべき性質 $\sum_{k=0}^{n} p_k = 1$ が成り立っているのは，予備知識 0.1(6) の 2 項定理において $a = p,\ b = 1 - p$ とすると

$$\sum_{k=0}^{n} {}_nC_k p^k (1-p)^{n-k} = (p + 1 - p)^n = 1$$

となることから確認できる．また，例 **3.14**(p.20) で計算したように，平均 $E[X]$ や分散 $V[X]$，および積率母関数 $M_X(t)$ は次のようになる．

【公式 16】 $X \sim B(n, p)$ の時，$E[X] = np,\ V[X] = np(1-p),\ M_X(t) = (1-p+pe^t)^n$.

確率の計算の簡単な例は，すでに 演習 **3.1**(p.23) でやったが，大きな n に対しては後で述べる中心極限定理を用いるとよい．

4.2　ポアソン分布

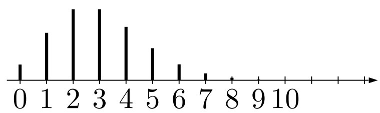
離散確率変数 X の確率関数 $p_k = P(X = k)$ が

$$p_k = \frac{\lambda^k}{k!} e^{-\lambda}, \quad k = 0, 1, 2, \ldots, \lambda > 0$$

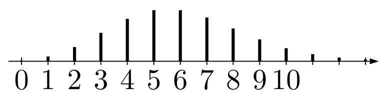
で与えられるとき，X は**ポアソン分布**に従うといい，$X \sim Po(\lambda)$ と表す．

例 **3.17**(p.22) では，積率母関数と確率分布の対応を利用して（ 公式 15(p.21) ），$p = \frac{\lambda}{n}$ である 2 項分布の確率関数が，$n \to \infty$ のとき，$\frac{\lambda^k}{k!}e^{-\lambda}$，つまり，ポアソン分布の確率関数に収束することを示した．$n \to \infty$ の時，$p \to 0$ なので，p が小さく，n が大きい時の 2 項分布がポアソン分布であると解釈できる．つまり，1 回観測しただけではめったに起こらない事象を多数回観測した時に，その事象が起こる回数がポアソン分布に従うことがわかる．たとえば，1 か月の交通事故の件数や，一日の来客数などはポアソン分布に従うと考えられる．

下のグラフが示すように，ポアソン分布は歪んだ分布であるが，λ が大きくなると，分布の山が右に移動することがわかる．



$P_o(3)$



$P_o(6)$

ポアソン分布の積率母関数 $M_X(t)$ は，例 **3.17**(p.22) で求めたように，$M_X(t) = e^{\lambda(e^t-1)}$ となる．公式 14(p.20) より，$M'_X(0) = E[X] = \lambda$, $M''_X(0) = E[X^2] = \lambda^2 + \lambda$ となる．これを使うと，$V[X] = E[X^2] - \{E[X]\}^2 = \lambda$ となることがわかる．

【公式 17】 $X \sim P_o(\lambda)$ のとき，$E[X] = \lambda$, $V[X] = \lambda$, $M_X(t) = e^{\lambda(e^t-1)}$.

4.3 幾何分布

離散確率変数 X の確率関数 $p_k = P(X = k)$ が

$$p_k = qp^k, \quad 0 < p < 1,\ q = 1 - p$$

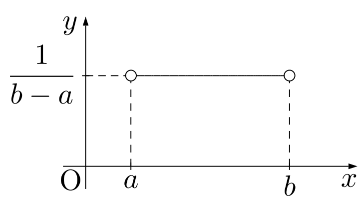
で与えられるとき，X は**幾何分布**に従うといい，$X \sim Ge(p)$ と表す．生起確率が p である事象が起こらなくなるまで観測し続けた時のその事象が起こった回数の分布が $Ge(p)$ である．等比級数の和の公式より，次が導ける (演習 **3.10**(p.24) 参照).

【公式 18】 $X \sim Ge(p)$ の時，$E[X] = \frac{p}{1-p}$, $V[X] = \frac{p}{(1-p)^2}$, $M_X(t) = \frac{1-p}{1-pe^t}$.

4.4 一様分布

連続確率変数 X の p.d.f. $f(x)$ が

$$f(x) = \begin{cases} \dfrac{1}{b-a} & (a < x < b) \\ 0 & (x \leqq a,\ b \leqq x) \end{cases}$$



一様分布 $U(a,b)$ の p.d.f.

で与えられるとき，X は**一様分布**に従うといい，$X \sim U(a,b)$ と表す．一様分布の確率計算の例は，すでに例 **3.4**(p.15) で見た．また，例 **3.9**(p.17)，例 **3.12**(p.19) で考えたように，期待値と分散は以下のようになる．

【公式 19】 $X \sim U(a,b)$ の時，$E[X] = \frac{a+b}{2}$, $V[X] = \frac{(a-b)^2}{12}$.

4.5 ベータ分布

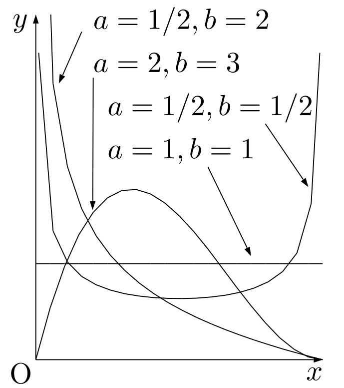
連続確率変数 X の p.d.f. $f(x)$ が

$$f(x) = \begin{cases} \frac{1}{B(a,b)} x^{a-1}(1-x)^{b-1} & (0 < x < 1) \\ 0 & (x \leqq 0) \end{cases}$$

で与えられるとき，X は**ベータ分布**に従うといい，$X \sim Be(a,b)$ と表す．ただし，$a > 0, b > 0$ で，$B(a,b)$ は

$$B(a,b) = \int_0^1 t^{a-1}(1-t)^{b-1} dt$$

で定義されるベータ関数と呼ばれる関数である．$B(1,1) = U(0,1)$ なので，ベータ分布は一様分布 $U(0,1)$ を含む分布族であり，右上のグラフからわかるように a, b の値を変化させると色々な形状の p.d.f. になることがわかる．平均と分散は次のようになることが知られている (演習 **4.10**(p.32)):



ベータ分布 $Be(a,b)$ の p.d.f.

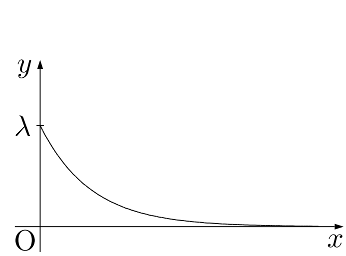
$$E[X] = \frac{a}{a+b}, \quad V[X] = \frac{ab}{(a+b)^2(a+b+1)}$$

4.6 指数分布

連続確率変数 X の p.d.f. $f(x)$ が

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & (0 < x) \\ 0 & (x \leqq 0) \end{cases}, \quad \lambda > 0,$$

で与えられるとき，X は**指数分布**に従うといい，$X \sim Ex(\lambda)$ と表す．



指数分布 $Ex(\lambda)$ の p.d.f.

連続確率変数の例として，例 **3.5**(p.16) で考えた蛍光灯の寿命 X はまさに $Ex(\lambda)$ に従う確率変数で，その期待値や分散は 例 **3.10**(p.17)，例 **3.13**(p.20) で求めた．また，例 **3.15**(p.21) では，積率母関数を求め，そこから平均と分散を求めた．

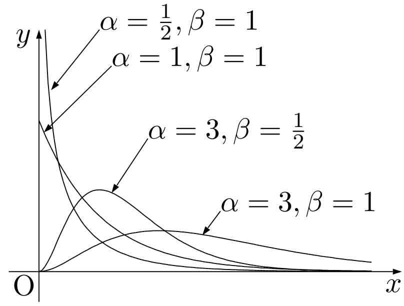
【公式 20】 $X \sim Ex(\lambda)$ の時，$E[X] = \frac{1}{\lambda}$, $V[X] = \frac{1}{\lambda^2}$, $M_X(t) = \frac{\lambda}{\lambda - t}$.

4.7 ガンマ分布

連続確率変数 X の p.d.f. $f(x)$ が

$$f(x) = \begin{cases} \dfrac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-x/\beta} & (x > 0) \\ 0 & (x \leqq 0) \end{cases}$$

で与えられるとき，X は**ガンマ分布**に従うといい，$X \sim Ga(\alpha, \beta)$ と表す．ただし，$\alpha > 0, \beta > 0$ であり，$\Gamma(\alpha)$ はガンマ関数と呼ばれる α の関数で，次のように定義される：



ガンマ分布 $Ga(\alpha, \beta)$ の p.d.f.

$$\Gamma(\alpha) = \int_0^\infty u^{\alpha-1} e^{-u} du, \quad \alpha > 0.$$

ガンマ分布の密度関数 $f(x)$ は右上図のように，α と β の値によって，様々な形状になるため，正の値しかとらない測定値のモデルとして応用範囲が広い．特に，$\alpha = 1$, $\beta = 1/\lambda$ の時の密度関数は $Ex(\lambda)$ のそれと一致するので，指数分布はガンマ分布の特別な場合であると考えられる．

べき乗 x^{a-1} と指数関数 $e^{-x/b}$ の積 $x^{a-1}e^{-x/b}$ の積分は，ガンマ関数で表すことができる．

【公式 21】 $\displaystyle\int_0^\infty x^{a-1} e^{-x/b} dx = b^a \Gamma(a)$, $b > 0$.

〔証明〕右辺において，$u = x/b$ と変数変換すると，$x = bu$, $dx = bdu$ なので

$$\int_0^\infty x^{a-1} e^{-x/b} dx = \int_0^\infty (bu)^{a-1} e^{-u} bdu = b^a \int_0^\infty u^{a-1} e^{-u} du = b^a \Gamma(a).$$

この公式 21 で，$a = \alpha$, $b = \beta$ として，両辺 $\Gamma(\beta)\beta^\alpha$ で割ると，ガンマ分布の p.d.f. $f(x)$ が確率密度関数の性質 $\int_{-\infty}^\infty f(x)dx = 1$ を満たすことが確認できる．また，

$$M_X(t) = E[e^{tX}] = \int_0^\infty e^{tx} \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-x/\beta} dx = \frac{1}{\Gamma(\alpha)\beta^\alpha} \int_0^\infty x^{\alpha-1} e^{-\frac{1-t\beta}{\beta}x} dx$$

だから，$\beta t < 1$ のとき，公式 21　で $a = \alpha$, $b = \frac{\beta}{1-t\beta}$ とおくと，$M_X(t) = \frac{\Gamma(\alpha)b^\alpha}{\Gamma(\alpha)\beta^\alpha} = (b/\beta)^\alpha = (1-\beta t)^{-\alpha}$ が導ける．演習 **3.8**(p.23) で計算したように，この積率母関数から，$E[X] = \alpha\beta$，$V[X] = \alpha\beta^2$ が導ける．

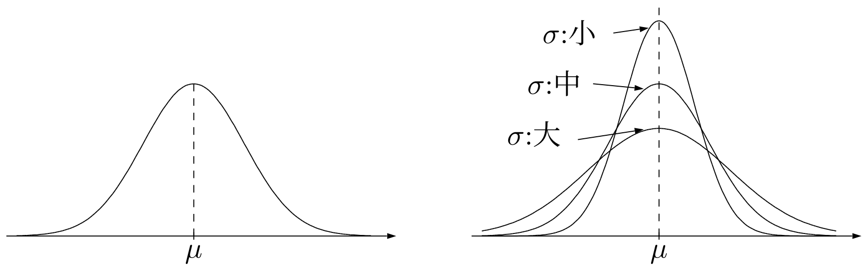
【公式 22】 $X \sim Ga(\alpha, \beta)$ の時，$E[X] = \alpha\beta$，$V[X] = \alpha\beta^2$，$M_X(t) = (1-\beta t)^{-\alpha}$.

4.8 正規分布

連続確率変数 X の p.d.f. $f(x)$ が

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad -\infty < \mu < \infty,\ \sigma > 0$$

で与えられるとき，X は**正規分布**に従うといい，$X \sim N(\mu, \sigma^2)$ と表す．左下図のように，$f(x)$ は直線 $x = \mu$ に関して対称であり，μ の大小により右左へ平行移動する．また，右下図のように，σ が大きいと左右に広がり，σ が小さいと中心が高くなる．



$N(\mu, \sigma^2)$ の p.d.f. $f(x)$　　σ の大小による $f(x)$ の違い

正規分布の確率密度関数が満たすべき次の公式が成り立つ．

【公式 23】 任意の $s > 0$ と任意の実数 m に対して，$\displaystyle\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi s^2}} e^{-\frac{1}{2s^2}(x-m)^2} dx = 1.$

この公式は理系の分野全般でよく知られている．証明は，演習 **4.8**(p.32) を参照．この公式で，$s = \sigma,\ m = \mu$ と置くと，$\int_{-\infty}^{\infty} f(x)dx = 1$ がわかる．また，

$$M_X(t) = E[e^{tx}] = \int_{-\infty}^{\infty} e^{tx} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{tx - \frac{(x-\mu)^2}{2\sigma^2}} dx$$

であり，指数関数の肩の部分を x に関して平方完成すると，

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

となるので，

$$M_X(t) = e^{\mu t + \frac{\sigma^2}{2}t^2} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}\left\{x - (\mu + \sigma^2 t)\right\}^2} dx.$$

公式 23 で $s = \sigma,\ m = \mu + \sigma^2 t$ と置くと，最後の項の積分は 1 となるので，$M_X(t) = \exp\left(\mu t + \frac{\sigma^2}{2}t^2\right)$ が導ける．演習 **3.9**(p.24) で求めたように，この積率母関数から $E[X] = \mu,\ V[X] = \sigma^2$ が得られる．

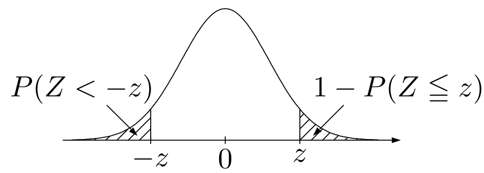
【公式 24】 $X \sim N(\mu, \sigma^2)$ のとき，$E[X] = \mu$, $V[X] = \sigma^2$, $M_X(t) = \exp\left(\mu t + \frac{\sigma^2}{2}t^2\right)$.

$Z \sim N(0,1)$ の時，Z は**標準正規分布**に従うという．この時，$P(Z \geqq z) = \alpha$ となる z を**上側** 100α **パーセント点**と呼び，$z(\alpha)$ と表す．応用上よく用いる α に対する $z(\alpha)$ の値は，**付表.2** に与えられている．また，$P(Z \leqq z) = \alpha$ となる z を**下側** 100α **パーセント点**と呼ぶが，正の z と $P(Z \leqq z)$ の関係は，**付表.1** のようになる．

【公式 25】 $Z \sim N(0,1)$ のとき，

$$P(Z < -z) = P(Z > z) = 1 - P(Z \leqq z)$$

であり，特に，$P(Z < 0) = P(Z > 0) = \frac{1}{2}$.



標準正規分布 $N(0,1)$

〔証明〕Z の p.d.f. は $x = 0$ で対称で，右図のようになる．図から結果は明らか．

《例 **4.1**》 **付表.1** より，$P(Z \leqq 1.56) = 0.9406$ であり，$P(Z \leqq z) = 0.9864$ となる $z = 2.21$ であることがわかる．また， 公式 25 より，$P(Z < -1.56) = 1 - P(Z \leqq 1.56) = 1 - 0.9406 = 0.0594$ であり，

$$P(-1.56 \leqq Z \leqq 2.21) = 0.9864 - 0.0594 = 0.927$$

と求められる．さらに，$P(Z \leqq z) = 0.99$ となる z は**付表.1** には載っていないが，0.99 に最も近い値について調べると，$P(Z \leqq 2.32) = 0.9898$ と $P(Z \leqq 2.33) = 0.9901$ であることがわかる．これらを用いて，

$$z = 2.32 + (2.33 - 2.32) \times \frac{0.99 - 0.9898}{0.9901 - 0.9898} = 2.32666\cdots\cdots$$

のように比例配分して，$P(Z \leqq z) = 0.99$ となる z の近似値を求めることができる．あるいは，$P(Z \leqq z) = 0.99$ を満たす z は下側 99% 点であるが，言い換えると $P(Z \geqq z) = 0.01$ を満たす上側 1% 点，つまり，$z = z(0.01)$ であることがわかるので，**付表.2** より，$z = z(0.01) = 2.326$ がわかる．

【公式 26】 $X \sim N(\mu, \sigma^2)$ の時，$aX + b \sim N(a\mu + b, a^2\sigma^2)$．特に，$(X - \mu)/\sigma \sim N(0,1)$.

〔証明〕$Y = aX + b$ とおくと，Y の積率母関数 $M_Y(t)$ は

$$\begin{aligned} M_Y(t) =& E[e^{t(aX+b)}] = e^{bt}E[e^{(at)X}] = e^{bt}M_X(at) \\ =& e^{bt}e^{\mu(at) + \frac{\sigma^2}{2}(at)^2} = e^{(a\mu+b)t + \frac{(a\sigma)^2}{2}t^2}. \end{aligned}$$

これは，$N(a\mu + b, a^2\sigma^2)$ に従う確率変数の積率母関数と一致するので，Y の分布も $N(a\mu + b, a^2\sigma^2)$ であることがわかる．$(X - \mu)/\sigma = \frac{1}{\sigma}X - \frac{\mu}{\sigma}$ なので，$a = \frac{1}{\sigma}$, $b = -\frac{\mu}{\sigma}$ とすると，$a\mu + b = 0$, $a^2\sigma^2 = 1$ となり，$(X - \mu)/\sigma \sim N(0,1)$ がわかる．

《例 **4.2**》 $X \sim N(2,16)$ の時，$Z=(X-2)/4 \sim N(0,1)$ だから，

$$P(X \leqq 3)=P\left(\frac{X-2}{4} \leqq \frac{1}{4}\right)=P\left(Z \leqq 0.25\right)=0.5987$$

となる．また，Y の積率母関数が $M_Y(t)=e^{3t+2t^2}$ の時，これは $N(3,4)$ の積率母関数と一致するので，$Y \sim N(3,4)$.

《例 **4.3**》 高校生男子の身長の度数分布は，$N(170,36)$ の確率密度関数で近似できるものとする．無作為に選んだ高校生男子の身長を X とする．このとき，$P(167 \leqq X \leqq 176)$ である確率は，$Z=(X-170)/\sqrt{36}=(X-170)/6 \sim N(0,1)$ であることを利用すると，

$$P(167 \leqq X \leqq 176)=P\left(\frac{167-170}{6} \leqq Z \leqq \frac{176-170}{6}\right)=P(-0.5 \leqq Z \leqq 1)$$
$$=P(Z \leqq 1)-(1-P(Z \leqq 0.5))=0.8413-(1-0.6915)=0.5328$$

と求められる．

演習問題

演習 4.1 $Z \sim N(0,1)$ のとき，

(1) $P(Z \leqq 1.96)$ の値，$P(Z \leqq z)=0.9901$ を満たす z を**付表.1** から求めよ．また，$z(0.05)$，$z(0.025)$，$z(0.005)$ の値を**付表.2** から求め，さらに，$P(|Z| \leqq z)=0.99$，$P(Z>w)=0.95$，$P(|Z|>x)=0.05$ となる z, w, x の値を求めよ．

(2) $X=3Z+2$ のとき，X が従う分布を答えよ．また，$E[X], V[X]$ を答え，$E[X^2]$ を求めよ．

演習 4.2 ある工業製品の組み立て時間 (分) は，$N(20,9)$ に従うことが知られている．組立時間が 15 分 30 秒以上，24 分 30 秒以下である確率を求めよ．

演習 4.3 センター試験の英語の得点について，その相対度数のヒストグラムを描くと，平均 μ，分散 30^2 の正規分布 $N(\mu,30^2)$ の p.d.f. $f(x)$ で近似できると考えられている．150 点の人が上位 10 パーセントに入るとき，μ がとる値の範囲を求めよ．

演習 4.4 Y の積率母関数 $M_Y(t)$ が $M_Y(t)=e^{-2t+8t^2}$ のとき，Y の従う確率分布を答えよ．また，$P(-1 \leqq Y \leqq 2)$ を求めよ．

演習 4.5 $X \sim B(n,p)$ のとき，$E[X]$ と $E[X(X-1)]$ を定義に従って求めよ．これらより，$V[X]$ を求めよ．ヒント：$\sum_{k=0}^{n} kp_k=\sum_{k=1}^{n} kp_k=\sum_{k'=0}^{n-1}(k'+1)p_{k'+1}$ として，

$(a{+}b)^{n-1}$ を2項定理で展開した式を使う．また，$\sum_{k=0}^{n} k(k{-}1)p_k = \sum_{k=2}^{n} k(k{-}1)p_k = \sum_{k'=0}^{n-2}(k'+2)(k'+1)p_{k'+2}$ と変形して，$(a+b)^{n-2}$ を 2 項定理で展開した式を使う．

演習 4.6　$X \sim P(\lambda)$ のとき，$E[X]$ と $E[X(X-1)]$ を定義に従って求めよ．これらより，$V[X]$ を求めよ．ヒント：演習 **4.5** と同様に変形して，2 項定理の代わりに指数関数のテイラー展開を使う．

演習 4.7　$X \sim Ga(\alpha,\beta)$ のとき，$E[X]$, $E[X^2]$ を定義にしたがって求めよ．ヒント：公式 21(p.28) と演習 **4.9** の (1) を使え．

演習 4.8

(1) $\displaystyle\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi s^2}} e^{-\frac{(x-m)^2}{2s^2}} dx = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz$ を示せ.

(2) $z = r\cos\theta$, $w = r\sin\theta$ という変数変換を使って，次を示せ.

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-\frac{z^2+w^2}{2}} dzdw = \int_{0}^{\infty}\int_{0}^{2\pi} re^{-\frac{r^2}{2}} drd\theta.$$

(3) $\displaystyle\frac{d}{dr}e^{-\frac{r^2}{2}}$ を求めよ．これを利用して，上の右辺の積分の値を求めよ.

(4) 公式 23(p.29) を証明せよ．ヒント：(2) の左辺は $\left\{\int_{-\infty}^{\infty} e^{-\frac{z^2}{2}} dx\right\}^2$ である.

演習 4.9

(1) $\Gamma(1) = 1$ を示せ．また，$\Gamma(\alpha+1) = \alpha\,\Gamma(\alpha)$ を示せ．ヒント：部分積分の公式.

(2) $\displaystyle\int_{0}^{\infty} e^{-\frac{x^2}{2}} dx = \sqrt{\frac{\pi}{2}}$ を示せ．ヒント：公式 23(p.29) を使え.

(3) $\Gamma\left(\dfrac{1}{2}\right) = \sqrt{\pi}$ を示せ．ヒント: 上の積分において，$u = \frac{x^2}{2}$ と変数変換する.

演習 4.10

(1) $B(a,b) = \dfrac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$ を示せ．ヒント：$\Gamma(a)\Gamma(b)$ を 2 重積分で表現して，適当な変数変換を行え.

(2) $X \sim Be(a,b)$ の時，$E[X]$, $V[X]$ を求めよ.

5 多変量確率変数とその期待値

5.1 2 変量離散確率変数

2 つの変数 X, Y の取り得る値がどちらも非負整数 $\{0, 1, 2, \ldots,\}$ であり，非負整数の組み (k, j) すべてに対して，確率 $P(X = k, Y = j)$ が定義されているとき，変数の組 (X, Y) を **2 変量離散確率変数**と呼び，$p(k, j) = P(X = k, Y = j)$ を**同時確率関数**と呼ぶ．2 変量確率変数 (X, Y) の同時確率関数が $p(k, j)$ であることを簡単に $(X, Y) \sim p(k, j)$ と表す．

$X = k$ であるが，Y は非負整数のどれかである確率 $P(X = k,\ Y = 0, 1, 2, \ldots)$ を簡単に $P(X = k)$，あるいは，$p_X(k)$ と表し，X の**周辺確率関数**と呼ぶ．この時，

$$\begin{aligned} &P(X = k, Y = 0, 1, 2, \ldots) \\ =&P(X = k, Y = 0) + P(X = k, Y = 1) + P(X = k, Y = 2) + \cdots \end{aligned}$$

が成り立つので，

$$p_X(k) = \sum_{j=0}^{\infty} p(k, j)$$

のように，同時確率関数 $p(k, j)$ から周辺確率関数 $p_X(k)$ を求めることができる．また，$X = k$ となる場合に限定した $Y = j$ となる確率 $\frac{P(X=k,Y=j)}{P(X=k)} = \frac{p(k,j)}{p_X(k)}$ を $X = k$ の下での Y の**条件付き確率関数**と呼び，$p_{Y|X}(j|k)$ と表す．つまり，

$$p_{Y|X}(j|k) := \frac{p(k, j)}{p_X(k)}.$$

また，

$$\sum_{j=0}^{\infty} p_{Y|X}(j|k) = \sum_{j=0}^{k} \frac{p(k, j)}{p_X(k)} = \frac{p_X(k)}{p_X(k)} = 1$$

が成り立つので，条件付き確率関数 $p_{Y|X}(j|k)$ は確率関数としての性質を満たすことがわかる．Y の周辺確率関数 $P(Y = j) = p_Y(j)$ や $Y = j$ の下での X の条件付き確率関数 $p_{X|Y}(k|j)$ も同様に定義する．

任意の非負整数の組 (k, j) に対して，$p(k, j) = p_X(k)p_Y(j)$ が成り立つ時，X と Y は**独立**であるといい，$X \perp\!\!\!\perp Y$ と表す:

$$X \perp\!\!\!\perp Y \overset{def}{\iff} p(k, j) = p_X(k)p_Y(j).$$

また，$X \perp\!\!\!\perp Y$ の時，$p(k, j) = p_X(k)p_Y(j)$ だから，

$$p_{X|Y}(k|j) = \frac{p_X(k)p_Y(j)}{p_Y(j)} = p_X(k),\ \ p_{Y|X}(j|k) = \frac{p_X(k)p_Y(j)}{p_X(k)} = p_Y(j)$$

であり，逆に $p_{X|Y}(k|j) = p_X(k)$ のとき，

$$p(k,j) = p_{X|Y}(k|j)p_Y(j) = p_X(k)p_Y(j)$$

なので，$X \perp\!\!\!\perp Y$ である．同様に，$p_{Y|X}(j|k) = p_Y(j)$ の時，$X \perp\!\!\!\perp Y$ も導ける．以上より，

$$X \perp\!\!\!\perp Y \iff p_{X|Y}(k|j) = p_X(k) \iff p_{Y|X}(j|k) = p_Y(j)$$

がわかる．

《例 **5.1**》 下の表は，例 **3.1**(p.14) で考えた地域の家族数の相対度数分布を東部と西部，あるいは，都市部とそれ以外に層別したものである．その地域から無作為に 1 世帯を選んだ時，その世帯の家族数を X とし，それが東部であれば $Y = 0$, 西部であれば $Y = 1$ とする．また，都市部であれば $Z = 0$，都市部以外であれば $Z = 1$ とする．

このとき，X と Y の同時確率関数 $p_{XY}(k,j)$ は下の相対度数分布の上半分であり，X と Z の同時確率関数 $p_{XZ}(k,j)$ は下半分である．また，X の周辺分布は $P(X = 1) = p_X(1) = 0.35$, $P(X = 2) = p_X(2) = 0.25\ldots$ などとなり，Y の周辺分布は $P(Y = 0) = p_Y(0) = 0.4$, $P(Y = 1) = p_Y(1) = 0.6$，Z の周辺分布は $P(Z = 0) = p_Z(0) = 0.6$, $P(Z = 1) = p_Z(1) = 0.4$ である．明らかに，$p_{XY}(k,j) = p_X(k)p_Y(j)$ が成り立つが，$p_{XZ}(k,j) = p_X(k)p_Z(j)$ は成り立たないので，X と Y は独立であり，X と Z は独立でない．表から $p_{X|Y}(k|j) = p_X(k)$ が容易に確認でき，また，$p_{X|Z}(1|0) = 0.3 \div 0.6 = 0.5$, $p_{X|Z}(1|1) = 0.05 \div 0.4 = 0.125$ などが求められる．

家族数 (X)	1	2	3	4	5	小計
東 ($Y = 0$)	0.14	0.1	0.08	0.06	0.02	0.4
西 ($Y = 1$)	0.21	0.15	0.12	0.09	0.03	0.6
都市部 ($Z = 0$)	0.3	0.18	0.07	0.04	0.01	0.6
その他 ($Z = 1$)	0.05	0.07	0.13	0.11	0.04	0.4
小計	0.35	0.25	0.2	0.15	0.05	1

《例 **5.2**》 0 の球が 1 個，1 の球が 3 個，2 の球が 2 個入った壺 (⓪, ①, ①, ①, ②, ②) から無作為に取り出した球の数字を X とすると，$p_X(0) = 1/6$, $p_X(1) = 1/2$, $p_X(2) = 1/3$ である．取り出した球を戻して，もう一度取り出した球の数字を Y とすると，$p_X(k) = p_Y(k)$ であり，X と Y は独立だから $p(k,j) = p_X(k)p_Y(j)$ が成り立つ．

$j \backslash k$	0	1	2	$p_Y(j)$
0	0	1/10	1/15	1/6
1	1/10	1/5	1/5	1/2
2	1/15	1/5	1/15	1/3
$p_X(k)$	1/6	1/2	1/3	1

1 つ目を戻さないで，もう 1 つ取り出した場合は，$p(0,0) = \frac{1}{6} \times 0 = 0$, $p(0,1) = \frac{1}{6} \times \frac{3}{5} = \frac{1}{10}$ などとなり，他の組み合わせの同時確率 $p(k,j)$ は左表のようになる．縦方向の小計が $p_X(k)$, 横方向のが $p_Y(j)$ であり，

$p(k,j) = p_X(k)p_Y(j)$ が成立しない．つまり，X と Y は独立ではない．表より条件付き確率関数 $p_{X|Y}(0|2) = 1/15 \div (1/3) = 1/5$, $p_{X|Y}(1|2) = 1/5 \div (1/3) = 3/5$, $p_{X|Y}(2|2) = 1/15 \div (1/3) = 1/5$ などが容易に求められる．

5.2 2 変量連続確率変数

X と Y のとりうる値がどちらも実数全体 $(-\infty, \infty)$ であり，$c_1 \leqq X \leqq c_2$ かつ $d_1 \leqq Y \leqq d_2$ である確率が,

$$P(c_1 \leqq X \leqq c_2,\ d_1 \leqq Y \leqq d_2) = \int_{c_1}^{c_2} \int_{d_1}^{d_2} f(x,y)dxdy$$

と定義される時，(X,Y) を **2 変量連続確率変数**と呼ぶ．ここで，$f(x,y)$ は, 非負で $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)dxdy = 1$ を満たす関数であり，**同時確率密度関数** (**同時 p.d.f.**) と呼ばれる．(X,Y) が同時確率密度関数 $f(x,y)$ を持つ 2 変量連続確率変数であることを, $(X,Y) \sim f(x,y)$ と表す.

離散確率変数の場合と同様に,

$$f_X(x) := \int_{-\infty}^{\infty} f(x,y)dy, \quad f_{Y|X}(y|x) := \frac{f(x,y)}{f_X(x)}$$

で定義される $f_X(x)$ を X の**周辺確率密度関数**，$f_{Y|X}(y|x)$ を $X = x$ の下での Y の**条件付き確率密度関数**と呼ぶ．$P(c_1 \leqq X \leqq c_2, -\infty < Y < \infty)$ を簡単に $P(c_1 \leqq X \leqq c_2)$ と表すと,

$$P(c_1 \leqq X \leqq c_2) = \int_{c_1}^{c_2} \left\{ \int_{-\infty}^{\infty} f(x,y)dy \right\} dx = \int_{c_1}^{c_2} f_X(x)dx$$

なので，$f_X(x)$ は X の確率密度関数であることがわかる．また,

$$\int_{-\infty}^{\infty} f_{Y|X}(y|x)dy = \frac{1}{f_X(x)} \int_{-\infty}^{\infty} f(x,y)dy = \frac{f_X(x)}{f_X(x)} = 1$$

なので，$f_{Y|X}(y|x)$ は確率密度関数の性質を満たすことがわかる．Y の周辺確率密度関数 $f_Y(y)$，$Y = y$ の下での X の条件付き確率密度関数 $f_{X|Y}(x|y)$ も同様に定義する．

任意の実数 x, y に対して，$f(x,y) = f_X(x)\,f_Y(y)$ が成り立つ時，X と Y は**独立**であると言い，$X \perp\!\!\!\perp Y$ と表す:

$$X \perp\!\!\!\perp Y \ \overset{def}{\Longleftrightarrow}\ f(x,y) = f_X(x)f_Y(y).$$

「X と Y が独立であること」と「条件付き確率が周辺確率と等しいこと」が同値であるのは，離散確率変数の場合と同様に証明できる：

$$X \perp\!\!\!\perp Y \iff f_{X|Y}(x|y) = f_X(x) \iff f_{Y|X}(y|x) = f_Y(y).$$

《例 **5.3**》 $0<x<1,\ 0<y<1$ の時，$f(x,y)=\frac{1}{5}(1+3x+2y+6xy)$ であり，それ以外の時は $f(x,y)=0$ である同時確率密度関数 $f(x,y)$ を持つ確率変数 (X,Y) に対して，

$$P\left(\frac{1}{3}<X<\frac{2}{3},\ 0<Y<\frac{1}{2}\right)=\int_{1/3}^{2/3}\int_0^{1/2}\frac{1}{5}(1+3x+2y+6xy)dxdy$$
$$=\int_{1/3}^{2/3}\frac{1}{5}\left[(1+3x)y+y^2+3xy^2\right]_0^{\frac{1}{2}}dx=\int_{1/3}^{2/3}\frac{1}{5}\left(\frac{(1+3x)}{2}+\frac{1}{4}+\frac{3}{4}x\right)dx=\frac{1}{8}$$

である．また，

$$f_X(x)=\int_0^1\frac{1}{5}(1+3x+2y+6xy)dy=\frac{1}{5}\left[(1+3x)y+y^2+3xy^2\right]_0^1=\frac{2}{5}(3x+1),$$
$$f_Y(y)=\int_0^1\frac{1}{5}(1+3x+2y+6xy)dx=\frac{1}{5}\left[(2y+1)x+\frac{3x^2}{2}+3x^2y\right]_0^1=\frac{1}{2}(2y+1)$$

である．$f(x,y)=\frac{1}{5}(3x+1)(2y+1)$ と因数分解できるので，$f(x,y)=f_X(x)f_Y(y)$ が成り立ち，X と Y は独立であることがわかる．

《例 **5.4**》 成人男子の身長の度数分布は $N(171,5.5^2)$ で，体重は $N(68,9^2)$ で表せるものとする．無作為に選んだ成人男子の身長を X, 体重を Y とすると，$X\sim N(171,5.5)$, $Y\sim N(68,9)$ であり，$Z=(X-171)/5.5$, $W=(Y-68)/9$ と標準化すると公式 26(p.30) より $Z\sim N(0,1)$, $W\sim N(0,1)$ となる．さらに，Z, W の同時確率密度関数は $|r|<1$ なる r を用いて，

$$f(z,w)=\frac{1}{2\pi\sqrt{1-r^2}}\exp\left\{-\frac{z^2-2rzw+w^2}{2(1-r^2)}\right\}$$

と表せるものとする．このとき，(Z,W) は 2 変量正規分布に従うという．また，(X,Y) も 2 変量正規分布に従うという．さて，(Z,W) の同時確率密度関数は

$$f(z,w)=\frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{w^2}{2}\right\}\frac{1}{\sqrt{2\pi(1-r^2)}}\exp\left\{-\frac{(z-rw)^2}{2(1-r^2)}\right\}\cdots(※)$$

と変形できて，公式 23(p.29) において $m=rw$, $s=\sqrt{1-r^2}$, $x=z$ とすると，

$$\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi(1-r^2)}}\exp\left\{-\frac{(z-rw)^2}{2(1-r^2)}\right\}dz=1$$

となるので，$f_W(w)=\int_{-\infty}^{\infty}f(z,w)dz=\frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{w^2}{2}\right\}$ が導ける．つまり，W の周辺分布が $N(0,1)$ であることが示せる．さらに，(※) より

$$f_{Z|W}(z|w)=\frac{f(z,w)}{f_W(w)}=\frac{1}{\sqrt{2\pi(1-r^2)}}\exp\left\{-\frac{(z-rw)^2}{2(1-r^2)}\right\},$$

つまり，$W=w$ の下での Z の条件付き分布は $N(rw, 1-r^2)$ であることがわかる．特に，$r=0$ の時，$f_{Z|W}(z|w)=\frac{1}{\sqrt{2\pi}}e^{-\frac{z^2}{2}}$ なので，$f_{Z|W}(z|w)=f_Z(z)$ であり，Z と W が独立であることがわかる．

5.3 期待値と共分散

2 変量確率変数 (X,Y) と関数 $h(x,y)$ に対して，

$$E[h(X,Y)]:=\begin{cases}\displaystyle\sum_{k=0}^{\infty}\sum_{j=0}^{\infty}h(k,j)p(k,j) & \text{(離散の時)}\\ \displaystyle\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}h(x,y)f(x,y)dxdy & \text{(連続の時)}\end{cases}$$

と定義し，$h(X,Y)$ の**期待値**と呼ぶ．さらに，関数 $g(x,y)$ に対して，

$$Cov[h(X,Y),g(X,Y)]:=E[\{h(X,Y)-\mu_h\}\{g(X,Y)-\mu_g\}]$$

と定義して，$h(X,Y)$ と $g(X,Y)$ の**共分散**と呼ぶ．ここで，$\mu_h=E[h(X,Y)]$, $\mu_g=E[g(X,Y)]$ である．さらに，

$$V[h(X,Y)]:=Cov[h(X,Y),h(X,Y)].$$

と定義される $V[h(X,Y)]$ を $h(X,Y)$ の**分散**と呼ぶ．特に，$Cov[X,Y]=:\sigma_{XY}$ と表し，X と Y の**共分散**と呼び，

$$r[X,Y]:=\frac{\sigma_{XY}}{\sigma_X\sigma_Y}$$

を X と Y の**相関係数**と呼ぶ．ただし，$\sigma_X=\sqrt{V[X]}$, $\sigma_Y=\sqrt{V[Y]}$ である．どんな確率変数 (X,Y) の相関係数 $r[X,Y]$ でも，$-1\leqq r[X,Y]\leqq 1$ を満足することに注意しよう．

【公式 27】

$$E[h(X)]=\begin{cases}\displaystyle\sum_{k=0}^{\infty}h(k)p_X(k) & (X,Y)\sim p(k,j)\ \text{のとき}\\ \displaystyle\int_{-\infty}^{\infty}h(x)f_X(x)dx & (X,Y)\sim f(x,y)\ \text{のとき}\end{cases}$$

〔証明〕$(X,Y)\sim p(k,j)$ の時に証明する．

$$E[h(X)]=\sum_{k=0}^{\infty}\sum_{j=0}^{\infty}h(k)p(k,j)=\sum_{k=0}^{\infty}h(k)\left(\sum_{j=0}^{\infty}p(k,j)\right)=\sum_{k=0}^{\infty}h(k)p_X(k).$$

条件付き確率関数 $p_{X|Y}(k|y)$ や確率密度関数 $f_{X|Y}(x|y)$ による期待値

$$E[h(X,Y)) \mid Y=y] = \begin{cases} \displaystyle\sum_{k=0}^{\infty} h(k,y)p_{X|Y}(k|y) & (X,Y) \sim p(k,j) \text{ のとき} \\ \displaystyle\int_{-\infty}^{\infty} h(x,y)f_{X|Y}(x|y)dx & (X,Y) \sim f(x,y) \text{ のとき} \end{cases}$$

を $Y=y$ の下での $h(X,Y)$ の**条件付き期待値**と呼ぶ．$X=x$ の下での $g(X,Y)$ の条件付き期待値も同様に $p_{Y|X}(j|x)$ や $f_{Y|X}(y|x)$ による期待値として定義する．$E[h(X,Y)|\ Y=y]$ は和や積分をする時，X, Y の同時確率から決まる条件付き確率を用いるが，結果は y にのみ依存する，つまり，y の関数である．仮に $H(y) := E[h(X,Y)|\ Y=y]$ とおいて，この関数の y に確率変数 Y を代入したものを簡単に $E[h(X,Y)|Y]$ と表すことにする．つまり，$E[h(X,Y)|\ Y] = H(Y)$ であり，$E[h(X,Y)|\ Y]$ は $p_{X|Y}(k|Y)$ または $f_{X|Y}(x|Y)$ による $h(X,Y)$ の期待値である．

《例 **5.5**》　例 **5.2**(p.34) では，1 つ目を戻しても戻さなくても $p_X(0)=p_Y(0)=1/6$, $p_X(1)=p_Y(1)=1/2$, $p_X(2)=p_Y(2)=1/3$ であった．よって，公式 27(p.37) より，

$$E[X] = \sum_{k=0}^{2} kp_X(k) = 0 \times \frac{1}{6} + 1 \times \frac{1}{2} + 2 \times \frac{1}{3} = \frac{7}{6}.$$

$$E[X^2] = \sum_{k=0}^{2} k^2 p_X(k) = 0^2 \times \frac{1}{6} + 1^2 \times \frac{1}{2} + 2^2 \times \frac{1}{3} = \frac{11}{6}.$$

分散公式 (公式 11(p.18)) より，$V[X] = \frac{11}{6} - \left(\frac{7}{6}\right)^2 = \frac{17}{36}$．$p_X(k) = p_Y(k)$ なので，$E[Y] = \frac{7}{6}$, $E[Y^2] = \frac{11}{6}$, $V[Y] = \frac{17}{36}$．さらに，1 つ目を戻さない場合は

$$E[XY] = 0 \times 0 \times 0 + 0 \times 1 \times \frac{1}{10} + 0 \times 2 \times \frac{1}{15} + \cdots + 2 \times 2 \times \frac{1}{15} = \frac{19}{15}.$$

$$Cov[X,Y] = (0-\frac{7}{6}) \times (0-\frac{7}{6}) \times 0 + \cdots + (2-\frac{7}{6}) \times (2-\frac{7}{6}) \times \frac{1}{15} = -\frac{17}{180}.$$

$$r[X,Y] = \frac{-\frac{17}{180}}{\sqrt{\frac{17}{36} \times \frac{17}{36}}} = -\frac{1}{5}.$$

また，$p_{X|Y}(0|0)=0$, $p_{X|Y}(1|0)=3/5$, $p_{X|Y}(2|0)=2/5$ なので

$$E[X|\ Y=0] = 0 \times 0 + 1 \times \frac{3}{5} + 2 \times \frac{2}{5} = \frac{7}{5}.$$

《例 **5.6**》　例 **5.3**(p.36) の確率変数 (X,Y) に対して，$f_X(x) = \frac{2}{5}(3x+1)$, $f_Y(y) =$

$\frac{1}{2}(2y+1)$ であった．よって，公式 27(p.37) より，

$$E[X]=\int_0^1 xf_X(x)dx=\int_0^1 x\times\frac{2}{5}(3x+1)\,dx=\frac{3}{5},$$
$$E[Y]=\int_0^1 yf_Y(y)dy=\int_0^1 y\times\frac{1}{2}(2y+1)\,dy=\frac{7}{12},$$

同様に，$E[X^2]=\int_0^1 x^2f_X(x)dx=\frac{13}{30}$，$E[Y^2]=\int_0^1 y^2f_Y(y)dy=\frac{5}{12}$．また，分散公式 (公式 11(p.18)) を使うと，$V[X]=E[X^2]-(E[X])^2=\frac{11}{150}$, $V[Y]=E[Y^2]-(E[Y])^2=\frac{11}{144}$．さらに，

$$E[XY]=\int_0^1\int_0^1 xy\cdot\frac{1}{5}(1+3x+2y+6xy)dxdy=\int_0^1\frac{3}{10}(2y^2+y)dy=\frac{7}{20}.$$

さらに，X と Y は独立だったので，$f_{X|Y}(x|y)=f_X(x)$ であり，$E[X|Y=y]=E[X]=\frac{3}{5}$ となる．

5.4 期待値と共分散の性質

任意の関数 $h(x,y)$, $g(x,y)$, $h_1(x,y)$, $h_2(x,y)$，$g_1(x)$, $g_2(y)$，および任意の定数 a, b に対して，以下が成り立つ．

【公式 28】 $E[a]=a$，$Cov[h(X,Y),a]=0$. (**定数の期待値と共分散**)
〔証明〕$(X,Y)\sim p(k,j)$ の時だけ証明する．

$$E[a]=\sum_{k=0}^{\infty}\sum_{j=0}^{\infty}ap(k,j)=a\sum_{k=0}^{\infty}\sum_{j=0}^{\infty}p(k,j)=a\times 1=a.$$

$(X,Y)\sim f(x,y)$ の時も同様に示せる．さらに，$\mu_h=E[h(X,Y)]$ とおくと，

$$Cov[h(X,Y),a]=E[\{h(X,Y)-\mu_h\}\{a-E[a]\}]=E[0]=0.$$

【公式 29】 $E[a\,h(X,Y)+b\,g(X,Y)]=aE[h(X,Y)]+bE[g(X,Y)]$. (**期待値の線形性**)
〔証明〕$(X,Y)\sim f(x,y)$ の時だけ証明する．このとき，定義より，

$$E[a\,h(X,Y)+b\,g(X,Y)]=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\{ah(x,y)+bg(x,y)\}f(x,y)dxdy$$
$$=a\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}h(x,y)f(x,y)dxdy+b\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}g(x,y)f(x,y)dxdy$$
$$=aE[h(X,Y)]+bE[g(X,Y)].$$

【公式 30】 $Cov[h(X,Y),g(X,Y)] = Cov[g(X,Y),h(X,Y)]$. (**共分散の対称性**)
〔証明〕$E[\{h(X,Y)-\mu_h\}\{g(X,Y)-\mu_g\}] = E[\{g(X,Y)-\mu_g\}\{h(X,Y)-\mu_h\}]$ だから，共分散の定義より明らか.

【公式 31】 $Cov[a\,h_1 + b\,h_2, g] = a\,Cov[h_1,g] + b\,Cov[h_2,g]$.
$Cov[h, a\,g_1 + b\,g_2] = a\,Cov[h,g_1] + b\,Cov[h,g_2]$. (**共分散の双線形性**)
〔証明〕$\tilde{h}_1(X,Y) = h_1(X,Y) - E[h_1(X,Y)]$, $\tilde{h}_2(X,Y) = h_2(X,Y) - E[h_2(X,Y)]$, $\tilde{g}(X,Y) = g(X,Y) - E[g(X,Y)]$ とおくと，

$$\begin{aligned}
&Cov[a\,h_1 + b\,h_2, g] = E[\{a\,h_1 + b\,h_2 - E[a\,h_1 + b\,h_2]\}\{g - E[g]\}]\\
=&E[\left\{a\,\tilde{h}_1 + b\,\tilde{h}_2\right\}\tilde{g}] = E[a\,\tilde{h}_1\tilde{g} + b\,\tilde{h}_2\tilde{g}] = aE[\tilde{h}_1\tilde{g}] + bE[\tilde{h}_2\tilde{g}]\\
=&aE[\{h_1 - E[h_1]\}\{g - E[g]\}] + bE[\{h_2 - E[h_2]\}\{g - E[g]\}]\\
=&a\,Cov[h_1,g] + b\,Cov[h_2,g].
\end{aligned}$$

この結果と公式 30 より，

$$\begin{aligned}
&Cov[h, a\,g_1 + b\,g_2] = Cov[a\,g_1 + b\,g_2, h]\\
=&a\,Cov[g_1,h] + b\,Cov[g_2,h] = a\,Cov[h,g_1] + b\,Cov[h,g_2].
\end{aligned}$$

【公式 32】 $E[(h-a)(g-b)] = Cov[h,g] + \{E[h]-a\}\{E[g]-b\}$.
特に, $Cov[h,g] = E[hg] - E[h]E[g]$.(**共分散公式**)
〔証明〕演習 **5.5**(p.44)．

【公式 33】 $Cov[aX+b, cY+d] = ac\,Cov[X,Y]$, $ac > 0$ の時，$r[aX+b, cY+d] = r[X,Y]$，$ac < 0$ の時，$r[aX+b, cY+d] = -r[X,Y]$.
〔証明〕演習 **5.5**(p.44)．

【公式 34】 $V[aX+bY] = a^2\,V[X] + 2ab\,Cov[X,Y] + b^2\,V[Y]$. (**分散の展開**)
〔証明〕演習 **5.5**(p.44)．

【公式 35】 $X \perp\!\!\!\perp Y$ ならば，$E[g_1(X)g_2(Y)] = E[g_1(X)]\,E[g_2(Y)]$.
特に，$E[XY] = E[X]E[Y]$，$Cov[X,Y] = 0$, $V[aX+bY] = a^2V[X] + b^2V[Y]$.
〔証明〕演習 **5.6**(p.44)．

【公式 36】

(1) $E[h_1(X)h_2(Y)|\,Y] = h_2(Y)E[h_1(X)|\,Y]$, $E[E[h(X,Y)|\,Y]] = E[h(X,Y)]$.

(2) $E[(Y - E[Y|X])E[Y|X]] = 0$, $E[Y^2] = E[(E[Y|X])^2] + E[(Y - E[Y|X])^2]$.

〔証明〕演習 **5.9**(p.44)．

《例 **5.7**》 例 **5.2**(p.34) の X, Y に対して，例 **5.5**(p.38) で $E[X] = E[Y] = \frac{7}{6}$ がわかった．これと公式 29(p.39) から，$E[4X+2Y] = 4E[X]+2E[Y] = 4\times\frac{7}{6}+2\times\frac{7}{6} = 7$．また，$V[X] = \frac{17}{36}$, $Cov[X,Y] = -\frac{17}{180}$ だったので，公式 31 を使って，$Cov[X+5Y, 3X] = 3\,Cov[X,X]+15\,Cov[Y,X] = 3V[X]+15\,Cov[X,Y] = 3\times\frac{17}{36}+15\times\left(-\frac{17}{180}\right) = 0$, さらに $V[Y] = \frac{17}{36}$ と公式 34 を使って，$V[X+5Y] = V[X]+10\,Cov[X,Y]+5^2V[Y] = 26\times\frac{17}{36}-10\times\frac{17}{180} = \frac{34}{3}$ などが計算できる．

《例 **5.8**》 例 **5.3**(p.36) の X と Y に対して, 例 **5.6**(p.38) で $E[X] = \frac{3}{5}$, $V[X] = \frac{11}{150}$, $E[Y] = \frac{7}{12}$, $V[Y] = \frac{11}{144}$, $E[XY] = \frac{7}{20}$ を求めたが，これらの結果と上の公式を利用すると色々な期待値や共分散が計算できる．たとえば，公式 29(p.39) より，$E[5X+6Y] = 5E[X]+6E[Y] = 5\times\frac{3}{5}+6\times\frac{7}{12} = \frac{13}{2}$．また，公式 32(p.40) より $Cov[X,Y] = E[XY]-E[X]E[Y] = \frac{7}{20}-\frac{3}{5}\times\frac{7}{12} = 0$ であるが，$X \perp\!\!\!\perp Y$ なので 公式 35(p.40) から $Cov[X,Y] = 0$ は計算しなくてもわかる．さらに，公式 35(p.40) から，$V[5X+6Y] = 25V[X]+36V[Y] = 25\times\frac{11}{150}+36\times\frac{11}{144} = \frac{55}{12}$．

5.5 多変量確率変数

n 個の変数 $X_1, \ldots, X_n$ の取り得る値が非負の整数 $\{0, 1, 2, \ldots\}$ であり，任意の非負整数の組 $(k_1, \ldots, k_n)$ に対して $X_1 = k_1$, …, $X_n = k_n$ が同時に成り立つ確率 $P(X_1 = k_1, \ldots, X_n = k_n)$ が定義されている時，組 $(X_1, \ldots, X_n)$ を n **変量離散確率変数**と呼ぶ．$P(X_1 = k_1, \ldots, X_n = k_n)$ を $p(k_1, \ldots, k_n)$ と表し，**同時確率関数**と呼ぶ．$X_1 = k_1$ であり，$X_2, \ldots, X_n$ はどれも非負整数のいずれかである確率 $P(X_1 = k_1, X_2 = 0, 1, 2, \ldots, \cdots, X_n = 0, 1, 2, \ldots)$ を簡単に $P(X_1 = k_1)$ と表すと，

$$P(X_1 = k_1) = \sum_{k_2=0}^{\infty}\cdots\sum_{k_n=0}^{\infty} P(X_1 = k_1, X_2 = k_2, \ldots, X_n = k_n)$$

が成り立つ．$P(X_1 = k_1)$ を $p_1(k_1)$ と表し，X_1 の**周辺確率関数**と呼ぶ．同時確率関数を用いると，

$$p_1(k_1) := \sum_{k_2=0}^{\infty}\cdots\sum_{k_n=0}^{\infty} p(k_1, k_2, \ldots, k_n)$$

と定義される．$X_2, \ldots, X_n$ のそれぞれに関する周辺確率関数 $P(X_2 = k_2), \ldots, P(X_n = k_n)$ も同様に定義して，$p_2(k_2), \ldots, p_n(k_n)$ と表すことにする．

n 変量離散確率変数 $(X_1, \ldots, X_n)$ が，任意の非負整数の組 $(k_1, \ldots, k_n$ に対して，

$$p(k_1, \ldots, k_n) = p_1(k_1)\cdots p_n(k_n)$$

を満たす時，$(X_1, \ldots, X_n)$ は**独立**であるという．

一方，n 個の変数 $X_1, \dots, X_n$ の取りうる値の範囲はどれも実数全体 $(-\infty, \infty)$ であり，確率 $P(c_1 \leqq X_1 \leqq d_1, \dots, c_n \leqq X_n \leqq d_n)$ が

$$P(c_1 \leqq X_1 \leqq d_1, \dots, c_n \leqq X_n \leqq d_n) = \int_{c_1}^{d_1} \cdots \int_{c_n}^{d_n} f(x_1, \dots, x_n) dx_1 \cdots dx_n$$

と定義されている時，組 $(X_1, \dots, X_n)$ を n **変量連続確率変数**と呼ぶ．f は**同時確率密度関数 (p.d.f.)** と呼び，$\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x_1, \dots, x_n) dx_1 \cdots dx_n = 1$ を満たす非負の関数である．$P(c_1 \leqq X_1 \leqq d_1, -\infty < X_2 < \infty, \dots, -\infty < X_n < \infty)$ を簡単に $P(c_1 \leqq X_1 \leqq d_1)$ と表すと，

$$\begin{aligned} P(c_1 \leqq X_1 \leqq d_1) &= \int_{c_1}^{d_1} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x_1, x_2, \dots, x_n) dx_1 dx_2 \cdots dx_n \\ &= \int_{c_1}^{d_1} \left\{ \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x_1, x_2, \dots, x_n) dx_2 \cdots dx_n \right\} dx_1. \end{aligned}$$

ここで，右辺の中括弧の中は $x_2, \dots, x_n$ に関する定積分なので，x_1 にのみ依存する，つまり，x_1 の関数である．それを

$$f_1(x_1) := \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x_1, x_2, \dots, x_n) dx_2 \cdots dx_n$$

と定義すると，$P(c_1 \leqq X_1 \leqq d_1) = \int_{c_1}^{d_1} f_1(x_1) dx_1$ と表されるので，$f_1(x_1)$ は X_1 の確率密度関数である．このように定義された，$f_1(x_1)$ を X_1 の**周辺確率密度関数**，簡単に，**周辺 p.d.f.** と呼ぶ．$X_2, \dots, X_n$ のそれぞれの周辺 p.d.f. $f_2(x_2), \dots, f_n(x_n)$ も同様に定義する．

また，n 変量連続確率変数 $(X_1, \dots, X_n)$ が，任意の実数の組 $(x_1, \dots, x_n)$ に対して，

$$f(x_1, \dots, x_n) = f_1(x_1) \cdots f_n(x_n)$$

を満たす時も，$(X_1, \dots, X_n)$ は**独立**であるという．

5.6 多変量確率変数の期待値と分散

n 変量離散確率変数 $(X_1, \dots, X_n) \sim p(k_1, \dots, k_n)$ と任意の関数 $h(x_1, \dots, x_n)$ に対して，

$$E[h(X_1, \dots, X_n)] := \sum_{k_1=0}^{\infty} \cdots \sum_{k_n=0}^{\infty} h(k_1, \dots, k_n) p(k_1, \dots, k_n)$$

と定義する．同様に，n 変量離散確率変数 $(X_1, \dots, X_n) \sim f(x_1, \dots, x_n)$ と任意の関数 $h(x_1, \dots, x_n)$ に対しては，

$$E[h(X_1, \dots, X_n)] := \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} h(x_1, \dots, x_n) f(x_1, \dots, x_n) dx_1 \cdots dx_n$$

と定義する．これらを $h(X_1,\dots,X_n)$ の期待値と呼び，μ_h と表す．2つの関数 $h(X_1,\dots,X_n)$ と $g(X_1,\dots,X_n)$ の共分散 $Cov[h,g]$ や $h(X_1,\dots,X_n)$ の分散 $V[h]$ を，2変量の時と同様に定義すると以下のような公式が成り立つ．

【公式 37】 定数 a_1, …, a_n に対して，

$$E[a_1X_1+\cdots+a_nX_n]=a_1E[X_1]+\cdots+a_nE[X_n] \quad (\textbf{期待値の線形性})$$

【公式 38】 定数 a_1, …, a_n に対して，

$$\begin{aligned}V[a_1X_1+\cdots+a_nX_n]=&a_1^2V[X_1]+\cdots+a_n^2V[X_n]+2a_1a_2\,Cov[X_1,X_2]\\&\cdots+2a_{n-1}a_n\,Cov[X_{n-1},X_n].\end{aligned}$$

特に，$X_1,\dots,X_n$ が独立である時，

$$V[a_1X_1+\cdots+a_nX_n]=a_1^2V[X_1]+\cdots+a_n^2V[X_n].$$

演習問題

演習 5.1 確率関数 $p(k,j)$ が下表で与えられる時，下の問いに答えなさい．ただし，$0<r<\frac{1}{3}$ であり，表にない (k,j) に対しては $p(k,j)=0$ とする．

$j\backslash k$	1	2	3
1	r	$\frac{1}{3}-r$	$\frac{1}{3}$
2	$\frac{1}{12}$	$\frac{r}{2}$	$\frac{1-2r}{4}$

(1) X と Y の周辺確率関数 $p_X(k)$, $k=1,2,3$ と $p_Y(j)$, $j=1,2$ を求めよ．

(2) $p_{X|Y}(k|1)$，$k=1,2,3$ を求めよ．

(3) X と Y が独立であるとき，r の値を求めよ．

演習 5.2 $0<x<1$, $0<y<1$ の時，$f(x,y)=\frac{1}{3}(x+ay+2xy+b)$ であり，それ以外の時は $f(x,y)=0$ である同時確率密度関数を持つ確率変数 (X,Y) に対して，以下の問いに答えなさい．ただし，a, b はある定数．

(1) a, b の満たすべき関係式を求めよ．

(2) $P(0<X<\frac{1}{2},\frac{1}{3}<Y<\frac{2}{3})$ の値を a, b を用いず表せ．

(3) X の周辺確率密度関数 $f_X(x)$ と $Y=y$ の下での X の条件付き確率密度関数 $f_{X|Y}(x|y)$ が $f_X(x)=f_{X|Y}(x|y)$ を満たすように，a, b の値を定めよ．

演習 5.3 上の **演習 5.1** において，$0 < r < \frac{1}{3}$ を満たす任意の r に対して，$E[X]$, $E[Y]$, $E[XY]$ を求めよ．さらに，それらを用いて，$Cov[X, Y]$ を求めよ．また，$E[X \mid Y = 1]$ を求めよ．

演習 5.4 上の演習 **5.2** において，$E[X]$, $E[Y]$，$E[XY]$ を求めよ．また，$Cov[X, Y]$ を求めよ．さらに，$Cov[X, Y] = 0$ の時，b の値を求め，その時 X と Y が独立であることを示せ．

演習 5.5 公式 28(p.39) ～公式 31(p.40) やそれより前の公式を用いて，公式 32(p.40) ～公式 34(p.40) を証明せよ．

演習 5.6 2 変量連続確率変数 (X, Y) に対して，公式 35(p.40) を示せ．

演習 5.7 独立確率変数列 X_1, X_2, X_3 が $E[X_1] = E[X_2] = E[X_3] = \mu$, $V[X_1] = V[X_2] = V[X_3] = \sigma^2$ を満たす時，$Y_1 = \frac{X_1+2X_2+X_3}{4}$, $Y_2 = \frac{X_1+2X_2+3X_3}{6}$ とする．$E[Y_1]$, $E[Y_2]$ を求めよ．また，$V[Y_1]$, $V[Y_2]$ を求め，どちらが小さいか答えよ．

演習 5.8 $x > 0$, $y > 0$ の時，$f(x, y) = \frac{1}{2}(x + y)e^{-x-y}$ であり，それ以外の時は $f(x, y) = 0$ である同時 p.d.f. をもつ (X, Y) に対して，以下の問いに答えよ．

(1) X の周辺 p.d.f. $f_X(x)$ を求めよ．また，$X = x$ の下で Y の条件付き p.d.f. $f_{Y|X}(y|x)$ を求めよ．

(2) $E[X]$, $Cov[X, Y]$ を求めよ．

演習 5.9 公式 36(p.40) を $(X, Y) \sim p(k, j)$ の場合に証明せよ．

演習 5.10 定数 a, b と 2 変量確率変数 (X, Y) に対して，$f(t) = E[(t(X-a)-(Y-b))^2]$ とおき，$f(t)$ を最小にする t の値を t_0 とする．

(1) t_0, $f(t_0)$ を求めよ．

(2) $f(t_0) \geqq 0$ であること利用して，$\{E[(X-a)(Y-b)]\}^2 \leqq E[(X-a)^2]\,E[(Y-b)^2]$ を示せ．

(3) $-1 \leqq r[X, Y] \leqq 1$ を示せ．

6　無作為標本の標本平均と標本分散

6.1　無作為標本

推測統計では，実験や調査を何度も繰り返し行った時に得られるだろう仮想的な集合を考えてそれを**母集団**と呼ぶのであった．実験や調査を行うことを母集団から無作為に数値を抽出することであると考え，抜き出された数値，つまり，実験や調査によって得られるデータを**標本**と呼んだ．標本は抽出されるまではどのような値になるかわからないが，母集団の度数分布 (母集団分布) によって，どのような値が出やすいかは決まる．したがって，実験や調査の前のデータや抽出前の標本を考える時，それらは確率変数であり，実際に得られたデータは確率変数の実現値である．本節では，母集団の要素が無数にある場合に*，n 回の抽出により得られる (確率変数としての) 標本 $X_1, \ldots, X_n$ について考える．なお，母集団の平均と分散をそれぞれ μ と σ^2 で表し，**母平均**，**母分散**と呼ぶことにする．

データが個数や回数を数えた計数値などのような非負整数である時，母集団は非負整数の集合であり，母集団における非負整数 $k = 0, 1, 2, \ldots$ の相対度数を p_k と表すと，１つの標本 X は $P(X = k) = p_k$ を満たす離散確率変数である．このとき，

$$\mu = \sum_{k=0}^{\infty} kp_k = E[X], \quad \sigma^2 = \sum_{k=0}^{\infty} (k-\mu)^2 p_k = V[X]$$

となり，母平均 μ と母分散 σ^2 は確率変数である１つの標本 X の平均 (期待値) $E[X]$ と分散 $V[X]$ と一致する．

n 回の無作為抽出により得られる標本 $X_1, \ldots, X_n$ に対しては，母集団の要素が無数なので，前の抽出結果が後の抽出に影響を与えず，どの X_i も確率関数が p_k だから，

$$P(X_1 = k_1, \ldots, X_n = k_n) = p_{k_1} \times \cdots \times p_{k_n}$$

なる関係が成り立つと考えられる．このことは，n 個の標本 $X_1, \ldots, X_n$ は独立な確率変数列であり，すべての周辺分布 $P(X_i = k)$ は p_k に等しいと考えることを意味する．このような周辺分布がすべて等しい，独立な離散確率変数列 $(X_1, \ldots, X_n)$ を**独立同一分布に従う**離散確率変数列 (**i.i.d.** 離散確率変数列)，または，(離散) **無作為標本**と呼び，$X_1, \ldots, X_n \overset{i.i.d.}{\sim} p_k$ と表す．p_k が代表的な離散確率変数の確率関数で，その確率分布に $B(n, p)$ などのような省略記号がある時は，$X_1, \ldots, X_n \overset{i.i.d.}{\sim} B(n, p)$ のように表す.

《例 **6.1**》　例 **3.2**(p.14)　で考えた不良品の抜取検査では，製造される製品全体の不良品率を p とし，不良品に 1，良品に 0 を割り当てた．このとき，母平均 $\mu = 1 \times p + 0 \times (1-p) = p$ であり，母分散 $\sigma^2 = (1-\mu)^2 \times p + (0-\mu)^2 \times (1-p) =$

* このような母集団を**無限母集団**と呼ぶ

$(1-p)^2p+p^2(1-p)=p(1-p)$ となる．また，1 つの製品を抜き取った結果を X とすると $P(X=1)=p$, $P(X=0)=1-p$ であった．また，例 **3.7**(p.17) で求めたように，$E[X]=p, V[X]=p(1-p)$ であり，母平均 μ と母分散 σ^2 に一致する．n 個の製品を抜き取った結果を $X_1,\ldots,X_n$ とすると，X_i の周辺確率関数は $P(X_i=1)=p$, $P(X_i=0)=1-p$ であり，同時確率関数はそれらの積になる．たとえば，$n=3$ の時，$P(X_1=0,X_2=1,X_3=1)=P(X_1=0)\times P(X_2=1)\times P(X_3=1)=(1-p)p^2$ となる．例 **3.14**(p.20) で考えたように，抜き取った n 個の製品の中の不良品の数 $Y=X_1+\cdots+X_n$ は，2 項分布 $B(n,p)$ に従い，$E[Y]=np, V[Y]=np(1-p)$ であった．

データが大きさや重さなどの計量値である時は，母集団は無数の実数からなる集合であり，その相対度数分布は階級幅を小さくするとある確率密度関数 $f(x)$ で近似できる．したがって，1 つの標本 X は確率密度関数が $f(x)$ である連続確率変数とみなせる．母集団が非負整数からなる時と同様に，1 つの標本 X の平均 $E[X]$ と分散 $V[X]$ と母平均 μ と母分散 σ^2 は一致する．つまり，

$$\mu=\int_{-\infty}^{\infty}xf(x)dx=E[X],\quad \sigma^2=\int_{-\infty}^{\infty}(x-\mu)^2f(x)dx=V[X].$$

n 個の無作為抽出による標本 $X_1,\ldots,X_n$ に対しては，母集団が非負整数からなる時と同様に，前の抽出結果が後の抽出に影響しないと考える．すなわち，どの X_i も確率密度関数が $f(x)$ である連続確率変数であり，その同時確率密度関数が，

$$f(x_1)\times\cdots\times f(x_n)$$

で与えられると考える．つまり，n 個の標本 $X_1,\ldots,X_n$ は独立で，すべての周辺確率密度関数が $f(x)$ であるある連続確率変数であると考えることになる．このような，$X_1,\ldots,X_n$ を**独立同一分布に従う**連続確率変数列 (**i.i.d.** 連続確率変数列)，または，連続**無作為標本**と呼び，$X_1,\ldots,X_n\overset{i.i.d.}{\sim}f(x)$ と表す．f が代表的な連続確率変数の p.d.f. で，その確率分布に $N(\mu,\sigma^2)$ のような省略記号がある時は，$X_1,\ldots,X_n\overset{i.i.d.}{\sim}N(\mu,\sigma^2)$ のように表す．

《例 **6.2**》 農産物 A の 1 グラムに含まれる有害物質 B の量を検査するものとする．生産されるすべての農産物 A を検査して得られる測定値全体 (母集団) が，正規分布で十分近似できる相対度数分布 (母集団分布) を持つ時，n 個の検査値 $X_1,\ldots,X_n$ は $N(\mu,\sigma^2)$ に従う (連続) 無作為標本である: $X_1,\ldots,X_n\overset{i.i.d.}{\sim}N(\mu,\sigma^2)$．また，$E[X_i]=\mu$, $V[X_i]=\sigma^2$ である．ただし，μ, σ^2 は母平均と母分散であり，生産される農産物 A を片っぱしから検査した場合に得られるであろう測定値の平均と分散を表す．

【公式 39】 *母平均 μ, 母分散 σ^2 の母集団からの無作為標本 $X_1,\ldots,X_n$ に対して，$E[X_1]=\cdots=E[X_n]=\mu, V[X_1]=\cdots=V[X_n]=\sigma^2$. $i\neq j$ の時，$Cov[X_i,X_j]=0$.*

6.2 標本平均と標本分散・標本不偏分散

標本 $X_1,\dots,X_n$ に対して，$\bar{X} := \frac{1}{n}\sum_{i=1}^n X_i$ を**標本平均**，$S^2 := \frac{1}{n}\sum_{i=1}^n (X_i-\bar{X})^2$ を**標本分散**と呼ぶ．実験や調査などで実際に得られたデータ $x_1,\dots,x_n$ は，n 変量確率変数 $X_1,\dots,X_n$ の実現値であると考えるが，それに対応して，実際のデータから計算した $\bar{x} = \frac{1}{n}\sum_{i=1}^n x_i$ や $s_{xx} = \frac{1}{n}\sum_{i=1}^n (x_i-\bar{x})^2$ は，確率変数である標本平均 $\bar{X}$ と標本分散 S^2 の実現値であると考える．

一つの標本 X_i の期待値 (平均) $E[X_i]$ や分散 $V[X_i]$ は定数であるが，$\bar{X}$, S^2 は確率変数であることに注意しよう．したがって，$\bar{X}$ や S^2 の平均 (期待値) や分散が定義できるが，無作為標本に対しては次のような結果になる．

【公式 40】 母平均 μ, 母分散 σ^2 の母集団からの無作為標本 $X_1,\dots,X_n$ に対して，

$$E[\bar{X}] = \mu, \qquad V[\bar{X}] = \frac{1}{n}\sigma^2, \qquad E[S^2] = \frac{n-1}{n}\sigma^2.$$

〔証明〕$\bar{X} = \frac{1}{n}X_1 + \cdots + \frac{1}{n}X_n$ なので，公式 37(p.43) より $E[\bar{X}] = \frac{1}{n}E[X_1] + \cdots + \frac{1}{n}E[X_n] = \frac{1}{n}\mu + \cdots + \frac{1}{n}\mu = \mu$. 公式 38(p.43) における独立な場合の展開式から，$V[\bar{X}] = \left(\frac{1}{n}\right)^2 V[X_1] + \cdots + \left(\frac{1}{n}\right)^2 V[X_n] = \frac{1}{n^2}\sigma^2 + \cdots + \frac{1}{n^2}\sigma^2 = \frac{1}{n}\sigma^2$. さらに，公式 3(p.9) から，$\frac{1}{n}\sum_{i=1}^n (X_i-\mu)^2 = S^2 + (\bar{X}-\mu)^2$ が成り立つので，

$$E[S^2] = E\left[\frac{1}{n}\sum_{i=1}^n (X_i-\mu)^2\right] - E[(\bar{X}-\mu)^2] = \frac{1}{n}\sum_{i=1}^n V[X_i] - V[\bar{X}] = \sigma^2 - \frac{1}{n}\sigma^2.$$

つまり，$E[S^2] = \frac{n-1}{n}\sigma^2$.

この結果から，標本平均 $\bar{X}$ は平均的に母平均 μ と一致するが，その分散 $V[\bar{X}]$ は母分散 σ^2 の $\frac{1}{n}$ 倍であることがわかる．つまり，実験や調査から得たデータの平均 $\bar{X}$ は，実験や調査をするごとに変動するが，平均的には母平均と一致し，そのバラツキの大きさである分散は，平均するデータの個数 n とともに小さくなることがわかる．

また，標本平均 S^2 は平均的に母分散 σ^2 と一致せず，少し小さくなる．つまり，$E[S^2] = \frac{n-1}{n}\sigma^2 < \sigma^2$. そこで，$U^2 := \frac{1}{n-1}\sum_{i=1}^n (X_i-\bar{X})^2$ と定義すると，$U^2 = \frac{n}{n-1}S^2$ なので，

$$E[U^2] = E\left[\frac{n}{n-1}S^2\right] = \frac{n}{n-1}E[S^2] = \frac{n}{n-1} \times \frac{n-1}{n} \times \sigma^2 = \sigma^2$$

であり，U^2 は平均的に σ^2 に一致する．つまり，平均的に偏りがない．この意味で，U^2 を**標本不偏分散**と呼ぶ．

《例 **6.3**》 例 **6.1**(p.45) で説明したように，例 **3.2**(p.14) で考えた不良品の抜取検査では，$\mu = E[X_i] = p$, $\sigma^2 = V[X_i] = p(1-p)$ であった．したがって，$E[\bar{X}] = p$,

$V[\bar{X}] = \frac{p(1-p)}{n}$ である．$\bar{X} = \frac{1}{n}Y$ であり，$E[Y] = np$, $V[Y] = np(1-p)$ だから，$E[\bar{X}] = \frac{1}{n}E[Y] = p$, $V[\bar{X}] = \frac{1}{n^2}V[Y] = \frac{p(1-p)}{n}$ のように計算してもよい．

《例 **6.4**》 サイコロを 50 回ふった時，i 回目に出た目を X_i とすると，例 **3.8**(p.17) や例 **3.11**(p.19) で計算したように $E[X_i] = \frac{7}{2}$, $V[X_i] = \frac{35}{12}$ だから，$E[\bar{X}] = \frac{7}{2}$, $V[\bar{X}] = \frac{1}{50} \times \frac{35}{12} = \frac{7}{120}$ である．

《例 **6.5**》 あるメーカーの蛍光灯 A の寿命は例 **3.5**(p.16) で述べたような指数分布 $Ex(\lambda)$ に従うものとする．無作為に選んだ n 本の蛍光灯 A の寿命 $X_1, \ldots, X_n$ に対して，例 **3.10**(p.17)，例 **3.13**(p.20)，例 **3.15**(p.21) で考えたように $E[X_i] = \frac{1}{\lambda}$, $V[X_i] = \frac{1}{\lambda^2}$ となるので，$E[\bar{X}] = \frac{1}{\lambda}$, $V[\bar{X}] = \frac{1}{n\lambda}$ である．

6.3 大標本の時の標本平均

【公式 41】 (中心極限定理) 母平均 μ, 母分散 σ^2 の母集団からの無作為標本 $X_1, \ldots, X_n$ に対して，$n \to \infty$ の時，

$$Z = \frac{\bar{X} - \mu}{\sqrt{\frac{\sigma^2}{n}}} = \sqrt{n}\frac{\bar{X} - \mu}{\sigma} \Rightarrow N(0,1) \quad (n \to \infty)$$

ここで，$Z \Rightarrow N(0,1)$ は，確率 $P(a < Z < b)$ は $n \to \infty$ のとき，$Z \sim N(0,1)$ として求めたものに収束することを意味する．

《例 **6.6**》 例 **6.3**(p.47) で考えた不良品の抜取検査において，$\bar{X}$ は抜き取った n 個の製品の中の不良品率を表す．この $\bar{X}$ に対して，$Z = \sqrt{n}\frac{\bar{X}-p}{\sqrt{p(1-p)}} \to N(0,1)$, $n \to \infty$ であることがわかる．今，$n = 100$, $p = \frac{1}{10}$ とすると，$Z = \sqrt{100}(\bar{X} - \frac{1}{10})/\sqrt{\frac{9}{100}} = \frac{10}{3}(10\bar{X} - 1)$ が近似的に $N(0,1)$ に従うと考えられ，

$$\begin{aligned} P(0.07 < \bar{X} < 0.13) =& P\left(\frac{10}{3}(10 \times 0.07 - 1) < \frac{10}{3}(10\bar{X} - 1) < \frac{10}{3}(10 \times 0.13 - 1)\right) \\ =& P(-1 < Z < 1) = P(Z < 1) - P(Z < -1) \\ \simeq& 0.8413 - (1 - 0.8423) = 0.6836. \end{aligned}$$

6.4 正規母集団からの標本平均と標本分散

母集団分布 (母集団の相対度数分布) が正規分布で十分近似できる母集団を**正規母集団**と呼ぶことにする．母平均 μ, 母分散 σ^2 の正規母集団からの無作為標本 $X_1, \ldots, X_n$

に対して，$X_i \sim N(\mu, \sigma^2)$ であるが，標本平均 $\bar{X}$ や標本分散 S^2 はどのような確率分布に従うのかを考える．

【公式 42】 正規母集団 $N(\mu, \sigma^2)$ からの無作為標本 $X_1, \dots, X_n$ に対して，$X_1 + \cdots + X_n \sim N(n\mu, n\sigma^2)$ であり，$\bar{X} \sim N(\mu, \frac{1}{n}\sigma^2)$．さらに，

$$Z = \frac{\bar{X} - \mu}{\sqrt{\frac{\sigma^2}{n}}} = \sqrt{n}\frac{\bar{X} - \mu}{\sigma} \sim N(0, 1)$$

〔証明〕X_1，X_2 は独立だから，公式 35(p.40) より，$M_{X_1+X_2}(t) = E[e^{t(X_1+X_2)}] = E[e^{tX_1}e^{tX_2}] = E[e^{tX_1}]E[e^{tX_2}] = M_{X_1}(t)M_{X_2}(t)$．同様に

$$M_{X_1+X_2+X_3}(t) = E[e^{t(X_1+X_2+X_3)}] = E[e^{t(X_1+X_2)}]E[e^{tX_3}] = M_{X_1}(t)M_{X_2}(t)M_{X_3}(t).$$

これを繰り返すと，$M_{X_1+\cdots+X_n}(t) = M_{X_1}(t) \times \cdots \times M_{X_n}(t)$．また，公式 24(p.30) より，$X_i$ の積率母関数 $M_{X_i}(t) = \exp\left(\mu t + \frac{\sigma^2}{2}t^2\right)$．したがって，

$$M_{X_1+\cdots+X_n}(t) = e^{\mu t + \frac{\sigma^2}{2}t^2} \times \cdots \times e^{\mu t + \frac{\sigma^2}{2}t^2} = e^{n\mu t + \frac{n\sigma^2}{2}t^2}.$$

これは，$X_1 + \cdots + X_n \sim N(n\mu, n\sigma^2)$ を意味する．この結果を公式 26(p.30) に適用すると，$\bar{X} = \frac{1}{n}(X_1 + \cdots + X_n) \sim N(\mu, \frac{1}{n}\sigma^2)$．さらに，公式 26(p.30) に適用すると，$Z = \frac{\sqrt{n}}{\sigma}\bar{X} - \frac{\sqrt{n}}{\sigma}\mu \sim N\left(\frac{\sqrt{n}}{\sigma}\mu - \frac{\sqrt{n}}{\sigma}\mu, \left(\frac{\sqrt{n}}{\sigma}\right)^2\frac{\sigma^2}{n}\right) = N(0, 1)$.

《例 **6.7**》 $X_1, \dots, X_{25} \overset{i.i.d.}{\sim} N(3, 16)$ の時，$Z = \sqrt{25}(\bar{X} - 3)/\sqrt{16} \sim N(0, 1)$ だから，

$$P(\bar{X} < 4) = P\left(\sqrt{25}\frac{\bar{X} - 3}{\sqrt{16}} < \sqrt{25}\frac{4 - 3}{\sqrt{16}}\right) = P(Z < 1.25) = 0.8944$$

$Z_1, \dots, Z_\nu$ が正規母集団 $N(0, 1)$ からの無作為標本，つまり，$Z_1, \dots, Z_\nu \overset{i.i.d.}{\sim} N(0, 1)$ の時，$W = Z_1^2 + \cdots + Z_\nu^2$ は **自由度 ν のカイ 2 乗分布** に従うといい，$W \sim \chi^2_\nu$ と表す．

$$\boxed{\chi^2_\nu := \underbrace{\{N(0,1)\}^2 + \cdots + \{N(0,1)\}^2}_{\nu\text{個}} \quad (\textbf{カイ 2 乗分布})}$$

$P(W > x) = \alpha$ となる x を $\chi^2_\nu(\alpha)$ と表し，**自由度 ν のカイ 2 乗分布の上側** $100\alpha\%$ **点**という．いくつかの自由度 ν と応用上よく用いる α の組み合わせについて，上側 $100\alpha\%$ 点 $\chi^2_\nu(\alpha)$ の値が**付表.3** にある．

【公式 43】 $\chi^2_\nu = Ga(\nu/2, 2)$．したがって，$W \sim \chi^2_\nu$ のとき，$E[W] = \nu$, $V[W] = 2\nu$.

〔証明〕$W \sim \chi_\nu$ の時，$W = Z_1^2 + \cdots + Z_\nu^2$, $Z_1, \ldots, Z_\nu \overset{i.i.d.}{\sim} N(0,1)$ と表せるので，公式 35(p.40) より，

$$M_W(t) = E[e^{tW}] = E[e^{t(Z_1^2+\cdots+Z_\nu^2)}] = E[e^{tZ_1^2}\cdots e^{tZ_\nu^2}] = E[e^{tZ_1^2}] \times \cdots \times E[e^{tZ_\nu^2}].$$

ここで，$Z_i \sim N(0,1)$ なので，

$$E[e^{tZ_i^2}] = \int_{-\infty}^{\infty} e^{tz^2} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1-2t}{2}z^2} dz$$

公式 23(p.29) において，$m = 0$, $\frac{1}{s^2} = 1 - 2t$ とすると，

$$= s \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi s^2}} e^{-\frac{1}{2s^2}z^2} dz = s = (1-2t)^{-1/2}$$

$E[e^{tZ_1^2}] = \cdots = E[e^{tZ_\nu^2}] = (1-2t)^{-1/2}$ なので，$M_W(t) = (1-2t)^{-\nu/2}$. 公式 22(p.28) より，これは $Ga(\nu/2, 2)$ の積率母関数であることがわかり，よって，$W \sim Ga(\nu/2, 2)$. ガンマ分布 $Ga(\alpha, \beta)$ の期待値は $\alpha\beta$, 分散は $\alpha\beta^2$ なので，$E[W] = \nu$, $V[W] = 2\nu$.

【公式 44】 正規母集団 $N(\mu, \sigma^2)$ からの無作為標本 $X_1, \ldots, X_n$ に対して，

(1) $\displaystyle \frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i - \mu)^2 \sim \chi_n^2.$

(2) $\displaystyle \frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i - \bar{X})^2 \sim \chi_{n-1}^2$ であり，$\displaystyle \frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i - \bar{X})^2 \perp\!\!\!\perp \frac{\bar{X}-\mu}{\sigma}.$

〔証明〕(i) $Z_i = \dfrac{X_i - \mu}{\sigma}$ とすると，$Z_1, \ldots, Z_n \overset{i.i.d.}{\sim} N(0,1)$ であり，

$$\frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i - \mu)^2 = Z_1^2 + \cdots + Z_n^2 \sim \chi_n^2.$$

(ii) $W = \dfrac{1}{\sigma^2}\displaystyle\sum_{i=1}^{n}(X_i - \mu)^2$, $Y = n\left(\dfrac{\bar{X}-\mu}{\sigma}\right)^2$ とおくと，(i) から $W \sim \chi_n^2$ であり，公式 42(p.49) から $\sqrt{n}(\bar{X}-\mu)/\sigma \sim N(0,1)$ なので，$Y \sim \chi_1^2$ であることがわかる．公式 43(p.49) とガンマ分布 $Ga(\alpha,\beta)$ の積率母関数が $(1-\beta t)^{-\alpha}$ であることから，W と Y の積率母関数は $M_W(t) = (1-2t)^{-n/2}$, $M_Y(t) = (1-2t)^{-1/2}$. また，$V = \dfrac{1}{\sigma^2}\displaystyle\sum_{i=1}^{n}(X_i - \bar{X})^2$ とおく．公式 3(p.9) において，$x_i = X_i$，$a = \mu$ とすると，$\frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2 + (\bar{X}-\mu)^2$ が成り立つので，$W = V + Y$.

もし，V と Y が独立なら，$M_W(t) = M_V(t)M_Y(t)$ なので，$M_V(t) = M_W(t)/M_Y(t) = (1-2t)^{-(n-1)/2}$．つまり，$V \sim \chi^2_{n-1}$. ここでは，$V$ と Y が独立であることを仮定したが，その独立性の証明は省略する.

$Z \sim N(0,1)$ と $V \sim \chi^2_\nu$ が独立であるとき, $T = \dfrac{Z}{\sqrt{\frac{V}{\nu}}}$ は自由度 ν のティー分布に従うといい，$T \sim t_\nu$ と表す.

$$\boxed{t_\nu = \frac{N(0,1)}{\sqrt{\frac{\chi^2_\nu}{\nu}}} \quad (\textbf{ティー分布})}$$

$P(T > x) = \alpha$ となる x を $t_\nu(\alpha)$ と表し, **自由度 ν のティー分布の上側** $100\alpha\%$ **点**という. **付表.4** を参照.

【公式 45】正規母集団 $N(\mu, \sigma^2)$ からの無作為標本 $X_1, \ldots, X_n$ に対して，$T = \dfrac{\bar{X} - \mu}{\sqrt{\frac{U^2}{n}}} \sim t_{n-1}$.
ただし，$U^2 = \dfrac{1}{n-1}\displaystyle\sum_{i=1}^{n}(X_i - \bar{X})^2$.

〔証明〕公式 42(p.49) より，$Z := \sqrt{n}\frac{\bar{X}-\mu}{\sigma} \sim N(0,1)$ であり，公式 44(p.50) (2) より，$W := \frac{(n-1)}{\sigma^2}U^2 \sim \chi^2_{n-1}$ で Z と W は独立である．$T = Z/\sqrt{W/(n-1)}$ なので，ティー分布の定義から，$T \sim t_{n-1}$.

《例 **6.8**》 $X_1, \ldots, X_{16} \overset{i.i.d.}{\sim} N(\mu, 3)$ に対して，$P(U^2 \leqq u) = 0.95$ となる u は，$W = \frac{16-1}{3}U^2 \sim \chi^2_{15}$ なので，$P(W \leqq 5u) = 0.95$. つまり，$5u = \chi^2_{15}(0.05) = 24.996$ なので，$u = 4.9992$ である. また，$V[W] = 2 \times 15 = 30$ なので，$V[U^2] = V[\frac{1}{5}W] = \frac{1}{25} \times 30 = \frac{5}{6}$．さらに，$T = \dfrac{\bar{X} - \mu}{\sqrt{\frac{U^2}{16}}} \sim t_{15}$ なので，$P(|T| < s) = 0.95$ となる s は，ティー分布が原点対称だから，$P(T > s) = (1 - 0.95)/2 = 0.025$ である．したがって，$s = t_{15}(0.025) = 2.131$ である.

演習問題

演習 6.1 サイコロの出る目を X とする．また，サイコロを n 回振ったとき，i 回目に出た目を X_i とし，$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$, $S^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$ とする．$n = 105$ の時，以下の問いに答えよ.

(1) $E[\bar{X}], V[\bar{X}], E[S^2]$ を求めよ.

(2) $P(3.4 \leqq \bar{X} \leqq 3.6)$ の値の近似値を求めよ.

演習 6.2 全有権者における内閣支持率を p , 無作為に抽出した有権者 n 人の内閣支持率を $\bar{X}$ とする. $n = 400$ として以下の問いに答えよ.

(1) $p = 1/5$ の時, $E[\bar{X}], V[\bar{X}]$ を求めよ. また, $P(\bar{X} > A) = 0.05$ となる A を中心極限定理を用いて求めよ.

(2) $p = 1/4$ の時, (1) で求めた A に対して $P(\bar{X} \geqq A)$ の値を中心極限定理を用いて求めよ. ただし, $\sqrt{3} = 1.7$ として計算せよ.

演習 6.3 センター試験の国語の得点分布は, 平均 μ, 分散 σ^2 の正規分布であるとする. 無作為に抽出した受験生 n 人の国語の得点を $X_1, \ldots, X_n$ として, その平均を $\bar{X}$, 標本不偏分散を U^2 , $T = \sqrt{n}(\bar{X} - \mu)/\sigma$ とする. $n = 25$ の時, 以下の問いに答えよ.

(1) $\sigma = 30$ のとき, $V[\bar{X}]$, $E[S^2]$, $V[S^2]$ を求めよ.

(2) $\mu = 110, \sigma = 30$ のとき, $\bar{X}$ が 122 以上になる確率を求めよ.

(3) $P(T < a) = 0.95$ となる a の値を求めよ.

(4) $P(U^2 < b) = 0.99$ となる b の値を σ^2 を用いて表せ.

(5) $P\left(\bar{X} - c\dfrac{U}{\sqrt{n}} \leqq \mu \leqq \bar{X} + c\dfrac{U}{\sqrt{n}}\right) = 0.95$ となる c の値を求めよ.

(6) $P\left(\sigma^2 \leqq \dfrac{(n-1)U^2}{d_1}\right) = 0.975$, $P\left(\dfrac{(n-1)U^2}{d_2} \leqq \sigma^2\right) = 0.025$ となる d_1, d_2 の値を求めよ.

演習 6.4 母平均 μ, 母分散 σ^2 の母集団からの無作為標本 $X_1, \ldots, X_n$ に対して, $Y_i = (X_i - \mu)/\sigma$ とおく. 以下の問いに答えよ.

(1) $M_{Y_i}(0)$, $M'_{Y_i}(0)$, $M''_{Y_i}(0)$ を求めよ.

(2) 一般に 2 階微分可能な関数 $f(x)$ に対して, 0 と x の間の実数 c を用いて, $f(x) = f(0) + f'(0)x + f''(c)x^2$ が成り立つ. このことを利用して, $M_{Y_i}(t)$ を t の多項式で表せ.

(3) $Z = \sqrt{n}(\bar{X} - \mu)/\sigma$ とすると, $Z = \frac{1}{\sqrt{n}}\sum_{i=1}^{n} Y_i$ が成り立つことを示せ. また, $M_Z(t) = M_{Y_1}(t/\sqrt{n}) \times \cdots \times M_{Y_n}(t/\sqrt{n}) = (M_{Y_1}(t/\sqrt{n}))^n$ を示せ.

(4) $n \to \infty$ のとき, $M_Z(t) \to e^{\frac{1}{2}t^2}$ を示せ. このことから, $Z \to N(0,1)$ がわかる.

7 統計的推定

7.1 点推定

母集団 Ω の未知なる特徴量の中で，推測の対象を**未知母数**と呼び，θ と表す．Ω からの無作為標本 $\boldsymbol{X} = (X_1, \ldots, X_n)$ が得られる時，無作為標本のある関数 $\hat{\theta}(\boldsymbol{X})$ を用いて,『θ の値は $\hat{\theta}(\boldsymbol{X})$ であろう』と推測することを**点推定**と呼ぶ．確率変数である標本 $\boldsymbol{X}$ の関数 $\hat{\theta}(\boldsymbol{X})$ を **(点) 推定量** と呼び，標本 $\boldsymbol{X}$ にその実現値 $\boldsymbol{x}$ を代入した値 $\hat{\theta}(\boldsymbol{x})$ を **推定値**と呼ぶ．もちろん，$\hat{\theta}(\boldsymbol{X})$ は確率変数であり，$\hat{\theta}(\boldsymbol{x})$ は定数である．

《例 **7.1**》 農産物 A の 1 グラムに含まれる有害物質 B の量を検査した n 個の測定値 $X_1, \ldots, X_n$ は正規母集団 $N(\mu, \sigma^2)$ からの無作為標本であると，例 **6.2**(p.46) で考えた．このとき，μ が母平均で，σ^2 が母分散である．

推測対象が母平均 μ であるとき，$\theta = \mu$ であり，$\theta = \mu$ に対する点推定量 $\hat{\theta}(\boldsymbol{X})$ として通常標本平均 $\bar{X}$ を用いる．$E[\hat{\theta}(\boldsymbol{X})] = E[\bar{X}] = \mu = \theta$ であり，$\hat{\theta}(\boldsymbol{X}) = \bar{X}$ は平均的に $\theta = \mu$ に一致する．このような (確率変数としての) 推定量の性質を**不偏性**とよび，望ましい性質の 1 つと考えられる．$\bar{X}$ を $\hat{\mu}$ と表すこともある．

推測対象が母分散 σ^2 であるときは，$\theta = \sigma^2$ であり，$\theta = \sigma^2$ に対する点推定量 $\hat{\theta}(\boldsymbol{X})$ として通常標本不偏分散 $U^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$ を用いる．このとき，$E[\hat{\theta}(\boldsymbol{X})] = E[U^2] = \sigma^2$ であったので，$\hat{\theta}(\boldsymbol{X}) = U^2$ は不偏性を持つ．U^2 を $\hat{\sigma}^2$ と表すこともある．

《例 **7.2**》 不良品の抜取検査について 例 **6.1**(p.45) , 例 **6.3**(p.47) などで考えた．このとき，母集団 Ω は不良品に対応する 1 と良品に対応する 0 からなり，1 の相対度数 p が不良品率であった．

推測対象は不良品率 p なので，$\theta = p$ であり，n 個の抜取検査における不良品率 $\bar{X}$ が $\theta = p$ の点推定量 $\hat{\theta}(\boldsymbol{X})$ である．もちろん，$E[\hat{\theta}(\boldsymbol{X})] = E[\bar{X}] = p = \theta$ であり，$\hat{\theta} = \bar{X}$ は不偏性を持つ．$\bar{X}$ を $\hat{p}$ と書くこともある．

一般に，不偏性は推定量として望ましい性質の 1 つであるが，不偏性に加えて $V[\hat{\theta}(\boldsymbol{X})] = E[(\hat{\theta}(\boldsymbol{X}) - \theta)^2]$ が小さいことも望ましい性質であろう．$V[\hat{\theta}(\boldsymbol{X})]$ は実現値 $\hat{\theta}(\boldsymbol{x})$ ごとの θ からの離れ具合の平均的な大きさを表すので，これが小さいことが望ましいのは明らかである．$V[\hat{\theta}(\boldsymbol{X})]$ が小さいとき，点推定量 $\hat{\theta}(\boldsymbol{X})$ は有効性があるという．有効性やその他の推定量の性質についての詳細は，専門書を参照すること．

また，推定量の構成方法には最尤法，モーメント法，最小 2 乗法などと呼ばれる様々な方法があり，推測対象の未知母数 θ や母集団分布の形状に応じて，色々な推定量が存在する．ただし，この講義では母平均 μ と母分散 σ^2 に対する点推定量 $\bar{X}$，U^2，あるいは，不良品の抜取検査で考えたような 1 と 0 のみからなる母集団 (**2 項母集団**) の 1 の比率 p(**母比率**) に対する標本の中の 1 の比率 $\bar{X}$(**標本比率**) を主に考える．

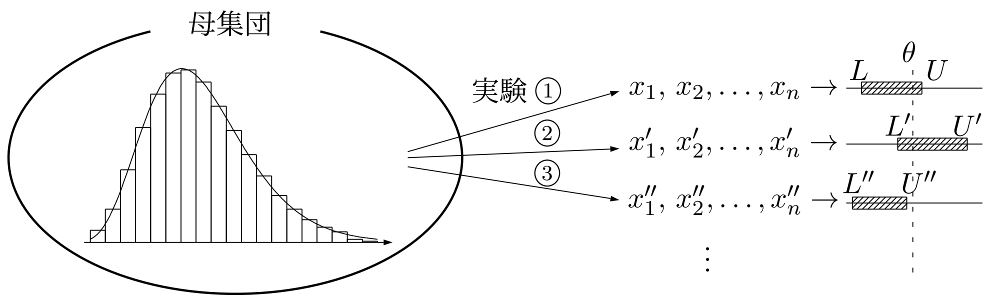
7.2 区間推定

未知母数 θ を一点 $\hat{\theta}(\boldsymbol{X})$ で推定するのを点推定と呼んだが，標本 $\boldsymbol{X}$ から区間 $[L(\boldsymbol{X}),\ U(\boldsymbol{X})]$ を求め,『θ は区間 $[L(\boldsymbol{X}),\ U(\boldsymbol{X})]$ の中にある』と推測することを **区間推定**と呼ぶ．また，区間推定に用いる区間 $[L(\boldsymbol{X}),\ U(\boldsymbol{X})]$ を**信頼区間**と呼ぶ．さらに，信頼区間 $[L(\boldsymbol{X}), U(\boldsymbol{X})]$ の中に θ が含まれることが望ましいが，その確率

$$P\left(\theta \in [L(\boldsymbol{X}), U(\boldsymbol{X})]\right) = P\left(L(\boldsymbol{X}) \leqq \theta \leqq U(\boldsymbol{X})\right)$$

を**信頼度**，あるいは，**信頼係数**と呼び，$1-\alpha$ と表す．応用上は $\alpha = 0.05, 0, 01$ とする，つまり，信頼度 $1-\alpha$ を $0.95(=95\%)$，$0.99(=99\%)$ とすることが多い．信頼度が $1-\alpha$ である信頼区間を簡単に，$1-\alpha$ 信頼区間，あるいは，$100(1-\alpha)\%$ 信頼区間と呼ぶ.

信頼度 $P\{L(\boldsymbol{X}) \leqq \theta \leqq U(\boldsymbol{X})\}$ が $1-\alpha$ であることは，確率的に変動する区間 $[L(\boldsymbol{X}), U(\boldsymbol{X})]$ が θ を含む確率が $1-\alpha$ であることを意味している．このことは，データセットが $\boldsymbol{x} = (x_1, \ldots, x_n)$, $\boldsymbol{x}' = (x'_1, \ldots, x'_n)$, $\boldsymbol{x}'' = (x''_1, \ldots, x''_n)$, $\ldots$ のように何組も得られたと仮想する時，それぞれのデータセットに対して信頼区間 $[L(\boldsymbol{x}), U(\boldsymbol{x})]$, $[L(\boldsymbol{x}'), U(\boldsymbol{x}')]$，$[L(\boldsymbol{x}''), U(\boldsymbol{x}'')]$，$\ldots$ を計算すると，それらの中に θ を含むものの割合は $100(1-\alpha)\%$ であると言い換えることができる．つまり，信頼度とは，どんなデータセットに対しても一定の計算方法で求める信頼区間 $[L(\boldsymbol{X}), U(\boldsymbol{X})]$ の性能 (命中率) を表していると考えられる.



実際に得られたデータが $\boldsymbol{x} = (x_1, \ldots, x_n)$ である時，それらから求められる信頼区間 $[L(\boldsymbol{x}), U(\boldsymbol{x})]$ の下限 $L(\boldsymbol{x})$ と上限 $U(\boldsymbol{x})$ は $L(\boldsymbol{x}) < U(\boldsymbol{x})$ を満たす定数である．未知母数 θ は標本 $\boldsymbol{X}$ の実現値 $\boldsymbol{x}$ がどのような値になるかとは無関係に一定の値をとるので，3 つの定数 θ, $L(\boldsymbol{x})$, $U(\boldsymbol{x})$ は，① $\theta < L(\boldsymbol{x}) < U(\boldsymbol{x})$, ② $L(\boldsymbol{x}) \leqq \theta \leqq U(\boldsymbol{x})$, ③ $L(\boldsymbol{x}) < U(\boldsymbol{x}) < \theta$ のいずれかの関係を満たす．与えられたデータから未知母数 θ の値を当てるために信頼区間を作るのだから，② であることが望ましいが，θ が未知である以上，どの関係が成り立つかはわからない.

したがって，信頼区間 $[L(\boldsymbol{X}), U(\boldsymbol{X})]$ が理論上仮想する無数のデータセットに対して高い性能 (信頼度 $100(1-\alpha)\%$) を持つことを根拠に,『その実現値 $[L(\boldsymbol{x}), U(\boldsymbol{x})]$ が

θ を含む (② が成り立つ) 』と $100(1-\alpha)\%$ 信じることにするというのが，区間推定の考え方である．

7.3 母平均の信頼区間

母平均 μ の信頼区間 $[L(\boldsymbol{X}),\ U(\boldsymbol{X})]$ は，点推定量である標本平均 $\bar{X}$ に幅を持たせて，

$$[L(\boldsymbol{X}),\ U(\boldsymbol{X})] = [\bar{X}-c, \bar{X}+c]$$

のように構成するのが一般的である．ただし，要求される信頼度 $100(1-\alpha)\%$ を達成するために c の値をどのように設定するかは，母集団分布の特徴や標本の大きさ n によって異なるため，以下のように場合分けして考える．

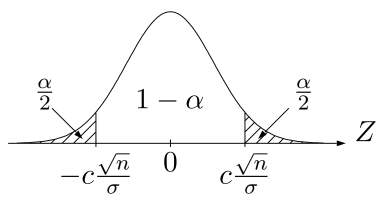
7.3.1 母分散既知の正規母集団

母分散 σ^2 の値がわかっている正規母集団 $N(\mu, \sigma^2)$ からの無作為標本 $\boldsymbol{X} = (X_1, \ldots, X_n)$ が得られる場合について考える．母平均 μ の点推定量は $\bar{X}$ であり，$Z = \sqrt{n}(\bar{X}-\mu)/\sigma$ と変形すると，公式 42(p.49) により $Z \sim N(0,1)$ であることを思い出そう．さて，

$$\begin{aligned}
\bar{X}-c \leqq \mu \leqq \bar{X}+c &\iff \bar{X}-\mu \leqq c,\quad -c \leqq \bar{X}-\mu \\
&\iff \sqrt{n}\frac{\bar{X}-\mu}{\sigma} \leqq c\frac{\sqrt{n}}{\sigma},\quad -c\frac{\sqrt{n}}{\sigma} \leqq \sqrt{n}\frac{\bar{X}-\mu}{\sigma} \\
&\iff Z \leqq c\frac{\sqrt{n}}{\sigma},\quad -c\frac{\sqrt{n}}{\sigma} \leqq Z \\
&\iff |Z| \leqq c\frac{\sqrt{n}}{\sigma}
\end{aligned}$$

なので，$N(0,1)$ の p.d.f. $f(x)$ が $x=0$ で対称であることを考慮すると，信頼度は

$$\begin{aligned}
P\left(\bar{X}-c \leqq \mu \leqq \bar{X}+c\right) =& P\left(|Z| \leqq c\frac{\sqrt{n}}{\sigma}\right) \\
=& 1-2P\left(Z > c\frac{\sqrt{n}}{\sigma}\right)
\end{aligned}$$



と表される．したがって，信頼度を $1-\alpha$ にするためには，$P(Z > c\frac{\sqrt{n}}{\sigma}) = \frac{\alpha}{2}$ とすればよい．つまり，$c = z\left(\frac{\alpha}{2}\right)\frac{\sigma}{\sqrt{n}}$ が信頼度を $1-\alpha$ とする c である．以上より，$[L(\boldsymbol{X}),\ U(\boldsymbol{X})]$ として，

$$\left[\bar{X} - z\left(\frac{\alpha}{2}\right) \times \frac{\sigma}{\sqrt{n}},\ \ \bar{X} + z\left(\frac{\alpha}{2}\right) \times \frac{\sigma}{\sqrt{n}}\right] \qquad \textbf{(母平均の信頼区間 } \mathrm{CI}_{\mu,1}\textbf{)}$$

とすると，$P\{L(\boldsymbol{X}) \leqq \mu \leqq U(\boldsymbol{X})\} = 1-\alpha$ であることがわかり，上のように定めた $[L(\boldsymbol{X}),\ U(\boldsymbol{X})]$ が信頼度 $100(1-\alpha)\%$ の信頼区間であることがわかる．

《例 **7.3**》 センター試験の国語の得点分布は，例年 $N(\mu, 30^2)$ で十分近似できるものとする．この時 μ の信頼区間は $\mathrm{CI}_{\mu,1}$ を用いると構成できる．

例えば，無作為に選ぶ 25 人の受験生の平均 $\bar{X}$ を用いて信頼度 95% の信頼区間を作るものとしよう．この場合，$n = 25$ で $1-\alpha = 0.95$，つまり $\alpha/2 = 0.025$ である．**付表.2** より $z(0.025) = 1.96$ なので，$\mathrm{CI}_{\mu,1}$ は $[\bar{X} - 1.96 \times \frac{30}{\sqrt{25}}, \bar{X} + 1.96 \times \frac{30}{\sqrt{25}}] = [\bar{X} - 11.76, \bar{X} + 11.76]$ となる．

実際に 25 人を選んで平均点を計算すると 110 であった時，$\mathrm{CI}_{\mu,1}$ は

$$[110 - 11.76, \quad 110 + 11.76] = [98.24, \quad 121.76]$$

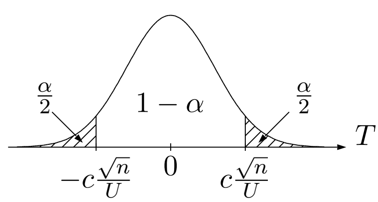
となる．

7.3.2 母分散未知の正規母集団

母分散 σ^2 も未知の正規母集団 $N(\mu, \sigma^2)$ からの無作為標本 $X_1, \ldots, X_n$ が得られる時，母平均 μ の点推定量 $\bar{X}$ を $T = \sqrt{n}(\bar{X} - \mu)/U$ と変形すると，公式 45(p.51) により $T \sim t_{n-1}$ である．前節で計算したように，

$$\begin{aligned}
\bar{X} - c \leqq \mu \leqq \bar{X} + c &\iff \bar{X} - \mu \leqq c, \quad -c \leqq \bar{X} - \mu \\
&\iff \sqrt{n}\frac{\bar{X} - \mu}{U} \leqq c\frac{\sqrt{n}}{U}, \quad -c\frac{\sqrt{n}}{U} \leqq \sqrt{n}\frac{\bar{X} - \mu}{U} \\
&\iff T \leqq c\frac{\sqrt{n}}{U}, \quad -c\frac{\sqrt{n}}{U} \leqq T \iff |T| \leqq c\frac{\sqrt{n}}{U}
\end{aligned}$$

となる．ティー分布が原点で対称であることを考慮すると，$P(\bar{X} - c \leqq \mu \leqq \bar{X} + c) = P(|T| \leqq c\sqrt{n}/U) = 1 - 2P(|T| > c\sqrt{n}/U)$ となり，信頼度を $1-\alpha$ とするためには，$P(T > c\sqrt{n}/U) = \alpha/2$, つまり，$c = t_{n-1}\left(\frac{\alpha}{2}\right)\frac{U}{\sqrt{n}}$ とするよい．以上から，



$$\left[\bar{X} - t_{n-1}\left(\frac{\alpha}{2}\right) \times \frac{U}{\sqrt{n}}, \quad \bar{X} + t_{n-1}\left(\frac{\alpha}{2}\right) \times \frac{U}{\sqrt{n}}\right] \qquad \textbf{(母平均の信頼区間 } \mathrm{CI}_{\mu,2}\textbf{)}$$

が，μ に対する 信頼度 $100(1-\alpha)\%$ の信頼区間であることがわかる．

《例 **7.4**》 センター試験の国語の得点分布が $N(\mu, \sigma^2)$ に従うものと仮定する時，$\mathrm{CI}_{\mu,2}$ を用いて μ の信頼区間を構成できる．例 **7.3** と同様に，無作為に選んだ受験生 25 人の平均 $\bar{X}$ とその標本不偏分散 U^2 を用いて信頼度 95% の信頼区間を作るものとしよ

う．このとき，$n=25$ で $1-\alpha=0.95$，つまり，$\alpha/2=0.025$ である．**付表.4** から $t_{24}(0.025)=2.064$ がわかるので，$\mathrm{CI}_{\mu,2}$ の信頼区間は $[\bar{X}-2.064\times\frac{U}{\sqrt{25}}, \bar{X}+2.064\times\frac{U}{\sqrt{25}}]=[\bar{X}-0.4128\times U, \bar{X}+0.4128\times U]$ となる．

実際に調査した結果，標本平均 $\bar{X}=110$ と標本不偏分散 $U^2=30^2$ を得た場合，$\mathrm{CI}_{\mu,2}$ は

$$[110-0.4128\times 30,\quad 110+0.4128\times 30]=[97.616,\quad 122.384]$$

である．例 **7.3**(p.56) の結果と比べると，信頼区間が少し広くなっていることに注意しよう．これは，分散の推定誤差を考慮した結果であると解釈できる．

7.3.3 大標本で任意母集団

正規母集団とは限らない任意の母集団からの無作為標本 $X_1, \ldots, X_n$ が得られるとき，母平均 μ の点推定量は標本平均 $\bar{X}$ であり，$V[\bar{X}]=\sigma^2/n$ である．公式 41(p.48) により，n が大きい時 $Z=\sqrt{n}(\bar{X}-\mu)/\sigma$ は近似的に $N(0,1)$ に従う．また，n が大きい時に $U^2\approx\sigma^2$ が知られているので，$T=\sqrt{n}(\bar{X}-\mu)/U\approx Z$ であり，T も近似的に $N(0,1)$ に従うことがわかる．このことを利用して，7.3.1 節 (p.55) と同様に考えると，

$$\left[\bar{X}-z\Big(\frac{\alpha}{2}\Big)\times\frac{U}{\sqrt{n}},\quad \bar{X}+z\Big(\frac{\alpha}{2}\Big)\times\frac{U}{\sqrt{n}}\right] \qquad \textbf{(母平均の信頼区間 }\mathrm{CI}_{\mu,3}\textbf{)}$$

が，μ に対する信頼区間であり，その信頼度は n が大きい時近似的に $1-\alpha$ であることがわかる．さらに，n が大きい時は，$U^2\approx S^2$ なので $T'=\sqrt{n}(\bar{X}-\mu)/S$ も近似的に $N(0,1)$ に従うと考えられ，信頼区間　の U を S で置き換えた

$$\left[\bar{X}-z\Big(\frac{\alpha}{2}\Big)\times\frac{S}{\sqrt{n}},\quad \bar{X}+z\Big(\frac{\alpha}{2}\Big)\times\frac{S}{\sqrt{n}}\right] \qquad \textbf{(母平均の信頼区間 }\mathrm{CI}_{\mu,4}\textbf{)}$$

でもよい．

《例 **7.5**》　センター試験の国語の得点分布がどのような分布か仮定できないものとする．こんな時でも，標本の大きさ n が大きければ，$\mathrm{CI}_{\mu,3}$ または $\mathrm{CI}_{\mu,4}$ を用いて信頼区間を構成できる．

たとえば，受験生 100 人を無作為に選び，その平均点 $\bar{X}$ と標本不偏分散 U^2 を求め，それらから 95% 信頼区間を作るものとする．このとき，$1-\alpha=0.95$ であり，$z(\alpha/2)=z(0.025)=1.96$ だから，$\mathrm{CI}_{\mu,3}$ は $[\bar{X}-1.96\times U/\sqrt{100}, \bar{X}+1.96\times U/\sqrt{100}]=[\bar{X}-0.196\times U, \bar{X}+0.196\times U]$ となる．

実際に受験生 100 人を無作為に選んだ結果，$\bar{X}=110$，$U^2=35^2$ であった時，$\mathrm{CI}_{\mu,3}$ は，

$$[110-0.196\times 35,\quad 110+0.196\times 35]=[103.14,\quad 116.86]$$

である．

《例 **7.6**》 工業製品の抜取検査において，不良品率 p は母集団における 1(不良品) の比率であり，例 **6.1**(p.45) でみたように母平均 μ であった．また，母分散 $\sigma^2 = p(1-p)$ であった．標本 X_i は 0 か 1 のどちらかの値をとり，$X_1 + \cdots + X_n$ は抜き取った n 個の製品の中の不良品の個数，$\bar{X}$ は n 個の製品の中の不良品率である．

$X_i = 1, 0$ なので，$X_i^2 = X_i$ であり，$\frac{1}{n}\sum_{i=1}^n X_i^2 = \frac{1}{n}\sum_{i=1}^n X_i = \bar{X}$ となる．したがって，標本分散 $S^2 = \frac{1}{n}\sum_{i=1}^n X_i^2 - \bar{X}^2 = \bar{X} - \bar{X}^2 = \bar{X}(1-\bar{X})$ であることがわかる．これから，$\mu = p$ の信頼区間 $\mathrm{CI}_{\mu,4}$ は

$$\left[\bar{X} - z\Big(\frac{\alpha}{2}\Big) \times \frac{\sqrt{\bar{X}(1-\bar{X})}}{\sqrt{n}},\ \ \bar{X} + z\Big(\frac{\alpha}{2}\Big) \times \frac{\sqrt{\bar{X}(1-\bar{X})}}{\sqrt{n}}\right] \quad \textbf{(母比率の信頼区間 } \mathrm{CI}_p\textbf{)}$$

となり，n が大きい時信頼度は近似的に $1-\alpha$ である．

たとえば，100 個の製品を無作為に抜き取り，その不良品率 $\bar{X}$ に基づいて 99% 信頼区間を作るものとする．このとき，$1-\alpha = 0.99$ なので，**付表.2** から $z(\alpha/2) = z(0.005) = 2.576$ ともとまり，$n = 100$ なので，CI_p は，$[\bar{X} - 2.576 \times \frac{\sqrt{\bar{X}(1-\bar{X})}}{\sqrt{100}}, \bar{X} + 2.576 \times \frac{\sqrt{\bar{X}(1-\bar{X})}}{\sqrt{100}}] = [\bar{X} - 0.2576 \times \sqrt{\bar{X}(1-\bar{X})}, \bar{X} + 0.2576 \times \sqrt{\bar{X}(1-\bar{X})}]$ となる．

実際に抜取検査をした結果，$\bar{X} = 0.3$ であった時，$\sqrt{\bar{X}(1-\bar{X})} = \sqrt{21}/10 = 0.4583$ であり，CI_p を小数第 3 位まで求めると (第 4 位を四捨五入して)

$$[0.3 - 0.2576 \times 0.4583,\ 0.3 + 0.2576 \times 0.4583] = [0.182,\ 0.418]$$

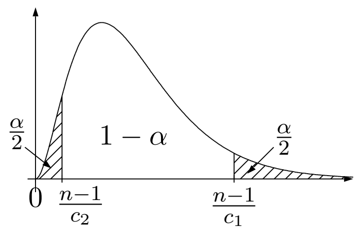
となる．

7.4 母分散の信頼区間

正規母集団 $N(\mu, \sigma^2)$ からの無作為標本 $X_1, \ldots, X_n$ が得られるとき，母分散 σ^2 の点推定量は U^2 であり，公式 44(p.50) の (2) より $V = \frac{1}{\sigma^2}\sum_{i=1}^n (X_i - \bar{X})^2 = \frac{n-1}{\sigma^2}U^2 \sim \chi^2_{n-1}$ であった．σ^2 の信頼区間として，$[c_1U^2, c_2U^2]$，$0 < c_1 < 1$, $c_2 > 1$ という形のものを考えよう．$c_1U^2 \leqq \sigma^2 \leqq c_2U^2$ を V を用いて表してみよう．

$$\begin{aligned} c_1U^2 \leqq \sigma^2 \leqq c_2U^2 &\iff \frac{1}{\sigma^2}U^2 \leqq \frac{1}{c_1},\ \ \frac{1}{c_2} \leqq \frac{1}{\sigma^2}U^2 \\ &\iff \frac{n-1}{\sigma^2}U^2 \leqq \frac{n-1}{c_1},\ \ \frac{n-1}{c_2} \leqq \frac{n-1}{\sigma^2}U^2 \\ &\iff \frac{n-1}{c_2} \leqq V \leqq \frac{n-1}{c_1}. \end{aligned}$$

したがって，$[c_1U^2, c_2U^2]$ の信頼度が $1-\alpha$ の時，



$$
\begin{aligned}
1-\alpha =& P\left(c_1U^2 \leqq \sigma^2 \leqq c_2U^2\right)\\
=& P\left(\frac{n-1}{c_2} \leqq V \leqq \frac{n-1}{c_1}\right)\\
=& 1-P\left(V < \frac{n-1}{c_2}\right) - P\left(\frac{n-1}{c_1} < V\right)
\end{aligned}
$$

よって，$P(V < \frac{n-1}{c_2}) = \frac{\alpha}{2}$, $P(\frac{n-1}{c_1} < V) = \frac{\alpha}{2}$ となるように c_1, c_2 を選べばいい．したがって，$\frac{n-1}{c_2} = \chi^2_{n-1}\big(1-\frac{\alpha}{2}\big)$, $\frac{n-1}{c_1} = \chi^2_{n-1}\big(\frac{\alpha}{2}\big)$，つまり，$c_1 = \frac{n-1}{\chi^2_{n-1}(\frac{\alpha}{2})}$, $c_2 = \frac{n-1}{\chi^2_{n-1}(1-\frac{\alpha}{2})}$.
以上のことから，

$$
\left[\frac{(n-1)}{\chi^2_{n-1}(\frac{\alpha}{2})}U^2,\ \frac{(n-1)}{\chi^2_{n-1}(1-\frac{\alpha}{2})}U^2\right] \qquad \textbf{(母分散の信頼区間 } \mathrm{CI}_{\sigma^2}\textbf{)}
$$

が，σ^2 に対する信頼度 $100\alpha\%$ の信頼区間であることがわかる．

《例 **7.7**》 正規母集団からの大きさ 20 の無作為標本に基づく，信頼度 95% の σ^2 の信頼区間を構成しよう．$1-\alpha = 0.95$ だから，$\alpha/2 = 0.025$, $1-\alpha/2 = 0.975$ であり，**付表.3** から $\chi^2_{20-1}(\alpha/2) = \chi^2_{19}(0.025) = 32.852$, $\chi^2_{20-1}(1-\alpha/2) = \chi^2_{19}(0.975) = 8.907$ を得る．したがって，CI_{σ^2} は $[\frac{(20-1)}{32.852}U^2, \frac{(20-1)}{8.907}U^2] = [0.578\times U^2, 2.133\times U^2]$ である．ただし，小数第 4 位を四捨五入した．

標本不偏分散 U^2 の実現値が 10 であったとき，CI_{σ^2} は，

$$
[0.578\times 10, 2.133\times 10] = [5.78,\ \ 21.33]
$$

となる．

演習問題

演習 7.1 あるスポーツ飲料水 (500ml) に含まれるカリウムの質量はペットボトルごとに異なり，無作為に選んだペットボトル n 本の測定値 $X_1, \dots, X_n$ は，互いに独立で同一の正規分布 $N(\mu, \sigma^2)$ に従うものとする．以下の問いに答えよ．

(1) $\sigma^2 = 216$ であることがわかっていて，$n = 24$ の時，μ の 95% 信頼区間を $\bar{X}$ を用いて表せ．また，$\bar{X} = 102$ が得られた時，その信頼区間を数値で表せ．

(2) σ^2 が未知であり，$n = 16$ の時の μ の 95% 信頼区間を $\bar{X}$, U^2 を用いて表せ．また， $\bar{X} = 102$，$U^2 = 169$ であった時，その信頼区間を数値で表せ．

演習 7.2　演習 **7.1** において，$X_1, \dots, X_n$ は互いに独立で正規分布以外の同一分布に従うものとする．σ^2 が未知で $n = 96$ である時，n が十分大きいものとして，μ の 99% 信頼区間を $\bar{X}$ と U^2 を用いて表せ．また，$\bar{X} = 102, U^2 = 216$ の時，その信頼区間を数値で表せ.

演習 7.3　いびつなサイコロの 1 の目が出る確率を p とする．n 回サイコロをふった時に 1 の目が出た割合を $\bar{X}$ とする．$n = 121$ の時，p の 99% 信頼区間を $\bar{X}$ で表せ.
また，$\bar{X} = 4/13$ の時に，p の信頼区間を計算せよ.

演習 7.4　$P(X_i = 1) = p, P(X_i = 0) = 1 - p, i = 1, \dots, n$ である独立同一分布に従う確率変数列 $X_1, \dots, X_n$ に対して，$\bar{X} = \frac{1}{n}(X_1 + \cdots + X_n)$ とする．このとき，どんな $\bar{X}$ の値に対しても，$\sqrt{\dfrac{\bar{X}(1-\bar{X})}{n}} \le \dfrac{1}{2\sqrt{n}}$ が成り立つことを示せ．これを用いて，区間 $\left[\bar{X} - \dfrac{1}{\sqrt{n}},\ \bar{X} + \dfrac{1}{\sqrt{n}}\right]$ が p を含む確率はいくつ以上であるか答えよ.

演習 7.5　演習 **7.1**(p.59) において，$n = 15$ の時の σ^2 の 95% 信頼区間を U^2 を用いて表せ．また，$U^2 = 100$ の時，その信頼区間を数値で表せ.

演習 7.6　n が大きい時，$\sqrt{n}(\bar{X} - \mu)/S$ が近似的に $N(0, 1)$ に従うことを用いて，μ の信頼区間 $\mathrm{CI}_{\mu,4}$ の信頼度が $1 - \alpha$ であることを証明せよ.

演習 7.7　$X_1, \dots, X_n \overset{i.i.d.}{\sim} Ex(1/\mu)$ の時，$E[X_i] = \mu$ であり，$M_{X_i}(t) = (1 - \mu t)^{-1}$ だから，$\bar{X} = \frac{1}{n}\sum_{i=1}^n X_i$ に対して，$\frac{2n}{\mu}\bar{X} \sim \chi^2_{2n}$ であることが証明できる．このことを利用して，μ の $1 - \alpha$ 信頼区間を作れ．また，$n = 15$ の時の 99% 信頼区間を $\bar{X}$ を用いて表せ．さらに，$\bar{X} = 6$ の時，μ の信頼区間を求めよ.

8 仮説検定

8.1 基本的な考え方

母集団 Ω の特性である未知母数 θ の取りうる値の範囲を**母数空間**と呼び， Θ と表すことにする．母数空間 Θ が Θ_0 と Θ_1 の二つに分割されている時，θ が Θ_0 に入っているという主張を H_0, Θ_1 に入っているという主張を H_1 と表すことにする：

$$\begin{cases} H_0 \ : \ \theta \in \Theta_0 \\ H_1 \ : \ \theta \in \Theta_1 \end{cases}$$

これらの主張は一方が正しければ，もう一方が偽であるという関係にあり，θ が未知定数である状況では，どちらの主張が真であるか明らかではないので，それぞれは**仮説**でしかない．

この 2 つの仮説のどちらが正しいか判定するために，Ω からの無作為標本 $\boldsymbol{X} = (X_1, \ldots, X_n)$ の適当な関数 $T(\boldsymbol{X})$ と適当な範囲 R を用意し，標本 $\boldsymbol{X}$ の実現値 $\boldsymbol{x}$ が得られた時,

$$\begin{cases} T(\boldsymbol{x}) \in R \text{ の時，} H_1 \text{ が正しい} \\ T(\boldsymbol{x}) \notin R \text{ の時，} H_0 \text{ が正しい} \end{cases}$$

と**判定する**推測手法を**仮説検定**と呼ぶ．仮説 H_0 を **帰無仮説**， H_1 を**対立仮説**と呼び，関数 $T(\boldsymbol{X})$ を **検定統計量**と呼ぶ．また，$T(\boldsymbol{X})$ の実現値 $T(\boldsymbol{x})$ が R に含まれる時，H_1 を正しいと判定し，H_0 を否定するため，R を H_1 の**採択域**，または，H_0 の**棄却域**と呼ぶ.

《例 **8.1**》 ある機械部品を製造する工場において，製造された部品の大きさが規格 μ_0 に合致しているかどうかを調べるために，製造された部品を無作為に n 個抜き取って，大きさの平均 $\bar{X}$ を測定することにした．通常，一定の基準値 c を設けて，$\mu_0 \pm c$ の範囲に $\bar{X}$ が入っている時，規格に合致していると判断する．

このとき，製造されるすべての機械部品の大きさが母集団 Ω であり，その母平均 μ が未知母数 θ であると考える．μ の取りうる値の範囲を $\Theta_0 = \{\mu_0\}$ と $\Theta_1 = (0, \mu_0) \cup (\mu_0, \infty)$ と分割すると，帰無仮説 $H_0 \ : \ \theta \in \Theta_0$ は $H_0 \ : \ \mu = \mu_0$ であり，対立仮説 $H_1 \ : \ \theta \in \Theta_1$ は $H_1 \ : \ \mu \neq \mu_0$ となる．また，検定統計量 $T(\boldsymbol{X})$ は $T(\boldsymbol{X}) = \bar{X}$ であり，棄却域 R は $R = \{\bar{X} \mid \bar{X} < \mu_0 - c,\ \mu_0 + c < \bar{X}\}$ である．

《例 **8.2**》 ある疾患に対する新薬の治癒率 p が目標値 p_0 を上回っているかどうか調べるために，無作為抽出された患者に新薬を投与し，その治癒率 $\bar{X}$ を調べることにした．通常，誤差を考慮して，$\bar{X}$ が一定の基準値 c 以上 p_0 を上回る時，p が p_0 より大きいと判断する．

このとき，その疾患を持つすべての患者に新薬を投与した結果が母集団 Ω であり，治癒した患者に 1, そうでない患者に 0 を割り当てると，2 項母集団になる．1 の

割合である母比率 p が検査したい治癒率であり，未知母数 θ である．$\Theta_0 = [0, p_0]$, $\Theta_1 = (p_0, 1]$ であり，帰無仮説 $H_0 : \theta \in \Theta_0$ は $H_0 : p \leqq p_0$ であり，対立仮説 $H_1 : \theta \in \Theta_1$ は $H_1 : p > p_0$ となる．p が p_0 より小さくなることがないと考えられるときは，$H_0 : p = p_0$ とすることもある．

また，検定統計量 $T(\boldsymbol{X})$ は $\bar{X}$ であり，棄却域を $R = \{\bar{X} \mid \bar{X} > p_0 + c\}$ である．

また，例 **8.1**(p.61) や例 **8.2**(p.61) の H_1 のように，複数の値を含む仮説を**複合仮説**，例 **8.1**(p.61) の H_0 のように，1 つの値しか含まない仮説を**単純仮説**と呼ぶ．複合仮説の中で，特定の値 θ_0 より大きいことを表す $\theta > \theta_0$ や，逆に小さいことを表す $\theta < \theta_0$ を**片側仮説**，さらに，特定の値 θ_0 ではないことを表す $\theta \neq \theta_0$ を**両側仮説**と呼ぶ．例 **8.1**(p.61) の H_1 は両側仮説であり，例 **8.2**(p.61) の H_1 は片側仮説である．

8.2　2 つの過誤

仮説検定では，**第 1 種の過誤**と呼ばれる『帰無仮説 H_0 が正しいのに，対立仮説 H_1 が正しいと判定する』誤りと，**第 2 種の過誤**と呼ばれる『対立仮説 H_1 が正しいのに，帰無仮説 H_0 が正しいと判定する』誤りの，どちらか一方が起こりうる．

第 1 種の過誤が起こる確率は，H_0 が正しいにもかかわらず，検定統計量 $T(\boldsymbol{X})$ が棄却域 R に入り，H_1 が正しいと判定する確率なので，$P(T(\boldsymbol{X}) \in R \mid H_0)$，または，$\alpha$ と表すことにする．また，第 2 種の過誤が起こる確率は，H_1 が正しいにもかかわらず，検定統計量 $T(\boldsymbol{X})$ が棄却域 R に入らないで，H_0 が正しいと判定する確率なので，$P(T(\boldsymbol{X}) \notin R \mid H_1)$，または，$\beta$ と表すことにする．

		真実	
		H_0	H_1
判定	H_0	正	β
	H_1	α	正

《例 **8.3**》 製造された機械部品が規格 μ_0 に合致しているかどうかの例 **8.1**(p.61) を再び考えよう．規格 μ_0 は $\mu_0 = 20$ とし，製造される機械部品の大きさは $N(\mu, 16)$ に従うと仮定する．そのとき，無作為に抜き取った部品の大きさ $X_1, \ldots, X_9$ は $N(\mu, 16)$ に従う無作為標本であり，公式 42(p.49) から $Z = \sqrt{9}(\bar{X} - \mu)/\sqrt{16} = \frac{3}{4}(\bar{X} - \mu) \sim N(0, 1)$ であることがわかる．基準値 $c = 2$ とすると，棄却域は $R = \{\bar{X} \mid \bar{X} < 18, 22 < \bar{X}\}$ である．

まず，第 1 種の過誤が起こる確率 $\alpha = P(\bar{X} \in R \mid H_0) = P(\bar{X} < 18, 22 < \bar{X} \mid \mu = 20)$ について考えよう．H_0 が正しい時，つまり，$\mu = \mu_0 = 20$ の時，$Z = \frac{3}{4}(\bar{X} - 20)$ なので，

$$
\begin{aligned}
\alpha =& P(\bar{X} < 18, 20 < \bar{X} \mid \mu = 20) = P\left(Z < \frac{3}{4}(18 - 20), \frac{3}{4}(22 - 20) < Z\right) \\
=& P(Z < -1.5, Z > 1.5) = 2 \times (1 - P(Z \leqq 1.5)) \\
=& 2 \times (1 - 0.9332) = 0.1336.
\end{aligned}
$$

つぎに，第2種の過誤が起こる確率 $\beta = P(\bar{X} \notin R \mid H_1) = P(18 \leqq \bar{X} \leqq 22 \mid \mu \neq 20)$ について考える．$18 \leqq \bar{X} \leqq 22 \iff \frac{3}{4}(18-\mu) \leqq \frac{3}{4}(\bar{X}-\mu) \leqq \frac{3}{4}(22-\mu)$ であり，$Z = \frac{3}{4}(\bar{X}-\mu)$ なので，

$$\beta = P(18 \leqq \bar{X} \leqq 22 \mid \mu \neq 20) = P\left(\frac{3}{4}(18-\mu) \leqq Z \leqq \frac{3}{4}(22-\mu)\right)$$

である．たとえば，$\mu = 19$ の時は $\beta = P(-0.75 \leqq Z \leqq 2.25) = 0.7612$ であり，$\mu = 22$ の時は $\beta = P(-3 \leqq Z \leqq 0) = 0.4987$ である．このように，H_1 が真であるとしても，$\mu \neq 20$ であることしかわからず，μ の値は特定されないので，第2種の過誤が起こる確率 β は一定ではない．

また，基準値 $c = 2$ とした時の第1種の過誤を犯す確率 α は，$\alpha = 0.1336$ であったが，c を大きくすると α が小さくなることがわかる．たとえば，$c = 3$ とすると，$R = \{\bar{X} \mid \bar{X} < 17, 23 < \bar{X}\}$ であり，

$$\begin{aligned}\alpha =& P(\bar{X} < 17, 23 < \bar{X} \mid \mu = 20) = P\left(Z < \frac{3}{4}(17-20), \frac{3}{4}(23-20) < Z\right)\\ =& P(Z < -2.25, 2.25 < Z) = 2 \times (1 - P(Z \leqq 2.25)) = 0.0244\end{aligned}$$

となる．一方，

$$\beta = P(17 \leqq \bar{X} \leqq 23 \mid \mu \neq 20) = P\left(\frac{3}{4}(17-\mu) \leqq Z \leqq \frac{3}{4}(23-\mu)\right)$$

なので，$\mu = 19$ の時は $\beta = P(-1.5 \leqq Z \leqq 3) = 0.9319$ であり，$\mu = 22$ の時は $\beta = P(-3.75 \leqq Z \leqq 0.75) = 0.7733$ となる (ここで，$P(Z < 3.75) = 0.9999$ であることを使った)．つまり，$c = 2$ の時より，β の値は大きくなることがわかる．

この例でみたように，仮説検定は次の特徴をもつ：

(1) 単純仮説が真である時，それを否定するような過誤が起こる確率は計算が可能であるが，複合仮説が真である場合にそれを否定するような過誤が起こる確率は特定できない．

(2) 基準値 c を大きくすると，第1種の過誤を犯す確率 α は小さくなるが，第2種の過誤を犯す確率 β は大きくなる．

8.3 有意水準

仮説検定では，第1種の過誤を犯す確率 $\alpha = P(T(\boldsymbol{X}) \in R \mid H_0)$ と第2種の過誤を犯す確率 $\beta = P(T(\boldsymbol{X}) \notin R \mid H_1)$ を同時に小さくすることが困難である．そこで，次のような方針で棄却域 R を選ぶ：

(1) 第 1 種の過誤を犯す確率 α を小さくすることを優先し，α が一定の基準値 α_0 以下になるように棄却域 R を設定する．

(2) そのような棄却域が複数ある場合は，その中で第 2 種の過誤を犯す確率 β が小さいものを選ぶ．

このとき，α_0 は検定手順の信頼性を表し，**有意水準**という．R は有意水準 α_0 の棄却域という．実用上は，有意水準 α_0 は通常 0.05 か 0.01 に設定することが多い．

このように，仮説検定では，第 1 種の過誤を犯す確率は小さくなるように設計されているので，対立仮説 H_1 が正しいと主張しても誤っている可能性が低く (α_0 以下)，自信を持って主張できる．このような意味を込めて，$T(\boldsymbol{x}) \in R$ の時，H_1 **は (**$100\alpha\%$**) 有意である**という．または，**(有意水準** $100\alpha\%$ **で)**H_0 **を棄却する**ともいう．

一方，帰無仮説 H_0 を誤って主張する可能性 β は一定の値以下になるようには設計されていないので，H_0 が正しいという主張は自信を持ってできない．このような事情を反映して，$T(\boldsymbol{X}) \notin R$ の時，H_0 は有意であるとは言わず，H_0 **を棄却しない**，または，H_1 **は有意ではない**というに留める．

対立仮説が $H_1 : \theta < \theta_0$ のような片側仮説であるとき，帰無仮説は $H_0 : \theta \geqq \theta_0$ が自然かもしれない．ところが，帰無仮説がそのような複合仮説である場合は，それを誤って棄却する確率 $\alpha = P(T(\boldsymbol{X}) \in R \mid H_0)$ は特定できないので，一定の基準である有意水準 α_0 以下にするように棄却域を設定することが難しい．このような理由で，以下の節では帰無仮説として $H_0 : \theta = \theta_0$ のような単純仮説を考える．ただし，そのように設計した棄却域を，帰無仮説が $\theta \geqq \theta_0$ というような複合仮説に用いても，第 1 種の過誤を犯す確率は有意水準 α_0 以下となることが多い．その意味でも，帰無仮説として単純仮説だけを考えることにする．

8.4 母平均の検定

母平均 μ に対する

$$(B)\begin{cases} H_0 : \mu = \mu_0 \\ H_1 : \mu \neq \mu_0 \end{cases}, \quad (U)\begin{cases} H_0 : \mu = \mu_0 \\ H_1 : \mu > \mu_0 \end{cases}, \quad (L)\begin{cases} H_0 : \mu = \mu_0 \\ H_1 : \mu < \mu_0 \end{cases}$$

の 3 種類の仮説の組について考える．いずれの帰無仮説 H_0 も $\mu = \mu_0$ という単純仮説である．これらの仮説の組に対しては，次の棄却域が自然であろう．

$$\begin{aligned} (B) \ \cdots \quad R_{\mu\neq\mu_0} &:= \{\bar{X} \mid \bar{X} < \mu_0 - c, \mu_0 + c < \bar{X}\}, \\ (U) \ \cdots \quad R_{\mu>\mu_0} &:= \{\bar{X} \mid \bar{X} > \mu_0 + c\}, \\ (L) \ \cdots \quad R_{\mu<\mu_0} &:= \{\bar{X} \mid \bar{X} < \mu_0 - c\}. \end{aligned}$$

どの棄却域でも，c の値を大きくすると，第 1 種の過誤を犯す確率 α が小さくなり，要求される有意水準 α_0 以下となるが，大きくしすぎると第 2 種の過誤を犯す確率 β

が大きくなる．そこで，以下では $\alpha=\alpha_0$ となるように c を選ぶようにする．ただし，母集団の特徴によって，c の選び方が異なるため，信頼区間の時のように，場合分けして説明する．

8.4.1 母分散既知の正規母集団

$X_1,\dots,X_n \overset{i.i.d.}{\sim} N(\mu,\sigma^2)$ で母分散 σ^2 が既知である時，$Z=\sqrt{n}(\bar{X}-\mu)/\sigma\sim N(0,1)$ であった (公式 42(p.49) 参照)．また，$Z_0=\sqrt{n}(\bar{X}-\mu_0)/\sigma$ とおくと，$\mu=\mu_0$ のとき，$Z_0=Z\sim N(0,1)$ であることに注意しよう．

まず，対立仮説 H_1 が両側仮説である場合，つまり，

$$(B)\begin{cases} H_0 \ : \ \mu=\mu_0 \\ H_1 \ : \ \mu\neq\mu_0 \end{cases}$$

の場合を考えよう．このとき，棄却域は $R_{\mu\neq\mu_0}=\{\bar{X}\mid\bar{X}<\mu_0-c,\mu_0+c<\bar{X}\}$ という形になる．したがって，

$$\begin{aligned}
\bar{X}\in R_{\mu\neq\mu_0} &\iff \bar{X}<\mu_0-c,\ \mu_0+c<\bar{X} \\
&\iff \sqrt{n}\frac{\bar{X}-\mu_0}{\sigma}<-c\frac{\sqrt{n}}{\sigma},\ c\frac{\sqrt{n}}{\sigma}<\sqrt{n}\frac{\bar{X}-\mu_0}{\sigma} \\
&\iff |Z_0|>c\frac{\sqrt{n}}{\sigma}
\end{aligned}$$

である．H_0 が真の時 $(\mu=\mu_0)$，$Z_0=Z\sim N(0,1)$ であり，

$$\alpha_0=\alpha=P\left(|Z_0|>c\frac{\sqrt{n}}{\sigma}\right)$$

だから，$c\frac{\sqrt{n}}{\sigma}=z\left(\frac{\alpha_0}{2}\right)$，つまり，$c=z\left(\frac{\alpha_0}{2}\right)\frac{\sigma}{\sqrt{n}}$ とすればよいことがわかる．よって，

$$R_{\mu\neq\mu_0,1}:=\left\{\bar{X}\ \middle|\ \bar{X}<\mu_0-\frac{z(\frac{\alpha_0}{2})\sigma}{\sqrt{n}},\ \ \mu_0+\frac{z(\frac{\alpha_0}{2})\sigma}{\sqrt{n}}<\bar{X}\right\}=\left\{Z_0\ \middle|\ |Z_0|>z\left(\frac{\alpha_0}{2}\right)\right\}$$

が対立仮説が両側仮説である場合 (B) の有意水準 α_0 の棄却域である．棄却域はもともと統計量 $\bar{X}$ に対して表現したが，Z_0 に対して表現すると最後の式のようになり，簡単になる．Z_0 は $\bar{X}$ と μ_0 の差を標準化した量なので，Z_0 に対する最後の表現は $\bar{X}$ が μ_0 から一定の距離だけ離れていることを意味することがわかる．

次に，対立仮説が複合仮説 $\mu>\mu_0$(上側仮説とも言う)，つまり，

$$(U)\begin{cases} H_0 \ : \ \mu=\mu_0 \\ H_1 \ : \ \mu>\mu_0 \end{cases}$$

について考える．このときは，棄却域は $R_{\mu>\mu_0} = \{\bar{X} \mid \bar{X} > \mu_0 + c\}$ となるが，

$$\begin{aligned} \bar{X} \in R_{\mu>\mu_0} &\iff \bar{X} > \mu_0 + c \\ &\iff \sqrt{n}\frac{\bar{X}-\mu_0}{\sigma} > c\frac{\sqrt{n}}{\sigma} \iff Z_0 > c\frac{\sqrt{n}}{\sigma} \end{aligned}$$

である．帰無仮説 H_0 が真の時 $(\mu = \mu_0)$，$Z_0 \sim N(0,1)$ であり，

$$\alpha_0 = \alpha = P\left(Z_0 > c\frac{\sqrt{n}}{\sigma}\right)$$

なので，$c\frac{\sqrt{n}}{\sigma} = z(\alpha_0)$，つまり，$c = z(\alpha_0)\frac{\sigma}{\sqrt{n}}$ とすればよいことがわかる．よって，

$$R_{\mu>\mu_0,1} = \left\{\bar{X} \;\middle|\; \mu_0 + \frac{z(\alpha_0)\sigma}{\sqrt{n}} < \bar{X}\right\} = \left\{Z_0 \;\middle|\; Z_0 > z(\alpha_0)\right\}$$

が，対立仮説が $\mu > \mu_0$ である (U) に対する有意水準 α_0 の棄却域である．Z_0 は $\bar{X}$ と μ_0 を標準化した量なので，最後の Z_0 に対する棄却域の表現は，$\bar{X}$ が μ_0 より一定の値以上に大きいことを意味していて，$\bar{X}$ に対する棄却域と同じことを表していることがわかる．$R_{\mu\neq\mu_0,1}$ では $z\left(\frac{\alpha_0}{2}\right)$ を用いたが，$R_{\mu>\mu_0,1}$ では $z(\alpha_0)$ であることに注意しよう．

同様に，対立仮説が複合仮説 $\mu < \mu_0$(下側仮説ともいう) の場合，つまり，

$$(L)\begin{cases} H_0 \ : \ \mu = \mu_0 \\ H_1 \ : \ \mu < \mu_0 \end{cases}$$

に対しても，

$$R_{\mu<\mu_0,1} = \left\{\bar{X} \;\middle|\; \bar{X} < \mu_0 - \frac{z(\alpha_0)\sigma}{\sqrt{n}}\right\} = \left\{Z_0 \;\middle|\; Z_0 < -z(\alpha_0)\right\}$$

は有意水準 α_0 の棄却域であることがわかる．

8.4.2 母分散未知の正規母集団

$X_1, \ldots, X_n \overset{i.i.d.}{\sim} N(\mu, \sigma^2)$ で σ^2 が未知の時，$T = \sqrt{n}(\bar{X} - \mu)/U \sim t_{n-1}$ であった (公式 45(p.51) 参照)．$T_0 = \sqrt{n}(\bar{X} - \mu_0)/U$ として，前節の母分散が既知である場合とまったく同様に考えると，

$$(B)\begin{cases} H_0 \ : \ \mu = \mu_0 \\ H_1 \ : \ \mu \neq \mu_0 \end{cases}$$

対しては,

$$\begin{aligned} R_{\mu\neq\mu_0,2} &= \left\{ \bar{X} \;\middle|\; \bar{X} < \mu_0 - \frac{t_{n-1}(\frac{\alpha_0}{2})U}{\sqrt{n}},\;\; \mu_0 + \frac{t_{n-1}(\frac{\alpha_0}{2})U}{\sqrt{n}} < \bar{X} \right\} \\ &= \left\{ T_0 \;\middle|\; |T_0| > t_{n-1}\left(\frac{\alpha_0}{2}\right) \right\} \end{aligned}$$

が,有意水準 α_0 の棄却域である.T_0 は $\bar{X}$ と μ_0 の差を正の数で割った量なので,T_0 に対する最後の棄却域の表現は,$\bar{X}$ が μ_0 と一定の値以上離れていることを意味しており,$\bar{X}$ に対する棄却域の表現と同じ内容を表していることがわかる.

また,

$$(U) \begin{cases} H_0 \;:\; \mu = \mu_0 \\ H_1 \;:\; \mu > \mu_0 \end{cases}$$

対しては,

$$R_{\mu>\mu_0,2} = \left\{ \bar{X} \mid \bar{X} > \mu_0 + \frac{t_{n-1}(\alpha_0)U}{\sqrt{n}} \right\} = \left\{ T_0 \;\middle|\; T_0 > t_{n-1}(\alpha_0) \right\}$$

さらに,

$$(L) \begin{cases} H_0 \;:\; \mu = \mu_0 \\ H_1 \;:\; \mu < \mu_0 \end{cases}$$

対しては,

$$R_{\mu<\mu_0,2} = \left\{ \bar{X} \mid \bar{X} < \mu_0 - \frac{t_{n-1}(\alpha_0)U}{\sqrt{n}} \right\} = \left\{ T_0 \;\middle|\; T_0 < -t_{n-1}(\alpha_0) \right\}$$

が,有意水準 α_0 の棄却域である.

8.4.3 大標本で任意母集団

$X_1, \ldots, X_n$ が i.i.d. の時,$X_i \sim N(\mu, \sigma^2)$ でなくても,n が大きければ,$Z = \sqrt{n}(\bar{X} - \mu)/\sigma$ は近似的に $N(0,1)$ に従う (公式 41(p.48) 参照).また,n が大きい時に $U^2 \approx \sigma^2$ なので,$T = \sqrt{n}(\bar{X} - \mu)/U \approx Z$ であり,T も 近似的に $N(0,1)$ に従うことがわかる.$T_0 = \sqrt{n}(\bar{X} - \mu_0)/U$ として,前々節の母分散が既知である場合とまったく同様に考えると,

$$(B) \begin{cases} H_0 \;:\; \mu = \mu_0 \\ H_1 \;:\; \mu \neq \mu_0 \end{cases} \quad \text{対しては,} \quad R_{\mu\neq\mu_0,3} = \left\{ T_0 \;\middle|\; |T_0| > z\left(\frac{\alpha_0}{2}\right) \right\}$$

が,(近似的に) 有意水準 α_0 の棄却域である.

また,

$$(U) \begin{cases} H_0 \;:\; \mu = \mu_0 \\ H_1 \;:\; \mu > \mu_0 \end{cases} \quad \text{対しては,} \quad R_{\mu>\mu_0,3} = \left\{ T_0 \;\middle|\; T_0 > z(\alpha_0) \right\}$$

さらに,

$$(L)\left\{\begin{array}{lcl} H_0 & : & \mu=\mu_0 \\ H_1 & : & \mu<\mu_0 \end{array}\right. \quad \text{対しては,} \quad R_{\mu<\mu_0,3}=\left\{T_0 \;\middle|\; T_0<-z(\alpha_0)\right\}$$

が,(近似的に) 有意水準 α_0 の棄却域である.

2 項母集団では,$\sigma=\sqrt{p(1-p)}$ のように σ が p で表現できるので,Z が近似的に $N(0,1)$ に従うことを利用することにする.$Z_0=\sqrt{n}(\bar{X}-p_0)/\sqrt{p_0(1-p_0)}$ として,前節の母分散が既知である場合とまったく同様に考えると,

$$(B)\left\{\begin{array}{lcl} H_0 & : & p=p_0 \\ H_1 & : & p\neq p_0 \end{array}\right. \quad \text{対しては,} \quad R_{p\neq p_0}=\left\{Z_0 \;\middle|\; |Z_0|>z\left(\frac{\alpha_0}{2}\right)\right\}$$

が,(近似的に) 有意水準 α_0 の棄却域である.

また,

$$(U)\left\{\begin{array}{lcl} H_0 & : & p=p_0 \\ H_1 & : & p>p_0 \end{array}\right. \quad \text{対しては,} \quad R_{p>p_0}=\left\{Z_0 \;\middle|\; Z_0>z(\alpha_0)\right\}$$

さらに,

$$(L)\left\{\begin{array}{lcl} H_0 & : & p=p_0 \\ H_1 & : & p<p_0 \end{array}\right. \quad \text{対しては,} \quad R_{p<p_0}=\left\{Z_0 \;\middle|\; Z_0<-z(\alpha_0)\right\}$$

が,(近似的に) 有意水準 α_0 の棄却域である.

《例 **8.4**》 例 **8.1**(p.61) で考えた仮説検定について,$X_1,\ldots,X_n \overset{i.i.d.}{\sim} N(\mu,\sigma^2)$ と仮定する.このとき,σ^2 が既知なら,棄却域は $R_{\mu\neq\mu_0,1}$ を用いればよいし,σ^2 が未知なら,棄却域は $R_{\mu\neq\mu_0,2}$ を用いればよい.

たとえば,$\alpha_0=0.05$, $n=25$, $\mu_0=5$, $\sigma^2=2.5^2$ の時,$z(\alpha_0/2)=z(0.025)=1.96$ だから,$R_{\mu\neq\mu_0,1}=\{Z_0 \mid |Z_0|>1.96\}$ であり,$\bar{X}=6$ であれば,$Z_0=2$ なので,H_0 は棄却される.つまり,μ は 5 と有意な差がある.

また,$\alpha_0=0.05$, $n=25$, $\mu_0=5$, σ^2 が未知の時,$t_{n-1}(\alpha_0/2)=t_{24}(0.025)=2.064$ だから,$R_{\mu\neq\mu_0,2}=\{T_0 \mid |T_0|>2.064\}$ であり,$\bar{X}=6$, $U^2=2.5^2$ であれば,$T_0=2$ なので,H_0 は棄却されない.つまり,μ は 5 と有意な差はない.

《例 **8.5**》 例 **8.2**(p.61) で考えた仮説検定については,$R_{p>p_0}$ を用いればよい.たとえば,$\alpha_0=0.05$,$n=100$,$p_0=1/5$ の時,**付表.2** より $z(0.05)=1.645$, $R_{p>p_0}=\{Z_0 \mid Z_0>1.645\}$ であり,$\sqrt{p_0(1-p_0)}=2/5$ だから,$Z_0=25\bar{X}-5$ である.したがって,$R_{p>p_0}=\{\bar{X} \mid 25\bar{X}-5>1.645\}=\{\bar{X} \mid \bar{X}>1/5+0.0658\}$ と表すこともできる.実際,$\bar{X}=1/4=0.25$ であった時は,$\bar{X}<1/5+0.0658=0.2658$ なので,治癒率 p は目標値 $p_0=1/5$ を有意に上回るとは言えない.

8.5 母分散の検定

$X_1, \ldots, X_n \overset{i.i.d.}{\sim} N(\mu, \sigma^2)$ の時，$V = \frac{1}{\sigma^2}\sum_{i=1}^n (X_i - \bar{X})^2 = \frac{n-1}{\sigma^2}U^2 \sim \chi^2_{n-1}$ であった (公式 44(p.50) の (2) 参照).

定数 σ_0^2 に対して，$V_0 = (n-1)U^2/\sigma_0^2$ とおくと，対立仮説 H_1 が両側仮説である

$$\begin{cases} H_0 \ : \ \sigma^2 = \sigma_0^2 \\ H_1 \ : \ \sigma^2 \neq \sigma_0^2 \end{cases}$$

に対しては,

$$\begin{aligned} R_{\sigma^2 \neq \sigma_0^2} &= \left\{ U^2 \ \middle| \ U^2 < \frac{\chi^2_{n-1}\left(1 - \frac{\alpha_0}{2}\right)}{n-1}\sigma_0^2, \ \frac{\chi^2_{n-1}\left(\frac{\alpha_0}{2}\right)}{n-1}\sigma_0^2 < U^2 \right\} \\ &= \left\{ V_0 \mid V_0 < \chi^2_{n-1}\left(1 - \frac{\alpha_0}{2}\right), \ \chi^2_{n-1}\left(\frac{\alpha_0}{2}\right) < V_0 \right\} \end{aligned}$$

が有意水準 α_0 の棄却域であることがわかる．V_0 は U^2 と σ_0^2 の比に正の数をかけた量なので，V_0 に対する最後の棄却域の表現は U^2 が σ_0^2 より一定の大きさだけ離れていることを意味することが分かる．以下でも，棄却域は V_0 を用いて表す.

片側仮説に対しては,

$$\begin{cases} H_0 \ : \ \sigma^2 = \sigma_0^2 \\ H_1 \ : \ \sigma^2 > \sigma_0^2 \end{cases} \quad \text{対しては,} \quad R_{\sigma^2 > \sigma_0^2} = \left\{ V_0 \ \middle| \ V_0 > \chi^2_{n-1}(\alpha_0) \right\}$$

または,

$$\begin{cases} H_0 \ : \ \sigma^2 = \sigma_0^2 \\ H_1 \ : \ \sigma^2 < \sigma_0^2 \end{cases} \quad \text{対しては,} \quad R_{\sigma^2 < \sigma_0^2} = \left\{ V_0 \ \middle| \ V_0 < \chi^2_{n-1}(1 - \alpha_0) \right\}$$

が有意水準 α_0 の棄却域であることがわかる

《例 **8.6**》 ある機械部品を製造する工程において，部品の大きさの分散が基準値 0.01 より小さいことが要求される．このことを検定するために，帰無仮説 $H_0 \ : \ \sigma^2 = 0.01$, 対立仮説 $H_1 \ : \ \sigma^2 < 0.01$ を立てて，検定することになった．ここで，σ^2 は製造されるすべての部品の大きさの分散を表す．このとき，$R_{\sigma^2 < \sigma_0^2}$ が有意水準 α_0 の棄却域である.

たとえば, $\alpha_0 = 0.05, n = 16$ のとき，**付表.3** より $\chi^2_{n-1}(1-\alpha_0) = \chi^2_{15}(0.95) = 7.261$ なので，$R_{\sigma^2 < \sigma_0^2} = \{V_0 \mid V_0 < 7.261\}$ となる．実際データを採って，$U^2 = 0.005$ であった時，$V_0 = 15 \times 0.005/0.001 = 7.5$ なので，H_0 は棄却されず，部品の分散は有意に 0.01 より小さいとは言えない.

演習問題

演習 8.1 $X_1, \ldots, X_{16} \overset{i.i.d.}{\sim} N(\mu, 25)$ の時, 帰無仮説 $H_0 : \mu = 2$, 対立仮説 $H_1 : \mu < 2$ に対する仮説検定を考える.

(1) 棄却域 $R = \{\bar{X} \mid \bar{X} < 1\}$ に対する第 1 種の過誤を犯す確率 α を求めよ．また，$\mu = 1$ の時，第 2 種の過誤を犯す確率 β_1 を求めよ．さらに，$\mu = 0$ の時の，第 2 種の過誤を犯す確率 β_2 を求めよ.

(2) 棄却域 $R' = \{\bar{X} \mid \bar{X} > 3\}$ に対する第 1 種の過誤を犯す確率 α' が (1) の α と等しいことを示せ．また，$\mu = 1$ の時，第 2 種の過誤を犯す確率 β'_1 を求め，(1) の β_1 と比較せよ．さらに，$\mu = 0$ の時の第 2 種の過誤を犯す確率 β'_2 を求め，(1) の β_2 と比較せよ．これらの結果から，R と R' のどちらが優れていると考えられるか答えよ.

(3) 対立仮説が $H'_1 : \mu = 1$ の時，第 1 種の過誤を犯す確率と第 2 種の過誤を犯す確率が等しくなるような棄却域を求め，その時の第 1 種の過誤を犯す確率を求めよ.

演習 8.2 $R_{\mu<\mu_0,1}$ が有意水準 α_0 の棄却域であることを示せ．また，$R_{\mu\neq\mu_0,2}$ が有意水準 α_0 の棄却域であることを示せ.

演習 8.3 ある工場で製造されるスナック菓子 1 袋の内容量が 100g と表示されている．平均内容量 μ が 100g であるかどうか調べるために，帰無仮説 $H_0 : \mu = 100$, 対立仮説 $H_1 : \mu \neq 100$ を立てて，無作為選んだ 16 袋の内容量 $X_1, \ldots, X_{16}$ から仮説検定することにした．$X_1, \ldots, X_{16} \overset{i.i.d.}{\sim} N(\mu, \sigma^2)$ と仮定して以下の問いに答えよ.

(1) $\sigma^2 = 9$ であることがわかっている時，有意水準 5% で仮説検定するための棄却域を求めよ．また，$\bar{X} = 98$ であった時，仮説検定の結果を述べよ.

(2) σ^2 がわからない時，有意水準 5% の棄却域 R を求めよ．また $\bar{X} = 98$ で，不偏分散 $U^2 = 9$ であった時，有意水準 5% で仮説検定を行った結果を述べよ．また，有意水準 1% ではどうか述べよ.

(3) 帰無仮説 $H_0 : \sigma^2 = 8$ と対立仮説 $H_1 : \sigma^2 \neq 8$ に対して，有意水準 5% の棄却域を求めよ．また，$U^2 = 9$ の時，仮説検定の結果を述べよ．また，1% ではどうか述べよ.

演習 8.4 $R_{\sigma^2\neq\sigma_0^2}$ が有意水準 α_0 の棄却域であることを示せ.

演習 8.5 ある製薬会社の総合鼻炎薬 B の新商品が，従来の総合鼻炎薬による改善率 55% と比べて，優れているかどうか検討するために，鼻炎患者 99 人に処方することにした．

(1) 改善率が 55% より優れているかどうかを判定するので，帰無仮説 H_0 : $\boxed{}$ と対立仮説 H_1 : $\boxed{}$ のように立てればよい．

(2) 上の仮説を有意水準 $100\alpha_0\%$ で検定するための棄却域 R は

$$R = \{Z_0 \mid Z_0 > z(\alpha_0)\}, \quad Z_0 = \boxed{}\bar{X} - \boxed{}$$

である．ただし，$\bar{X}$ は 99 人中の改善した患者の比率，$z(\alpha)$ は標準正規分布 $N(0,1)$ の上側 $100\alpha\%$ 点である．

(3) 処方した結果，鼻炎の改善が見られた患者は 64 人であった．このとき，$Z_0 = \boxed{}$ である．この新商品は従来品と比べて優れているかどうか，有意水準 5% で仮説検定すると，$z(\alpha_0) = \boxed{}$ だから，Z_0 は $z(\alpha_0)$ より { 小さく，大きく }，H_0 は { 棄却される， 棄却されない }.

(4) 上と同様の結果である時，有意水準 1% で検定すると，$z(\alpha_0) = \boxed{}$ だから，Z_0 は $z(\alpha_0)$ より { 小さく，大きく }, H_0 は { 棄却される， 棄却されない }.

演習 8.6 一辺の長さが 1 の正方形が 6 面，一辺の長さが 1 の正三角形が 8 面からなる 14 面体のサイコロを考える．正三角形の面が出る確率 p を調べるために，192 回投げたところ，正三角形の面が 48 回出た．第 i 回目に正三角形が出たとき $X_i = 1$, 正方形が出たとき $X_i = 0$ とすると，$A = X_1 + \cdots + X_n = \boxed{}$, $\bar{X} = A/n = \boxed{}$ である．ただし，$n = 192$ とする．

(1) 正規近似による p の信頼度 95% の信頼区間は，小数第 2 位を四捨五入すると，

$$\left[\ \boxed{}\ ,\ \boxed{}\ \right]$$

(2) 正三角形の面が出る確率が 1/5 かどうか，有意水準 5% で仮説検定した結果を述べよ．

演習 8.7 T 社の風邪薬 20 錠に含まれるある成分の含有量を測定したところ，標本平均が $7(mg)$ であり，標本不偏分散が $8(mg^2)$ だった．成分の含有量が正規分布に従うものとして，次の問いに答えよ．ただし，$\sqrt{10} = 3.16$ として計算せよ．

(1) 母平均 μ の 99% 信頼区間を求めよ．ただし，小数第 2 位を四捨五入して答えよ．

(2) 母平均 μ が 5.5(mg) と等しいかどうか，仮説を立てて検定を行え (有意水準 5% と 1% で).

(3) 母分散 σ^2 の 99% 信頼区間を求めよ．ただし，結果は小数第 2 位を四捨五入して答えよ．

(4) 母分散 σ^2 が 16 より有意に小さいかどうか，仮説を立て，検定を行え (有意水準 5% と 10% で).